Lineare Algebra und Analytische Geometrie 1
Wintersemester 2006/07

David J. Green
Mathematisches Institut
Friedrich-Schiller-Universität Jena

01. Februar 2007

Inhaltsverzeichnis

1 Mengen

Logik

Nur weil Mäuse Vierbeiner sind, heißt es noch lange nicht, dass jeder Vierbeiner eine Maus ist. In diesem kurzen Abschnitt geht es darum, welche Schlussfolgerungen zulässig sind. Außerdem wird einiges an Notation eingeführt.

Aussagen Aussagen sind entweder wahr oder falsch. „Der Dachs ist ein Säugetier“ ist eine wahre Aussage, dagegen ist die Aussage „Die Eins ist eine negative Zahl“ falsch.

Aussagen können verneint werden: aus „Ich mag Schockolade“ wird etwa „Ich mag Schockolade nicht“. Ist P eine Aussage, so bezeichnet man mit $\neg P$ die verneinte Aussage. Bei der Verneinung von komplexen Aussagen, wie etwa „Ist der Koch blau, dann ist die Suppe versalzen oder die Bohnen sind nicht gar“, muss man vorsichtig sein.

Man kann zwei Aussagen zu einer Aussage zusammenfassen, etwa „Der Eisverkäufer ist da, und ich habe gerade Taschengeld bekommen“, oder „Gibst du mir eins deiner Lollis, so bin ich dein bester Freund„. Die verschiedenen Wege, Aussagen zusammenzufassen, kann man gut mit *Wahrheitstafeln* voneinander unterscheiden. Vier Zusammenfassungsregeln sind:

a) Und $\wedge$ Ich ging einkaufen, und die Sonne schien.

b) Oder $\vee$ Entweder mähe ich den Rasen, oder ich jäte Unkraut – oder beides.

c) Implikation $\Rightarrow$: Wenn es morgen regnet, dann bleibe ich zu Hause – aber vielleicht tue ich das sowieso.

d) Äquivalenz $\Leftrightarrow$ Wenn du sofort dein Zimmer aufräumst, dann darfst du eine halbe Stunde fern gucken – sonst aber nicht.

Wahrheitstafeln:

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
F	F	W	F	F	W	W
F	W	W	F	W	W	F
W	F	F	F	W	F	F
W	W	F	W	W	W	W

Beachten Sie: $P \vee Q$ ist nur falsch, wenn P, Q beide falsch sind; und $P \Rightarrow Q$ ist nur falsch, wenn P wahr und Q falsch ist. Die Aussagen $(\neg P) \vee Q$ und $P \Rightarrow Q$ sind sogar äquivalent.

Will man zeigen, dass alle Vierbeiner Mäuse sind, so reicht es nicht aus, um nachzuweisen, dass alle Mäuse Vierbeiner sind. Dagegen ist die Aussage „Alle Vierbeiner sind Mäuse“ äquivalent zur (falschen) Aussage „Alles, was kein Maus ist, ist auch kein Vierbeiner„. In Symbolen ausgedrückt will ich sagen, dass die Folgerung $(P \Rightarrow Q) \Rightarrow (Q \Rightarrow P)$ nicht immer gilt, die Äquivalenz $(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$ dagegen schon. Diese beiden Behauptungen lassen mit Wahrheitstafeln nachweisen.

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$(P \Rightarrow Q) \Rightarrow (Q \Rightarrow P)$
F	F	W	W	W
F	W	W	F	**F**
W	F	F	W	W
W	W	W	W	W

P	Q	$\neg P$	$\neg Q$	$P \Rightarrow Q$	$\neg Q \Rightarrow \neg P$	$(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$
F	F	W	W	W	W	W
F	W	W	F	W	W	W
W	F	F	W	F	F	W
W	W	F	F	W	W	W

Aussageformen und Quantifikatoren Eine Aussageform ist eine Aussage mit einem oder mehreren freien Variablen. Es hängt vom Wert der Variable(n) ab, ob die Aussage wahr oder falsch ist. Schreibt man etwa $P(x)$ für die Aussageform „x ist eine reelle Zahl, die $x^2 - x = 0$ erfüllt“, so ist $P(0)$ eine wahre und $P(3)$ eine falsche Aussage.

Gerade machten wir die Aussageform $P(x)$ zu einer Aussage, indem wir den Wert $x = 3$ einsetzen. Man kann aber auch Quantifikatoren benutzen, um aus Aussageformen Aussagen zu machen. Wir benötigen drei Quantifikatoren:

a) $\forall$ „für alle” Beispiel: „$\forall x\colon x^2 \geq 0$“ ist die Aussage, dass jedes x die Ungleichung $x^2 \geq 0$ erfüllt.

b) $\exists$ „es gibt (mindestens) ein“ Beispiel: „$\exists x\colon x$ ist eine ganze Zahl und $x^2 = 2$“ ist die (falsche) Aussage, dass $\sqrt{2}$ eine ganze Zahl ist.

c) $\exists!$ „es gibt genau ein“ Beispiel: „$\exists! x\colon x$ ist eine ganze Zahl und $x^2 = y$“ ist eine Aussageform, die für $y = 0$ wahr ist, aber für $y = 3$ bzw. $y = 4$ falsch ist, denn dort gibt es keine bzw. zwei Möglichkeiten für x.

Mengen

Wer Mathematik-Vorlesungen besucht, wird mit sehr vielen Definitionen konfrontiert. Eine zufriedenstellende Definition des Mengenbegriffs ist aber nicht dabei, stattdessen wird erwartet, dass man ungefähr weiß, was eine Menge ist. Die vier wichtigsten Punkte:

a) Mengen wurden von Cantor eingeführt. Traditionell gibt man anstelle einer formalen Definition das folgende Zitat wieder:

> *Unter einer ‚Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objecten m unsrer Anschauung oder unseres Denkens (welche die ‚Elemente' von M genannt werden) zu einem Ganzen.* (G. Cantor, 1895)

b) Somit bilden die rationalen Zahlen eine Menge, und die Städte Thüringens mit mehr als 50.000 Einwohnern bilden eine andere.

> *Eine Menge ensteht durch Zusammenfassung von Einzeldingen zu einem Ganzen. Eine Menge ist eine Vielheit, als Einheit gedacht.* (F. Hausdorff, 1927)

c) Allerdings sind manche Zusammenfassungen, die man sich vorstellen kann, – die Zusammenfassung aller Mengen, zum Beispiel – so groß, dass man sie nicht als Mengen durchgehen lassen kann (Stichwort: Russellsche Paradoxon).

d) Aber dieser Problematik müssen wir uns – wenn überhaupt – erst dann stellen, wenn wir im Hauptstudium eine Axiomatische Mengenlehre hören. Denn die Methoden, die wir benutzen, um neue Mengen zu konstruieren, liefern als Ergebnis tatsächlich immer Mengen.

Mengen bestehen aus Elementen. Ist M eine Menge und x ein Element dieser Menge, so drückt man diese Tatsache durch die Bezeichung $x \in M$ aus. So ist zum Bespiel $1 \in \mathbb{Z}$, wobei $\mathbb{Z}$ die Menge aller ganzen Zahlen ist. Die Verneinung von $x \in M$ ist $x \notin M$. Es ist etwa $\frac{1}{2} \notin \mathbb{Z}$.

Eine Menge M heißt *endliche*, wenn sie nur endlich viele Elemente enthält. Sind $a_1, a_2, \ldots, a_n$ diese Elemente, so schreibt man $M = \{a_1, a_2, \ldots, a_n\}$. Die Anzahl der Elemente nennt man die *Mächtigkeit* $|M|$ der Menge. Beispiel: Die Menge $M = \{0, 3, 7, 15\}$ hat Mächtigkeit $|M| = 4$. Insbesondere im Fall $n = 0$ erhält man die *leere* Menge $\emptyset = \{\}$, d.h. die Menge, die *keine* Elemente enthält.

Einige wichtige Mengen:

$\mathbb{N}$	Die Menge $\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\}$ aller natürlichen Zahlen[1]
$\mathbb{N}_0$	Die Menge $\mathbb{N}_0 = \{0, 1, 2, 3, 4, 5, \ldots\}$ aller natürlichen Zahlen einschl. 0
$\mathbb{Z}$	Die Menge $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ aller ganzen Zahlen
$\mathbb{Q}$	Die Menge der rationalen Zahlen
$\mathbb{R}$	Die Menge der reellen Zahlen
$\mathbb{C}$	Die Menge der komplexen Zahlen (kommt noch)

Definition (Gleichheit von Mengen)
Zwei Mengen M, N heißen genau dann gleich, wenn gilt: jedes Element aus M ist gleichzeitig ein Element aus N, und auch umgekehrt. In Symbolen:

$$(M = N) \quad \Longleftrightarrow \quad (\forall x\colon (x \in M) \Leftrightarrow (x \in N))\,.$$

Beachten Sie hierzu:

- Die Mengen $\{1, 2, 3\}$ und $\{1, 2, 2, 3\}$ sind gleich, denn jedes Element der ersten Menge ist Element der zweiten, und auch umgekehrt. Beide Mengen haben auch genau drei Elemente.
- Dedekind stellte sich eine Menge vor wie einen geschlossenen Sack, der die Elemente der Menge enthält. Stelle ich mir aber einen roten und einen blauen Sack vor, die die gleichen Elemente enthalten, so stellen nach obiger Definition diese beiden Säcke die gleiche Menge dar: die farbliche Unterschied ist unerheblich.

Eine Menge N heißt eine *Teilmenge* der Menge M, wenn jedes Element von N auch ein Element aus M ist. Bezeichnung: $N \subseteq M$. In Symbolen:

$$(N \subseteq M) \quad \Longleftrightarrow \quad (\forall x\colon (x \in N) \Rightarrow (x \in M))\,.$$

Somit gilt:

Hilfssatz 1 *Es ist genau dann $M = N$, wenn sowohl $N \subseteq M$ als auch $M \subseteq N$ gelten.* ∎

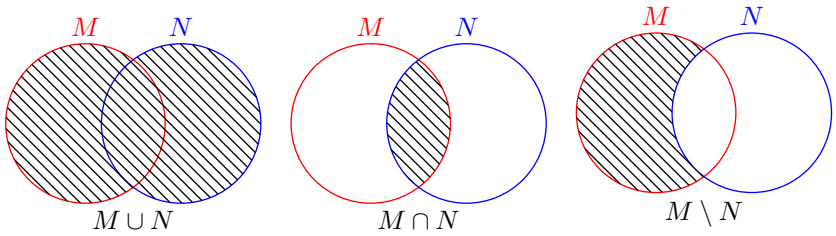
Es gibt einige Wege, um neue Mengen zu konstruieren. Ist M eine Menge und $P(x)$ eine Aussagenform in einer freien Variable, so bilden die Elemente x aus M, für die die Aussage $P(x)$ wahr ist, eine Teilmenge von M. Diese Teilmenge bezeichnet man so: $\{x \in M \mid P(x)\}$, manchmal auch so: $\{x \in M\colon P(x)\}$. Ein Beispiel: $\{x \in \mathbb{Z} \mid x^4 - 5x^2 + 4 = 0\} = \{-2, -1, 1, 2\}$.

Warnhinweis: Dagegen ist die Zusammenfassung $\{x \mid P(x)\}$ von *allen* x, die die Aussage $P(x)$ erfüllen, nicht immer eine Menge, wie B. Russell 1901 entdeckte[2]. Russells Beispiel: $L = \{x \mid x \text{ ist selbst eine Menge, und } x \notin x\}$. Wäre nämlich L eine Menge, so müsste entweder $L \in L$ oder $L \notin L$ gelten. Aber aus $L \in L$ folgt $L \notin L$, und auch umgekehrt. Ein Widerspruch liegt also vor, d.h. L kann keine Menge sein.

[2]Dieses Russellsche Paradoxon erschütterte die Anstrengungen des Jenaer Mathematikers und Philosophen Gottlob Frege, der als erster versuchte. der Mathematik eine axiomatische Grundlage zu geben. Frege wurde durch Ernst Abbe gefördert, aber von der mathematischen Welt ignoriert oder – sogar von Cantor – unfair kritisiert. Heutzutage gilt Frege als ein bedeutender Philosoph.

Sind M, N Mengen, so kann man folgende neue Mengen bilden: die *Vereinigung* $M \cup N$, die emphSchnitt $M \cup N$, und die *Differenzmenge* $M \setminus N$. Das geht so:

$$M \cup N = \{x \mid (x \in M) \vee (x \in N)\}$$
$$M \cap N = \{x \mid (x \in M) \wedge (x \in N)\}$$
$$M \setminus N = \{x \in M \mid x \notin N\}$$



Ist $M \cap N = \emptyset$, so heißen die Mengen M, N *disjunkt*.

Lemma 1.1 (Distributivgesetze) *Sind A, B, C Mengen, so gelten*

a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Sind A, B Teilmengen einer Menge X, so gelten

c) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$;

d) $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Beweis. Zuerst zeigen wir $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$: Ist $x \in A \cup (B \cap C)$, dann $x \in A$ oder $x \in B \cap C$. Ist $x \in A$, dann $x \in A \cup B$ und $x \in A \cup C$, weshalb $x \in (A \cup B) \cap (A \cup C)$. Ist dagegen $x \in B \cap C$, dann ist $x \in B$ und $x \in C$. Also ist $x \in A \cup B$ wegen $x \in B$, und $x \in A \cup C$ wegen $x \in C$. Also $x \in (A \cup B) \cap (A \cup C)$.

Jetzt zeigen wir $(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)$. Ist $x \in (A \cup B) \cap (A \cup C)$ dann $x \in A \cup B$ und $x \in A \cup C$. Ist $x \in A$, dann $x \in A \cup (B \cap C)$. Ist $x \notin A$, so gilt $x \in B$ wegen $x \in A \cup B$, und $x \in C$ gilt wegen $x \in A \cup C$. Aslo $x \in B \cap C$, weshalb $x \in A \cup (B \cap C)$.

Somit gilt $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. Der Beweis der zweiten Aussage verläuft ähnlich. ∎

Die Potenzmenge $\mathscr{P}(M)$ einer Menge M ist per Definition die Menge aller Teilmengen von M.

$$\mathscr{P}(M) = \{N \colon N \subseteq M\}.$$

Lemma 1.2 *Ist M eine endliche Menge, so ist $|\mathscr{P}(M)| = 2^{|M|}$.*

Beweis. Wir führen einen Induktionsbeweis über $|M|$. Das heißt, wir beweisen

- Die Aussage gilt im Fall $|M| = 0$ (*Induktionsanfang*); und
- Für jede ganze Zahl $n \geq 0$ gilt: stimmt die Aussage im Fall $|M| = n$, so stimmt sie auch im Fall $|M| = n + 1$ (*Induktionsschritt*).

Das *Prinzip der mathematischen Induktion* besagt dann, dass das Lemma bewiesen ist.

Induktionsanfang: Ist $|M| = 0$, so hat M keine Elemente. Deshalb gilt übrigens $M = \emptyset$. Ist $N \subseteq M$ eine Teilmenge, so hat auch N keine Elemente – denn solche müssten auch Elemente von M sein –, weshalb $N = M$ gilt. Andererseits gilt immer $M \subseteq M$. Somit gilt: Ist $|M| = 0$, so hat M genau eine Teilmenge, nämlich M selbst. Das heißt, $\mathscr{P}(M) = \{M\}$ hat genau ein Element.

Induktionsschritt: Sei $n \geq 0$ eine ganze Zahl. Wir setzen voraus, dass für n-elementige Mengen die Aussage gilt. Sei M eine Menge mit $n+1$ Elementen. Sei b ein Element aus M, und sei N das n-elementige Menge $M \setminus \{b\}$. Sei $T \subseteq M$ eine Teilmenge. Es gibt zwei Fälle:

a) Ist $b \notin T$, dann $T \subseteq N$. Umgekehrt ist jedes $S \subseteq N$ ein solches T. Da es nach Annahme 2^n Teilmengen $S \subseteq N$ gibt, gibt es genau 2^n Teilmengen $T \subseteq M$ dieser ersten Art.

b) Ist $b \in T$, so ist die Menge $T' = T \setminus \{b\}$ eine Teilmenge von N, und es ist $T = T' \cup \{b\}$. Umgekehrt ist $S \cup \{b\}$ ein solches T für jedes $S \subseteq N$, und das T' zu diesem T ist dann S. Somit stimmt die Anzahl solcher Teilmengen $T \subseteq M$ mit der Anzahl der Teilmengen $S \subseteq N$ überein, d.h. es gibt 2^n Teilmengen dieser zweiten Art.

Somit gibt es ingesamt $2^n + 2^n = 2^{n+1} = 2^{|M|}$ Teilmengen von M. Der Induktionsschritt ist also bewiesen, das Lemma auch. ∎

Definition Sind A, B zwei Mengen, so definiert man das *direkte Produkt* $A \times B$ als die Menge aller *geordnete Paare* (a, b) mit $a \in A$ und $b \in B$. Das Wort „geordnet“ bedeutet, dass z.B. $(1, 2) \neq (2, 1)$.

Allgemeiner definiert man für Mengen $A_1, A_2, \ldots, A_n$ das direkte Produkt

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in A_i \text{ für alle } i\}.$$

Die Elemente dieses Produkts nennt man (geordnete) n-Tupel[3].

[3]Wenn man später die Mengenlehre axiomatisieren will, stellt man sich die Frage, was für ein Objekt ein geordnetes Paar sein soll. Die heute allgemein akzeptierte Antwort lautet: das geordnete Paar (a, b) ist die Menge $\{\{a\}, \{a, b, \}\}$. Beachten Sie, dass diese Menge im Fall $a = b$ aus nur einem Element besteht.

Abbildungen

Definition Seien X, Y Mengen. Eine *Abbildung* $f\colon X \to Y$ von X nach Y ist eine Vorschrift, die zu jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet. Man spricht auch von der Abbildung $x \mapsto f(x)$.

Die Menge aller Abbildungen von X nach Y wird mit Y^X oder $\mathrm{Abb}(X, Y)$ bezeichnet.

Bemerkung Eine äquivalente Definition lautet so: eine Abbildung $f\colon X \to Y$ ist eine Teilmenge $F \subseteq X \times Y$ mit der folgenden Eigenschaft: zu jedem $x \in X$ gibt es genau ein Element $y \in Y$ derart, dass das Paar (x, y) in F liegt. Es ist dann $y = f(x)$.

Beispiel Die Vorschrift „$f(x) =$ das y mit $y^2 = x$“ definiert aus zwei Gründen keine Abbildung $f\colon \mathbb{R} \to \mathbb{R}$. Erstens wird zu negativen Zahlen wie z.B. -1 keinen Wert $f(x)$ zugeordnet; und zweitens wird zu positiven Zahlen mehr als einen Wert zugeordnet, z.B. $f(1) = 1$ und $f(1) = -1$.

Definition Sind $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen, so wird deren *Verknüpfung* $g \circ f\colon X \to Z$ so definiert: für jedes $x \in X$ ist $(g \circ f)(x) = g(f(x))$.

Hilfssatz 2 *Verknüpfung von Abbildungen ist assoziativ, d.h. sind $f\colon X \to Y$, $g\colon Y \to Z$ und $h\colon Z \to W$ Abbildungen, so sind die Abbildungen $h \circ (g \circ f)$ und $(h \circ g) \circ f\colon X \to W$ gleich.*

Beweis. In jedem $x \in X$ nehmen beide Abbildungen den Wert $h(g(f(x)))$. ∎

Definition Sei $f\colon X \to Y$ eine Abbildung.

a) Sei $A \subseteq X$ eine Teilmenge. Die eingeschränkte Abbildung $f|_A\colon A \to Y$ wird definiert durch $f|_A(x) = f(x)$ für jedes $x \in A$.

b) Ist $A \subseteq X$ eine Teilmenge, so bezeichnet man mit $f(A)$ die Bildmenge $f(A) = \{f(x) \mid x \in A\}$ von A. Insbesondere nennt man $f(X)$ das *Bild*[4] von f, es ist also $\mathrm{Bild}(f) = \{f(x) \mid x \in X\}$.

c) Ist $B \subseteq Y$, so definiert man die *Urbildmenge* $f^{-1}(B) \subseteq X$ durch

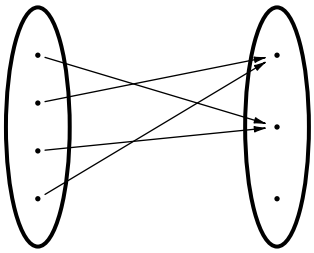
$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}\,.$$

Ist B die Teilmenge $\{b\}$, die aus einem Element $b \in Y$ besteht, so schreibt man $f^{-1}(b)$ statt $f^{-1}(\{b\})$. Beachten Sie aber, dass $f^{-1}(b)$ kein Element von X ist, sondern eine Teilmenge von X.

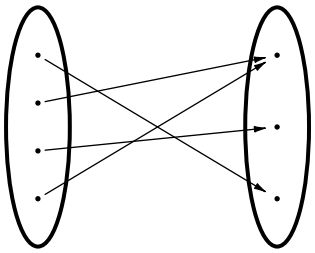
[4] Auf Englisch *Image*

Beispiel Sei $f\colon \mathbb{R} \to \mathbb{R}$ die Abbildung $x \mapsto x^2$. Dann $f^{-1}(\{-2, 0, 9\}) = \{-3, 0, 3\}$, $f^{-1}(-1) = \emptyset$ und $f^{-1}(4) = \{-2, 2\}$.

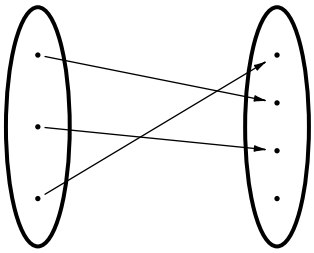
Definition Eine Abbildung $f\colon X \to Y$ heißt *injektiv*, wenn keine zwei Elemente von X das gleiche Bild in Y haben, d.h. wenn jedes $y \in Y$ höchstens ein Urbild in X hat.



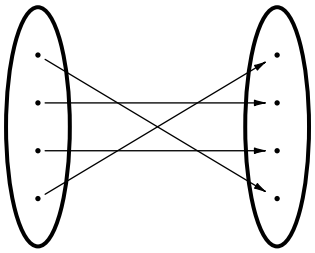
Weder injektiv noch surjektiv



Surjektiv aber nicht injektiv



Injektiv aber nicht surjektiv



Bijektiv

Lemma 1.3 *Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen.*

a) Ist die Verknüpfung $g \circ f$ injektiv, so muss f injektiv sein.

b) Ist die Verknüpfung $g \circ f$ surjektiv, so muss g surjektiv sein.

