LINEARE ALGEBRA UND ANALYTISCHE GEOMETRIE I (WS 09/10)

BERNHARD HANKE

Die lineare Algebra ist neben der Analysis die zweite wichtige Säule der mathematischen Grundausbildung. Ihre Konzepte, Begriffe und Methoden durchziehen die gesamte moderne Mathematik.

Wir beginnen die Vorlesung mit einer kurzen Einführung in die heute übliche mengentheoretische Sprache. Nach einem Abriss der wichtigsten algebraischen Grundstrukturen Gruppen, Ringe und Körper stellen wir den Aufbau des Zahlensystems vor.

Unser weiteres Vorgehen wird von einer der zentralen Fragestellungen der linearen Algebra motiviert: Die Beschreibung der Struktur der Lösungsmengen linearer Gleichungssysteme.

Die dazu entwickelte Theorie besteht aus der Untersuchung von Vektorräumen einerseits und der linearen Abbildungen zwischen diesen andererseits. Letzteres führt auf den Matrizenkalkül, ein zentrales Thema dieser Vorlesung.

Danach diskutieren wir Determinanten. Diese sind durch den elementargeometrischen Begriffs des Volumens eines Parallelogramms motiviert.

Wir beschließen diese Vorlesung mit einer ausführlichen Behandlung der Polynomringe, die später in der Eigenwerttheorie wichtig werden, und dem Beweis des Fundamentalsatzes der Algebra, einem der wichtigsten Resultate der gesamten Mathematik.

Die lineare Algebra schlägt auch eine Brücke von der Algebra zur Geometrie. Auf diesen Zusammenhang werden wir in der Vorlesung immer wieder stoßen.

Wir beziehen uns in dieser Vorlesung an verschiedenen Stellen auf die Quelle

Gerd Fischer, *lineare Algebra und analytische Geometrie*, 15. Auflage, Vieweg-Verlag,

die wir mit [Fischer] abkürzen.

0. Formale Grundlagen der Mathematik

Logik und Mengenlehre bilden das Fundament der modernen Mathematik. Die Logik stellt den formalen Rahmen (=Sprache) bereit, (mathematische) Aussagen zu formulieren und neue Aussagen aus bereits gegebenen abzuleiten. Die Mengenlehre hingegen formalisiert, beschreibt und untersucht die

grundlegendste Struktur der Mathematik, gefasst im Begriff der Menge. Deren einziges Strukturdatum ist die Beziehung „x ist Element der Menge A“. Alle anderen Objekte der Mathematik lassen sich aus dem Mengenbegriff mit Hilfe formaler Definitionen entwickeln.

Die Mengenlehre in ihrer heutigen Form basiert auf einem von Zermelo und Fraenkel stammenden, je nach Zählung etwa acht Aussagen umfassenden Axiomensystem (Axiom=unbewiesene grundlegende Aussage, die am Anfang einer Theorie steht), das präzise festlegt, welche Eigenschaften Mengen haben sollen.

Der berühmte zweite Unvollständigkeitssatz von K. Gödel besagt, dass mit den Mitteln der Mathematik weder gezeigt noch widerlegt werden kann, ob es überhaupt eine mathematische Struktur (=Mengenlehre) gibt, die diese Axiome erfüllt. Dies wird aber von niemandem in Zweifel gezogen. Insofern reichen die Grundlagen der Mathematik über die Mathematik selbst hinaus und berühren Gebiete wie Philosphie und Wissenschaftstheorie, aber auch Kognitionsforschung.

Wir wollen in diesem Kapitel kurz auf die durch die Mengentheorie geprägte Sprache der Mathematik eingehen. Zunächst werden wir aber darüber nachdenken, was wir unter mathematischen Aussagen überhaupt verstehen wollen. Hier beschränken wir uns auf eine informelle Einführung und überlassen die genaue Diskussion formaler Sprachen anderen Vorlesungen.

Wir verstehen unter einer *Aussage* eine (evtl. umgangssprachliche) Formulierung, der wir - wenigstens prinzipiell - in eindeutiger Weise einen Wahrheitswert „wahr (w)“ oder „falsch (f)“ zuordnen können (wir gehen also hier von einer klassischen zweiwertigen Logik aus). Beispielsweise sind „Die Zahl 8 ist gerade“ und „es gibt außerirdisches Leben“ Aussagen, „man nehme einen Teelöffel Zucker“ jedoch nicht.

Aussagen mit dem gleichen Wahrheitswert gelten vom Standpunkt der Aussagenlogik her als *äquivalent* (gleichwertig).

Aussagen können *negiert* werden. Wir notieren dies durch Voranstellen des Zeichens $\neg$. Diese Operation kehrt den Wahrheitswert um (d.h. eine wahre Aussage wird falsch und umgekehrt). Wir können zwei Aussagen durch die *Disjunktion* $\vee$ („(einschließendes) oder“) und die *Konjunktion* $\wedge$ (“und“) verbinden. Die Abhängigkeit der Wahrheitswerte der so erhaltenen Aussagen von den Wahrheitswerten der ursprünglichen Aussagen entnimmt man sogenannten *Wahrheitstafeln*. Aus den Grundoperationen $\neg$, $\vee$ und $\wedge$ kann man andere aus der Aussagenlogik geläufige Verknüpfungen wie „ausschließendes oder“ oder „impliziert“ erhalten und entsprechend die Wahrheitswerte der so erhaltenen Aussagen mit Hilfe der Wahrheitstafeln der Grundoperationen $\vee$, $\wedge$ und $\neg$ berechnen.

Sind Aussagen A und B gegeben, so schreiben wir zum Beispiel

$$A \Rightarrow B$$

gelesen „aus A folgt B“ für die Aussage $\neg(A \wedge (\neg B))$. Man überlegt, dass die Aussage $A \Rightarrow B$ genau dann falsch ist, falls A wahr und B falsch sind.

Um die Gültigkeit der Implikation $A \Rightarrow B$ zu zeigen, genügt es also zu zeigen, dass unter der Annahme, dass A wahr ist, auch B wahr ist. Dies entspricht einerseits genau unserer Intuition und schließt andererseits das wichtige logische Prinzip „ex falso quodlibet“ ein, nach dem die Implikation $A \Rightarrow B$ auf jeden Fall wahr ist, falls A falsch ist. So ist zum Beispiel die Aussage „Wenn 7 eine gerade Zahl ist, dann gibt es außerirdisches Leben“ in jedem Fall wahr. Als weitere abkürzende Schreibweise führen wir noch

$$A \Leftrightarrow B$$

für die zusammengesetzte Aussage

$$(A \Rightarrow B) \wedge (B \Rightarrow A)$$

ein.

Sind zwei beliebige Aussagen A und B gegeben, so ist die Aussage

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$$

eine *Tautologie*, d.h. sie ist für alle (insgesamt vier) möglichen Wahrheitsbelegungen von A und B immer wahr (dies kann man wieder mit einer Wahrheitstafel überprüfen). Dies formalisiert das *Beweisprinzip durch Kontraposition* (oder auch *Widerspruchsbeweis*): Wollen wir die Implikation $A \Rightarrow B$ zeigen, so ist es genauso gut, die Gültigkeit der Implikation $\neg B \Rightarrow \neg A$ zu überprüfen.

Manchmal enthalten Formulierungen *freie Variablen* (häufig bezeichnet mit x oder y) und werden erst dann zu Aussagen, wenn diese durch die *Quantoren* $\exists$ („es existiert“) oder $\forall$ („für alle“) *gebunden* werden. Man muss dann in der Regel noch spezifizieren, auf welche Klasse von Objekten man sich bezieht. So ist zum Beispiel

$$x > 3 \Rightarrow x > 5$$

noch keine Aussage. Wir können aber durch Voranstellen von Quantoren Aussagen gewinnen:

$$\forall x \in \mathbb{N} : x > 3 \Rightarrow x > 5$$

„für alle natürlichen Zahlen x gilt die Implikation $x > 3 \Rightarrow x > 5$ (diese Aussage ist falsch). Durch Binden der Variablen x mit dem Existenzquantor erhalten wir die wahre (!) Aussage

$$\exists x \in \mathbb{N} : x > 3 \Rightarrow x > 5\,.$$

Beim Negieren von Aussagen, die Quantoren enthalten, müssen Existenz- gegen Allquantoren ausgetauscht werden und umgekehrt:

Beispiel. Die Negation der Aussage „An allen Tagen gibt es einen Zeitpunkt, an dem die Sonne scheint.“ lautet: „Es gibt Tage, an denen nie die Sonne scheint.“.

Weitere Regeln für das Umformen von Aussagen finden sich in den Übungen.

Neben der Logik ist die Mengenlehre die zweite grundlegende mathematische Disziplin.

Wir betrachten zunächst die folgende naive von G. Cantor, dem Vater der modernen Mengenlehre, im Jahre 1895 vorgeschlagene „Definition“:

Eine Menge M ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche Elemente von M genannt werden) zu einem Ganzen.

Diese Beschreibung nimmt jedoch nicht ausschließlich auf Begriffe Bezug, die bereits exakt eingeführt wurden (was ist zum Beispiel eine „Zusammenfassung“?), und kann daher nicht als Definition im strengen Sinne akzeptiert werden. Man könnte diese Kritik als überzogen bezeichnen, wenn nicht B. Russell im Jahre 1903 auf folgende Paradoxie hingewisen hätte:

(Russellsche Antinomie) *Es sei M die Menge derjenigen Mengen, die sich nicht selbst als Element enthalten. Enthält sich dann M selbst als Element oder nicht?*

Nun prüft man leicht nach, dass die Annahme, M enthielte sich als Element, die Folgerung erzwingt, dass sich M nicht als Element enthält, und umgekehrt. Die Russelsche Konstruktion der Menge M, die ja nach Cantor möglich sein sollte, ergibt also keinen Sinn.

Wir werden jedoch nicht mit solch widersprüchlichen Gebilden zu tun haben und können daher problemlos mit dem naiven, von Cantor eingeführten Begriff arbeiten.

Mengen werden in der Regel mit großen Buchstaben bezeichnet, ihre Elemente meist mit Kleinbuchstaben. Wir schreiben kurz

$$a \in A$$

wenn die Menge A das Element a enthält. Ist eine Liste von Elementen gegeben, so können wir diese mit Hilfe von Mengenklammern zu einer Menge zusammenfassen. Z.B. bezeichnet

$$\{1, 2, 3, 5, 9\}$$

die Menge, die als Elemente die Zahlen 1, 2, 3, 5 und 9 enthält. Es kommt dabei nicht auf die Reihenfolge an und Wiederholungen sind irrelevant, d.h.

$$\{1, 2, 3, 5, 9\}$$

und

$$\{2, 9, 1, 1, 3, 5\}$$

bezeichnen die gleichen Mengen.

Das *Extensionalitätsaxiom* besagt, dass jede Menge durch ihre Elemente bestimmt ist. Für zwei gegebenen Mengen A und B gilt also genau dann $A = B$, falls für alle $x \in A$ auch $x \in B$ gilt und für alle $x \in B$ auch $x \in A$ gilt. Dies entspricht natürlich genau unserer intuitiven Vorstellung von Mengen.

Sind zwei Mengen A und B gegeben, und ist jedes Element, das in A enthalten ist, auch in B enthalten, so sagen wir „A ist eine *Teilmenge* von B“ und schreiben $A \subset B$. In diesem Fall sagt man auch, B ist eine *Obermenge* von A.

Damit können wir das Extensionalitätsaxiom wie folgt fassen:

Für zwei Mengen A und B gilt genau dann $A = B$, falls $A \subset B$ und $B \subset A$.

Diese Tatsache ist oft nützlich, wenn man konkret beweisen will, dass zwei Mengen gleich sind: Man beweist zunächst, dass $A \subset B$ (d.h. dass jedes Element in A auch in B liegt) und dann - möglicherweise getrennt davon - $B \subset A$. Weiter unten gibt es ein Beispiel dazu.

Eine besonders wichtige Menge ist die *Menge der natürlichen Zahlen*

$$\mathbb{N} = \{0, 1, 2, 3, \ldots\}.$$

Diese werden wir in dieser Vorlesung nicht konstruieren, sondern - zusammen mit der Addition und Multiplikation der natürlichen Zahlen - als gegeben voraussetzen. Wir machen des öfteren Gebrauch vom *Prinzip der vollständigen Induktion*, das wir in der Mengensprache wie folgt formulieren können:

Es sei $M \subset \mathbb{N}$ eine Teilmenge mit den folgenden beiden Eigenschaften:

- $0 \in M$.
- Ist $n \in M$, so auch $n + 1$.

Dann gilt $M = \mathbb{N}$.

Die weiteren Zahlbereiche $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ kann man aus $\mathbb{N}$ durch explizite Konstruktionen gewinnen. Darauf werden wir weiter unten eingehen.

Ist eine Menge M gegeben und ist $P(x)$ eine Formulierung mit der freien Variablen x, die durch Einsetzen von Elementen aus M zu einer Aussage wird, so können wir nach dem *Komprehensionsaxiom* die Menge

$$\{m \in M \mid P(m)\}$$

bilden, die genau aus denjenigen Elemente m aus M besteht, so dass $P(m)$ wahr ist. Setzen wir $P(x) := x$ ist Primzahl und $Q(x) := x$ ist gerade, dann besteht die Menge

$$\{n \in \mathbb{N} \mid P(n) \wedge Q(n)\}$$

also genau aus den geraden Primzahlen. Aus dem Komprehensaxiom folgt auch die Existenz einer Menge, die gar keine Elemente enthält, diese wird *leere Menge* genannt und mit $\emptyset$ bezeichnet. Eine mögliche Definition ist

$$\emptyset := \{n \in \mathbb{N} \mid n \neq n\}.$$

Als Elemente von Mengen können durchaus auch selbst Mengen vorkommen. Solche Mengen bezeichnet man manchmal als *Mengensysteme*. Z.B. bezeichnet

$$\{\{1, 2\}, \{1, 2, 3\}, \{2, 3, 4\}, \{4\}\}$$

eine Menge mit vier Elementen, die jeweils wieder Mengen, jeweils bestehend aus natürlichen Zahlen, sind.

Falls A und B Mengen sind, so schreiben wir

$$\begin{aligned} A \cap B &:= \{x \mid x \in A \text{ und } x \in B\} \\ A \cup B &:= \{x \mid x \in A \text{ oder } x \in B\} \\ A \setminus B &:= \{x \mid x \in A \text{ und } x \notin B\} \end{aligned}$$

für die *Schnitt-*, *Vereinigungsmenge* und *Komplementmenge* von A und B.

Beispiel. Sind zwei Mengen A und B gegeben, so gilt

$$A \setminus B = (A \cup B) \setminus B\,.$$

Wir weisen zunächst die Inklusion $A \setminus B \subset (A \cup B) \setminus B$ nach: Sei $x \in A \setminus B$. Dann liegt x in A aber nicht in B, d.h. x liegt sicher auch in $A \cup B$, aber nicht in B. Also gilt $x \in (A \cup B) \setminus B$. Dies zeigt die behauptete Inklusion.

Gilt umgekehrt $x \in (A \cup B) \setminus B$, so liegt x in A oder in B, jedoch nicht in B. Also liegt x in A, aber nicht in B, kurz $x \in A \setminus B$. Dies zeigt die Inklusion $(A \cup B) \setminus B \subset A \setminus B$.

Wichtige und bekannte Rechenregeln für Mengenoperationen, wie zum Beispiel die *Assoziativgesetze* $(A \cup B) \cup C = A \cup (B \cup C)$ und $(A \cap B) \cap C = A \cap (B \cap C)$, die *de Morganschen Regeln* $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$, $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$ und die *Distributivgesetze* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ lassen sich ähnlich beweisen.

Interessant ist hier die Parallelität zur Aussagenlogik. Zum Beispiel entsprechen die de Morgansche Regeln den Tautologien

$$\begin{aligned} \neg(A \vee B) &\Leftrightarrow \neg A \wedge \neg B \\ \neg(A \wedge B) &\Leftrightarrow \neg A \vee \neg B \end{aligned}$$

die auch bei der Verneinung von umganssprachlichen Aussagen, die durch Und oder Oder verknpft sind, Anwendung finden.

Die Menge aller Teilmengen einer gegebenen Menge A wird mit

$$\mathcal{P}(A)$$

bezeichnet und heißt *Potenzmenge* von A. Die Existenz dieser Menge sichert das *Potenzmengenaxiom*. Es gilt also z.B.

$$\mathcal{P}(\emptyset) = \{\emptyset\}\,,$$

insbesondere ist die Potenzmenge der leeren Menge nicht wieder die leere Menge.

Sind zwei Mengen A und B gegeben, so existiert nach dem *Paarmengenaxiom* die Menge

$$A \times B\,,$$

das *kartesische Produkt* von A und B. Die Menge $A \times B$ besteht aus allen *geordneten Paaren* (a, b), wobei $a \in A$ und $b \in B$. Wir sagen „geordnetes“ Paar deshalb, weil $(a, b) \neq (b, a)$, falls $a \neq b$. Ist $x \in A \cap B$, so ist $(x, x) \in A \times B$.

Definition. *Es seien A und B Mengen. Eine Teilmenge*

$$R \subset A \times B$$

heißt Relation zwischen A und B. *Ist hier $A = B$, so sprechen wir auch von einer Relation* auf A.

Erfüllt ein Paar $(a, b) \in A \times B$ eine gegebene Relation R (d.h. $(a, b) \in R$), so schreibt man auch $R(a, b)$ oder - noch häufiger - aRb.

Beispiel. Die Relation

$$\{(x, y) \in \mathbb{N} \times \mathbb{N} \mid \exists \alpha \in \mathbb{N} \text{ mit } x + \alpha = y\} \subset \mathbb{N} \times \mathbb{N}\,.$$

wird üblicherweise mit „$\leq$“ bezeichnet. Erfüllt ein Paar (x, y) diese Relation, so schreibt man „$x \leq y$“.

Definition. *Eine Relation $R \subset A \times B$ zwischen A und B heißt* Abbildung *oder* Funktion von A nach B, *falls für jedes Element $a \in A$ genau ein Element $b \in B$ existiert, so dass $(a, b) \in R$. In diesem Fall heißt A* Definitionsbereich *(oder* Quelle*) und B der* Wertebereich *(oder* Ziel*) von R. Die Relation selbst wird oft auch als* Graph *der Abbildung R bezeichnet.*

Man kann sich eine Abbildungen von A nach B anschaulich als eine Vorschrift vorstellen, die jedem Element aus A ein Element aus B zuordnet.

Abbildungen bezeichnet man in der Regel mit Kleinbuchstaben. Ist die Relation $f \subset A \times B$ eine Abbildung von A nach B, so schreiben wir $f : A \to B$ oder auch $A \xrightarrow{f} B$ und ist in dieser Situation $(x, y) \in f$, so schreiben wir $f(x) = y$ oder auch $x \overset{f}{\mapsto} y$. Nach obiger Definition sind zwei Abbildungen $f, g : A \to B$ genau dann gleich (d.h. f und g sind durch die gleiche Teilmenge von $A \times B$ gegeben), falls $f(a) = g(a)$ für alle $a \in A$ gilt.

Beispiel. Von den Relationen

$$\begin{aligned} \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x = 1\} &\subset \mathbb{N} \times \mathbb{N} \\ \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid y = 1\} &\subset \mathbb{N} \times \mathbb{N} \end{aligned}$$

ist die zweite eine Abbildung $\mathbb{N} \to \mathbb{N}$, die erste jedoch nicht.

Relationen und Abbildungen kann man gut durch (Pfeil-)Diagramme darstellen. Für jede Menge A gibt es eine besonders einfache Abbildung, die *Identität* auf A:

$$\mathrm{id}_A : A \to A\,,\ a \mapsto a\,.$$

Sind $f : A \to B$ und $g : B \to C$ Abbildungen, so bezeichnen wir die *Komposition* oder *Hintereinanderausführung* von g und f mit

$$g \circ f : A \to C\,,\ a \mapsto (g \circ f)(a) := g(f(a))\,.$$

Ist $f : A \to B$ eine Abbildung und $X \subset A$ eine Teilmenge, so bezeichnet $f|_X : X \to B$, $x \mapsto f(x)$, die *Einschränkung* von f auf X.

Definition. *Es seien A und B Mengen und $f : A \to B$ eine Abbildung.*

- *f heißt* injektiv*, falls für alle $x, y \in A$ mit $x \neq y$ immer $f(x) \neq f(y)$ gilt.*
- *f heißt* surjektiv*, falls für alle $y \in B$ ein $x \in A$ mit $f(x) = y$ existiert.*
- *f heißt* bijektiv*, falls f injektiv und surjektiv ist.*

Definition. *Es sei $f : A \to B$ eine Abbildung.*

- *Für $X \subset A$ heißt*

  $$f(X) := \{y \in B \mid \exists x \in X \text{ mit } f(x) = y\} \subset B$$

  das Bild *von X unter f. Falls $X = A$, so nennen wir $f(A)$ einfach das* Bild *von f.*
- *Für $Y \subset B$ heißt*

  $$f^{-1}(Y) := \{x \in A \mid f(x) \in Y\} \subset A$$

  das Urbild *von Y unter f.*

Eine Abbildung $f : A \to B$ ist also genau dann injektiv (surjektiv, bijektiv), falls für alle $b \in B$ das Urbild $f^{-1}(\{b\}) \subset A$ aus höchstens (mindestens, genau) einem Element besteht.

Der Übergang zum Urbild ist mit Mengenoperationen verträglich, d.h. ist $f : X \to Y$ eine Abbildung und sind $A, B \subset Y$, dann gilt

- $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$
- $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$
- $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$
- $f^{-1}(Y) = X$.

Beim Übergang zum Bild muss man aber aufpassen: Sind $A, B \subset X$, so gilt zwar immer noch $f(A \cup B) = f(A) \cup f(B)$, aber im allgemeinen ist $f(A \cap B)$ eine echte Teilmenge von $f(A) \cap f(B)$ und $f(X \setminus A)$ eine echte Obermenge von $f(X) \setminus f(A)$ (es sei denn, f ist injektiv) und $f(X)$ eine echte Teilmenge von Y (es sei denn, f ist surjektiv).

Die Komposition von Abbildungen ist assoziativ: Sind $f : A \to B$, $g : B \to C$ und $h : C \to D$ Abbildungen, so gilt

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Für den Beweis reicht nach dem oben Gesagten zu zeigen, dass für alle $a \in A$ die Gleichung $(h \circ (g \circ f))(a) = ((h \circ g) \circ f)(a)$ gilt. Aber die linke Seite ist gleich $h((g \circ f)(a)) = h(g(f(a)))$ und die rechte Seite gleich $(h \circ g)(f(a)) = h(g(f(a)))$, also sind beide Seiten gleich.

Hier noch zwei Beispiele, wie man mit diesen Begriffen umgeht.

Proposition 0.1. *Es seien $f : A \to B$ und $g : B \to A$ Abbildungen, so dass $g \circ f = \mathrm{id}_A$. Dann ist f injektiv und g surjektiv. Eine Abbildung $f : A \to B$ ist genau dann bijektiv, falls es eine Abbildung $g : B \to A$ gibt, so dass*

$$g \circ f = \mathrm{id}_A\,, \;\; f \circ g = \mathrm{id}_B\,.$$

Beweis. Zur ersten Aussage. Angenommen, f ist nicht injektiv. Dann gibt es $x, y \in A$ mit $x \neq y$ und $f(x) = f(y)$. Es folgt $g(f(x)) = g(f(y))$ und damit nach Voraussetzung $x = y$. Dies steht im Widerspruch zur Annahme. Damit ist f injektiv. Sei nun $a \in A$. Nach Annahme ist dann $g(f(a)) = a$ und somit ist g surjektiv. Damit haben wir auch schon gezeigt, dass f bijektiv ist, falls eine Abbildung g mit den Eigenschaften wie in der zweiten Aussage existiert. Es sei nun f bijektiv. Ist $b \in B$, so gibt es also genau ein $a \in A$ mit $f(a) = b$. Wir definieren $g(b) := a$. Dies definiert eine Abbildung $g : B \to A$ mit den beschriebenen Eigenschaften. Für weitere Details verweisen wir auf das Tutorium. □

In der Situation der letzten Proposition bezeichnet man die Abbildung g in der Regel mit f^{-1} und nennt diese Abbildung die *Umkehrabbildung* von f. In diesem Falle hat also f^{-1} zwei verschiedene Bedeutungen: Einmal als Abbildung $\mathcal{P}(B) \to \mathcal{P}(A)$ (Bildung des Urbildes) und ein anderes Mal als Abbildung $B \to A$.

Definition. *Eine Menge A heißt* endlich*, falls $A = \emptyset$ oder falls es eine natürliche Zahl n und eine Bijektion*

$$A \to \{1, 2, 3, \ldots, n\}$$

gibt. Die Menge A heißt unendlich*, falls sie nicht endlich ist.*

Ist $A \neq \emptyset$ eine endliche Menge, so kann man zeigen, dass in der letzten Definition die Zahl n eindeutig bestimmt ist. Diese heißt *Mächtigkeit* oder auch *Kardinalität* der Menge A, geschrieben $|A| = n$. Die Mächtigkeit der leeren Menge setzen wir als 0 fest, d.h. $|\emptyset| := 0$. Eine fundamentale Erkenntnis von Cantor ist, dass auch unendliche Mengen durchaus verschiedene „Anzahlen“ von Elementen haben können.

Sind A und B Mengen, so bezeichnen wir mit B^A oder auch mit $\mathrm{Abb}(A, B)$ die Menge aller Abbildungen $A \to B$. Die Schreibweise B^A kommt daher, dass man die n-fachen kartesischen Produkte

$$X^n = X \times X \times \ldots \times X$$

(n Faktoren) mit der Menge aller Abbildungen $\{1, 2, 3, \ldots, n\} \to X$ identifizieren kann. Allgemeiner ist manchmal folgender Gesichtspunkt nützlich: Sind I und A Mengen, so können wir die Menge A^I aller Abbildungen $I \to A$ als geordnete Tupel in A auffassen, die mit Hilfe der Elemente aus I indiziert sind. Für $I = \{0, 1\}$ ist also A^I nichts anderes als die Menge der geordneten Paare von Elementen aus A. Die Elemente aus A^I schreiben wir manchmal auch als *durch I indizierte Familien* $(x_i)_{i \in I}$, wobei $x_i \in A$ für alle $i \in I$.

Definition. *Es sei $R \subset A \times A$ eine Relation auf einer Menge A.*

- *R heißt* symmetrisch*, falls für alle $a, b \in A$ die Aussagen aRb und bRa äquivalent sind.*
- *R heißt* reflexiv*, falls für alle $a \in A$ die Aussage aRa gilt.*
- *R heißt* transitiv*, wenn für alle $a, b, c \in A$ die Aussagen aRb und bRc die Aussage aRc implizieren.*
- *R heißt* antisymmetrisch*, falls für $a, b \in A$ aus den Aussagen aRb und bRA folgt, dass $a = b$.*
- *R heißt* Ordnung*, falls R reflexiv, transitiv und antisymmetrisch ist.*
- *R heißt* Äquivalenzrelation*, falls R reflexiv, transitiv und symmetrisch ist.*

Ordnungen werden oft mit dem Zeichen $\leq$ benannt (d.h. statt aRb schreibt man $a \leq b$) und Äquivalenzrelationen häufig mit dem Zeichen $\sim$.

Beispiel. Ist A eine Menge, so wird durch $X \leq Y :\Leftrightarrow X \subset Y$ eine Ordnung $\leq$ auf der Potenzmenge $\mathcal{P}(A)$ definiert. Auf der Menge X der Hörer der Vorlesung „Lineare Algebra“ definieren wir eine Relation $\sim$ durch $x \sim y :\Leftrightarrow$ „x und y haben das gleiche Nebenfach“. Falls jeder Studierende genau ein Nebenfach hat, ist dies eine Äquivalenzrelation.

Ordnungen spielen besonders in der Analysis eine wichtige Rolle. Hier wollen wir uns zunächst mit den Äquivalenzrelationen genauer befassen.

Der folgende Satz zeigt, dass jede Äquivalenzrelation auf einer Menge A diese Menge in nichtleere Teilmengen zerlegt, die paarweise disjunkt sind, d.h. paarweise leeren Schnitt haben. Wir sprechen auch von einer *Partition* von A.

Proposition 0.2. *Es sei $R \subset A \times A$ eine Äquivalenzrelation. Für jedes $x \in A$ sei*

$$[x] := R[x] := \{a \in A \mid a \sim x\} \subset A$$

die Teilmenge der zu x äquivalenten Elemente in A. Dann gilt

- *Für alle $x \in A$ ist $[x] \neq \emptyset$,*
- *Für $x, y \in A$ gilt entweder $[x] = [y]$ oder $[x] \cap [y] = \emptyset$,*
- *$A = \bigcup_{x \in A} [x]$.*

Beweis. Die erste Aussage folgt aus $x \sim x$ für alle $x \in A$. Zum Beweis der zweiten Aussage nehmen wir an, wir haben $x, y \in A$ gegeben und $[x] \cap [y] \neq \emptyset$. Wir müssen $[x] = [y]$ zeigen. Wähle ein $z \in [x] \cap [y]$. Gilt nun $a \in [x]$, also $a \sim x$, so haben wir $a \sim x$ und $z \sim x$, also mit Symmetrie und Transitivität von R auch $a \sim z$. Da aber auch $z \sim y$, folgt wieder mit Transitivität $a \sim y$ und damit $a \in [y]$. Insgesamt folgt also $[x] \subset [y]$. Die andere Inklusion zeigt man analog. Die letzte Aussage der Proposition folgt daraus, dass $x \in [x]$ für alle $x \in A$ gilt, d.h. $A \subset \bigcup_{x \in A} [x]$ (die andere Inklusion ist klar). $\square$

Ist $\sim$ eine Äquivalenzrelation auf A, so bezeichnen wir die Menge der Äquivalenzklassen mit $A/\!\sim$. Dies ist ein Mengensystem, dessen Elemente

paarweise disjunkte Teilmengen von A sind. Ist $[x], [y] \in A/\sim$, so gilt genau dann $[x] \cap [y] \neq \emptyset$ (also $[x] = [y]$ nach der letzten Proposition), falls $x \sim y$.

Wer mehr über die Mengenlehre und deren historische Entwicklung wissen möchte, dem empfehle ich das Buch

Oliver Deiser: *Einführung in die Mengenlehre*, 2. Auflage (2004), Springer-Verlag.

1. Algebraische Grundstrukturen: Gruppen, Ringe, Körper

Definition. *Eine* Gruppe *ist ein Tripel $(G, \circ, e)$ bestehend aus einer Menge G, einer Verknüpfung (d.h. Abbildung) $\circ : G \times G \to G$ und einem Element $e \in G$, so dass folgende* Gruppenaxiome *erfüllt sind:*

- *(Assoziativgesetz) $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in G$.*
- *(Neutrales Element) $e \circ a = a$ für alle $a \in G$.*
- *(Existenz des Inversen) Für alle $a \in G$ existiert ein Element a^{-1} mit $a^{-1} \circ a = e$.*

Wir ziehen zunächst weitere einige Folgerungen über Gruppen.

