Skript zur Vorlesung

Lineare Algebra I
Aufbaukurs

Wintersemester 2005/2006
(dreistündig)

Prof. Dr. Annette Werner

Inhaltsverzeichnis

1 Mengen und Abbildungen

Um ganz genau zu sein, müßten wir erst über die axiomiatische Begründung der Mengenlehre sprechen. Das dauert jedoch zu lange, daher geben wir uns mit einem naiven Mengenbegriff zufrieden.
Eine Menge X ist eine Zusammenfassung gewisser Objekte, diese nennt man Elemente. Wir schreiben x in X (bzw. $x \notin X$), wenn x ein Element von X (bzw. kein Element von X) ist.
Die folgenden Mengen betrachten wir als gegeben:

$$\begin{aligned}
\emptyset &= \text{leere Menge} \\
\mathbb{N} &= \{1, 2, 3, \dots\}, \text{ die Menge der natürlichen Zahlen} \\
\mathbb{N}_0 &= \{0, 1, 2 \dots\}, \text{ die Menge der natürliche Zahlen mit Null} \\
\mathbb{Z} &= \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}, \text{ die Menge der ganzen Zahlen} \\
\mathbb{Q} &= \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}, \text{ die Menge der rationalen Zahlen} \\
\mathbb{R} &= \text{die Menge der reellen Zahlen.}
\end{aligned}$$

Eine Menge ist vollständig durch ihre Elemente bestimmt. Eine Menge X ist eine Teilmenge von Y (wir schreiben $X \subset Y$), falls alle Elemente von X auch Elemente von Y sind. So gilt etwa $\emptyset \subset \mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.
Zwei Mengen sind gleich genau dann, wenn sie dieselben Elemente haben. Also ist $X = Y$ gleichbedeutend mit $X \subset Y$ und $Y \subset X$.
Um Widersprüche zu vermeiden, schreiben die Mengenaxiome für die Bildung von Mengen gewisse Regeln vor. Beispielsweise darf eine Menge niemals sich selbst als Element enthalten. Aus diesem Grund ist auch die Gesamtheit aller Mengen keine Menge! (Sondern eine Klasse). Ohne solche Vorsichtsmaßnahmen greift das Russell'sche Paradox: Wir betrachten alle Mengen A mit $A \notin A$. Angenommen, man könnte alle diese A zu einer Menge X zusammenfassen. Dann ist entweder $X \in X$ oder $X \notin X$. Im ersten Fall widerspricht das der Tatsache, dass für alle $A \in X$ die Bedingung $A \notin A$ gilt. Im zweiten Fall widerspricht das der Tatsache, dass alle $A \notin X$ die Bedingung $A \in A$ erfüllen.
Folgende Prozesse zur Mengenbildung sind hingegen erlaubt.

i) Teilmengen ausschneiden durch eine Bedingung: Ist X eine Menge und $P(x)$ eine Bedingung an die Elemente von X, so ist auch $Y = \{x \in X : P(x)\}$ eine Menge. Ein Beispiel ist die Menge der positiven reellen Zahlen $\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$.

ii) Potenzmengen bilden: Ist X eine Menge, so ist auch $\mathcal{P}(X) = \{Y : Y \subset X\}$ eine Menge, die sogenannte Potenzmenge von X.

iii) Vereinigung und Durchschnitt: Sei X eine Menge und I eine weitere Menge, deren Elemente wir als Indizes verwenden wollen. Ist für jedes $i \in I$ eine Teilmenge $X_i \subset X$ gegeben, so sei

$$\bigcup_{i \in I} X_i = \{x \in X : \text{ es gibt ein } i \in I \text{ mit } x \in X_i\}$$

die Vereinigung der X_i und

$$\bigcap_{i \in I} X_i = \{x \in X : x \in X_i \text{ für alle } i \in I\}$$

der Durchschnitt der X_i. In beiden Fällen ist dies wieder eine Teilmenge von X.

Ist $I = \{1, 2, \ldots, n\}$ endlich, so schreiben wir auch $X_1 \cup \cdots \cup X_n$ bzw. $X_1 \cap \cdots \cap X_n$.

Zwei Teilmengen X_1 und X_2 von X heißen disjunkt, falls $X_1 \cap X_2 = \emptyset$ gilt.

iv) Differenz: Sind X_1, X_2 Teilmengen von X, so ist $X_1 \setminus X_2 = \{x \in X_1 : x \notin X_2\}$ eine Teilmenge von X, die Differenz von X_1 und X_2.

v) Kartesisches Produkt: Sind $X_1, \ldots, X_n$ Mengen, so ist die Menge der n-Tupel mit Komponenten $x_i \in X_i$, also

$$X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n) : x_1 \in X_1, \ldots, x_n \in X_n\},$$

wieder eine Menge. Hier gilt $(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n)$ genau dann, wenn $x_1 = x'_1, \ldots, x_n = x'_n$ ist. Wir schreiben auch $\prod_{i=1}^{n} X_i = X_1 \times \cdots \times X_n$. Sind alle $X_i = X$, so schreiben wir auch X^n statt $\underbrace{X \times \cdots \times X}_{n\text{-mal}}$.

Ist I eine beliebige Indexmenge und ist für alle $i \in I$ eine Menge X_i gegeben, so ist auch $\prod_{i \in I} X_i = \{(x_i)_{i \in I} : x_i \in X_i\}$ wieder eine Menge. Mit $(x_i)_{i \in I}$ meinen wir ein Tupel, das für jeden Index $i \in I$ eine Komponente hat.

Eine **Abbildung** $\varphi : X \to Y$ zwischen zwei Mengen ist eine Vorschrift, die jedem $x \in X$ ein eindeutig bestimmtes Bild $\varphi(x) \in Y$ zuordnet. Statt Abbildung sagt man auch Funktion (und meint dasselbe). Wir nennen X den Definitionsbereich und Y den Wertebereich oder Zielbereich der Abbildung φ. Nach Definition ist jedes Bildelement $\varphi(x)$ in Y enthalten, es ist aber möglich, dass es Elemente $y \in Y$ gibt, die nicht Bild eines $x \in X$ unter φ sind.

Wir schreiben manchmal kurz zur Angabe der Funktionsvorschrift $x \mapsto y$ anstatt $\varphi(x) = y$. Zu jeder Menge X gibt es die identische Abbildung

$$\begin{aligned} \mathrm{id}_X \quad &: \quad X \to X, \text{ definiert durch} \\ & \quad x \mapsto x. \end{aligned}$$

Ist $\varphi : X \to Y$ eine Abbildung und $\psi : Y \to Z$ eine Abbildung, dann können wir φ und ψ komponieren (hintereinander ausführen) und erhalten eine Abbildung

$$\begin{aligned} \psi \circ \varphi \quad &: \quad X \to Z, \text{ gegeben durch} \\ & \quad x \mapsto \psi(\varphi(x)). \end{aligned}$$

Das ist nur dann definiert, wenn der Definitionsbereich von ψ mit dem Wertebereich von φ übereinstimmt.
Das **Bild** einer Abbildung $\varphi : X \to Y$ ist die Menge

$$\varphi(X) = \{y \in Y : \text{ es gibt ein } x \in X \text{ mit } \varphi(x) = y\}.$$

$\varphi(X)$ ist eine Teilmenge von Y. Wir schreiben auch Bild(φ) (in der englischen Literatur auch im(φ)) für $\varphi(X)$.
Falls das Bild von φ gleich Y ist, d.h. falls $\varphi(X) = Y$ gilt, so nennen wir φ **surjektiv**. Also ist φ genau dann surjektiv, wenn jedes $y \in Y$ von der Form $\varphi(x)$ für ein $x \in X$ ist.
Die Abbildung φ heißt **injektiv**, wenn je zwei verschiedene Elemente aus X auch verschiedene Bilder in Y haben. Das ist äquivalent dazu, dass aus $\varphi(x_1) = \varphi(x_2)$ schon $x_1 = x_2$ folgt. Man kann also nach Anwenden der Abbildung φ verschiedene Elemente von X noch auseinanderhalten.
Eine Abbildung, die injektiv und surjektiv ist, heißt **bijektiv**.
Ist $\varphi : X \to Y$ eine Abbildung, so nennen wir eine Abbildung $\psi : Y \to X$ eine **Umkehrabbildung** oder inverse Abbildung, zu φ, falls $\varphi \circ \psi : Y \to Y$ und $\psi \circ \varphi : X \to X$ jeweils die identischen Abbildungen sind, d.h. falls für alle $y \in Y \quad \varphi(\psi(y)) = y$ und für alle $x \in X \quad \psi(\varphi(x)) = x$ gilt. In diesem Fall bezeichnen wir φ auch mit ψ^{-1}.

Lemma 1.1 Eine Abbildung $\varphi : X \to Y$ hat genau dann eine Umkehrabbildung, wenn sie bijektiv ist.

Beweis : Wir nehmen zunächst an, dass φ eine Umkehrabbildung $\psi : Y \to X$ besitzt, und zeigen, dass φ injektiv und surjektiv ist. Angenommen, es gilt $\varphi(x_1) = \varphi(x_2)$. Dann folgt wegen $\psi \circ \varphi = \mathrm{id}_X$:

$$x_1 = \psi(\varphi(x_1)) = \psi(\varphi(x_2)) = x_2.$$

Also ist φ injektiv.
Ferner gilt für jedes $y \in Y$:

$$y = \varphi(\psi(y)),$$

also liegt y im Bild von φ. Daher ist φ auch surjektiv.
Nun nehmen wir umgekehrt an, φ sei bijektiv. Dann ist φ surjektiv, d.h. jedes $y \in Y$ ist von der Form $y = \varphi(x)$ für ein $x \in X$. Da φ injektiv ist, gibt es nur ein $x \in X$, das dieser Gleichung genügt. Wir definieren $\psi(y)$ als dieses eindeutig bestimmte $x \in X$ mit $y = \varphi(x)$. Dann gilt offenbar $\varphi(\psi(y)) = y$ und $\psi(\varphi(x)) = x$ für alle $y \in Y$ und alle $x \in X$. (Prüfen Sie das!) Also ist ψ eine Umkehrabbildung zu φ. □

Es sei $\varphi : X \to Y$ eine Abbildung von Mengen und U eine Teilmenge von Y. Das **Urbild** von U ist definiert als die Menge

$$\varphi^{-1}(U) = \{x \in X : \varphi(x) \in U\}$$

Achtung: Diese Menge ist auch dann definiert, wenn φ kein Inverses hat! $\varphi^{-1}(U)$ ist eine Teilmenge von X.

Lemma 1.2 Es sei $\varphi : X \to Y$ eine Abbildung von Mengen.

i) Für jede Teilmenge V von Y ist $\varphi(\varphi^{-1}(V)) \subset V$. Ist φ surjektiv, so gilt $\varphi(\varphi^{-1}(V)) = V$.

ii) Für jede Teilmenge U von X ist $U \subset \varphi^{-1}(\varphi(U))$. Ist φ injektiv, so gilt $U = \varphi^{-1}(\varphi(U))$.

Beweis :

i) Wir zeigen zunächst $\varphi(\varphi^{-1}(V)) \subset V$. Ist $y \in \varphi(\varphi^{-1}(V))$, so gibt es ein $x \in \varphi^{-1}(V)$ mit $\varphi(x) = y$. Also ist $\varphi(x) \in V$, d.h. $y \in V$. Ist φ surjektiv, so zeigen wir zusätzlich $V \subset \varphi(\varphi^{-1}(V))$. Es sei $y \in V$. Da φ surjektiv ist, existiert ein $x \in X$ mit $\varphi(x) = y$. Aus $\varphi(x) \in V$ folgt $x \in \varphi^{-1}(V)$. Also folgt $y \in \varphi(\varphi^{-1}(V))$.

ii) Wir zeigen zunächst $U \subset \varphi^{-1}(\varphi(U))$. Sei $x \in U$. Dann ist $\varphi(x) \in \varphi(U)$, also $x \in \varphi^{-1}(\varphi(U))$. Ist φ injektiv, so zeigen wir zusätzlich $\varphi^{-1}(\varphi(U)) \subset U$. Ist $x \in \varphi^{-1}(\varphi(U))$, so gilt $\varphi(x) \in \varphi(U)$. Also existiert ein $x' \in U$ mit $\varphi(x) = \varphi'(x')$. Da φ injektiv ist, folgt $x = x' \in U$.

□

Eine Menge X heißt endlich, falls sie endlich viele Elemente enthält. In diesem Fall bezeichnen wir die Anzahl der Elemente von X auch mit $|X|$ und nennen diese Zahl die Kardinalität oder die Mächtigkeit der Menge X. Ist X unendlich, so schreiben wir $|X| = \infty$.

Satz 1.3 Es sei $\varphi : X \to Y$ eine Abbildung zwischen zwei endlichen Mengen.

i) Ist φ injektiv, so ist $|X| \leq |Y|$.

ii) Ist φ surjektiv, so ist $|X| \geq |Y|$.

iii) Ist $|X| = |Y|$, so ist φ entweder bijektiv oder weder injektiv noch surjektiv.

Beweis :

i) Wir schreiben $X = \{x_1, \ldots, x_n\}$ mit $n = |X|$. Für jedes $i = 1, \ldots, n$ sei $y_i = \varphi(x_i)$. Dann sind die Elemente $y_1, \ldots, y_n$ von Y paarweise verschieden. Ist nämlich $y_i = y_j$, d.h. $\varphi(x_i) = \varphi(x_j)$, so folgt aus der Injektivität von φ, dass $x_i = x_j$, d.h. $i = j$ gilt. Also folgt $|Y| \geq n = |X|$.

ii) Es sei $Y = \{y_1, \ldots, y_m\}$ mit $m = |Y|$. Da φ surjektiv ist, existiert zu jedem $i = 1, \ldots, m$ ein $x_i \in X$ mit $\varphi(x_i) = y_i$. Die Elemente $x_1, \ldots, x_m$ von X sind paarweise verschieden, denn aus $x_i = x_j$ folgt $y_i = \varphi(x_i) = \varphi(x_j) = y_j$, d.h. $i = j$. Also ist $|X| \geq m = |Y|$.

iii) Es sei $|X| = |Y|$ und φ nicht bijektiv. Ist φ nicht injektiv, so seien $x, x' \in X$ Elemente mit $\varphi(x) = \varphi(x')$. Angenommen, φ ist surjektiv. Dann betrachten wir $\tilde{\varphi} = \varphi|_{X \setminus \{x'\}} : X \setminus \{x'\} \to Y$. Diese Abbildung ist immer noch surjektiv, d.h. nach ii) ist $|X| - 1 = |X \setminus \{x'\}| \geq |Y|$, was der Voraussetzung $|X| = |Y|$ widerspricht. Also ist φ auch nicht surjektiv. Ist φ nicht surjektiv, so sei $y \in Y$ ein Element, das nicht in Bild(φ) liegt. Angenommen, φ ist injektiv. Dann betrachten wir die Abbildung $\tilde{\varphi} : X \to Y \setminus \{y\}$, definiert als $\tilde{\varphi}(x) = \varphi(x)$. $\tilde{\varphi}$ ist immer noch injektiv, nach i) gilt also $|X| \leq |Y \setminus \{y\}| = |Y| - 1$, was $|X| = |Y|$ widerspricht. Also ist φ auch nicht injektiv.