Beweis. S. Übungsblatt. ∎

Sei X eine Menge. Die Identitätsabbildung $\mathrm{Id}_X\colon X \to X$ wird so definiert: für jedes $x \in X$ ist $\mathrm{Id}_X(x) = x$.

Lemma 1.4 *Sei $f\colon X \to Y$ eine Abbildung.*

a) Ist f injektiv und X nicht leer, so gibt es eine – nach Lemma 1.3 zwangsweise surjektive – Abbildung $g\colon Y \to X$ mit $g \circ f = \mathrm{Id}_X$.

b) Ist f surjektiv, so gibt es eine – nach Lemma 1.3 zwangsweise injektive – Abbildung $g\colon Y \to X$ mit $f \circ g = \mathrm{Id}_Y$.

c) *f ist genau dann bijektiv, wenn es eine Umkehrabbildung $f^{-1}\colon Y \to X$ gibt mit den beiden Eigenschaften $f \circ f^{-1} = \mathrm{Id}_Y$ und $f^{-1} \circ f = \mathrm{Id}_X$.*

Außerdem gilt: ist f bijektiv, so wird f^{-1} durch jede dieser beiden Eigenschaften charakterisiert[5]*.*

Beweis. a) Da X nicht leer ist, dürfen wir ein Element $x_0 \in X$ wählen. Sei $y \in Y$. Ist $y \in \mathrm{Bild}(f)$, so gibt es wegen Injektivität genau ein $x \in X$ mit $f(x) = y$, und wir setzen $g(y) =$ dieses x. Ist $y \notin \mathrm{Bild}(f)$, so können wir $g(y) = x_0$ setzen. Somit ist g überall definiert, und es ist $g \circ f = \mathrm{Id}_X$.

b) Für jedes $y \in Y$ ist die Urbildmenge $f^{-1}(y) = \{x \in X \mid f(x) = y\}$ nicht leer, weshalb wir zu jedem $y \in Y$ ein Element $x_y \in f^{-1}(y)$ wählen dürfen[6]. Durch $g\colon y \mapsto x_y$ ist dann eine Abbildung $g\colon Y \to X$ definiert, und es ist $(g \circ f)(y) = g(f(y)) = g(x_y) = y$.

c) Da f bijektiv ist, gibt es zu jedem $y \in Y$ genau ein Element $x_y \in X$ mit $f(x_y) = y$. Definieren wir $g\colon Y \to X$ durch $g(y) = x_y$, so gilt $f \circ g = \mathrm{Id}_Y$ sofort, denn $f \circ g(y) = f(x_y) = y$. Sei jetzt $x \in X$ und $z = g(f(x))$. Es ist $f(z) = f(g(f(x))) = (f \circ g)(f(x)) = f(x)$, denn $f \circ g = \mathrm{Id}_Y$. Also $z = x$, denn f ist injektiv. Das heißt, $g \circ f = \mathrm{Id}_X$. Somit gilt $f^{-1} = g$.

Man sieht, dass Identitätsabbildungen injektiv und surjektiv sind. Somit ist $g \circ f$ injektiv, weshalb nach Teil 1) auch f injektiv ist. Außerdem ist $f \circ g$ und deshalb auch f surjektiv.

Eindeutigkeit von f^{-1}: Sei $h\colon Y \to X$ eine Abbildung mit $h \circ f = \mathrm{Id}_X$. Da f surjektiv ist, folgt $h = g$: denn ist $y \in Y$, so gibt es ein $x \in X$ mit $f(x) = y$, also $h(y) = h(f(x)) = x = g(f(x)) = g(y)$. Analog folgt $h = g$ aus $f \circ h = \mathrm{Id}_Y$, denn f ist injektiv. ∎

Familien Man könnte die Geburtstage der Kinder aus der Klasse 1b als eine Menge auffassen. Dies hat aber zwei Nachteile. Erstens erfährt man zwar, dass jemand am 13.8. Geburtstag hat, aber nicht, dass es sich hier um Julie handelt. Zweitens kann zwar die Anzahl der Kinder mit der Anzahl der Geburstage vergleichen, und so feststellen, dass zwei Kinder den gleichen Geburtstag haben muss: aber ob dieser doppelte Geburtstag der 27.3. oder der 4.7. ist, kann man nicht erkennen.

[5] Das heißt, f^{-1} ist die einzige Abbildung, die die eine *oder* die andere Eigenschaft hat.

[6] Damit diese Vorgehensweise – für *jedes* $y \in Y$ die Wahl eines *beliebigen* Urbildelements – auch nach Axiomatisierung der Mengenlehre möglich ist, bedarf es ein eigens hierfür konzipiertes Axiom, das Auswahlaxiom. Dieses Axiom braucht man, um je eine Socke aus unendlich viele Paare auszuwählen, aber nicht, um je einen Schuh aus unendlich viele Paare auszuwählen: denn bei Schuhen kann man z.B. immer den linken wählen, dagegen sind beide Socken eines Paares identisch.

Am besten fasst man die Geburtstage als eine *Familie* auf. Zu einer Familie gehört eine *Indexmenge*, I; in diesem Fall ist I die Menge der Kinder aus der 1b. Eine Familie von Elementen einer Menge M mit Indexmenge I ist eigentlich eine Abbildung $f\colon I \to M$, die man aber als $(m_i)_{i\in I}$ schreibt, wobei $m_i = f(i)$ ist. Zum Beispiel $\text{Geburtstag}_{\text{Julie}} = 13.8$.

Es gibt auch Familien von Mengen, eine typische Schreibweise wäre $(A_i)_{i\in I}$. Zum Beispiel könnte A_i die Menge der AGs sein, die Kind i besucht: etwa

$$A_{\text{Nils}} = \{\text{Badminton}, \text{Hockey}, \text{Schach}\}\,.$$

Liegt eine Familie von Mengen vor, so kann man die Vereinigung und den Schnitt aller Mengen in der Familie bilden:

$$\bigcup\{A_i \mid i \in I\} = \bigcup_{i\in I} A_i = \{x \mid \exists i \in I\colon x \in A_i\}$$
$$\bigcap\{A_i \mid i \in I\} = \bigcap_{i\in I} A_i = \{x \mid \forall i \in I\colon x \in A_i\}$$

Im obigen Beispiel ist $\bigcup\{A_i \mid i \in I\}$ die Menge aller AGs, die von irgendeinem Kind aus der 1b besucht werden, und $\bigcap_{i\in I} A_i$ ist die (vermutlich leere) Menge aller AGs, die jedes Kind aus der 1b besucht.

Äquivalenz- und Ordnungsrelationen

Wir führen Relationen ein. Die wichtigsten Relationen sind Äquivalenzrelationen (etwa „Peter und Sabine haben den gleichen Geburtstag“) sowie Ordnungsrelationen (etwa „Jena hat weniger Einwohner als Erfurt“).

Definition Sei X eine Menge. Eine (binäre) *Relation* auf X ist eine Teilmenge $R \subseteq X\times X$. Sind $x, y \in X$, so schreibt man meistens $x\ R\ y$ anstelle von $(x, y) \in R$. Somit ist $x\ R\ y$ eine Aussage, und deshalb entweder wahr oder falsch.

Eine Relation heißt

- *reflexiv*, falls $x\ R\ x$ gilt für jedes $x \in X$.
- *symmetrisch*, falls $x\ R\ y$ genau dann gilt, wenn $y\ R\ x$ gilt.
- *antisymmetrisch*, falls $x = y$ gelten muss, wenn $x\ R\ y$ und $y\ R\ x$ gleichzeitig gelten.
- *transitiv*, falls $x\ R\ z$ aus $x\ R\ y$ und $y\ R\ z$ folgt.
- *linear*, falls mindestens eins aus $x\ R\ y$ und $y\ R\ x$ gelten muss.

Eine *Äquivalenzrelation* ist eine Relation, die reflexiv, symmetrisch und transitiv ist. Ist $\sim$ eine Äquivalenzrelation auf der Menge X, so nennt man die Menge $[x] = [x]_\sim = \{y \in X \mid x \sim y\}$ die *Äquivalenzklasse* eines Elements $x \in X$.

Eine *Ordnung* ist eine Relation, die reflexiv, transitiv, linear und antisymmetrisch ist. Verlangt man Linearität nicht, so erhält man eine *Teilordnung*. Will man die Linearität unterstreichen, so spricht man von einer *linearen* Ordnung.

Lemma 1.5 *Sei $\sim$ eine Äquivalenzrelation auf der Menge X. Dann ist $x \in [x]$ für jedes $x \in X$, wegen Reflexivität. Außerdem sind für $x, y \in X$ folgende drei Aussagen äquivalent:*

a) $x \sim y$

b) $[x] \cap [y] \neq \emptyset$.

c) $[x] = [y]$

Beweis. Wir zeigen 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 1). Ist $x \sim y$, dann $y \in [x]$. Wegen Reflexivität ist aber $y \in [y]$. Also $y \in [x] \cap [y]$.

Jetzt nehmen wir an, es gibt ein $z \in [x] \cap [y]$. Ist $w \in [x]$, dann $x \sim w$. Wegen $z \in [x]$ ist $x \sim z$, also $z \sim x$ wegen Symmetrie. Aus $z \sim x$ und $x \sim w$ folgt $z \sim w$ wegen Transitivität. Aufgrund von $z \in [y]$ gilt $y \sim z$, also $y \sim w$ wegen Transitivität, also $w \in [y]$. Wir haben gezeigt, dass $[x] \subseteq [y]$ aus $[x] \cap [y] \neq \emptyset$ folgt. Analog folgt $[y] \subseteq [x]$, also $[y] = [x]$.

Ist dagegen $[x] = [y]$, dann $y \in [y] = [x]$, weshalb $x \sim y$. ∎

Bemerkung Somit stellen die Äquivalenzklassen von $\sim$ eine *Partition* der Menge X dar, d.h. X ist die Vereinigung von disjunkten, nichtleeren Äquivalenzklassen.

Definition Ist $\sim$ eine Äquivalenzrelation auf der Menge X, so bezeichnet man mit $X/\sim$ die Menge der Äquivalenzklassen: $X/\sim = \{[x] \mid x \in X\}$.

Lemma 1.6 *Sei $\sim$ eine Äquivalenzrelation auf der Menge X. Sei $f\colon X \to Y$ eine Abbildung mit folgender Eigenschaft: für alle $x, x' \in X$ mit $x \sim x'$ ist $f(x) = f(x')$. Dann wird durch $\bar{f}([x]) = f(x)$ eine wohldefinierte Abbildung $\bar{f}\colon X/\sim \to Y$ erklärt. So erhält man eine bijektive Korrespondenz*

$$\begin{aligned} \{f \in \mathrm{Abb}(X,Y) \mid \forall x, x' \in X\colon x \sim x' \Rightarrow f(x) = f(x')\} &\longrightarrow \mathrm{Abb}(X/\sim, Y) \\ f &\longmapsto \bar{f} \end{aligned}$$

Beweis. Für Wohldefiniertheit müssen wir zeigen, dass $\bar{f}$ repräsentantenunabhängig ist. Ist $[x] = [x']$, dann $x \sim x'$, also $f(x) = f(x')$, weshalb $\bar{f}([x]) = \bar{f}([x'])$ gilt, wie erwünscht.

Die bijektive Korrespondenz: Ist $g \in \mathrm{Abb}(X/\sim, Y)$, so definieren wir $\hat{g}\colon X \to Y$ durch $\hat{g}(x) = g([x])$. Ist $x \sim x'$, dann $\hat{g}(x) = g([x]) = g([x']) = \hat{g}(x')$. Man rechnet jetzt nach, dass $\bar{\hat{g}} = g$ und $\hat{\bar{f}} = f$ gelten. Somit ist die Abbildung $f \mapsto \bar{f}$ eine Bijektion, denn sie hat in $g \mapsto \hat{g}$ ein beidseitiges Inverses. ∎

2 Gruppen und Körper

Definition Eine Gruppe $G = (G, *)$ besteht aus einer Menge G und einer Operation $*\colon G \times G \to G$ mit den folgenden drei Eigenschaften:

a) Assoziativität: $x * (y * z) = (x * y) * z$ für alle $x, y, z \in G$.

b) Existenz eines neutralen Elements: Es gibt ein $e \in G$, so dass $x*e = e*x = x$ gilt für jedes $x \in G$.

c) Existenz von Inversen: Zu jedem $x \in G$ gibt es ein $x' \in G$ mit der Eigenschaft, dass $x * x' = x' * x = e$ gilt.

Eine Gruppe G heißt *abelsch*, falls $x * y = y * x$ gilt für alle $x, y \in G$.

Beispiele a) $G = \mathbb{Z}$ mit $x * y = x + y$. Es ist $e = 0$, $x' = -x$.

b) $G = \mathbb{R}^\times = \{x \in \mathbb{R} \mid x \neq 0\}$ mit $x * y = xy$ (Multiplikation). Es ist $e = 1$, $x' = \frac{1}{x}$.

c) Die Menge $G = \{+1, -1\}$, wieder mit Multiplikation. Es ist $e = 1$, $x' = \frac{1}{x} = x$. Diese ersten drei Beispiele sind abelsch.

d) Die symmetrische Gruppe S_n oder $\mathrm{Sym}(\Omega)$, s. unten. Bereits S_3 ist nichtabelsch.

e) Die Gruppe $GL_2(\mathbb{R})$ aller invertierbaren 2×2-Matrizen mit Einträgen aus $\mathbb{R}$, bezüglich Matrixmultipliktation. Nichtabelsch.

Lemma 2.1 *Sei G eine Gruppe.*

a) Es gibt in G genau ein neutrales Element e.

b) Zu jedem $x \in G$ gibt es genau ein Inverses $x' \in G$.

c) Es ist $(x * y)' = y' * x'$ für alle $x, y \in G$.

d) Für jedes $x \in G$ gilt $(x')' = x$.

e) Es ist $e' = e$.

Beweis. a) Sind e_1, e_2 zwei neutrale Elemente, so muss $e_1 * e_2 = e_1$ gelten (da e_2 neutral) und gleichzeitig $e_1 * e_2 = e_2$ (da e_1 neutral). Also $e_1 = e_2$.

b) Sind a, b zwei Inversen von x, so ist $a * x = x * b = e$. Also

$$a = a * e = a * (x * b) = (a * x) * b = e * b = b\,.$$

c) Wegen Assoziativität gilt

$(x*y)*(y'*x') = ((x*y)*y')*x' = (x*(y*y'))*x' = (x*e)*x' = x*x' = e\,.$

Analog gilt $(y'*x')*(x*y) = e$. Wegen Teil b) folgt $(x*y)' = y'*x'$.

d) Folgt aus $x'*x = x*x' = e$ wegen Eindeutigkeit von $(x')'$.

e) Folgt bereits aus $e*e = e$. ∎

Permutationen und die symmetrische Gruppe Für eine beliebige Menge Ω erklärt man die *symmetrische Gruppe* $\mathrm{Sym}(\Omega)$ als die Menge aller Bijektionen von Ω nach sich selbst.

$$\mathrm{Sym}(\Omega) = \{f\colon \Omega \to \Omega \mid f \text{ ist eine Bijektion}\}\,.$$

Dies ist eine Gruppe bezüglich Verküpfung: $f*g = f\circ g$. Die Identitätsabbildung $\mathrm{Id}\colon \Omega \to \Omega$ ist das neutrale Element, und $f' = f^{-1}$. Die Elemente von $\mathrm{Sym}(\Omega)$ nennt man *Permutationen* von Ω.

Häufig beschäftigt man sich mit dem Fall $\Omega = \{1, 2, 3, \ldots, n\}$. In diesem Fall schreibt man einfach S_n für die symmetrische Gruppe $\mathrm{Sym}(\Omega)$. Es gibt verschiedene Wege, um Permutationen $\pi \in S_n$ hinzuschreiben. Ein Weg ist

$$\pi = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n) \end{pmatrix}.$$

So bezeichnet zum Beispiel $\left(\begin{smallmatrix}1&2&3\\3&2&1\end{smallmatrix}\right)$ die Permutation $\sigma\colon \{1,2,3\} \to \{1,2,3\}$ mit $\sigma(1) = 3$, $\sigma(2) = 2$ und $\sigma(3) = 1$.

Sei $\tau = \left(\begin{smallmatrix}1&2&3\\2&1&3\end{smallmatrix}\right)$, ein weiteres Element von S_3. Es ist

$$\sigma\tau = \sigma\circ\tau = \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} \qquad \tau\sigma = \tau\circ\sigma = \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},$$

also $\sigma\tau \neq \tau\sigma$: somit ist S_3 nichtabelsch.

Gruppenhomomorphismen

Definition Seien G, H Gruppen. Eine Abbildung $f\colon G \to H$ heißt ein *Homomorphismus*, falls $f(x*y) = f(x)*f(y)$ gilt für alle $x, y \in G$.

Bemerkung Auf dem Übungsblatt werden Sie zeigen: ist f ein Homomorphismus, so gelten $f(e_G) = e_H$ und $f(x') = [f(x)]'$.

Beispiele a) $f\colon \mathbb{Z} \to \mathbb{Z}$, $f(n) = 2n$.

b) Das Vorzeichen einer Permutation $\varepsilon\colon S_n \to \{+1, -1\}$. *Kommt noch, im Kapitel über Determinanten.*

c) $f\colon \mathbb{R} \to \mathbb{C}^\times$, $f(x) = \exp(2\pi i x)$ ist ein Homomorphismus: $f(x+y) = f(x)f(y)$.

Ringe

Definition Ein *Ring* $R = (R, +, \cdot)$ besteht aus einer abelschen Gruppe $(R, +)$ mit neutralem Element 0 zusammen mit einer Abbildung $\cdot\colon R \times R \to R$, $(x, y) \to x \cdot y = xy$, die die folgenden Bedingungen erfüllt:

a) $\cdot$ ist assoziativ: $(xy)z = x(yz)$.

b) $\cdot$ hat ein neutrales Element 1.

c) Distributivgesetze: $x(y + z) = xy + xz$ und $(x + y)z = xz + yz$ für alle $x, y, z \in R$.

Der Ring heißt *kommutativ*, falls $xy = yx$ gilt für alle $x, y \in R$.

Beispiele a) $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ sind Ringe bezüglich der üblichen Addition und Multiplikation.

b) Ist R ein Ring und $n \geq 1$, so ist die Menge $M_n(R)$ aller $n \times n$-Matrizen über R ein Ring bezüglich Matrixaddition und -multiplikation (s. unten). $M_2(\mathbb{R})$ ist nicht kommutativ, denn $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, aber $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

c) Der Nullring $R = \{0\}$ mit $0 + 0 = 0 \cdot 0 = 0$ ist ein Ring, mit $1_R = 0_R = 0$.

d) Die Menge $\mathbb{R}[X]$ aller Polynome mit Koeffizienten aus $\mathbb{R}$ ist ein Ring, denn man kann Polynome addieren und multiplizieren.

Lemma 2.2 *Sei R ein Ring. Für alle $x, y \in R$ gelten dann:*

a) $0 \cdot x = x \cdot 0 = 0$.

b) $x \cdot (-y) = (-x) \cdot y = -(xy)$.

c) $(-1)^2 = 1$.

Beweis. a) Wegen $0 + 0 = 0$ gilt $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$. Zieht man $0 \cdot x$ von beiden Seiten ab, so folgt $0 \cdot x = 0$. Analog zeigt man $x \cdot 0 = 0$.

b) Für $x \cdot (-y) = -(xy)$ reicht es zu zeigen: $xy + x \cdot (-y) = 0$. Aber $xy + x(-y) = x(y + (-y)) = x \cdot 0 = 0$. Analog zeigt man $(-x)y = -(xy)$.

c) Wegen b) gilt $(-1)^2 = -(1 \cdot (-1)) = -(-(1 \cdot 1)) = -(-1) = 1$. ∎

Matrizen

Hier ist eine (3×2)-Matrix mit Einträgen aus $\mathbb{R}$: $\begin{pmatrix} 1 & 2 \\ 4 & -7 \\ 0,5 & \pi \end{pmatrix}$.

Definition Sei R ein Ring und n, m ganze Zahlen ≥ 1. Eine $(m \times n)$-Matrix mit Einträgen aus R besteht aus mn Elemente von R, aufgestellt in m Zeilen und n Spalten.

Ist A eine solche Matrix, so bezeichnet man mit A_{ij} der Eintrag an der Stelle (i, j), d.h. in der $i$ten Zeile und der $j$ten Spalte.

Beispiel Für $A = \begin{pmatrix} 1 & 3 & 7 & 4 \\ 3 & 1 & 2 & 9 \\ 8 & 0 & 7 & 3 \end{pmatrix}$ ist $A_{23} = 2$.

Bezeichnung Die Menge der $(m \times n)$-Matrizen mit Einträgen aus R werden wir mit $M(m \times n, R)$ bezeichnen. Besonders wichtig ist der Fall $m = n$, man schreibt $M_n(R)$ für $M(n \times n, R)$. Eine Matrix heißt *quadratisch*, wenn die Anzahl der Zeilen und der Spalten gleich sind.

Matrixaddition und -multiplikation Für Matrizen $A, B \in M(m \times n, R)$ wird die Summe $A + B \in M(m \times n, R)$ definiert durch $(A + B)_{ij} = A_{ij} + B_{ij}$ für alle i, j.

Beispiel $\begin{pmatrix} 1 & 2 & 4 \\ 2 & 5 & 7 \end{pmatrix} + \begin{pmatrix} 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 5 \\ 3 & 6 & 8 \end{pmatrix}$.

Sind $A \in M(m \times n, R)$ und $B \in M(n \times p, R)$ Matrizen, so wird das Produkt $AB \in M(m \times p, R)$ definiert durch

$$(AB)_{ik} = A_{i1}B_{1k} + A_{i2}B_{2k} + A_{i3}B_{3k} + \cdots + A_{in}B_{nk} =: \sum_{j=1}^{n} A_{ij}B_{jk} \, .$$

Beispiel

$$\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 0 \cdot 1 & 1 \cdot 2 + 0 \cdot 3 & 1 \cdot 4 + 0 \cdot 5 \\ 3 \cdot 0 + 1 \cdot 1 & 3 \cdot 2 + 1 \cdot 3 & 3 \cdot 4 + 1 \cdot 5 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 2 & 4 \\ 1 & 9 & 17 \end{pmatrix} .$$

Bemerkung Beachten Sie: das Produkt AB ist nur dann definiert, wenn A die gleiche Anzahl von Spalten hat, wie B Zeilen hat.