Man sieht, dass auch $a \circ e = a$ für alle $a \in G$ gilt, wie aus der Gleichung

$$a^{-1}(ae) = (a^{-1}a)e = ee = e = a^{-1}a$$

durch Multiplikation von links mit $(a^{-1})^{-1}$ folgt. (Wir lassen im Folgenden häufig das Verknüpfungssymbol $\circ$ weg). Daraus folgt auch, dass e durch die obigen Bedingungen eindeutig bestimmt ist, denn ist e' ein weiteres neutrales Element, so gilt $e' = e \circ e' = e$ nach dem gerade Gezeigten. Nun sieht man auch, dass für $a \in G$ das Inverse a^{-1} die Gleichung $aa^{-1} = e$ erfüllt, wie aus der Gleichung

$$a^{-1}(aa^{-1}) = (a^{-1}a)a^{-1} = ea^{-1} = a^{-1}$$

durch Multiplikation mit $(a^{-1})^{-1}$ von links folgt. Damit ist für alle $a \in G$ das Inverse a^{-1} eindeutig bestimmt, denn ist $b \in G$ ein weiteres Inverses von a, also $ba = e$, so folgt nach Multiplikation mit a^{-1} von rechts mit dem gerade Gezeigten $b = a^{-1}$. Daraus folgt insbesondere, dass das Inverse von a^{-1} das Element a selbst ist, denn es ist ja $aa^{-1} = e$. Wichtig ist auch die Gleichung

$$(ab)^{-1} = b^{-1}a^{-1}$$

denn die Gleichung $b^{-1}a^{-1}ab = e$ zeigt, dass $b^{-1}a^{-1}$ das Inverse von ab ist.

Da die Verknüpfung in einer Gruppe assoziativ ist, können wir bei Ausdrücken der Art $(a \circ b) \circ c$ auf die Klammern verzichten: Es ist ja gleichgültig, welche Verknüpfung man zuerst durchführt.

Gruppen beschreiben Symmetrien mathematischer (oder auch physikalischer) Objekte. Auf diesen Gesichtspunkt werden wir noch oft zurückkommen. Das folgende Beispiel ist aber in dieser Hinsicht fundamental.

Beispiel. Es sei X eine Menge. Die Menge der bijektiven Abbildungen $f : X \to X$ bildet zusammen mit der Hintereinanderausführung von Abbildungen (d.h. die Verknüpfung ist gegeben durch $(f, g) \mapsto f \circ g$) und der Identität als neutralem Element eine Gruppe, die *symmetrische Gruppe* von X, die mit Sym_X bezeichnet wird. Die Assoziativität der Gruppenverknüpfung folgt aus der Assoziativität der Komposition von Abbildungen. Ist $X = \{1, 2, \ldots, n\}$, so schreiben wir statt Sym_X auch oft Sym_n oder S_n.

Hat eine Gruppe G endlich viele Elemente, so bezeichnet man ihre Anzahl $|G|$ als *Ordnung* der Gruppe. Die Ordnung der symmetrischen Gruppe S_n ist gleich $n!$.

Ist $\sigma \in S_n$, so stellen wir dieses Element als

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

dar. Zum Beispiel ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

diejenige Permutation von $\{1, 2, 3\}$, die 1 auf 3, 2 auf 2 und 3 auf 1 abbildet.

Definition. *Eine Gruppe $(G, \circ, e)$ heißt* abelsch *oder* kommutativ*, falls für alle $a, b \in G$ die Gleichung $a \circ b = b \circ a$ gilt.*

In einer abelschen Gruppe wird die Verknüpfung oft mit $+$, das neutrale Element oft mit 0 und das Inverse eines Elementes a mit $-a$ (anstatt a^{-1}) bezeichnet. Statt $a + (-b)$ schreibt man in der Regel $a - b$. Es gelten dann also die gewohnten Rechenregeln $a + b = b + a$ und $-(a + b) = -a - b$. Man beachte aber, dass die Symbole $+$ und 0 verschiedene Bedeutung haben, wenn wir es mit verschiedenen Gruppen zu tun haben.

Für Gruppen kleiner Ordnung kann man die Verknüpfung in *Gruppentafeln* darstellen. Eine Gruppe ist genau dann abelsch, wenn die zugehörige Gruppentafel spiegelsymmetrisch zur Diagonalen ist. Die symmetrische Gruppe S_n ist genau dann abelsch, falls $n \leq 2$.

Die wichtigste abelsche Gruppe ist die Gruppe der ganzen Zahlen. Wir wollen im folgenden die ganzen Zahlen explizit aus den natürlichen Zahlen konstruieren. Dies ist auch eine gute Gelegenheit, das Konzept der Äquivalenzrelationen zu illustrieren.

Wir setzen im folgenden voraus, dass wir die Menge $\mathbb{N} = \{0, 1, 2, \ldots\}$ der natürlichen Zahlen gegeben haben und darauf eine Verknüpfung („Addition“) mit den folgenden Eigenschaften definiert ist:

- Sie ist kommutativ: Für alle $a, b \in \mathbb{N}$ ist $a + b = b + a$.
- Sie ist assoziativ: Für alle $a, b, c \in \mathbb{N}$ gilt $(a + b) + c = a + (b + c)$.
- Für alle $a \in \mathbb{N}$ gilt die Gleichung $0 + a = a$.
- Es gilt die folgende *Kürzungsregel*: Sind $x, y, n \in \mathbb{N}$ und gilt $x + n = y + n$, so ist $x = y$. (Mit anderen Worten: Die Abbildung $\mathbb{N} \to \mathbb{N}$, $x \mapsto x + n$, ist injektiv).

Auf die Konstruktion der Menge $\mathbb{N}$ und der Verknüpfung $+$ im Rahmen der axiomatischen Mengenlehre gehen wir an dieser Stelle nicht näher ein.

Wir konstruieren nun die Menge $\mathbb{Z}$ der ganzen Zahlen zusammen mit der gewohnten Addition wie folgt. Auf der Menge $\mathbb{N} \times \mathbb{N}$ betrachten wir die Relation $\sim$ gegeben durch

$$(x, y) \sim (x', y') :\Leftrightarrow x + y' = y + x' .$$

Man überlegt sich, dass es sich um eine Äquivalenzrelation handelt. Reflexivität und Symmetrie sind klar. Zur Transitivität nehmen wir an, dass $(x, y) \sim (x', y')$ und $(x', y') \sim (x'', y'')$. Somit ist $x+y' = y+x'$ und $x'+y'' = y'+x''$. Addition der Gleichungen führt auf $x+y'+x'+y'' = y+x'+y'+x''$, woraus mit der Kürzungsregel folgt, dass $x + y'' = y + x''$. Somit ist auch $(x, y) \sim (x'', y'')$.

Es sei nun $M := \mathbb{N} \times \mathbb{N}/\sim$ die Menge der Äquivalenzklassen. Die Äquivalenzklasse von $(x, y) \in \mathbb{N} \times \mathbb{N}$ schreiben wir als $[x, y]$ (anstatt $[(x, y)]$). Auf $\mathbb{N} \times \mathbb{N}$ definieren wir eine Verknüpfung

$$\circ : (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N}) \to (\mathbb{N} \times \mathbb{N})$$

durch die Setzung

$$(x, y) \circ (a, b) := (x + a, y + b)$$

und zeigen:

Proposition 1.1. *Falls $(x, y) \sim (x', y')$ und $(a, b) \sim (a', b')$, so gilt $(x, y) \circ (a, b) = (x', y') \circ (a', b')$.*

Beweis. Wir nehmen also an, dass $x + y' = y + x'$ und $a + b' = b + a'$. Dann folgt $x + a + y' + b' = y + b + x' + a'$ durch Addition der Gleichungen und mit dem Kommutativgesetz, somit also auch wie gewünscht $(x + a, y + b) \sim (x' + a', y' + b')$. □

Diese Proposition zeigt, dass die Verknüpfung $\circ$ eine Verknüpfung auf M durch die Setzung

$$([x, y], [a, b]) \mapsto [x + a, y + b]$$

induziert. Man sagt auch, die so gegebene Verknüpfung ist *wohldefiniert*. Wir bezeichnen diese Verknüpfung ebenfalls mit $\circ$.

Proposition 1.2. *Das Tripel $(M, \circ, [0, 0])$ ist eine abelsche Gruppe.*

Beweis. Das Kommutativ- und Assoziativgesetz folgen direkt aus der Definition. Dies gilt auch für die Tatsache, dass $[0, 0]$ ein neutrales Element ist. Interessant ist die Existenz von Inversen: Sei also $[x, y] \in \mathbb{N} \times \mathbb{N}/\sim$ gegeben. Wir behaupten, dass $[y, x]$ invers dazu ist, d.h. wir müssen die Gleichung

$$[x, y] \circ [y, x] = [0, 0]$$

zeigen. Es ist aber $[x, y] \circ [y, x] = [x + y, y + x]$ und $[x + y, y + x] = [0, 0]$, da $x + y + 0 = y + x + 0$. Damit ist alles gezeigt. □

Die so definierte Gruppe sieht noch nicht so aus wie die ganzen Zahlen, wie wir sie gewohnt sind. Zunächst versuchen wir, die natürlichen Zahlen in M wiederzufinden.

Proposition 1.3. *Die Abbildung $\phi : \mathbb{N} \to M$, $x \mapsto [x, 0]$ ist injektiv. Es gilt $\phi(x + y) = \phi(x) \circ \phi(y)$ für alle $x, y \in \mathbb{N}$.*

Beweis. Es sei $\phi(x) = \phi(y)$, also $[x, 0] = [y, 0]$. Das bedeutet nach Definition der Äquivalenzrelation gerade $x + 0 = y + 0$ und somit gilt nach den vorher genannten Eigenschaften der Addition auf $\mathbb{N}$ (0 als neutrales Element und Kommutativgesetz) die Gleichung $x = y$. Der Rest der Proposition ist klar. □

Aufgrund der Injektivität von ϕ können wir uns die Menge $\mathbb{N}$ als Teilmenge von M vorstellen (indem wir $n \in \mathbb{N}$ mit $\phi(n) \in M$ identifizieren). Da unter dieser Identifikation die Verknüpfung $+$ auf $\mathbb{N}$ mit der Verknüpfung $\circ$ auf M übereinstimmt, und da M ohnehin abelsch ist, benennen wir die Verknüpfung $\circ$ auf M in $+$ um. Die abelsche Gruppe $(M, +, 0)$ bezeichnen wir fortan mit dem Symbol $\mathbb{Z}$. Dies ist die *Gruppe der ganzen Zahlen*.

Die folgende Proposition gibt noch genauer Auskunft darüber, wie man sich die Menge $\mathbb{Z}$ vorstellen kann.

Proposition 1.4. *Für alle Elemente $m \in \mathbb{Z}$ gilt genau einer der drei folgenden Fälle:*

- $m = 0$.
- *Es gibt ein $n \in \mathbb{N} \setminus \{0\}$ mit $m = [n, 0]$ (d.h. $m \in \mathbb{N}$). Wir schreiben in diesem Fall $m > 0$ und nennen m* positiv.
- *Es gibt ein $n \in \mathbb{N} \setminus \{0\}$ mit $m = [0, n]$. Wir schreiben dann $m < 0$ und nennen m* negativ. *In diesem Fall ist also $-m \in \mathbb{N}$.*

Die Zahl n im zweiten und dritten Fall ist durch m eindeutig bestimmt.

Für den Beweis verwenden wir die folgende Eigenschaft der Addition auf $\mathbb{N}$, die wir ebenfalls nicht beweisen werden: Sind $a, b \in \mathbb{N}$, so ist mindestens eine der Gleichungen $a + x = b$ oder $b + x = a$ lösbar. (Die Lösung x ist dann nach der Kürzungsregel jeweils eindeutig bestimmt). Falls beide Gleichungen lösbar sind, so gilt $a = b$.

Beweis. Es sei $m = [a, b]$ gegeben. Falls $m \neq 0 = [0, 0]$, so gilt $a \neq b$. Nach dem eben Gesagten ist genau eine der Gleichungen $a + x = b$ oder $b + x = a$ lösbar. Angenommen, die zweite Gleichung ist lösbar mit Lösung $x \in \mathbb{N}$. Dann ist $a + 0 = x + b$ und somit $[a, b] = [x, 0]$. Falls die erste Gleichung lösbar ist mit Lösung $x \in \mathbb{N}$, so erhalten wir $a + x = b + 0$ und somit $[a, b] = [0, x]$.

Dass die Zahl n durch m eindeutig bestimmt ist, folgt aus der Injektivität von ϕ aus der vorherigen Proposition. Ganz analog zeigt man, dass $n \mapsto [0, n]$ eine injektive Abbildung $\mathbb{N} \to M$ definiert.

Es bleibt noch zu zeigen, dass sich die drei angegebenen Fälle gegenseitig ausschließen. Der einzig nichttriviale Teil ist, dass sich der zweite und dritte

Fall ausschließen. Seien also $x, y \in \mathbb{N} \setminus \{0\}$. Wir müssen zeigen, dass $[x, 0] \neq [0, y]$ und nehmen an, dass dies doch so ist. Dann gilt $x + y = 0 + 0 = 0$. Also sind die beiden Gleichungen $x + \xi = 0$ und $0 + \xi = x$ in $\xi \in \mathbb{N}$ lösbar (die Lösung der ersten Gleichung ist nach Annahme y, die der zweiten ist x). Nach der Bemerkung vor dem Beweis folgt daraus $x = 0$, und dies steht im Widerspruch zur Annahme. □

Jedes Element in $\mathbb{Z}$ ist also positiv, negativ oder gleich 0 und diese Fälle schließen sich gegenseitig aus. Sind $x, y \in \mathbb{Z}$, so schreiben wir im folgenden statt $x + (-y)$ kürzer $x - y$. Damit gelten auf $\mathbb{Z}$ die bekannten Rechenregeln. Inbesondere haben wir die suggestive Gleichung

$$[x, y] = [x, 0] + [0, y] = [x, 0] - [y, 0] = x - y$$

und die Äquivalenz

$$x - y = x' - y' \Leftrightarrow x + y' = y + x'$$

führt uns wieder auf die obige Äquivalenzrelation in $\mathbb{N} \times \mathbb{N}$.

In vielen algebraischen Strukturen hat man es mit zwei Operationen gleichzeitig zu tun:

Definition. *Ein* Ring *ist ein Quadrupel* $(R, +, 0, \cdot)$*, wobei*

- $(R, +, 0)$ *eine abelsche Gruppe ist.*
- $\cdot : R \times R \to R$ *eine assoziative Verknüpfung ist.*
- *das Distributivgesetz gilt: Sind* $a, b, c \in R$*, so gilt* $a \cdot (b + c) = a \cdot b + a \cdot c$ *und* $(b + c) \cdot a = b \cdot a + c \cdot a$*.*

Wir nennen einen Ring R kommutativ*, falls für alle* $a, b \in R$ *die Gleichung* $a \cdot b = b \cdot a$ *gilt. Ist in einem Ring* $(R, +, 0, \cdot)$ *eine Element* $1 \in R$ *gegeben, so dass* $1 \cdot a = a \cdot 1 = a$ *für alle* $a \in R$*, so spricht man von einem* Ring mit Einselement*. Fast alle Ringe, die uns begegnen werden, sind Ringe mit Einselement.*

Man beachte, dass wir an dieser Stelle nicht fordern, dass es für die Verknüpfung $\cdot$ auf R inverse Elemente gibt. In jedem Ring R gilt die Gleichung $0 \cdot a = 0$ für alle $a \in R$, wie aus der Gleichung $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$ durch Abziehen (d.h. Addition des Negativen) von $0 \cdot a$ folgt.

Die gewöhnliche Mulitplikation $\cdot$ auf $\mathbb{N}$ kann so zu einer Multiplikation $\cdot$ auf $\mathbb{Z}$ erweitert werden, dass $(\mathbb{Z}, +, 0, \cdot, 1)$ ein kommutative Ring mit 1 ist (dies geschieht in den Übungen). Das ist vielleicht das wichtigste Beispiel eines kommutativen Ringes.

Definition. *Ein Ring* R *heißt* nullteilerfrei*, falls aus der Gleichung* $a \cdot b = 0$ *folgt, dass* $a = 0$ *oder* $b = 0$*.*

Der Ring der ganzen Zahlen ist nullteilerfrei, wie in den Übungen bewiesen wird.

Bevor wir später weitere Ringe (auch nicht nullteilerfreie und nichtkommutative) kennenlernen, bringen wir noch die dritte wichtige Definition

in diesem Abschnitt. Ist R ein Ring, so schreiben wir im folgenden R^* statt $R \setminus \{0\}$.

Definition. *Es sei $(R, +, 0, \cdot, 1)$ ein kommutativer Ring mit Einselement. Wir nennen R einen Körper, falls $0 \neq 1$ und falls für jedes $a \in R^*$ ein Element $b \in R$ existiert mit $ba = 1$.*

Anders ausgedrückt: Ein Körper ist ein kommutativer Ring mit Einselement, in dem 0 von 1 verschieden ist und in dem jedes von Null verschiedene Element ein multiplikatives Linksinverses besitzt (das dann auch rechtsinvers ist, da $\cdot$ kommutativ ist). Ist R ein Körper und sind $a, b \in R^*$, so ist $ab \neq 0$ (ansonsten führt die Multiplikation von links mit einem Inversen von a zu einem Widerspruch). Damit ist jeder Körper nullteilerfrei. Insbesondere induziert $\cdot$ eine Verknüpfung auf R^* und wir sehen so, dass in jedem Körper das Tripel $(R^*, \cdot, 1)$ eine abelsche Gruppe ist. Also ist das Inverse jedes Elementes $x \neq 0$ (wie in jeder Gruppe) eindeutig bestimmt (jedenfalls innerhalb R^*; man überlegt sich aber noch leicht, dass es nicht gleich 0 sein kann, weil dann die Gleichung $0 = 0 \cdot x = 1$ zu einem Widerspruch zur Annahme $0 \neq 1$ führt.)

Ist k ein Körper, so wird das multiplikative Inverse von $x \in k^*$ in der Regel mit x^{-1} oder auch $\frac{1}{x}$ bezeichnet. Da in einem Körper die Multiplikation nach Definition kommutativ ist, gilt in jedem Körper die Gleichung $(xy)^{-1} = x^{-1}y^{-1}$. Sind $y \in k$ und $x \in k^*$, so benutzen wir statt $y \cdot x^{-1}$ die Bruchschreibweise $\frac{y}{x}$. In einem Körper sind also die „vier Grundrechenarten“ erklärt und es gelten die aus der Schule bekannten Rechenregeln. Insbesondere erfolgt die Addition zweier Brüche nach der Regel

$$\frac{x}{y} + \frac{a}{b} = \frac{xb + ya}{by}\,,$$

also durch Bilden des „Hauptnenners“. Eine Begründung für diese Regel steckt in der Gleichung

$$xy^{-1} + ab^{-1} = (xb)(b^{-1}y^{-1}) + (ay)(y^{-1}b^{-1}) = (xb + ay)(by)^{-1}\,.$$

9.11.09

Wir wollen aus $\mathbb{Z}$ den Körper der rationalen Zahlen konstruieren. Zunächst machen wir uns Gedanken, welche Zahlen in $\mathbb{Z}$ überhaupt ein multiplikatives Inverses besitzen.

Proposition 1.5. *Es sei $x \in \mathbb{Z}$. Dann besitzt x genau dann ein multiplikatives Inverses in $\mathbb{Z}$, falls $x = \pm 1$.*

Im folgenden Beweis machen wir davon Gebrauch, dass wir auf $\mathbb{Z}$ eine Ordnung $\leq$ gegeben haben, so dass die üblichen Rechenregeln für Ungleichungen gelten (diese lassen sich aus den bisher bewiesenen Tatsachen leicht ableiten).

Beweis. Falls $x = \pm 1$, so besitzt x offensichtlich ein multiplikatives Inverses. Seien nun umgekehrt $x, y \in \mathbb{Z}$ mit $yx = 1$. Dann gilt $x \neq 0$, denn $\mathbb{Z}$ ist

nullteilerfrei. Ohne Beschränkung der Allgemeinheit (indem wir, wenn nötig, x und y durch $-x$ und $-y$ ersetzen) ist $x > 0$. Dann gilt auch $y > 0$ (andernfalls wäre $yx < 0$ im Widerspruch zu $yx = 1 > 0$). Wir wollen $x = 1$ zeigen. Angenommen $x \geq 2$. Dann haben wir $x(y-1) = xy - x = 1 - x < 0$. Da $y > 0$, haben wir $y - 1 \geq 0$ und damit $y - 1 > 0$ (im Falle $y - 1 = 0$ wäre $x(y-1) = 0$ im Widerspruch zu $x(y-1) < 0$.) Aus der Gleichung $x(y-1) < 0$ folgt wegen $y - 1 > 0$ dann aber $x < 0$ und das widerspricht unserer Annahme $x > 0$. □

Wenn wir aus $\mathbb{Z}$ einen Körper konstruieren wollen, müssen wir also alle ganzen Zahlen ungleich $0, \pm 1$ „künstlich“ invertierbar machen. Dies geschieht ähnlich wie beim Übergang von $\mathbb{N}$ nach $\mathbb{Z}$ durch Betrachtung einer geschickt gewählten Äquivalenzrelation.

Wir definieren auf der Menge $\mathbb{Z} \times \mathbb{Z}^*$ eine Relation $\sim$ durch

$$(x, y) \sim (x', y') :\Leftrightarrow xy' = yx' .$$

Dies ist eine Äquivalenzrelation: Symmetrie und Reflexivität sind klar, die Transitivität wird im Tutorium behandelt.

Wir setzen nun für $[x, y], [a, b] \in (\mathbb{Z} \times \mathbb{Z}^*) / \sim$

$$[x, y] \oplus [a, b] := [xb + ya, yb]$$

und

$$[x, y] \odot [a, b] := [xa, yb]$$

Diese Verknüpfungen sind wohldefiniert, denn sind $y \neq 0 \neq b$, so gilt auch $yb \neq 0$ (denn $\mathbb{Z}$ ist nullteilerfrei). Gilt außerdem $(x, y) \sim (x', y')$, d.h. $xy' = yx'$, so haben wir

$$(xb + ya)(y'b) = xby'b + yay'b = x'byb + y'ayb = (x'b + y'a)yb$$

und dies zeigt $(x, y) \oplus (a, b) \sim (x', y') \oplus (a, b)$. Außerdem ist

$$xay'b = ybx'a$$

woraus $(x, y) \odot (a, b) \sim (x', y') \odot (a, b)$ folgt. Für $(a, b) \sim (a', b')$ zeigt man leicht die entsprechenden Aussagen. Man beweist nun, dass die Abbildung

$$\phi : \mathbb{Z} \to (\mathbb{Z} \times \mathbb{Z}^*) / \sim , \ z \mapsto [z, 1]$$

injektiv ist und die Gleichungen $\phi(x+y) = \phi(x) \oplus \phi(y)$, $\phi(x \cdot y) = \phi(x) \odot \phi(y)$ gelten. In diesem Sinne fassen wir $\mathbb{Z}$ als Teilmenge von $(\mathbb{Z} \times \mathbb{Z}^*) / \sim$ auf (d.h. wir identifizieren $z \in \mathbb{Z}$ mit $[z, 1] \in (\mathbb{Z} \times \mathbb{Z}^*) / \sim$) und benennen die Verknüpfungen $\oplus$ und $\odot$ in $+$ und $\cdot$ um. Die Menge $(\mathbb{Z} \times \mathbb{Z}^*) / \sim$ wird fortan mit $\mathbb{Q}$ bezeichnet.

Proposition 1.6. $(\mathbb{Q}, +, 0, \cdot, 1)$ *ist ein Körper.*

Wir nennen diesen Körper den *Körper der rationalen Zahlen.*

Beweis. Die Ringaxiome prüft man unmittelbar nach. Es sei nun $[x, y] \neq 0$. Dann ist $x \neq 0$ (andernfalls wäre $(x, y) \sim (0, 1)$, also $[x, y] = 0$). Damit ist die rationale Zahl $[y, x]$ definiert. Man prüft leicht nach, dass es sich um ein multiplikatives Inverses handelt. □

Da $[1, y]$ invers zu $y = [y, 1]$ ist und außerdem die Gleichung $[x, y] = [x, 1] \cdot [1, y] = xy^{-1}$ gilt, schreiben wir in Zukunft $\frac{x}{y}$ statt $[x, y]$.

Wir werden in den Übungen sehen, dass man durch die Setzung

- $\frac{x}{y} > 0 :\Leftrightarrow xy > 0$.
- $\frac{x}{y} < 0 :\Leftrightarrow xy < 0$

den Körper $\mathbb{Q}$ zu einem Archimedisch angeordneten Körper macht.

Aus der Sicht der Analysis besteht nun das Problem, dass in $\mathbb{Q}$ nicht jede Cauchy-Folge konvergiert, bzw. nicht jede Intervallschachtelung eine rationale Zahl definiert. Insbesondere ist die Gleichung $x^2 = 2$ in $\mathbb{Q}$ nicht lösbar.

Wieder durch eine explizite Konstruktion (Betrachtung von Äquivalenzklassen von Cauchy-Folgen) kann man den Körper $\mathbb{Q}$ zum Körper $\mathbb{R}$ der reellen Zahlen erweitern, in dem diese Probleme nicht auftreten. Der Körper $\mathbb{R}$ ist ein Archimedisch angeordneter, vollständiger Körper und Hauptgegenstand der Analysis. Allerdings ist in $\mathbb{R}$ die Gleichung $x^2 = -1$ nicht lösbar.

Auf dem kartesischen Produkt $\mathbb{R} \times \mathbb{R}$ kann man nun durch die Setzungen $(x, y)+(a, b) := (x+a, y+b)$ und $(x, y)\cdot(a, b) := (xa-yb, xb+ya)$ ebenfalls die Struktur eines Körpers definieren (mit Nullelement $(0, 0)$ und Einselement $(1, 0)$) und erhält so den *Körper der komplexen Zahlen* $\mathbb{C}$. Statt (x, y) schreibt man $x + iy$. Es gilt damit die Gleichung $i^2 = -1$.

Bei der Konstruktion der verschiedenen Zahlsysteme haben wir uns immer von der gleichen Problematik leiten lassen: In $\mathbb{N}$ ist die Gleichung $x + 1 = 0$ nicht lösbar; dies führte zur Konstruktion von $\mathbb{Z}$. In $\mathbb{Z}$ ist die Gleichung $2x = 1$ nicht lösbar (siehe Proposition 1.5) und dies führte zur Konstruktion von $\mathbb{Q}$. In $\mathbb{Q}$ ist die Gleichung $x^2 = 2$ nicht lösbar; dies führte zur Konstruktion von $\mathbb{R}$. Und die Nichtlösbarkeit der Gleichung $x^2 = -1$ in $\mathbb{R}$ führte schließlich zur Konstruktion von $\mathbb{C}$. Dieser Körper ist nun unter dem Aspekt der Lösbarkeit von Gleichungen optimal, denn es gilt

Satz 1.7 (Fundamentalsatz der Algebra)**.** *Jede Gleichung der Form*

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0$$

mit $a_0, \ldots, a_n \in \mathbb{C}$, $n \geq 1$ *und* $a_n \neq 0$ *hat mindestens eine Lösung* $x \in \mathbb{C}$.

Der Beweis dieses Resultates benötigt Hilfsmittel aus der Analysis und wird am Ende der Vorlesung nachgeholt.

Nachdem wir die algebraischen Strukturen Gruppe, Ring und Körper eingeführt haben, wollen wir die strukturerhaltenden Abbildungen zwischen diesen studieren.

Definition. *Es seien $(G, \circ_G, e_G)$ und $(H, \circ_H, e_H)$ Gruppen. Eine Abbildung $\phi : G \to H$ heißt* Gruppenhomomorphismus*, falls $\phi(g \circ_G h) = \phi(g) \circ_H \phi(h)$ für alle $g, h \in G$. Wir nennen ϕ einen* Gruppenisomorphismus*, wenn ϕ ein Gruppenhomomorphismus und bijektiv ist. Existiert ein Gruppenisomorphismus $G \to H$, so nennen wir G und H* isomorph *und wir schreiben $G \cong H$.*

Aus der Sicht der Gruppentheorie werden isomorphe Gruppen als im wesentlichen gleich angesehen.

Beispiel. Die Abbildung $\mathbb{Z} \to \mathbb{Z}$, $x \mapsto 3x$, ist ein Gruppenhomomorphismus, die Abbildung $\mathbb{Z} \to \mathbb{Z}$, $x \mapsto x + 1$, jedoch nicht. Ist G eine Gruppe, so ist $G \to G$, $g \mapsto g^{-1}$, genau dann ein Gruppenhomomorphismus, wenn G abelsch ist.

Ist ϕ ein Gruppenhomomorphismus, so gilt $\phi(e_G) = e_H$ und $\phi(g^{-1}) = \phi(g)^{-1}$ für alle $g \in G$ und ist ϕ ein Gruppenisomorphismus, so ist die Umkehrabbildung ϕ^{-1} ebenfalls ein Gruppenhomomorphismus. Die Beweise sind leicht und werden in [Fischer] ausgeführt (siehe S. 48).

Ist $\phi : G \to H$ ein Gruppenhomomorphismus, so hat das Bild $\phi(G) \subset H$ eine besonders schöne Eigenschaft.

Definition. *Es sei $(G, \circ, e)$ eine Gruppe. Eine* Untergruppe *von G ist eine nichtleere Teilmenge $K \subset G$, so dass $g^{-1} \in K$ und $g \circ h \in K$ für alle $g, h \in K$. Wir schreiben in diesem Falle $K < G$.*

Gilt $K < G$, so ist $e \in K$, denn wegen $K \neq \emptyset$ gibt es gibt ein $g \in K$ und damit ist $g \circ g^{-1} = e \in K$.

Für alle $n \in \mathbb{N}$ ist $n\mathbb{Z} := \{n \cdot z \mid z \in \mathbb{Z}\} \subset \mathbb{Z}$ eine Untergruppe. Insbesondere ist $\{0\}$ eine Untergruppe von $\mathbb{Z}$.

Proposition 1.8. *Es sei $\phi : G \to H$ ein Gruppenhomomorphismus. Dann ist $\phi(G) \subset H$ eine Untergruppe von H.*

Der Beweis ist leicht.

Definition. *Es sei $\phi : G \to H$ ein Gruppenhomomorphismus. Wir definieren den* Kern *von ϕ durch*

$$\ker(\phi) := \phi^{-1}(\{e_H\}) \subset G$$

Wichtig ist die folgende Aussage.

Proposition 1.9. $\ker(\phi)$ *ist eine Untergruppe von G. Die Abbildung ϕ ist genau dann injektiv, wenn $\ker(\phi) = \{e_G\} \subset G$.*

Beweis. Falls ϕ injektiv ist, so gilt $\phi^{-1}(\{e_H\}) = \{e_G\}$. Sei nun umgekehrt $\ker \phi = \{e_G\}$. Angenommen $\phi(x) = \phi(y)$. Es folgt $\phi(xy^{-1}) = \phi(x)\phi(y)^{-1} = e_H$ und somit wegen $\ker(\phi) = \{e_G\}$ die Gleichung $xy^{-1} = e_G$, d.h. $x = y$. □

Wichtige Beispiele von Ringen sind die *Restklassenringe*. Dazu betrachten wir zunächst die Teilbarkeitsrelation auf $\mathbb{Z}$.

Definition. *Es seien $a, b \in \mathbb{Z}$. Wir schreiben $a|b$ („a teilt b“), falls ein $z \in \mathbb{Z}$ existiert mit $za = b$. Insbesondere gilt also $a|0$ für alle $a \in \mathbb{Z}$ und $0|0$.*

Die Teilbarkeitsrelation ist reflexiv und transitiv.

Es sei nun $n \in \mathbb{N} \setminus \{0\}$. Wir betrachten die Relation $\sim$ auf $\mathbb{Z}$ gegeben durch:

$$a \sim b :\Leftrightarrow n|(a-b)$$

Man überlegt sich leicht, dass dies eine Äquivalenzrelation auf $\mathbb{Z}$ ist. Statt $a \sim b$ schreiben wir auch $a = b \mod n$. Die Menge der Äquivalenzklassen wird mit $\mathbb{Z}/n\mathbb{Z}$ oder auch $\mathbb{Z}/n$ bezeichnet. Die Äquivalenzklassen heißen *Restklassen modulo n*. Die Menge $\mathbb{Z}/n$ hat genau n Elemente, nämlich die Restklassen $[0], [1], \ldots [n-1]$. Wir schreiben im folgenden oft $\overline{x}$ statt $[x]$. Es gilt $[x] = \{x + kn \mid k \in \mathbb{Z}\}$.