□

Die Umkehrung von Satz 1.3 i) nennt man auch das
Schubfachprinzip: Ist $|X| > |Y|$, so ist φ nicht injektiv.
Verteilt man etwa 27 Socken in 20 Schubladen, so enthält eine Schublade mindestens zwei Socken.
Auch für unendliche Mengen kann man noch verschiedene Größen (Kardinalitäten) definieren. Wir nennen eine unendliche Menge X abzählbar, falls es eine bijektive Abbildung

$$\varphi : \mathbb{N} \to X$$

von der Menge der natürlichen Zahlen nach X gibt. (Wir können X also mit $\varphi(x_1), \varphi(x_2), \ldots$ komplett aufzählen). Gibt es keine solche Abbildung, so heißt X überabzählbar.

Satz 1.4 Die Menge $\mathbb{R}$ der reellen Zahlen ist überabzählbar.

Beweis : Der folgende Beweis heißt auch Cantor'sches Diagonalargument. Wir benutzen die Tatsache, dass jede reelle Zahl eine Dezimaldarstellung der Form

$$ab_1b_2b_3\ldots$$

mit einer ganzen Zahl a und abzählbar unendlich vielen Nachkommastellen $b_1, b_2, b_3\ldots$ hat. Dabei gibt es reelle Zahlen, die zwei Dezimalentwicklungen haben, so ist etwa $0,99999\cdots = 1,00000\ldots$.
Angenommen, $\varphi : \mathbb{N} \to \mathbb{R}$ ist eine Abbildung. Dann schreiben wir jede der reellen Zahlen $\varphi(1), \varphi(2), \varphi(3)\ldots$ in Dezimaldarstellung. Wir lassen alle diejenigen Zahlen weg, für die in der Dezimalentwicklung eine Neun oder eine Null vorkommt. So erhalten wir eine Liste $r_1, r_2, r_3\ldots$ reeller Zahlen. Wenn φ surjektiv ist, so muss jede reelle Zahl, deren Dezimalentwicklung keine Neun und keine Null enthält, in dieser Liste vorkommen. Das wollen wir jetzt widerlegen, indem wir eine Zahl x konstruieren, die in der Liste nicht vorkommen kann. Es existiert nämlich eine Zahl der Form

$$x = 0, b_1b_2b_3\ldots,$$

wobei alle $b_i \neq 0$ und $\neq 9$ sind und so gewählt sind, dass b_1 ungleich der ersten Nachkommastelle von r_1, b_2 ungleich der zweiten Nachkommastelle von r_2, b_3 ungleich der dritten Nachkommastelle von r_3 usw. ist. Also ist x eine reelle Zahl, in deren Dezimalentwicklung weder Nullen noch Neunen auftreten und die trotzdem nicht auf unserer Liste steht. □

Wir wollen jetzt das sogenannte Zorn'sche Lemma kennenlernen, das ein nützliches Hilfsmittel für die Handhabung überabzählbarer Mengen ist. Das Zorn'sche Lemma ist logisch äquivalent zum sogenannten Auswahlaxiom, das Teil der Mengenaxiome ist. Wir werden es hier nicht beweisen. Zunächst erklären wir einige Begriffe.

Definition 1.5 i) Es sei X eine Menge. Eine Relation auf X ist eine Teilmenge $R \subset X \times X$. Die Menge R besteht also aus gewissen Paaren (x, x') für $x \in X, x' \in X$.

ii) Eine Relation R auf X heißt partielle Ordnung, und wir schreiben auch $x \leq x'$ statt $(x, x') \in R$, wenn folgende Bedingungen gelten:

a) $x \leq x$ für alle $x \in X$

b) Aus $x \leq x'$ und $x' \leq x''$ folgt $x \leq x''$.

c) Aus $x \leq x'$ und $x' \leq x$ folgt $x = x'$.

Eine partielle Ordnung heißt Totalordnung, falls zusätzlich gilt:

d) Für alle $x, x' \in X$ ist $x \leq x'$ oder $x' \leq x$.

In einer partiellen Ordnung müssen zwei Elemente also nicht vergleichbar sein. So definiert zum Beispiel die Relation

$$A \leq B \text{ genau dann, wenn } A \subset B \text{ ist,}$$

eine partielle Ordnung auf der Potenzmenge $\mathcal{P}(\mathbb{R})$ von $\mathbb{R}$, aber keine Totalordnung. (Überlegen Sie sich das!) Die gewöhnliche Ordnung $\leq$ auf $\mathbb{R}$ ist hingegen eine Totalordnung.
Ist A eine Teilmenge einer Menge X und $\leq$ eine partielle Ordnung auf X, so heißt ein $s \in X$ obere Schranke von A, falls für alle $a \in A$ gilt: $a \leq s$.
Es muss also jedes Element von A

i) mit s vergleichbar und

ii) $\leq s$ sein.

(Aber s muss nicht in A liegen.) Ein $m \in X$ heißt maximales Element von X, wenn für alle $x \in X$ aus $m \leq x$ schon $m = x$ folgt. Das bedeutet: Sind m und x vergleichbar, so ist $x \leq m$. Es kann aber natürlich Elemente x geben, die nicht mit m vergleichbar sind. Hat X ein maximales Element m, so muss m daher keine obere Schranke von X sein. Betrachtet man etwa die Menge X der Teilmengen von $\{1, \ldots, n\}$, die $\neq \{1, \ldots, n\}$ sind, versehen mit der partiellen Ordnung $A \leq B$ genau dann, wenn $A \subset B$ ist, so besitzt X die maximalen Elemente $\{2, 3, \ldots, n\}, \{1, 3, 4, \ldots, n\}, \{1, 2, 4, \ldots, n\}$ etc., aber keine obere Schranke (Prüfen Sie das!).

Satz 1.6 (Zorn'sches Lemma) Es sei $(X, \leq)$ eine partiell geordnete Menge. Besitzt jede Teilmenge $A \subset X$, die bezüglich der Einschränkung von $\leq$ auf A total geordnet ist, eine obere Schranke in M, so hat X ein maximales Element.

Wir werden dieses Prinzip etwa anwenden, um zu zeigen, dass jeder Vektorraum eine Basis besitzt.
Wir wollen zum Abschluss dieses Paragraphen noch das Prinzip der vollständigen Induktion besprechen. Dies ist eine Methode, um eine Folge von Aussagen $(P_n)_{n \in \mathbb{N}}$ zu beweisen. P_n ist hier eine Behauptung, die von einer natürlichen Zahl n abhängt. Um die Aussage P_n für alle $n \in \mathbb{N}$ zu beweisen, genügt es nach dem Prinzip der vollständigen Induktion, zwei Dinge zu tun:

i) **Induktionsanfang:** Man zeigt, dass P_1 gilt.

ii) **Induktionsschluss:** Man zeigt für alle $k \in \mathbb{N}$, dass aus der Gültigkeit von P_k die Gültigkeit von P_{k+1} folgt. Mit anderen Worten, man zeigt für all $k \in \mathbb{N}$ die Folgerung $P_k \Rightarrow P_{k+1}$. Die Aussage P_k nennt man dann die Induktionsvoraussetzung.

Man muss sich klarmachen, dass ii) alleine nichts darüber sagt, ob P_k gilt oder nicht. Es besagt nur: **Wenn** P_k stimmt, so stimmt auch P_{k+1}. Zum Beispiel ist es leicht zu zeigen, dass für die Aussage

$$P_k : k(k+1) \text{ ist eine ungerade Zahl}$$

die Bedingung ii) erfüllt ist. Ist nämlich $k(k+1)$ ungerade, so ist $(k+1)(k+2) = (k+1)k + (k+1)2$ als Summe einer ungeraden und einer geraden Zahl ebenfalls ungerade. Trotzdem ist $k(k+1)$ für jede natürliche Zahl gerade, d.h. P_k ist immer falsch! (Wieso?) Hier fehlt nämlich der Induktionsanfang.
Wir wollen jetzt ein Beispiel für einen Induktionsbeweis sehen. Es sei $\mathcal{P}_k(n)$ die Menge aller k-elementigen Teilmengen von $\{1, 2, \ldots, n\}$. Ferner sei für jede natürliche Zahl r die Zahl $r!$ (sprich „ r Fakultät"), definiert als $r! := 1 \cdot 2 \cdots r$. Wir setzen außerdem $0! = 1$.

Satz 1.7 Für alle natürlichen Zahlen k und n mit $k \leq n$ ist

$$|\mathcal{P}_k(n)| = \frac{n!}{k!(n-k)!}.$$

Die Zahl auf der rechten Seite bezeichnet man auch mit $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ (sprich „ n über k "). Aus Satz 1.7 folgt insbesondere, dass $\binom{n}{k}$ eine natürliche Zahl ist, was man dem Bruch nicht auf den ersten Blick ansieht.

Beweis : Wir zeigen für jede natürliche Zahl n die Aussage

$$P_n : \text{ Für alle } 1 \leq k \leq n \text{ ist } |\mathcal{P}_k(n)| = \frac{n!}{k'(n-k)!}$$

mit Induktion nach n.
Induktionsanfang: Ist $n = 1$, so muss auch $k = 1$ sein. Offenbar ist $|\mathcal{P}_1(1)| = 1$. Andererseits ist $\frac{1!}{1!0!} = 1$, die Aussage P_1 stimmt also.
Induktionsschluss: Wir zeigen für alle $n \geq 2$, dass $P_{n-1} \Rightarrow P_n$ gilt.
Ist $T \subset \{1 \ldots, n\}$ eine beliebige k-elementige Teilmenge, so ist entweder $n \notin T$ und somit $T \subset \{1, \ldots, n-1\}$ oder aber $n \in T$ und somit $T \setminus \{n\}$ eine $(k-1)$-elementige Teilmenge von $\{1, \ldots, n-1\}$. Umgekehrt ist jede k-elementige Teilmenge

von $\{1\ldots,n-1\}$ auch eine k-elementige Teilmenge von n, und für jede $(k-1)$-elementige Teilmenge T von $\{1,\ldots,n-1\}$ ist $T\cup\{n\}$ eine k-elementige Teilmenge von $\{1,\ldots,n\}$. Daher ist

$$|\mathcal{P}_k(n)| = |\mathcal{P}_k(n-1)| + |\mathcal{P}_{k-1}(n-1)|.$$

Nach Induktionsvorausetzung ist $|\mathcal{P}_k(n-1)| = \binom{n-1}{k}$ und $|\mathcal{P}_{k-1}(n-1)| = \binom{n-1}{k-1}$. Also gilt

$$\begin{aligned} |\mathcal{P}_k(n)| &= \binom{n-1}{k} + \binom{n-1}{k-1} = \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \\ &= \frac{(n-k)(n-1)! + k(n-1)!}{k!(n-k)!} = \frac{n(n-1)!}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k}. \end{aligned}$$

Also gilt P_n. □

Eine Variante des Prinzips der Vollständigen Induktion ist folgendes Prinzip: Um eine Aussage P_n für alle $n\in\mathbb{N}$ zu beweisen, genügt es, für alle $n\in\mathbb{N}$ zu zeigen: Gilt P_k für alle $k<n$, so gilt auch P_n. Man beachte, dass man so P_1 mit beweist, da es keine natürlichen Zahlen <1 gibt. Manchmal ist dieser Ansatz praktischer, wie etwa im Beweis des folgenden Satzes.

Satz 1.8 Jede natürliche Zahl $n>1$ lässt sich als Produkt von Primzahlen schreiben.

Beweis : Wir zeigen für alle $n>1$:
Lässt sich jedes k mit $1<k<n$ als Produkt von Primzahlen schreiben, so lässt sich auch n als Produkt von Primzahlen schreiben. Ist n selbst eine Primzahl, so sind wir fertig. Andernfalls hat n einen Teiler a mit $1<a<n$. Dann gilt $n=ab$ mit einer Zahl b, für die ebenfalls $1<b<n$ gilt. Wir können die Induktionsvoraussetzung also auf a und b anwenden und beide Zahlen als Produkt von Primzahlen schreiben. Dann ist aber auch n das Produkt von Primzahlen. □

2 Matrizen

Wir wollen zeigen, dass die Zeilenstufenform, in die man eine gegebene Matrix A durch elementare Zeilenumformungen überführen kann, eindeutig bestimmt ist. Es kommt also nicht auf die gewählte Folge elementarer Umformungen an.

Satz 2.1 Sei C eine $m \times n$-Matrix. Sind A und B Matrizen in Zeilenstufenform, die aus C durch eine Folge elementarer Zeilenumformungen hervorgehen, so gilt $A = B$.

Beweis : Da man alle elementaren Zeilenumformungen mit elementaren Zeilenverformungen vom selben Typ umkehren kann, müssen wir nur zeigen: Sind A und B zwei Matrizen in Zeilenstufenform, so dass man A mit einer Folge elementarer Zeilenumformungen in B überführen kann, so gilt $A = B$.
Nach Voraussetzung gibt es Elementarmatrizen $L_1, \ldots, L_r$ mit $L_r L_{r-1} \cdots L_1 A = B$. Es sei $L = L_r \cdots L_1$ das Produkt der Elementarmatrizen, d.h. es gilt $LA = B$. Ist $A = 0$, so folgt $B = 0$, also $B = A$. Wir können also $A \neq 0$ annehmen. Da L invertierbar ist, ist dann auch $B \neq 0$. Beide Matrizen enthalten also mindestens einen Pivot. Es seien $(1, j_1), \ldots, (r, j_r)$ die Positionen der Pivots in A, d.h. A hat die Form

$$\left(\begin{array}{ccccccccccc}
0 & \cdots & 0 & 1 & * \ldots * & 0 & * & \cdots & * & 0 & * \\
\vdots & & \vdots & 0 & \vdots & 1 & \vdots & & \vdots & \vdots & \vdots \\
 & & & \vdots & & 0 & & & & & \\
 & & & & & \vdots & & & & \vdots & \\
 & & & & & & & & & 0 & \\
\vdots & & & & & & & & & 1 & \vdots \\
0 & \cdots & & & & & & & & \cdots & 0 \\
\vdots & & & & & & & & & & \vdots \\
0 & \cdots & & & & & & & & \cdots & 0 \\
\underbrace{\qquad}_{j_1 - 1} & & & \uparrow j_1 & & \uparrow j_2 & & & & \uparrow j_r &
\end{array}\right) \begin{array}{l} \Big\} r \\ \\ \Big\} m - r \end{array}$$

Mit $(1, k_1), \ldots, (s, k_s)$ bezeichnen wir die Positionen der Pivots in B. Die ersten $(j_1 - 1)$ Spalten von $B = LA$ sind Nullspalten, da die ersten $(j_1 - 1)$ Spalten von A Nullspalten sind. Wir bezeichnen für alle $\nu \in \{1, \ldots, m\}$ mit e_ν den Spaltenvektor $e_\nu = (0, \ldots, 0, \underbrace{1}_{\nu}, 0, \ldots 0)^t$. Die Spalte j_1 in B ist gerade Le_1. Das kann nicht Null sein, da L invertierbar ist, also ist dies die Spalte mit dem ersten Pivot in B, d.h. $j_1 = k_1$ und $Le_1 = e_1$.
Nun ist aber Le_1 gerade die erste Spalte von L. Also gilt für die Koeffizienten (l_{ij}) von L : $l_{i1} = 0$ für $i > 1$. Wir zeigen nun mit Induktion nach $\nu \in \{1, \ldots, r\}$, dass gilt

i) $j_\nu \leq k_\nu$ oder $\nu > s$.

ii) In den ersten ν Spalten von L sind höchstens die ersten ν Einträge $\neq$ Null, d.h. $l_{ij} = 0$ für $j \leq \nu$ und $i > \nu$.