Lemma 2.3 *Matrixmultiplikation ist assoziativ.*

Beweis. Sei $A \in M(m \times n, R)$, $B \in M(n \times p, R)$ und $C \in M(p \times q, R)$. Für $1 \leq i \leq m$ und $1 \leq \ell \leq q$ ist

$$\begin{aligned}[(AB)C]_{i\ell} &= \sum_{k=1}^{p} (AB)_{ik} C_{k\ell} = \sum_{k=1}^{p} \left(\sum_{j=1}^{n} A_{ij} B_{jk} \right) C_{k\ell} = \sum_{j=1}^{n} \sum_{k=1}^{p} A_{ij} B_{jk} C_{k\ell} \\ &= \sum_{j=1}^{n} A_{ij} \left(\sum_{k=1}^{p} B_{jk} C_{k\ell} \right) = \sum_{j=1}^{n} A_{ij} (BC)_{j\ell} = [A(BC)]_{i\ell} \,. \end{aligned}$$

■

Auf ähnlicher Weise zeigt man, dass Matrixaddition und -multiplikation die Distributivgesetze erfüllen. Dann kann man nachweisen, dass $M_n(R)$ tatsächlich ein Ring ist. Das Einselement dieses Rings ist die *Einheitsmatrix* E_n, gegeben durch

$$(E_n)_{ij} = \delta_{ij} \,, \quad \text{wobei} \quad \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} .$$

Körper

Definition Ein *Körper* k ist ein kommutativer Ring, die zwei weiteren Begindungen erfüllt:

a) In k gilt $1 \neq 0$.

b) Die Menge $k^\times := k \setminus \{0\}$ ist eine Gruppe[7] bezüglich Multiplikation, d.h. zu jedes $0 \neq x \in k$ gibt es ein $x^{-1} \in k$ mit $xx^{-1} = 1$.

In einem Körper gelten genau die Regeln, die man von Skalaren erwartet.

Beispiele a) $\mathbb{R}$ und $\mathbb{Q}$ sind Körper.

b) Die komplexen Zahlen bilden einen Körper $\mathbb{C}$ (s. unten).

c) $\mathbb{Z}$ ist kein Körper, denn 2 liegt in $\mathbb{Z}$ aber $\frac{1}{2}$ nicht.

d) Der Nullring $\{0\}$ ist kein Körper, denn in diesem Fall gilt $1 = 0$.

e) $M_2(\mathbb{R})$ ist kein Körper, denn Matrixmultiplikation ist nicht kommutativ.

f) Die Menge $\{0, 1\}$ bildet einen Körper, genannt $\mathbb{F}_2$, mit Addition und Multiplikation gegeben durch

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

[7] Notwendigerweise abelsch, da der Ring kommutativ ist.

g) Die Menge $\{0,1,2\}$ bildet einen Körper, genannt $\mathbb{F}_3$, mit Addition und Multiplikation gegeben durch

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

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

Der Körper $\mathbb{C}$ der komplexen Zahlen

Definition Der Körper $\mathbb{C}$ der komplexen Zahlen besteht aus der Menge $\mathbb{R}^2$, zusammen mit der üblichen komponentenweise Addition und mit der folgenden Multiplikationsregel:

$$(a,b)(c,d) := (ac - bd, ad + bc)\,. \qquad (*)$$

Ist $z = (a,b) \in \mathbb{C}$, so nennt man a bzw. b den reellen Teil $\Re(z)$ bzw. den imaginären Teil $\Im(z)$ von z. Durch $|z| := \sqrt{a^2+b^2}$ wird der Betrag von z definiert. Man schreibt $i = (0,1) \in \mathbb{C}$ und identifiziert $a \in \mathbb{R}$ mit $(a,0) \in \mathbb{C}$. Dann kann man $a + bi$ für (a,b) schreiben. Insbesondere gilt $i^2 = -1$.

Für $z = (a,b)$ wird die komplex konjugierte komplexe Zahl $\bar{z}$ durch $\bar{z} = (a,-b)$ definiert. Das heißt, für $a,b \in \mathbb{R}$ ist $\overline{a+bi} = a - bi$.

Es gilt jetzt, zu prüfen, dass $\mathbb{C}$ tatsächlich ein Körper ist. Wir wissen bereits, dass $(\mathbb{R}^2,+)$ eine abelsche Gruppe ist. Dass $(a,b)(c,d) = (c,d)(a,b)$ gilt, sieht man aus der Multiplikationsregel (*). Es lässt sich nachrechnen, dass die Multiplikation assoziativ ist, und zwar:

$$\begin{aligned}
&(a_1,b_1)[(a_2,b_2)(a_3,b_3)]\\
&= (a_1a_2a_3 - a_1b_2b_3 - b_1a_2b_3 - b_1b_2a_3, b_1a_2a_3 + a_1b_2a_3 + a_1a_2b_3 - b_1b_2b_3)\\
&= [(a_1,b_1)(a_2,b_2)](a_3,b_3)\,.
\end{aligned}$$

Distributivität lässt sich auch nachweisen (vgl. Übungsblatt). Das Einselement ist $(1,0)$, das Nullelement ist $(0,0)$. Man rechnet nach, dass jedes $z \neq 0$ ein Inverses z^{-1} hat: $(a,b)^{-1} = \left(\frac{a}{\sqrt{a^2+b^2}}, -\frac{b}{\sqrt{a^2+b^2}}\right)$. Somit ist $\mathbb{C}$ tatsächlich ein Körper.

Hilfssatz 3 *Seien $z,w \in \mathbb{C}$. Dann*

a) $\overline{z+w} = \bar{z} + \bar{w}$.

b) $\overline{zw} = \bar{z} \cdot \bar{w}$.

c) Es ist $z \in \mathbb{R}$ genau dann, wenn $z = \bar{z}$.

d) $z \cdot \bar{z} = |z|^2$. *Für $z \neq 0$ gilt also* $z^{-1} = \frac{\bar{z}}{|z|^2}$.

Beweis. Die ersten beiden Teile stellen eine Übungsaufgabe auf Blatt 5 dar.

c) $z = (a, b)$ liegt genau dann in $\mathbb{R}$, wenn $b = 0$ gilt. Dies ist genau dann der Fall, wenn $b = -b$ gilt, d.h. wenn $(a, b) = (a, -b)$ gilt.

d) Es ist $(a, b)(a, -b) = (a^2 + b^2, 0)$. Diese komplexe Zahl wird mit $a^2 + b^2 \in \mathbb{R}$ identifiziert. ∎

3 Vektorräume; Basis und Dimension

Definition Sei k ein Körper. Ein k-Vektorraum $V = (V, +, \cdot)$ besteht aus einer abelschen Gruppe $(V, +)$ zusammen mit einer Abbildung $\cdot\colon k \times V \to V$, $(\lambda, v) \mapsto \lambda v$, die die folgenden Eigenschaften erfüllt:

a) Assoziativität: $(\lambda\mu)v = \lambda(\mu v)$ für alle $\lambda, \mu \in k$ und für alle $v \in V$.

b) Distributivität: $(\lambda+\mu)v = \lambda v + \mu v$ und $\lambda(v+w) = \lambda v + \lambda w$ für alle $\lambda, \mu \in k$ und für alle $v, w \in V$.

c) Normierung: $1v = v$ für jedes $v \in V$.

Die Operation + bzw. $\cdot$ nennt man Addition bzw. Skalarmultiplikation. Häufig schreibt man 0 sowohl für die Null 0_k des k als auch für die Null 0_V des V.

Beispiele 1 Reelle Vektorräume, d.h. $k = \mathbb{R}$

a) Geometrische Vektoren im dreidimensionalen Anschauungsraum.

b) $\mathbb{R}^n$

c) $C^0(\mathbb{R})$

d) Alle Folgen in $\mathbb{R}$. Auch alle konvergente Folgen.

e) $\mathbb{C}$

f) Die Lösungsmenge eines homogenen linearen Gleichungssystems.

g) Alle Polynome, auch alle vom Grad höchstens n.

Bemerkung Ist V ein k-Vektorraum, so gelten $\lambda \cdot 0_V = 0_k \cdot v = 0_V$ und $(-\lambda)v = \lambda(-v) = -(\lambda v)$ für alle $\lambda \in k$ und für alle $v \in V$. Der Beweis ist der gleiche wie für Lemma 2.2.

Lemma 3.1 *Sei V ein k-Vektorraum. Erfüllen $\lambda \in k$ und $v \in V$ die Gleichung $\lambda v = 0$, so ist $\lambda = 0_k$ oder $v = 0_V$.*

Beweis. Wir zeigen: ist $\lambda v = 0$ aber $\lambda \neq 0$, dann gilt $v = 0$. Da λ ein Element $\neq 0$ des Körpers k ist, gibt es ein $\mu \in k$ mit $\mu\lambda = 1$. Es ist dann

$$0_V = \mu 0_V = \mu(\lambda v) = (\mu\lambda)v = 1v = v\,. \qquad \blacksquare$$

Beispiele 2 Andere zugrunde liegende Körper k.

a) k^n für beliebiges k.

b) $M(m \times n; k)$ für beliebiges k.

c) Alle analytische Funktionen auf $\mathbb{C}$, für $k = \mathbb{C}$.

d) Alle Bytes (d.h. Folgen von 8 Bits) für $k = \mathbb{F}_2$; die Addition ist die Operation XOR, die Skalarmultiplikation ist durch $1v = v$, $0v = 0$ definiert.

e) Die Potenzmenge $\mathscr{P}(M)$ einer Menge M, für $k = \mathbb{F}_2$. Die Addition ist die symmetrische Differenz $\triangle$ von Mengen:

$$A \triangle B := \{x \mid x \text{ liegt in genau einer der Mengen } A, B\} = (A \cup B) \setminus (A \cap B)\,.$$

Diese Operation ist assoziativ, denn

$$(T_1 \triangle T_2) \triangle T_3 = T_1 \triangle (T_2 \triangle T_3)$$
$$= \left\{ x \mid x \text{ liegt entweder in genau einer der Mengen } T_1, T_2, T_3, \text{ oder in allen drei} \right\}.$$

Der Nullvektor ist $\emptyset$, und $-T = T$. Die Skalarmultiplikation ist $1T = T$, $0T = \emptyset$.

f) $\mathbb{R}$ für $k = \mathbb{Q}$.

Bemerkung Die Vorstellung, dass die Elemente eines Vektorraums Vektoren sind, ist manchmal hilfreich – und manchmal nicht.

Linearkombinationen, lineare Abhängigkeit

Definition Sei V ein k-Vektorraum, und seien $v, v_1, v_2, \ldots, v_n$ Vektoren aus V. Gibt es Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n \in k$ derart, dass

$$v = \sum_{i=1}^{n} \lambda_i v_i \text{ gilt,} \qquad \text{das heißt } v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n\,,$$

so heißt v eine *Linearkombination* von $v_1, v_2 \ldots, v_n$.

Beispiele a) Jedes $v \in \mathbb{R}^3$ ist eine $\mathbb{R}$-lineare Kombination von $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ und $e_3 = (0, 0, 1)$, denn $(x, y, z) = xe_1 + ye_2 + ze_3$.

b) In $V = \mathbb{R}^3$ sei $v_1 = (1, 3, 1)$ und $v_2 = (0, 2, 1)$. Der Vektor $u = (2, 0, -1)$ ist eine Linearkombination von v_1, v_2, denn $u = 2v_1 - 3v_2$. Dagegen ist $w = (1, 0, 1)$ keine Linearkombination von v_1, v_2: wäre nämlich $w = \lambda v_1 + \mu v_2$, dann $(\lambda, 3\lambda + 2\mu, \lambda + \mu) = (1, 0, 1)$. Ein Koeffizientenvergleich ergibt $\lambda = 1$, $\lambda + \mu = 1$ (weshalb $\mu = 0$) und $3\lambda + 2\mu = 0$, weshalb $3 = 0$, ein Widerspruch.

c) Der Nullvektor 0_V ist eine Linearkombination jedes Systems $v_1, v_2, \ldots, v_n$ von Vektoren: man setzt $\lambda_i = 0$ für jedes i.

d) In $V = k^4$ sei $v_1 = (1,0,1,0)$, $v_2 = (0,1,0,1)$, $v_3 = (1,1,1,0)$, $v_4 = (1,1,0,0)$ und $v_5 = (0,0,1,1)$. Der Vektor $v = (1,1,1,1,)$ hat zwei verschiedene Darstellungen als eine Linearkombination von v_1, v_2, v_3, v_4, v_5: $v = v_1 + v_2$ und $v = v_4 + v_5$.

Somit ist $v_1 + v_2 - v_4 - v_5$ der Nullvektor. Das heißt, es gibt zwei verschiedene Darstellungen des Nullvektors als eine Linearkombination: die triviale Darstellung $0 = 0v_1 + 0v_2 + 0v_3 + 0v_4 + 0v_5$ sowie die nichttriviale Darstellung $0 = 1v_2 + 1v_2 + 0v_3 + (-1)v_4 + (-1)v_5$.

Definition Vektoren $v_1, v_2, \ldots, v_n$ heißen linear abhängig, wenn der Nullvektor sich als eine nichttriviale Linearkombination von ihnen darstellen lässt, d.h. wenn es Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n \in k$ gibt, die nicht alle Null sind und trotzdem $\sum_{i=1}^n \lambda_i v_i = 0$ erfüllen.

Beispiele a) Im obigen Beispiel sind die Vektoren v_1, v_2, v_3, v_4, v_5 linear abhängig.

b) In $V = C^0(\mathbb{R})$ sind die Funktionen $f(x) = 2x$, $g(x) = 3x$ und $h(x) = x^2$ linear abhängig, denn $3f - 2g + 0h = 0$.

Definition Sei V ein k-Vektorraum. Vektoren $v_1, v_2, \ldots, v_n \in V$ heißen linear unabhängig, wenn sie nicht linear abhängig sind.

Das heißt: $v_1, v_2, \ldots, v_n$ sind linear unabhängig, wenn gilt: sind $\lambda_1, \lambda_2, \ldots, \lambda_n \in k$ Skalare mit der Eigenschaft, dass $\sum_{i=1}^n \lambda_i v_i = 0$ gilt, so muss $\lambda_i = 0$ gelten für jedes i.

Beispiele a) Das System $v_1 = (1,3,1)$, $v_2 = (0,2,1)$ ist linear unabhängig in $\mathbb{R}^3$, denn: ist $\lambda v_1 + \mu v_2 = 0$, dann $(\lambda, 3\lambda + 2\mu, \lambda + \mu) = (0,0,0)$. Komponentenvergleich: $\lambda = 0$; $\lambda + \mu = 0$, also $\mu = 0$.

Sogar das System $v_1, v_2, v_3 = (1,0,1)$ ist linear unabhängig: ist $\lambda v_1 + \mu v_2 + \nu v_3 = 0$, dann $(\lambda + \nu, 3\lambda + 2\mu, \lambda + \mu + \nu) = (0,0,0)$. Komponentenvergleich:

$$\lambda + \nu = 0 \quad \text{(I)} \qquad 3\lambda + 2\mu = 0 \quad \text{(II)} \qquad \lambda + \mu + \nu = 0 \quad \text{(III)}\,.$$

(III) − (I): $\mu = 0$. Aus (II) folgt $\lambda = 0$, aus (I) folgt jetzt $\nu = 0$. Also $\lambda = \mu = \nu = 0$.

Das System $v_1, v_2, v_3, v_4 = (0,1,0)$ dagegen ist nicht linear unabhängig, denn $v_1 = v_3 + 3v_4$, also $1 \cdot v_1 + 0 \cdot v_2 + (-1) \cdot v_3 + (-3) \cdot v_4 = 0$.

b) In $C^0(\mathbb{R})$ sei $f(x) = 1$, $g(x) = x^2$ und $h(x) = e^x - 1$. Diese drei Funktionen sind linear unabhängig, denn: angenommen es ist $\lambda f(x) + \mu g(x) + \nu h(x) = 0$ für jedes $x \in \mathbb{R}$. Wir setzen $x = 0$ ein und erhalten $\lambda \cdot 1 + \mu \cdot 0 + \nu \cdot 0 = 0$, d.h. $\lambda = 0$. Jetzt setzen wir $x = 1$ und $x = -1$ ein:

$$\mu + \nu(e - 1) = 0 \quad \text{(I)} \qquad\qquad \mu + \nu(e^{-1} - 1) = 0 \quad \text{(II)}$$

(I) − (II): $\nu(e - e^{-1}) = 0$, also $\nu = 0$. Aus (I) folgt dann $\mu = 0$. Also $\lambda = \mu = \nu = 0$.

Unterräume, Erzeugendensysteme

Definition Sei V ein k-Vektorraum und $U \subseteq V$ eine Teilmenge. Ist U selbst ein Vektorraum, und zwar mit der gleichen Addition und Skalarmultiplikation wie V, dann nennt man U einen *Untervektorraum* oder kürzer einen *Unterraum* von V. Dies hat insbesondere zur Folge, dass U bezüglich der Addition und Skalarmultiplikation auf V abgeschlossen ist – das heißt, dass $u + v$ und λu in U liegen $\forall \lambda \in k$, $\forall u, v \in U$.

Anders gesagt: $U \subseteq V$ ist ein Unterraum, falls 0_V in U liegt, und außerdem $u + v$, $-u$ und λu in U liegen $\forall u, v \in U$, $\forall \lambda \in k$. Die Gesetze (Assoziativität usw.) gelten dann in U, denn sie gelten sogar in V.

Lemma 3.2 *Eine Teilmenge U eines k-Vektorraums V ist genau dann ein Unterraum, wenn folgende Bedingungen gelten:*

a) $U \neq \emptyset$ b) $\lambda u + \mu v \in U$ *für alle* $\lambda, \mu \in k$, $\forall u, v \in U$.

Beweis. Bedingungen notwendig: Wegen $0_V \in U$ ist $U \neq \emptyset$. Da U bezüglich der Skalarmultiplikation auf V abgeschlossen ist, liegen λu, μv in U. Da U auch bezüglich der Addition abgeschlossen ist, liegt $\lambda u + \mu v$ in U.

Bedingungen hinreichend: Wegen $U \neq \emptyset$ gibt es ein $u_0 \in U$. Dann ist $0_V = u_0 - u_0 = 1u_0 + (-1)u_0$ ein Element von U. Sind $u, v \in U$, dann liegen auch $u + v = 1u + 1v$, $-u = (-1)u + 1 \cdot 0_V$ und $\lambda u = \lambda u + 1 \cdot 0_V$ in U. ∎

Beispiele a) Aufgrund der Grenzwertsätze ist die Menge aller konvergenten Folgen in $\mathbb{R}$ ein Unterraum des $\mathbb{R}$-Vektorraums aller Folgen.

b) Die Menge P_5 aller Polynome vom Grad höchstens 5 ist ein Unterraum des Vektorraums aller Polynome. Die Menge aller Polynome vom Grad 5, die außerdem eine Nullstelle in $x = 2$ haben, ist widerum ein Unterraum von P_5.

c) Die Lösungsmenge des linearen Gleichungssystems $x_1 + x_2 - 3x_3 - x_4 = x_2 + 5x_4 = 0$ ist ein Unterraum des $\mathbb{R}^4$.

Definition Sei V ein k-Vektorraum und $v_1, v_2, \ldots, v_n$ ein System von Elementen aus V. Man schreibt

$$\text{Spann}(v_1, v_2, \ldots, v_n) := \{v \in V \mid v \text{ ist eine Linearkombination von } v_1, \ldots, v_n\}.$$

Diese Menge, die von $v_1, v_2, \ldots, v_n$ aufgespannt wird, nennt man auch die *lineare Hülle* dieser Vektoren.

Lemma 3.3 $\text{Spann}(v_1, \ldots, v_n)$ *ist ein Unterraum von* V*.*

Beweis. Indem man $\lambda_i = 0$ für jedes i setzt, sieht man, dass 0_V im Spann liegt. Sind $\lambda, \mu \in k$ und $u, v \in \text{Spann}(v_1, \ldots, v_n)$, so gibt es Skalare $a_1, a_2, \ldots, a_n$, $b_1, b_2, \ldots, b_n \in k$ mit $u = \sum_{i=1}^n a_i v_i$, $v = \sum_{i=1}^n b_i v_i$. Es ist dann $\lambda u + \mu v = \sum_{i=1}^n c_i v_i$, wobei $c_i \in k$ für jedes $1 \leq i \leq n$ durch $c_i = \lambda a_i + \mu b_i$ gegeben wird. Folglich liegt auch $\lambda u + \mu v$ im Spann. Das Ergebnis folgt aus Lemma 3.2. ∎

Definition Sei U ein Unterraum des k-Vektorraums V. Seien $v_1, v_2, \ldots, v_n$ Elemente aus V. Gilt $U = \text{Spann}(v_1, v_2, \ldots, v_n)$, so heißt $v_1, v_2, \ldots, v_n$ ein *Erzeugendensystem* von U.

Häufig betrachtet man den Fall $U = V$. Gibt es ein $n \geq 0$ und Vektoren $v_1, v_2, \ldots, v_n \in V$ derart, dass $\text{Spann}(v_1, v_2, \ldots, v_n) = V$ ist, so heißt V ein *endlich erzeugter* k-Vektorraum. Häufig sagt man *endlich-dimensional* anstelle von endlich-erzeugt.

In dieser Vorlesung beschäftigen wir uns vorwiegend mit endlich erzeugten Vektorräumen.

Bemerkung Setzt man $\lambda_j = 1$ und $\lambda_i = 0$ für $i \neq j$, so sieht man, dass $v_j = \sum_{i=1}^n \lambda_i v_i$ im Spann liegt. Hieraus folgt: ist $v_1, v_2, \ldots, v_n$ ein Erzeugendensystem für U, so muss insbesondere jedes v_i in U liegen.

Basen

Definition Eine *Basis* für einen Vektorraum ist ein linear unabhängiges Erzeugendensystem.

Beispiel Sei V der Vektorraum k^n für einen beliebigen Körper k. Für $1 \leq i \leq n$ definieren wir $e_i \in k^n$ als der Vektor $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ dessen $i$ten Komponente 1 ist und alle weitere Komponenten Null sind. Die Vektoren $e_1, e_2, \ldots, e_n$ bilden dann eine Basis für k^n, die sogenannte *Standardbasis*.