11.11.09

Proposition 1.10. *Die Addition und Multiplikation auf $\mathbb{Z}$ induzieren Verknüpfungen auf $\mathbb{Z}/n$. Mit diesen Verknüpfungen wird $(\mathbb{Z}/n, +, [0], \cdot, [1])$ ein kommutativer Ring mit 1.*

Beweis. Sei $a \sim a'$, also $n|(a-a')$. Ist $b \in \mathbb{Z}$, so gilt $n|((a+b)-(a'+b))$, somit ist $a+b \sim a'+b$. Entsprechend zeigt man $a+b \sim a+b'$, falls $b \sim b'$. Aus der Tatsache, dass $\sim$ eine Äquivalenzrelation (somit transitiv) ist, folgt, dass $a+b \sim a'+b'$, falls $a \sim a'$ und $b \sim b'$, und die Addition auf $\mathbb{Z}/n$ ist somit wohldefiniert. Ist $a \sim a'$, so gilt $n|(ab-a'b)$ (beachte $ab-a'b = (a-a')b$) und somit $ab \sim a'b$. Ebenso zeigt man $ab \sim ab'$, falls $b \sim b'$. Damit ist auch die Multiplikation wohldefiniert. Der Rest der Proposition folgt durch direktes Nachrechnen. □

Es gelten also die Rechenregeln $\overline{x} + \overline{y} = \overline{x+y}$, $\overline{x} \cdot \overline{y} = \overline{xy}$ in $\mathbb{Z}/n$. Zum Beispiel ist $[3]+[21] = [3] = [10]$ in $\mathbb{Z}/7$. Man beachte, dass hier das Zeichen $+$ verschiedene Bedeutungen hat: Auf der linken Seite der Gleichungen ist es eine Verknüpfung in $\mathbb{Z}/n$, auf der rechten Seite in $\mathbb{Z}$.

Wir nennen eine von Null und Eins verschiedene Zahl $p \in \mathbb{N}$ eine *Primzahl*, falls für alle $a, b \in \mathbb{Z}$ gilt

$$p|(ab) \Rightarrow p|a \vee p|b\,.$$

Proposition 1.11. *Der Ring $\mathbb{Z}/n$ ist genau dann ein Körper, falls n eine Primzahl ist.*

Beweis. Für $\overline{x} \in \mathbb{Z}/n$ betrachten wir den Gruppenhomomorphismus der additiven Gruppe $\mathbb{Z}/n$

$$\phi_{\overline{x}} : \mathbb{Z}/n \to \mathbb{Z}/n\,,\ \overline{z} \mapsto \overline{x} \cdot \overline{z}$$

Der Ring $\mathbb{Z}/n$ ist genau dann ein Körper, falls für alle $\overline{x} \neq 0$, das Element $\overline{1}$ im Bild von $\phi_{\overline{x}}$ liegt. Es sei nun n eine Primzahl. Ist $\overline{x} \in (\mathbb{Z}/n)^*$ [1], so ist $\phi_{\overline{x}}$ injektiv: Falls $\overline{y} \in \ker(\phi_{\overline{x}})$, so gilt $n|(yx)$. Da n eine Primzahl ist, gilt also $n|y$

[1] zur Erinnerung: Ist R ein Ring, so setzen wir $R^* := R \setminus \{0\}$

oder $n|x$. Wegen $\overline{x} \neq 0$ scheidet der zweite Fall aus, also ist $\overline{y} = 0$ und es ist $\ker(\phi_{\overline{x}}) = 0$ (hier bezeichnet 0 die Untergruppe von $\mathbb{Z}/n$ bestehend aus dem neutralen Element). Somit ist $\phi_{\overline{x}}$ injektiv. Da die Menge $\mathbb{Z}/n$ endlich ist, ist $\phi_{\overline{x}}$ auch surjektiv. Insbesondere gilt $\overline{1} \in \operatorname{im}(\phi_{\overline{x}})$ und $\mathbb{Z}/n$ ist ein Körper.

Angenommen, n ist keine Primzahl. Dann gibt es $a, b \in \mathbb{Z}$ mit $n \not| a$ und $n \not| b$, aber $n|(ab)$. Inbesondere ist also $\overline{a} \neq 0$ und $\overline{b} \neq 0$, aber $\overline{a}\overline{b} = 0$. Dann ist also $\mathbb{Z}/n$ nicht nullteilerfrei und somit auch kein Körper. □

Dieser Beweis gibt keine Auskunft darüber, wie wir das multiplikative Inverse eines Elementes $\overline{x} \in (\mathbb{Z}/n)^*$, n prim, explizit bestimmen können. Auf diesen Punkt kommen wie später zurück.

Definition. *Es seien $(R, +_R, 0_R, \cdot_R)$ und $(S, +_S, 0_S, \cdot_S)$ Ringe. Ein* Ringhomomorphismus *$f : R \to S$ ist eine Abbildung, so dass f ein Gruppenhomomorphismus $(R, +_R, 0_R) \to (S, +_S, 0_S)$ ist und außerdem $f(x \cdot_R y) = f(x) \cdot_S f(y)$ für alle $x, y \in R$. Sind R und S Ringe mit 1, so fordern wir außerdem, dass $f(1_R) = 1_S$. Sind R und S* Körper*, so nennen wir einen Ringhomomorphismus $f : R \to S$ auch* Körperhomomorphismus*. Ein Ringhomomorphismus heißt* Ringisomorphismus*, falls er zusätzlich bijektiv ist. In diesem Falle heißen die beteiligten Ringe (Körper)* isomorph*.*

Sind R und S Körper und $f : R \to S$ ein Ringhomomorphismus, so gilt $f(x^{-1}) = f(x)^{-1}$ für alle $x \in R^*$, wie aus der Gleichung $f(x)f(x)^{-1} = f(xx^{-1}) = f(x)f(x^{-1})$ durch Multiplikation mit $f(x)^{-1}$ folgt. Man zeigt ähnlich wie im Falle von Gruppen, dass die Umkehrabbildung eines Ringisomorphismus wieder ein Ringisomorphismus ist.

Proposition 1.12. *Jeder Ringhomomorphismus $\mathbb{Z} \to \mathbb{Z}$ und $\mathbb{Q} \to \mathbb{Q}$ ist gleich der Identität.*

Beweis. Es sei $\phi : \mathbb{Z} \to \mathbb{Z}$ ein Ringhomomorphismus. Da $\mathbb{Z}$ ein Ring mit 1 ist, haben wir $\phi(1) = 1$ und da ϕ ein Gruppenhomomorphismus ist (von der Gruppe $(\mathbb{Z}, +, 0)$ in die Gruppe $(\mathbb{Z}, +, 0)$), gilt dann auch $\phi(-1) = -1$. Wir stellen nun eine beliebige ganze Zahl $z \in \mathbb{Z}$ als $1 + 1 + \ldots + 1$, bzw. $-1 - 1 \ldots - 1$ dar (jeweils $|z|$ Summanden), um zu folgern, dass $\phi(z) = z$ für alle $z \in \mathbb{Z}$.

Ist $\phi : \mathbb{Q} \to \mathbb{Q}$ ein Ringhomomorphismus, so haben wir gerade gezeigt, dass $\phi|_{\mathbb{Z}} = \mathrm{id}_{\mathbb{Z}}$. Sei nun $y \in \mathbb{Z} \setminus \{0\}$. Wir erhalten $1 = \phi(yy^{-1}) = \phi(y)\phi(y^{-1})$. Hieraus folgt (wegen $0 \cdot q = 0$ für alle $q \in \mathbb{Q}$), dass $\phi(y) \neq 0$ und Multiplikation der letzten Gleichung mit $\phi(y)^{-1}$ zeigt $\phi(y^{-1}) = \phi(y)^{-1} = y^{-1}$ (hier benutzen wir die Gleichung $\phi(y) = y$, die wir ja für $y \in \mathbb{Z}$ bereits gezeigt haben). Ist nun $\frac{x}{y} \in \mathbb{Q}$, so folgern wir $\phi(\frac{x}{y}) = \phi(x \cdot y^{-1}) = x \cdot y^{-1} = \frac{x}{y}$. Also ist $\phi = \mathrm{id}_{\mathbb{Q}}$. □

2. Lineare Gleichungssysteme

Es sei K ein Körper, also z.B. $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ oder $K = \mathbb{Z}/p$ mit p prim. Ein Grundproblem der Algebra besteht in der Bestimmung der Lösungen

algebraischer Gleichungen der Form

$$\Phi(x_1, \ldots, x_n) = 0$$

wobei $x_1, \ldots, x_n$ Unbestimmte sind und $\Phi(x_1, \ldots, x_n)$ ein Ausdruck ist, der die Verknüpfungen in K, die Elemente aus K, sowie die Unbestimmten $x_1, \ldots, x_n$ enthält. Ein Beispiel ist die Gleichung

$$\frac{x_1}{x_2} + x_1^2 - x_1 x_3 + 7 = 0$$

in $\mathbb{Q}$. Im Allgemeinen ist die Bestimmung solcher Lösungmengen sehr schwierig. Jede Lösung der Gleichung

$$x_1^n + x_2^n = x_3^n$$

in $\mathbb{Q}$ führt zum Beispiel durch Multiplikation mit dem Hauptnenner auf eine Gleichung der Form

$$z_1^n + z_2^n = z_3^n$$

mit $z_1, z_2, z_3 \in \mathbb{Z}$. Falls $n \geq 3$, so konnte A. Wiles im Jahre 1993 zeigen, dass diese Gleichung keine (ganzzahligen) Lösungen besitzt, falls wir $z_1, z_2, z_3 \neq 0$ fordern (großer Fermatscher Satz). Hier kommen allerdings sehr fortgeschrittene Techniken der modernen Mathematik zum Einsatz. Also ist auch für jede Lösung $(x_1, x_2, x_3) \in \mathbb{Q}^3$ der vorigen Gleichung $x_1 = 0$ oder $x_2 = 0$ oder $x_3 = 0$, falls $n \geq 3$.

Allgemeiner kann man auch *Gleichungssysteme* bestehend aus endlich vielen Gleichungen obiger Form in den Unbestimmen $x_1, \ldots, x_n$ betrachten. Die Lösungsmenge eines solchen Gleichungssystems ist die Schnittmenge der Lösungsmengen der einzelnen Gleichungen. Lösungsmengen solch allgemeiner Gleichungssysteme untersucht die algebraische Geometrie.

In der lineare Algebra beschänken wir uns auf *lineare Gleichungssysteme*, also Gleichungssysteme der Form

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

Dies sind m *lineare Gleichungen* in n *Unbekannten* mit *Koeffizienten* a_{ij}, b_i, wobei $1 \leq i \leq m$, $1 \leq j \leq n$. Wir sprechen auch von einem *linearen Gleichungssystem*. Dieses Gleichungssystem heißt *homogen*, falls $b_1 = b_2 = \ldots = b_m = 0$. Folgende Fragen liegen nahe:

- Unter welchen Voraussetzungen sind derartige Gleichungssysteme lösbar?
- Falls Lösungen existieren, welche Struktur hat die Lösungsmenge $L \subset K^n$?
- Wie kann man L effektiv berechnen?

Die lineare Algebra gibt auf diese und viele weitere Fragen sehr befriedigende Antworten. Erstmal ist klar, dass jedes homogene Gleichungssystem die *triviale Lösung* $(0,\dots,0) = 0 \in \mathbb{R}^n$ hat.

Im allgemeinen ist es hilfreich, den geometrischen Gehalt obiger Gleichungen (auch im nicht-homogenen Fall) zu beleuchten. Als Illustration betrachten wir die Gleichung

$$2x_1 + x_2 = 1$$

über dem Körper $\mathbb{R}$. Diese Gleichung beschreibt eine *Gerade* durch die Punkte $(\frac{1}{2}, 0)$ und $(0, 1)$ im $\mathbb{R}^2$. Die Lösungemenge der Gleichung $0x_1 + 0x_2 = 0$ ist der ganze $\mathbb{R}^2$ und die Lösungsmenge der Gleichung $0x_1 + 0x_2 = 1$ ist leer. Im Allgemeinen ist die Lösungsmenge jeder Gleichung

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i$$

eine *Hyperebene* im $\mathbb{R}^n$, falls mindestens ein $a_{ij} \neq 0$, $j = 1,\dots,n$. Diese Hyperebenen sind gewisse „$(n-1)$-dimensionale Teilräume“ im $\mathbb{R}^n$ (d.h. eine Gerade, falls $n = 2$, eine Ebene, falls $n = 3$ etc.). Lösungsmenge des gesamten Gleichungssystems ist der Schnitt dieser Hyperebenen. Falls alle $b_1,\dots,b_n = 0$, so enthalten alle Hyperebenen den Punkt $(0,\dots,0) \in \mathbb{R}^n$ und somit auch deren Schnitt. Falls es außerdem weniger Gleichungen als Unbekannte gibt, sollte die „Dimension“ dieses Schnittes größer als 0 sein. Insbesondere sollte es auch Lösungen $(x_1,\dots,x_n) \neq (0,\dots,0)$ geben.

Die lineare Algebra macht genau diese geometrische Intuition präzise. Man kann mit dieser Intuition auch schon gut verstehen, welche Antworten auf die obigen Fragen zu erwarten sind: Falls wir 3 lineare Gleichungen in 2 Unbestimmten betrachten, so ist im Allgemeinen die Lösungsmenge die Schnittmenge dreier Geraden im $\mathbb{R}^2$ und somit leer. Ist $b_1,\dots,b_n = 0$ und ist die Anzahl der Gleichungen kleiner als die Anzahl der Unbekannten, so sollte es mindestens eine Lösung $(x_1,\dots,x_n) \neq 0$ des Gleichungssystems geben, denn es werden ja m Hyperebenen im $\mathbb{R}^n$ geschnitten, wobei $m < n$.

16.11.09

Wir beschreiben den folgenden Algorithmus zur Lösung linearer Gleichungssysteme, das sogenannte *Gaußsche Eliminationsverfahren*. Der Übersichtlichkeit halber fassen wir die Koeffizienten $a_{ij} \in K$ des obigen Systems zur *Koeffizientenmatrix*

$$A := (a_{ij})_{1\leq i\leq m,\ 1\leq j\leq n} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

mit m Zeilen und n Spalten zusammen. Wir sprechen auch von einer $(m \times n)$-Matrix. Wir betrachten daneben auch die *erweiterte Koeffizientenmatrix*

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

also eine$(m \times (n+1))$-Matrix. Wir werden später Matrizen im Zusammenhang mit linearen Abbildungen noch genauer untersuchen. Hier dienen sie nur der bequemen Notation. Es ist klar, dass jede $(m \times (n+1))$-Matrix die erweiterte Koeffizientenmatrix genau eines lineares Gleichungssystems mit m Gleichungen und n Unbekannten ist. Wir betrachten nun die folgenden Operationen, genannt *elementare Zeilenumformungen* auf der erweiterten Koeffizientenmatrix

i. Vertauschung zweier Zeilen.
ii Addition des λ-fachen der i_1-ten Zeile zur i_2-ten Zeile, wobei $\lambda \in K$ und $1 \leq i_1, i_2 \leq m$, $i_1 \neq i_2$.

Diese ändern die Lösungsmenge des zugrunde liegenden Gleichungssystems nicht:

Proposition 2.1. *Angenommen, die erweiterten Koeffizientenmatrizen $(A|b)$ und (A', b') gehen durch elementare Zeilenumformungen auseinander hervor. Dann stimmen die Lösungsmengen der entsprechenden linearen Gleichungssysteme überein.*

Beweis. Dies ist offensichtlich bei der Vertauschung zweier Zeilen, denn dann werden einfach zwei Gleichungen vertauscht. Angenommen $(x_1, \dots, x_n)$ löst das System $(A|b)$, und $(A'|b')$ gehe aus $(A|b)$ durch Addition des λ-fachen der i_1-ten zur i_2-ten Zeile hervor. Dann lautet die neue i_2-te Gleichung:

$$(\lambda a_{i_1 1} + a_{i_2 1})x_1 + \ldots + (\lambda a_{i_1 n} + a_{i_2 n})x_n = \lambda b_{i_1} + b_{i_2}$$

Da $(x_1, \dots, x_n)$ die i_1-te und die i_2-te Gleichung löst, ist aber $(x_1, \dots, x_n)$ auch eine Lösung dieser neuen Gleichung. Damit ist die Lösungsmenge des Gleichungssystems zu $(A|b)$ in dem zu (A', b') enthalten. Die andere Inklusion zeigt man analog, denn $(A|b)$ entsteht durch Addition des $-\lambda$-fachen der i_1-ten Zeile zur i_2-ten Zeile aus $(A'|b')$. □

Die elementaren Zeilenumformungen sind deshalb nützlich, weil man mit ihrer Hilfe jedes lineare Gleichungssystem in *Zeilenstufenform* bringen kann.

Definition. *Ein lineares Gleichungssystem ist in* Zeilenstufenform*, falls die (nicht erweiterte) Koeffizientenmatrix A in* Zeilenstufenform *vorliegt: Entweder hat diese Matrix nur 0 als Einträge oder es gibt ein $1 \leq r \leq m$ und eine Folge $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ mit den folgenden Eigenschaften:*

- *$a_{ij} = 0$, falls $j < j_i$, $1 \leq i \leq r$, oder falls $i > r$*
- *$a_{ij_i} \neq 0$ falls $1 \leq i \leq r$.*

(Insbesondere sind alle Zeilen unterhalb der r-ten gleich 0.)

Die von Null verschiedenen Elemente a_{ij_i}, $i = 1, \ldots, r$ heißen dann Pivotelemente *des Gleichungssystems, bzw. der Koeffizientenmatrix.*

Proposition 2.2. *Jedes lineare Gleichungssystem lässt sich durch elementare Zeilenumformungen auf Zeilenstufenform bringen.*

Beweis. Falls die Koeffizientenmatrix A nur 0 als Einträge hat, sind wir schon fertig. Andernfalls wählen wir einen minimalen Spaltenindex j_1, $1 \leq j_1 \leq n$, mit der Eigenschaft, dass es einen Zeilenindex $i_1 \in \{1, \ldots, m\}$ gibt mit $a_{i_1 j_1} \neq 0$. Wir vertauschen nun in $(A|b)$ die erste mit der i_1-ten Zeile. Dann ist der j_1-te Eintrag der ersten Zeile ungleich 0 und alle Einträge in der Matrix A links von der j_1-ten Spalte sind gleich 0. Indem wir in $(A|b)$ geeignete Vielfache der ersten zur zweiten bis m-ten Zeile addieren, machen wir alle a_{ij_1} zu Null, falls $i > 1$. Dieses Verfahren wiederholen wir für die Teilmatrix von $(A|b)$ bestehend aus der zweiten bis zur m-ten Zeile. □

Ist ein lineares Gleichungssystem in Zeilenstufenform gegeben, so lässt sich dieses sehr einfach lösen.

Angenommen es existiert ein $b_i \neq 0$ mit $i > r$. Dann ist die Lösungsmenge leer.

Andernfalls bestimmen wir die Lösungsmenge wie folgt: Für jede beliebige Wahl der $n - r$ Zahlen $x_j \in K$ für $1 \leq j \leq n$, $j \neq j_1, j_2, \ldots, j_r$ („freie Parameter“) existiert genau eine Wahl der verbleibenden Komponenten $x_{j_1}, \ldots, x_{j_r}$, so dass $(x_1, \ldots, x_n)$ das Gleichungssystem löst. Denn durch die r-te Gleichung ist wegen $a_{rj_r} \neq 0$ und $a_{rj} = 0$ für $j < j_r$ die Komponente x_{j_r} eindeutig durch die Komponenten $x_{j_r+1}, \ldots, x_n, b_r$ festgelegt:

$$x_{j_r} = \frac{1}{a_{rj_r}}(b_r - a_{r\,j_r+1}x_{j_r+1} - \ldots - a_{rn}x_n)$$

Danach legt die $(r-1)$-te Gleichung die Komponente $x_{j_{r-1}}$ eindeutig fest und so weiter. Umgekehrt sind natürlich durch jede Lösung $(x_1, \ldots, x_n)$ des Gleichungssystems die Komponenten x_j, $j \neq j_1, \ldots, j_r$ eindeutig bestimmt. Damit gilt:

Satz 2.3. *Es sei ein lineares Gleichungssystem in Zeilenstufenform gegeben. Es sei L die Lösungsmenge. Dann existiert eine bijektive Abbildung*

$$\phi : K^{n-r} \to L$$

indem wir jedes $(n-r)$-Tupel $(\lambda_1, \ldots, \lambda_{n-r}) \in K^{n-r}$ auf die durch diese Elemente eindeutig bestimmte Lösung $(x_1, \ldots, x_n)$ des Gleichungssystems abbilden, bei der die freien Parameter x_j, $j \neq j_1, \ldots, j_r$ gleich $\lambda_1, \ldots, \lambda_{n-r}$ gesetzt wurden.

Dieses Theorem erlaubt es, die Lösungsmenge des Gleichungssystems in der sogenannten *Parameterform* anzugeben:

$$L = \{\phi(\lambda_1, \ldots, \lambda_{n-r}) \mid \lambda_1, \ldots, \lambda_{n-r} \in K\}$$

Da wir jedes lineare Gleichungssystem auf Zeilenstufenform bringen können ohne die Lösungsmenge zu ändern, haben wir somit eine effektive Methode gefunden, beliebige lineare Gleichungssystem zu lösen.

Als Beispiel betrachten wir das lineare Gleichungssystem

$$\begin{aligned} 3x_6 + x_7 &= 2 \\ 2x_2 + 4x_4 + 6x_5 + 5x_7 &= 3 \\ 2x_2 + x_3 + 7x_4 + 8x_5 + x_6 + 5x_7 &= 4 \\ 2x_2 + 4x_4 + 6x_5 + 3x_6 + 6x_7 &= 5 \end{aligned}$$

über $\mathbb{R}$. Dieses hat die erweiterte Koeffizientenmatrix

$$(A|b) := \left(\begin{array}{ccccccc|c} 0 & 0 & 0 & 0 & 0 & 3 & 1 & 2 \\ 0 & 2 & 0 & 4 & 6 & 0 & 5 & 3 \\ 0 & 2 & 1 & 7 & 8 & 1 & 5 & 4 \\ 0 & 2 & 0 & 4 & 6 & 3 & 6 & 5 \end{array}\right)$$

Durch elementare Zeilenumformungen wird daraus die erweiterte Koeffizientenmatrix

$$\left(\begin{array}{ccccccc|c} 0 & 2 & 0 & 4 & 6 & 0 & 5 & 3 \\ 0 & 0 & 1 & 3 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

deren (nicht erweiterte) Koeffizientenmatrix in Zeilenstufenform vorliegt. Die Lösungsmenge L dieses Gleichungssystems ist in Parameterform gegeben durch

$$L = \{ \begin{pmatrix} \lambda_1 \\ \frac{3}{2} - 2\lambda_2 - 3\lambda_3 - \frac{5}{2}\lambda_4 \\ \frac{1}{3} - 3\lambda_2 - 2\lambda_3 + \frac{1}{3}\lambda_4 \\ \lambda_2 \\ \lambda_3 \\ \frac{2}{3} - \frac{1}{3}\lambda_4 \\ \lambda_4 \end{pmatrix} \mid \lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}\} \subset \mathbb{R}^7$$

wobei wir die Elemente in $\mathbb{R}^7$ in Spaltenform notieren (d.h. die Komponenten untereinander statt nebeneinander schreiben), vgl. [Fischer], S. 25.

Aus unseren allgemeinen Betrachtungen über das Gaußsche Eliminationsverfahren erhalten wir:

Korollar 2.4. *Es sei ein homogenes lineares Gleichungssystem bestehend aus m Gleichungen mit n Unbestimmten gegeben. Falls $m < n$, so besitzt dieses Gleichungssystem mindestens eine nichttriviale Lösung ungleich $(0, \ldots, 0) \in K^n$.*

Beweis. Da das Gleichungssystem homogen ist, hat es auf jeden Fall Lösungen (z.B. die triviale). Ist das Gleichungssystem auf Zeilenstufenform gebracht, so muss es wegen $m < n$ mindestens einen freien Parameter x_j,

$1 \leq j \leq n$ geben. Für diesen wählen wir einfach ein Element ungleich 0 in K. □

18.11.09

Es tauchen die folgenden theoretischen Fragen auf:

- Ist die Zahl $n - r$ der freien Parameter (d.h. die Anzahl r der Pivotelemente) durch das Gleichungssystem eindeutig festgelegt? Oder könnte es sein, dass verschiedene Verfahren, das ursprüngliche Gleichungssystem auf Zeilenstufenform zu bringen, zu unterschiedlichen Anzahlen von Pivotelementen führen?
- Kann man der Abbildung ϕ eine bessere algebraische Struktur geben?

Diese Fragen beantworten wir mit der Theorie der Vektorräume. Innerhalb dieser Theorie können wir insbesondere auch die algebraische Struktur, die hinter linearen Gleichungssystemen und ihren Lösungen steckt, klären.

3. Theorie der Vektorräume

Definition. *Es sei K ein Körper. Ein* Vektorraum *über K oder auch K-*Vektorraum *ist eine abelsche Gruppe $(V, +, 0)$ zusammen mit einer Verknüpfung*

$$\cdot : K \times V \to V$$

genannt Skalarenmultiplikation*, die folgende Eigenschaften für alle $\lambda, \mu \in K$ und $v, w \in V$ hat:*

- $1 \cdot v = v$.
- $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$
- $\lambda(v + w) = \lambda v + \lambda w$.
- $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$ *(man beachte, dass hier $\cdot$ in verschiedenen Rollen auftritt).*

Die Elemente eines Vektorraumes nennt man *Vektoren*, die Elemente von K *Skalare*.

Beispiel.

- K ist mit den Verknüpfungen aus K selbst ein K-Vektorraum.
- Die abelsche Gruppe $\{0\}$ mit einem Element ist ein Vektorraum über jedem Körper K, genannt *Nullvektorraum*. Diesen bezeichnen wir auch mit 0. Die Vektorraumstrukturen sind aber für verschiedene K verschieden, denn der zugrundeliegende Köper ist Teil der Struktur eines Vektorraumes.
- Die Menge K^n ist mit den *komponentenweisen* Verknüpfungen

$$\begin{aligned}(x_1, \ldots, x_n) + (y_1, \ldots, y_n) &:= (x_1 + y_1, \ldots, x_n + y_n)\\ \lambda \cdot (x_1, \ldots, x_n) &:= (\lambda x_1, \ldots, \lambda x_n)\end{aligned}$$

ein Vektorraum über K. Dieser Vektorraum wird uns sehr häufig begegnen. Er heißt manchmal auch n-dimensionaler *Koordinatenraum*.

Oft schreiben wir die Elemente von K^n auch in Spaltenform. Wir definieren noch $K^0 := 0$ als den Nullvektorraum über K.

- Es sei $K^{m\times n}$ die Menge der $m \times n$-Matrizen mit Einträgen in K. Hier erhalten wir durch komponentenweise Addition und Skalarenmultiplikation ebenfalls die Struktur eines Vektorraumes.
- Die Menge der komplexen Zahlen $\mathbb{C}$ ist mit der üblichen Addition und der Skalarenmultiplikation $\mathbb{R} \times \mathbb{C} \to \mathbb{C}$, $\lambda \cdot v := \lambda v$ (hier fassen wir $\mathbb{R} \subset \mathbb{C}$ auf) ein Vektorraum über $\mathbb{R}$. Allgemeiner ist jeder $\mathbb{C}$-Vektorraum auch ein $\mathbb{R}$-Vektorraum.
- Ist K ein Körper und I eine nichtleere Menge, so ist die Menge der Abbildungen $K^I = \{(x_i)_{i\in I} \mid x_i \in K\}$ mit der komponentenweisen Verknüpfung ein Vektorraum.

In Vektorräumen gelten einige weitere Rechenregeln, siehe [Fischer], S. 77.

Da in einem Vektorraum V über K die Addition assoziativ ist, ist für alle $v_1, \ldots, v_r$ und $\lambda_1, \ldots, \lambda_r \in K$ der Ausdruck

$$\sum_{i=1}^{r} \lambda_i \cdot v_i \in V$$

ohne Klammersetzung eindeutig definiert. Das Gleiche gilt für $\lambda_1 \cdot \ldots \cdot \lambda_r \cdot v$, falls $v \in V$.

Definition. *Es sei V ein K-Vektorraum und $W \subset V$ eine Teilmenge. Wir nennen W einen* Untervektorraum *von V, falls W eine Untergruppe von V ist (insbesondere ist $0 \in W$) und für alle $\lambda \in K$ und $w \in W$ auch $\lambda w \in W$, d.h. W ist abgeschlossen unter Skalarenmultiplikation.*

Ein Untervektorraum von V ist also mit den Verknüpfungen von V selbst wieder ein Vektorraum. Man studiere die Beispiele in [Fischer], S. 78 oben. Ein sehr wichtiges Beispiel ist das folgende.

Proposition 3.1. *Es sei ein homogenes lineares Gleichungssystem mit n Unbestimmten über K gegeben. Dann ist die Lösungsmenge dieses Gleichungssystems ein Untervektorraum des K^n.*

Beweis. Dies rechnet man direkt nach. □

Man beachte aber, dass die Lösungsmenge eines inhomogenen Gleichungssystems kein Untervektorraum des K^n ist (denn in diesem Fall ist $(0, \ldots, 0)$ nicht in der Lösungsmenge enthalten).

Definition. *Es sei V ein Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren aus V (dabei ist I eine beliebige Indexmenge). Eine* Linearkombination *dieser Vektoren ist eine Summe der Gestalt*

$$\sum_{i\in I} \lambda_i v_i$$

wobei alle $\lambda_i \in K$ sind und fast alle (d.h. alle bis auf endlich viele) $\lambda_i = 0$ (anders ausgedrückt: Nur endlich viele λ_i sind ungleich 0). Falls $I = \{1, 2, \ldots, r\}$ endlich ist, d.h. $(v_i)_{i\in I} = (v_1, \ldots, v_r)$, kann man auf diese Bedingung natürlich verzichten.

Es sei nun $W \subset V$. Falls $W \neq \emptyset$, setzen wir

$$\operatorname{span}(W) := \{v \in V \mid v \text{ ist Linearkombination aus Vektoren in } W\} \subset V\,.$$

Ist $W = \emptyset$, so setzen wir $\operatorname{span}(W) := \{0\} \subset V$. *Entsprechend definieren wir* $\operatorname{span}(v_i)_{i\in I} \subset V$.

Proposition 3.2. *Für alle $W \subset V$ ist* $\operatorname{span}(W) \subset V$ *ein Untervektorraum.*

Beweis. Dies ist klar, denn eine Linearkombination aus Linearkombinationen von Vektoren aus W ist wieder eine Linearkombinationen von Vektoren aus W. □

Wir nennen $\operatorname{span}(W)$ den von W *erzeugten* oder *aufgespannten* Vektorraum. Dies ist der kleinste Untervektorraum von V, der W enthält.

Definition. *Es sei V ein Vektorraum und $(v_i)_{i\in I}$ eine Familie aus Vektoren in V.*

Wir nennen $(v_i)_{i\in I}$ linear unabhängig, *falls folgendes gilt: Ist eine Linearkombination $\sum_{i\in I} \lambda_i v_i = 0$, so gilt $\lambda_i = 0$ für alle $i \in I$. Ist $(v_i)_{i\in I}$ nicht linear unabhängig, so heißt* $(v_i)_{i\in I}$ linear abhängig.

Wir nennen $(v_i)_{i\in I}$ ein Erzeugendensystem *von V, falls* $\operatorname{span}(v_i) = V$. *In diesem Fall können wir also jeden Vektor aus V als Linearkombination in den v_i, $i \in I$, schreiben.*

Und schließlich heißt $(v_i)_{i\in I}$ eine Basis *von V, falls (v_i) linear unabhängig und ein Erzeugendensystem ist.*

23.11.09

Beispiel.