Den Anfang $\nu = 1$ haben wir gerade erledigt. Also nehmen wir an, i) und ii) gilt für j_ν und zeigen es für $j_{\nu+1}$.
Sind (b_{ij}) die Koeffizienten von B, so sehen die ersten Einträge der $(\nu+1)$-ten Zeile von B folgendermaßen aus:

$$b_{\nu+1j} = ((\nu+1)\text{-te Zeile von } L) \cdot (j\text{-te Spalte von } A).$$

Ist $j < j_{\nu+1}$, so folgt

$$b_{\nu+1j} = (\underbrace{0 \cdots 0}_{\nu} * \cdots *) \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ \vdots \\ 0 \end{pmatrix} \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix}\right\} \nu \quad = 0$$

Da $k_{\nu+1}$ für $\nu + 1 \leq s$ die Position des ersten von Null verschiedenen Eintrags in der $(\nu+1)$-ten Zeile von B ist, folgt $k_{\nu+1} \geq j_{\nu+1}$ oder aber $\nu + 1 > s$, falls die $(\nu+1)$-te Zeile keinen Pivot enthält. Da die $j_{\nu+1}$-te Spalte von A gerade $e_{\nu+1}$ ist, ist die $j_{\nu+1}$-te Spalte von B gleich $Le_{\nu+1}$, also gleich der $(\nu+1)$-ten Spalte von L. Aus $j_{\nu+1} \leq k_{\nu+1}$ folgt, dass die Einträge in dieser Spalte unterhalb der $(\nu+1)$-ten Zeile alle Null sind. Damit ist i) und ii) auch für $\nu + 1$ gezeigt und die Induktion beendet.
Nun entsteht auch A aus B durch elementare Zeilenumformungen, d.h. wir wissen aus Symmetriegründen auch:

$$k_\nu \leq j_\nu \text{ für alle } \nu \leq \min\{r, s\}.$$

Also folgt $k_\nu = j_\nu$ für alle $\nu \leq \min\{r, s\}$., Wir nehmen nun $r \leq s$ an, d.h. die Anzahl der Pivots in A sei $\leq$ der Anzahl der Pivots von B. Ist $s \leq r$, so funktioniert der Rest des Beweises genauso, indem man die Rollen von A und B vertauscht. Dann ist für alle $\nu \leq r$ die j_ν-te Spalte von A und B gerade e_ν. Aus $B = LA$ folgt, dass die ν-te Spalte von L genau derselbe Spaltenvektor ist. Also sieht L so aus:

$$L = \begin{pmatrix} E_r & * \\ 0 & * \end{pmatrix}.$$

Da A genau r Pivots hat, sind die letzten $(m - r)$ Zeilen von A Null, d.h. A hat die Form $\begin{pmatrix} A_1 \\ 0 \end{pmatrix}$ für eine $r \times n$-Matrix A_1. Man kann nun leicht berechnen, dass

$$B = LA = \begin{pmatrix} E_r & * \\ 0 & * \end{pmatrix} \begin{pmatrix} A_1 \\ 0 \end{pmatrix} = \begin{pmatrix} A_1 \\ 0 \end{pmatrix} = A$$

gilt. □

3 Permutationen

Wir wollen uns nun etwas ausführlicher mit der sogenannten Leibniz-Formel oder vollständigen Entwicklung der Determinante beschäftigen.

Definition 3.1 Sei $n \in \mathbb{N}$. Eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ heißt **Permutation** von $\{1, \ldots, n\}$. Die Menge aller Permutationen von $\{1, \ldots, n\}$ nennt man auch „ symmetrische Gruppe“ und bezeichnet sie mit $\mathcal{S}_n$. Eine Permutation τ, die zwei Zahlen $i \neq j$ vertauscht und alle anderen festlässt, heißt **Transposition**. Man schreibt eine Permutation manchmal in der Form $\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$ In der unteren Zeile steht jede der Zahlen $1, \ldots, n$ genau einmal.
Beispiel:

i) $\begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix}$ ist eine Permutation.

ii) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$ ist eine Transposition.

Lemma 3.2 $|\mathcal{S}_n| = n!$.

Beweis : mit Induktion nach n. Offenbar ist $\mathcal{S}_1 = \{\mathrm{id}\}$, also $|\mathcal{S}_1| = 1! = 1$. Angenommen, $|\mathcal{S}_{n-1}| = (n-1)!$ für ein $n \geq 2$. Für jedes $i \in \{1, \ldots, n\}$ bildet die Menge aller Permutationen σ von $\{1, \ldots, n\}$ mit $\sigma(n) = i$ die Menge $\{1, \ldots, n-1\}$ bijektiv auf $\{1, \ldots, n\} \setminus \{i\}$ ab. Also gibt es nach Induktionsvoraussetzung $|\mathcal{S}_{n-1}| = (n-1)!$ viele Permutationen σ mit $\sigma(n) = i$. Da es n Wahlen für die Zahl i gibt, folgt $|\mathcal{S}_n| = n \cdot (n-1)! = n!$. □

Sind $\sigma, \tau : \{1, \ldots, n\} \to \{1, \ldots, n\}$ Permutationen, so ist

$$\begin{aligned} \sigma\tau = \sigma \circ \tau : \{1, \ldots, n\} &\to \{1, \ldots, n\} \\ k &\mapsto \sigma(\tau(k)) \end{aligned}$$

ebenfalls eine Permutation. Wir nennen $\sigma\tau$ auch das Produkt von σ und τ.

Lemma 3.3 Jedes $\sigma \in \mathcal{S}_n$ lässt sich als Produkt von Transpositionen schreiben.

Beweis : Für jede Permutation σ sei $r(\sigma) \geq 0$ diejenige Zahl, für die $\sigma(i) = i$ für $i = 1, \ldots, r(\sigma)$ und $\sigma(r(\sigma)+1) \neq r(\sigma)+1$ ist. Schreiben wir

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix},$$

so ist $r(\sigma)$ die Stelle, bis zu der unten dasselbe wie oben steht. Ist $\sigma(1) \neq 1$, so ist $r(\sigma) = 0$. Wir führen für festes n eine absteigende Induktion nach $r(\sigma)$.
Ist $r(\sigma) = n$, so ist $\sigma = \mathrm{id}$ und somit das leere Produkt von Transpositionen. Das ist der Induktionsanfang. Wir nehmen nun für den Induktionsschluss an, die Behauptung gelte für alle σ mit $r(\sigma) > r$ und zeigen sie für alle σ mit $r(\sigma) = r < n$. Aus $r(\sigma) = r$ folgt $\sigma(r+1) \neq r+1$. Da σ auf $\{1, \ldots, r\}$ die Identität ist, folgt $\sigma(r+1) > r+1$. Ist τ_1 die Vertauschung von $r+1$ und $\sigma(r+1)$, so lässt $\tau_1\sigma$ die Zahlen $1, 2, \ldots, r+1$ unverändert, d.h. es ist $r(\tau_1\sigma) \geq r+1$. Nach Induktionsvoraussetzung ist also $\tau_1\sigma = \tau_2 \cdots \tau_s$ das Produkt von Transpositionen $\tau_2, \ldots, \tau_s$. Da $\tau_1^2 = \mathrm{id}$ ist, folgt hieraus $\sigma = \tau_1\tau_2 \cdots \tau_s$ durch Multiplikation mit τ_1 auf beiden Seiten, d.h. σ ist ebenfalls Produkt von Transpositionen. □

Diese Darstellung von σ als Produkt von Transpositionen ist allerdings nicht eindeutig. Wir könnten etwa an beliebiger Stelle noch $\tau^2 = \mathrm{id}$ für eine Transposition τ einfügen.
Wie wir im Basiskurs gelernt haben (Definition 3.22), ist eine Permutationsmatrix eine Matrix, so dass in jeder Zeile und in jeder Spalte genau eine 1 und sonst nur Nullen stehen.
Jede Permutation $\sigma \in \mathcal{S}_n$ definiert eine $(n \times n)$-Matrix

$$P_\sigma = \sum_{i=1}^{n} e_{\sigma(j)j}$$

Sind p_{ij} die Koeffizienten von P_σ, so gilt also

$$P_{ij} = \begin{cases} 0 & \text{für } \sigma(j) \neq i \\ 1 & \text{für } \sigma(j) = i \end{cases}$$

Offenbar steht hier in allen Spalten genau eine 1 (in Zeile $\sigma(j)$) und sonst Nullen.
Da σ bijektiv ist, existiert nach Lemma 1.1 die Umkehrabbildung $\sigma^{-1} : \{1, \ldots, n\} \to \{1, \ldots, n\}$, die ebenfalls eine Permutation ist. In der i-ten Zeile von P_σ steht somit auch genau eine 1 (nämlich in der Spalte $\sigma^{-1}(i)$) und sonst Nullen. Also ist P_σ eine Permutationsmatrix. Ist σ eine Transposition, so ist P_σ eine Elementarmatrix vom Typ II (Übungsaufgabe).

Wir bezeichnen wieder mit e_j den n-dimensionalen Spaltenvektor, der in der j-ten Zeile den Eintrag 1 hat und sonst nur aus Nullen besteht. Jede Permutationsmatrix P definiert eine Permutation $\sigma = \sigma^P : \{1, \cdots, n\} \to \{1, \ldots, n\}$ durch $Pe_j = e_{\sigma^P(j)}$.(Übungsaufgabe: σ^P ist bijektiv).

Lemma 3.4 Die Zuordnung

$$\begin{array}{rcl} \mathcal{S}_n & \to & \{n \times n\text{-Permutationsmatrizen}\} \\ \sigma & \mapsto & P_\sigma \end{array}$$

ist bijektiv mit Umkehrabbildung $P \mapsto \sigma^P$.

Beweis : Nach Lemma 1.1 müssen wir nur nachrechnen, dass $\sigma^{P\sigma} = \sigma$ für alle $\sigma \in \mathcal{S}_n$ und $P_{\sigma^P} = P$ für alle Permutationsmatrizen P gilt. Nun hat P_σ von links nach rechts gelesen die Spalten $e_{\sigma(1)}, \ldots, e_{\sigma(n)}$. Also ist $P_\sigma(e_j) = e_{\sigma(j)}$, woraus $\sigma^{P_\sigma} = \sigma$ folgt.
Ist P gegeben, so gilt für $\sigma = \sigma^P$, dass $Pe_j = e_{\sigma(j)}$ gilt. Also hat P von links nach rechts gelesen die Spalten $e_{\sigma(1)}, \ldots, e_{\sigma(n)}$. Daher ist $P = P_\sigma$. □

Definition 3.5 Das **Signum** $\text{sgn}(\sigma)$ der Permutation σ ist definiert als die Determinante der zugehörigen Permutationsmatrix:

$$\text{sgn}(\sigma) = \det P_\sigma.$$

Man sieht leicht, dass $P_{\sigma_\tau} = P_\sigma P_\tau$ ist (siehe Proposition 3.25, Basiskurs).
Nach Lemma 3.3 gibt es für jedes $\sigma \in \mathcal{S}_n$ Transpositionen $\tau_1, \ldots, \tau_s$ mit $\sigma = \tau_1 \cdot \ldots \cdot \tau_s$. Dann ist $P_\sigma = P_{\tau_1} \cdots P_{\tau_s}$. Da $P_{\tau_1} \ldots, P_{\tau_s}$ Elementarmatrizen vom Typ II sind, entsteht P_σ durch s Zeilenvertauschungen aus der Einheitsmatrix. Nach Satz 3.12 (Basiskurs) folgt also

$$\text{sgn}(\sigma) = \det P_\sigma = (-1)^s.$$

Dies zeigt, dass zumindest die Parität der Anzahl der Transpositionen, in die man σ zerlegen kann, nur von σ und nicht von der Auswahl der Transpositionen abhängt.
Wir wollen jetzt eine direkte Formel für $\text{sgn}(\sigma)$ beweisen, die ohne den Umweg über Permutationsmatrizen auskommt. Ein Paar (i, j) mit $i, j \in \{1, \ldots, n\}$, so dass $i < j$, aber $\sigma(i) > \sigma(j)$ ist, nennt man einen Fehlstand von σ. Mit $f(\sigma)$ bezeichnen wir die Anzahl der Fehlstände von σ, d.h.

$$f(\sigma) = |\{(i, j) \in \{1, \ldots, n\}^2 : i < j, \sigma(i) > \sigma(j)\}|$$

Satz 3.6 Für alle $\sigma \in \mathcal{S}_n$ gilt

i) $\operatorname{sgn}(\sigma) = (-1)^{f(\sigma)}$

ii) $\operatorname{sgn}(\sigma) = \prod\limits_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i}$

Beweis :

1) Wir zeigen zunächst

$$(-1)^{f(\sigma)} = \prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}.$$

Es gilt

$$\begin{aligned}
\prod_{1\le i<j\le n} (j-i) &\overset{\sigma \text{ bijektiv}}{=} \prod_{\sigma(k)<\sigma(l)} (\sigma(l)-\sigma(k)) \\
&= \prod_{k<l,\sigma(k)<\sigma(l)} (\sigma(l)-\sigma(k)) \cdot \prod_{k>l,\sigma(k)<\sigma(l)} (\sigma(l)-\sigma(k)) \\
&= \prod_{i<j,\sigma(i)<\sigma(j)} (\sigma(j)-\sigma(i)) \cdot \prod_{i<j,\sigma(j)<\sigma(i)} (\sigma(i)-\sigma(j)) \\
&= \prod_{i<j} (-1)^{s_{ij}} (\sigma(j)-\sigma(i)) \text{ mit}
\end{aligned}$$

$$s_{ij} = \begin{cases} 0 & \text{falls } \sigma(i) < \sigma(j) \\ 1 & \text{falls } \sigma(i) > \sigma(j) \end{cases}$$

Also folgt

$$\begin{aligned}
\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i} &= \frac{\prod\limits_{i<j}(\sigma(j)-\sigma(i))}{\prod\limits_{i<j}(j-i)} \\
&= \frac{\prod\limits_{i<j}(\sigma(j)-\sigma(i))}{\prod\limits_{i<j}(-1)^{s_{ij}}(\sigma(j)-\sigma(i))} = \prod_{i<j}(-1)^{s_{ij}} \\
&= (-1)^{f(\sigma)}
\end{aligned}$$

denn $\sum\limits_{i<j} s_{ij}$ zählt, wie oft für $i<j$ die Ungleichung $\sigma(i)>\sigma(j)$ gilt.