Erzeugendensystem: Es ist $\sum_{i=1}^n \lambda_i e_i = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, und jedes $v \in k^n$ hat diese Form. Linear unabhängig: ist $\sum_{i=1}^n \lambda_i e_i = 0_V = (0, 0, \ldots, 0)$, so zeigt ein Koeffizientenvergleich, dass $\lambda_i = 0$ für alle i.

Bemerkung Will man zeigen, dass $v_1, \ldots, v_n \in V$ eine Basis des Unterraums $U \subseteq V$ sind, so muss man drei Bedingungen prüfen:

a) $v_i \in U$ für alle i;

b) Jedes $u \in U$ ist eine Linearkombination von $v_1, \ldots, v_n$;

c) Das System $v_1, \ldots, v_n$ ist linear unabhängig.

Es ist nämlich $\text{Spann}(v_1, \ldots, v_n) \subseteq U$ wegen der ersten Bedingung zzgl. Hilfssatz 4 unten. Die zweite Bedingung liefert dann die umgekehrte Inklusion, d.h. $v_1, \ldots, v_n$ ist ein Erzeugendensystem für U.

Hilfssatz 4 *Sei $v_1, \ldots, v_n$ ein System von Vektoren eines k-Vektorraums V. Liegen alle Vektoren v_i in einem Unterraum $U \subseteq V$, so gilt* $\text{Spann}(v_1, \ldots, v_n) \subseteq V$.

Beweis. Induktion über n. $n = 0$: $\text{Spann}(\emptyset) = \{0\} \subseteq U$. $n = 1$: U ist abgeschlossen bezüglich Skalarmultiplikation. Induktionsschritt $n-1 \to n$: Ist $v \in \text{Spann}(v_1, \ldots, v_n)$, dann $v = \sum_{i=1}^{n} \lambda_i v_i$ für Skalare $\lambda_1, \ldots, \lambda_n \in k$. Wegen der Induktionsannahme liegt $a := \sum_{i=1}^{n-1} \lambda_i v_i$ in U; der Fall $n = 1$ zeigt, dass $b = \lambda_n v_n \in U$; und deshalb liegt auch $v = a + b$ in U, da U bezüglich Addition abgeschlossen ist. ∎

Beispiel Sei $V \subseteq \mathbb{R}^3$ der Lösungsraum der Gleichung $x_1 + x_2 + x_3 = 0$. Eine Basis für V ist $v_1 = (-1, 1, 0)$, $v_2 = (-1, 0, 1)$. Man sieht, dass v_1, v_2 beide in V liegen. Ist $(x, y, z) \in V$, dann $x = -y - z$, weshalb $(x, y, z) = yv_1 + zv_2$. Jedes Element von U liegt also in $\text{Spann}(v_1, v_2)$. Schließlich zeigt ein Komponentenvergleich, dass v_1, v_2 linear unabhängig sind.

Eine weitere Basis für V besteht aus $w_1 = (1, 0, -1)$ und $w_2 = (0, 1, -1)$.

Beispiel Sei V der Raum aller Polynome vom Grad ≤ 3, die eine Nullstelle in $x = -1$ haben. Drei solche Polynome sind $p_1(x) = x + 1$, $p_2(x) = x^2 + x$ und $p_3(x) = x^3 + x^2$. Diese drei Polynome sind auch linear unabhängig: ist $\lambda p_1(x) + \mu p_2(x) + \nu p_3(x) = 0$ für alle x, dann $\lambda = 0$ ($x = 0$ einsetzen), also $\mu p_2(x) + \nu p_3(x) = 0$. Leiten wir diese Gleichung einmal ab und setzen wir dann $x = 0$ ein, erhalten wir $\mu = 0$, denn $p_2'(x) = 2x + 1$, $p_3'(x) = 3x^2 + 2x$. Also $\nu p_3(x) = 0$. Setzen wir $x = 1$ ein so erhalten wir $2\nu = 0$, weshalb $\lambda = \mu = \nu = 0$.

Diese drei Polynome erzeugen auch V, denn: ist $q(x) \in V$, so ist $q(x) = ax^3 + bx^2 + cx + d$ mit $d = c - b + a$. Also $q(x) = ap_3(x) + (b-a)p_2(x) + (c-b+a)p_1(x)$. Damit ist p_1, p_2, p_3 eine Basis für V.

Unendliche Systeme

Nicht alle Vektorräume sind endlich erzeugt. Ein Beispiel ist der Vektorraum aller Folgen in $\mathbb{R}$. Um mit unendlich-dimensionalen Vektorräume umgehen zu können, müssen wir die Begriffe Linearkombination, Erzeugendensystem und lineare Unabhängigkeit auch für unendliche Systeme erklären.

Hat man endlich viele Vektoren, so nennt man sie $v_1, \ldots, v_n$. Bei unendlich vielen Vektoren ist es nicht immer möglich bzw. sinnvoll, sie ausschöpfend als $v_1, v_2, v_3, \ldots$ zu bezeichnen: denn manche unendliche Mengen wie z.B. $\mathbb{R}$ sind unvergleichbar großer als die natürlichen Zahlen $\mathbb{N}$ (Stichwort Überabzählbarkeit). Deswegen gibt man ein unendliches System von Vektoren als $(v_i)_{i\in I}$ an, wobei die Indexmenge I eine beliebig große Menge sein kann.

Beispiel Für $a \in \mathbb{R}$ sei $f_a\colon \mathbb{R} \to \mathbb{R}$ die stetige Funktion $f_a(x) = x^2 - ax$. So erhält man ein durch $\mathbb{R}$ indiziertes unendliches System $(f_a(x))_{a\in\mathbb{R}}$ von stetigen Funktionen auf $\mathbb{R}$.

Definition Sei V ein k-Vektorraum und $(v_i)_{i\in I}$ ein beliebiges System von Vektoren aus V.

a) Ein Vektor $v \in V$ heißt eine Linearkombination von $(v_i)_{i\in I}$, wenn es endlich viele Elemente $i_1, i_2, \ldots, i_n \in I$ gibt derart, dass v eine Linearkombination des endlichen Untersystems $v_{i_1}, v_{i_2}, \ldots, v_{i_n}$ ist.

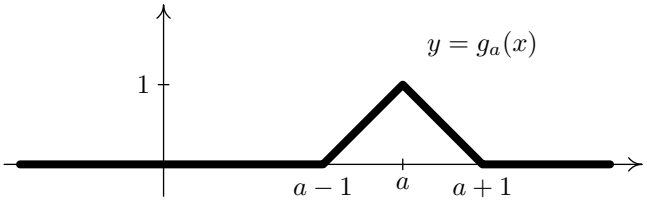
b) Ein unendliches System heißt genau dann linear unabhängig, wenn jedes endliches Untersystem linear unabhängig ist. Folglich heißt das System linear abhängig, wenn irgendein endliches Untersystem linear abhängig ist.

c) Wieder wird die lineare Hülle $\mathrm{Spann}(v_i \mid i \in I)$ als die Menge aller Linearkombinationen definiert, und $(v_i)_{i\in I}$ heißt ein Erzeugendensystem des Unterraums $U \subseteq V$, falls U die lineare Hülle ist.

d) Die Definition, dass eine Basis ein linear unabhängiges Erzeugendensystem ist, muss nicht angepasst werden.

Das heißt, auch für unendliche Systeme beschäftigt man sich nur mit endlichen Summen: in vielen Vektorräumen gibt es auch gar keinen Grenzwertbegriff, von unendlichen Summen kann also keine Rede sein.

Beispiel Für jedes $a \in \mathbb{R}$ sei $g_a\colon \mathbb{R} \to \mathbb{R}$ folgende stetige Funktion: es ist $g_a(x) = 0$ für $x \le a - 1$ und für $x \ge a + 1$; $g_a(a) = 1$; und g_a ist linear auf den Intervallen $[a-1, a]$ und $[a, a+1]$.



Wir werden zeigen, dass dieses System linear unabhängig ist, und dass die Funktion $f(x) = x^2$ keine Linearkombination der g_a ist.

Ist $g_{a_1}, g_{a_2}, \ldots, g_{a_n}$ ein endliches Untersystem, so dürfen wir problemlos die a_i in aufsteigender Reihenfolge wählen. Es ist also $a_n > a_i$ für alle $i \leq n-1$. Für $\varepsilon > 0$ klein genug ist dann $a_n + 1 - \varepsilon > a_i + 1$ für alle $i \leq n-1$, weshalb $g_{a_n}(a_n + 1 - \varepsilon) > 0$ und $g_{a_i}(a_n + 1 - \varepsilon) = 0$ für alle $i \leq -1$. Ist $\sum_{i=1}^{n} \lambda_i g_{a_i}(x) = 0$ für alle x, dann folgt $\lambda_n = 0$ durch Einsetzen $x = a_n + 1 - \varepsilon$. Per Induktion über n folgt dann: es ist $\lambda_i = 0$ für alle i, d.h. $g_{a_1}, \ldots, g_{a_n}$ sind linear unabhängig, also ist das ganze System $(g_a)_{a \in \mathbb{R}}$ linear unabhängig.

Wäre

$$f(x) = \sum_{i=1}^{n} \lambda_i g_{a_i}(x) \qquad (*)$$

für alle x, so setzen wir $x = x_0$ ein für ein x_0 mit $x_0 \geq 1$ und $x_0 \geq a_i + 1$ für alle $1 \leq i \leq n$. Wegen $x_0 \geq a_i + 1$ ist $g_{a_i}(x) = 0$ für alle i, weshalb die rechte Seite von (*) den Wert 0 annimmt. Wegen $x_0 \geq 1$ ist $f(x_0) \geq 1$, ein Widerspruch. Somit ist $f(x)$ keine Linearkombination der Funktionen $g_a(x)$.

Lemma 3.4 *Die neuen Definitionen – für beliebig große Systeme – der Begriffe Linearkombination, lineare Unabhängigkeit, Erzeugendensystem und Basis stimmen im Fall eines endlichen Systems mit den alten Definitionen überein. Auch für ein beliebig großes System von Vektoren aus einem k-Vektorraum V ist der Spann ein Untervektorraum von V.*

Beweis. Sei $v_1, \ldots, v_n$ ein endliches System in V. Linearkombinationen: in der neuen Definition können einige Vektoren v_i weggelassen werden; dies erreicht man aber auch, indem man die entsprechenden λ_i gleich Null setzt. Lineare Unabhängigkeit: Lässt ein Untersystem von $v_1, \ldots, v_n$ eine nichttriviale Darstellung des Nullvektors zu, so erhält man, indem man wieder die restlichen λ_i gleich Null setzt, eine nichttriviale Darstellung des Nullvektors durch die ursprünglichen Vektoren $v_1, \ldots, v_n$. Die anderen beiden Begriffe – Erzeugendensystem und Basis – werden anhand der Begriffe Linearkombinationen und lineare Unabhängigkeit definiert.

Nun sei $(v_\alpha)_{\alpha \in A}$ ein beliebig großes System von Vektoren in V. Der Nullvektor ist als Linearkombination des leeren Systems immer im Spann. Sind u, w Elemente des Spanns und λ, μ Skalare, so gibt es Elemente $\beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n \in A$ derart, dass u bzw. w eine Linearkombination des Systems $v_{\beta_1}, \ldots v_{\beta_m}$ bzw. des Systems $v_{\gamma_1}, \ldots, v_{\gamma_n}$ ist. Dann ist aber $\lambda u + \mu w$ eine Linearkombination des vereinigten Systems $v_{\beta_1}, \ldots v_{\beta_m}, v_{\gamma_1}, \ldots, v_{\gamma_n}$ und folglich ein Element von $\mathrm{Spann}(v_\alpha \mid \alpha \in A)$. Nach Lemma 3.2 ist der Spann ein Unterraum. ∎

Bemerkung Die bereits angekündigte Tatsache, dass der Raum aller Folgen in $\mathbb{R}$ nicht endlich erzeugt ist, werden wir erst mit Hilfe des Steinitzschen Austauschsatzes beweisen. Ein unendliches linear unabhängiges System lässt sich

aber gut konstruieren. Für $m \geq 1$ sei $a(m)$ die Folge $a(m)_1, a(m)_2, a(m)_3, \ldots$ mit $a(m)_m = 1$ und $a(m)_n = 0$ sonst. Das unendliche System $(a(m))_{m\in\mathbb{N}}$ von Folgen ist linear unabhängig: denn jedes endliche Untersystem hat die Gestalt $a(m_1), a(m_2), \ldots, a(m_r)$ für natürliche Zahlen $m_1 < m_2 < \cdots < m_r$, und für alle Skalare $\lambda_1, \ldots, \lambda_r$ ist die Linearkombination $a = \sum_{j=1}^{r} \lambda_j a(m_j)$ die Folge mit $a_{m_j} = \lambda_j$ für jedes j und $a_i = 0$, falls i kein m_j ist. Folglich ist a nur dann die Nullfolge, wenn jedes $\lambda_j = 0$ ist, was wir auch zeigen wollten.

Dass der $\mathbb{F}_2$-Vektorraum aller Folgen im endlichen Körper $\mathbb{F}_2$ nicht endlich erzeugt ist, lässt sich dagegen schon jetzt zeigen. Da der Körper $\mathbb{F}_2$ nur zwei Elemente hat, haben n Elemente eines $\mathbb{F}_2$-Vektorraums höchstens 2^n verschiedene Linearkombinationen. Da der Raum aller Folgen im $\mathbb{F}_2$ unendlich viele Elemente hat, kann es kein endliches Erzeugendensystem haben.

Sätze über Basen

Über Basen sind viele allgemeine Aussagen möglich.

- Existenz: Jeder Vektorraum hat eine Basis.
- Basisergänzungssatz: Jedes linear unabhängige System lässt sich zu einer Basis fortsetzen.
- Steinitzscher Austauschsatz: Hat zur Folge, dass alle Basen gleich lang sind.
- Charakterisierung von Basen: die Begriffe „Basis“, „maximales linear unabhängiges System“ und „minimales Erzeugendensystem“ sind gleichbedeutend.

Wir beschränken uns hier auf endlich erzeugten Vektorräume: wegen des robusten Dimensionsbegriffs ist dieser Fall besonders wichtig, dagegen sind die Beweise im unendlich dimensionalen Fall komplizierter (Zornsche Lemma).

Lemma 3.5 *Sei V ein k-Vektorraum.*

a) Ist $\sum_{j=1}^{n} \lambda_j v_j = 0$ für Skalare $\lambda_j \in k$ und Vektoren $v_j \in V$, und ist $\lambda_i \neq 0$, so ist v_i eine Linearkombination der Vektoren $v_1, \ldots, \hat{v}_i, \ldots, v_n$.

b) Ist $v \in V$ eine Linearkombination des Systems $v_1, \ldots, v_n$, so ist das System $v_1, \ldots, v_n, v$ linear abhängig.

Beweis. a) Es ist $\lambda_i v_i = -\sum_{\substack{j=1 \\ j\neq i}}^{n} \lambda_j v_j$, weshalb

$$v_i = -\sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\lambda_j}{\lambda_i} v_j \,.$$

b) Schreiben wir $v_{n+1} = v$. Da v eine Linearkombination ist, gibt es Skalare $\mu_1, \ldots, \mu_n \in k$ mit $v = \sum_{i=1}^{n} \mu_i v_i$. Nun sei $\lambda_i = -\mu_i$ für $i \leq n$ und $\lambda_{n+1} = 1$. Dann $\sum_{i=1}^{n+1} \lambda_i v_i = v - \sum_{i=1}^{n} \mu_i v_i = 0$. Da $\lambda_{n+1} \neq 0$ ist, ist das System $v_1, \ldots, v_n, v$ linear abhängig.

■

Lemma 3.6 *Sei $v_1, v_2, \ldots, v_n$ ein System von Vektoren im k-Vektorraum V.*

a) Ist $v_1, \ldots, v_n$ ein linear abhängiges Ergzeugendensystem für V, so gibt es mindestens ein $1 \leq i \leq n$ mit folgdender Eigenschaft: auch nachdem man v_i aus dem System streicht, liegt ein Erzeugendensystem weiherhin vor.

Zusatz: Sind $v_1, \ldots, v_r$ linear unabhängig, so kann man ein solches i in $\{r+1, \ldots, n\}$ finden.

b) Ist $v_1, \ldots, v_n$ ein linear unabhängiges System in V aber kein Erzeugendensystem, so kann man ein $v_{n+1} \in V$ derart finden, dass auch $v_1, \ldots, v_n, v_{n+1}$ ein linear unabhängiges System ist.

Beweis. a) Aufgrund linearer Abhängigkeit gibt es Skalare $\lambda_1, \ldots, \lambda_n \in k$, die $\sum_{j=1}^{n} \lambda_j v_j = 0$ erfüllen und nicht alle Null sind. Sei $i \in \{1, \ldots, n\}$ einer der Indizes mit $\lambda_i \neq 0$. Für den Zusatz merken wir an dieser Stelle an, dass man $i \geq r+1$ wählen kann: denn ist $\lambda_j = 0$ für jedes $j \geq r+1$, so ist $\sum_{j=1}^{r} \lambda_j v_j = 0$, weshalb auch $\lambda_j = 0$ für jedes $j \leq r$ wegen der linearen Unabhängigkeit.

Nach Lemma 3.5 liegt v_i in $\mathrm{Spann}(v_1, \ldots, \hat{v}_i, \ldots, v_n)$, also liegt das System $v_1, \ldots, v_n$ in diesem Spann. Nach Hilfssatz 4 auf S. 24 folgt dann: es ist

$$V = \mathrm{Spann}(v_1, \ldots, v_n) \subseteq \mathrm{Spann}(v_1, \ldots, \hat{v}_i, \ldots, v_n) \subseteq V\,,$$

weshalb $\mathrm{Spann}(v_1, \ldots, \hat{v}_i, \ldots, v_n) = V$.

b) Ist $v_1, \ldots, v_n$ kein Erzeugendensystem, so gibt es ein $v \in V$, das keine Linearkombination von $v_1, \ldots, v_n$ ist. Wir setzen $v_{n+1} := v$. Seien nun $\lambda_1, \ldots, \lambda_{n+1} \in k$ Skalare mit $\sum_{i=1}^{n+1} \lambda_i v_i = 0$. Wir müssen zeigen, dass $\lambda_i = 0$ ist für jedes i. Wäre $\lambda_{n+1} \neq 0$, dann wäre v_{n+1} nach Lemma 3.5 doch eine Linearkombination von $v_1, \ldots, v_n$ ist, ein Widerspruch. Somit ist $\lambda_{n+1} = 0$, weshalb $\sum_{i=1}^{n} \lambda_i v_i = 0$. Da $v_1, \ldots, v_n$ linear unabhängig sind, folgt auch $\lambda_i = 0$ für alle $i \leq n$.

■

Satz 3.7 (Existenz von Basen) *Jeder endlich erzeugter Vektorraum hat mindestens eine Basis.*

1. Zusatz: Jedes endliches Erzeugendensystem lässt sich durch Streichungen zu einer Basis kürzen.

2. Zusatz: Bilden die ersten r Vektoren ein linear unabhängiges System, so kann man diese bei den Streichungen schonen.

Beweis. Die Aussage folgt aus dem 1. Zusatz. Zu diesem Zusatz: Sei $v_1, \ldots, v_n$ ein Erzeugendensystem des k-Vektorraums V. Wir zeigen die Aussage per Induktion über n. Ist das System linear unabhängig, so liegt eine Basis bereits vor. Ist das System linear abhängig, so können wir nach Lemma 3.6 ein geeignetes v_i streichen, um ein Erzeugendensystem der Länge $n-1$ zu erhalten. Dies ist der Induktionsschritt. Der Induktionsanfang ist der Fall $n = 0$: ein System von keinen Vektoren ist immer linear unabhängig! Auch im Fall $n = 1$ ist das System v_1 nur dann ein linear abhängiges Erzeugendensystem, wenn $v_1 = 0$ und $V = \{0\}$ ist – und in diesem Fall ist die leere Menge eine Basis.

Der 2. Zusatz folgt aus dem Zusatz zu Lemma 3.6 Teil a). ∎

Lemma 3.8 (Charakterisierung von Basen) *Für ein System $v_1, \ldots, v_n$ im k-Vektorraum V sind die folgenden drei Aussagen äquivalent:*

a) $v_1, \ldots, v_n$ ist eine Basis für V, d.h. ein linear unabhängiges Erzeugendensystem.

b) $v_1, \ldots, v_n$ ist ein minimales Erzeugendensystem für V.
Ein Erzeugendensystem ist minimal, wenn für jedes $1 \leq i \leq n$ gilt: streicht man v_i, so liegt kein Erzeugendensystem mehr vor.

c) $v_1, \ldots, v_n$ ist ein maximales linear unabhängiges System in V.
Maximal heißt, dass, egal wie man $v_{n+1} \in V$ wählt, das System $v_1, \ldots, v_{n+1}$ linear abhängig ist.

Beweis. Minimales Erzeugendensystem ⇒ Basis: Folgt aus Satz 3.7, denn wegen Minimalität sind keine Streichungen möglich.

Basis ⇒ minimales Erzeugendensystem: Ist $v_1, \ldots, v_n$ linear unabhängig, so ist nach Lemma 3.5 keiner der Vektoren v_i eine Linearkombination der anderen Vektoren, d.h. linear unabhängige Erzeugendensysteme sind minimal.

Basis ⇒ maximal linear unabhängig: Ist $v_1, \ldots, v_n$ ein Erzeugendensystem, so ist nach Lemma 3.5 $v_1, \ldots, v_n, v_{n+1}$ linear abhängig für jedes $v_{n+1} \in V$.