- Die leere Familie ist linear unabhängig, denn für $I = \emptyset$ ist die Bedingung „$\lambda_i = 0$ für alle $i \in I$“ stets erfüllt. Diese Familie ist genau dann eine Basis von V, falls $V = \{0\}$, d.h. falls V der Nullvektorraum ist.
- Im K-Vektorraum K^n setzen wir $e_i := (0, \ldots, 0, 1, 0, \ldots, 0)$, wobei die 1 genau an der i-ten Stelle steht, für $i = 1, \ldots, n$. Dann ist $(e_1, \ldots, e_n)$ eine Basis von K^n, die sogenannte *Standardbasis*.
- Es sei eine Matrix $A \in K^{m\times n}$ in Zeilenstufenform gegeben. Es sei $r \geq 0$ die Anzahl der von 0 verschiedenen Zeilen. Dann sind die Zeilenvektoren $(a_{11}, \ldots, a_{1n}), \ldots, (a_{r1}, \ldots, a_{rn})$ linear unabhängig im K^n. Ebenso sind die Spaltenvektoren

$$\begin{pmatrix} a_{1j_1} \\ \vdots \\ a_{mj_1} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1j_r} \\ \vdots \\ a_{mj_r} \end{pmatrix}$$

linear unabhängig im K^m, wobei $j_1, \ldots, j_r$ die Pivotspalten sind, vergleiche [Fischer], S. 83 oben.

Proposition 3.3. *Es sei $(v_i)_{i\in I}$ eine Familie von Vektoren aus einem Vektorraum V. Dann gilt:*

a) *(v_i) ist linear unabhängig $\Leftrightarrow$ jeder Vektor in V kann auf höchstens eine Weise als Linearkombination in den v_i dargestellt werden.*
b) *(v_i) ist ein Erzeugendensystem von $V \Leftrightarrow$ jeder Vektor in V kann auf mindestens eine Weise als Linearkombination in den v_i dargestellt werden.*
c) *(v_i) ist eine Basis von $V \Leftrightarrow$ jeder Vektor in V kann in eindeutiger Weise als Linearkombination in den v_i dargestellt werden.*

Beweis. Angenommen, (v_i) ist nicht linear unabhängig. Dann kann der Vektor $0 \in V$ auf mindestens zwei Arten als Linearkombination in den v_i dargestellt werden: Als triviale Linearkombination (alle Koeffizienten gleich Null) und als nichttriviale Linearkombination (mindestens ein Koeffizient von Null verschieden), wegen der linearen Abhängigkeit der (v_i).

Es sei nun (v_i) linear unabhängig, $v \in V$ und $v = \sum \lambda_i v_i = \sum \lambda_i' v_i$ (die Summation ist immer über $i \in I$). Subtraktion der Gleichungen führt auf $\sum(\lambda_i - \lambda_i')v_i = 0$ und wegen der linearen Unabhängigkeit folgt $\lambda_i - \lambda_i' = 0$ für alle $i \in I$. Teil a) ist damit gezeigt.

Teil b) folgt direkt aus der Definition von Erzeugendensystem. Teil c) folgt aus a) und b). □

Folgendes Kriterium für lineare Unabhängigkeit ist ebenfalls nützlich.

Proposition 3.4. *Eine Familie (v_i) aus Vektoren des Vektorraums V ist genau dann linear abhängig, falls es ein $i_0 \in I$ gibt, so dass v_{i_0} als Linearkombination in den v_i mit $i \neq i_0$ geschrieben werden kann, d.h. $v_{i_0} \in \operatorname{span}(v_i)_{i\in I, i\neq i_0}$. Der Fall $I \setminus \{i_0\} = \emptyset$ ist erlaubt. Insbesondere gilt also:*

- *Falls I aus nur einem Element besteht, so ist $(v_i)_{i\in I} = (v)$ genau dann linear unabhängig, falls $v \neq 0$.*
- *Gibt es ein $i \in I$ mit $v_i = 0$, so ist $(v_i)_{i\in I}$ linear abhängig.*
- *Gibt es $i, j \in I$ mit $i \neq j$ und $v_i = v_j$, so ist $(v_i)_{i\in I}$ linear abhängig.*
- *Ist $I' \subset I$ eine Teilmenge und ist $(v_i)_{i\in I'}$ linear abhängig, so ist auch $(v_i)_{i\in I}$ linear abhängig.*

Beweis. Existiert ein i_0, so dass $v_{i_0} = \sum_{i\neq i_0} \lambda_i v_i$ (falls $I \setminus \{i_0\} = \emptyset$, so ist die rechte Seite gleich 0), so haben wir $1 \cdot v_{i_0} - \sum_{i\neq i_0} \lambda_i v_i = 0$ und wegen $1 \neq 0$ ist (v_i) somit linear abhängig.

Umgekehrt sei (v_i) linear abhängig, also $\sum \lambda_i v_i = 0$ mit einem $i_0 \in I$ mit $\lambda_{i_0} \neq 0$. Indem wir die Gleichung $\lambda_{i_0} v_{i_0} = -\sum_{i\neq i_0} \lambda_i v_i$ durch λ_{i_0} teilen, können wir v_{i_0} als Linearkombination in den v_i, $i \neq i_0$, schreiben. □

Folgende Charakterisierung von Basen ist ebenfalls wichtig.

Proposition 3.5. *Es sei (v_i) ein Familie von Vektoren in V. Dann sind äquivalent:*

a) *(v_i) ist eine Basis von V.*
b) *(v_i) ist ein* unverkürzbares Erzeugendensystem *von V, d.h. (v_i) ist ein Erzeugendensystem und für jede echte Teilmenge $I' \subsetneq I$ ist $(v_i)_{i\in I'}$ kein Erzeugendensystem.*
c) *(v_i) ist* unverlängerbar linear unabhängig*, d.h. (v_i) ist linear unabhängig und ist $I' \supset I$ eine echte Obermenge, d.h. $I \subsetneq I'$, und erweitern wir $(v_i)_{i\in I}$ zu einer Familie $(v_i)_{i\in I'}$, so ist $(v_i)_{i\in I'}$ linear abhängig.*

Beweis. Es sei (v_i) eine Basis. Dann ist nach Definition (v_i) ein Erzeugendensystem und linear unabhängig.

Angenommen, (v_i) ist ein verkürzbares Erzeugendensystem. Es sei $I' \subsetneq I$ und $(v_i)_{i\in I'}$ ebenfalls Erzeugendensystem. Ist $i_0 \in I \setminus I'$, so können wir somit v_{i_0} als Linearkombination in den $(v_i)_{i\neq i_0}$ schreiben. Dann ist aber $(v_i)_{i\in I}$ linear abhängig und somit keine Basis.

Angenommen, (v_i) ist verlängerbar linear unabhängig. Sei $I \subsetneq I'$ und $(v_i)_{i\in I'}$ ebenfalls linear unabhängig. Es sei $i_0 \in I' \setminus I$. Dann kann v_{i_0} keine Linearkombination in den v_i, $i \in I$ sein, d.h. (v_i) kein Erzeugendensystem und somit keine Basis.

Die Implikationen a) $\Rightarrow$ b) und a) $\Rightarrow$ c) sind damit klar.

Es gelte nun b). Wir müssen zeigen, dass (v_i) linear unabhängig ist. Angenommen (v_i) ist linear abhängig. Dann gibt es ein $i_0 \in I$, so dass $v_{i_0} = \sum_{i\neq i_0} \mu_i v_i$. Dann ist aber $(v_i)_{i\in I\setminus\{i_0\}}$ ebenfalls Erzeugendensystem von V, denn wegen

$$\sum_{i\in I} \lambda_i v_i = \sum_{i\neq i_0} \lambda_i v_i + \lambda_{i_0} \sum_{i\neq i_0} \mu_i v_i$$

ist jede Linearkombination in den v_i, $i \in I$ auch eine Linearkombination in den v_i, $i \in I \setminus \{i_0\}$. Also ist $(v_i)_{i\in I}$ doch verkürzbar, im Widerspruch zu b).

Schließlich gelte c). Wir müssen zeigen, dass (v_i) ein Erzeugendensystem ist. Angenommen, dies ist nicht so. Dann gibt es ein $v \in V$, das nicht Linearkombination der v_i, $i \in I$ ist. Wir setzen nun $I' := I \cup \{i_0\}$ und $v_{i_0} := v$. Wir behaupten, dass $(v_i)_{i\in I'}$ immer noch linear unabhängig ist. Sei $\sum_{i\in I'} \lambda_i v_i = 0$. Falls $\lambda_{i_0} \neq 0$, so können wir diese Gleichung nach v_{i_0} auflösen und somit v_{i_0} doch als Linearkombination in den v_i, $i \in I$, schreiben. Also ist $\lambda_{i_0} = 0$. Dann ist aber $0 = \sum_{i\in I'} \lambda_i v_i = \sum_{i\in I} \lambda_i v_i$ und damit auch $\lambda_i = 0$ für alle $i \in I$, denn $(v_i)_{i\in I}$ ist nach Annahme linear unabhängig. Insgesamt ist also $(v_i)_{i\in I'}$ linear unabhängig und somit $(v_i)_{i\in I}$ doch verlängerbar, im Widerspruch zu c). □

Die letzten drei Propositionen sind fundamental und werden in Zukunft ohne weitere Referenz benutzt.

Definition. *Ein Vektorraum V heißt* endlichdimensional*, falls er ein endliches Erzeugendensystem besitzt. Ansonsten heißt er* unendlichdimensional.

Satz 3.6 (Basisauswahlsatz)**.** *Es sei V ein endlichdimensionaler Vektorraum mit Erzeugendensystem $(v_1, \ldots, v_k)$. Dann bildet eine Teilfamilie dieser Familie eine Basis von V. Insbesondere hat jeder endlichdimensionale Vektorraum eine Basis.*

Beweis. Falls die gegebene Familie noch keine Basis ist, so ist sie ein verkürzbares Erzeugendensystem. D.h. wir können ein v_i, $i = 1, \ldots, k$, wegnehmen, so dass die verbleibende Familie $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ immer noch ein Erzeugendensystem ist. Entweder diese reduzierte Familie ist nun linear unabhängig und somit eine Basis, oder wir wiederholen obiges Argument. Nach endlich vielen Schritten erhalten wir auf diesem Weg eine linear unabhängige Familie und somit eine Basis von V, denn spätestens nach k Schritten sind wir bei der leeren Familie angekommen und diese ist linear unabhängig. □

Dieser Beweis funktioniert nicht, wenn das Erzeugendensystem von V nicht endlich ist: Wenn wir eine unendliche Menge um ein Element reduzieren, so bleibt diese Menge unendlich und der Prozess des schrittweisen Verkleinerns eines gegebenen Erzeugendensystems resultiert eventuell nie in einer linear unabhängigen Familie.

25.11.09

Zum Basisauswahlsatz gibt es ein Gegenstück, der *Basisergänzungssatz*. Das nächste Resultat bildet das Fundament dazu:

Proposition 3.7. *Es sei V ein endlichdimensionaler K-Vektorraum mit Basis $(v_1, \ldots, v_n)$, $n \in \mathbb{N}$ (falls $n = 0$, meinen wir hier die leere Familie). Dann ist jede Familie in V, die aus mindestens $n+1$ Vektoren besteht, linear abhängig.*

Beweis. Es genügt, Familien $(w_1, \ldots, w_{n+1})$ bestehend aus genau $n+1$ Vektoren in V zu betrachten. Jeder Vektor w_i, $i = 1, \ldots, n+1$ hat eine Darstellung der Form

$$w_i = \lambda_{1i} v_1 + \ldots + \lambda_{ni} v_n$$

mit $\lambda_{1i}, \ldots, \lambda_{ni} \in K$. Wir studieren die Gleichung

$$\mu_1 w_1 + \ldots + \mu_{n+1} w_{n+1} = 0\,,$$

wobei wir $\mu_1, \ldots, \mu_{n+1}$ als Unbekannte in K betrachten. Diese Gleichung ist äquivalent zur Gleichung

$$\big(\mu_1 \lambda_{11} + \ldots + \mu_{n+1} \lambda_{1\,n+1}\big) v_1 + \ldots + \big(\mu_1 \lambda_{n1} + \ldots + \mu_{n+1} \lambda_{n\,n+1}\big) v_n = 0$$

und diese - wegen der linearen Unabhängigkeit der Familie $(v_1, \ldots, v_n)$ - äquivalent zum Gleichungssystem

$$\begin{aligned} \lambda_{11}\mu_1 + \ldots + \lambda_{1\,n+1}\mu_{n+1} &= 0 \\ \lambda_{21}\mu_1 + \ldots + \lambda_{2\,n+1}\mu_{n+1} &= 0 \\ \vdots \qquad\qquad & \quad \vdots \\ \lambda_{n1}\mu_1 + \ldots + \lambda_{n\,n+1}\mu_{n+1} &= 0 \end{aligned}$$

mit den Koeffizienten (λ_{ij}) und Unbestimmten $\mu_1, \dots, \mu_{n+1}$. Dies ist ein homogenes Gleichungssystem aus n Gleichungen in den Unbestimmten $\mu_1, \dots, \mu_{n+1}$. Da es mehr Unbestimmte als Gleichungen gibt, hat dieses System eine nichttriviale Lösung $(\mu_1, \dots, \mu_{n+1}) \neq (0, \dots, 0)$, vgl. Korollar 2.4. Damit ist die Familie $(w_1, \dots, w_{n+1})$ linear abhängig. □

Wir notieren die folgende wichtige Folgerung:

Korollar 3.8. *Es sei V ein endlichdimensionaler Vektorraum. Dann sind alle Basen von V endlich und haben die gleiche Länge (d.h. bestehen aus der gleichen Anzahl an Vektoren).*

Beweis. Es sei $(v_1, \dots, v_n)$ eine Basis von V (falls $n = 0$, ist dies wieder die leere Familie). Ist nun $(w_i)_{i \in I}$ ebenfalls eine Basis, so ist nach der letzten Proposition I endlich mit $|I| \leq n$. Daraus folgt nun umgekehrt (wieder mit der letzten Proposition), dass $n \leq |I|$. □

In [Fischer] und vielen anderen Lehrbüchern über lineare Algebra wird dieses Resultat mit Hilfe des Austauschsatzes von Steinitz gezeigt. Das obige Argument benutzt hingegen nur die elementare Lösungstheorie linearer Gleichungssysteme basierend auf dem Gaußschen Eliminationsverfahren.

Definition. *Es sei V ein K-Vektorraum. Falls V endlichdimensional ist, so hat V nach dem Basisauswahlsatz eine Basis $(v_1, \dots, v_n)$ und wir definieren* $\dim V := n$, *die* Dimension *von V, als die Länge dieser Basis.*

Falls V nicht endlichdimensional ist, setzen wir $\dim V := \infty$.

Möchte man betonen, über welchem Grundkörper man arbeitet, so schreibt man $\dim_K V$ *statt* $\dim V$.

Für den Null-Vektorraum 0 *(der nur aus dem neutralen Element für die Addition besteht) über K erhalten wir wir*

$$\dim_K 0 := 0\,,$$

denn die leere Familie bildet eine Basis von 0.

Da alle Basen in einem endlichdimensionalen Vektorraum die gleiche Länge haben, ist diese Definition sinnvoll (d.h. $\dim V$ hängt nicht von der Wahl der Basis $(v_1, \dots, v_n)$ ab).

Ist V ein n-dimensionaler Vektorraum, so wählen wir als Indexfamilie I für Basen von V in der Regel die Menge $\{1, \dots, n\}$. Dies wird auch durch die (bereits oben verwendete) Tupelschreibweise $(v_1, \dots, v_n)$ deutlich. Ist also zum Beispiel $\dim V = 3$ und ist (u, v, w) eine Basis von V, dann ist auch (v, u, w) eine Basis, diese ist jedoch von (u, v, w) zu unterscheiden, da die zu Grunde liegenden Tripel von Vektoren verschieden sind.

Beispiel. Als Beispiel haben wir $\dim_K K^n = n$ für alle $n \in \mathbb{N}$

Satz 3.9 (Basisergänzungssatz)**.** *Es sei V ein endlichdimensionaler Vektorraum der Dimension n und $(v_1, \ldots, v_k)$ eine linear unabhängige Familie. Dann gilt $k \leq n$. Falls $k = n$, so ist die gegebene Familie bereits eine Basis von V. Falls $k < n$, so lässt sich diese Familie durch Vektoren $v_{k+1}, \ldots, v_n \in V$ zu einer Basis von V ergänzen.*

Beweis. Die Aussage $k \leq n$ haben wir bereits weiter oben gezeigt. Falls $k = n$, so ist $(w_1, \ldots, w_k)$ nicht verlängerbar, denn je $n + 1$ Vektoren in V sind linear abhängig. Somit ist $(w_1, \ldots, w_k)$ eine Basis. Falls $k < n$, so kann $(w_1, \ldots, w_k)$ keine Basis sein, denn alle Basen haben die gleiche Länge. Somit ist $(w_1, \ldots, w_k)$ linear unabhängig verlängerbar. Indem wir dieses Argument wiederholen, sehen wir, dass wir nach genau $n - k$ solcher Verlängerungen zu einer Basis von V gelangen. □

Korollar 3.10. *Es sei V ein endlichdimensionaler Vektorraum und $W \subset V$ ein Untervektorraum. Dann ist auch W endlichdimensional und es gilt* $\dim W \leq \dim V$*. Falls* $\dim W = \dim V$*, so ist $W = V$.*

Beweis. Es sei $\dim V = n$. Angenommen, W ist nicht endlichdimensional. Dann finden wir eine Familie $(w_1, \ldots, w_{n+1})$ von in W linear unabhängigen Vektoren: Jede linear unabhängige Familie von Vektoren in W ist linear unabhängig verlängerbar (da sie keine Basis ist), somit können wir ausgehend von der leeren Familie die Vektoren $w_1, \ldots, w_{n+1}$ induktiv finden. Da die Familie $(w_1, \ldots, w_{n+1})$ auch in V linear unabhängig ist (man mache sich dies klar!) haben wir einen Widerspruch zu Proposition 3.7. Somit ist W doch endlichdimensional.

Es sei $(w_1, \ldots, w_k)$ eine Basis von W. Diese Familie ist linear unabhängig in V, und somit folgen die weiteren Aussagen des Korollars aus dem Basisergänzungssatz. □

Der letzte Teil des Korollars gilt nicht für unendlichdimensionale Vektorräume. Auf dem Übungsblatt findet sich ein Gegenbeispiel dazu.

Da die Lösungsmengen homogener Gleichungssysteme mit n Unbestimmten Untervektorräume des K^n sind, sind diese Lösungsmengen also endlichdimensional und haben Dimension $\leq n$. Genauer gilt folgendes: Es sei ein homogenes lineares Gleichungssystem mittels einer Matrix $A \in K^{m \times n}$ gegeben.

Definition. *Wir defineren den* Zeilenrang *einer Matrix A, geschrieben* ZRang(A) *als die Dimension des von den Zeilenvektoren aufgespannten Untervektorraums des K^n.*

Proposition 3.11. *Elementare Zeilenumformungen von A verändern den Zeilenrang* ZRang(A) *nicht.*

Beweis. Es seien $h_1, \ldots, h_m \in K^n$ die Zeilenvektoren von A. Wir wollen zeigen:

$$\begin{aligned} \operatorname{span}(h_1, \ldots, h_i, \ldots, h_j, \ldots, h_m) &= \operatorname{span}(h_1, \ldots, h_j, \ldots, h_i, \ldots, h_m) \\ \operatorname{span}(h_1, \ldots, h_i, \ldots, h_j, \ldots, h_m) &= \operatorname{span}(h_1, \ldots, h_i, \ldots, h_j + \lambda h_i, \ldots, h_m) \end{aligned}$$

falls $1 \leq i \neq j \leq m$ und $\lambda \in K$. Wir zeigen also nicht nur, dass die Dimensionen der von den Zeilen aufgespannten Untervektorräumen gleich sind, sondern dass tatsächlich die Unterräume identisch sind. Die erste Gleichung ist klar. Für die zweite beachten wir die Gleichung

$$\mu_i h_i + \mu_j h_j = (\mu_i - \mu_j \lambda) h_i + \mu_j (h_j + \lambda h_i)$$

um eine beliebige Linearkombination der Vektoren in der Familie links in eine Linearkombination der Vektoren in der Familie rechts umzurechnen. Dies zeigt die Inklusion $\subset$. Die andere Inklusion sieht man entsprechend. □

Wenn wir A auf Zeilenstufenform gebracht haben und diese aus r Zeilen ungleich 0 besteht, so haben wir also

$$r = \mathrm{ZRang(A)}\,.$$

Insbesondere können wir nun eine der früher gestellten Fragen beantworten:

Proposition 3.12. *Ist ein homogenes lineares Gleichungssystem gegeben und bringen wir dieses auf Zeilenstufenform, so hängt die Anzahl der Pivotelemente (d.h. die Anzahl der von* 0 *verschiedenen Zeilen) nur vom linearen Gleichungssystem ab, aber nicht von der speziellen Zeilenstufenform.*

Wir besprechen nun Konstruktionen neuer Vektorräume aus schon gegebenen.

Definition. *Es sei $(V_i)_{i\in I}$ eine Familie von K-Vektorräumen. Die* direkte Summe $\bigoplus_{i\in I} V_i$ *ist der K-Vektorraum*

$$\{(v_i)_{i\in I} \mid v_i \in V_i, \text{ nur endlich viele } v_i \neq 0\}$$

versehen mit der komponentenweisen Addition und Skalarenmultiplikation.

Beispiel. Betrachten wir $I = \{1, 2\}$ mit den $\mathbb{R}$-Vektorräumen $V_1 = \mathbb{R}^n$ und $V_2 = \mathbb{R}^m$, so ist

$$\bigoplus_{i\in I} V_i = V_1 \oplus V_2 = \mathbb{R}^n \oplus \mathbb{R}^m = \mathbb{R}^{n+m}.$$

Proposition 3.13. *Es sei $(V_1, \ldots, V_r)$ eine endliche Familie endlichdimensionaler K-Vektorräume. Dann ist $V_1 \oplus \ldots \oplus V_r$ ebenfalls endlichdimensional und*

$$\dim_K(V_1 \oplus \ldots \oplus V_r) = \dim_K V_1 + \ldots + \dim_K V_r\,.$$

Beweis. Es sei $n_i := \dim V_i$ und $(v_{i1}, \ldots, v_{in_i})$ sei eine Basis von V_i , $i = 1, \ldots, r$. Dann ist die Familie

$$(v_{ij})_{1\leq i\leq r, 1\leq j\leq n_i}$$

eine Basis von $V_1 \oplus \ldots \oplus V_r$. Diese hat offensichtlich $n_1 + \ldots + n_r$ Elemente. Wir identifizieren dabei das Element $v_{ij} \in V_i$ mit dem Element $(0, \ldots, v_{ij}, \ldots, 0) \in V_1 \oplus \ldots \oplus V_r$, wobei v_{ij} an der i-ten Stelle steht. □

30.11.09

Wir haben hier mit mehreren verschiedenen Vektorräumen hantiert. Im folgenden betrachten wir Untervektorräume eines fest gegebenen Vektorraumes.

Definition. *Es sei V ein K-Vektorraum und $(W_i)_{i\in I}$ eine Familie von Untervektorräumen von V. Die* Summe *der W_i ist definiert als der Untervektorraum*

$$\sum_{i\in I} W_i := \operatorname{span}(\bigcup_{i\in I} W_i) \subset V$$

Falls die gegebene Familie endlich ist, d.h. von der Form $(W_1, \ldots, W_r)$, so schreiben wir für die Summe

$$W_1 + \ldots + W_r$$

Die Summe $\sum_{i\in I} W_i$ ist der kleinste Untervektorraum von V, der alle W_i, $i \in I$, enthält. Man kann leicht zeigen, dass $\sum_{i\in I} W_i$ genau aus denjenigen Vektoren $v \in V$ besteht, die sich als endliche Summe $v = w_{i_1} + \ldots + w_{i_k}$ schreiben lassen, wobei $w_{i_j} \in W_{i_j}$, $1 \le j \le k$ und $k \ge 1$.

Es stellt sich die Frage, wann diese Darstellungen eindeutig sind.

Definition. *Es sei $(W_i)_{i\in I}$ eine Familie von Untervektorräumen von V. Wir nennen diese Familie* direkt*, falls folgendes gilt: Ist $(w_i)_{i\in I}$ eine Familie von Vektoren $w_i \in W_i$, von denen nur endlich viele ungleich 0 sind, und gilt*

$$\sum_{i\in I} w_i = 0\,,$$

so folgt $w_i = 0$ für alle $i \in I$.

Ist V ein Vektorraum und ist $(v_i)_{i\in I}$ eine Familie von Vektoren in V, so sind also äquivalent:

- Die Familie (v_i) ist linear unabhängig.
- Die Familie von Untervektorräumen $(\operatorname{span}(v_i))_{i\in I}$ von V ist direkt.

Ganz ähnlich wie für linear unabhängige Familien zeigt man:

Proposition 3.14. *Es sei $(W_i)_{i\in I}$ eine Familie von Untervektorräumen von V. Dann sind äquivalent:*

a) *Die Familie (W_i) ist direkt.*
b) *Jeder Vektor $w \in \sum_{i\in I} W_i$ besitzt eine eindeutige Darstellung als $\sum_{i\in I} w_i$ mit $w_i \in W_i$ und $w_i \neq 0$ für nur endlich viele i.*
c) *Für alle $i_0 \in I$ gilt $W_{i_0} \cap (\sum_{i\in I\setminus\{i_0\}} W_i) = 0$*

Beweis. Angenommen, die Familie (W_i) ist direkt. Es sei

$$w = \sum_{i\in I} w_i = \sum_{i\in I} w_i'\,,$$

wobei in jeder Summe nur endlich viele Vektoren ungleich 0 sind und $w_i, w_i' \in W_i$ für alle $i \in I$. Dann haben wir $\sum_{i\in I}(w_i - w_i') = 0$ und aus der Direktheit der gegebenen Familie folgt $w_i = w_i'$ für alle $i \in I$.

Angenommen, die in b) beschriebene Eigenschaft ist erfüllt. Dann besitzt insbesondere der Vektor $0 \in \sum_i W_i$ eine endeutige Darstellung und damit ist die Familie $(W_i)_{i\in I}$ direkt.

Die Äquivalenz von a) und b) zu c) wird im Tutorium behandelt. □

Ist (W_i) eine direkte Familie von Untervektorräumen, so wird ihre Summe auch mit

$$\bigoplus_{i\in I} W_i$$

bezeichnet und diese Summe wird *direkt* genannt. Falls wir eine endliche direkte Familie $(W_1, \ldots, W_r)$ vorliegen haben, so schreiben wir

$$W_1 \oplus \ldots \oplus W_r \,.$$

Diese Sprechweise und insbesondere der Zusammenhang zu der vorhin eingeführten direkten Summe einer beliebigen Familie von K-Vektorräumen wird im nächsten Abschnitt geklärt (siehe das Beispiel auf S. 44).

Wir betrachten nun endliche Familien von Untervektorräumen und stellen einige Dimensionsberechnungen an.

Proposition 3.15. *Es seien $W_1, \ldots, W_r \subset V$ endlichdimensionale Untervektorräume. Dann ist deren Summe $W_1 + \ldots + W_r$ ebenfalls endlichdimensional und es gilt die Ungleichung*

$$\dim(W_1 + \ldots + W_r) \leq \dim W_1 + \ldots + \dim W_r \,.$$

In dieser Formel tritt genau dann Gleichheit ein, falls die Familie $(W_1, \ldots, W_r)$ direkt ist.

Beweis. Es sei $n_i := \dim W_i$ und $(w_{i1}, \ldots, w_{in_i})$ eine Basis von W_i, $1 \leq i \leq r$. Dann ist $(w_{11}, \ldots, w_{1n_1}, \ldots, w_{r1}, \ldots, w_{rn_r})$ (dies ist eine Familie, keine Vereinigungsmenge!) ein Erzeugendensystem von $W_1 + \ldots + W_r$. Dies zeigt die Endlichdimensionalität der Summe und aus dem Basisauswahlsatz folgt die behauptete Ungleichung.

Es gilt genau dann Gleichheit, falls diese Familie nicht nur ein Erzeugendensystem, sondern darüberhinaus linear unabhängig ist.

Angenommen, diese Familie ist linear unabhängig. Wir behaupten, dass dann die Familie $(W_1, \ldots, W_r)$ direkt ist. Wir nehmen dazu $0 = w_1 + \ldots + w_r$ an mit $w_i \in W_i$ für $i = 1, \ldots, r$. Jedes w_i kann in eindeutiger Weise als Linearkombination

$$w_i = \sum_{j=1}^{n_i} \lambda_{ij} w_{ij}$$

geschrieben werden. Setzen wir dies in die vorige Gleichung ein, so folgt aus der linearen Unabhängigkeit der Familie $(w_{1n_1}, \ldots, w_{11}, \ldots, w_{r1}, \ldots, w_{rn_r})$, dass alle $\lambda_{ij} = 0$, $1 \leq i \leq r$, $1 \leq j \leq n_i$, und somit $w_1 = \ldots = w_r = 0$.

Sei umgekehrt die Familie $(W_1, \ldots, W_r)$ direkt. Haben wir

$$\sum_{i=1}^{r} \sum_{j=1}^{n_i} \lambda_{ij} w_{ij} = 0$$

so folgt daraus

$$\sum_{j=1}^{n_i} \lambda_{ij} w_{ij} = 0$$

für alle $i = 1, \ldots, r$. Und aus der linearen Unabhängigkeit der Familie $(w_{i1}, \ldots, w_{in_i})$ für alle $i = 1, \ldots, r$ erhalten wir $\lambda_{ij} = 0$ für alle $i = 1, \ldots, r$, $j = 1, \ldots, n_i$. Somit ist die Familie $(w_{1n_1}, \ldots, w_{11}, \ldots, w_{r1}, \ldots, w_{rn_r})$ linear unabhängig. □

Wir untersuchen noch die Dimension einer Summe für den Fall, wenn die Summe der Untervektorräume nicht direkt ist.

Proposition 3.16. *Es sei $W_1, W_2 \subset V$ endlichdimensionale Untervektorräume. Dann gilt die Gleichung*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)\,.$$

(Hier beachte man, dass der Schnitt einer Familie von Untervektorräumen eines gegebenen Vektorraumes wieder ein Untervektorraum ist).

Der Beweis findet sich in [Fischer], S. 100 f.

Definition. *Es sei V ein K-Vektorraum und $W \subset V$ ein Untervektorraum. Wir nennen einen Untervektorraum $X \subset V$ ein* Komplement *von W, falls $V = W \oplus X$. (Diese Schreibweise deutet insbesondere an, dass die Familie (W, X) direkt ist).*

Ist V endlichdimensional und $W \subset V$ ein Untervektorraum, so zeigt der Basisergänzungssatz, dass W ein Komplement besitzt. Dieses ist jedoch fast nie eindeutig: Der Untervektorraum $\mathbb{R} \times 0 \subset \mathbb{R}^2$ hat die Komplemente $\mathrm{span}((0,1))$ und $\mathrm{span}((1,1))$.

Wir besprechen noch eine weitere Konstruktion neuer Vektorräume aus bereits gegebenen. Diese ist abstrakter als die bisherigen.

Es sei V ein K-Vektorraum und $W \subset V$ ein Untervektorraum. Wir definieren eine Relation $\sim$ auf V wie folgt:

$$v_1 \sim v_2 :\Leftrightarrow v_1 - v_2 \in W$$

Es ist nicht schwer zu sehen, dass es sich um eine Äquivalenzrelation handelt. Die Menge der Äquivalenzklassen bezeichnen wir mit V/W. Ist $v \in V$, so ist also die Äquivalenzklasse $[v]$ von V gleich der Menge

$$v + W := \{v + w \mid w \in W\}$$

und wir haben genau dann $v_1 + W = v_2 + W$, falls $v_1 - v_2 \in W$.