2) Sei nun τ eine Transposition. Wir zeigen jetzt, dass $(-1)^{f(r)} = -1$ gilt. τ vertausche die Zahlen i und j mit $i<j$, d.h. es gilt

$$\tau = \begin{pmatrix} 1 & 2 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & 2 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}$$

Die Fehlstände von τ liegen einerseits bei den Paaren $(i, i+1), (i, i+2), \ldots, (i, j)$ und andererseits in $(i+1, j), (i+2, j), \ldots, (j-1, j)$. Das sind zusammen $(j - i) + (j - 1 - i) = 2j - 2i - 1$ Stück, woraus $(-1)^{f(\tau)} = -1$ folgt.

3) Jetzt zeigen wir für alle $\rho, \sigma \in \mathcal{S}_n$, dass $(-1)^{f(\rho\sigma)} = (-1)^{f(\sigma)+f(\sigma)}$ ist. Es gilt

$$\begin{aligned}(-1)^{f(\rho\sigma)} &\overset{1)}{=} \prod_{i<j} \frac{\rho\sigma(j) - \rho\sigma(i)}{j - i} \\ &= \prod_{i<j} \frac{\rho\sigma(j) - \rho\sigma(i)}{\sigma(j) - \sigma(i)} \cdot \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \\ &\overset{1)}{=} (-1)^{f(\rho)+f(\sigma)},\end{aligned}$$

falls $\prod\limits_{i<j} \frac{\rho\sigma(j)-\rho\sigma(i)}{\sigma(j)-\sigma(i)} = \prod\limits_{i<j} \frac{\rho(j)-\rho(i)}{j-i}$ gilt.

Das rechnen wir jetzt nach:

$$\begin{aligned}&\prod_{i<j} \frac{\rho\sigma(j) - \rho\sigma(i)}{\rho(j) - \sigma(i)} \\ = &\prod_{i<j,\sigma(i)<\sigma(j)} \frac{\rho\sigma(j)) - \rho(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{i<j,\sigma(i)>\sigma(j)} \frac{\rho(\sigma(j)) - \rho(\sigma(i))}{\sigma(j) - \sigma(i)} \\ = &\prod_{k<l,\sigma^{-1}(k)<\sigma^{-1}(l)} \frac{\rho(l) - \rho(k)}{l - k} \cdot \prod_{k<l,\sigma^{-1}(k)>\sigma^{-1}(l)} \frac{\rho(k) - \rho(l)}{k - l} \\ = &\prod_{k<l} \frac{\rho(l) - \rho(k)}{l - k}.\end{aligned}$$

4) Sei $\sigma \in \mathcal{S}_n$. Nach Lemma 3.3 ist $\sigma = \tau_1 \cdot \ldots \cdot \tau_s$ mit Transpositionen $\tau_1, \ldots, \tau_s$. Es gilt $\text{sgn}(\sigma) = (-1)^s$, wie wir oben gesehen haben. Nach 2) und 3) gilt

$$(-1)^{f(\sigma)} = (-1)^{f(\tau_1)+\cdots+f(\tau_s)} = (-1)^s = \text{sgn}(\sigma)$$

Das beweist i). Aus 1) folgt damit auch die Formel ii).

□

Die Formeln aus Satz 3.6 liefern eine Möglichkeit, das Signum einer Permutation ohne den Umweg über Permutationsmatrizen zu berechnen. Jetzt wollen wir die Leibniz-Formel für die Determinante beweisen - und zwar auf zwei verschiedenen Wegen.

Satz 3.7 (siehe auch Satz 3.27 (Basiskurs)) Für jede $n \times n$-Matrix A ist

$$\begin{aligned} \det A &= \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)_n} \\ \det A &= \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}. \end{aligned}$$

Es genügt offenbar, eine der beiden Formeln zu beweisen, indem man $\det A = \det A^t$ verwendet.

Beweis (Variante 1) : Wir rechnen nach, dass die Funktion $d : \mathbb{R}^{n \times n} \to \mathbb{R}$, definiert durch $d(A) = \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ die drei Eigenschaften aus Satz 3.16 (Basiskurs) erfüllt:

i) Ist $A = E_n$, so ist $a_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$,

woraus $a_{i\sigma(i)} = 0$ folgt, falls $\sigma(i) \neq i$ ist. Also bleibt nur der Summand für $\sigma = \text{id}$ übrig, und es folgt

$$d(E_n) = \text{sgn}(\text{id}) = 1.$$

ii) Ist $A = \begin{pmatrix} | \\ z_i \\ | \end{pmatrix}, B = \begin{pmatrix} | \\ z_i \\ | \end{pmatrix}$ und $C = \begin{pmatrix} | \\ z_i + \tilde{z}_i \\ | \end{pmatrix}$, so gilt $a_{k\sigma(k)} = b_{k\sigma(k)} = c_{k\sigma(k)}$ für alle $\sigma \in \mathcal{S}_n$ und alle $k \neq i$, sowie $a_{i\sigma(i)} + b_{i\sigma(i)} = c_{i\sigma(i)}$ für alle $\sigma \in \mathcal{S}_n$.

Somit folgt

$$\begin{aligned} d(C) &= \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) c_{1\sigma(1)} \cdots (a_{i\sigma(i)} + b_{i\sigma(i)}) \cdot \ldots \cdot c_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)} = \det A + \det B. \end{aligned}$$

Analog zeigt man $d \begin{pmatrix} | \\ az_i \\ | \end{pmatrix} = a \;\; d \begin{pmatrix} | \\ z_i \\ | \end{pmatrix}$.

Also ist d linear in den Zeilen.

iii) Sind zwei benachbarte Zeilen in $A = (a_{ij})$ gleich, so gilt für einen Zeilenindex i: $a_{ij} = a_{i+1j}$ für alle $j = 1, \ldots, n$. Es sei τ die Transposition, die i und $i + 1$

vertauscht. Mit σ durchläuft auch $\sigma\tau$ die Menge $\mathcal{S}_n$. Also ist

$$\begin{aligned} d(A) &= \sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma\tau) a_{1\sigma\tau(1)}\cdots a_{n\sigma\circ\tau(n)} \\ &\overset{3.6}{=} \sum_{\sigma\in\mathcal{S}_n} -\text{sgn}(\sigma) a_{1\sigma\tau(1)}\cdots a_{n\sigma\tau(n)} \\ &\overset{\tau \text{ ausrechnen}}{=} \sum_{\sigma\in\mathcal{S}_n} -\text{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{i\sigma(i+1)} a_{i+1\sigma(i)}\cdots a_{n\sigma(n)} \\ &= -\sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{i+1\sigma(i+1)} a_{i\sigma(i)}\cdots a_{n\sigma(n)} \\ &= -d(A), \end{aligned}$$

woraus aus $d(A) = 0$ folgt.

Mit Satz 3.16 (Basiskurs) folgt also $d(A) = \det A$ für alle $A \in \mathbb{R}^{n\times n}$. □

Beweis (Variante 2) : Wir benutzen die Entwicklungsformeln für Determinanten und führen Induktion nach n. Wir fixieren ein $j \in \{1,\ldots,n\}$ und definieren eine Abbildung

$$\varphi_j : \mathcal{S}_{n-1} \to \mathcal{S}_n,$$

indem wir einem $\sigma' \in \mathcal{S}_{n-1}$ folgende Permutation $\varphi_j(\sigma')$ zuordnen:

$$\varphi_j(\sigma')(k) = \begin{cases} \sigma'(k) & \text{falls } \sigma'(k) < j, \\ \sigma'(k)+1 & \text{falls } \sigma'(k) \geq j \\ j & k = n \end{cases}$$

(Übungsaufgabe: Das ist wirklich eine Permutation von $\{1,\ldots,n\}$.)
Das Bild von φ_j ist also enthalten in der Menge aller $\sigma \in \mathcal{S}_n$ mit $\sigma(n) = j$. Diese Teilmenge von $\mathcal{S}_n$ hat $(n-1)!$ viele Elemente. Ferner ist φ' offenbar injektiv (Übungsaufgabe), also mit Satz 1.3 auch surjektiv. Für ein beliebiges $\sigma \in \mathcal{S}_n$ mit $\sigma(n) = j$ gibt es also genau ein $\sigma' \in \mathcal{S}_{n-1}$ mit $\varphi_j(\sigma') = \sigma$. Nach Satz 3.6 ist $\text{sgn}(\sigma) = (-1)^{f(\sigma)}$, wobei $f(\sigma)$ die Anzahl der Fehlstände von σ ist. Nun gibt offenbar jeder Fehlstand von σ' einen Fehlstand von σ in der Menge $\{1,\ldots,n-1\}$. Da $\sigma(n) = j$ ist, hat σ außerdem noch Fehlstände in allen Paaren (i,n) mit $i < n$ und $\sigma(i) > j$, d.h. in allen (i,n) mit $i \in \sigma'^{-1}\{j,\ldots,n-1\}$. Davon gibt es $n-j$ Stück. Also ist $f(\sigma) = f(\sigma') + n - j$, woraus nach Satz 3.6

$$\text{sgn}(\sigma) = (-1)^{n-j}\text{sgn}(\sigma')$$

folgt.

Wir nehmen nun an, dass $\sigma = \varphi_j(\sigma')$ ist. Dann gilt für $A = (a_{kl})_{k,l=1\cdots n}$ und $A_{nj} = (b_{kl})_{k,l=1\cdots n-1}$, dass $a_{k\sigma(k)} = b_{k\sigma'(k)}$ für alle $k \neq n$ ist.
Nun zeigen wir $\det(A) = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ mit Induktion nach n.
Der Anfang $n = 1$ ist klar (Übungsaufgabe).
Angenommen, die Behauptung gilt für alle $(n-1) \times (n-1)$-Matrizen. Sei $A = (a_{ij})$ eine $n \times n$-Matrix. Wir entwickeln $\det A$ nach der letzten Zeile (Satz 3.21 Basiskurs) und erhalten

$$\det A = \sum_{j=1}^{n} (-1)^{j+n} a_{nj} \det(A_{nj}).$$

Nach Induktionsvoraussetzung gilt für $A_{nj} = (b_{kl})_{k,l=1,\ldots,n-1}$

$$\begin{aligned} \det(A_{nj}) &= \sum_{\sigma'\in\mathcal{S}_{n-1}} \operatorname{sgn}(\sigma') b_{1\sigma'(1)} \cdots b_{n-1\sigma'(n-1)} \\ &= \sum_{\sigma\in\mathcal{S}_n:\sigma(n)=j} (-1)^{n-j} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n-1\sigma(n-1)} \end{aligned}$$

nach unseren obigen Überlegungen. Also folgt

$$\begin{aligned} \det A &= \sum_{j=1}^{n} (-1)^{j+n} a_{nj} \sum_{\sigma\in\mathcal{S}_n:\sigma(n)=j} (-1)^{n-j} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n-1\sigma(n-1)} \\ &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \end{aligned}$$

□

4 Gruppen und Körper

Wir zeigen zunächst

Lemma 4.1 Eine Menge G mit einer Verknüpfung $m(a,b) = ab$ ist genau dann eine Gruppe, wenn gilt:

i) m ist assoziativ.

ii)′ Es existiert ein $e \in G$ mit $ea = a$ für alle $a \in G$.

iii)′ Für alle $a \in G$ existiert ein $b \in G$ mit $ba = e$.

Beweis : Wir müssen zeigen, dass unter diesen Voraussetzungen die stärkeren Bedingungen aus Definition 4.1 (Basiskurs) gelten. Für jedes $a \in G$ existiert nach iii)' ein „Linksinverses" b mit $ba = e$. Zu b existiert ebenfalls ein „Linksinverses" c mit $cb = e$. Also folgt:

$$ab \overset{ii)'}{=} (ea)b = (cb)ab \overset{i)}{=} c(ba)b = c(eb) \overset{ii)'}{=} cb = e.$$

Also gilt $ab = ba = e$, d.h. es gilt iii) aus Definition 4.1 (Basiskurs.)
Nun sei $a \in G$ und $b \in G$ ein Element mit $ab = ba = e$. Dann gilt:

$$ae = a(ba) \overset{i)}{=} (ab)a = ea \overset{ii)'}{=} a,$$

also ist $ae = ea = a$, d.h. ii) aus Definition 4.1 (Basiskurs) gilt. □

Lemma 4.2 Es sei G eine Gruppe mit Verknüpfung $m(a, b) = ab$. Dann ist das neutrale Element e und für jedes $a \in G$ auch das Element b mit $ab = ba = e$ eindeutig bestimmt.

Beweis : Hat $e' \in G$ ebenfalls die Eigenschaft, dass $ae' = e'a = a$ für alle $a \in G$ gilt, so folgt $e = e'e = e'$, da e ein neutrales Element ist. Also ist e eindeutig bestimmt.
Sei $a \in G$ und seien b und b' zwei Elemente in G mit $ab = ba = e$ und $ab' = b'a = e$. Dann folgt $b = be = b(ab') = (ba)b' = eb' = b'$, also ist das Inverse zu a eindeutig bestimmt. □

Wir wollen nun als Beispiel für den Untergruppenbegriff alle Untergruppen der abelschen Gruppen $(\mathbb{Z}, +)$ bestimmen. Dazu brauchen wir die Division mit Rest.

Lemma 4.3 Es sei $q \in \mathbb{N}$. Jedes Element $a \in \mathbb{Z}$ lässt sich auf eindeutige Weise als

$$a = kq + r$$

mit einem $k \in \mathbb{Z}$ und einem $r \in \{0, 1, \ldots, q-1\}$ darstellen.

Beweis : Wir betrachten zu a und q die Menge $R = \{a - kq : k \in \mathbb{Z}\} \cap \mathbb{N}_0$. Offenbar ist $R \neq \emptyset$. Es sei r die kleinste Zahl in R. Diese ist in $\{0, 1, \ldots, q-1\}$ enthalten, denn sonst liegt $r - q$ auch in R. Also gibt es ein $k \in \mathbb{Z}$ und ein $r \in \{0, 1, \ldots, q-1\}$ mit $a = kq + r$. Ist $a = k'q + r'$ mit $k' \in \mathbb{Z}$ und $r' \in \{0, 1, \ldots, q-1\}$ eine zweite solche Darstellung von a, so folgt $r' = k'q - a \in R$, also $r' \geq r$. Außerdem gilt $0 = (k'q + r') - (kq + r) = (k' - k)q + (r' - r)$, d.h. $r' - r = (k - k')q$. Nun ist $r' - r \in \{0, \ldots, q-1\}$, woraus $r' - r = 0 = (k - k')$ folgt. Also ist $r = r'$ und $k = k'$, d.h. die Darstellung ist eindeutig. □

Satz 4.4 Für jede ganze Zahl b ist die Teilmenge $b\mathbb{Z} := \{bk : k \in \mathbb{Z}\}$ von $\mathbb{Z}$ mit der Verknüpfung $+$ eine Untergruppe von $\mathbb{Z}$. Jede Untergruppe von $(\mathbb{Z}, +)$ ist von der Form $b\mathbb{Z}$ für ein $b \in \mathbb{Z}$.