Maximal linear unabhängig ⇒ Basis: Sei $v \in V$. Mit $v_{n+1} = v$ ist das System $v_1, \ldots, v_{n+1}$ linear abhängig wegen Maximalität. Es gibt also Skalare $\lambda_1, \ldots, \lambda_{n+1}$, die $\sum_{i=1}^{n+1} \lambda_i v_i = 0$ erfüllen und nicht alle Null sind. Da $v_1, \ldots, v_n$ linear unabhängig sind, darf $\lambda_{n+1} = 0$ nicht sein. Also $\lambda_{n+1} \neq 0$, weshalb v nach Lemma 3.5 eine Linearkombination der Vektoren $v_1, \ldots, v_n$ ist. ∎

Basisergänzungssatz *Sei V ein endlich erzeugter k-Vektorraum. Jedes linear unabhängige System $v_1, \ldots, v_r$ in V lässt sich zu einer Basis $v_1, \ldots, v_n$ für V fortsetzen.*

Beweis. Sei $w_1, \ldots, w_m$ ein Erzeugendensystem für V. Dann ist $v_1, \ldots, v_r$, $w_1, \ldots, w_m$ auch ein Erzeugendensystem. Nach den beiden Zusätzen zu Satz 3.7 dürfen wir dieses lange Erzeugendensystem zu einer Basis zusammenstreichen, wobei wir ausschließlich Vektoren der Art w_j streichen, d.h. jedes v_i kommt in der Basis vor. ∎

Der Austauschsatz

Der Steinitzsche Austauschsatz *Gegeben seien sowohl ein linear unabhängiges System $a_1, a_2, \ldots, a_m$ als auch ein Erzeugendensystem $b_1, b_2, \ldots, b_n$ im k-Vektorraum V. Dann ist $m \leq n$, und ferner gilt: nachdem man evtl. die Reihenfolge der b_j geändert hat, ist auch $a_1, \ldots, a_m, b_{m+1}, \ldots, b_n$ ein Erzeugendensystem von V.*

Der Austauschsatz wird per Induktion mit Hilfe des folgenden Lemmas bewiesen:

Lemma 3.9 *Sei V ein k-Vektorraum. Für ein $m \geq 0$ seien $a_1, a_2, \ldots, a_{m+1}$ linear unabhängige Elemente von V; und für ein $n \geq m+1$ seien $b_{m+1}, b_{m+2}, \ldots, b_n$ Elemente von V derart, dass $a_1, a_2, \ldots, a_m, b_{m+1}, b_{m+2}, \ldots, b_n$ ein Erzeugendensystem für V bilden. Dann gibt es ein j_0 zwischen $m+1$ und n derart, dass die Elemente $a_1, a_2, \ldots, a_{m+1}$, $b_{m+1}, b_{m+2}, \ldots, \widehat{b_{j_0}}, \ldots, b_n$ ein Erzeugendensystem von V bilden.*

Beweis des Lemmas. Der Vektor a_{m+1} ist eine Linearkombination von den Vektoren $a_1, a_2, \ldots, a_m, b_{m+1}, b_{m+2}, \ldots, b_n$, denn dieses ist ein Erzeugendensystem. Nach Lemma 3.5 Teil b) ist also $a_1, a_2, \ldots, a_{m+1}, b_{m+1}, \ldots, b_n$ ein linear abhängiges Erzeugendensystem. Da $a_1, \ldots, a_{m+1}$ linear unabhängig sind, besagt der Zusatz zu Lemma 3.6 Teil a), dass wir ein geeignetes b_{j_0} streichen dürfen, ein Erzeugendensystem wird weiterhin vorliegen. ∎

Beweis des Austauschsatzes. Mit Hilfe des Lemmas ersetzen wir im Erzeugendensystem ein b_{j_0} durch a_1. Nach einer Umnummerierung wird es b_1 sein, der ersetzt wird. Dann ersetzen wir eins der übriggebliebenen b_j – nach Umnummerierung wird es b_2 sein – durch a_2. Nach und nach wird also jedes b_i durch ein a_i ersetzt; ein Erzeugendensystem der Länge n liegt weiterhin vor.

Dieses Ersetzen hört erst dann auf, wenn entweder die a_i oder die b_j ausgeschöpft sind. Sind die a_i ausgeschöpft, so haben wir das erwünschte Ergebnis. Ist aber $m > n$, dann gehen die b_j zuerst aus. In diesem Fall ist $a_1, \ldots, a_n$ ein Erzeugendensystem, weshalb a_{n+1} eine Linearkombination von $a_1, \ldots, a_n$ sein muss. Nach Lemma 3.5 Teil b) ist das System $a_1, \ldots, a_{n+1}$ linear abhängig. Das kann aber nicht sein, denn $a_1, \ldots, a_m$ ist linear unabhängig. ∎

Korollar 3.10 *Kein endlich erzeugter Vektorraum enthält ein unendliches linear unabhängiges System.*

Beweis. Ist $b_1, \ldots, b_m$ ein Erzeugendensystem und $(a_i)_{i \in I}$ ein unendliches linear unabhängiges System im k-Vektorraum V, so können wir $m+1$ verschiedene Elemente $i_1, \ldots, i_{m+1} \in I$ wählen. Dann ist $a_{i_1}, \ldots, a_{i_{m+1}}$ ein linear unabhängiges System in V, das länger als ein Erzeugendensystem ist. Diesem Zustand widerspricht der Austauschsatz. ∎

Der Dimensionsbegriff

Satz 3.11 (Alle Basen gleich lang) *Ist der k-Vektorraum V endlich erzeugt, so gilt: Jede Basis von V ist endlich, und alle Basen haben die gleiche Länge.*

Beweis. Nach Satz 3.7 hat V mindestens eine Basis. Nach Korollar 3.10 ist jedes linear unabhängiges System und deshalb jede Basis endlich. Da die a_i linear unabhängig sind, und die b_j ein Erzeugendensystem bilden, folgt $m \leq n$ aus dem Austauschsatz. Da die b_j linear unabhängig sind, und die a_i ein Erzeugendensystem bilden, folgt auch $n \leq m$ aus dem Austauschsatz. Also $n = m$. ∎

Definition Sei V ein endlich erzeugter k-Vektorraum. Nach Satz 3.7 hat V mindestens eine Basis. Nach Satz 3.11 sind alle Basen gleich lang. Diese gemeinsame Länge aller Basen heißt die *Dimension* $\dim(V)$ von V.

Aus diesem Grund spricht man häufig von endlich dimensionalen Vektorräumen, statt von endlich erzeugten Vektorräumen.

Satz 3.12 *Sei V ein k-Vektorraum und $U \subseteq V$ ein Unterraum. Dann ist auch U endlich dimensional, und es ist $\dim(U) \leq \dim(V)$. Ferner gilt: es ist $\dim(U) = \dim(V)$ genau dann, wenn $U = V$ ist.*

Beweis. Sei $n = \dim(V)$. Sei $v_1, \ldots, v_r$ ein linear unabhängiges System in U. Dieses System ist auch in V linear unabhängig, kann also (Basisergänzungssatz) zu einer Basis $v_1, \ldots, v_n$ von V fortgesetzt werden. Also $r \leq n$.

Nachdem wir das linear unabhängiges System $v_1, \ldots, v_r$ in U evtl. neu gewählt haben, dürfen wir annehmen, dass es mindestens so lang ist wie jedes andere linear unabhängige System in U, d.h. die Länge r nimmt ihre größtmöglichen Wert an. In diesem Fall ist $v_1, \ldots, v_r$ ein maximales linear unabhängiges System in U, d.h. eine Basis für U. Also $\dim(U) = r \leq n = \dim(V)$.

Nach wie vor kann man $v_1, \ldots, v_r$ zu einer Basis $v_1, \ldots, v_n$ von V fortsetzen. Ist $r = n$, dann ist $v_1, \ldots, v_r$ bereits diese Basis. Also $U = \text{Spann}(v_1, \ldots, v_r) = V$. ∎

Beispiel Aufgabe: Zeigen Sie, dass $v_1 = (2, -1, 1, 0, 0)$, $v_2 = (-2, 1, 1, 1, 1)$ und $v_3 = (2, -2, 0, 1, -1)$ eine Basis des folgenden Lösungsraums U bilden:

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid x_1 - 2x_3 + x_4 + 3x_5 = x_1 + x_2 - x_3 + x_4 + x_5 = 0\}.$$

Lösung: Man rechnet leicht nach, dass die drei Vektoren in U liegen. Ferner sind sie linear unabhängig: ist $\lambda v_1 + \mu v_2 + \nu v_3 = 0$, dann $\lambda = \mu = \nu = 0$ (Komponentenvergleich für die 3., 4. und 5. Komponenten). Sei $W \subseteq U$ der Unterraum $\text{Spann}(v_1, v_2, v_3)$. Dann ist v_1, v_2, v_3 eine Basis für W, weshalb $\dim(W) = 3$. Wir werden zeigen: es ist $\dim(U) = 3$. Nach Satz 3.12 folgt dann $W = U$, und wir sind fertig.

Nach Satz 3.12 bedeutet $W \subseteq U$, dass $\dim(U) \geq 3$ ist. Sei $T \subseteq \mathbb{R}^5$ der Lösungsraum der Gleichung $x_1 - 2x_3 + x_4 + 3x_5 = 0$. Dann $U \subseteq T \subseteq \mathbb{R}^5$. Nun, $e_1 = (1,0,0,0,0)$ liegt in $\mathbb{R}^5$ aber nicht in T. Also $T \subsetneq \mathbb{R}^5$. Aus Satz 3.12 folgt $\dim(T) < \dim(\mathbb{R}^5) = 5$, d.h. $\dim(T) \leq 4$. Ferner liegt $e_2 = (0,1,0,0,0)$ in T aber nicht in U. Also $\dim(U) < \dim(T)$, woraus folgt $\dim(U) \leq 3$. Also $\dim(U) = 3$, wie erwünscht.

Unterräume: Dimensionsformel, Komplemente

Definition Seien U, W Unterräume eines k-Vektorraums V.

a) Die Summe $U + W$ wird als der Unterraum

$$U + W := \{v \in V \mid \text{Es gibt } u \in U,\, w \in W \text{ mit } v = u + w\}$$

von V definiert.

b) Gilt $U \cap W = \{0\}$, so nennt man die Summe $U + W$ eine *direkte* Summe, Bezeichnung $U \oplus W$.

c) Ist die Summe $U + W$ direkt, und gilt ferner $U \oplus W = V$, so heißt W ein *Komplement* von U.

Hilfssatz 5 *Der Schnitt $U \cap W$ und die Summe $U + W$ sind tatsächlich Unterräume von V.*

Die Räume U, W sind Unterräume der Summe $U+W$. Diese Summe ist genau dann direkt, wenn jedes $v \in U+W$ genau eine Darstellung $v = u+w$ mit $u \in U$, $w \in W$ hat.

Beweis. 1. Teil: s. Blatt 6. 2. Teil: Summe direkt $\Rightarrow$ Darstellung eindeutig: Hat v zwei solche Darstellungen $v = u + w = u' + w'$, dann $u - u' = w' - w$. Die linke Seite dieser Gleichung liegt in U, die rechte in W, also liegen beide Seiten in $U \cap W = \{0\}$. Das heißt, $u = u'$ und $w = w'$.

Inklusionen, sowie Darstellung eindeutig $\Rightarrow$ Summe direkt:
Jedes $u \in U$ und jedes $w \in W$ liegen in $U + W$, denn $u = u + 0$ und $w = 0 + w$. Ist $v \in U \cap W$, dann hat v zwei Darstellungen: $v = v + 0 = 0 + v$. Da diese beiden gleich sein müssen, muss $v = 0$ sein. ∎

Definition Sind $U_1, U_2, \ldots, U_r$ Unterräume eines k-Vektorraums V, so kann man problemlos die Summe $U_1 + U_2 + \cdots + U_r$ rekursiv als $(U_1 + \cdots + U_{r-1}) + U_r$ definieren. Per Induktion weist man dann nach: es ist

$$U_1 + U_2 + \cdots + U_r = \{u_1 + u_2 + \cdots + u_r \mid u_i \in U_i \text{ für jedes } 1 \leq i \leq r\}.$$

Diese mehrfache Summe nennt man direkt, wenn jede Summe in der rekursiven Definition direkt ist, d.h. wenn jedes $v \in U_1 + \cdots + U_r$ genau eine Darstellung $v = u_1 + \cdots + u_r$ hat.

Beispiel Sei U_1 bzw. U_2 bzw. U_3 der eindimensionale Unterraum des $\mathbb{R}^2$ mit Basis $(1,0)$ bzw. $(0,1)$ bzw. $(1,1)$. Für alle $i \neq j \in \{1,2,3\}$ ist $U_i + U_j$ direkt, aber $U_1 + U_2 + U_3$ ist nicht direkt.

Satz 3.13 (Dimensionsformel für Unterräume) *Für Unterräume U_1, U_2 eines endlich dimensionalen k-Vektorraums V gilt*

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)\,.$$

Beweis. Sei $r = \dim(U_1 \cap U_2)$, $m = \dim(U_1)$ und $n = \dim(U_2)$. Nach Satz 3.12 ist $r \leq \min(m,n)$. Sei $a_1, \ldots, a_r$ eine Basis von $U_1 \cap U_2$. Der Basisergänzungssatz besagt, dass es $b_{r+1}, \ldots, b_m \in U_1$ gibt derart, dass $a_1, \ldots, a_r, b_{r+1}, \ldots, b_m$ eine Basis für U_1 ist. Analog hat U_2 eine Basis der Form $a_1, \ldots, a_r, c_{r+1}, \ldots, c_n$. Wir werden zeigen, dass das System $a_1, \ldots, a_r,\ b_{r+1}, \ldots, b_m,\ c_{r+1}, \ldots, c_n$ eine Basis für $U_1 + U_2$ ist. Daraus folgt der Satz, denn dieses System hat die Länge $r + (m - r) + (n - r) = m + n - r$.

Erzeugendensystem: Jedes $v \in U_1 + U_2$ ist eine Summe $v = u_1 + u_2$ mit $u_1 \in U_1$, $u_2 \in U_2$. Da u_1 eine Linearkombination der a_i und der b_j ist, und u_2 eine Linearkombination der a_i and der c_ℓ ist, ist v eine Linearkombination der a_i, b_j und c_ℓ.

Linear unabhängig: Ist $\sum_{i=1}^{r} \lambda_i a_i + \sum_{j=r+1}^{m} \mu_j b_j + \sum_{\ell=r+1}^{n} \nu_\ell c_\ell = 0$, dann

$$\sum_{i=1}^{r} \lambda_i a_i + \sum_{j=r+1}^{m} \mu_j b_j = -\sum_{\ell=r+1}^{n} \nu_\ell c_\ell\,. \qquad (*)$$

Da die linke bzw. rechte Seite in U_1 bzw. in U_2 liegt, liegen beide Seiten in $U_1 \cap U_2$. Somit gibt es $\rho_1, \ldots, \rho_r \in k$ mit

$$\text{Beide Seiten von } (*) = \sum_{i=1}^{r} \rho_i a_i\,.$$

Betrachten wir die rechte Seite von (*), erhalten wir

$$\sum_{i=1}^{r} \rho_i a_i + \sum_{\ell=r+1}^{n} \nu_\ell c_\ell = 0\,.$$

Da die a_i und die c_ℓ eine Basis von U_2 bilden, folgt $\rho_i = \nu_\ell = 0$ für alle i, ℓ. Somit ist auch die linke Seite von (*) gleich Null. Da die a_i und die b_j eine Basis von U_1 bilden, folgt $\lambda_i = \mu_j = 0$ für alle i, j. Also bilden die a_i, b_j und c_ℓ ein linear unabhängiges System. ∎

Korollar 3.14 *Ist die Summe direkt, so gilt* $\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2)$.

Beweis. In diesem Fall ist $U_1 \cap U_2 = \{0\}$. Die leere Menge stellt eine Basis dar, die Dimension ist also 0. ∎

Satz 3.15 (Existenz eines Komplements) *In einem endlich dimensionalen k-Vektorraum V besitzt jeder Unterraum U mindestens ein Komplement. Jedes Komplement W erfüllt*

$$\dim(W) = \dim(V) - \dim(U)\,.$$

Beweis. Wir nehmen eine Basis $v_1, \ldots, v_r$ von U. und setzen diese zu einer Basis $v_1, \ldots, v_n$ von V fort. Dann setzen wir $W = \text{Spann}(v_{r+1}, \ldots, v_n)$.

$V = U + W$: Ist $v \in V$, dann gibt es $\lambda_1, \ldots, \lambda_n \in k$ mit $v = \sum_{i=1}^{n} \lambda_i v_i = a + b$, wobei $a = \sum_{i=1}^{r} \lambda_i v_i \in U$ und $b = \sum_{i=r+1}^{n} \lambda_i v_i \in W$. Also $v \in U + W$.

Summe direkt: Ist $v \in U \cap W$, dann gibt es $\mu_1, \ldots, \mu_r$ und $\nu_{r+1}, \ldots, \nu_n \in k$ mit $v = \sum_{i=1}^{r} \mu_i v_i = \sum_{i=r+1}^{n} \nu_i v_i$. Es ist also $\sum_{i=1}^{n} \lambda_i v_i = 0$, wobei $\lambda_i = \mu_i$ für $i \leq r$ und $\lambda_i = -\nu_i$ für $i \geq r+1$. Da $v_1, \ldots, v_n$ linear unabhängig ist, müssen also alle μ_i und ν_i gleich Null sein, also $v = 0$.

Dimension: Folgt aus Korollar 3.14 ∎

Beispiel Sowohl $\text{Spann}(e_2)$ als auch $\text{Spann}((1,1))$ ist ein Komplement in $\mathbb{R}^2$ von $\text{Spann}(e_1)$.

Beispiel Wählt man einen Ursprung O, so wird der Anschauungsraum zu einem dreidimensionalen $\mathbb{R}$-Vektorraum. Die eindimensionalen Unterräume sind die Geraden durch O, und die zweidimensionale Unterräume sind die Ebenen, die O enthalten. Zwei Geraden G_1, G_2 durch O spannen eine Ebene E auf: in der Sprache dieser Vorlesung gilt dann $G_1 + G_2 = E$. Ist E eine Ebene, die O enthält, und ist G eine Gerade durch O, die nicht in der Ebene liegt, dann ist die Summe $G + E$ der ganze Anschauungsraum. In diesem Fall ist E ein Komplement von G und auch umgekehrt.

Im Anschauungsraum hat die Ebene E ein bevorzugtes Komplement, nämlich die Gerade, die senkrecht auf dieser Ebene steht. Erst im Kapitel über euklidischen Räumen werden wir uns mit senkrechten Vektoren beschäftigen. Bis dahin gelten alle Komplemente eines Unterraums als gleichberechtigt.

4 Lineare Abbildungen und deren Matrizen

Definition Seien V, W zwei k-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt eine *lineare Abbildung*, falls für alle $a, b \in V$ und für alle $\lambda, \mu \in k$ gilt: $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$.

Hilfssatz 6 *a) Es ist $f(0_V) = 0_W$, $f(a+b) = f(a) + f(b)$, $f(\lambda a) = \lambda f(a)$ und $f(-a) = -f(a)$ für $a, b \in V$ und $\lambda \in k$.*

b) Für alle $v_1, \ldots, v_n$ in V und für alle $\lambda_1, \ldots, \lambda_n \in k$ gilt

$$f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i f(v_i)\,.$$

c) Mit $f\colon V \to W$ und $g \to U$ ist auch $g \circ f\colon V \to U$ linear. Auch die Identitätsabbildung $\mathrm{Id}_V\colon V \to V$, $v \mapsto v$ ist linear.

Beweis. a) $f(a+b) = f(1 \cdot a + 1 \cdot b) = 1 \cdot f(a) + 1 \cdot f(b) = f(a) + f(b)$. $f(\lambda a) = f(\lambda a + 0_k \cdot 0_V) = \lambda f(a) + 0_k \cdot f(0_V) = \lambda f(a) + 0_W = \lambda f(a)$. $f(0_V) = f(0_k \cdot 0_V) = 0_k \cdot f(0_V) = 0_W$. $f(-a) = f((-1)a) = (-1)f(a) = -f(a)$.

b) Induktion über n. Die Fälle $n = 0, 1, 2$ folgen aus a). Induktionsschritt wie im Beweis von Hilfssatz 4 auf S. 24.

c) Es ist $g(f(\lambda a + \mu b)) = g(\lambda f(a) + \mu f(b)) = \lambda g(f(a)) + \mu g(f(b))$. ∎

Lemma 4.1 *Ist $f\colon V \to W$ bijektiv und linear, so ist auch $f^{-1}\colon W \to V$ linear. Eine bijektive lineare Abbildung heißt ein (linearer)* Isomorphismus.

Beweis. Die Abbildung f^{-1} ist auf jedem Fall bijektiv. Seien $c, d \in W$ und $\lambda, \mu \in k$. Sei $a = f^{-1}(c)$, $b = f^{-1}(d)$. Dann $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b) = \lambda c + \mu d$, also $f^{-1}(\lambda c + \mu d) = \lambda a + \mu b = \lambda f^{-1}(c) + \mu f^{-1}(d)$. ∎

Beispiele a) Sei P_n der $\mathbb{R}$-Vektorraum aller Polynome von Grad $\leq n$. Dann ist die Ableitung $D\colon P_n \to P_{n-1}$, $f(x) \mapsto f'(x)$ eine lineare Abbildung.

b) Für $a \in \mathbb{R}$ sei $\Phi_a\colon C^0(\mathbb{R}) \to \mathbb{R}$ die Abbildung, die eine stetige Funktion $f(x)$ in $x = a$ auswertet: $\Phi_a(f) := f(a)$. Diese Abbildung Φ_a ist linear.

c) Sei $U \subseteq k^3$ der Lösungsraum der Gleichung $x_1 + x_2 + x_3 = 0$. Sei $f\colon k^2 \to U$ die Abbildung $f(a, b) = (a, b - a, -b)$. Diese Abbildung f ist linear und bijektiv, d.h. f ist ein Isomorphismus.

d) *(Wichtig)* Sei $A \in M(m \times n, k)$ eine Matrix. Fassen wir Elemente $v \in k^n$ als Spaltenvektoren d.h. als $(n \times 1)$-Matrizen auf, so können wir das Produkt $A \cdot v$ bilden. Das Ergebnis ist eine $(m \times 1)$-Matrix, d.h. ein als Spaltenvektor aufgefasstes Element von k^m. Somit erhalten wir eine Abbildung $L_A \colon k^n \to k^m$, $v \mapsto A \cdot v$. Diese Abbildung ist linear. Die Abbildung $k^2 \to k^3$ in c) ist L_A für $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}$.

Bezeichnung Gibt es einen Isomorphismus $f \colon V \to W$, so heißen V, W isomorph. Bezeichnung: $V \cong W$. Zusammen mit dem Hilfssatz bedeutet Lemma 4.1, dass Isomorphie einen Äquivalenzrelation auf der Klasse von Vektorräumen ist.