Proposition 3.17. *Durch die Setzung*

$$\begin{aligned} (v_1 + W) + (v_2 + W) &:= v_1 + v_2 + W \\ \lambda \cdot (v + W) &:= \lambda v + W \end{aligned}$$

wird V/W zu einem K-Vektorraum.

Beweis. Es ist leicht zu überprüfen, dass die gegebenen Verknüpfungen wohldefiniert sind. Die Vektorraumaxiome folgen mit direkten Rechnungen. □

Wir nennen diesen Vektorraum den *Quotientenvektorraum* von V nach W.

Definition. *Es sei V ein Vektorraum. Ein* affiner Unterraum *von V ist eine Teilmenge von V der Form $v + W$, wobei $v \in V$ und $W \subset V$ ein Untervektorraum ist. Der Punkt $v \in V$ heißt auch* Aufhängepunkt *des affinen Unterraumes.*

Man beachte, dass ein affiner Unterraum in der Regel kein Untervektorraum von V ist. Dies ist genau dann der Fall, falls $0 \in v + W$, d.h. falls $v \in W$.

Die Äquivalenzklassen der oben eingeführten Äquivalenzrelation auf V sind nach Definition affine Unterräume von V. Diese affinen Unterräume sind genau die Elemente des Quotientenvektorraumes V/W.

7.12.09

Ist $W \subset V$ ein Untervektorraum und sind $v_1, v_2 \in V$, so nennen wir die affinen Unterräume $v_1 + W$ und $v_2 + W$ *parallel*.

Proposition 3.18. *Es sei V ein K-Vektorraum und $A \subset V$ ein affiner Unterraum. Schreiben wir $A = v+W$ mit $v \in V$ und einem Untervektorraum $W \subset V$, so ist W durch A eindeutig bestimmt. Der Vektor v ist eindeutig bestimmt bis auf Addition von Vektoren in W.*

Insbesondere hat jeder affine Unterraum $A = v + W \subset V$ einen eindeutig bestimmten parallelen Untervektorraum $W = 0 + W \subset V$.

Beweis. Es sei

$$A = v_1 + W_1 = v_2 + W_2\,,$$

wobei $v_1, v_2 \in V$ und $W_1, W_2 \subset V$ Untervektorräume sind. Da $v_1 \in v_2 + W_2$, folgt $v_1 - v_2 \in W_2$, also $[v_1] = [v_2] \in W/W_2$ und somit $v_1 + W_2 = v_2 + W_2$. Wir erhalten $A = v_1 + W_1 = v_1 + W_2$ und somit

$$W_1 = \{v \in V \mid \exists a \in A \text{ mit } v = a - v_1\} = W_2$$

wie behauptet. Setzen wir $W := W_1 = W_2$, so haben schon weiter oben gesehen, dass $v_1 - v_2 \in W$. Dies zeigt die zweite Behauptung. □

Damit ist die folgende Definition sinnvoll.

Definition. *Es sei $A = a + W \subset V$ ein affiner Unterraum. Die* Dimension *von A wird definiert als die Dimension von W (W ist ja durch A eindeutig bestimmt). Affine Unterräume der Dimension* 1 *heißen* Geraden*, der Dimension* 2 Ebenen *und der Dimension $n-1$* Hyperebenen *in V.*

Wir können jetzt die Struktur der Lösungsmengen inhomogener linearer Gleichungssysteme folgendermaßen beschreiben.

Proposition 3.19. *Es sei ein lineares Gleichungssystem durch die erweiterte Koeffizientenmatrix $(A|b) \in K^{m\times(n+1)}$ gegeben. Es sei $L \subset K^n$ die Lösungsmenge dieses linearen Gleichungssysstems. Dann tritt genau einer der folgenden beiden Fälle ein:*

- $L = \emptyset$.
- $L \subset K^n$ *ist ein affiner Teilraum mit* $\dim L = \dim L'$, *wobei* $L' \subset K^n$ *die Lösungsmenge des zugehörigen durch A gegebenen homogenen Systems ist (insbesondere ist $L' \subset K^n$ ein Untervektorraum). Genauer gilt folgendes: Ist $l \in V$ eine (beliebige) Lösung des inhomogenen Systems, so gilt $L = l + L'$.*

Beweis. Es sei $L \neq \emptyset$ und $l \in L$ eine Lösung des durch $(A|b)$ gegebenen lineraren Gleichungssystems. Ist nun $v \in V$, so gilt $v \in L$ genau dann, wenn $v - l$ das zugehörige homogene System löst. Somit ist $L = l + L'$ wie behauptet. □

Wir werden in Kürze sehen, dass man umgekehrt jeden affinen Teilraum in K^n als Lösungsmenge eines linearen Gleichungssystems beschreiben kann.

Wir zitieren am Schluss dieses Abschnitts noch das folgende Resultat über unendlichdimensionale Vektorräume. Diese lassen sich aus dem Auswahlaxiom der Mengenlehre (bzw. dem dazu äquivalenten Zornschen Lemma) ableiten, worauf wir allerdings an dieser Stelle nicht weiter eingehen.

Proposition 3.20. *Jeder K-Vektorraum V besitzt eine Basis. Jede linear unabhängige Familie in V lässt sich zu einer Basis ergänzen. Insbesondere hat jeder Untervektorraum von V ein Komplement. Jedes Erzeugendensystem in V enthält eine Basis.*

Da $\mathbb{R}$ ein Vektorraum über $\mathbb{Q}$ ist, besitzt also zum Beispiel $\mathbb{R}$ auch eine Basis als $\mathbb{Q}$-Vektorraum! Diese kann man sich kaum konkret vorstellen.

4. Lineare Abbildungen und Matrizen

Wie wir schon bei der Diskussion der algebraischen Grundstrukturen gesehen haben, sind nicht nur diese Strukturen selbst, sondern auch die strukturerhaltenden Abbildungen zwischen ihnen wichtig.

Definition. *Es seien K ein Körper und V und W Vektorräume über K. Eine Abbildung $f : V \to W$ heißt* linear *oder* Vektorraumhomomorphismus, *falls*

- $f(v + w) = f(v) + f(w)$ *für alle $v, w \in V$, d.h. f ist ein Gruppenhomomorphismus $(V, +, 0) \to (W, +, 0)$.*
- $f(\lambda v) = \lambda f(v)$ *für alle $v \in V$ und $\lambda \in K$.*

Ist f zusätzlich bijektiv, so nennt man f einen Vektorraumisomorphismus. *Lineare Abbildungen $V \to V$ heißen* Endomorphismen *von V, bijektive Endomorphismen nennt man* Automorphismen. *Existiert ein Isomorphismus $V \to W$, so heißen V und W* isomorph *und wir schreiben $V \cong W$.*

Man zeigt leicht, dass die Umkehrabbildung eines Vektorraumisomorphismus wieder linear und somit ebenfalls ein Vektorraumisomorphismus ist. Ist $f : V \to W$ linear und $(v_i)_{i\in I}$ eine Familie von Vektoren in V, so gilt für Linearkombinationen in V

$$f(\sum_{i\in I} \lambda_i v_i) = \sum_{i\in I} \lambda_i f(v_i)$$

Proposition 4.1. *Es sei $(v_i)_{i\in I}$ eine Basis des K-Vektorraumes V und W ein beliebiger K-Vektorraum. Es sei $(w_i)_{i\in I}$ eine ebenfalls durch I parametrisierte Familie von Vektoren in W. Dann existiert genau eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für alle i.*

Mit anderen Worten: Eine lineare Abbildung ist eindeutig dadurch festgelegt, dass man ihre Werte auf einer Basis vorgibt. Diese Vorgabe ist beliebig.

Beweis. Ist $v \in V$, so besitzt v eine eindeutige Darstellung als Linearkombination $v = \sum_{i\in I} \lambda_i v_i$. Wir setzen nun

$$f(v) := \sum_{i\in I} \lambda_i w_i \,.$$

Man prüft leicht nach, dass die so definierte Abbildung $f : V \to W$ linear ist. Andererseits muss jede lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für alle $i \in I$ die obige Gleichung erfüllen. Daraus folgt die Eindeutigkeit. □

Mit Hilfe dieser Proposition sehen wir: Es sei V ein endlichdimensionaler K-Vektorraum mit Basis $\mathcal{B} = (v_1, \ldots, v_n)$. Dann gibt es genau eine lineare Abbildung $\Phi_\mathcal{B} : K^n \to V$, die den i-ten kanonischen Basisvektor $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ (mit 1 an der $i$ten Stelle) auf v_i abbildet. Diese Abbildung erfüllt die Gleichung

$$\Phi_\mathcal{B}(\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} \lambda_i v_i \,.$$

Da $(v_1, \ldots, v_n)$ eine Basis ist, ist $\phi_\mathcal{B}$ bijektiv, also ein Vektorraumisomorphismus. Insbesondere ist also $K^n \cong V$.

Wir nennen $\Phi_\mathcal{B}$ das zur Basis $\mathcal{B}$ gehörende *Koordinatensystem* (hier wird klar, warum wir K^n als den n-dimensionalen Koordinatenraum bezeichnet haben). Wir nennen die Koeffizienten $(\lambda_1, \ldots, \lambda_n)$ in einer Linearkombination $v = \sum_{i=1}^{n} \lambda_i v_i$ die *Koordinaten* von v bezüglich der Basis $\mathcal{B}$. Man beachte, dass $\Phi_\mathcal{B}$ von der Basis $\mathcal{B}$ abhängt.

Beispiel.

- Es sei V ein beliebiger K-Vektorraum und 0 der Nullvektorraum (über K). Dann ist die Abbildung $V \to 0$, $v \mapsto 0$, linear.
- Ist $\lambda \in K$, so ist die Abbildung $\mu \mapsto \lambda \cdot \mu$ eine lineare Abbildung $K \to K$, wobei wir K wir üblich als eindimensionalen Vektorraum über K ansehen.

- Ist allgemeiner V ein beliebiger K-Vektorraum und $\lambda \in K$, so ist die Abbildung $V \to V$, $v \mapsto \lambda v$ linear. Dies ist eine *Streckung* um den Faktor λ.
- Noch allgemeiner gilt folgendes: Es sei $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$. Dann ist die Abbildung $f : K^n \to K^m$ gegeben durch

$$(x_1, \dots x_n) \mapsto (\sum_{j=1}^{n} a_{1j} x_j, \dots, \sum_{j=1}^{n} a_{mj} x_j)$$

linear. Wir bezeichnen diese Abbildung ebenfalls mit dem Buchstaben A. Wenn wir lineare Abbildungen durch Matrizen darstellen, schreiben wir obige Zuordnung in der Regel in der Form

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{pmatrix}$$

Der Grund dafür wird weiter unten klar.

9.12.09

Wir können umgekehrt jede lineare Abbildung $f : K^n \to K^m$ durch eine $(m \times n)$-Matrix A beschreiben, indem wir den Wert $f(e_j) \in K^m$ des j-ten Einheitsvektors mit $1 \leq j \leq n$ in die j-te Spalte der Matrix A schreiben. Die so entstehende lineare Abbildung $A : K^n \to K^m$ stimmt dann auf der kanonischen Basis $(e_1, \dots, e_n)$ von K^n mit f überein, also dann auch auf ganz K^n. Es gilt also die Merkregel

Die j-te Spalte von A ist das Bild von e_j.

Zum Beispiel beschreibt die Matrix

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

eine Drehung des $\mathbb{R}^2$ um den Winkel t gegen den Uhrzeigersinn.

- Es sei ein homogenes lineares Gleichungssystems durch eine Matrix $A \in K^{m \times n}$ gegeben. Wir haben schon früher gesehen, dass sich die Lösungsmenge $L \subset K^n$ dieses Gleichungssystems bei elementaren Zeilenumformungen nicht ändert. Wir nehmen also an, dass A in Zeilenstufenform mit Pivotspalten $j_1, \dots, j_r$ gegeben ist. Dann ist die in Satz 2.3 konstruierte bijektive Abbildung

$$\phi : K^{n-r} \to L$$

linear (dies sollte man überprüfen!) und somit ein Isomorphismus von Vektorräumen. Eine Basis von L ist also gegeben, indem man nacheinander genau einen der $n-r$ freien Parameter x_j, $j \neq j_1, \dots, j_r$ gleich 1 und die anderen gleich 0 setzt und das Gleichungssystem für

diese Wahlen löst. Die so entstehenden $n - r$ Vektoren im $\mathbb{R}^n$ sind dann eine Basis von L.

- Es sei $C([0,1])$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen $[0,1] \to \mathbb{R}$. Dann ist die Abbidung
$$C([0,1]) \to \mathbb{R}\,, f \mapsto \int_0^1 f(x)dx$$
linear.
- Es sei $C^\infty(\mathbb{R})$ der Vektorraum der unendlich oft differenzierbaren Funktionen $\mathbb{R} \to \mathbb{R}$. Dann ist die Abbildung $\frac{d}{dx} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, $f \mapsto \frac{df}{dx}$, linear. Diese Abbildung ist sogar surjektiv, denn die Abbildung
$$\chi : \phi \mapsto f \text{ mit } f(x) := \int_0^x \phi(t)dt$$
ist rechtinsinvers zu $\frac{d}{dx}$ (d.h. $\frac{d}{dx} \circ \chi = \mathrm{id}$) nach dem Hauptsatz der Differential- und Integralrechnung. Die Abbildung ist aber nicht injektiv. Genauer gilt $\ker \frac{d}{dx} = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist konstant}\}$ wie in der Analysisvorlesung gezeigt werden wird.

Da jede lineare Abbildung $f : V \to W$ auch ein Gruppenhomomorphismus ist, sind - wie für jeden Gruppenhomomorphismus - die Teilmengen

- $\operatorname{im} f = f(V) \subset W$ (Bild von f) und
- $\ker f := f^{-1}(\{0\}) \subset V$ (Kern von f)

Untergruppen. Diese sind sogar Untervektorräume von W, bzw. V, denn ist z.B. $w = f(v) \in \operatorname{im} f$ und $\lambda \in K$, so haben wir $\lambda w = \lambda f(v) = f(\lambda v) \in \operatorname{im} f$.

Allgemeiner ist für jeden Untervektorraum $Z \subset W$ das Urbild $f^{-1}(Z) \subset V$ wieder ein Untervektorraum und für jeden Untervektorraum $U \subset V$ das Bild $f(V) \subset W$ ebenfalls ein Untervektorraum.

Ist $A \in K^{m\times n}$, so ist $\ker A \subset K^n$ genau die Lösungsmenge des durch A bestimmten homogenen Gleichungssystems. Mit Hilfe des Gaußschen Eliminationsverfahrens können wir also eine Basis von $\ker A$ bestimmen.

Ist $f : V \to W$ linear und $b \in W$, so ist das Urbild $f^{-1}(\{b\}) \subset V$ entweder leer oder ein affiner Unterraum der Dimension $\dim \ker f$. Denn ist $x \in f^{-1}(\{w\})$, so gilt für alle $v \in V$ die Äquivalenz
$$f(v) = b \Leftrightarrow f(v - x) = 0$$
und somit die Äquivalenz
$$v \in f^{-1}(\{w\}) \Leftrightarrow v \in x + \ker f\,.$$
Falls $A \in K^{m\times n}$, so ist $A^{-1}(\{b\}) \subset K^n$ genau die Lösungsmenge des durch $(A|b)$ gegebenen linearen Gleichungssysstems. Wir sehen also erneut, dass diese Lösungsmenge entweder leer oder ein affiner Teilraum ist, dessen Aufhängepunkt als irgendeine Lösung des gegebenen Systems gewählt werden kann und dessen Untervektorraum gerade die Lösungsmenge des zugehörigen homogenen Systems ist.

Da jede lineare Abbildung ein Gruppenhomomorphismus ist, folgt direkt aus Proposition 1.9

Proposition 4.2. *Es sei $f : V \to W$ linear. Dann ist f genau dann injektiv, wenn* $\ker f = 0 \subset V$.

Beispiel. Es sei V ein K-Vektorraum und $(W_i)_{i\in I}$ eine direkte Familie von Untervektorräumen (siehe S. 36). Dann ist die kanonische Abbildung

$$\bigoplus_{i\in I} W_i \to \sum_{i\in I} W_i \subset V\,, (w_i)_{i\in I} \mapsto \sum_{i\in I} w_i$$

ein Vektorraumisomorphismus. Dabei steht links die (abstrakte) direkte Summe der Vektorräume W_i (vgl. die Definition auf S. 35). Die Injektivität folgt aus der vorgehenden Proposition und der Definition von direkten Familien. Die Surjektivität folgt aus der Definition der Summe von Vektorräumen.

Damit ist der Zusammenhang der direkten Summe von abstrakten Vektorräumen und der Summe einer direkten Familie von Untervektorräumen eines festen Vektorraumes geklärt (man vergleiche die Bemerkung auf S. 37 oben).

Proposition 4.3. *Es sei $f : V \to W$ eine lineare Abbildung und $(v_i)_{i\in I}$ eine Familie von Vektoren in V.*

i. *Ist (v_i) linear abhängig in V, so ist $(f(v_i))$ linear abhängig in W.*
ii. *Ist (v_i) ein Erzeugendensystem von V, so ist $(f(v_i))$ ein Erzeugendensystem von $f(V)$, insbesondere gilt* $\dim f(V) \leq \dim V$ *(hier benutzen wir die Konvention $n \leq \infty$ für alle $n \in \mathbb{N} \cup \{\infty\}$).*
iii. *f ist genau dann injektiv, wenn f alle linear unabhängigen Familien von Vektoren in V in linear unabhängige Familien von Vektoren in W überführt.*
iv. *Es sei $(v_i)_{i\in I}$ eine Basis. Dann ist f genau dann ein Isomorphismus, wenn $(f(v_i))_{i\in I}$ eine Basis von W ist.*

Beweis. i. ist klar, da f eine Lineakombination von Vektoren in V in eine Linearkombination von Vektoren in W überführt (mit den gleichen Koeffizienten).

Für ii. sei $w = f(v) \in f(V)$. Wir können $v = \sum_{i\in I} \lambda v_i$ als Linearkombination schreiben und haben dann $f(v) = \sum \lambda_i f(v_i)$. Somit ist $(f(v_i))$ wirklich ein Erzeugendensystem von $f(V)$.

Zu iii. Es sei f injektiv und $(v_i)_{i\in I}$ eine linear unabhängige Familie in V. Angenommen $\sum_{i\in I} \lambda_i f(v_i) = 0$. Da $\sum_{i\in I} \lambda_i f(v_i) = f(\sum_{i\in I} \lambda_i v_i)$, folgt aus $\ker f = 0$, dass $\sum_{i\in I} \lambda_i v_i = 0$ und daraus wegen der linearen Unabhängigkeit von (v_i), dass alle $\lambda_i = 0$.

Wir kommen nun zu iv. Es sei f ein Isomorphismus. Dann ist $(f(v_i))$ nach ii. ein Erzeugendensystem von $f(V) = W$ und nach iii. linear unabhängig, also ist $(f(v_i))$ eine Basis von W. Falls umgekehrt $(f(v_i))$ eine Basis von W

ist, so ist f ein Isomorphismus: Es sei $v \in \ker f$. Wir schreiben $v = \sum_i \lambda_i v_i$ und die Gleichung $0 = f(v) = \sum_i \lambda_i f(v_i)$ und die Voraussetzung, dass $(f(v_i))$ linear unabhängig ist, zeigen $\lambda_i = 0$ für alle i. Somit ist $\ker f = 0$ und f injektiv. Da $(f(v_i))$ ein Erzeugendensystem von W ist, können wir jedes $w \in W$ in der Form $\sum_i \lambda_i f(v_i) = f(\sum_i \lambda_i v_i)$ schreiben und somit ist f auch surjektiv. □

Proposition 4.4. *Es seien V und W endlichdimensionale Vektorräume über K. Dann sind V und W genau dann isomorph, falls* $\dim V = \dim W$.

Beweis. Falls V und W isomorph sind, existiert ein Isomorphismus $f : V \to W$. Dieser bildet nach Proposition 4.3 iv. jede Basis von V auf eine Basis von W ab. Somit haben die Basen in V und W gleiche Länge.

Sei umgekehrt $\dim V = \dim W = n$. Nach Wahl von Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W haben wir Koordinatensysteme

$$\Phi_{\mathcal{B}} : K^n \xrightarrow{\cong} V \,, \Phi_{\mathcal{C}} : K^n \xrightarrow{\cong} W \,.$$

Die Komposition $\Phi_{\mathcal{C}} \circ \Phi_{\mathcal{B}}^{-1}$ ist dann ein Isomorphismus $V \xrightarrow{\cong} W$. □

14.12.09

Es seien V und W beliebige K-Vektorräume. Wir bezeichnen mit

$$\mathrm{Hom}(V, W) \subset \mathrm{Abb}(V, W)$$

die Menge der Vektorraumhomomorphismen $V \to W$. Die Menge $\mathrm{Hom}(V, W)$ ist mit der punktweisen Addition und Skalarenmultiplikation linearer Abbildungen (man überzeuge sich, dass diese Operationen lineare Abbildungen wieder in solche überführen) ein K-Vektorraum.

Falls $V = W$, so wird durch die Verknüpfung

$$\mathrm{Hom}(V, V) \times \mathrm{Hom}(V, V) \to \mathrm{Hom}(V, V) \,, \ (f, g) \mapsto f \circ g$$

die abelsche Gruppe $\mathrm{Hom}(V, V)$ zu einem Ring mit 1 (man beachte hier, dass die Verknüpfung zweier linearer Abbildungen wieder linear ist). Das Assoziativgesetz für die Multiplikation folgt aus der entsprechenden Regel für die Komposition von Abbildungen. Die Distributivgesetze, d.h. $(f + g) \circ h = f \circ h + g \circ h$ und $h \circ (f + g) = h \circ f + h \circ g$ für $f, g, h \in \mathrm{Hom}(V, V)$ rechnet man direkt nach. (Dabei braucht man für die zweite Gleichung die Linearität von h). Das Nullelement (d.h. das neutrale Element für die Addition) ist die Nullabbildung $V \to V$, $v \mapsto 0$, und das Einselement (d.h. das neutrale Element für die Komposition) ist die Identität $\mathrm{id}_V : V \to V$. Wir bezeichnen $(\mathrm{End}(V), +, 0, \circ, \mathrm{id}_V)$ als den *Endomorphismenring* von V.

Es sei $\mathrm{Aut}(V) \subset \mathrm{End}(V)$ die Teilmenge der Isomorphismen $V \to V$. Das Tripel $(\mathrm{Aut}(V), \circ, \mathrm{id}_V)$ bildet dann eine Gruppe, die *Automorphismengruppe* von V. Man beachte dass $\mathrm{Aut}(V)$ kein Ring ist, denn die Summe zweier Automorphismen braucht kein Automorphismus zu sein (z.B. sind $\mathrm{id}, -\mathrm{id} : \mathbb{R} \to \mathbb{R}$ Automorphismen von $\mathbb{R}$, aber $\mathrm{id} + (-\mathrm{id})$ ist die Nullabbildung).

Um ein konkreteres Bild dieser algebraischen Strukturen zu erhalten, vertiefen wir den Zusammenhang zwischen linearen Abbildungen und Matrizen.

Sind $A, B \in K^{m\times n}$, so entspricht die Summe $A + B$ (durch komponentenweise Addition) genau der Summe der linearen Abbildungen $A, B : K^n \to K^m$. Ebenso entspricht für $\lambda \in K$ das Produkt $\lambda \cdot A \in K^{m\times n}$ (berechnet durch komponentenweise Skalarenmultiplikation) der linearen Abbildung $\lambda \cdot A : K^n \to K^m$. Somit haben wir:

Proposition 4.5. *Die K-Vektorräume $K^{m\times n}$ und $\mathrm{Hom}(K^n, K^m)$ sind kanonisch isomorph.*

Dabei bedeutet das Adjektiv *kanonisch*, dass es einen bevorzugten Isomorphismus gibt (nämlich denjenigen, der einer Matrix einfach die entsprechende lineare Abbildung zuordnet).

Korollar 4.6. $\dim_K \mathrm{Hom}(K^n, K^m) = m \cdot n$.

Beweis. Eine Basis von $K^{m\times n}$ ist durch Matrizen gegeben, die genau an einer Stelle eine 1 und sonst überall 0 haben. Daher gilt $\dim_K K^{m\times n} = m \cdot n$. □

Wir müssen nun noch die Komposition linearer Abbildungen in die Matrix-Sprechweise übersetzen.

Es seien K ein Körper und lineare Abbildungen, d.h. Matrizen

$$B = (b_{jk})_{1\le j\le n, 1\le k\le r} : K^r \to K^n\,,\ A = (a_{ij})_{1\le i\le m, 1\le j\le n} : K^n \to K^m$$

gegeben. Insbesondere ist also $B \in K^{n\times r}$, $A \in K^{m\times n}$. Welche Matrix in $K^{m\times r}$ entspricht der Komposition $A\circ B : K^r \to K^m$? Offensichtlich wird diese lineare Abbildung beschrieben durch eine Matrix $C = (c_{ik})_{1\le i\le m, 1\le k\le r}$. Die Komponente c_{ik} dieser Matrix lässt sich bestimmen, indem wir den k-ten Einheitsvektor ($1 \le k \le r$) im K^r in die Abbildung $A \circ B$ einsetzen und die i-te Zeile ($1 \le i \le m$) des resultierenden Spaltenvektors im K^m ablesen. Wir erhalten

$$Ce_k = A(Be_k) = A \begin{pmatrix} b_{1k} \\ \vdots \\ \vdots \\ b_{nk} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} b_{jk} \\ \vdots \\ \sum_{j=1}^n a_{mj} b_{jk} \end{pmatrix}.$$

Die (ik)-te Komponente von $C = (c_{ik})$ ist also gegeben durch $\sum_{j=1}^n a_{ij} b_{jk}$. Dies nehmen wir zum Anlass folgender Definition

Definition. *Es seien $A = (a_{ij}) \in K^{m\times n}$ und $B = (b_{jk}) \in K^{n\times r}$. Dann ist das* Produkt $A \cdot B \in K^{r\times m}$ *die Matrix mit Komponenten c_{ik}, $1 \le i \le m$, $1 \le k \le r$, wobei*

$$c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}\,.$$

Es gilt also die Merkregel

Der (ik)-te Eintrag von $A \cdot B$ entsteht durch Multiplikation der i-ten Zeile von A mit der k-ten Spalte von B.

Man beachte, dass das Produkt $A \cdot B$ genau dann gebildet werden kann, wenn die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist.

Wir wir bereits oben gesehen haben, entspricht das Produkt $A \cdot B$ genau der Verkettung $A \circ B$ der durch A und B gegebenen linearen Abbildungen.

In $K^{n\times n}$, dem K-Vektorraum der *quadratischen* Matrizen, können also je zwei Elemente multipliziert werden. Zusammen mit der Struktur von $K^{n\times n}$ als abelsche Gruppe (durch komponentenweise Addition der Matrixkomponenten) erhalten wir so die Struktur eines Ringes mit 1 auf $K^{n\times n}$. Diesen bezeichnen wir mir $\mathrm{Mat}(n, K)$ und nennen ihn den *Ring der* $(n \times n)$*-Matrizen.* Das bezüglich der Addition neutrale Element ist die Nullmatrix $0 \in K^{n\times n}$, die nur 0 als Einträge hat, die Eins ist die n-dimensionale *Einheitsmatrix*

$$E_n := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

mit dem Eintrag 1 auf der Diagonalen und 0 sonst. Die Ringaxiome sind leicht zu zeigen, nur das Assoziativgesetz für die Matrixmultiplikation verdient etwas mehr Aufmerksamkeit: Eine direkte Rechnung mit dem Produkt von Matrizen ist mühsam. Da aber die Matrixmultiplikation genau der Komposition der entsprechenden linearen Abbildungen entspricht, folgt die Assoziativität leicht aus dem Assoziativgesetz für die Komposition von Abbildungen.

Wir bezeichnen mit

$$\mathrm{GL}(n, K) \subset \mathrm{Mat}(n, K)$$

die Teilmenge derjenigen $(n \times n)$-Matrizen, die ein multiplikatives Inverses besitzen. Diese Menge ist mit der Multiplikation von Matrizen und der Einheitsmatrix als neutralem Element eine Gruppe, die *allgemeine lineare Gruppe* über K.

Aus unseren Betrachtungen folgt:

Proposition 4.7. *Der Endomorphismenring* $(\mathrm{End}(K^n), +, 0, \circ, \mathrm{id}_{K^n})$ *und der Matrixring* $(\mathrm{Mat}(n, K), +, 0, \cdot, E_n)$ *sind kanonisch isomorph.*

Die Gruppen $(\mathrm{Aut}(K^n), \circ, \mathrm{id}_{K^n})$ *und* $(\mathrm{GL}(n, K), \cdot, E_n)$ *sind ebenfalls kanonisch isomorph.*

Sind V und W beliebige endlichdimensionale K-Vektorräume, so können wir lineare Abbildungen $V \to W$, d.h. Elemente in $\mathrm{Hom}(V, W)$ mit Hilfe von Koordinatensystemen durch Matrizen beschreiben,

Es sei dazu $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \dots, w_m)$ eine Basis von W. Wir erhalten Koordinatensysteme $\Phi_{\mathcal{B}} : K^n \to V$ und $\Phi_{\mathcal{C}} : K^m \to W$. Diese sind Isomorphismen, also existiert die Abbildung

$$\Phi_{\mathcal{C}}^{-1} \circ f \circ \Phi_{\mathcal{B}} : K^n \to K^m$$

Diese Abbildung ist linear und kann somit durch eine Matrix $A \in K^{m\times n}$ dargestellt werden.

Nach Definition sind die Spalten dieser Matrix durch die folgende Gleichung bestimmt:

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

wobei $1 \leq j \leq n$, d.h. wir erhalten die Merkregel

Die j-te Spalte der darstellenden Matrix ist der Koordinatenvektor des Bildes des j-ten Basisvektors.

Definition. *Die soeben definierte Matrix $M = M_{\mathcal{C}}^{\mathcal{B}}(f)$ heißt die* darstellende Matrix *von f bezüglich der Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W.*

Die Matrix $M_{\mathcal{C}}^{\mathcal{B}}(f)$ hat also die Eigenschaft, dass das folgende *Diagramm* von Vektorräumen und linearen Abbildungen *kommutiert*, d.h. starten wir an einem Punkt (in diesem Fall oben links) und laufen wir auf verschiedenen Wegen zu einem anderen Punkt (in diesem Fall unten rechts), so sind die entsprechenden Kompositionen der Abbildungen gleich.

$$\begin{array}{ccc} K^n & \xrightarrow{M_{\mathcal{C}}^{\mathcal{B}}(f)} & K^m \\ \cong \downarrow \Phi_{\mathcal{B}} & & \cong \downarrow \Phi_{\mathcal{C}} \\ V & \xrightarrow{f} & W \end{array}$$

Wir werden in $M_{\mathcal{C}}^{\mathcal{B}}(f)$ manchmal auf die hoch- und tiefgestellten Indizes verzichten, wenn klar ist, bezüglich welcher Basen wir arbeiten. Betrachten wir eine Matrix $A \in K^{m\times n}$ als lineare Abbildug $K^n \to K^m$, so ist A die darstellende Matrix dieser linearen Abbildung bezüglich der kanonischen Basen von K^m und K^n.

Beispiel. Es seien die Basen $\mathcal{B} = (\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$ von $\mathbb{R}^3$ und $\mathcal{C} = (\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix})$ von $\mathbb{R}^2$ gegeben. Dann ist die darstellende Matrix der linearen Abbildung

$$A := \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

bezüglich dieser Basen gegeben durch die Matrix

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Hierzu wendet man die Abbildung A einfach auf die Basisvektoren in $\mathcal{B}$ an und drückt die Ergebnisse als Linearkombinationen in der Basis $\mathcal{C}$ aus. Die

entstehenden Koordinatenvektoren (die man im allgemeinen als Lösungen linearer Gleichungssysteme erhält) sind die Spalten der darstellenden Matrix.