Beweis : Wir prüfen zunächst, dass $b\mathbb{Z}$ eine Untergruppe ist. Ist bk_1 und $bk_2 \in b\mathbb{Z}$, so ist $bk_1 + bk_2 = b(k_1 + k_2) \in b\mathbb{Z}$. Außerdem ist $0 = b \cdot 0 \in b\mathbb{Z}$ und mit $bk \in b\mathbb{Z}$ liegt auch $-bk = b(-k)$ in $b\mathbb{Z}$. Nach Definition 4.3 (Basiskkurs) ist $b\mathbb{Z}$ also eine Untergruppe von $\mathbb{Z}$.
Nun zeigen wir, dass jede Untergruppe H von dieser Form ist. Da H eine Untergruppe von $(\mathbb{Z}, +)$ ist, gilt $0 \in H$. Ist 0 das einzige Element in H, so ist $H = \{0\} = 0\mathbb{Z}$ von der gewünschten Gestalt.
Enthält H ein $a \neq 0$ aus $\mathbb{Z}$, so ist entweder $a \in \mathbb{N}$ oder $-a \in \mathbb{N}$. Da H mit a auch das Inverse $-a$ enthält, existiert also eine natürliche Zahl $n \in H$. Wir definieren b als die kleinste natürliche Zahl in H und behaupten $H = b\mathbb{Z}$.
„ $\supset$“: Sei $k \in \mathbb{Z}$. Ist $k > 0$, so ist $kb = \underbrace{b + \cdots + b}_{k\text{-mal}} \in H$, da H eine Untergruppe ist. Ist $k = 0$, so ist $kb = 0 \in H$. Ist $k < 0$, so haben wir schon gesehen, dass $(-k)b \in H$ it, also folgt $kb = -(-k)b \in H$.
„ $\subset$“: Sei m eine beliebige Zahl in H. Division mit Rest durch die Zahl $b \in H$ liefert $k \in \mathbb{Z}$ und $r \in \{0, 1, \ldots, b-1\}$ mit

$$m = kb + r.$$

Nun ist $kb \in b\mathbb{Z} \subset H$, wie wir oben gezeigt haben. Also ist $-kb \in H$ und daher ist auch

$$r = m - kb$$

in H. Nun ist $r < b$, wegen der Minimalität von b folgt also $r = 0$. Somit ist $m = kb \in H$. □

Lemma 4.5 Es ist $d\mathbb{Z} = d'\mathbb{Z}$ genau dann, wenn $d = d'$ oder $d = -d'$ ist.

Beweis : Die Richtung „$\Leftarrow$“ ist klar. Wir zeigen „$\Rightarrow$“ d.h. es gelte $d\mathbb{Z} = d'\mathbb{Z}$. Dann ist $d \in d'\mathbb{Z}$, d.h. $d = d'k$ für ein $k \in \mathbb{Z}$. Ferner ist $d' \in d\mathbb{Z}$, d.h. $d' = dl$ für ein $l \in \mathbb{Z}$. Somit folgt $d = d'k = dlk$. Ist $d \neq 0$, so impliziert dies $1 = lk$, d.h. $l, k \in \{\pm 1\}$ und $d = \pm d'$. Ist $d = 0$, so ist auch $d' = dl = 0$, d.h. $d = d'$. □

Die Elemente der Untergruppe $b\mathbb{Z}$ von $\mathbb{Z}$ sind offenbar genau diejenigen ganzen Zahlen, die durch b teilbar sind.

Nun betrachten wir zwei ganze Zahlen a und b, die nicht beide Null sind. Es sei

$$\begin{aligned} a\mathbb{Z} + b\mathbb{Z} &= \{z \in \mathbb{Z} : z = r + s \text{ für } r \in a\mathbb{Z} \text{ und } s \in b\mathbb{Z}\} \\ &= \{z \in \mathbb{Z} : z = ak + bl \text{ für } k, l \in \mathbb{Z}\}. \end{aligned}$$

Diese Menge ist eine Untergruppe von $(\mathbb{Z}, +)$ (Übungsaufgabe). Sie heißt die von a und b erzeugte Untergruppe von $\mathbb{Z}$.

Lemma 4.6 $a\mathbb{Z} + b\mathbb{Z}$ ist die kleinste Untergruppe von $\mathbb{Z}$, die a und b enthält, d.h. ist $H \subset \mathbb{Z}$ eine Untergruppe mit $a \in H$ und $b \in H$, so folgt $a\mathbb{Z} + b\mathbb{Z} \subset H$.

Beweis : Da H eine Untergruppe ist, die a enthält, liegen 0 sowie alle $na = \underbrace{a + \cdots + a}_{n\text{-mal}}$ für $n \in \mathbb{N}$ in H.
Da H mit jedem Element sein Inverses enthält, liegt auch $(-n)a = -na \in H$. Insgesamt gilt also $a\mathbb{Z} \subset H$. Genauso folgt $b\mathbb{Z} \subset H$. Da H abgeschlossen unter $+$ ist, folgt $a\mathbb{Z} + b\mathbb{Z} \subset H$. □

Nach Satz 4.4 gibt es also ein $d \in \mathbb{Z}$, so dass die Untergruppe $a\mathbb{Z} + b\mathbb{Z}$ von $\mathbb{Z}$ mit $d\mathbb{Z}$ übereinstimmt. Da a und b nicht beide Null sind, ist auch $d \neq 0$. Nach Lemma 4.5 gibt es genau ein $d \in \mathbb{N}$ mit

$$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}.$$

Wir wollen diese Zahl d jetzt untersuchen.

Satz 4.7 Es seien $a, b \in \mathbb{Z}$ nicht beide Null und $d \in \mathbb{N}$ die Zahl mit

$$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}.$$

Dann gilt

i) d lässt sich schreiben als $d = ax + by$ für geeignete $x, y \in \mathbb{Z}$

ii) d teilt a und b.

iii) Teilt eine ganze Zahl e beide Zahlen a und b, so teilt e auch d.

Beweis :

i) folgt aus $d \in d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$.

ii) $a \in a\mathbb{Z} \subset a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$, also ist $a = dk$ für ein $k \in \mathbb{Z}$. Daher ist d ein Teiler von a. Genauso zeigt man, dass d ein Teiler von b ist.

iii) Ist e ein Teiler von a und b, so gilt $a = ek$ und $b = el$ für $k, l \in \mathbb{Z}$. Also ist $a \in e\mathbb{Z}$ und $b \in e\mathbb{Z}$, d.h. $e\mathbb{Z}$ ist eine Untergruppe von $\mathbb{Z}$, die a und b enthält. Da $d\mathbb{Z}$ nach Lemma 4.6 die kleinste solche Untergruppe ist, folgt $d\mathbb{Z} \subset e\mathbb{Z}$. Also gilt $d = em$ für ein $m \in \mathbb{Z}$, d.h. e teilt d.

Aufgrund der Eigenschaften ii) und iii) von Satz 4.7 nennen wir die eindeutig bestimmte natürliche Zahl d mit

$$d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$$

auch den **größten gemeinsamen Teiler** von a und b und schreiben $d = \text{ggT}(a, b)$. Nach Satz 4.7 i) kann man den größten gemeinsamen Teiler von a und b „linear aus a und b kombinieren", d.h. es gibt $x, y \in \mathbb{Z}$ mit $d = ax + by$.

Das Problem, zu gegebenen ganzen Zahlen a und b den $\text{ggT}(a, b)$ und die Zahlen x und y zu berechnen, lässt sich mit dem Euklidischen Algorithmus lösen. □

Satz 4.8 (Euklidischer Algorithmus) Es seien $a, b \in \mathbb{Z}$ zwei Zahlen ungleich Null. Wir definieren induktiv eine Folge $(r_k)_{k=1,2\ldots}$ natürlicher Zahlen durch: $r_0 = |a|, r_1 = |b|$. Für alle $k = 1, 2, \ldots$ mit $r_k \neq 0$ sei r_{k+1} der Rest von r_{k-1} bei Division durch r_k, d.h. es gilt $r_{k-1} = q_k r_k + r_{k+1}$ für $r_{k+1} \in \{0, 1, \ldots, r_k - 1\}$ und ein $q_k \in \mathbb{Z}$. Nach endlich vielen Schritten erhält man so erstmals ein $r_n = 0$. Für dieses n gilt

$$\text{ggT}(a, b) = r_{n-1}.$$

Nach Konstruktion ist $r_1 > r_2 > \cdots r_k > \cdots \geq 0$. Also erreichen wir irgendwann ein $r_n = 0$.

Beweis : Wir zeigen induktiv für alle $k \geq 0$ mit $0 \notin \{r_1, \ldots, r_k\}$, dass $\text{ggT}(a, b) = \text{ggT}(r_k, r_{k+1})$ gilt.
Der Induktionsanfang $(k = 0)$ ist klar, da $\text{ggT}(a, b) = \text{ggT}(|a|, |b|)$ gilt. Angenommen, $0 \notin \{r_1, \ldots, r_k\}$ und es gilt nach Induktionsvoraussetzung $\text{ggT}(a, b) = \text{ggT}(r_k, r_{k-1})$. Da $r_{k-1} = q_k r_k + r_{k+1}$ für ein $q_k \in \mathbb{Z}$ gilt, stimmt die Menge der gemeinsamen Teiler von r_k und r_{k-1} mit der Menge der gemeinsamen Teiler von r_k und r_{k+1} überein. Also folgt $\text{ggT}(r_k, r_{k-1}) = \text{ggT}(r_k, r_{k+1})$ und damit die Behauptung.
Wir nehmen nun an, $0 \notin \{r_1, \ldots, r_{n-1}\}$, aber $r_n = 0$. Dann ist $\text{ggT}(a, b) = \text{ggT}(r_{n-1}, r_n) = \text{ggT}(r_{n-1}, 0) = r_{n-1}$, wie behauptet. □

Beispiel: Wir berechnen $\text{ggT}(112, 86)$ mit dem euklidischen Algorithmus. Es ist $r_0 = 112, r_1 = 86, r_2 = 26, r_3 = 8, r_4 = 2, r_5 = 0$. Also folgt $\text{ggT}\ (112, 86) = 2$.

Wir können auch die Koeffizienten x und y in der Linearkombination

$$d = xa + yb$$

berechnen, indem wir den Euklidischen Algorithmus erweitern.

Satz 4.9 (Erweiterter Euklidischer Algorithmus) Seien $a, b \in \mathbb{Z}$ ungleich Null. Ferner nehmen wir $a > 0$ und $b > 0$ an. Es sei $(r_k)_{k=0,\ldots,n}$ die Folge natürlicher Zahlen aus dem Euklidischen Algorithmus, d.h.

$$r_1, \ldots, r_{n-1} \neq 0, r_n = 0, r_{k-1} = q_k r_k + r_{k+1}$$

für $k = 1, \ldots, n-1$, sowie $r_{n-1} = \text{ggT}(a, b)$.
Es sei $x_0 = 1, x_1 = 0, y_0 = 0$ und $y_1 = 1$. Ferner sei für alle $k = 1, \ldots, n-2$

$$\begin{aligned} x_{k+1} &= q_k x_k + x_{k-1} \text{ und} \\ y_{k+1} &= q_k y_k + y_{k-1}. \end{aligned}$$

Dann ist

$$\text{ggT}(a, b) = (-1)^{n-1} x_{n-1} a + (-1)^n y_{n-1} b.$$

Beweis : Wir zeigen mit vollständiger Induktion, dass für alle $k = 0, \ldots, n-1$ gilt:

$$r_k = (-1)^k x_k a + (-1)^{k+1} y_k b.$$

Induktionsanfang ($k = 0$). Es ist $r_0 = |a| = a, x_0 = 1, x_1 = 0$, also stimmt die Behauptung.
Für den Induktionsausschluss nehmen wir an, für ein $k < n-1$ ist $r_e = (-1)^e x_e a + (-1)^{e+1} y_e b$ für alle $0 \leq e \leq k$. Dann ist

$$\begin{aligned} r_{k+1} &= r_{k-1} - q_k r_k \\ &\overset{\text{i.V.}}{=} [(-1)^{k-1} x_{k-1} a + (-1)^k y_{k-1} b] - q_k [(-1)^k x_k a + (-1)^{k+1} y_k b] \\ &= (-1)^{k+1} (x_{k-1} + q_k x_k) a + (-1)^{k+1} (y_{k-1} + q_k y_k) b \\ &= (-1)^{k+1} x_{k+1} a + (-1)^{k+2} y_{k+1} b, \end{aligned}$$

wie behauptet.
Insbesondere gilt $\text{ggT}(a, b) = r_{n-1} = (-1)^{n-1} x_{n-1} a + (-1)^n y_{n-1} b$. □

Beispiel: Für $a = 112$ und $b = 86$ wie oben berechnen wir:

$$\begin{aligned}
112 &= r_0 = 1r_1 + r_2, \quad x_2 = 1, \quad y_2 = 1 \\
86 &= r_1 = 3r_2 + r_3, \quad x_3 = 3, \quad y_3 = 4 \\
26 &= r_2 = 3r_3 + 2, \quad x_4 = 10, \quad y_4 = 13 \\
8 &= r_3 = 4r_4 + 0, \quad x_5 = 43, \quad y_5 = 53 \\
2 &= r_4 \\
0 &= r_5.
\end{aligned}$$

Tatsächlich gilt

$$\begin{aligned}
2 &= x_4 \cdot 112 - y_4 \cdot 86 \\
&= 1120 - 1118.
\end{aligned}$$

Die Bedingung $a > 0$ und $b > 0$ in Satz 4.9 ist keine echte Einschränkung. Es ist nämlich $\text{ggT}(a, b) = \text{ggT}(|a|, |b|)$, und aus

$$\text{ggT}(|a|, |b|) = x|a| + y|b|$$

folgt

$$\text{ggT}(a, b) = \text{ggT}(|a|, |b|) = \left(\frac{|a|}{a}\right) xa + \left(\frac{|b|}{b}\right) yb$$

durch Einfügen der Vorzeichen $\frac{|a|}{a}$ bzw. $\frac{|b|}{b}$.
Der Euklidische Algorithmus ist nützlich, um das multiplikative Inverse einer Zahl in einem endlichen Körper $\mathbb{F}_p$ zu berechnen.
Beispiel: Wir wollen 11^{-1} in $\mathbb{F}_{17}$ berechnen. (Versuchen Sie das mal „ zu Fuß“!).
Aus $1 = \text{ggT}(11, 17) = x11 + y17$ folgt, dass $5^{-1} = \overline{x}$ ist, wobei $\overline{x}$ der Rest von x bei Division durch 17 ist. Wir müssen also nur die Folge der x_k im verallgemeinerten Euklidischen Algorithmus berechnen:

$$\begin{aligned}
r_0 = 11, r_1 = 17, x_0 = 1, x_1 = 0 & \\
r_0 = 11 = 0r_1 + 11, \quad d.h. \quad & q_1 = 0, r_2 = 11, x_2 = 1, \\
r_1 = 17 = 1r_2 + 6, \quad d.h. \quad & q_2 = 1, r_3 = 6, x_3 = 1, \\
r_2 = 11 = 1r_3 + 5, \quad d.h. \quad & q_3 = 1, r_4 = 5, x_4 = 2, \\
r_3 = 6 = 1r_4 + 1, \quad d.h. \quad & q_4 = 1, r_5 = 1, x_5 = 3, \\
r_4 = 4 = 4r_5 + 0, \quad d.h. \quad & r_6 = 0.
\end{aligned}$$

An der Stelle $n = 6$ bricht der Erweiterte Euklidische Algorithmus also ab. Wir erhalten $1 = r_5 = (-1)x_5 5 + y_5 17$, woraus $11^{-1} = \overline{-3} = \overline{14}$ in $\mathbb{F}_{17}$ folgt.
Zum Abschluss dieses Paragraphen wollen wir noch folgende Rechenregeln in beliebigen Körpern beweisen.