Lemma 4.2 *Isomorphe Vektorräume haben die gleiche Dimension.*

Beweis. Sei $f \colon V \to W$ ein Isomorphismus und $a_1, \ldots, a_n$ eine Basis von V. Für $1 \leq i \leq n$ setze $b_i := f(a_i) \in W$. Behauptung: $b_1, \ldots, b_n$ ist eine Basis von W.

Erzeugendensystem: Ist $w \in W$, dann f surjektiv $\Rightarrow$ es gibt ein $v \in V$ mit $f(v) = w$. Da $a_1, \ldots, a_n$ eine Basis von V ist, gibt es $\lambda_1, \ldots, \lambda_n \in k$ mit $\sum_{i=1}^n \lambda_i a_i = v$. Nach dem Hilfssatz ist also $\sum_{i=1}^n \lambda_i b_i = w$.

Linear unabhängig: Ist $\sum_{i=1}^n \lambda_i b_i = 0_W$, dann $f \colon \sum_{i=1}^n \lambda_i a_i \mapsto 0_W = f(0_V)$. Wegen Injektivität folgt also $\sum_{i=1}^n \lambda_i a_i = 0_V$. Da die a_i linear unabhängig sind, folgt $\lambda_i = 0$ für alle i.

Dieses Argument lässt sich ohne große Schwierigkeiten an dem Fall anpassen, wo die Basis beliebig groß sein kann. ∎

Definition Sei $f \colon V \to W$ eine k-lineare Abbildung. Der Kern und das Bild von f werden so definiert:

$$\begin{aligned} \mathrm{Kern}(f) &= \{v \in V \mid f(v) = 0_W\} \subseteq V \\ \mathrm{Bild}(f) &= \{w \in W \mid \exists v \in V \text{ mit } f(v) = w\} \subseteq W\,. \end{aligned}$$

Lemma 4.3 *Ist $f \colon V \to W$ eine k-lineare Abbildung, so ist* $\mathrm{Kern}(f)$ *ein Unterraum von V, und* $\mathrm{Bild}(f)$ *ist ein Unterraum von W.*

Die Abbildung f ist genau dann injektiv bzw. surjektiv, wenn $\mathrm{Kern}(f) = \{0_V\}$ *bzw.* $\mathrm{Bild}(f) = W$ *gilt.*

Beweis. Da $f(0_V) = 0_W$ gelten $0_V \in \mathrm{Kern}(f)$ und $0_W \in \mathrm{Bild}(f)$, somit sind Kern und Bild nicht leer. Liegen v_1, v_2 im Kern, so gilt für $\lambda_1, \lambda_2 \in k$: es ist $f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 \cdot 0 + \lambda_2 \cdot 0 = 0$, weshalb $\lambda_1 v_1 + \lambda_2 v_2$ auch im Kern liegt. Liegen w_1, w_2 im Bild, so gibt es $v_1, v_2 \in V$ mit $f(v_1) = w_1$, $f(v_2) = w_2$. Dann gilt

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 w_1 + \lambda_2 w_2\,,$$

weshalb auch $\lambda_1 w_1 + \lambda_2 v_2$ im Bild liegt. Nach Lemma 3.2 sind Kern und Bild Unterräume.

Dass Surjektivität gleichbedeutend mit $\mathrm{Bild}(f) = W$ ist, ist klar. Ist f injektiv, so folgt $v = 0_V$ aus $f(v) = 0_W = f(0_V)$. Ist $\mathrm{Kern}(f) = \{0_V\}$ und $f(v_1) = f(v_2)$, dann $f(v_2 - v_1) = 0_W$, weshalb $v_2 - v_1 \in \mathrm{Kern}(f)$, d.h. $v_1 = v_2$. ∎

Dimensionsformel für lineare Abbildungen *Seien V und W zwei k-Vektorräume und $f\colon V \to W$ eine lineare Abbildung. Ferner sei V endlich dimensional. Dann ist* $\mathrm{Bild}(f)$ *endlich dimensional, und es gilt*

$$\dim(V) = \dim \mathrm{Kern}(f) + \dim \mathrm{Bild}(f)\,.$$

Bezeichnung Man nennt $\dim \mathrm{Bild}(f)$ den *Rang* von f.

Bemerkung Beachten Sie, dass $\dim(W)$ nicht in der Dimensionsformel vorkommt. Manche Fehler entstehen dadurch, dass man die Rollen von $\dim(V)$ und $\dim(W)$ vertauscht. Man sollte aber doch beachten, dass $\dim \mathrm{Bild}(f) \leq \dim(W)$ gelten muss.

Beweismethode Nr. 1. Sei $U \subseteq V$ ein Komplement von $\mathrm{Kern}(f)$. Nach der anderen Dimensionsformel gilt $\dim(V) = \dim \mathrm{Kern}(f) + \dim(U)$. Wir müssen also zeigen, es ist $\dim \mathrm{Bild}(f) = \dim(U)$.

Betrachten wir hierfür die Enschränkunng $f|_U$. Ist $f(u) = 0_W$ für $u \in U$, dann $u \in U \cap \mathrm{Kern}(f) = \{0_V\}$, also $u = 0_V$. Somit ist $f|_U$ injektiv. Außerdem ist $f(U) = \mathrm{Bild}(f)$: für jedes $v \in V$ ist $v = v' + u$ für $v' \in \mathrm{Kern}(f)$, $u \in U$, also $f(v) = f(u) \in f(U)$. Somit ist $f|_U\colon U \to \mathrm{Bild}(f)$ eine bijektive lineare Abbildung, d.h. ein Isomorphismus. Nach Lemma 4.2 gilt also $\dim(U) = \dim \mathrm{Bild}(f)$. ∎

Beweismethode Nr. 2. Bild endlich erzeugt: Ist $v_1, \ldots, v_n$ eine Basis von V, so ist $f(v_1), \ldots, f(v_n)$ ein Erzeugendensystem von $\mathrm{Bild}(f)$.

Dimensionsformel: Sei $a_1, \ldots, a_r$ eine Basis von $\mathrm{Bild}(f)$. Dann gibt es Elemente $b_1, \ldots, b_r \in V$ mit $f(b_i) = a_i$ für jedes i. Sei $c_1, \ldots, c_s$ eine Basis von $\mathrm{Kern}(f)$. Wir werden zeigen, dass $b_1, \ldots, b_r, c_1, \ldots, c_s$ eine Basis von V ist. Daraus folgt, dass $\dim(V) = r + s = \dim \mathrm{Bild}(f) + \dim \mathrm{Kern}(f)$.

Erzeugendensystem: Ist $v \in V$, so ist $f(v)$ eine Linearkombination der a_i. Sei $v' \in V$ die entsprechende Linearkombination der b_i, dann $f(v') = f(v)$. Also liegt $v - v'$ im Kern und ist deshalb eine Linearkombination der c_j. Also ist $v = v' + (v - v')$ eine Linearkombination der b_i und c_j.

Linear unabhängig: Ist $\sum_{i=1}^{r} \lambda_i b_i + \sum_{j=1}^{s} \mu_j c_j = 0_V$, dann wenden wir f auf beide Seiten an und erhalten $\sum_{i=1}^{r} \lambda_i a_i = 0$. Da die a_i linear unabhängig sind, folgt $\lambda_i = 0$ für alle i, weshalb $\sum_{j=1}^{s} \mu_j c_j = 0$. Da auch die c_j linear unabhängig sind, folgt $\mu_j = 0$. ∎

Satz von der linearen Fortsetzung *Sei $b_1, \dots, b_n$ eine Basis des k-Vektorraums V. Seien $w_1, \dots, w_n$ beliebige Elemente des k-Vektorraums W. Dann gibt es genau eine lineare Abbildung $f\colon V \to W$, die $f(b_i) = w_i$ für jedes $1 \leq i \leq n$ erfüllt.*

Anders gesagt: jede Abbildung von der Menge $\{b_1, \dots, b_n\}$ nach W lässt sich auf genau einer Weise zu einer linearen Abbildung von V nach W fortsetzen.

Beweis. Zuerst halten wir fest, dass jedes $v \in V$ sich auf genau einer Weise als eine Linearkombination $v = \sum_{i=1}^n \lambda_i b_i$ der Basiselemente schreiben lässt. Existenz: Basen sind Erzeugendensysteme. Eindeutigkeit: Gilt auch $v = \sum_{i=1}^n \mu_i b_i$, dann $\sum_{i=1}^n (\mu_i - \lambda_i) b_i = 0$ wegen linearer Unabhängigkeit, weshalb $\lambda_i = \mu_i$.

Eindeutigkeit der Fortsetzung: Ist f eine solche lineare Abbildung und $v = \sum_{i=1}^n \lambda_i b_i \in V$, dann $f(v) = \sum_{i=1}^n \lambda_i f(b_i) = \sum_{i=1}^n \lambda_i w_i$. Es gibt also höchstens eine solche Abbildung f.

Existenz der Fortsetzung: Wir müssen also zeigen, dass das einzige Kandidat $f(\sum_{i=1}^n \lambda_i b_i) := \sum_{i=1}^n \lambda_i w_i$ repräsentantenunabhängig und linear ist. Repräsentantenunabhängig ist es, da jedes $v \in V$ sich auf nur einer Weise als eine Linearkombination der Basis darstellen lässt. Linearität: Ist $v = \sum_{i=1}^n \lambda_i b_i$ und $v' = \sum_{i=1}^n \lambda_i' b_i$, dann $\mu v + \mu' v' = \sum_{i=1}^n (\mu\lambda_i + \mu'\lambda_i') b_i$, weshalb

$$f(\mu v + \mu' v') = \sum_{i=1}^n (\mu\lambda_i + \mu'\lambda_i') w_i = \mu f(v) + \mu' f(v')\,. \qquad \blacksquare$$

Korollar 4.4 *Zwei endlich-dimensionale k-Vektorräume sind genau dann isomorph, wenn Sie die gleiche Dimension haben. Insbesondere ist jeder n-dimensionale k-Vektorraum zu k^n isomorph.*

Beweis. In Lemma 4.2 sahen wir, dass isomorphe Vektorräume die gleiche Dimension haben. Nun seien V, W zwei n-dimensionale k-Vektorräume. Sei $a_1 \dots, a_n$ bzw. $b_1, \dots, b_n$ eine Basis von V bzw. von W. Nach dem Satz von der linearen Fortsetzung gibt es lineare Abbildungen $f\colon V \to W$ und $g \to W \to V$ mit $f(a_i) = b_i$ und $g(b_i) = a_i$ für alle $1 \leq i \leq n$. Es ist also $g(f(a_i)) = a_i$ und $f(g(b_i)) = b_i$, weshalb die Eindeutigkeitsaussage desselben Satzes bedeutet, dass $g \circ f = \mathrm{Id}_V$ und $f \circ g = \mathrm{Id}_W$ gelten. Das heißt, f ist bijektiv und $g = f^{-1}$. $\blacksquare$

Korollar 4.5 *Ist V ein endlich dimensionaler k-Vektorraum und $U \subseteq V$ ein Unterraum, so gibt es einen endlich dimensionalen Vektorraum W und eine surjektive lineare Abbildung $f\colon V \to W$ mit $\mathrm{Kern}(f) = U$.*

Anders gesagt: Jeder Kern ist ein Unterraum, und auch umgekehrt.

Beweis. Nach Satz 3.12 ist U endlich dimensional. Sei $v_1, \dots, v_r$ eine Basis von U. Nach dem Basisergänzungssatz können wir dieses linear unabhängiges System zu einer Basis $v_1, \dots, v_n$ von V fortsetzen. Sei $\phi\colon V \to V$ die lineare

Abbildung mit $\phi(v_i) = 0$ für alle $i \leq r$ und $\phi(v_i) = v_i$ für alle $i \geq r+1$. Für jedes $v = \sum_{i=1}^{n} \lambda_i v_i \in V$ ist dann $\phi(v) = \sum_{i=r+1}^{n} \lambda_i v_i$. Dies ist genau dann 0, wenn $\lambda_i = 0$ ist für jedes $i \geq r+1$, d.h. wenn $v = \sum_{i=1}^{r} \lambda_i v_i$, d.h. wenn $v \in U$. Es ist also $U = \mathrm{Kern}(\phi)$. Nun sei $W = \mathrm{Bild}(\phi) \subseteq V$ und $f\colon V \to W$ die Abbildung $f(v) = \phi(v)$. Dann ist f surjektiv und $\mathrm{Kern}(f) = \mathrm{Kern}(\phi) = U$.

Übrigens: Für jedes v_i gilt $\phi(\phi(v_i)) = \phi(v_i)$. Der Satz von der linearen Fortsetzung besagt also, dass $\phi \circ \phi = \phi$ gilt. Eine solche lineare Abbildung $\phi\colon V \to V$ nennt man eine *Projektion*. ∎

Die Matrix einer linearen Abbildung

Zur Erinnerung: Sei k ein Körper und $A \in M(m \times n, k)$ eine Matrix. Fassen wir Elemente $v \in k^n$ als Spaltenvektoren d.h. als $(n \times 1)$-Matrizen auf, so können wir das Produkt $A \cdot v$ bilden. Das Ergebnis ist eine $(m \times 1)$-Matrix, d.h. ein als Spaltenvektor aufgefasstes Element von k^m. Somit erhalten wir eine Abbildung $L_A\colon k^n \to k^m$, $v \mapsto A \cdot v$. Diese Abbildung ist linear, und es gilt $L_{AB} = L_A \circ L_B$ (vgl. Übungsserie 8). Ist $A = (a_{ij})_{\substack{i \leq m \\ j \leq n}}$ und $v = (v_1, \ldots, v_n)$, so ist $L_A(v) = (w_1, \ldots, w_m)$ mit $w_i = \sum_{j=1}^{n} a_{ij} v_j$.

Beispiel Die Matrix $A = \left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{smallmatrix}\right) \in M(3 \times 2, k)$ induziert $L_A\colon k^2 \to k^3$. Es ist

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -x+y \\ -y \end{pmatrix},$$

weshalb $L_A(x, y) = (x, y - x, -y)$.

Die Matrix $B = \left(\begin{smallmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{smallmatrix}\right) \in M(2 \times 3, k)$ induziert $L_B\colon k^3 \to k^2$. Es ist

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x-y \\ y-z \end{pmatrix},$$

weshalb $L_B(x, y, z) = (x - y, y - z)$.

Bezeichnung Sind V, W k-Vektorräume, so bezeichnet man mit $L(V, W)$ die Menge der k-linearen Abbildungen $f\colon V \to W$.

Lemma 4.6 *Sei k ein Körper.*

a) Jede lineare Abbildung $f\colon k^n \to k^m$ hat die Gestalt $f = L_A$ für A eine Matrix in $M(m \times n, k)$.

b) Die Mengen $M(m \times n, k)$ und $L(V, W)$ – für beliebige k-Vektorräume V, W – haben natürliche Vektorraumstrukturen.

c) *Die Zuordnung $A \mapsto L_A$ ist ein Isomorphismus von $M(m \times n, k)$ nach $L(k^n, k^m)$.*

Beweis. a) Für die Standardbasisvektoren $e_1, \ldots, e_n \in k^n$ gilt:

$$L_A(e_j) = j\text{te Spalte von } A.$$

Wir benutzen diese Eigenschaft, um A zu konstruieren.

Wir definieren also Zahlen a_{ij} für $1 \leq i \leq m$ und $1 \leq n$ so: a_{ij} ist der Koeffizient von e_i in $f(e_j)$. Anders gesagt: $f(e_j) = \sum_{i=1}^{m} a_{ij} e_i$. Dann aber gilt auch $L_A(e_j) = \sum_{i=1}^{m} a_{ij} e_i$, d.h. $L_A(e_j) = f(e_j)$ für alle $1 \leq j \leq n$. Nach dem Satz von der linearen Fortsetzung gilt $L_A = f$.

b) Matrizen: Addition ist Matrixaddition, und Skalarmultiplikation wird definiert durch $(\lambda A)_{ij} = \lambda \cdot A_{ij}$, d.h. jeder Eintrag der Matrix A wird mit dem Skalar λ multipliziert.

Lin. Abbildungen: Addition und Skalarmultiplikation definiert man punktweise, d.h. $f+g$ wird durch $(f+g)(v) = f(v)+g(v)$ für alle $v \in V$ definiert; und $(\lambda f)(v) = \lambda \cdot f(v)$ für alle $v \in V$ definiert λf.

Die Vektorraumaxiome lassen sich nachweisen.

c) In a) sahen wir, dass die Abbildung $A \mapsto L_A$ surjektiv ist.

Linearität: Sei $v = (v_1, \ldots, v_n) \in k^n$. Dann $L_{\lambda A + \mu B}(v) = (w_1, \ldots, w_m)$ für

$$w_i = \sum_{j=1}^{n} (\lambda A + \mu B)_{ij} v_j = \sum_{j=1}^{n} (\lambda A_{ij} + \mu B_{ij}) v_j = \lambda \sum_{j=1}^{n} A_{ij} v_j + \mu \sum_{j=1}^{n} B_{ij} v_j$$
$$= i\text{te Komponente von } \lambda L_A(v) + \mu L_B(v).$$

Das heißt, $L_{\lambda A+\mu B}(v) = \lambda L_A(v) + \mu L_B(v) = (\lambda L_A + \mu L_B)(v)$.

Injektiv: Wegen Lemma 4.3 reicht es, zu zeigen: Ist L_A die Nullabbildung, so ist A die Nullmatrix. Dies ist tatsächlich der Fall, denn $A_{ij} = i$te Komponente in $L_A(e_j)$.

■

Definition Sei $f\colon V \to W$ eine k-lineare Abbildung. Sei $b_1, \ldots, b_n$ eine Basis von V und $c_1, \ldots, c_m$ eine Basis von W. Da jedes $w \in W$ sich auf genau einer Weise als eine Linearkombination der Basis $c_1, \ldots, c_m$ ausdrücken lässt, können wir eine Matrix $A \in M(m \times n, k)$ definieren durch

$$f(b_j) = \sum_{i=1}^{m} A_{ij} c_i \quad \text{für alle } 1 \leq j \leq n. \tag{*}$$

Diese Matrix A nennt man die Matrix der linearen Abbildung f bezüglich den gewählten Basen von V und W.

Bezeichnung Diese Matrix werden wir mit ${}_B\mathrm{M}_C(f)$ bezeichnen, wobei B die Basis $b_1, \ldots, b_n$ von V und C die Basis $c_1, \ldots, c_m$ von W bezeichnet.

In Lemma 4.7 leiten wir eine nützliche Alternativbeschreibung von ${}_B\mathrm{M}_C(f)$ her. Zuerst benötigen wir den Begriff Koordinatenvektor.

Bezeichnung Ist $v \in V$, so lässt sich V auf genau einer Weise als eine Linearkombination $v = \sum_{j=1}^{n} \lambda_j b_j$ schreiben. Wir nennen $\underline{\lambda} = (\lambda_1, \ldots, \lambda_m) \in k^n$ den *Koordinatenvektor* von $v \in V$ bezüglich der Basis B.

Beachten Sie, dass die Abbildung $v \mapsto \underline{\lambda}$ ist der Isomorphismus $V \to k^n$ aus Korollar 4.4, den man durch Wahl der Basis B für V und der Standardbasis für k^n erhält.

Lemma 4.7 *a)* ${}_B\mathrm{M}_C(f)$ *ist die einzige Matrix* $A \in M(m \times n, k)$ *mit der folgenden Eigenschaft:*

> *Ist* $\underline{\lambda} = (\lambda_1, \ldots, \lambda_n) \in k^n$ *der* Koordinatenvektor *von* $v \in V$ *bezüglich der Basis* B*, dann ist* $L_A(\underline{\lambda})$ *der Koordinatenvektor von* $f(v) \in W$ *bezüglich der Basis* C *von* W*.*

b) Insbesondere gilt dann ${}_C\mathrm{M}_D(g) \cdot {}_B\mathrm{M}_C(f) = {}_B\mathrm{M}_D(g \circ f)$*.*

c) $f \mapsto {}_B\mathrm{M}_C(f)$ *ist ein Isomorphismus* $L(V, W) \to M(m \times n, k)$*.*

Beweis. a) Bezeichnen wir mit $h_B \colon V \to k^n$ den Isomorphismus, der durch $h_B(b_j) = e_j$ für $1 \leq j \leq n$ gegeben wird. Es ist dann $h_B(v) =$ Koordinatenvektor von v – vgl. den Beweis von Korollar 4.4. Sei $g \colon k^n \to k^m$ die Abbildung

$$g\colon \text{Koordinatenvektor von } v \longmapsto \text{Koordinatenvektor von } f(v).$$

Dann

$$g = h_C \circ f \circ h_B^{-1}\,. \tag{**}$$

Also g ist linear und es gibt nach Lemma 4.6 genau eine Matrix A mit $g = L_A$. Mit $A = {}_B\mathrm{M}_C(f)$ gilt auf jedem Fall $L_A(e_j) = (A_{1j}, A_{2j}, \ldots, A_{mj})$. Nach (*) ist dies der Koordinatenvektor von $f(e_j)$. Also $L_A(e_j) = g(e_j)$ für alle $1 \leq j \leq n$. Also $L_A = g$.

b) Setze $A(f) = {}_B\mathrm{M}_C(f)$ und $A(g) = {}_C\mathrm{M}_D(g)$. Sei $\underline{\lambda} \in k^n$ der Koordinatenvektor von $v \in V$. Dann ist $L_{A(f)}(\underline{\lambda})$ der Koordinatenvektor von $f(v)$, und $L_{A(g)}L_{A(f)}(\underline{\lambda})$ der Koordinatenvektor von $g(f(v))$. Also ist $L_{A(g)A(f)}(\underline{\lambda})$ der Koordinatenvektor von $g(f(v))$. Nach der Eindeutigeitsaussage gilt dann $A(g)A(f) = {}_B\mathrm{M}_D(g \circ f)$.

c) In Anlehnung an (**) definieren wir $\Phi\colon L(V,W) \to L(k^n, k^m)$ durch $\Phi(f) = h_C \circ f \circ h_B^{-1}$. Sei $\Psi\colon L(k^n, k^m) \to M(m \times n, k)$ die Umkehrabbildung des Isomorphismus $A \mapsto L_A$. Nach Konstruktion gilt ${}_B\mathrm{M}_C(f) = \Psi(\Phi(f))$. Wir wissen aus Lemma 4.6, dass Ψ ein Isomorphismus ist. Es reicht also zu zeigen, dass Φ ein Isomorphismus ist.