16.12.09

Auch in diesem allgemeineren Bild entspricht die Komposition linearer Abbildungen der Multiplikation von Matrizen. Genauer:

Proposition 4.8. *Es seien U, V und W endlichdimensional und es seien Basen $\mathcal{B}$ von U, $\mathcal{C}$ von V und $\mathcal{D}$ von W gewählt, und es seien $f : U \to V$, $g : V \to W$ linear. Es seien A und B die darstellenden Matrizen von f und g bzgl. der gewählten Basen. Dann wird die lineare Abbildung $g \circ f : U \to W$ durch die Matrix $B \cdot A$ dargestellt.*

Beweis. Da $A = \Phi_{\mathcal{C}}^{-1} \circ f \circ \Phi_{\mathcal{B}}$ und $B = \Phi_{\mathcal{D}}^{-1} \circ g \circ \Phi_{\mathcal{C}}$ haben wir

$$B \circ A = (\Phi_{\mathcal{D}}^{-1} \circ g \circ \Phi_{\mathcal{C}}) \circ (\Phi_{\mathcal{C}}^{-1} \circ f \circ \Phi_{\mathcal{B}}) = \Phi_{\mathcal{D}}^{-1} \circ (g \circ f) \circ \Phi_{\mathcal{B}}$$

wie gewünscht. □

Wir erhalten also

Proposition 4.9. *Es seien V und W endlichdimensionale K-Vektorräume mit $\dim V = n$, $\dim W = m$. Dann gilt*

- $\mathrm{Hom}(V, W) \cong K^{m \times n}$ *als K-Vektorräume,*
- $\mathrm{End}(V) \cong \mathrm{Mat}(n, K)$ *als Ringe mit 1 und*
- $\mathrm{Aut}(V) \cong \mathrm{GL}(n, K)$ *als Gruppen.*

Diese Isomorphismen sind jedoch nicht kanonisch, sondern hängen von der Wahl einer Basis von V (und ggf. von W) ab.

Da die Beschreibung linearer Abbildungen $V \to W$ durch Matrizen von der Wahl von Basen von V und W abhängt, stellt sich die Frage, ob wir für eine gegebene lineare Abbildung $V \to W$ durch Wahl besonders geschickter Basen eine besonders einfache Darstellung durch eine Matrix erhalten. Wir formulieren konkreter die beiden folgenden Fragen:

- Seien V und W endlichdimensionale K-Vektorräume und $f : V \to W$ linear. Man wähle Basen von V und W, so dass die entsprechende darstellende Matrix von f besonders einfache Gestalt hat.
- Es sei V ein endlichdimensionaler K-Vektorraum. Man wähle **eine** Basis von V, so dass die entsprechende darstellende Matrix einer gegebenen linearen Abbildung $f : V \to V$ eine besonders einfache Gestalt hat.

Da man im ersten Fall zwei Basen getrennt voneinander wählen kann, hat man hier mehr Flexibilität als im zweiten Fall. Wir werden für die erste Frage in Kürze eine befriedigende Antwort finden.

Das zweite Problem ist schwieriger, weil wir nur eine einzige Basis wählen können. Dies führt auf Themen wie Eigenwerttheorie, Spektralzerlegungen

und die Jordansche Normalform, die später in dieser Vorlesung behandelt werden.

Wir untersuchen jetzt noch die Frage, wie sich die darstellende Matrix einer linearen Abbildung $V \to W$ ändert, wenn wir die Basen von V und W ändern. Hier müssen wir verstehen, wie sich die Koordinaten eines festen Vektors eines Vektorraumes ändern, wenn wir von einer Basis zu einer anderen übergehen. Diese Frage können wir aber in der oben entwickelten Sprache sofort beantworten:

Definition. *Es sei V ein endlichdimensionaler K-Vektorraum und es seien $\mathcal{B}$ und $\mathcal{C}$ zwei Basen von V. Die darstellende Matrix der Identität* $\mathrm{id}_V : V \to V$ *bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$ heißt* Matrix der Koordinatentransformation bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$*, geschrieben $T_{\mathcal{C}}^{\mathcal{B}}$.*

Es gilt also folgendes: Ist $\dim V = n$ und ist $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ der Koordinatenvektor eines Vektors $v \in V$ bzg. der Basis $\mathcal{B}$, so hat v bezüglich der Basis $\mathcal{C}$ den Koordinatenvektor $T_{\mathcal{C}}^{\mathcal{B}} \cdot x$. Insbesondere sind die Spalten von $T_{\mathcal{C}}^{\mathcal{B}}$ genau die Koordinatenvektoren der Basisvektoren in $\mathcal{B}$ bezüglich der Basis $\mathcal{C}$.

Man beachte, dass nach Konstruktion

$$T_{\mathcal{C}}^{\mathcal{B}} = \Phi_{\mathcal{C}}^{-1} \circ \Phi_{\mathcal{B}} \in \mathrm{GL}(n, K)$$

wobei $\Phi_{\mathcal{B}} : K^n \to V$ und $\Phi_{\mathcal{C}} : K^n \to V$ wieder die entsprechenden Koordinatensysteme bezeichnen. Wir erhalten so auch die Formel

$$(T_{\mathcal{C}}^{\mathcal{B}})^{-1} = T_{\mathcal{B}}^{\mathcal{C}} .$$

Das allgemeine Resultat zum Basiswechsel lautet nun wie folgt:

Proposition 4.10. *Es seien V und W endlichdimensionale K-Vektorräume mit $\dim V = n$ und $\dim W = m$. Es seien $\mathcal{B}, \mathcal{B}'$ Basen von V und $\mathcal{C}, \mathcal{C}'$ Basen von W. Es sei $f : V \to W$ linear. Dann gilt die Gleichung*

$$M_{\mathcal{C}'}^{\mathcal{B}'}(f) = T_{\mathcal{C}'}^{\mathcal{C}} \cdot M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot T_{\mathcal{B}}^{\mathcal{B}'} .$$

Der Beweis folgt direkt aus den Definitionen der beteiligten Objekte: Ist $v \in V$ und $x \in K^n$ der Koordinatenvektor von v bzgl. $\mathcal{B}'$, so ist $T_{\mathcal{B}}^{\mathcal{B}'} \cdot x$ der Koordinatenvektor von v in der Basis $\mathcal{B}$. Wenden wir darauf $M_{\mathcal{C}}^{\mathcal{B}}(f)$ an, so erhalten wir also die Koordinaten von $f(v)$ in der Basis $\mathcal{C}$. Anwenden von $T_{\mathcal{C}'}^{\mathcal{C}}$ rechnet die Koordinaten dann noch in die Basis $\mathcal{C}'$ um. Damit ist $T_{\mathcal{C}'}^{\mathcal{C}} \cdot M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot T_{\mathcal{B}}^{\mathcal{B}'} \cdot x$ der Koordinatenvektor von $f(v)$ in der Basis $\mathcal{C}'$. Das ist aber genau gleich $M_{\mathcal{C}'}^{\mathcal{B}'}(f) \cdot x$.

Definition. *Es sei $f : V \to W$ eine lineare Abbildung. Der* Rang *von f, geschrieben* rang(f)*, ist die Dimension des Bildes von f, also des K-Vektorraumes* $\mathrm{im}\, f = f(V) \subset W$.

Sind V und W endlichdimensional und stellen wir die lineare Abbildung f durch eine Matrix $A \in K^{m\times n}$ dar, so ist der Rang von f gleich dem *Spaltenrang* $\mathrm{SR}(A)$ von A, d.h. gleich der Dimension des von den Spaltenvektoren aufgespannten Untervektorraumes von K^n.

Satz 4.11 (Dimensionsformel für Kern und Bild)**.** *Es seien V und W endlichdimensionale K-Vektorräume und $f : V \to W$ eine lineare Abbildung. Dann gilt*

$$\dim \ker(f) + \mathrm{rang(f)} = \dim \mathrm{V}\,.$$

Beweis. Es sei $(v_1, \ldots, v_k)$ eine Basis von $\ker f$. Insbesondere ist also $k = \dim \ker f$. Wir ergänzen diese Familie zu einer Basis $(v_1, \ldots, v_k, v_{k+1}, \ldots v_n)$ von V und zeigen: Die Familie $(f(v_{k+1}), \ldots, f(v_n))$ ist eine Basis von $\mathrm{im}\, f$. Daraus folgt dann die Behauptung des Satzes.

Es sei $w \in \mathrm{im}\, f$, d.h. $w = f(v)$ mit einem $v \in V$. Wir schreiben $v = \sum_{i=1}^{n} \lambda_i v_i$ und erhalten wegen $v_1, \ldots, v_k \in \ker f$

$$w = \sum_{i=k+1}^{n} \lambda_i f(v_i)\,.$$

Somit ist die gegebene Familie ein Erzeugendensystem von $\mathrm{im}\, f$.

Es sei nun $\sum_{i=k+1}^{n} \lambda_i f(v_i) = 0$. Damit ist $\sum_{i=k+1}^{n} \lambda_i v_i \in \ker f$. Da aber bereits $(v_1, \ldots, v_k)$ eine Basis von $\ker f$ ist, folgt aus der eindeutigen Darstellbarkeit von Vektoren in V als Linearkombination in $(v_1, \ldots, v_n)$, dass $\lambda_{k+1}, \ldots, \lambda_n = 0$. Somit ist die gegebene Familie auch linear unabhängig. □

21.12.09

Indem wir im letzten Beweis $V' := \mathrm{span}(v_{k+1}, \ldots, v_n)$ setzen, haben wir eine direkte Summenzerlegung

$$V = \ker f \oplus V'$$

und $f|_{V'} : V' \to \mathrm{im}\, f$ ist ein Isomorphismus. Da wir in der Regel $(v_1, \ldots, v_k)$ auf viele Arten durch Vektoren $(v_{k+1}, \ldots, v_n)$ zu einer Basis von V ergänzen können, ist allerdings der Unterraum V' komplementär zu $\ker f$ nicht eindeutig bestimmt.

Das folgende Korollar ist ein wichtiger Spezialfall der Dimensionsformel:

Korollar 4.12. *Es sei V ein endlichdimensionaler K-Vektorraum und $f : V \to V$ linear. Dann sind äquivalent:*

- $f \in \mathrm{Aut}(V)$.
- $\ker f = 0$.
- $\mathrm{im}\, f = V$.

Ist $A \in K^{m\times n}$, so ist der Rang von A offensichtlich gleich dem *Spaltenrang*, geschrieben $\mathrm{SR}(A)$, definiert als die Dimension des von den Spaltenvektoren aufgespannten Untervektorraumes des K^m. (Man erinnere sich, dass die Spaltenvektoren von A die Bilder der kanonischen Basisvektoren des K^n sind).

Wir haben bereits früher den Begriff des Zeilenranges einer Matrix $A \in K^{m\times n}$ kennengelernt, definiert als die Dimension des von den Zeilenvektoren aufgespannten Untervektorraumes des K^n.

Wir können diese beiden Begriffe mit Hilfe des Transponierten einer Matrix leicht in Beziehung setzen:

Definition. *Ist $A = (a_{ij}) \in K^{m\times n}$, so bezeichnen wir mit $A^T := (a_{ji}) \in K^{n\times m}$ die* Transponierte *der Matrix A. Diese entsteht aus A einfach durch Vertauschung der Rollen von Zeilen und Spalten.*

Offensichtlich ist

$$\mathrm{ZR}(A) = \mathrm{SR}(A^T)\,,\ \mathrm{SR}(A) = \mathrm{ZR}(A^T)\,.$$

Aus dem Rangsatz folgern wir nun:

Korollar 4.13. *Für alle $A \in K^{m\times n}$ gilt* $\mathrm{ZR}(A) = \mathrm{SR}(A)$.

Beweis. Zunächst halten wir fest, dass $\mathrm{ZR}(A) = n - \dim\ker A$. Denn wenden wir elementare Zeilenumformungen auf A an, so ändert sich der Zeilenrang offensichtlich nicht. Haben wir aber A auf diese Weise auf Zeilenstufenform gebracht, so ist der Zeilenrang (d.h. die Anzahl der von 0 verschiedenen Zeilen der Matrix in Zeilenstufenform) gleich $n-$Dimension des Lösungsraumes des durch A gegebenen homogenen Gleichungssystems, also gleich $n - \dim\ker A$.

Da $\mathrm{SR}(A) = \dim\mathrm{im}\, A$, folgt nun die Behauptung des Korollars aus der Dimensionsformel. □

Die wahre Bedeutung der Aussage $\mathrm{ZR}(A) = \mathrm{SR}(A)$ und der transponierten Matrix kann man erst im Rahmen der Dualitätstheorie von Vektorräumen verstehen, auf die wir im nächsten Kapitel eingehen.

Wir erwähnen noch das folgende Kriterium für die Lösbarkeit linearer Gleichungssysteme:

Proposition 4.14. *Es sei $A \in K^{m\times n}$ und $b \in K^m$. Dann sind äquivalent:*

- *Das durch $(A|b)$ gegebene lineare Gleichungssystem ist lösbar.*
- $\mathrm{rang(A)} = \mathrm{rang(A|b)}$.

Beweis. Man kann diese Aussage über den Spaltenrang oder über den Zeilenrang beweisen.

Im ersten Fall bedeutet die Aussage $\mathrm{rang(A)} = \mathrm{rang(A|b)}$, dass $\dim\mathrm{im}\, A = \dim\mathrm{im}\,(A|b)$ und dies ist wegen $\mathrm{im}\,(A|b) = \mathrm{im}\, A + \mathrm{span}(b)$ äquivalent zur Aussage $b \in \mathrm{im}\, A$, also zur Lösbarkeit des durch $(A|b)$ gegebenen Systems.

Betrachten wir den Zeilenrang, so bringen wir $(A|b)$ durch elementare Zeilenumformungen auf Zeilenstufenform mit $r = ZR(A)$ von 0 verschiedenen Zeilen. Dies ändert $ZR(A)$ und $ZR(A|b)$ nicht. Die Aussage $ZR(A) = ZR(A|b)$ bedeutet dann genau, dass in der b-Spalte keine von 0 verschiedenen Einträge unterhalb der r-ten Zeile auftreten und dies ist äquivalent zur Lösbarkeit des durch $(A|b)$ gegebenen Systems. □

Definition. *Eine Matrix $A \in K^{m\times n}$ ist in* Spaltenstufenform*, wenn die transponierte Matrix $A^T \in K^{n\times m}$ in Zeilenstufenform ist.*

Durch *elementare Spaltenumformungen* (d.h. Vertauschung zweier Spalten und Addition des λ-fachen ($\lambda \in K$) einer Spalte zu einer anderen Spalte) können wir jede Matrix $A \in K^{m\times n}$ auf Spaltenstufenform bringen. Daraus können wir sowohl eine Basis des von den Spaltenvektoren aufgespannten Untervektorraums, also von $\operatorname{im} A$, ablesen - dies ist genau die Familie der von 0 verschiedenen Spaltenvektoren der erhaltenen Matrix in Spaltenstufenform - als auch eine Familie finden, die diese Basis zu einer Basis von K^m ergänzt (Spaltenvektoren mit genau einer 1 in „Nicht-Pivot-Zeilen“ und Nullen sonst).

Hier benutzen wir die elementare Tatsache, dass elementare Spaltenumformungen einer Matrix $A \in K^{m\times n}$ das Bild $\operatorname{im} A \subset K^m$ nicht ändern.

Wir beschäftigen uns nun noch mit dem folgenden Problem: Es sei $A \in \operatorname{Mat}(n, k)$. Wie sehen wir, ob A invertierbar ist, d.h. $A \in \operatorname{GL}(n, K)$? Wie können wir das Inverse von A effektiv berechnen?

Man beachte, dass diese Frage von Bedeutung ist, wenn wir lineare Gleichungssysteme mit ebenso vielen Gleichungen wir Unbekannten lösen wollen (die also durch eine erweiterte Matrix $(A|b) \in K^{n\times(n+1)}$ dargestellt werden): Nach Korollar 4.12 besitzt ein Gleichungssystem dieser Art ja genau dann eine eindeutige Lösung, wenn $A \in \operatorname{GL}(n, K)$. Genauer gilt folgendes: Ist $A \in \operatorname{GL}(n, K)$ und $b \in K^n$, so erhält man die eindeutige Lösung der Gleichung

$$Ax = b$$

durch Multiplikation mit A^{-1} von links:

$$x = A^{-1}b\,.$$

Ist $A \in K^{n\times n}$ gegeben, so folgt aus der Lösungstheorie linearer Gleichungssysteme, dass A genau dann invertierbar ist, wenn A durch elementare Zeilenumformungen auf eine Matrix in Zeilenstufenform mit genau n von 0 verschiedenen Zeilen gebracht werden kann. Diese Matrix ist dann also von der Gestalt

$$\begin{pmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & * \end{pmatrix}$$

wobei die Einträge auf der Diagonalen alle von 0 verschieden sind. Wir führen nun zu den beiden schon besprochenen elementaren Zeilenumformungen noch eine weitere elementare Zeilenumformung ein, und zwar **Multiplikation einer Zeile mit einem $\lambda \in K$ mit $\lambda \neq 0$**. Dann erhalten wir das folgende Ergebnis:

Proposition 4.15. *Eine quadratische Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn sie mit Hilfe elementarer Zeilenumformungen in die Einheitsmatrix E_n überführt werden kann.*

Beweis. Wir haben bereits gesehen, dass $A \in \mathrm{GL}(n, K)$ genau dann, wenn A durch elementare Zeilenumformungen in eine Matrix der vor dieser Proposition beschriebenen Art überführt werden kann.

Wir nehmen nun an, es sei eine Matrix in dieser Form gegeben. Wir müssen noch zeigen, dass sie durch elementare Zeilenumformungen in die Einheitsmatrix E_n verwandelt werden kann. Durch Multiplikation der Zeilen mit geeigneten $0 \neq \lambda \in K$ können wir erreichen, dass auf der Diagonale überall 1 steht. Anschließend führt Addition von geeigneten Vielfachen einer Zeile zu den darüber liegenden Zeilen dazu, dass alle Einträge, die nicht auf der Diagonalen liegen, gleich 0 sind. Das ist aber genau die Einheitsmatrix. □

Wir erhalten nun das folgende einfache Verfahren zur Bestimmung von A^{-1}.

Proposition 4.16. *Es sei $A \in \mathrm{GL}(n, K)$. Wir erhalten A^{-1} dadurch, dass wir die elementaren Zeilenumformungen, die A in E_n überführen, auf die Einheitsmatrix E_n anwenden.*

Beweis. Elementare Zeilenumformungen können dadurch beschrieben werden, dass man die gegebene Matrix **von links** mit sogenannten *Elementarmatrizen* multipliziert. Genauer entspricht

- eine Vertauschung zweier Zeilen einer Multiplikation mit der Elementarmatrix P_i^j,
- der Addition des λ-fachen ($\lambda \in K$) der Zeile j zur Zeile i einer Multiplikation mit der Elementarmatrix $Q_i^j(\lambda)$,
- der Multiplikation der i-ten Zeile mit $\lambda \neq 0$ einer Multiplikation mit der Elementarmatrix $S_i(\lambda)$.

Diese Elementarmatrizen sind in [Fischer], S. 163 f. angegeben.

Da wir A durch elementare Zeilenumformungen in die Matrix E_n überführen können, erhalten wir eine Gleichung der Form

$$E_n = B_s \cdot \ldots \cdot B_1 \cdot A$$

mit Elementarmatrizen $B_1, \ldots, B_s$. Dabei enspricht die Matrix B_k, $1 \leq k \leq s$, genau der k-ten elementaren Zeilenumformung. Multiplizieren wir diese Gleichung von rechts mit A^{-1}, so erhalten wir

$$A^{-1} = B_s \cdot \ldots \cdot B_1 = B_s \cdot \ldots \cdot B_1 \cdot E_n \,.$$

Indem wir die Multiplikation mit Elementarmatrizen als elementare Zeilenumformungen interpretieren, folgt die Behauptung. □

Wir illustrieren dieses Verfahren an dem in [Fischer], S. 168, angegebenen Beispiel.

Wir weisen darauf hin, dass die Elementarmatrizen invertierbar sind. Die Inversen kann man leicht berechnen (z.B. mit dem gerade beschriebenen Verfahren).

23.12.09

Das folgende Ergebnis zeigt nun, dass wir lineare Abbildungen durch besonders einfache Matrizen darstellen können, wenn wir die Basen von Quelle und Ziel unabhängig voneinander wählen können.

Satz 4.17 (Normalform linearer Abbildungen)**.** *Es seien V und W endlichdimensionale K-Vektorräume, $n = \dim V$, $m = \dim W$, und es sei $f : V \to W$ linear. Dann existieren Basen von V und von W, so dass bezüglich dieser Basen die Abbildung f durch eine Matrix der Form*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in K^{m\times n}$$

dargestellt wird, wobei $r = \mathrm{rang(f)}$.

Beweis. Es sei $(v_{r+1}, \ldots, v_n)$ eine Basis von $\ker f$ (man erinnere sich, dass nach der Dimensionsformel $\dim \ker f + r = n$). Wir ergänzen diese zu einer Basis $\mathcal{B} = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ von V. Dann ist $(f(v_1), \ldots, f(v_r))$ linear unabhängig in W. Sei zum Beweis dieser Aussage $\sum_{i=1}^r \lambda_i f(v_i) = 0$, d.h. $\sum_{i=1}^r \lambda_i v_i \in \ker f$. Da $\ker f = \mathrm{span}(v_{r+1}, \ldots, v_n)$, folgt aus der eindeutigen Darstellbarkeit von Vektoren in V als Linearkombination in $(v_1, \ldots, v_n)$, dass $\lambda_1 = \ldots = \lambda_r = 0$ wie behauptet.

Ergänzen wir nun $(f(v_1), \ldots, f(v_r))$ zu einer Basis $\mathcal{C} := (f(v_1), \ldots, f(v_r), w_{r+1}, \ldots, w_m)$ von W, so hat die darstellende Matrix von f bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$ die behauptete Form. □

11.1.10

5. Dualräume

Durch das Gaußsche Eliminationsverfahren sind wir in der Lage, zu einem linearen Gleichungssystem $Ax = b$, $A \in K^{m\times n}$, $b \in K^m$, die Lösungsmenge $L \subset K^n$ in *Parameterform* zu bestimmen, d.h. in der Form

$$L = \{l + \lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} \mid \lambda_1, \ldots, \lambda_{n-r} \in K\},$$

wobei $l \in K^n$ eine spezielle Lösung ist, $r = \mathrm{rang(A)}$, und $(v_1, \ldots, v_{n-r})$ eine Basis von $\ker A \subset K^n$ ist. Haben wir ein weiteres Gleichungssystem $A'x = b'$ mit $A' \in K^{m'\times n}$, $b' \in K^{m'}$ und Lösungsmenge $L' \subset K^n$ gegeben (d.h. die Anzahl der Unbestimmten bleibt gleich), so können wir das Gleichungssystem

$$\begin{pmatrix} A \\ B \end{pmatrix} x = \begin{pmatrix} b \\ b' \end{pmatrix}$$

mit n Unbestimmten und $m + m'$ Gleichungen betrachten. Die Lösungsmenge dieses Gleichungssystems ist offensichtlich gegeben durch den affinen Unterraum $L \cap L' \subset K^n$. Wir können also den Schnitt von affinen Teilräumen L und L' im K^n in Parameterform leicht berechnen, wenn L und L' als Lösungsmengen von explizit gegebenen Gleichungssystemen vorliegen.

Es stellt sich die Frage, wie wir $L \cap L'$ in Parameterform berechnen können, wenn L und L' selbst in Parameterform, d.h. in der Form $L = l + V$, $L' =$

$l'+V'$ mit Untervektorräumen $V, V' \subset K^n$ gegeben sind. Diese geometrische Grundaufgabe tritt oft auf.

Beispiel. Man berechne den Schnitt der Ebenen

$$E = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mid \lambda_1, \lambda_2 \in \mathbb{R} \right\} \subset \mathbb{R}^3$$

und

$$F = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mid \mu_1, \mu_2 \in \mathbb{R} \right\} \subset \mathbb{R}^3$$

in Parameterform.

Eine Möglichkeit besteht darin, zunächst L und L' als Lösungsmengen von Gleichungssystemen $Ax = b$ und $A'x = b'$ zu schreiben und anschließend so vorzugehen wie oben beschrieben.

Die Aufgabe, zu einem affinen Unterraum $L \subset \mathbb{R}^n$ eine Matrix $A \in K^{m \times n}$ zu finden sowie ein $b \in K^m$, so dass $L = \{x \in \mathbb{R}^n \mid Ax = b\}$ kann man z.B. mit der Theorie dualer Vektorräume lösen.

Definition. *Es sei V ein K-Vektorraum. Der Vektorraum* $\mathrm{Hom}_K(V, K)$ *der K-linearen Abbildungen $V \to K$ in den eindimensionalen K-Vektorraum K heißt der zu V* duale *Vektorraum und wird mit V^* bezeichnet. Die Elemente von V^* heißen* Linearformen *auf V.*

Beispiel.

- Wir betrachten den $\mathbb{R}$-Vektorraum $C([0,1])$ der stetigen Funktionen $[0,1] \to \mathbb{R}$. Dann definiert die Zuordnung
$$f \mapsto \int_0^1 f(x)dx$$
eine Linearform auf V, also ein Element in V^*.
- Ist X eine Menge und $V = \mathbb{R}^X$ der $\mathbb{R}$-Vektorraum der Abbildungen $X \to \mathbb{R}$, so definiert für jedes $x \in X$ die Abbildung
$$\phi_x : \mathbb{R}^X \to \mathbb{R}\,,\ f \mapsto f(x)$$
eine Linearform auf V.
- Ist $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n$, so definiert die Abbildung
$$K^n \to K\,,\ (x_1, \ldots, x_n) \mapsto a_1x_1 + \ldots + a_nx_n$$
ein Element in $(K^n)^*$. Der Kern dieser Linearform ist genau die Lösungsmenge der Gleichung
$$a_1x_1 + \ldots + a_nx_n = 0\,.$$

Wir bezeichnen diese Linearform im folgenden auch mit dem Zeilenvektor $(a_1, \ldots, a_n)$. Schreiben wir Vektoren in K^n als Spaltenvektoren, dann ist obige Linearform einfach durch die Vorschrift $x \mapsto (a_1, \ldots, a_n) \cdot x$ gegeben.

Wir werden im folgenden die Dualitätstheorie nur für endlichdimensionale Vektorräume entwickeln und dabei insbesondere das letzte Beispiel auf eine systematische Grundlage stellen.

In der Analysis betrachtet man dagegen oft unendlichdimensionale Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$ (siehe die ersten beiden Beispiele), versieht aber dann die auftretenden Vektorräume zusätzlich mit der Struktur einer vollständigen Norm, so dass man es mit Banachräumen zu tun hat. In diesem Kontext definiert man als das Duale eines Banachraumes den Vektorraum der stetigen linearen Funktionale in den Grundkörper.

Es sei V ein endlichdimensionaler K-Vektorraum. Haben wir eine Basis $\mathcal{B} := (v_1, \ldots, v_n)$ von V gegeben, so können wir die Elemente von V^* mit den $(1 \times n)$-Matrizen (d.h. Zeilenvektoren in K^n) identifizieren, indem wir für $f : V \to K$ die darstellende Matrix $M_{\mathcal{C}}^{\mathcal{B}}(f)$ betrachten. Hier bezeichnet $\mathcal{C} = (1)$ die kanonische Basis von $K = K^1$. Diese Überlegung zeigt, dass $\dim V = \dim V^*$ und dass somit $V \cong V^*$ als K-Vektorräume. Wir erhalten wie folgt einen expliziten Isomorphismus.

Proposition 5.1. *Für $j = 1, \ldots, n$ sei $v_j^* : V \to K$ gegeben durch $v_j^*(v_i) := \delta_{ij}$. Dann ist $(v_1^*, \ldots, v_n^*)$ eine Basis von V^* und die Abbildung $v_i \mapsto v_i^*$ induziert einen Isomorphismus $\Psi_{\mathcal{B}} : V \to V^*$.*

Der Beweis findet sich in [Fischer], S. 332. Die so konstruierte Basis $(v_1^*, \ldots, v_n^*)$ von V^* heißt die zu $\mathcal{B}$ *duale* Basis von V^* und wird mit $\mathcal{B}^*$ bezeichnet.

Wir betonen, dass der so erhaltene Isomorphismus $\Psi_{\mathcal{B}} : V \to V^*$ von der Basis $\mathcal{B}$ abhängt. Man vergleiche das Beispiel in [Fischer], S. 332 f. Insbesondere ist die Schreibweise v_i^* gefährlich, weil diese Linearform nicht nur von dem einzelnen Vektor v_i, sondern auch von den anderen Basisvektoren der Basis $\mathcal{B}$ abhängt.

Korollar 5.2. *Ist V endlichdimensional und $0 \neq v \in V$, so existiert ein $\phi \in V^*$ mit $\phi(v) \neq 0$.*

Beweis. Es sei $n = \dim V$. Da nach Voraussetzung die Familie (v) linear unabhängig ist, können wir diese zu einer Basis $\mathcal{B} = (v, v_2, \ldots, v_n)$ ergänzen. Die bezüglich dieser Basis gebildete Linearform $v^* = \Psi_{\mathcal{B}}(v)$ leistet das Gewünschte. □

Wir wollen nun nicht nur Vektorräume, sondern auch die linearen Abbildungen zwischen ihnen dualisieren.

Definition. *Es sei $f : V \to W$ linear. Wir bezeichnen die durch die Vorschrift*

$$\phi \mapsto \phi \circ f$$

gegebene lineare Abbildung $W^* \to V^*$ mit f^* und nennen sie die zu f duale *Abbildung.*

Man beachte, dass nach Dualisieren die Rollen von V und W bei Quelle und Ziel vertauscht werden.

Zentral ist das folgende Resultat.

Proposition 5.3. *Es seien V und W endlichdimensional und $\mathcal{B}$ eine Basis von V und $\mathcal{C}$ eine Basis von W. Es sei $f : V \to W$ linear. Dann gilt*

$$M_{\mathcal{B}^*}^{\mathcal{C}^*}(f^*) = M_{\mathcal{C}}^{\mathcal{B}}(f)^T .$$

Oder kurz: Die darstellende Matrix der dualen Abbildung ist die Transponierte der darstellenden Matrix.

Der Beweis findet sich in [Fischer], S. 334.

13.1.10

Ist V ein Vektorraum, so können wir sein Dual V^* wieder dualisieren und erhalten das *Doppeldual* $V^{**} := (V^*)^*$. Man hat das Gefühl, dass wir dieses Spiel beliebig weit fortsetzen können. Dem ist aber nicht so, denn V^{**} ist im Wesentlichen wieder der Vektorraum V, falls V endlichdimensional ist. Dies wird im folgenden präzisiert.

Ist $v \in V$, so definiert die Auswertung bei v

$$\iota_v : V^* \to K\,,\ \phi \mapsto \phi(v)$$

ein Element in V^{**}.

Proposition 5.4. *Es sei V endlichdimensional. Dann ist die Abbildung $\iota : V \to V^{**}$, $v \mapsto \iota_v$ einen Vektorraumisomorphismus.*

Beweis. Linearität sieht man leicht. Wir haben

$$\ker \iota = \{v \in V \mid \forall \phi \in V^* : \phi(v) = 0\} = 0$$

nach Korollar 5.2. Damit ist ι injektiv, also ein Isomorphismus, da $\dim V = \dim V^{**}$. □

Der Isomorphismus $\iota : V \to V^{**}$ hängt nun **nicht** von der Wahl einer Basis von V ab. Man sagt auch, V und V^{**} sind *kanonisch* isomorph.