Lemma 4.10 Sei K ein beliebiger Körper. Dann gilt für alle $a, b \in K$:

i) $0a = a0 = 0$

ii) $(-1)a = -a$

iii) Ist $ab = 0$, so folgt $a = 0$ oder $b = 0$.

iv) $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a_k b^{n-k}$ für alle $n \in \mathbb{N}$. Hier ist das Produkt der natürlichen Zahl $m = \binom{n}{k}$ mit dem Körperelement $c = a^k b^{n-k}$ wie immer definiert als

$$mc = \underbrace{c + \cdots + c}_{m\text{-mal}}.$$

Beweis :

i) Es ist $0a + 0a = (0+0)a = 0a$, woraus nach Addition von $-0a$ die Gleichung $0a = 0$ folgt. Genauso zeigt man $a0 = 0$.

ii) Es gilt $a + (-1)a = (1-1)a = 0a \overset{i)}{=} 0$, d.h. $(-1)a = -a$.

iii) Sei $ab = 0$ und $a \neq 0$. Dann besitzt a ein multiplikatives Inverses a^{-1}. Durch Multiplikation beider Seiten mit a^{-1} folgt: $b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 \overset{i)}{=} 0$.

iv) Eine leichte Induktion zeigt zunächst für alle $n \in \mathbb{N}$ und $1 \leq k \leq n$:

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

wobei

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} = \frac{n!}{k!(n-k)!} \text{ für } k \geq 1$$

und $\binom{n}{0} = 1$ ist.
Nun zeigen wir die Behauptung mit Induktion nach n.
Für $n = 1$ ist $(a+b) = \binom{1}{0} a^0 b^1 + \binom{1}{1} a^1 b^0$.
Für den Induktionsschritt nehmen wir an

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Dann folgt

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k+1} \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} a^k b^{n+1-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
&= a^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n+1-k} + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{k+1} b^{k+1},
\end{aligned}$$

wie behauptet. □

5 Vektorräume

Wir wollen zunächst die noch offenen Aussagen aus Lemma 5.2 (Basiskurs) beweisen. Diese formulieren wir noch einmal in folgendem Lemma:

Lemma 5.1 Es sei K ein Körper und V ein K-Vektorraum. Dann gilt:

i) Ist 0_V das Nullelement von $(V, +)$, so ist $a0_V = 0_V$ für alle $a \in K$.

ii) Ist $av = 0_V$ für $a \in K$ und $v \in V$, so ist $a = 0$ in K oder $v = 0$ in V.

Beweis :

i) Es ist $a0_V + a0_V = a(0_V + 0_V) = a0_V$, nach Addieren von $-a0_V$ also $a0_V = 0$.

ii) Angenommen, $av = 0_V$ und $a \neq 0$. Dann existiert a^{-1} im Körper K. Wir multiplizieren die Gleichung $0_V = av$ mit a^{-1} und erhalten $a^{-1}0_V = a^{-1}(av) = (a^{-1}a)v = v$. Nach i) ist $a^{-1}0_V = 0_V$, woraus $v = 0_V$ folgt.

□

Im Basiskurs haben wir lineare Abhängigkeit und Unabhängigkeit nur für endliche Mengen von Vektoren eingeführt. Das wollen wir jetzt allgemeiner definieren:

Definition 5.2 Eine beliebige Teilmenge M des K-Vektorraums V heißt **linear unabhängig**, falls jede endliche Teilmenge von M linear unabhängig ist. Also ist eine

beliebige Teilmenge M von V genau dann linear unabhängig, wenn für alle $n \in \mathbb{N}$ und alle $v_1, \ldots, v_n \in M$ gilt: Sind $a_1, \ldots, a_n$ Elemente in K mit

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

so folgt $a_1 = a_2 = \cdots = a_n = 0$.

Definition 5.3 Eine beliebige Teilmenge M von V heißt **linear abhängig**, falls M nicht linear unabhängig ist. M ist also genau dann linear abhängig, wenn es eine endliche Teilmenge von M gibt, die linear abhängig ist.
Dann gilt folgende Verallgemeinerung von Proposition 5.11 (Basiskurs):

Proposition 5.4 Sei $M \subset V$ eine beliebige, linear unabhängige Teilmenge und $v \in V$. Die Menge $M \cup \{v\}$ ist genau dann linear unabhängig, wenn $v \notin < M >$ gilt.
(Man beachte, dass wir die lineare Hülle $< M >$ in Definition 5.6 (Basiskurs) für beliebige Teilmengen M von V definiert haben).

Beweis : Ist $v \in < M >$, so gibt es eine natürliche Zahl n sowie $v_1, \ldots, v_n \in M$ und $a_1, \ldots, a_n \in K$ mit

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

Also ist $a_1 v_1 + \cdots + a_n v_n + (-1)v = 0$ eine Linearkombination der 0, in der nicht alle Koeffizienten Null sind. (Es kommt ja eine (-1) vor). Somit ist $\{v_1, \ldots, v_n, v\}$ eine endliche linear abhängige Teilmenge, d.h. $M \cup \{v\}$ ist linear abhängig.
Umgekehrt nehmen wir an, dass $M \cup \{v\}$ linear abhängig ist. Dann gibt es eine endliche Teilmenge von $M \cup \{v\}$, die linear abhängig ist. Also finden wir ein $n \in \mathbb{N}$ sowie $v_1, \ldots, v_n \in M$, so dass $\{v_1, \ldots, v_n, v\}$ linear abhängig ist. Es gibt daher $a_1, \ldots, a_n, b \in K$, die nicht alle Null sind, mit

$$a_1 v_1 + \cdots + a_n v_n + bv = 0.$$

Da $\{v_1, \ldots, v_n\}$ als endliche Teilmenge von M linear unabhängig ist, folgt $b \neq 0$. Also ist

$$v = -\frac{a_1}{b} v_1 - \ldots - \frac{a_n}{b} v_n \in < M >,$$

wie behauptet. □

Definition 5.5 Eine Teilmenge $M \subset V$ heißt **Basis von** V, falls gilt:

i) M ist linear unabhängig und

ii) $< M >= V$.

Wir wollen nun zeigen, dass auch Vektorräume, die nicht endlich erzeugt sind, eine Basis besitzen.
Nach Satz 5.24 (Basiskurs) ist ein Vektorraum genau dann nicht endlich erzeugt, wenn er eine unendliche, linear unabhängige Teilmenge besitzt.
Um zu zeigen, dass ein solcher Vektorraum auch eine Basis besitzt, benötigen wir das Zorn'sche Lemma (siehe Satz 1.6):
Eine partiell geordnete Menge, in der jede total geordnete Teilmenge eine obere Schranke besitzt, hat ein maximales Element.

Satz 5.6 Sei V ein beliebiger K-Vektorraum. Dann besitzt V eine Basis.

Beweis : Wir wissen, dass in einem endlich erzeugten Vektorraum jede Basis B eine maximale linear unabhängige Teilmenge ist, d.h. gilt $B \subset L$ für eine linear unabhängige Menge L, so folgt $B = L$. In einem beliebigen Vektorraum wollen wir ebenfalls eine Basis als „maximale linear unabhängige Teilmenge" definieren. Mit dem Zorn'schen Lemma zeigen wir, dass das geht.
Dazu sei M die Menge aller linear unabhängigen Teilmengen von V, d.h.

$$M = \{L \subset V : L \text{ linear unabhängig } \}.$$

M ist partiell geordnet bezüglich der Inklusion von Teilmengen, d.h. wir definieren

$$L_1 \leq L_2 \text{ genau dann, wenn } L_1 \subset L_2.$$

Es sei nun $\Sigma \subset M$ eine total geordnete Teilmenge. Für zwei beliebige linear unabhängige Teilmengen L_1, L_2 von V mit $L_1 \in \Sigma$ und $L_2 \in \Sigma$ gilt also $L_1 \subset L_2$ oder $L_2 \subset L_1$. Dann definieren wir S als Vereinigung aller $L \in \Sigma$, d.h.

$$S = \bigcup_{L \in \Sigma} L = \{v \in V : v \in L \text{ für ein } L \in \Sigma\}.$$

Wir behaupten, dass S eine linear unabhängige Teilmenge von V ist. Dazu müssen wir zeigen, dass jede endliche Teilmenge $\{v_1, \ldots, v_r\}$ von S linear unabhängig sind. Nach Konstruktion von S gibt es Mengen $L_1, \ldots, L_r \in \Sigma$ mit $v_1 \in L_1, \ldots, v_r \in L_r$. Wir zeigen nun, dass eine der Mengen L_i alle anderen enthält. Da Σ total geordnet ist, folgt aus $L_j \not\subset L_i$ bereits $L_i \subsetneq L_j$. Angenommen, keine der Mengen L_i enthält alle anderen, dann können wir also induktiv eine Kette $L_1 \subsetneq L_{i_1} \subsetneq L_{i_2} \subsetneq \cdots \subsetneq L_{i_k} \subsetneq \cdots$ definieren, wobei alle $L_{i_k} \in \{L_1, \ldots, , L_r\}$ sind. An einer Stelle dieser Kette müsste sich dann ein Element aus $\{L_1, \ldots, L_r\}$ wiederholen, was zu einem Widerspruch führt. Also gibt es ein i mit $L_j \subset L_i$ für alle $j = 1, \ldots, r$. Dann gilt $\{v_1, \ldots, v_r\} \subset L_i$. Da

L_i linear unabhängig ist, ist auch $\{v_1, \ldots, v_r\}$ linear unabhängig. Also haben wir gezeigt, dass S linear unabhängig ist. Das Element $S \in M$ ist somit eine obere Schranke von Σ.
Nach dem Zorn'schen Lemma besitzt M nun ein maximales Element, d.h. es gibt eine linear unabhängige Teilmenge $L \subset V$, so dass für jede linear unabhängige Teilmenge L' mit $L \subset L'$ schon $L = L'$ folgt.
Wir zeigen nun, dass dieses L eine Basis von V ist. Dazu bleibt zu zeigen, dass $< L >= V$ ist.
Angenommen, $v \in V$ ist ein Element mit $v \notin < L >$. Dann ist nach Proposition 5.4 auch $L \cup \{v\}$ linear unabhängig. Das widerspricht aber der Maximalität von L'. Also gilt $V =< L >$, d.h. L ist in der Tat eine Basis von V. □

Der Beweis von Satz 5.6 beweist zwar die Existenz einer Basis, er gibt aber keine Information darüber, wie eine solche Basis aussieht!
Typische Beispiele für Vektorräume, die nicht endlich erzeugt sind, sind Folgen- oder Funktionenräume.
Beispiel 1: Sei $\mathbb{R}^\infty = \{(a_k)_{k\in\mathbb{N}_o} : a_k \in \mathbb{R}\}$ die Menge aller reellen Folgen. Wir definieren $(a_k)_{k\in\mathbb{N}_0} + (b_k)_{k\in\mathbb{N}_0} = (a_k + b_k)_{k\in\mathbb{N}_0}$ sowie für $c \in K$ eine skalare Multiplikation $c(a_k)_{k\in\mathbb{N}_0} = (ca_k)_{k\in\mathbb{N}_0}$. Mit diesen Verknüpfungen wird $\mathbb{R}^\infty$ ein $\mathbb{R}$-Vektorraum.
Es sei für alle $i \in \mathbb{N}_0$ $e^{(i)}$ die Folge

$$e^{(i)} = (0, \ldots, 0, 1, 0, 0, \ldots) \text{ mit 1 an der } i\text{-ten Stelle},$$

d.h. es ist $e_i^{(i)} = 1$ und $e_k^{(i)} = 0$ für $k \neq i$.
Dann ist die unendliche Menge

$$M = \{e^{(i)} : i \in \mathbb{N}_0\}$$

linear unabhängig in $\mathbb{R}^\infty$.
Ist nämlich $\{e^{(i_1)}, \ldots, e^{(i_n)}\} \subset M$ eine endliche Teilmenge und gilt

$$a_1 e^{(i_1)} + \cdots + a_n e^{(i_n)} = 0$$

für $a_1, \ldots, a_n \in K$, so gilt für die einzelnen Folgenglieder, d.h. für alle $k \in \mathbb{N}_0$:

$$a_1 e_k^{(i_1)} + \cdots + a_n e_k^{(i_n)} = 0.$$

Ist $k = i_1$, so gilt also $0 = a_1 e_{i_1}^{(i_1)} + a_2 e_{i_1}^{(i_2)} + \cdots + a_n e_{i_1}^{(i_n)} = a_1$. Genauso zeigt man $a_2 = \ldots = a_n = 0$.
Da $\mathbb{R}^\infty$ die unendliche, linear unabhängige Teilmenge M besitzt, kann $\mathbb{R}^\infty$ nicht endlich erzeugt sein (siehe Satz 5.24 (Basiskurs)).