Φ ist bijektiv, denn $\Phi^{-1}(g) = h_C^{-1} \circ g \circ h_B$. Für Linearität von Φ sei $\underline{\lambda} \in k^n$, der Koordinatenvektor bezüglich Basis B des Vektors $v := h_B^{-1}(\underline{\lambda})$. Seien $f, f' \in L(V,W)$ und $\mu, \mu' \in k$. Sei $g = \mu f + \mu' f' \in L(V,W)$. Dann

$$\Phi(g)(\underline{\lambda}) = h_C(g(h_B^{-1}(\underline{\lambda}))) = h_C(g(v)) = h_C(\mu f(v) + \mu' f'(v))\,.$$

Da h_C linear ist, folgt

$$\Phi(g)(\underline{\lambda}) = \mu h_C(f(v)) + \mu' h_C(f'(v)) = \mu\Phi(f)(\underline{\lambda}) + \mu'\Phi(f')(\underline{\lambda})\,,$$

also $\Phi(\mu f + \mu' f') = \mu\Phi(f) + \mu'\Phi(f')$.

■

5 Lineare Gleichungssysteme

Einleitung Um das Gleichungssystem

$$x + y + z = 1 \tag{I}$$
$$x + 2y + 2z = 2 \tag{II}$$
$$2x + y + z = 1 \tag{III}$$

zu lösen, können wir (I) benutzen, um x aus (II) und (III) zu eliminieren: wir setzen $x = 1 - y - z$ in diesen beiden Gleichungen ein und erhalten

$$x + y + z = 1 \tag{I}$$
$$y + z = 1 \tag{II$'$}$$
$$-y - z = -1 \tag{III$'$}$$

Da die zweite und dritte Gleichungen jetzt äquivalent sind, können wir die dritte streichen. Das Gleichungssystem lautet jetzt

$$x + y + z = 1 \tag{I}$$
$$y + z = 1 \tag{II$'$}$$

Wir erkennen jetzt, dass der Wert von z frei wählbar ist. Nach der Wahl von z ist der Wert von y eindeutig durch (II$'$) bestimmt; und sind y, z bekannt, so ist der Wert von x eindeutig durch (I) angegeben. Man sagt, dass die Lösungen durch den Wert von z parametrisiert sind. Sei t der Wert, den wir für z wählen. Dann $y = 1 - z$ und $x = 1 - y - z = 0$. Die Lösungsmenge ist also $\{(0, 1 - t, t) \mid t \in \mathbb{R}\}$.

Wir hätten auch (II$'$) benutzen können, um y aus (I) zu eliminieren: wählen wir diesen Weg, so setzen wir $y = 1 - z$ in (I) ein und erhalten das Gleichungssystem

$$x = 0 \tag{I$'$}$$
$$y + z = 1 \tag{II$'$}$$

Von dieser Form des Gleichungssystems ausgehend, lässt sich die Lösungsmenge noch schneller hinschreiben.

In diesem Kapitel beschreiben wir den Gauß-Algorithmus, auch Gauß-Eliminationsverfahren genannt, der diesen Lösungsansatz systematisiert.

Lineare Gleichungssysteme

Sei k ein Körper. Unter ein lineares Gleichungssystem versteht man ein System

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array}$$

von m Gleichungen, in dem die Skalare $a_{ij}, b_i \in k$ alle bekannt sind. Es gilt, das Gleichungssystem für die n unbekannte Elemente $x_1, \ldots, x_n$ von k zu lösen.

Sei $A \in M(m \times n, k)$ die Matrix mit $A_{ij} = a_{ij}$ für alle i, j. Sei $b \in k^m$ der Spaltenvektor $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$. Sei $x \in k^n$ der Spaltenvektor $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, der sich aus den Unbekannten zusammensetzt. Dass obige Gleichungssystem lässt sich dann durch die folgende Vektorengleichung ausdrücken:

$$A \cdot x = b\,.$$

Aufgabe 5.1 *Sei k ein Körper, $A \in M(m \times n, k)$ eine Matrix und $b \in k^m$ ein (Spalten-)vektor. Man bestimme die Lösungsmenge*

$$\mathrm{LR}(A, b) := \{x \in k^n \mid A \cdot x = b\}$$

des „inhomogenen“ linearen Gleichungssystems $A \cdot x = b$.

Ein wichtiger Spezialfall ist das *homogene* Gleichungssystem $A \cdot x = 0$, d.h. der Fall $b =$ Nullvektor. In diesem Fall ist die Lösungsmenge $\mathrm{LR}(A, 0)$ der Kern der linearen Abbildung $L_A \colon k^n \to k^m$ und somit ein Unterraum des k^n.

Aufgabe 5.2 *Sei k ein Körper und $A \in M(m \times n, k)$ eine Matrix. Man bestimme eine Basis des Lösungsraums*

$$\mathrm{LR}(A, 0) := \{x \in k^n \mid A \cdot x = 0\}$$

des homogenen linearen Gleichungssystems $A \cdot x = 0$. Diesen Raum nennt man auch den Nullraum *von A.*

Wir beginnen mit Aufgabe 5.2. Das Ergebnis und die Methoden, die wir hierfür entwickeln, werden auch eine Lösung von Aufgabe 5.1 ermöglichen.

Elementare Zeilenoperationen

Definition Sei $A \in M(m \times n, k)$ eine Matrix. Die folgenden drei Wege, um aus A eine neue Matrix in $M(m \times n, k)$ zu machen, heißen elementare Zeilenoperationen.

Typ 1 Zwei Zeilen vertauschen.

Typ 2 Eine Zeile mit $\lambda \in k^\times := k \setminus \{0\}$

Typ 3 Das λ-fache von einer Zeile zu einer anderen Zeile addieren ($\lambda \in k$). multiplizieren.

Lemma 5.3 *a) Ändert man die Matrix A durch eine oder sogar durch eine Kette von elementaren Zeilenoperationen, so ändert sich der Nullraum* $\mathrm{LR}(A,0)$ *nicht. Auch wenn man Nullzeilen streicht oder hinzufügt, so ändert sich der Nullraum nicht.*

b) Jede elementare Zeilenoperation ist durch eine weitere elementare Zeilenoperation rückgängig zu machen. Jede elementare Zeilenoperation lässt sich durch Linksmultiplikation $A \mapsto H \cdot A$ durch eine invertierbare Matrix $H \in M_m(k) = M(m \times m, k)$ bewirken.

Zur Erinnerung: eine quadratische Matrix $H \in M_m(k)$ heißt genau dann invertierbar, wenn es eine Matrix $H' \in M_m(k)$ mit $HH' = H'H = E_m$ gibt, wobei E_m das Einselement des Rings $M_m(k)$ ist, d.h. E_m ist die Einheitsmatrix

$$(E_m)_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} .$$

Beweis. b) Für jede der drei Operationstypen definieren wir eine Matrix aus $M_m(k)$. Für $1 \le r, s \le m$ mit $r \neq s$ definieren wir $T = T(r,s)$ durch $T_{rs} = T_{sr} = 1$, $T_{ii} = 1$ für $i \neq r, s$, und $T_{ij} = 0$ sonst. Für $1 \le r \le m$ und $0 \neq \lambda \in k$ sei $M = M(r,\lambda)$ die Matrix mit $M_{rr} = \lambda$, $M_{ii} = 1$ für $i \neq r$ und $M_{ij} = 0$ für $i \neq j$. Für $r \neq s$ und $\lambda \in k$ definieren wir $F = F(r,\lambda,s)$ durch $F_{sr} = \lambda$, $F_{ii} = 1$ für alle i, und $F_{ij} = 0$ sonst. In $M_4(\mathbb{R})$ haben wir zum Beispiel

$$T(1,4) = \begin{pmatrix} 0&0&0&1 \\ 0&1&0&0 \\ 0&0&1&0 \\ 1&0&0&0 \end{pmatrix} \qquad M(3,\tfrac{1}{2}) = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&\frac{1}{2}&0 \\ 0&0&0&1 \end{pmatrix} \qquad F(2,-2,4) = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&-2&0&1 \end{pmatrix}$$

Man rechnet jetzt nach, dass $T(r,s)A$ bzw. $M(r,\lambda)A$ bzw. $F(r,\lambda,s)A$ das Ergebnis der entsprechenden elementaren Zeilenoperation vom Typ 1 bzw. Typ 2 bzw. Typ 3 ist. Wir rechnen Beispielhaft für Typ 3 nach. Sei $F = F(r,\lambda,s)$. Es ist $F_{ij} = \delta_{ij} + \lambda\delta_{is}\delta_{jr}$. Also

$$(FA)_{ij} = \sum_{a=1}^{m} F_{ia}A_{aj} = \sum_{a=1}^{m} \delta_{ia}A_{aj} + \sum_{a=1}^{m} \lambda\delta_{is}\delta_{ar}A_{aj} = A_{ij} + \delta_{is}\lambda A_{rj} .$$

Also FA und A stimmen außer für $i = s$ überein, d.h. außerhalb der $s$ten Zeile. Die $s$te Zeile von FA ist die $s$te Zeile von A plus das λ-fache der $r$ten Zeile von A.

Diese Matrizen bewirken also die Zeilenoperationen. Nun, $T(r,s)T(r,s) = E_m$, $M(r,\lambda)M(r,\mu) = M(r,\lambda\mu)$ und $F(r,\lambda,s)F(r,\mu,s) = F(r,\lambda+\mu,s)$: die linke Matrix im Produkt entspricht eine Zeilenoperation, man führe diese auf die rechte Matrix aus. Ferner gilt $M(r,1) = F(r,0,s) = E_m$. Somit sind die Matrizen invertierbar: $T(r,s)^{-1} = T(r,s)$, $M(r,\lambda)^{-1} = M(r,1/\lambda)$ und $F(r,\lambda,s)^{-1} = F(r,-\lambda,s)$. Auch die inverse Matrizen entsprechen Zeilenoperationen: diese machen die ursprünglichen Zeilenoperationen rückgängig.

a) Ist $A \cdot x = 0$, dann $HA \cdot x = H \cdot Ax = 0$, d.h. $\mathrm{LR}(A,0) \subseteq \mathrm{LR}(HA,0)$. Ist H invertierbar, dann $A = H^{-1} \cdot HA$, weshalb $\mathrm{LR}(HA,0) \subseteq \mathrm{LR}(A,0)$ nach dem gleichen Argument. Also $\mathrm{LR}(A,0) = \mathrm{LR}(HA,0)$. Bleibt der Nullraum bei jeder elementaren Operation erhalten, dann auch bei einer beliebigen Kette von elementaren Operationen.

Eine Nullzeile entspricht der Gleichung $0 = 0$. An der Lösungsmenge ändert sich nichts, wenn man diese Gleichung streicht bzw. hinzufügt. ∎

Definition Sei $a \in M(m \times n, k)$ eine Matrix mit m Zeilen und n Spalten. Der *Zeilenraum* ist der Unterraum des k^m, der von den Zeilen aufgespannt wird. Der *Spaltenraum* ist der Unterraum des k^n, der durch die Spalten erzeugt wird. Der *Spalten-* bzw. *Zeilenrang* ist die Dimension des Spalten- bzw. des Zeilenraums.

Später werden wir die wichtige Tatsache nachweisen, dass Zeilenrang und Spaltenrang gleich sind.

Lemma 5.4 *Sei $A \in M(m \times n, k)$ eine Matrix. Dann*

a) Der Spaltenraum stimmt mit dem Bild der linearen Abbildung $L_A\colon k^n \to k^m$, $v \mapsto A \cdot v$ überein. Somit stimmt der Spaltenrang von A mit dem Rang[8] *von L_A überein.*

b) Elementare Zeilenoperationen lassen den Zeilenraum und somit den Zeilenrang unverändert.

c) Elementare Zeilenoperationen lassen den Spaltenrang unverändert, auch wenn sie den Spaltenraum ändern.

Beweis. a) Der Spaltenraum wird von den Spalten von A erzeugt, das Bild wird von den Vektoren $L_A(e_j)$ für $1 \leq j \leq n$ erzeugt. Aber $L_A(e_j) = A \cdot e_j =$ $j$te Spalte von A.

b) Sind $z_1, \ldots, z_m \in k^n$ die Zeilen, so ist der Zeilenraum $\mathrm{Spann}(z_1, \ldots, z_n)$. Vertauscht man die Reihenfolge der Zeilen, so ändert sich nichts am Spann. Ersetzt man z_s durch $z'_s = z_s + \lambda z_r$ $(r \neq s)$, so ändert sich der Zeilenraum auch nicht: denn $z'_s = z_s + \lambda z_r$ ist eine Linearkombination von $z_1, \ldots, z_s, \ldots, z_m$, und $z_s = z'_s - \lambda z_r$ ist eine Linearkombination von $z_1, \ldots, z'_s, \ldots, z_m$. Auch wenn man z_r durch $z'_r = \lambda z_r$ für $\lambda \neq 0$ ersetzt, ändert sich der Spann nicht.

c) Wir wenden die Dimensionsformel auf $L_A\colon k^n \to k^m$ an. Wegen a) ist $\dim \mathrm{Bild}(L_A)$ der Spaltenrang. Beachten Sie, dass $\mathrm{Kern}(L_A)$ der Lösungsraum $\mathrm{LR}(A,0)$ ist. Also lautet die Dimensionsformel so:

$$\mathrm{Spaltenrang}(A) = n - \dim \mathrm{LR}(A,0)\,.$$

Nach Lemma 5.3 Teil a) bleibt aber die rechte Seite unverändert.

∎

[8] Rang einer linearen Abbildung = Dimension des Bildes

Das Gaußsche Eliminationsverfahren

Dieses Verfahren wendet Zeilenoperationen an, um eine Matrix A in einer besonderen Form zu bringen, die sich Zeilenstufenform nennt. Ist A in Zeilenstufenform, so kann man eine Basis des Lösungsraums $\mathrm{LR}(A, 0)$ mehr oder weniger ablesen. Auch lässt sich nachweisen, dass der Spaltenrang und Zeilenrang gleich sind. Hier sind vier Matrizen in Zeilenstufenform:

$$\begin{pmatrix} \underline{2} & 1 & 0 & 3 & 4 \\ 0 & 0 & \underline{3} & 1 & 1 \\ 0 & 0 & 0 & \underline{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & \underline{1} & 1 \\ 0 & 0 & \underline{1} \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} \underline{1} & 2 & 0 & 3 & 0 \\ 0 & 0 & \underline{1} & 2 & 0 \\ 0 & 0 & 0 & 0 & \underline{1} \end{pmatrix}$$

Die unterstrichene Einträge sind die sogenannte *Pivotstellen*. Nur die letzten beiden Matrizen sind in *strenger* Zeilenstufenform. Für Zeilenstufenform muss eine Matrix folgende Gestalt haben:

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

Die Pivotstellen * müssen Werte $\neq 0$ haben; die ?-Einträge können beliebige Werte haben. Für strenge Zeilenstufenform muss die Gestalt so sein:

$$\begin{pmatrix} 0 \dots 0 & 1 & ? \dots ? & 0 & ? \dots ? & 0 & ? \dots ? \\ 0 \dots 0 & 0 & 0 \dots 0 & 1 & ? \dots ? & 0 & ? \dots ? \\ \vdots & & & & & & \\ 0 \dots 0 & 0 & 0 \dots 0 & 0 & 0 \dots 0 & 1 & ? \dots ? \\ 0 \dots 0 & 0 & 0 \dots 0 & 0 & 0 \dots 0 & 0 & 0 \dots 0 \\ \vdots & & & & & & \\ 0 \dots 0 & 0 & 0 \dots 0 & 0 & 0 \dots 0 & 0 & 0 \dots 0 \end{pmatrix}$$

Definition Eine Matrix $A \in M(m \times n, k)$ ist in Zeilenstufenform, wenn es ein $0 \leq r \leq m$ und Zahlen $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ gibt mit

a) $A_{i,j_i} \neq 0$ für alle $1 \leq i \leq r$; und

b) $A_{i,j} = 0$ falls $i > r$, oder falls $i \leq r$ und $j < j_i$.

Die (i, j_i)-Einträge heißen die Pivotstellen. Für strenge Zeilenstufenform verlangt man außerdem, dass alle Pivotstellen den Wert 1 annehmen, und dass alle weitere Einträge in der Spalte einer Pivotstelle Null sind: $A_{i,j_i} = 1$ und $A_{a,j_i} = 0$ für $a \neq i$.

Algorithmus 5.5 (Das Gaußsche Eliminationsverfahren)
Input: Matrix $A \in M(m \times n, k)$.
Output: Auf A werden so lange elementare Zeilenoperationen ausgeführt, bis die Matrix in Zeilenstufenform ist.
Wahlweise kann auch die strenge Zeilenstufenform berechnet werden.

Erste Pivotstelle suchen

1) Sei j_1 die Nummer der ersten Spalte, die einen Eintrag $\neq 0$ enthält.
[Sind alle Einträge Null, so ist A bereits in strenger Zeilenstufenform.]

2) Ggf. durch Vertauschung der ersten mit einer anderen Zeile stellt man sicher, dass $A_{1,j_1} \neq 0$ ist. Dies ist die erste Pivotstelle.

3) Die erste Zeile mit $1/A_{1,j_1}$ multiplizieren, damit $A_{1,j_1} = 1$ gilt.

Unterhalb der Pivotstelle kehren

4) Für jedes $2 \leq i \leq m$ das A_{i,j_1}-fache der ersten Zeile von der $i$ten Zeile abziehen.

Nach weiteren Pivotstellen suchen

5) Sei B die Matrix, die aus den Zeilen 2 bis m von A besteht. Man bringe B auf Zeilenstufenform (Rekursion!) Jetzt ist A in Zeilenstufenform.

Ggf. strenge Zeilenstufenform berechnen.

6) Für jede Pivotstelle (i, j_i) kehre man oberhalb dieser Zeile: für jedes $1 \leq a < i$ ziehe man das A_{a,j_i}-fache der $i$ten Zeile von der $a$ten Zeile ab.

Lemma 5.6 *Mittels des Gauß-Algorithmus lässt sich jede Matrix durch eine Kette von elementaren Zeilenoperationen auf Zeilenstufenform und sogar auf strenge Zeilenstufenform bringen.*

Beweis. Jede Änderung zu A ist eine elementare Zeilenoperation. Da B eine Zeile weniger als A hat, bricht die Rekursion irgendwann ab. Nach Schritt 3) ist A_{1,j_1} der einzige Eintrag $\neq 0$ in den ersten j_1 Spalten. Somit liegen alle Pivotstellen von B in der $(j_1 + 1)$te Spalte oder später. Ist also B in Zeilenstufenform, dann A auch. Das kehren nach oben bringt dann die Matrix auf strenge Zeilenstufenform. ∎

Bemerkung Man kann sogar zeigen, dass jede Matrix genau eine strenge Zeilenstufenform hat.

Satz: Spaltenrang und Zeilenrang gleich *Der Zeilenrang und der Spaltenrang einer Matrix stimmen überein.*

Diesen gemeinsamen Wert nennt man den Rang *der Matrix. Beachten Sie: Für jede Matrix A gilt auch* $\mathrm{Rang}(A) = \mathrm{Rang}(L_A)$.

Beweis. Sei $A \in M(m \times n, k)$ eine Matrix. Wir zeigen zunächst

$$\text{Zeilenrang}(A) \leq \text{Spaltenrang}(A)\,.$$

Hieraus folgt die Gleichung Zeilenrang = Spaltenrang, denn die Ungleichung gilt auch für die durch $(A^T)_{ij} = A_{ji}$ definierte *transponierte* Matrix $A^T \in M(n \times m, k)$: und offensichtlich gelten Spaltenrang(A) = Zeilenrang(A^T), Zeilenrang(A) = Spaltenrang(A^T).

Wendet man den Gauß-Algorithmus an, so erhält man eine strenge Zeilenstufenform A' von A, d.h. A' ist in strenger Zeilenstufenform und geht aus A durch elementare Zeilenoperationen hervor. Nach Lemma 5.4 haben A, A' den gleichen Zeilenrang und den gleichen Spaltenrang. Es reicht also zu zeigen: Zeilenrang $\leq$ Spaltenrang gilt für jede Matrix in strenger Zeilenstufenform.

Sei also A eine Matrix in strenger Zeilenstufenform. Seien $A_{1,j_1}, \ldots, A_{r,j_r}$ die Pivotstellen, dann sind nur die ersten r Zeilen von A ungleich Null. Die ersten r Zeilen spannen den Zeilenraum auf. Diese Zeilen sind auch linear unabhängig: ist $\sum_{i=1}^{r} \lambda_i (i\text{te Zeile}) = 0$, dann $\sum_{i=1}^{r} \lambda_i A_{ij} = 0$ für jedes j. Setzen wir $j = j_1$, so folgt $\lambda_1 A_{1,j_1} = 0$ wegen Zeilenstufenform, also $\lambda_1 = 0$. Also $\sum_{i=2}^{r} \lambda_i A_{ij} = 0$, und $j = j_2$ führt zu $\lambda_2 = 0$, usw. Somit beträgt der Zeilenrang r. Da die j_i-te Spalte der Standardbasisvektor $e_i \in k^m$ ist für $1 \leq i \leq r$, folgt Spaltenrang$(A) \geq r =$ Zeilenrang(A).

Letzter Teil: Die Gleichheit $\mathrm{Rang}(A) = \mathrm{Rang}(L_A)$ folgt aus Lemma 5.4. ∎

Lemma 5.7 *Eine Matrix $A \in M_m(k)$ ist genau dann invertierbar, wenn ihr Rang m beträgt. Jede invertierbare Matrix lässt sich als ein Produkt von Elementarmatrizen schreiben, d.h. als ein Produkt von den Matrizen, die im Beweis des Lemmas 5.3 die elementaren Zeilenoperationen darstellen.*

Beweis. Ist A invertierbar, so ist L_A ein Isomorphismus, denn $L_{A^{-1}}$ ist die Umkehrabbildung. Also ist L_A injektiv, und Kern(L_A) hat Dimension 0. Nach der Dimensionsformel ist $\dim \mathrm{Bild}(L_A) = m$, also $\mathrm{Rang}(A) = m$.

Ist umgekehrt $\mathrm{Rang}(A) = m$, so ist L_A surjektiv, denn Bild$(L_A) \subseteq k^m$ und beide haben die gleiche Dimension. Nach der Dimensionsformel folgt außerdem $\dim \mathrm{Kern}(L_A) = 0$, also L_A ist auch injektiv, d.h. L_A ist ein Isomorphismus. Nach Lemma 4.6 Teil a) hat die Umkehrabbildung die Gestalt L_B für eine Matrix $B \in M_m(k)$. Dann $L_A B = L_A L_B = \mathrm{Id} = L_{E_m}$ und analog $L_B A = L_{E_m}$. Wegen Teil c) des gleichen Lemmas ist dann $AB = BA = E_m$, d.h. A ist invertierbar.