Identifizieren wir vermöge dieses Isomorphismus $v \in V$ mit $\iota_v \in V^{**}$, so erhalten wir die suggestive Gleichung

$$v(\phi) = \phi(v)\,.$$

Diese Beobachtung zeigt, dass wir nach nochmaliger Dualisierung V^{***} nichts Neues erhalten: Der kanonische Isomorphismus $\iota : V \to V^{**}$ induziert einen kanonischen Isomorphismus $\iota^* : V^{***} \to V^*$, d.h. wir können statt V^{***} wieder einfach das Dual V^* betrachten.

Proposition 5.5. *Es sei V endlichdimensional mit Basis $\mathcal{B}$. Wir erhalten Isomorphismen $\Psi_{\mathcal{B}} : V \to V^*$ mittels der Basis $\mathcal{B}$ und $\Psi_{\mathcal{B}^*} : V^* \to V^{**}$ mittels der Basis $\mathcal{B}^*$. Die Komposition dieser Isomorphismen*

$$V \to V^* \to V^{**}$$

stimmt mit dem eben definierten kanonischen Isomorphismus $V \to V^{**}$ überein.

Beweis. Die Basis $(v_1^{**}, \ldots, v_m^{**})$ von V^{**} hat nach Konstruktion die Eigenschaft

$$v_j^{**}(v_i^*) = \delta_{ij} = v_i^*(v_j) = \iota_{v_j}(v_i^*)$$

für alle $1 \leq i, j \leq n$. Daher gilt $v_j^{**} = \iota_{v_j}$ für alle $1 \leq j \leq n$. □

Korollar 5.6. *Es seien V und W endlichdimensional und $f : V \to W$ linear. Es seien $\mathcal{B}$ und $\mathcal{C}$ Basen von V und W und es sei M die darstellende Matrix von f bezüglich dieser Basen. Fassen wir diese Basen mit Hilfe der eben besprochenen Identifikation als Basen von V^{**} und W^{**} auf, so ist die Abbildung $f^{**} : V^{**} \to W^{**}$ bezüglich dieser Basen wieder durch die Matrix M gegeben. Insbesondere gilt $f^{**} = f$.*

Beweis. Die darstellende Matrix von f^{**} ist bezüglich der Basen $\mathcal{B}^{**}$ und $\mathcal{C}^{**}$ durch die Matrix $(M^T)^T = M$ gegeben. □

Definition. *Es sei V ein K-Vektorraum und $W \subset V$ ein Untervektorraum. Die Teilmenge*

$$W^0 := \{\phi \in V^* \mid \forall w \in W : \phi(w) = 0\}$$

heißt der Annulator *von W. Dies ist offensichtlich ein Untervektorraum von V^*.*

Proposition 5.7. *Es gilt* $\dim W^0 = \dim V - \dim W$.

Der Beweis findet sich in [Fischer], S. 333 unten. Insbesondere gilt: Je „größer“ W, desto „kleiner“ W^0.

Korollar 5.8. *Unter der kanonischen Identifikation $V^{**} = V$ gilt $(W^0)^0 = W$.*

Beweis. Nach der letzten Proposition gilt $\dim(W^0)^0 = \dim V^{**} - \dim W^0 = \dim V - (\dim V - \dim W) = \dim W$. Also genügt es, die Inklusion $W \subset (W^0)^0$ zu zeigen. Dies ist aber klar: Es sei $w \in W$ und $\phi \in W^0$. Dann gilt $\iota_w(\phi) = \phi(w) = 0$. □

Auf dem Niveau von linearen Abbildungen wird die Idee der Dualität präzisiert durch

Proposition 5.9. *Es seien V und W endlichdimensionale K-Vektorräume und $f : V \to W$ linear. Dann gilt* $\operatorname{im} f^* = (\ker f)^0$ *und* $\ker f^* = (\operatorname{im} f)^0$.

Beweis. Ist $\psi \in \operatorname{im} f^* \subset V^*$, so existiert ein $\phi \in W^*$ mit $\phi \circ f = \psi$. Ist nun $v \in \ker f$, so haben wir $\psi(v) = \phi \circ f(v) = 0$. Somit ist $\psi \in (\ker f)^0$.

Es sei nun umgekehrt $\psi \in (\ker f)^0$, also $\psi|_{\ker f} = 0$. Es sei $r = \operatorname{rang(f)}$ und $(w_1, \ldots, w_r)$ eine Basis von $\operatorname{im} f$. Wir ergänzen diese zu einer Basis $(w_1, \ldots, w_n)$ von W. Wir wählen Vektoren $v_1, \ldots, v_r \in V$ mit $f(v_i) = w_i$ und definieren $\phi : W \to K$ durch $\phi(w_i) := \psi(v_i)$, falls $1 \leq i \leq r$ und $\phi(w_i) := 0$ sonst. Diese Setzung ist wohldefiniert, denn gilt auch $f(v_i') = w_i$ so ist $v_i' - v_i \in \ker f$, also $\psi(v_i) = \psi(v_i')$. Nach Konstruktion ist $\psi = \phi \circ f$, somit $\psi \in \operatorname{im} f^*$.

Für die zweite Gleichung wenden wir die erste Gleichung auf die Abbildung $f^* : W^* \to V^*$ an und erhalten

$$\operatorname{im} f = \operatorname{im} f^{**} = ((\ker f^*)^0$$

somit

$$(\operatorname{im} f)^0 = ((\ker f^*)^0)^0 = \ker f^*$$

wie behauptet. □

Diese abstrakten Tatsachen können wir nun zur Lösung unserer ursprünglichen Aufgabe verwenden.

Es sei zunächst $V \subset K^n$ ein Untervektorraum mit Basis $(v_1, \ldots, v_k)$ gegeben. Wir wollen eine Matrix $A \in K^{m \times n}$ so bestimmen, dass $\ker A = V$. Dazu schreiben wir die Vektoren $v_1, \ldots, v_k$ als Spalten einer Matrix $B \in K^{n \times k}$. Dann gilt $\operatorname{im} B = V$ und somit nach der letzten Proposition $V^0 = \ker B^T$. Durch Anwenden des Gaußschen Eliminationsverfahrens auf B^T bestimmen wir nun eine Basis $(w_1, \ldots, w_m)$ von $\ker B^T = V^0$, wobei $w_1, \ldots, w_m \in K^n$. Fassen wir diese als Zeilen zu einer Matrix $A \in K^{m \times n}$ zusammen, so gilt $(\ker A)^0 = \operatorname{im} A^T = \ker B^T = (\operatorname{im} B)^0$ somit, nach Übergang zum Annulator auf beiden Seiten, $\ker A = \operatorname{im} B = V$ wie gewünscht.

Ist allgemeiner $L = a + V \subset K^n$ ein affiner Teilraum, so bestimmt man zunächst mittels des obigen Verfahrens eine Matrix $A \in K^{m \times n}$ mit $\ker A = V$. Dann ist die Lösungsmenge des inhomogenen Gleichungssystems

$$Ax = b\,,$$

wobei $b := Aa$, genau gleich L.

Beispiel. Wir behandeln das Beispiel vom Anfang des Kapitels. Wir berechnen dazu Basen der Kerne von

$$\begin{pmatrix} -1 & 3 & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

und von

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

mit Hilfe des Eliminationsverfahrens. Diese Basen sind gegeben durch

$$\left(\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}\right) \text{ und } \left(\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}\right)$$

Damit ist nach der obigen Bemerkung E die Lösungsmenge der Gleichung

$$\begin{pmatrix} 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4$$

und F die Lösungsmenge der Gleichung

$$\begin{pmatrix} -1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = -1$$

schließlich also $E \cap F$ die Lösungsmenge des Gleichungssysstems

$$\begin{pmatrix} 3 & 1 & 0 \\ -1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

und damit gleich der Geraden

$$\left\{ \begin{pmatrix} 7/5 \\ -1/5 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \subset \mathbb{R}^3$$

18.1.10

6. Determinanten

Ist eine quadratische Matrix $A \in K^{n \times n}$ gegeben, d.h. ein Endomorphismus von K^n, so suchen wir nach einem einfachen Kriterium, ob A invertierbar ist oder nicht. Wir haben bereits die Möglichkeit diskutiert, A auf Zeilenstufenform zu bringen und nachzusehen, ob es dort n von 0 verschiedene Zeilen gibt (genau dann ist A invertierbar). Eine andere Möglichkeit besteht darin, die *Determinante* von A zu berechnen. Diese liefert ein Element in K und dieses ist von 0 verschieden genau dann, wenn A invertierbar ist. In diesem Abschnitt werden wir Determinanten einführen. Sie bilden einen zentralen Gegenstand der klassischen linearen Algebra.

Die grundlegende Idee ist: $A \in \mathbb{R}^{n \times n}$ ist genau dann invertierbar, wenn das von den Spalten- (oder Zeilenvektoren) aufgespannte Parallelepiped in $\mathbb{R}^n$ positives Volumen hat. Wir werden auf diesen Gesichtspunkt später zurückkommen.

Als Vorbereitung benötigen wir den Begriff des Signums einer Permutation.

Wir definieren das *Signum* einer Permutation $\sigma \in S_n$ durch die Formel

$$\operatorname{sgn}(\sigma) := \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i} \in \mathbb{Q} \setminus \{0\}\,.$$

Wir werden beweisen, dass $\operatorname{sgn}(\sigma)$ nur die Werte ± 1 annehmen kann. Wichtig ist dabei die folgenden Beobachtung.

Proposition 6.1. *Die Abbildung* sgn *definiert einen Gruppenhomomorphismus*

$$(S_n, \circ, \mathrm{id}) \to (\mathbb{Q} \setminus \{0\}, \cdot, 1)\,.$$

Beweis. Offensichtlich ist $\operatorname{sgn}(\mathrm{id}) = 1$. Es bleibt noch zu zeigen, dass

$$\operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau) \cdot \operatorname{sgn}(\sigma)$$

für alle $\sigma, \tau \in S_n$ ist. Dies folgt mit einer expliziten Rechnung:

$$\operatorname{sgn}(\tau \circ \sigma) = \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j - i} = \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}\,.$$

Die zweite Gleichheit erhält man durch Erweitern. Da sich die auftretenden Brüche nicht ändern, wenn man i und j vertauscht, gilt

$$\prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} = \prod_{i,j \in \{1,\ldots,n\},\ \sigma(i) < \sigma(j)} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)}$$

und da σ eine Bijektion ist, stimmt dieser Ausdruck mit

$$\prod_{\alpha,\beta \in \{1,\ldots,n\}, \alpha < \beta} \frac{\tau(\beta) - \tau(\alpha)}{\beta - \alpha}\,,$$

also mit $\operatorname{sgn}(\tau)$ überein. Mit der vorhergehenden Rechnung gilt also in der Tat $\operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau) \operatorname{sgn}(\sigma)$. □

Wir nennen eine Permutation $\sigma \in S_n$ eine *Transposition*, wenn σ genau zwei Zahlen aus $\{1, 2, \ldots, n\}$ vertauscht und die anderen Zahlen festlässt.

Lemma 6.2. *Jede (von der Identität verschiedene) Permutation in S_n ist Produkt von Transpositionen.*

Beweis. Wir machen Induktion nach n. Offensichtlich ist die Aussage für $n = 1$ richtig. Es sei $\sigma \in S_n$, $\sigma \neq \mathrm{id}$, gegeben und es sei $\tau \in S_n$ die Permutation, die $\sigma(n)$ mit n vertauscht (dies ist entweder eine Transposition oder die Identität). Dann bildet

$$\tau \circ \sigma$$

das Element n auf sich ab. Die Permutationen in S_{n-1} und die Permutationen in S_n, die n auf sich abbilden, können in offensichtlicher Weise identifiziert werden. Nach Induktionsannahme gibt es also Transpositionen $\tau_1, \ldots, \tau_k \in S_{n-1} \subset S_n$ mit

$$\tau \circ \sigma = \tau_1 \circ \ldots \circ \tau_k\,.$$

Daher ist

$$\sigma = \tau^{-1} \circ \tau_1 \circ \ldots \circ \tau_k$$

selbst ein Produkt von Transpositionen. □

Wir müssen nun das Signum für konkrete Permutationen berechnen. Dies tun wir für Transpositionen.

Lemma 6.3. *Ist $\tau \in S_n$ eine Transposition, so gilt*

$$\operatorname{sgn}(\tau) = -1\,.$$

Beweis. Sei zunächst

$$\tau = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 2 & 1 & 3 & \ldots & n \end{pmatrix}.$$

In diesem Fall erhalten wir

$$\operatorname{sgn}(\tau) = \prod_{i=1,j=2} \frac{\tau(j)-\tau(i)}{j-i} \cdot \prod_{j\geq 3} \frac{\tau(j)-\tau(i)}{j-i} \cdot \prod_{j\geq 3} \frac{\tau(j)-\tau(i)}{j-i}\,.$$

Wir setzen nun die Definition von τ ein und erhalten

$$\operatorname{sgn}(\tau) = (-1) \cdot \prod_{i=1,j\geq 3} \frac{j-2}{j-1} \cdot \prod_{i=2,j\geq 3} \frac{j-1}{j-2} = -1\,.$$

Es seien nun $\tau \in S_n$ eine beliebige Transposition und $1 \leq k < l \leq n$ die beiden Zahlen, die τ vertauscht. Wir erhalten eine Permutation $\sigma \in S_n$, indem wir die injektive Abbildung $\{k,l\} \to \{1,2\}$, $k \mapsto 1$, $l \mapsto 2$, zu einer injektiven (und damit bijektiven) Abbildung $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ fortsetzen.

Damit erhalten wir

$$\tau = \sigma^{-1} \cdot \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 2 & 1 & 3 & \ldots & n \end{pmatrix} \cdot \sigma$$

also

$$\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)^{-1} \cdot (-1) \cdot \operatorname{sgn}(\sigma) = -1\,.$$

Hier haben wir die Tatsache benutzt, dass $\operatorname{sgn} : S_n \to (\{\pm 1\}, \cdot, 1)$ ein Gruppenhomomorphismus ist (siehe Proposition 6.1.) □

Kombiniert man dieses Lemma mit Proposition 6.1 und Lemma 6.2, so folgt nun in der Tat, das sgn nur die Werte ± 1 annimmt. Darüberhinaus gilt $\operatorname{sgn}(\sigma) = +1$ genau dann, falls sich σ als Produkt einer **geraden** Anzahl von Transpositionen schreiben lässt. Zum Beispiel ist das Signum der Permutation

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

gleich $+1$. Diejenigen Permutationen $\sigma \in S_n$ mit $\operatorname{sgn}(\sigma) = +1$ bezeichnet man auch als *gerade*, die anderen als *ungerade*. Als Nebenprodukt unserer Überlegungen erhalten wir also auch

Korollar 6.4. *Eine Permutation kann nicht zugleich Produkt einer geraden und einer ungeraden Anzahl von Transpositionen sein.*

20.1.10

Wir fassen die Eigenschaften des Signums in folgender Proposition zusammen:

Proposition 6.5. *Für alle* $n \in \{2, 3, \ldots\}$ *definiert die Abbildung* $\operatorname{sgn} : S_n \to (\{\pm 1\}, \cdot, 1)$ *einen surjektiven Gruppenhomomorphismus.*

Die Gruppe $(\{\pm 1\}, \cdot, 1)$ ist übrigens isomorph zur Gruppe $(\mathbb{Z}/2, +, 0)$ mittels des Isomorphismus, der durch $1 \mapsto 0$, $-1 \mapsto 1$ gegeben ist. Man fasst aber den Wert von sgn in der Regel als Element in der multiplikativen Gruppe $(\{\pm 1\}, \cdot, 1)$ auf.

Definition. *Es sei* $n \geq 1$. *Unter einer* Determinantenabbildung *verstehen wir eine Abbildung* $\det : \operatorname{Mat}(n, K) \to K$ *mit den folgenden Eigenschaften:*

i. det *ist linear in jeder Zeile.*
ii. $\det A = 0$, *wenn zwei Zeilen übereinstimmen.*
iii. $\det E_n = 1$.

Der Wert $\det A \in K$ *heißt in diesem Fall die* Determinante *von* A.

Satz 6.6. *Es gibt genau eine Determinantenabbildung* $\det : \operatorname{Mat}(n, K) \to K$.

Wir ziehen zunächst weitere Folgerungen aus der vorherigen Charakterisierung von Determinanten.

Proposition 6.7. *Es sei* $\det : \operatorname{Mat}(n, K) \to K$ *eine Determinantenabbildung. Dann gilt*

iv. *Der Wert von* det *ändert sich nicht, wenn wir das Vielfache einer Zeile zu einer anderen addieren.*
v. *Der Wert von* det *wird mit* -1 *multipliziert, wenn wir zwei Zeilen vertauschen.*

Beweis. Es entstehe A' aus A durch Addition des λ-fachen von Zeile i_1 zu Zeile i_2, wobei $1 \leq i_1, i_2 \leq n$, $i_1 \neq i_2$. Nach Punkt i. gilt dann

$$\det A' = \det A + \lambda \cdot \det \overline{A}\,,$$

wobei $\overline{A}$ aus A entsteht, indem wir Zeile i_2 durch Zeile i_1 ersetzen. Nach ii. ist aber $\det \overline{A} = 0$.

Es entstehe nun A' aus A durch Vertauschung zweier Zeilen. Wir betrachten die Matrix $\overline{A}$, wo wir sowohl Zeile i_1 als auch Zeile i_2 von A durch die Summe dieser Zeilen ersetzen. Dann ist

$$0 = \det(\overline{A}) = det \begin{pmatrix} * \\ a_{i_1 j} + a_{i_2 j} \\ * \\ a_{i_1 j} + a_{i_2 j} \\ * \end{pmatrix} = \det \begin{pmatrix} * \\ a_{i_1 j} \\ * \\ a_{i_1 j} + a_{i_2 j} \\ * \end{pmatrix} + \det \begin{pmatrix} * \\ a_{i_2 j} \\ * \\ a_{i_1 j} + a_{i_2 j} \\ * \end{pmatrix} = \det A + \det A'$$

wobei wir nur Zeilen i_1 und i_2 anschreiben und die Eigenschaften i., ii. und iv. verwenden. □

Wir können nun die Eindeutigkeit der Determinante beweisen.

Satz 6.8. *Es seien* $\det, \det' : \mathrm{Mat}(n, K) \to K$ *Determinantenfunktionen. Dann gilt* $\det = \det'$.

Beweis. Es sei $A \in \mathrm{Mat}(n, K)$. Nach den Eigenschaften iv. und v., die ja für det und det′ gelten, können wir annehmen, dass A in Zeilenstufenform vorliegt. Hier beachten wir, dass bei jeder dazu nötigen elementaren Zeilenumformung det und det′ gleich bleiben oder beide mit -1 multipliziert werden.

Angenommen, rang(A) < n. Dann ist die letzte Zeile von A gleich 0. Es entstehe A' aus A durch Multiplikation dieser Zeile mit 0. Dann ist $A = A'$, also insbesondere $\det(A') = \det(A)$. Aber nach Eigenschaft i. gilt auch

$$\det(A') = 0 \cdot \det(A) = 0\,.$$

Das gleiche Argument zeigt $\det'(A) = 0$.

Es sei nun rang(A) = n. Durch Anwendung elementarer Zeilenumformungen, die auch die Multiplikation einer Zeile mit $0 \neq \lambda \in K$ einschließen, können wir annehmen, dass A die Einheitsmatrix ist, denn bei diesen Operationen passiert mit det und det′ das Gleiche: Nichts, Multiplikation mit -1, oder Multiplikation mit $\lambda \neq 0$. Hier benutzen wir die Eigenschaften i., iv. und v., die für det und det′ gelten. Für die Einheitsmatrix ist aber $\det(E_n) = \det'(E_n) = 1$ nach Eigenschaft iii. □

Dieser Beweis zeigt auch:

Proposition 6.9. *Es sei* $\det : \mathrm{Mat}(n, K) \to K$ *eine Determinantenfunktion. Ist* $A \in \mathrm{Mat}(n, K)$*, so gilt* $\det A = 0$ *genau dann, wenn* rang(A) < n.

Die Existenz einer Determinante zeigen wir durch Angabe einer explizten Vorschrift.

Satz 6.10. *Es gibt eine Determinantenfunktion* $\det : \mathrm{Mat}(n, K) \to K$.

Gegeben sei $A \in \mathrm{Mat}(n, K)$. Wir setzen

$$\det A := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}\,.$$

Diese Vorschrift erfüllt Regeln i. und iii., wie man ziemlich leicht sieht (vgl. [Fischer], S. 194 Beweis von D1 und D3). Bevor wir Eigenschaft ii. zeigen, beweisen wir noch

Proposition 6.11. *Es sei* $A \in K^{n \times n}$*. Dann gilt* $\det(A) = \det(A^T)$.

Beweis. Es sei $A = (a_{ij})_{1 \leq i,j \leq n}$. Dann ist $A^T = (a_{ji})_{1 \leq i,j \leq n}$ und somit

$$\det(A^T) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)}\,.$$

Hier benutzen wir, dass für alle $\sigma \in S_n$

$$a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)}$$

denn links und rechts stehen die gleichen Faktoren - nur in anderer Reihenfolge. Durchläuft σ ganz S_n, so auch σ^{-1} und wir erhalten schließlich wegen $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$

$$\det(A^T) = \sum_{\sigma^{-1} \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)} \,.$$

Wenn σ ganz S_n durchläuft, so auch σ^{-1}. Daher ist die rechte Seite gleich $\det(A)$. □

25.1.10

Es sei nun $A = (a_{ij}) \in \operatorname{Mat}(n, K)$ und es seien die i_1-te und i_2-te Zeile von A gleich. Wir wollen zeigen, dass $\det(A) = 0$.

Für diese Rechnung ist folgende Definition praktisch.

Definition. *Es sei $n \in \mathbb{N}$, $n \geq 1$. Wir setzen*

$$A_n := \ker(\operatorname{sgn} : S_n \to \{\pm 1\}) < S_n \,.$$

Diese Untergruppe von S_n heißt alternierende Gruppe.

Die alternierende Gruppe A_n besteht also genau aus den geraden Permutationen in S_n.

Es sei nun $n \geq 2$ und $\tau \in S_n$ eine Transposition. Ist $\sigma \in S_n$ beliebig, so hat genau eine der Permutationen σ oder $\tau \circ \sigma$ Signum $+1$. Es ist also $\sigma \in A_n$ oder es gibt ein eindeutiges $\sigma' \in A_n$ mit $\sigma = \tau \circ \sigma'$ (nämlich $\sigma' := \tau \circ \sigma$). Also ist

$$S_n = A_n \dot\cup (\tau \circ A_n) \,,$$

wobei $\tau \circ A_n := \{\tau \circ \sigma \mid \sigma \in A_n\}$ und $\dot\cup$ die disjunkte Vereinigung bezeichnet. Somit gilt für die Ordnung (d.h. die Anzahl der Elemente) von A_n

$$|A_n| = \frac{1}{2}|S_n| = \frac{n!}{2} \,.$$

Wir kommen nun zum Beweis von $\det(A) = 0$, falls Zeilen i_1 und Zeile i_2 von A übereinstimmen. Da $\det(A) = \det(A^T)$, haben wir

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} \,.$$

Es sei $\tau \in S_n$ die Transposition, die i_1 und i_2 vertauscht. Wir können dann aufgrund der Zerlegung $S_n = A_n \dot\cup (\tau \circ A_n)$ schreiben

$$\det A = \sum_{\sigma \in A_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} + \sum_{\sigma \in A_n} \operatorname{sgn}(\tau \circ \sigma) a_{\tau(\sigma(1))1} \cdot \ldots \cdot a_{\tau(\sigma(n))n} \,.$$

Für jedes $\sigma \in A_n$ ist aber

$$\operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = -\operatorname{sgn}(\tau \circ \sigma) a_{\tau(\sigma(1))1} \cdot \ldots \cdot a_{\tau(\sigma(n))n}$$

da $\operatorname{sgn}(\sigma) = 1$, $\sigma(\tau \circ \sigma) = -1$ und die Produkte der auftretenden Komponenten von A gleich sind: τ vertauscht ja nur i_1 und i_2 und die Zeilen i_1 und i_2 in A stimmen überein. Da sich in $\det(A)$ also die Terme paarweise wegheben, ist also in der Tat $\det(A) = 0$.

Wir bemerken, dass man diesen Beweis auch direkt an Hand der Formel für $\det(A)$ zeigen kann, vgl. [Fischer], S. 194, Beweis von D2.

Da immer $\det(A) = \det(A^T)$ (siehe Proposition 6.11), halten wir noch fest:

Korollar 6.12. *Stimmen in A zwei Spalten überein, so gilt $\det(A) = 0$. Addieren wir ein Vielfaches einer Spalte zu einer anderen Spalte, so ändert sich die Determinante nicht.*

Die Determinante hat die folgende fundamentale Multiplikativitätseigenschaft.

Satz 6.13. *Für alle $A, B \in \operatorname{Mat}(n, K)$ gilt*

$$\det(B \cdot A) = \det B \cdot \det A\,.$$

Beweis. Falls rang(A) < n, so gilt auch rang(BA) < n und somit $\det(BA) = 0 = \det(B) \cdot \det(A)$ (vgl. Proposition 6.9).

Es sei nun rang(A) = n, also insbesondere $\det(A) \neq 0$. Wir betrachten die Abbildung

$$\phi : \operatorname{Mat}(n, K) \to K\,,\ M \mapsto \frac{\det(MA)}{\det A}$$

Die Funktion ϕ hat die obigen Eigenschaften i., ii. und iii., die eine Determinante charakterisieren. Dies zeigt man direkt durch Betrachtung des Produktes MA für verschiedene M und die entsprechenden Eigenschaften von det. Somit gilt nach Theorem 6.8, dass $\phi = \det$, also

$$\frac{\det(MA)}{\det A} = \det(M)$$

für alle $M \in \operatorname{Mat}(n, K)$. Daraus folgt die Behauptung, indem wir $M := B$ setzen. □

An diesem Beweis zeigt sich der Vorteil der axiomatischen Charakterisierung von det. Wenn wir hier direkt mit der Formel aus Theorem 6.10 von det arbeiten, treten hingegen sehr unübersichtliche Ausdrücke auf.

Korollar 6.14. *Die Abbildung* det *induziert einen Gruppenhomomorphismus (zur Erinnerung: Für einen Körper K bezeichnet K^* die Teilmenge $K \setminus \{0\}$).*

$$\det : \operatorname{GL}(n, K) \to (K^*, \cdot, 1)$$

Mit der expliziten Formel kann man für $n = 2$ und $n = 3$ die Determinante nach der sogenannten *Sarrus'schen Regel* berechnen, vgl. [Fischer], S. 195.

Für größere n ist die Berechnung von $\det(A)$ komplizierter. Wir diskutieren daher einige arithmetische Eigenschaften der Determinante. Diese zeigen, wie man die Berechnung von Determinanten einer Matrix A auf die Berechnung von Determinanten von Teilmatrizen von A zurückführen kann. Daneben haben diese arithmetischen Eigenschaften auch interessante theoretische Konsequenzen. Leider sind die Rechnungen etwas technisch. Aber der Aufwand lohnt sich.

Ist $A \in K^{n\times n}$ und $1 \leq i,j \leq n$, so bezeichnen wir mit A_{ij} die Matrix, die aus A durch Ersetzen des ij-ten Achsenkreuzes durch das Kreuz mit 1 an der Stelle ij und 0 sonst ensteht. Weiterhin bezeichne $A^\sharp := (a^\sharp_{ij})$ mit $a^\sharp_{ij} := \det A_{ji}$ (man beachte die Vertauschung der Indizes) die *Komplementärmatrix* von A, sowie $A'_{ij} \in K^{(n-1)\times(n-1)}$ die ij-te *Streichungsmatrix*, die durch Streichen der i-ten Zeile und j-ten Spalte aus A entsteht. Man vergleiche [Fischer], S. 201.

Lemma 6.15. *Für alle $1 \leq i,j \leq n$ gilt*

- $\det A_{ij} = (-1)^{i+j} \det A'_{ij}$.
- $\det A_{ij} = \det(A(\text{Spalte } j = e_i))$. *Dabei ist für jeden Vektor $a \in K^n$ die Matrix $A(\text{Spalte } j = a)$ diejenige Matrix, die aus A entsteht, indem wir die j-te Spalte durch den Vektor $a \in K^n$ ersetzen.*

Zum Beweis siehe [Fischer], S. 201 f.

Wir erhalten das folgende erstaunliche Resultat.

Proposition 6.16. *Es sei $A \in K^{n\times n}$. Dann ist*

$$A^\sharp \cdot A = A \cdot A^\sharp = (\det A) \cdot E_n \,.$$

Beweis. Der ik-te Eintrag von $A^\sharp \cdot A$ ist gegeben durch

$$\sum_{j=1}^n (\det A_{ji}) \cdot a_{jk} = \sum_{j=1}^n a_{jk} \det A(\text{Spalte } i = e_j) = \det A(\text{Spalte } i = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix})$$

und dies ist gleich $\delta_{ik} \det A$. Dabei benutzen wir die Tatsache, dass die Determinante einer Matrix gleich 0 ist, wenn in dieser Matrix zwei Spalten übereinstimmen. Damit ist die erste Gleichung bewiesen.

Die zweite Gleichung zeigt man analog: Der ik-te Eintrag von $A \cdot A^\sharp$ ist gegeben durch $\sum_{j=1}^n a_{ij}(\det A_{kj})$ und dies ist gleich

$$\sum_{j=1}^n a_{ij} \det A(\text{Zeile } k = e_j) = \det A(\text{Zeile } k = \begin{pmatrix} a_{i1} & \dots & a_{in} \end{pmatrix}) = \delta_{ik} \cdot \det A \,.$$

Dabei haben wir die analoge Aussage zu Lemma 6.15 benutzt, die sich auf Ersetzen der i-ten Zeile bezieht. □

Korollar 6.17. *Es sei $A \in \mathrm{GL}(n,K)$. Dann ist die inverse Matrix von A gleich*

$$A^{-1} = \frac{A^\sharp}{\det A} \,.$$

Im Fall $n = 2$ ergibt sich die folgende nützliche Tatsache: Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so ist A genau dann invertierbar, falls

$$\det A = ad - bc \neq 0$$

und in diesem Fall gilt

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Proposition 6.18 (Laplace'scher Entwicklungssatz)**.** *Es sei $A \in K^{n\times n}$. Dann gilt*

- *Entwicklung nach der i-ten Zeile: Für alle $1 \leq i \leq n$ ist*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij} \,.$$

- *Entwicklung nach der j-ten Spalte: Für alle $1 \leq j \leq n$ ist*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij} \,.$$

Beweis. Nach der zweiten Rechnung im Beweis von Proposition 6.16 und nach Lemma 6.15 haben wir

$$\det A = \delta_{ii} \cdot \det A = \sum_{j=1}^{n} a_{ij} \cdot \det A_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij} \,.$$

Dies ist die Entwicklung nach der i-ten Zeile. Die Entwicklung nach der j-ten Spalte zeigt man wieder analog (mit Hilfe der ersten Rechnung im Beweis von Proposition 6.16). □

Eine ganz ähnliche Rechnung zeigt folgende Methode zur Lösung linearer Gleichungssysteme.

Satz 6.19 (Cramersche Regel)**.** *Es sei $A \in \mathrm{GL}(n, K)$ und $b \in K^n$. Dann ist die (eindeutig bestimmte) Lösung des linearen Gleichungssystems*

$$Ax = b$$

gegeben durch den Vektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, wobei

$$x_j = \frac{\det(A(\text{Spalte } j = b))}{\det A}$$

für $j = 1, \ldots, n$.