Wir zeigen jetzt

$$\langle M\rangle = \{(a_k)_{k\in\mathbb{N}_0} \in \mathbb{R}^\infty : \text{ alle bis auf endlich viele Folgenglieder sind Null }\}.$$

Ist $a = (a_k) \in \langle M\rangle$, so gibt es endlich viele $e^{(i_1)}, \ldots, e^{(i_n)} \in M$ und $c_1, \ldots, c_n \in K$ mit $a = c_1e^{(i_1)} + \cdots + c_ne^{(i_n)}$.
Also ist $a_k = 0$ für alle $k \notin \{i_1, \ldots, i_n\}$, d.h. fast alle Folgenglieder in a sind Null. Ist umgekehrt $a = (a_k)_{k\in\mathbb{N}_0}$ eine Folge, so dass $a_k = 0$ ist für $k \geq k_0$, so ist $a = a_1e^{(1)} + \cdots + a_{k_0}e^{(k_0)} \in \langle M\rangle$.
Die Menge der konvergenten Folgen

$$U_{\text{konv.}} = \{(a_k)_{k\geq 0} \in \mathbb{R}^\infty, (a_k)_k \text{ konvergent}\}$$

und die Menge der Nullfolgen

$$U_{\text{NF}} = \{(a_k)_{k\geq 0} \in \mathbb{R}^\infty : a_k \to 0\}$$

sind Untervektorräume des $\mathbb{R}^\infty$ mit

$$\langle M\rangle \subset U_{\text{NF}} \subset U_{\text{konv.}} \subset \mathbb{R}^\infty.$$

Beispiel 2: Die Menge der Funktionen

$$\mathcal{F} = \{f : \mathbb{R} \to \mathbb{R}\}$$

ist zusammen mit der Addition $(f+g)(x) = f(x)+g(x)$ und der skalaren Multiplikation $(\alpha f)(x) = \alpha f(x)$ ein $\mathbb{R}$-Vektorraum (siehe Übungsaufgabe 1 auf dem Zusatzblatt).
Die Menge der stetigen Funktionen

$$\mathcal{F}_{\text{stet}} = \{f : \mathbb{R} \to \mathbb{R} \text{ stetig }\}$$

sowie die Menge der differenzierbaren Funktionen

$$\mathcal{F}_{\text{diff}} = \{f : \mathbb{R} \to \mathbb{R} \text{ differenzierbar }\}$$

sind Unterräume von $\mathcal{F}$ mit

$$\mathcal{F}_{\text{diff}} \subset \mathcal{F}_{\text{stet}} \subset \mathcal{F}.$$

Wir definieren nun einen weiteren Unterraum von $\mathcal{F}$.

Definition 5.7 Es sei K ein Körper. Der Polynomring $K[X]$ in einer Variablen über K ist definiert als die Menge aller (formalen) Summen der Form

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$$

für $a_0, \ldots, a_n \in K$ und beliebiges $n \in \mathbb{N}_0$.
(„Formal" bedeutet hier, dass wir das $+$ stur hinschreiben und uns über die Bedeutung keine Gedanken machen.)
Zwei solche formalen Summen $f(x) = a_0 + a_1X + \ldots + a_nX^n$ und $g(X) = b_0 + b_1X + \ldots + b_mX^m$ sind gleich, falls für alle $k \leq N = \max\{n, m\}$ gilt: $a_k = b_k$. Falls $n < N$, so setzen wir hier $a_{n+1} = \ldots = a_N = 0$ und analog, falls $m < N, b_{m+1} = \ldots = b_N = 0$.
Für $f(X)$ und $g(X)$ wie oben definieren wir ferner

$$f(X) + g(X) = (a_0 + b_0) + (a_1 + b_1)X + \ldots + (a_N + b_N)X^N,$$

wobei wir wieder die überzähligen Koeffizienten gleich Null gesetzt haben.
Für $c \in K$ definieren wir noch

$$cf(X) = ca_0 + ca_1X + \ldots + ca_nX^n.$$

Beispiel:

i) Für jedes $a \in K$ ist $f(X) = a$ ein Polynom.

ii) $f(X) = 3x^3 + 5x^2 + 7, g(X) = X^{30} - 2$ sind Polynome.

iii) Für jedes $k \in \mathbb{N}_0$ ist $l_k(X) = X^k$ ein Polynom (Dabei ist es $l_0(X) = 1$.)

Lemma 5.8 Mit der oben definierten Addition und skalaren Multiplikation ist $K[X]$ ein K-Vektorraum. Die Teilmenge $B = \{l_k : k \in \mathbb{N}_0\}$ ist eine Basis von $K[X]$.

Beweis : Wir erinnern uns (hoffentlich) an die Definition eines K-Vektorraums in Definition 5.1 (Basiskurs). Es ist leicht nachzurechnen, dass $(K[X], +)$ eine abelsche Gruppe mit neutralem Element $0(X) = 0$ und inversem Element $-f(X) = -a_0 - a_1X - \ldots - a_nX^n$ bildet. Ferner gilt für $f = a_0 + a_1X + \ldots + a_nX$ in $K[X]$ offenbar $1f = f$ und für $a, b \in K$ auch

$$\begin{aligned}(ab)f &= (ab)a_0 + (ab)a_1X + \ldots + (ab)a_nX^n \\ &= a(ba_0 + ba_1X + \ldots + ba_nX^n) \\ &= a(bf).\end{aligned}$$

Die Distributivgesetze $(a+b)f = af + bf$ sowie $a(f+g) = af + ag$ für $g \in K[X]$ und $a, b \in K$ rechnet man ebenfalls sofort nach.
Für jedes Polynom $f(X) = a_0 + a_1X + \ldots + a_nX^n$ ist $f(X) = a_0l_0(X) + a_1l_1(X) + \ldots + a_nl_n(X)$, also liegt f in $< B >$. Somit ist $< B >= K[X]$, d.h.. B ist ein Erzeugendensystem von $K[X]$. Wir prüfen jetzt, dass B auch linear unabhängig ist. Dazu

sei $\{l_{k_1}, \ldots, l_{k_n}\} \subset B$ eine endliche Teilmenge. Gegeben seien $a_1, \ldots, a_n \in K$ mit

$$a_1 l_{k_1} + \ldots + a_n l_{k_n} = 0.$$

Das bedeutet

$$a_1 X^{k_1} + \ldots + a_n X^{k_n} = 0,$$

woraus $a_1 = \ldots = a_n = 0$ folgt, da die $k_1, \ldots, k_n$ paarweise verschieden sind. Also ist B eine Basis von $K[X]$. □

Somit ist $K[X]$ ein Beispiel für einen K-Vektorraum, der eine abzählbar unendliche Basis hat.
Ist $K = \mathbb{R}$, so definiert jedes $f = a_0 + a_1 X + \ldots + a_n X^n \in \mathbb{R}[X]$ eine Abbildung

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ s &\mapsto a_0 + a_1 s + \ldots + a_n s^n. \end{aligned}$$

Diese Abbildung ist differenzierbar (Analysis), also ist $\mathbb{R}[X] \subset \mathcal{F}_{\text{diff}}$. Man rechnet leicht nach, dass $\mathbb{R}[X]$ ein Untervektorraum von $\mathcal{F}_{\text{diff}}$ ist.
Also haben wir eine Kette von $\mathbb{R}$-Vektorräumen

$$\mathbb{R}[X] \subset \mathcal{F}_{\text{diff}} \subset \mathcal{F}_{\text{stet}} \subset \mathcal{F},$$

wobei schon der kleinste Vektorraum $\mathbb{R}[X]$ nicht endlich erzeugt ist. Daher sind auch die $\mathbb{R}$-Vektorräume $\mathcal{F}_{\text{diff}}$, $\mathcal{F}_{\text{stet}}$ und $\mathcal{F}$ nicht endlich erzeugt.

6 Ringe

Am Ende des letzten Paragraphen haben wir für jeden Körper K den Polynomring $K[X] = \{a_0 + a_1 X + \cdots + a_n X^n : n \in \mathbb{N}_0, a_i \in K\}$ kennengelernt. $K[X]$ ist eine abelsche Gruppe bezüglich der Addition von Polynomen.
Für $f(X) = \sum_{i=0}^{n} a_i X^i$ und $g(X) = \sum_{j=0}^{m} b_j X^j$ definieren wir nun das Produkt $fg \in K[X]$ durch

$$fg(X) = \sum_{k=0}^{m+n} [\sum_{l=0}^{k} (a_l + b_{k-l})] X^k.$$

Dieses Produkt entsteht, indem man $f(X)g(X)$ formal ausmultipliziert und dann nach Potenzen von X sortiert.
Dann rechnet man leicht nach, dass $fg = gf$ sowie $f(gh) = (fg)h$ und $1f = f1 = f$ für das Polynom $a_0 = 1$ ist. Mit ein bisschen Schreibarbeit zeigt man außerdem für

$f, g, h \in K[X]$, dass man in bekannter Weise Klammern auflösen kann:

$$(f + g)h = fh + gh.$$

Also ist $K[X]$ ein „kommutativer Ring mit 1“, was in folgender Definition erklärt wird:

Definition 6.1 Ein **Ring** ist eine Menge R mit zwei Verknüpfungen

$$\begin{aligned} + : \quad & R \times R \to R \text{ und} \\ \cdot : \quad & R \times R \to R, \end{aligned}$$

so dass gilt:

i) $(R, +)$ ist eine abelsche Gruppe.

ii) Die Multiplikation $\cdot$ ist assoziativ und besitzt das neutrale Element 1, d.h. für alle $r \in R$ ist $1r = r1 = r$.

iii) Es gelten die Distributivgesetze, d.h. für alle $r, s, t \in R$ ist

$$\begin{aligned} r(s + t) &= rs + rt \text{ und} \\ (r + s)t &= rt + st. \end{aligned}$$

R heißt **kommutativer Ring** mit 1, falls zusätzlich die Multiplikation kommutativ ist.

Beispiel: $\mathbb{Z}$ ist ein kommutativer Ring mit 1.

In Lemma 4.3 haben wir gesehen, dass es in dem Ring $\mathbb{Z}$ eine Division mit Rest gibt. Ein analoges Resultat gilt im Polynomring $K[X]$.

Definition 6.2 Der **Grad** des Polynoms $0 \neq f(X) = a_0 + a_1X + \cdots + a_nX^n$ ist die größte Zahl k mit $a_k \neq 0$. Wir bezeichnen den Grad von f mit $\text{grad}(f)$ (in der englischen Literatur $\deg(f)$ für degree). Wir setzen $\text{grad}(0) = -\infty$.

Für $d = \text{grad}(f)$ können wir f also schreiben als $f = a_0 + a_1X + \cdots + a_dX^d$. Der „höchste“ Koeffizient a_d von f heißt auch **Leitkoeffizient**. Ist $a_d = 1$, so heißt f **normiert**.

Satz 6.3 Es sei $g(X) \in K[X]$ ein Polynom vom Grad $d \geq 0$. Dann gibt es für jedes $f \in K[X]$ eindeutig bestimmte Polynome $q(X)$ und $r(X) \in K[X]$ mit grad $r < d$, so dass

$$f = qg + r$$

gilt.

Beweis : Wir zeigen die Behauptung mit Induktion nach grad(f). Ist grad$(f) < d$, so können wir $q(X) = 1$ und $r(X) = 0$ wählen.
Ist $n = \text{grad}(f) \geq d$, so nehmen wir an, die Behauptung gilt für alle Polynome vom Grad $< n$. Es sei $f(X) = a_0 + a_1X + \cdots + a_nX^n$ und $g(X) = b_0 + b_1X + \cdots + b_dX^d$ mit $a_n \neq 0$ und $b_d \neq 0$.
Dann hat das Polynom

$$f(X) - \frac{a_n}{b_d}X^{n-d}g(X)$$

einen grad $< n$. Nach Induktionsvoraussetzung gilt also

$$f(X) - \frac{a_n}{b_d}X^{n-d}g(X) = \tilde{q}(X)g(X) + r(X)$$

für $\tilde{q}, r \in K[X]$ mit grad $r < d$.
Also gilt $f(X) = q(X)g(X) + r(X)$ für $q(X) = \tilde{q}(X) + \frac{a_n}{b_d}X^{n-d}$. Damit ist die Existenz der verlangten Gleichung gezeigt. Wir müssen nun noch die Eindeutigkeit nachweisen.
Gilt $f = qg + r = q'g + r'$ für r und r' vom Grad $< d$, so folgt

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

Die linke Seite ist also ebenfalls Polynom vom Grad $< d$. Da grad $g = d$ ist, folgt $q = q'$. Also gilt auch $r = r'$. □

Wir wollen dieses Resultat nun in einen allgemeineren Rahmen einordnen.

Definition 6.4 Ein kommutativer Ring R mit 1 heißt **Integritätsring**, wenn R keine Nullteiler $\neq 0$ besitzt, d.h. wenn aus $fg = 0$ in R bereits $f = 0$ oder $g = 0$ folgt.

Beispiel: $\mathbb{Z}$ und $K[X]$ sind Integritätsringe.

Definition 6.5 Ein Integritätsring R mit 1 heißt **euklidisch**, wenn es eine Funktion $\sigma : R \setminus \{0\} \to \mathbb{N}_0$ gibt, so dass für alle $g \neq 0$ aus R und alle $f \in R$ Elemente $q, r \in R$ mit $r = 0$ oder $\sigma(r) < \sigma(g)$ existieren, so dass

$$f = qg + r$$

gilt.

$\mathbb{Z}$ mit $\sigma = |\ |$ und $K[X]$ mit $\sigma =$ grad sind Beispiele für euklidische Ringe.
Euklidische Ringe sind also Ringe, in denen es eine Division mit Rest gibt.
Beispiel: Wir teilen $f(X) = 4X^5 + 4X^4 + 2X^2 + 1$ durch $g(X) = 2X^3 + X^2 - 1$ mit Rest.

Der Beweis von Satz 6.3 gibt uns dafür ein Verfahren an die Hand. Da grad$(g) <$ grad(f) ist, berechnen wir

$$\begin{aligned} f(X) - 2X^2 g(X) &= 4X^5 + 4X^4 + 2X^2 + 1 - 4X^5 - 2X^4 + 2X^2 \\ &= 2X^4 + 4X^2 + 1. \end{aligned}$$

Der Grad dieses Polynoms ist immer noch $>$ grad$g$, also berechnen wir

$$\begin{aligned} 2X^4 + 4X^2 + 1 - Xg(X) &= 2X^4 + 4X^2 + 1 - 2X^4 - X^3 + X \\ &= -X^3 + 4X^2 + X + 1. \end{aligned}$$

Und noch einmal:

$$\begin{aligned} -X^3 + 4X^2 + X + 1 + \frac{1}{2}g(X) &= -X^3 + 4X^2 + X + 1 + X^3 + \frac{1}{2}X^2 - \frac{1}{2} \\ &= \frac{9}{2}X^2 + X + \frac{1}{2}. \end{aligned}$$

Also ist $f(X) = (2X^2 - X + \frac{1}{2})g(X) + (\frac{9}{2}X^2 + X + \frac{1}{2})$.

Wir haben in § 5 schon gesehen, dass der Polynomring $K[X]$ über einem Körper K ein K-Vektorraum ist. Einen solchen Ring, der auch noch ein K-Vektorraum ist, nennt man K-Algebra.

Definition 6.6 Es sei K ein Körper. Ein Ring R mit 1, der außerdem eine skalare Multiplikation $K \times R \to R$ besitzt, so dass

i) $(R, +)$ mit dieser skalaren Multiplikation ein K-Vektorraum ist und

ii) $a(fg) = (af)g = f(ag)$ für alle $a \in K$, und $f, g \in R$ gilt,

heißt K-**Algebra** (genauer gesagt K-Algebra mit 1).

Beispiel 1: $K[X]$ ist eine K-Algebra.
Beispiel 2: Für alle $n \in \mathbb{N}$ ist $K^{n\times n}$ mit der Addition, Multiplikation und skalaren Multiplikation von Matrizen eine K-Algebra.

Der Polynomring ist eine besonders ausgezeichnete K-Algebra, wie wir jetzt erklären wollen.

Betrachtet man Abbildungen zwischen Strukturen wie Gruppen, Ringen, Vektorräumen oder Algebren, so nennt man eine solche Abbildung einen Homomorphismus, wenn sie mit den jeweiligen Rechenstrukturen verträglich ist. Genauer gesagt:

Definition 6.7

i) Eine Abbildung $f : G \to G'$ zwischen zwei (additiv geschriebenen) Gruppen G und G' heißt **Gruppenhomomorphismus**, falls

$$f(a+b) = f(a) + f(b)$$

für alle $a, b \in G$ gilt.

ii) Eine Abbildung $f : R \to R'$ zwischen zwei Ringen R und R' heißt **Ringhomomorphismus**, falls für alle $a, b \in R$

$$\begin{aligned} f(a+b) &= f(a) + f(b), \\ f(ab) &= f(a)f(b) \\ \text{sowie } f(1) &= 1 \end{aligned}$$

gilt.

iii) Eine Abbildung $f : A \to A'$ zwischen zwei K-Algebren A und A', die sowohl ein Ringhomomorphismus als auch eine lineare Abbildung ist, nennt man K-Algebrenhomomorphismus.

iv) Eine lineare Abbildung zwischen zwei Vektorräumen nennt man analog auch Homomorphismus von Vektorräumen.

Ein Gruppenhomomorphismus $f : G \to G'$ bildet automatisch das neutrale Element $0 \in G$ auf das neutrale Element $0 \in G'$ ab, denn es gilt $f(0) = f(0+0) = f(0) + f(0)$, woraus nach Addition von $-f(0)$ sofort $f(0) = 0$ folgt. Ebenso kann man zeigen, dass für alle $a \in G \;\; f(-a) = -f(a)$ gilt, d.h. dass f mit Inversenbildung vertauscht.

Nun gilt für den Polynomring:

Satz 6.8 Es sei A eine beliebige K-Algebra. Zu jedem $t \in A$ gibt es genau einen K-Algebrenhomomorphismus

$$\varphi : K[X] \to A$$

mit $\varphi(X) = t$. Es ist

$$\varphi(\sum_{i=0}^{n} a_i X^i) = \sum_{i=0}^{n} a_i t^i.$$

Wir nennen φ auch Einsetzungshomomorphismus.

Beweis : Wir definieren wie angegeben

$$\varphi(\sum_{i=0}^{n} a_i X^i) = \sum_{i=0}^{n} a_i t^i$$

für alle Polynome $\sum\limits_{i=0}^{n} a_i X^i \in K[X]$.
Hier wird $\sum\limits_{i=0}^{n} a_i t^i$ in der K-Algebra A ausgerechnet. Man rechnet leicht nach, dass φ mit der Addition, der Multiplikation und der skalaren Multiplikation vertauscht und dass $\varphi(1) = 1$ gilt. Also ist φ ein K-Algebrenhomomorhpismus mit $\varphi(X) = t$. Ist ψ ein weiterer K-Algebrenhomomorphismus mit $\psi(X) = t$, so gilt

$$\psi\left(\sum_{i=0}^{n} a_i X^i\right) = \sum_{i=0}^{n} \psi(a_i X^i) = \sum_{i=0}^{n} a_i \psi(X^i) = \sum_{i=0}^{n} a_i \psi(X)^i = \sum_{i=0}^{n} a_i t^i = \varphi\left(\sum_{i=0}^{n} a_i X\right)$$

Also ist φ eindeutig bestimmt. □

Beispiel: Sei $c \in K$. Dann definiert c einen K-Algebrenhomomorphismus

$$\begin{aligned} K[X] &\to K \\ f(X) = \sum_{i=0}^{n} a_i X^i &\mapsto \sum_{i=0}^{n} a_i c^i = f(c). \end{aligned}$$

Das Polynom $f(X)$ wird also auf die reelle Zahl abgebildet, die sich durch Einsetzen von c in f ergibt.
Eine Zahl $c \in K$ heißt **Nullstelle** von f, falls $f(c) = 0$ gilt.

Lemma 6.9 Sei $f \in K[X]$ und a eine Nullstelle von f. Dann gibt es ein $q \in K[X]$ mit

$$f(X) = (X - a)q(X).$$

Beweis : Wir teilen f durch $(X-a)$ mit Rest und erhalten $f(X) = (X-a)q(X)+r(X)$ für $q, r \in K[X]$ mit grad$r < 1$. Also ist $r(X) = c$ für ein $c \in K$. Da $0 = f(a) = r(a)$ ist, folgt $r = 0$. □

Gilt in einem Integritätsring R die Gleichung $f = pq$, so heißen p und q **Teiler** von f. Wir schreiben auch $p|f$ bzw. $q|f$.

Definition 6.10 Sei $f \in K[X]$. Für $a \in K$ sei $\text{ord}_a(f)$ definiert als die größte Zahl $k \in \mathbb{N}_0$, so dass $(X-a)^k | f$ gilt. Wir nennen $\text{ord}_a(f)$ die **Vielfachheit** der Nullstelle a von f.
Aus Gradgründen gilt $\text{ord}_a(f) \leq n$. Ist $f(a) \neq 0$, so ist $ord_a(f) = 0$. Offenbar ist $\text{ord}_a(fg) = \text{ord}_a(f) + \text{ord}_a(g)$.

Satz 6.11 Sei $f \neq 0$ in $K[X]$. Seien $a_1, \ldots, a_r$ paarweise verschiedene Nullstellen von f mit $k_i = \text{ord}_{a_i}(f)$. Dann gilt

$$f(x) = (X - a_1)^{k_1} \cdot \ldots \cdot (X - a_r)^{k_r} g(X)$$

mit einem $g \in K[X]$, so dass $g(a_i) \neq 0$ ist für $i = 1, \ldots, r$.

Beweis : Mit Induktion nach r.
Für $r = 1$ folgt durch wiederholtes Anwenden von Lemma 6.9, dass $f(X) = (X - a_1)^m \cdot g(X)$ für ein $m \geq 0$ und ein $g \in K[X]$ mit $g(a_1) \neq 0$ ist. Dann ist $\text{ord}_{a_1}(f) = m$, denn sonst hätten wir eine Gleichung der Form $f(X) = (X - a_1)^{m+1} h(X) = (X - a_1)^m g(X)$ für ein $h \in K[X]$. Da $K[X]$ nullteilerfrei ist, folgte $(X - a_1)h(X) = g(X)$, was $g(a_1) \neq 0$ widerspricht.
Für den Induktionsschluss können wir annehmen, dass $f(X) = (X - a_1)^{k_1} \cdot \ldots \cdot (X - a_{r-1})^{k_{r-1}} h(X)$ für $k_i = \text{ord}_{a_i}(f)$ und $h \in K[X]$ gilt.
Wir wenden den Induktionsanfang auf h und a_r an. Dann ist $h(X) = (X - a_r)^{k_r} g(X)$ für ein $g \in K[X]$ mit $k_r = \text{ord}_{a_r}(h) = \text{ord}_{a_r}(f)$ und $g(a_r) \neq 0$. Daraus folgt die gewünschte Darstellung $f(X) = (X - a_1)^{k_1} \cdot \ldots \cdot (X - a_r)^{k_r} g(X)$. □

Wir sagen, $f \neq 0$ zerfällt vollständig in Linearfaktoren über K, wenn sich f in $K[X]$ schreiben lässt als $f(X) = c(X - a_1) \cdot \ldots \cdot (X - a_n)$ mit $c \in K$ und nicht notwendig verschiedenen $a_i \in K$.
Dann ist $n = \text{grad}(f)$ und c der Leitkoeffizient.
Beispiel: $(X^3 - 1) = (X - 1)(1 + X + X^2)$ zerfällt über $\mathbb{R}$ nicht vollständig in Linearfaktoren (wohl aber über $\mathbb{C}$).
$(X^2 - 2) = (X - \sqrt{2})(X + \sqrt{2})$ zerfällt über $\mathbb{R}$ vollständig in Linearfaktoren, nicht über über $\mathbb{Q}$.

Korollar 6.12 Ein Polynom $f \neq 0$ in $K[X]$ vom Grad n hat höchstens n verschiedene Nullstellen in K.

Beweis : Das folgt sofort aus Satz 6.11. □

Beispiel:

i) $(X^2 + 1) \in \mathbb{R}[X]$ hat keine Nullstelle in $\mathbb{R}$.

ii) $(X^2 - 1) \in \mathbb{R}[X]$ hat zwei Nullstellen in $\mathbb{R}$.

Später werden wir den sogenannten **Fundamentalsatz der Algebra** kennenlernen, der besagt, dass jedes Polynom $f \neq 0$ in $\mathbb{C}[X]$ eine Nullstelle besitzt. Einen Körper mit dieser Eigenschaft nennt man übrigens **algebraisch abgeschlossen**.
Jetzt wollen wir Satz 6.8 benutzen, um Matrizen in Polynome einzusetzen.

Definition 6.13 Sei R ein kommutativer Ring mit 1 und sei $C = (c_{ij})_{i,j=1,\ldots,n} \in R^{n\times n}$ eine $(n \times n)$-Matrix mit Koeffizienten $a_{ij} \in R$.
Dann definieren wir die Determinante von C durch die Leibnizformel:

$$\det C = \sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma) c_{1\sigma(1)} \cdot \ldots \cdot c_{n\sigma(n)} \in R.$$

Auch für Determinanten von Matrizen in $R^{n\times n}$ gilt

i) det ist linear in den Zeilen und Spalten (vgl. Satz 3.5 (Basiskurs) und Korollar 3.20 (Basiskurs)).

ii) $\det(C) = 0$, falls C eine Nullzeile enthält (vgl. Korollar 3.6 (Basiskurs)).

iii) $\det(C) = -\det C'$, falls C' durch Vertauschen zweier Zeilen oder Spalten aus C hervorgeht (vgl. Satz 3.12 und Korollar 3.20 (Basiskurs)).

iv) $\det C = \det C^t$

v) die Entwicklungsformeln Satz 3.21 (Basiskurs).

Die Beweise kann man ähnlich wie über K führen.
Bei Korollar 3.18 (Basiskurs) und auch bei elementaren Umformungen vom Typ III muss man allerdings etwas aufpassen. Es gilt stattdessen: $C \in R^{n\times n}$ ist invertierbar genau dann, wenn $\det(C)$ invertierbar in R ist, d.h. wenn es ein $r \in R$ mit $\det(C)r = 1$ gibt.
Wir betrachten nun für $A = (a_{ij}) \in K^{n\times n}$ die Matrix

$$XE_n - A = \begin{pmatrix} X - a_{11} & -a_{12} & \ldots & a_{1n} \\ -a_{21} & X - a_{22} & \ldots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \ldots & X - a_{nn} \end{pmatrix} \in (K[X])^{n\times n}$$

und setzen $\chi_A(X) = \det(XE_n - A)$ (vgl. Definition 7.10 (Basiskurs)). Nach der Leibnizformel ist das ein Polynom in X, d.h. es ist $\chi_A(X) \in K[X]$.
Wir nennen $\chi_A(X)$ das **charakteristische Polynom** von A.

Lemma 6.14 Für alle $\alpha \in K$ ist $\chi_A(\alpha) = \det(\alpha E_n - A)$.

Beweis : $\chi_A(\alpha)$ ist das Bild von $\chi_A(X)$ unter dem Einsetzungshomomorphismus $\varphi : K[X] \to K$, der X auf α abbildet. Es ist $\chi_A(X) = \det(XE_n - A) = \sum_\sigma \text{sgn}(\sigma) b_{1\sigma(1)}(X) \cdots b_{n\sigma(n)}(X)$ für $XE_n - A = (b_{ij}(X))_{i,j}$ mit $b_{ij}(X) \in K[X]$. Da φ ein Ringhomomorphismus ist, folgt $\chi_A(\alpha) = \sum_\sigma \text{sgn}(\sigma) b_{1\sigma(1)}(\alpha) \cdots b_{n\sigma(n)}(\alpha) = \det(\alpha E_n - A)$. □

Nach Proposition 7.11 (Basiskurs) ist $\chi_A(\alpha) = 0$ genau dann, wenn α ein Eigenwert von A ist.
Wir betrachten nun für $A \in K^{n\times n}$ den Einsetzungshomomorphismus

$$\varphi : K[X] \mapsto K^{n\times n},$$

der nach Satz 6.8 durch $\varphi(X) = A$ definiert wird. Wir schreiben auch suggestiv $f(A) = \varphi(f)$.

Satz 6.15 (Satz von Cayley-Hamilton) Es ist $\chi_A(A) = 0$.

Beweis : $\chi_A(A)$ ist eine $n \times n$-Matrix über K. Es genügt zu zeigen, dass Kern $(\chi_A(A)) = K^n$ gilt. Sei $v \in K^n$ ein Vektor $\neq 0$. Wir betrachten $v, Av, A^2v, \ldots$ und wählen r minimal, so dass $(v, Av, \ldots, A^rv)$ linear unabhängig, aber $(v, Av, \ldots, A^{r+1}v)$ linear abhängig ist. Dann ist $b_0 = v, b_1 = Av, \ldots, b_r = A^rv$ eine Basis von $U = \langle v, Av, \ldots, A^rv\rangle$. Diese ergänzen wir zu einer Basis $B = (b_0, \ldots, b_r, b_{r+1}, \ldots, b_n)$ von K^n. Die lineare Abbildung

$$\begin{aligned} \lambda_A =: \lambda : K^n &\mapsto K^n \\ v &\mapsto Av \end{aligned}$$

erfüllt offenbar $\lambda(U) \subset U$. Also hat die Koordinatenmatrix $A_{\lambda,B,B}$ von λ bezüglich B die Gestalt

$$A_{\lambda,B,B} = \begin{pmatrix} C & * \\ 0 & D \end{pmatrix}$$

mit der $(r+1)\times(r+1)$-Matrix

$$C = \begin{pmatrix} 0 & 0 & & 0 & c_0 \\ 1 & 0 & & \vdots & c_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & & & 1 & c_r \end{pmatrix}$$

wobei $A^{r+1}v = c_0v + c_1Av + \cdots + c_rA^rv$ gilt, und einer $(n-r-1)\times(n-r-1)$-Matrix D.
Es ist

$$\chi_C(X) = \det(XE_n - C) = X^{r+1} - c_rX^r \cdots - c_1X - c_0,$$

wie man durch Entwickeln nach der letzten Spalte errechnet. Also ist

$$\chi_C(A)v = A^{r+1}v - c_rA^rv \cdots - c_1Av - c_0 = 0,$$

woraus nach Übungsaufgabe 20

$$\chi_A(A)v = (\chi_D(A) \cdot \chi_C(A))v = 0$$

folgt. □

Definition 6.16 Das Minimalpolynom $p_A(X) \in K[X]$ einer Matrix $A \in K^{n\times n}$ ist definiert als das normierte Polynom kleinsten Grades mit $p_A(A) = 0$.
Dann teilt $p_A(X)$ in $K[X]$ jedes Polynom $q(X)$ mit $q(A) = 0$, nach Satz 6.15 also insbesondere $\chi_A(X)$. Das zeigt man genau wie Lemma 6.9.