Beweis. Die Entwicklung von $\det A(\text{Spalte } j = b)$ nach der j-ten Spalte führt auf

$$\sum_{i=1}^{n}(-1)^{i+j}b_i \cdot \det A'_{ij} = \sum_{i=1}^{n}(\det A_{ij})b_i = \sum_{i=1}^{n} a^\sharp_{ji} b_i = (A^\sharp \cdot b)_j$$

wobei der Index j bedeutet, dass wir nur die j-te Komponente betrachten. Es gilt aber in K^n die Gleichung

$$A^\sharp \cdot b = (\det A) \cdot A^{-1} b$$

wie man durch Multiplikation der Gleichung $A \cdot A^\sharp \cdot b = (\det A) \cdot E_n \cdot b$ (die aus Proposition 6.16 folgt) von links mit A^{-1} sieht. Damit ist alles gezeigt. □

1.2.10

7. Polynomringe, der Fundamentalsatz der Algebra

Wir kehren in diesem Kapitel wieder zu den algebraischen Grundstrukturen zurück und diskutieren Polynomringe, die in der Algebra eine zentrale Rolle spielen. Polynome werden später in der Eigenwerttheorie wichtig.

Es sei $R = (R, +, 0, \cdot)$ ein kommutativer Ring. Der *Polynomring* $R[X]$ in der „Unbestimmten“ X (die auch oft mit dem Buchstaben t bezeichnet wird) besteht aus allen formalen Ausdrücken der Form

$$a_0 + a_1 \cdot X + \ldots + a_n \cdot X^n$$

wobei $a_0, \ldots, a_n \in R$. Die Addition und Multiplikation zweier solcher Ausdrücke erfolgt nach den üblichen Regeln:

$$(a_0 + \ldots + a_n X^n) + (b_0 + \ldots + b_n X^n) := (a_0 + b_0) + (a_1 + b_1)X + \ldots + (a_n + b_n)X^n$$

und

$$(a_0 + \ldots + a_n X^n) \cdot (b_0 + \ldots + b_m X^m) := a_0 b_0 + (a_0 b_1 + a_1 b_0)X + \ldots + a_n b_m X^{m+n}$$

Neutrales Element der Addition ist das *Nullpolynom* 0 und neutrales Element der Multiplikation ist das Polynom 1. Damit wird $R[X]$ wieder ein kommutativer Ring mit 1. Wichtig ist hier, dass X kein Element aus R bezeichnet, sondern zunächst nur ein Symbol ist, mit dem wir nach bestimmten Regeln rechnen können.

Dieses Vorgehen ist jedoch nicht ganz befriedigend. Es bleibt die Schwierigkeit, wann wir zwei Polynome der obigen Form als gleich ansehen wollen. Zunächst denkt man, dies sei genau dann der Fall, wenn genau alle Koeffizienten übereinstimmen. Das ist aber nicht ganz richtig: Wir wollen ja zum Beispiel die Polynome $1 + X$ und $1 + X + 0 \cdot X^2$ identifizieren. Mathematisch würde dies bedeuten, dass wir auf der Menge der obigen formalen Ausdrücke eine geeignete Äquivalenzrelation einführen. Dann müssten wir aber nachweisen, dass die oben definierten Operationen wohldefiniert sind. Außerdem haben wir überhaupt noch nicht geklärt, was wir unter „formalen Ausdrücken“ verstehen wollen.

Ein mathematisch sauberes Vorgehen ist wie folgt.

Es sei $R[X]$ die Menge der Folgen $(a_0, a_1, \ldots)$ (d.h. Abbildungen $\mathbb{N} \to R$), wobei $a_i \in R$ für alle $i \in \mathbb{N}$, bei denen nur endlich viele a_i von 0 verschieden sind. Das Symbol X hat an dieser Stelle noch keine eigenständige Bedeutung. Durch komponentenweise Addition wird $R[X]$ zu einer abelschen Gruppe mit neutralem Element $0 := (0, 0, \ldots)$. Wir definieren nun noch eine Multiplikation auf $R[X]$ durch die Formel

$$(a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) := (p_0, p_1, p_2, \ldots),$$

wobei $p_i := \sum_{j=0}^{i} a_j b_{i-j} = \sum_{j+k=i} a_j b_k$. Das Element $(p_0, p_1, \ldots)$ liegt dabei wieder in $R[X]$, denn ist $a_i = 0$ für alle $i > n$ und $b_j = 0$ für alle $j > m$, dann ist $p_k = 0$ für alle $k > m + n$. Wir setzen noch $1 := (1, 0, \ldots) \in R[X]$.

Proposition 7.1. *$(R[X], +, 0, \cdot, 1)$ ist ein kommutativer Ring mit 1.*

Beweis. Es ist klar, dass $(R[X], +, 0)$ eine abelsche Gruppe ist.

Weiterhin folgt

$$1 \cdot (a_0, a_1, \ldots) = (a_0, a_1, \ldots)$$

für alle $(a_0, a_1, \ldots) \in R[X]$ direkt aus der Definition der Multiplikation auf $R[X]$.

Die Kommutativität der Multiplikation auf $R[X]$ folgt aus der Definition der p_i sowie der Kommutativität der Multiplikation in R.

Wir zeigen nun das Assoziativgesetz für die Multiplikation: Sind $A := (a_0, a_1, \ldots), B := (b_0, b_1, \ldots), C := (c_0, c_1, \ldots) \in R[X]$, dann ist die i-te Komponente ($i \in \mathbb{N}$) in $(AB)C$ gleich

$$\sum_{j+k=m, m+l=i} (a_j b_k) c_l = \sum_{j+k+l=i} (a_j b_k) c_l$$

Ganz ähnlich ist die i-te Komponente von $A(BC)$ gleich $\sum_{j+k+l=i} a_j (b_k c_l)$, und das Assoziativgesetz in $R[X]$ folgt aus dem Assoziativgesetz in R.

Der Beweis des Distributivgesetzes ist recht einfach und wird hier nicht ausgeführt.

□

Wir betrachten nun die Abbildung

$$\phi : R \to R[X], a \mapsto (a, 0, \ldots)$$

Es ist leicht zu sehen, dass dies ein injektiver Ringhomomorphismus ist. Indem wir jedes $a \in R$ mit $\phi(a) \in R[X]$ identifizieren, können wir daher im folgenden R als Unterring von $R[X]$ ansehen.

Wir bezeichnen nun das Element $(0, 1, 0, \ldots) \in R[X]$ mit dem Symbol X. Durch Induktion nach $k \geq 1$ zeigt man, dass

$$X^k = (0, \ldots, 0, 1, 0, \ldots)$$

mit der 1 an der k-ten Komponente. Wie in jedem Ring setzen wir noch $X^0 := 1$. Es gilt in $R[X]$ dann für alle $k \geq 0$ und alle $a \in R \subset R[X]$ die Gleichung

$$aX^k = (0, \ldots, 0, a, 0, \ldots)$$

mit a an der k-ten Komponente. Dies zeigt man ebenfalls durch Anwenden der Formel für die Multiplikation in $R[X]$. Daher gilt in $R[X]$ die Gleichung

$$(a_0, a_1, \ldots, a_n, 0, 0, \ldots) = a_0 + a_1 X + \ldots + a_n X^n$$

und wir haben damit wieder unsere gewohnte Darstellung von Polynomen wiedergefunden.

Nun sind wir aber einen Schritt weiter, denn der Ausdruck auf der rechten Seite ist Element in einem sauber definierten algebraischen Objekt $R[X]$. Wir bezeichnen Elemente in $R[X]$ oft mit dem Buchstaben f oder auch mit dem Symbol $f(X)$, wenn wir betonen wollen, wie wir die „Unbekannte“ X nennen.

Definition. *Der Ring $R[X]$ heißt* Polynomring über R in der Unbestimmten X.

Diese Konstruktion können wir iterieren: Da $R[X]$ wieder ein kommutativer Ring mit Eins ist, erhalten wir einen Polynomring $(R[X])[Y]$ über $R[X]$ in der Unbestimmten Y. Ist $f \in R[X]$, so können wir also schreiben

$$f = a_0 + \ldots + a_n Y^n$$

wobei $a_i \in R[X]$ für alle i. Somit ist

$$f = \sum_{i,j \geq 0} a_{ij} X^i Y^j$$

wobei $a_{ij} \in R$ für alle i, j und nur endlich viele a_{ij} von 0 verschieden sind. Wir können dieses Element in $(R[X])[Y]$ auf kanonische Weise mit dem Element

$$\sum_{i,j \geq 0} a_{ij} Y^j X^i \in (R[Y])[X]$$

identifizieren und umgekehrt. Es ist nicht schwer zu sehen, dass diese Identifikation von $(R[X])[Y]$ und $(R[Y])[X]$ mit den algbraischen Strukturen verträglich ist, d.h. diese Ringe sind kanonisch isomorph. Wir lassen daher im folgenden die Klammern weg und definieren iterativ den Polynomring

$$R[X_1, \ldots, X_n]$$

in den Unbestimmten $X_1, \ldots, X_n$. Ein typisches Element in diesem Polynomring ist von der Gestalt

$$\sum_{i_1, \ldots, i_n \geq 0} a_{i_1 \ldots i_n} X_1^{i_1} \cdot \ldots \cdot X_n^{i_n}$$

mit $a_{i_1 \ldots i_n} \in R$ für alle $(i_1, \ldots, i_n) \in \mathbb{N}^n$, wobei nur endlich viele dieser Koeffizienten von 0 verschieden sind. Addition und Multiplikation solcher Polynome erfolgt in der üblichen Weise. Wir nennen Elemente der Gestalt

$$X_1^{i_1} \cdot \ldots \cdot X_n^{i_n}$$

Monome. Es folgt aus der Konstruktion, dass in $R[X_1, \ldots, X_n]$ die Gleichung

$$X_1^{i_1} \cdot \ldots \cdot X_n^{i_n} = X_1^{j_1} \cdot \ldots \cdot X_n^{j_n}$$

genau dann gilt, falls $(i_1, \ldots, i_n) = (j_1, \ldots, j_n) \in \mathbb{N}^n$ und dass die Relation

$$\sum_{i_1,\ldots,i_n \geq 0} a_{i_1 \ldots i_n} X_1^{i_1} \cdot \ldots \cdot X_n^{i_n} = 0$$

genau dann gilt, falls alle Koeffizienten $a_{i_1 \ldots i_n} = 0$.

Wir studieren nun noch einige Eigenschaften von Polynomringen in einer Unbestimmten. Ist $f \in R[X]$ und $f \neq 0$, so können wir mit einem eindeutig bestimmten $n \in \mathbb{N}$

$$f = a_0 + a_1 X + \ldots + a_n X^n$$

schreiben, wobei $a_n \neq 0$. Wir nennen n den *Grad* von f, geschrieben $\deg f$, und a_n den *Leitkoeffizienten* (oder *höchsten Koeffizienten*). Wir setzen noch $\deg 0 := -\infty$, wobei $-\infty$ nur als Symbol zu verstehen ist, das den üblichen Regeln $-\infty < n$ für alle $n \in \mathbb{N}$, sowie $(-\infty) + (-\infty) = -\infty$, $-\infty + n = -\infty$ für alle $n \in \mathbb{N}$ genügt.

Man beachte, dass $f \in R$ genau dann gilt, falls $\deg f \in \{-\infty, 0\}$ und $f \in R \setminus \{0\}$ genau dann, falls $\deg f = 0$. Sind $f, g \in R[X]$, so gilt mit diesen Setzungen die Ungleichung

$$\deg(f + g) \leq \max(\deg(f), \deg(g))$$

und hier kann Ungleichheit nur dann eintreten (muss aber nicht!), falls $\deg f = \deg g$.

Wir nehmen nun zusätzlich an, dass R nullteilerfrei ist (dies ist zum Beispiel der Fall, wenn R ein Körper ist). Haben wir dann Polynome

$$f = a_0 + \ldots + a_n X^n\,,\ g = b_0 + \ldots + b_m X^m$$

gegeben, wobei $a_n, b_m \neq 0$, so ist

$$f \cdot g = a_0 b_0 + (a_0 b_1 + a_1 b_0) X + \ldots + (a_n b_m) X^{n+m}$$

mit $a_n b_m \neq 0$, da R nullteilerfrei ist. Somit gilt

$$\deg(f \cdot g) = \deg f + \deg g\,.$$

Mit unserer obigen Konvention betreffend $-\infty$ bleibt diese Gleichung für alle $f, g \in R[X]$ gültig, wenn R nullteilerfrei ist.

3.2.10

Proposition 7.2. *Es sei R ein kommutativer Ring mit 1 und zudem nullteilerfrei. Dann ist auch $R[X]$ nullteilerfrei.*

Beweis. Es seien $f, g \in R[X]$ mit $fg = 0$, d.h. $\deg(fg) = -\infty$. Mit der Gradgleichung, die wir gerade diskutiert haben, ist dies nur möglich, wenn entweder $\deg f = -\infty$ oder $\deg g = -\infty$. □

Korollar 7.3. *Es sei R nullteilerfrei. Dann ist auch der Polynomring $R[X_1, \ldots, X_n]$ in n Unbestimmten nullteilerfrei.*

Wir betrachten nun der Einfachheit halber Polynome in einer Unbestimmten über einem Körper K. Das im nächsten Satz beschriebene Verfahren ist auch unter dem Namen *Polynomdivision* bekannt und analog zur Division mit Rest für ganze Zahlen. Man vergleiche Aufgabe 1 auf Übungsblatt 4.

Satz 7.4. *Es seien $f, g \in K[X]$ Polynome und $g \neq 0$. Dann gibt es eindeutig bestimmte Polynome $q, r \in K[X]$, so dass*

- $f = gq + r$ *und*
- $\deg r < \deg g$.

Beweis. Wir beweisen zunächst die Existenz von q und r. Falls $f = 0$, schreiben wir einfach $f = 0g + g$. Es sei nun $\deg f \geq 0$. Wir schreiben

$$\begin{aligned} f &= a_0 + \ldots + a_n X^n \\ g &= b_0 + \ldots + b_d X^d \,. \end{aligned}$$

wobei $a_n, b_d \neq 0$ (d.h. $n = \deg f$, $d = \deg g$). Wir führen Induktion nach n. Falls $n = 0$ und $\deg g > \deg f$, dann setzen wir $q := 0$ und $r := f$. Falls $\deg g = \deg f = 0$, dann setzen wir $q := \frac{a_0}{b_0}$ und $r := 0$.

Angenommen, die Aussage ist schon für alle f und g mit $0 \leq \deg f < n$ gezeigt. Es sei $\deg f = n$. Falls $\deg g > \deg f$, setzen wir wieder $q := 0$ und $r := f$. Sei also nun $\deg g \leq \deg f$. Dann ist

$$f_1 := f - \frac{a_n}{b_d} X^{n-d} g \in K[X]$$

ein Polynom mit $\deg f_1 < \deg f$. Falls $f_1 = 0$, setzen wir $q := \frac{a_n}{b_d} X^{n-d}$ und $r := 0$. Falls $\deg f_1 \geq 0$, erhalten wir nach Induktionsannahme Polynome $q_1, r \in K[X]$ mit

$$f_1 = q_1 g + r$$

und $\deg r < \deg g$. Setzen wir also $q := \frac{a_n}{b_d} X^{n-d} + q_1$, so haben wir

$$f = qg + r$$

und $\deg r < \deg g$. Damit ist die Existenz einer behaupteten Darstellung von f gezeigt.

Wir kommen nun zur Eindeutigkeit. Angenommen

$$f = q_1 g + r_1 = q_2 g + r_2 \,,$$

mit $\deg r_1, \deg r_2 < \deg g$. Durch Subtraktion erhalten wir

$$(q_1 - q_2) g = r_2 - r_1$$

Nach der Gradformel gilt aber andererseits

$$\deg((q_1 - q_2)g) = \deg(q_1 - q_2) + \deg g$$

Da $\deg(r_2 - r_1) < \deg g$ kann diese Gleichung nur gelten, falls $\deg(q_1 - q_2) = -\infty$, also $q_1 = q_2$. Durch Einsetzen in die obige Gleichung folgt daraus auch $r_1 = r_2 = 0$, also die behauptete Eindeutigkeit. □

Haben wir ein Polynom $f \in R[X]$ gegeben und schreiben wir

$$f = a_0 + \ldots + a_n X^n$$

so erhalten wir für jedes Element $\lambda \in R$ das Element

$$f(\lambda) := a_0 + a_1\lambda + \ldots + a_n\lambda^n \in R\,,$$

durch *Einsetzen* von λ in die Unbestimmte X. Wir erhalten auf diese Weise einen Ringhomomorphismus („Auswertung bei λ“)

$$\iota_\lambda : R[X] \to R\,,\ f \mapsto f(\lambda)\,.$$

Nach Definition gilt hier für das Nullpolynom $0(\lambda) = \iota_\lambda(0) := 0$. Wir erhalten weiterhin eine Abbildung

$$\tilde{}\colon R[X] \to \mathrm{Abb}(R,R)\,, f \mapsto \tilde{f}$$

wobei $\tilde{f} : R \to R$ durch $\lambda \mapsto f(\lambda) = \iota_\lambda(f)$ gegeben ist. Diese Abbildung $\tilde{f}$ ist die zu f gehörige *Polynomfunktion*. Diese müssen wir sorgfältig von $f \in R[X]$ unterscheiden, denn im allgemeinen ist die Abbildung $\tilde{}$ weder injektiv noch surjektiv.

Definition. *Es sei K ein Körper, $f \in K[X]$ und $\lambda \in K$. Wir nennen λ eine* Nullstelle *von f, falls $f(\lambda) = 0$.*

Die Existenz von Nullstellen ist ein zentrales Thema der Algebra.

Beispiel. Das Polynom $X^2 + 1 \in \mathbb{R}[X]$ besitzt keine Nullstellen. Fassen wir dieses Polynom jedoch als Element von $\mathbb{C}[X]$ auf, dann hat es die Nullstellen $-i$ und i. Ist $k = \{k_1, \ldots k_r\}$ ein endlicher Körper, so hat das Polynom $f := (X - k_1) \cdot \ldots \cdot (X - k_r) + 1$ keine Nullstellen.

Wir beschäftigen uns nun mit der Frage, wie viele Nullstellen ein Polynom höchstens besitzen kann.

Proposition 7.5. *Es sei $f \in K[X]$, $f \neq 0$, und $\lambda \in K$ eine Nullstelle von f, so gibt es ein eindeutig bestimmtes Polynom $g \in K[X]$ mit $f = (X-\lambda)\cdot g$ und* $\deg g = \deg f - 1$.

Der Beweis findet sich in [Fischer], S. 65.

Korollar 7.6. *Es sei $f \in K[X]$, $f \neq 0$ und k die Anzahl der Nullstellen von f. Dann gilt $k \leq \deg f$. Insbesondere ist die Anzahl der Nullstellen von f endlich.*

Mit anderen Worten: Hat ein Polynom f vom Grad $\leq k$ mehr als k Nullstellen, so gilt zwangsläufig schon $f = 0$. Der Beweis des Korollares findet sich ebenfalls in [Fischer], S. 65.

Beispielsweise folgt aus diesem Korollar, dass $\sin : \mathbb{R} \to \mathbb{R}$ nicht die zu einem Polynom in $\mathbb{R}[X]$ gehörige Polynomfunktion sein kann.

Wir ergänzen am Schluss dieser Vorlesung noch einige Dinge zu Determinanten mit Einträgen in einem kommutativen Ring R mit 1. Dies brauchen

wir später für den Fall, dass $R = K[X]$ ein Polynomring über einem Körper K ist.

Wir betrachten die Menge $R^{m\times n}$ der $m \times n$-Matrizen mit Einträgen in R. Falls $m = n$, so schreiben wir auch $\mathrm{Mat}(n, R)$. Dies ist mit der üblichen Matrixmultiplikation ein (i.a. nicht kommutativer) Ring mit Einselement E_n (die $n \times n$-Matrix mit Einsen auf der Diagonale und Nullen sonst).

Wir wollen wieder zeigen, dass es eine eindeutige *Determinantenabbildung*

$$\det : \mathrm{Mat}(n, R) \to R$$

gibt, die die Eigenschaften i., ii. und iii. hat, die wir auf Seite 63 für eine Determinante gefordert haben. Die Exististenz zeigt man wieder durch Angabe der expliziten Formel

$$\det A = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \,,$$

falls $A = (a_{ij})$ mit $a_{ij} \in R$. Diese Abbildung hat die drei geforderten Eigenschaften. Dies kann man genauso wie früher in Satz 6.10 zeigen, da wir dazu nicht durch Elemente aus R dividieren müssen.

Bei der Diskussion der Eindeutigkeit in Theorem 6.8 haben wir jedoch elementare Zeilenumformungen benutzt, die auch die Division durch gewisse Elemente benutzen. Hier brauchen wir also ein anderes Argument.

Wir schreiben dazu den i-ten Zeilenvektor $(a_{i1}, \ldots, a_{in})$ von A als

$$(a_{i1}, \ldots, a_{in}) = \sum_{j=1}^{n} a_{ij} e_j$$

mit dem kanonischen Einheitsvektor $e_j \in R^n$ (geschrieben als Zeilenvektor). Durch sukzessives Anwenden von Regel i. auf Zeile 1, Zeile 2, etc. erhalten wir für $\det A$

$$\sum_{j_1=1}^{n} a_{1j_1} \det A(\text{Zeile } 1 = e_{j_1}) = \ldots = \sum_{1\le j_1,\ldots,j_n\le n} a_{1j_1} \cdot \ldots \cdot a_{nj_n} \det \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} .$$

Die Determinante der Matrix

$$\begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix}$$

ist gleich 0, wenn $j_1, \ldots, j_n$ nicht paarweise verschieden sind, denn dann stimmen zwei Zeilen überein. Falls diese Indizes paarweise verschieden sind, ist die Determinante gleich $\mathrm{sgn}(\sigma)$, wobei $\sigma(i) := j_i$ für alle $1 \le i \le n$. Dies folgt aus den Regeln ii. und iii. Damit erhält man genau die obige Definition von $\det A$ und die Eindeutigkeit der Determinante ist gezeigt.

Der Determinantenmultiplikationssatz Theorem 6.13, Proposition 6.16 über die Komplementärmatrix und der Laplace'sche Entwicklungssatz 6.18

behalten für Matrizen in $\mathrm{Mat}(n, R)$ ebenfalls ihre Gültigkeit. Diese Tatsachen werden wir nicht benötigen und wir verzichten hier auf die Beweise.

8.2.10

Oft ist es günstig, Nullstellen mit Vielfachheiten zu zählen.

Definition. *Es sei $f \in K[X]$ und $\lambda \in K$. Wir setzen dann*

$$\mu(f, \lambda) := \max\{k \in \mathbb{N} \mid \exists g \in K[X] \text{ mit } f = (X - \lambda)^k g\}$$

Dies ist die Vielfachheit *der Nullstelle λ.*

$\lambda \in K[X]$ ist also offensichtlich genau dann eine Nullstelle von f, falls $\mu(f, \lambda) \geq 1$.

Ist $f \in K[X]$ und sind $\lambda_1, \ldots, \lambda_k \in K$ die Nullstellen von f und haben diese Vielfachheiten $\mu_1, \ldots, \mu_k \geq 1$, so gilt also nach Proposition 7.5

$$f = (X - \lambda_1)^{\mu_1} \cdot \ldots \cdot (X - \lambda_k)^{\mu_k} g$$

wobei g in Polynom ohne Nullstellen ist. Im Falle $K = \mathbb{C}$ hat dieses Polynom g den Grad 0 nach dem Fundamentalsatz der Algebra:

Satz 7.7 (Fundamentalsatz der Algebra, Gauß 1799)**.** *Es sei $p \in \mathbb{C}[X]$ ein nichkonstantes Polynom (d.h.* $\deg p \geq 1$*). Dann besitzt p eine Nullstelle in $\mathbb{C}$.*

Beweis. Der Beweis von Gauß stützt sich auf topologische Methoden. Wir geben hier einen Beweis, der auf d'Alembert (1746) und Argand (1806) zurückgeht und nur elementare Tatsachen der Analysis benutzt.

Wir zeigen zunächst: Eine stetige Funktion $f : \mathbb{C} \to \mathbb{R}$ nimmt auf jeder abgeschlossenen Kreisscheibe $D := \{z \in \mathbb{C} \mid |z| \leq R\}$ ihr Minimum an. Dabei heißt f *stetig* in $a \in \mathbb{C}$, wenn für jede konvergente Folge $z_n \to a$ komplexer Zahlen gilt: $f(z_n) \to f(a)$.

Sei $m \in \mathbb{R} \cup \{-\infty\}$ das Infimum von f auf D. Es gibt dann eine Folge (z_n) in D mit $f(z_n) \to m$. Da D beschränkt ist, sind die Real- und Imaginärteile von (z_n) beschränkt. Nach dem Satz von Bolzano-Weierstraß (angewandt zunächst auf den Realteil der gegebenen Folge und dann auf den Imaginärteil der erhalteten Teilfolge von (z_n)) besitzt $(z_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(z_{n_k})_{k\in\mathbb{N}}$. Sei $a \in \mathbb{C}$ der Grenzwert dieser Teilfolge. Da $|z_n| \leq R$ für alle n und die Betragsfunktion $|\cdot| : \mathbb{C} \to \mathbb{R}$ stetig ist, gilt $|a| \leq R$, also $a \in D$. Wegen $f(a) = \lim\limits_{k\to\infty} f(z_{n_k}) = m$ nimmt f das Minimum in $a \in D$ an.

Es sei nun ein nicht konstantes Polynom $p = \sum_{k=0}^{n} a_n z^n$ gegeben, wobei $a_i \in \mathbb{C}$, $n \geq 1$ und $a_n \neq 0$.

Für $0 \neq z \in \mathbb{C}$ haben wir

$$|p(z)| = |z^n| \cdot |a_n + \sum_{k=0}^{n-1} a_k z^{k-n}| \geq |z|^n \cdot (|a_n| - \sum_{k=0}^{n-1} |a_n z^{k-n}|)\,.$$

Es sei $K := \frac{|p(0)|}{|a_n|}$. Es gibt nun ein $R \in \mathbb{R}_{>0}$, so dass $|z|^n \geq 2K$ und außerdem $\sum_{k=0}^{n-1} |a_k z^{k-n}| \leq \frac{|a_n|}{2}$ für alle $|z| \geq R$ gilt (beachte, dass $k - n < 0$ für

$0 \leq k \leq n-1$.) Für $|z| \geq R$ ist also $|p(z)| \geq |p(0)|$. Die stetige Funktion $z \mapsto |p(z)|$ nimmt auf der abgeschlossenen Kreisscheibe $\{z \in \mathbb{C} \mid |z| \leq R\}$ ihr Minimum in einem Punkt z_0 an. Nach Wahl von R ist dann $|p(z_0)|$ auch ein Minimum der auf ganz $\mathbb{C}$ definierten Funktion $z \mapsto |p(z)|$.

Durch Übergang von $p(z)$ zu $p(z+z_0)$ können wir $z_0 = 0$ annehmen. Wir zeigen $p(0) = 0$, indem wir die Annahme $p(0) \neq 0$ zu einem Widerspruch führen. Sei also $p(0) \neq 0$. Wir haben dann eine Polardarstellung

$$p(0) = re^{i\phi}$$

mit $r \in \mathbb{R}_{>0}$ und $\phi \in \mathbb{R}$. Es sei $k \in \mathbb{N}$ minimal mit $k > 0$ und $a_k \neq 0$. Ohne Einschränkung der Allgemeinheit können wir $a_k = 1$ annehmen (indem wir p durch a_k teilen). Wir haben also

$$p(z) = a_n z^n + \ldots + a_{k+1} z^{k+1} + z^k + p(0)\,.$$

Für $\epsilon > 0$ setzen wir nun

$$z_\epsilon := \epsilon \sqrt[k]{r} \cdot e^{\frac{i}{k}(\phi+\pi)}\,.$$

Es gilt $z_\epsilon^k = \epsilon^k r e^{i(\phi+\pi)} = -\epsilon^k r e^{i\phi}$. Damit haben wir

$$p(z_\epsilon) - p(0) = \epsilon^k(-re^{i\phi} + r(\epsilon))\,,$$

wobei $\lim\limits_{\epsilon \to 0} |r(\epsilon)| = 0$. Wählen wir $\epsilon > 0$ klein genug, so ist also

$$|r(\epsilon)| \leq \frac{1}{2} \cdot r$$

und es folgt mit der Dreieckungleichung

$$|p(z_\epsilon)| = |re^{i\phi} - \epsilon^k r e^{i\phi} + \epsilon^k r(\epsilon)| \leq (1-\epsilon^k)r + \frac{1}{2} \cdot \epsilon^k r < r\,,$$

im Widerspruch zur Minimalität von $|p(0)| = r$. □

Korollar 7.8. *Jedes nichtkonstante Polynom $f \in \mathbb{C}[X]$ zerfällt in Linearfaktoren, d.h. es gibt Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, wobei $n = \deg f$, und $a \in \mathbb{C}\backslash\{0\}$ mit*

$$f = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n)\,.$$

10.2.10

Korollar 7.9. *Jedes nichtkonstane Polynom $f \in \mathbb{R}[X]$ kann als Produkt*

$$f = a \cdot (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_r) \cdot g_1 \cdot \ldots \cdot g_m$$

geschrieben werden. Dabei ist $a \in \mathbb{R} \setminus \{0\}$, $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und $g_1, \ldots, g_m$ sind normierte reelle Poynome vom Grad 2 (normiert bedeutet, dass der Leitkoeffizient gleich 1 ist) ohne reelle Nullstellen.

Beweis. Indem wir f durch den Leitkoeffizienten teilen, können wir annehmen, dass f normiert ist.

Nun fassen wir f als Polynom mit komplexen Nullstellen auf und erhalten nach dem Fundamentalsatz der Algebra

$$f = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_k)\,,$$

wobei $k = \deg f$. Auf dem Polynomring $\mathbb{C}[X]$ betrachten wir den Ringhomomorphismus (!)

$$\overline{\cdot} : \mathbb{C}[X] \to \mathbb{C}[X]\,,$$

der durch Konjugation der Koeffizienten gegeben ist

$$a_0 + \ldots + a_n X^n \mapsto \overline{a_0 + \ldots + a_n X^n} := \overline{a_0} + \ldots + \overline{a_n} X^n\,.$$

Diese Abbildung ist übrigens auch $\mathbb{R}$-linear, jedoch nicht $\mathbb{C}$-linear.

Ist $p \in \mathbb{C}[X]$, so gilt offensichtlich $\overline{p} = p$ genau dann, wenn p nur reelle Koeffizienten hat, also $p \in \mathbb{R}[X]$.

Für unser gegebenes Polynom gilt also

$$(X - \lambda_1) \cdot \ldots \cdot (X - \lambda_k) = f = \overline{f} = (X - \overline{\lambda_1}) \cdot \ldots \cdot (X - \overline{\lambda_k})\,.$$

Da die Polynome links und rechts gleich sind, haben sie die gleichen Nullstellen. Also gilt folgendes: Ist $\lambda \in \mathbb{C}$ eine Nullstelle von f, so auch $\overline{\lambda}$. Für jede Nullstelle λ von f gibt es also zwei Möglichkeiten: Entweder ist $\lambda \in \mathbb{R}$, oder $\lambda \in \mathbb{C} \setminus \mathbb{R}$ und in diesem Fall ist auch $\overline{\lambda}$ eine Nullstelle, wobei $\overline{\lambda} \neq \lambda$. Bezeichnen wir die reellen Nullstellen von f mit $\lambda_1, \ldots, \lambda_r$ und die echt komplexen Nullstellen mit $\mu_1, \overline{\mu_1}, \ldots, \mu_m, \overline{\mu_m}$, so ist also

$$f = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_r)(X - \mu_1)(X - \overline{\mu_1}) \cdot \ldots \cdot (X - \mu_k)(X - \overline{\mu_k})$$

Jedes Polynom $g_i := (X - \mu_i)(X - \overline{\mu_i}) = X^2 - (\mu_i + \overline{\mu_i})X + \mu_i \overline{\mu_i}$ ist vom Grad 2, hat reelle Koeffizienten (dies sieht man durch Anwenden der Konjugation), aber keine reellen Nullstelen. □