Sei $A \in M_m(k)$ invertierbar. Nach Lemma 5.3 und 5.6 gibt es Elementarmatrizen $H_1, \ldots, H_N \in M_m(k)$ derart, dass $A' := H_N H_{N-1} \cdots H_2 H_1 A$ eine strenge Zeilenstufenform von A ist. Da A' in Zeilenstufenform ist und

$$\text{Anzahl(Zeilen)} = \text{Anzahl(Spalten)} = \text{Rang} = m$$

gilt, muss jede Zeile und jede Spalte eine Pivotstelle enthalten, d.h. die Pivotstellen sind (i,i) für jedes $1 \leq i \leq m$. Da A' in strenger Zeilenstufenform ist, muss die $i$te Spalte dann der $i$te Standardbasisvektor e_i sein. Also $A' = E_m$. Hieraus folgt $A = H_1^{-1} H_2^{-1} \cdots H_{N-1}^{-1} H_N^{-1}$. Nach Lemma 5.3 ist auch jedes H_i^{-1} eine Elementarmatrix. ∎

Beispiel Um das Gleichungssystem

$$\begin{aligned} 2x_2 + 6x_3 - 4x_4 - 3x_5 &= 0 \\ x_1 + x_2 + 4x_3 + x_5 &= 0 \\ x_1 - x_2 - 2x_3 + 4x_4 - 2x_5 &= 0 \\ x_1 + 2x_2 + 7x_3 - 2x_4 + x_5 &= 0 \end{aligned}$$

zu lösen, fassen wir es als die Vektorengleichung

$$\begin{pmatrix} 0 & 2 & 6 & -4 & -3 \\ 1 & 1 & 4 & 0 & 1 \\ 1 & -1 & -2 & 4 & -2 \\ 1 & 2 & 7 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

auf. Nun wollen wir diese (4×5)-Matrix auf strenge Zeilenstufenform bringen. Wir vertauschen die ersten beiden Zeilen, dann liegt die erste Pivotstelle in $(1,1)$. Nachdem wir diese neue erste Zeile von den dritten und vierten Zeilen abgezogen haben, hat die Matrix die Gestalt

$$\begin{pmatrix} 1 & 1 & 4 & 0 & 1 \\ 0 & 2 & 6 & -4 & -3 \\ 0 & -2 & -6 & 4 & -3 \\ 0 & 1 & 3 & -2 & 0 \end{pmatrix}$$

Die zweite Pivotstelle liegt also bei $(2,2)$. Wir teilen die zweite Zeile durch 2, danach addieren wir sie zweimal zur dritten Zeile und ziehen sie von der vierten Zeile ab:

$$\begin{pmatrix} 1 & 1 & 4 & 0 & 1 \\ 0 & 1 & 3 & -2 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 & \frac{3}{2} \end{pmatrix}$$

Die dritte Pivotstelle liegt also bei $(3,5)$. Wir teilen die dritte Zeile durch -6, damit der Pivot-Eintrag 1 beträgt. Dann ziehen wir das $\frac{3}{2}$-fache der dritten von der vierten Zeile ab:

$$\begin{pmatrix} 1 & 1 & 4 & 0 & 1 \\ 0 & 1 & 3 & -2 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Die Matrix ist jetzt in Zeilenstufenform; das Gleichungssystem lautet jetzt

$$\begin{aligned} x_1 + x_2 + 4x_3 + x_5 &= 0 \\ x_2 + 3x_3 - 2x_4 - \frac{3}{2}x_5 &= 0 \\ x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

Die vierte Gleichung kann gestrichen werden. Jeder der ersten drei Gleichungen fängt mit einer anderen Unbekannten an: x_1 bzw. x_2 bzw. x_5. Somit sind die Werte von x_3, x_4 frei wählbar, und die Werte von x_1, x_2, x_5 sind eindeutig durch die Werte von x_3, x_4 bestimmt:

$$\begin{aligned} &x_5 = 0 \\ &x_2 = -3x_3 + 2x_4 + \tfrac{3}{2}x_5 = -3x_3 + 2x_4 \\ &x_1 = -x_2 - 4x_3 - x_5 = 3x_3 - 2x_4 - 4x_3 = -x_3 - 2x_4 \end{aligned}$$

Somit ist der Lösungsraum $\{(-s - 2t, -3s + 2t, s, t, 0) \mid s, r \in k\}$. Eine Basis erhalten wir, indem wir $s = 1, t = 0$ und $s = 0, t = 1$ setzen: die Lösungen $(-1, -3, 1, 0, 0)$ und $(-2, 2, 0, 1, 0)$ bilden eine Basis des Lösungsraums.

Alternativ hätten wir den Gauß-Algorithmus bis zur strengen Zeilenstufenform durchführen können. Zuerst ziehen wir die zweite von der ersten Zeile ab:

$$\begin{pmatrix} 1 & 0 & 1 & 2 & \frac{5}{2} \\ 0 & 1 & 3 & -2 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Dann ziehen wir das $\frac{5}{2}$-fache der dritten Zeile von der ersten Zeile ab, und addieren das $\frac{3}{2}$-fache der dritten Zeile zur zweiten Zeile hinzu:

$$\begin{pmatrix} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Die Matrix ist jetzt in strenger Zeilenstufenform, das Gleichungssystem lautet jetzt

$$\begin{aligned} x_1 + x_3 + 2x_4 &= 0 \\ x_2 + 3x_3 - 2x_4 &= 0 \\ x_5 &= 0 \end{aligned}$$

Klarer als zuvor ist jetzt zu erkennen: x_3, x_4 dürfen beliebig gewählt werden, und es ist $x_1 = -x_3 - 2x_4$, $x_2 = -3x_3 + 2x_4$, $x_5 = 0$.

Bestimmung des Lösungsraums

Zu Aufgabe 5.2: gesucht wird eine Basis des Lösungsraums $\mathrm{LR}(A, 0)$. Nach Lemma 5.6 können wir die Matrix auf strenge Zeilenstufenform bringen. Nachdem man Nullzeilen gestrichen hat, bleiben r Gleichungen übrig, für $r =$ Anzahl der Pivotstellen. Die $i$te Gleichung hat die Gestalt

$$x_{j_i} = - \sum_{\substack{j > j_i \\ j \text{ kein } j_\ell}} A_{ij} x_j \, .$$

Somit sind die x_j derart, dass es keine Pivotstelle in der $j$ten Spalte gibt, frei wählbar; hiervon gibt es $n - r$ Stück. Die restlichen x_j enstsprechen Pivotspalten und sind eindeutig durch die Werte der freien Variablen festgelegt.

Für $1 \leq i \leq n - r$ sei v_i die Lösung, die man erhält, wenn man die $i$te frei Variable gleich Eins und die restlichen freien Variablen gleich Null setzt. Durch ein Komponentenvergleich sieht man dann, dass diese Lösungen $v_1, \ldots, v_{n-r}$ linear unabhängig sind. Wegen der Dimensionsformel stimmt die Anzahl der v_i mit der Dimension des Lösungsraums überein. Also bilden $v_1, v_2, \ldots, v_{n-r}$ eine Basis des Lösungsraums.

Jetzt behandeln wir Aufgabe 5.1. Wir wollen die Gleichung $A \cdot x = b$ lösen. Nun,

$$A \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 S_1 + \lambda_2 S_2 + \ldots + \lambda_n S_n$$

für $S_j = j$te Spalte von A. Somit ist $A \cdot x = b$ genau dann lösbar, wenn $b \in k^m$ eine Linearkombination der Spalten von A ist, d.h. wenn b im Spaltenraum von A liegt. Bezeichnen wir mit $(A|b)$ die *erweiterte* Matrix mit m Zeilen und $n + 1$ Spalten, die man erhält, die man aus A erhält, wenn man b als zusätzliche, letzte Spalte hinzufügt. Liegt b im Spaltenraum von A, so haben A und $(A|b)$ den gleichen Spaltenraum und -rang. Liegt b nicht im Spaltenraum von A, so ist der Spaltenrang von $(A|b)$ um eins größer. Wir haben gezeigt:

Lemma 5.8 *Das Gleichungssystem $A \cdot x = b$ ist genau dann lösbar, wenn A und die erweiterte Matrix $(A|b)$ den gleichen Rang haben. Den Rang dieser beiden Matrizen kann man durch den Gaußschen Algorithmus berechnen.* ∎

Beispiel Betrachten wir das Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_2 + x_3 &= 2 \\ x_1 + 2x_2 + 2x_3 &= 2 \end{aligned}$$

Es ist $A = \left(\begin{smallmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 2 \end{smallmatrix}\right)$ und $b = \left(\begin{smallmatrix} 1 \\ 2 \\ 2 \end{smallmatrix}\right)$, also $(A|b) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 2 & 2 & 2 \end{array}\right)$. Bringen wir diese Matrix auf Zeilenstufenform, so erhalten wir $\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{array}\right)$, wobei die ersten drei Spalten eine Zeilenstufenform von A sind. Somit ist $\mathrm{Rang}(A) = 3$, $\mathrm{Rang}(A|b) = 4$ und das Gleichungssystem ist unlösbar.

Dagegen ist das folgende Gleichungssystem lösbar:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_2 + x_3 &= 2 \\ x_1 + 2x_2 + 2x_3 &= 3 \end{aligned}$$

Diesmal ist $(A|b) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 2 & 2 & 3 \end{array}\right)$ mit Zeilenstufenform $\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$ und strenge Zeilenstufenform $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$. Dies entspricht dem Gleichungssystem

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

Es gibt eine freie Variable: x_3. Die Lösungsmenge ist $\{(-1, 2-t, t) \mid t \in k\}$.

6 Die Determinante

Jede quadratische Matrix $A \in M_n(k)$ hat eine Determinante $\det(A) = |A| \in k$. Für $n = 2$ gilt $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$. Die Determinante hat einige Anwendungen, insbesondere ist A genau dann invertierbar, wenn $\det(A) \neq 0$ ist.

Definition Sei k ein Körper und $n \geq 1$. Eine Abbildung $\Delta\colon M_n(k) \to k$ heißt eine Determinantenfunktion, wenn folgende drei Bedingungen erfüllt sind:

(D1) $\Delta(A)$ ist n-fach multilinear in den Zeilen, d.h. für alle $1 \leq i \leq n$ ist

$$\Delta(\ldots, \underbrace{\lambda v + \mu w}_{i\text{te Zeile}}, \ldots) = \lambda\Delta(\ldots, \underbrace{v}_{i\text{te Zeile}}, \ldots) + \mu\Delta(\ldots, \underbrace{w}_{i\text{te Zeile}}, \ldots)\,.$$

(D2) $\Delta(A)$ ist alternierend: sind zwei Zeilen gleich, so ist $\Delta(A) = 0$:

$$\Delta(\ldots, \underbrace{v}_{i\text{te Zeile}}, \ldots, \underbrace{v}_{j\text{te Zeile}}, \ldots) = 0\,.$$

(D3) Normierung: $\Delta(E_n) = 1$.

Später in Satz 6.3 werden wir zeigen, dass es genau eine Determinantenfunktion gibt. Für die Berechnung der Determinante benutzt man Bedingungen (D1) – (D3) sowie folgende weitere Eigenschaften:

Lemma 6.1 *Sei $\Delta\colon M_n(k) \to k$ eine Determinantenfunktion. Dann:*

(D4) *Hat A eine Nullzeile, so ist $\Delta(A) = 0$.*

(D5) *Determinantenfunktionen und elementare Zeilenoperationen:*

- Typ 1: *Entsteht B aus A, indem man zwei Zeilen miteinander vertauscht, so ist $\Delta(B) = -\Delta(A)$.*
- Typ 2: *Entsteht B aus A, indem man eine Zeile mit $\lambda \neq 0$ multipliziert, so ist $\Delta(B) = \lambda\Delta(A)$.*
- Typ 3: *Entsteht B aus A, indem man das λ-fache der $i$ten Zeile zur $j$ten Zeile addiert, so ist $\Delta(B) = \Delta(A)$.*

Beweis. (D4): Linearität (D1) in dieser Nullzeile.
(D5): Typ 2: Linearität (D1) in dieser Zeile. Typ 3: Wegen (D1), (D2) ist

$$\begin{aligned}\Delta(\ldots, \underbrace{v}_{\text{Zeile } i}, \ldots, \underbrace{w + \lambda v}_{\text{Zeile } j}, \ldots) &= \Delta(\ldots, \underbrace{v}_{\text{Zeile } i}, \ldots, \underbrace{w}_{\text{Zeile } j}, \ldots) \\ &\quad + \lambda\Delta(\ldots, \underbrace{v}_{\text{Zeile } i}, \ldots, \underbrace{v}_{\text{Zeile } j}, \ldots) \\ &= \Delta(\ldots, \underbrace{v}_{\text{Zeile } i}, \ldots, \underbrace{w}_{\text{Zeile } j}, \ldots)\end{aligned}$$

Typ 1: Wegen (D1), (D2) und Typ 3 ist

$$\begin{aligned}
0 &= \Delta(\ldots, \underbrace{v+w}_{\text{Zeile } i}, \ldots, \underbrace{v+w}_{\text{Zeile } j}, \ldots) \\
&= \Delta(\ldots, \underbrace{v}_{\text{Zeile } i}, \ldots, \underbrace{v+w}_{\text{Zeile } j}, \ldots) + \Delta(\ldots, \underbrace{w}_{\text{Zeile } i}, \ldots, \underbrace{v+w}_{\text{Zeile } j}, \ldots) \\
&= \Delta(\ldots, \underbrace{v}_{\text{Zeile } i}, \ldots, \underbrace{w}_{\text{Zeile } j}, \ldots) + \Delta(\ldots, \underbrace{w}_{\text{Zeile } i}, \ldots, \underbrace{v}_{\text{Zeile } j}, \ldots)\,. \qquad \blacksquare
\end{aligned}$$

Jetzt fangen wir mit der Konstruktion der Determinante an. Sei $A \in M_n(k)$ und sei Δ eine Determinantenfunktion. Sei $Z_i \in k^n$ die $i$te Zeile von A, also $Z_i = \sum_{j=1}^n A_{ij} e_j$. Wegen Linearität in der ersten Zeile gilt

$$\Delta(A) = \Delta(Z_1, Z_2, Z_3, \ldots, Z_n) = \sum_{j=1}^{n} A_{1j} \Delta(e_j, Z_2, Z_3, \ldots, Z_n)\,.$$

Wegen Linearität in der zweiten Zeile gilt dann

$$\Delta(A) = \sum_{j=1}^{n} \sum_{\ell=1}^{n} A_{1j} A_{2\ell} \Delta(e_j, e_\ell, Z_3, \ldots, Z_n)\,.$$

Verfährt man so mit jeder Zeile, so erhält man

$$\Delta(A) = \sum_{f} A_{1,f(1)} A_{2,f(2)} \cdots A_{n,f(n)} \Delta(e_{f(1)}, e_{f(2)}, \ldots, e_{f(n)})\,,$$

wobei aufsummiert wird über alle Abbildungen $f\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Ist f nicht injectiv, so ist $\Delta(e_{f(1)}, e_{f(2)}, \ldots, e_{f(n)}) = 0$ wegen (D2). Die injektive Abbildungen f sind gleichzeitig surjektiv, und deswegen Permutationen $\in S_n = \mathrm{Sym}(\{1, \ldots, n\})$. Es ist also

$$\Delta(A) = \sum_{\sigma \in S_n} A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)} \Delta(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)})\,.$$

Da σ eine Bijektion ist, kommt jeder Standardbasisvektor e_i genau einmal als $e_{\sigma(j)}$ vor. Sei $N(\sigma)$ die Anzahl von Vertauschungen, die nötig sind, um $(e_1, \ldots, e_n)$ aus $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$ zu machen. Dann

$$\Delta(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = (-1)^{N(\sigma)} \Delta(e_1, \ldots, e_n) = (-1)^{N(\sigma)}\,,$$

wegen (D5) Typ 2 und (D3). Also

$$\Delta(A) = \sum_{\sigma \in S_n} (-1)^{N(\sigma)} A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}\,. \qquad (*)$$

Wir wollen (*) als die Definition der Determinante benutzen. Vorher untersuchen wir, inwiefern die Zahl $N(\sigma)$ oder zumindest $(-1)^{N(\sigma)}$ eindeutig durch die Permutation σ definiert wird.

Lemma 6.2 *Das* Vorzeichen $\varepsilon(\sigma)$ *einer Permutation* $\sigma \in S_n$ *definiert man durch*

$$\varepsilon(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}\,.$$

a) Es ist $\varepsilon(\sigma) = (-1)^N$ *für*

$$N = \textit{Anzahl der Paare } i, j \textit{ mit } i < j \textit{ und } \sigma(i) > \sigma(j)\,.$$

b) Für $\sigma, \tau \in S_n$ *ist* $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$.

c) Ist $\sigma \in S_n$ *eine* Transposition*, d.h. werden lediglich zwei Elemente vertauscht, so ist* $\varepsilon(\sigma) = -1$.

d) Jede Permutation $\sigma \in S_n$ *lässt sich als ein Produkt von Transpositionen schreiben. Lässt sich* σ *als ein Produkt von* N *Transpositionen schreiben, so ist* $\varepsilon(\sigma) = (-1)^N$.

Beweis. a) Jedes Paar $i < j$ kommt einmal als $j - i$ im Nenner und einmal als $\pm(j - i)$ im Zähler: $+$ falls $\sigma^{-1}(j) > \sigma^{-1}(i)$ und $-$ sonst.

b) Es ist $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$, also hängt $\frac{\sigma(j)-\sigma(i)}{j-i}$ nur vom *ungeordneten* Paar i, j ab. Nun,

$$\begin{aligned}\varepsilon(\sigma\tau) = \prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{j - i} &= \prod_{i<j} \left(\frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)}\right)\left(\frac{\tau(j) - \tau(i)}{j - i}\right) \\ &= \left(\prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)}\right)\varepsilon(\tau)\,,\end{aligned}$$

und $\prod_{i<j} \frac{\sigma\tau(j)-\sigma\tau(i)}{\tau(j)-\tau(i)} = \varepsilon(\sigma)$, denn für jedes ungeeordnete Paar a, b kommt $\frac{\sigma(b)-\sigma(a)}{b-a}$ genau einmal im Produkt vor.

c) Es gibt also $i < j$ mit $\sigma(i) = j$, $\sigma(j) = i$ und $\sigma(\ell) = \ell$ sonst. Die Anzahl der Paare in Teil a) ist also ungerade: das Paar i, j, sowie die beiden Paare i, ℓ und ℓ, j für jedes $i < \ell < j$.

d) Der letzter Teil folgt aus c) und dem ersten Teil. Den ersten Teil zeigt man per Induktion über n. Permutationen mit $\sigma(n) = n$ betrachten wir als Elemente von S_{n-1}. Ist $\sigma(n) \neq n$, so sei τ die Transposition, die n und $\sigma^{-1}(n)$ vertauscht. Dann $\sigma' := \sigma\tau^{-1} \in S_{n-1}$ und $\sigma = \sigma'\tau$. Nach Induktionsannahme lässt sich σ' als ein Produkt von Transpositionen schreiben.

■

Satz 6.3 *Sei k ein Körper und $n \geq 1$. Es gibt genau eine Determinantenfunktion* $\det\colon M_n(k) \to k$. *Diese wird gegeben durch*

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} \,.$$

Man schreibt auch $|A|$ für $\det(A)$.

Beweis. Wegen Lemma 6.2 ist $\varepsilon(\sigma)$ das $(-1)^{N(\sigma)}$ aus Gleichung (*). Die Überlegungen vor dieser Gleichung zeigen, dass det die einzige Möglichkeit für eine Determinantenfunktion ist. Andrerseits rechnet man nach, dass det die drei Bedinungen erfüllt. (D1) folgt aus der Definition von det. (D3) Das Produkt der $A_{i\sigma(i)}$ ist nur $\neq 0$ für $\sigma = \mathrm{Id}$, und $\varepsilon(\mathrm{Id}) = +1$. (D2): Angenommen, Zeilen i, j sind gleich für $i < j$. Sei τ die Transposition, die i, j vertauscht. Dann $A_{ia}A_{jb} = A_{ib}A_{ja}$, weshalb für jedes $\sigma \in S_n$ gilt $A_{1,\sigma(1)}A_{2,\sigma(2)} \cdots A_{n,\sigma(n)} = A_{1,\sigma\tau(1)}A_{2,\sigma\tau(2)} \cdots A_{n,\sigma\tau(n)}$. Diese beiden Summanden heben einander auf, denn deren Vorzeichen sind $\varepsilon(\sigma)$ und $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau) = -\varepsilon(\sigma)$. ∎

Lemma 6.4 *Zuzüglich zu (D1) bis (D5) verfügt die Determinante* $\det\colon M_n(k) \to k$ *auch über die folgenden Eigenschaften:*

(D6) *Hat die Matrix A* obere Dreiecksgestalt*, d.h. ist $A_{ij} = 0$ für all $i > j$, wie z.B. in* $\begin{pmatrix} 1&2&3&4 \\ 0&3&2&1 \\ 0&0&2&2 \\ 0&0&0&1 \end{pmatrix}$ *der Fall, dann ist* $\det(A)$ *das Produkt der Einträge auf der Hauptdiagonale:* $\det(A) = \prod_{i=1}^n A_{ii} = A_{11}A_{22}\cdots A_{nn}$.

(D7) *Ist A eine* Blockmatrix *mit der Gestalt* $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$*, wobei die Matrizen B, D quadratisch sind, so ist* $\det(A) = \det(B)\det(D)$.

(D8) *Es ist* $\det(A) \neq 0$ *genau dann, wenn* $\mathrm{Rang}(A) = n$ *ist, d.h. genau dann, wenn A invertierbar ist.*

(D9) *Produktregel: Für Matrizen $A, B \in M_n(k)$ ist* $\det(AB) = \det(A)\det(B)$.

(D10) *Für die durch $(A^T)_{ij} = A_{ji}$ definierte* transponierte *Matrix $A^T \in M_n(k)$ gilt* $\det(A^T) = \det(A)$.

(D11) *Eigenschaften (D1), (D2), (D4) und (D5) gelten auch für Spalten und Spaltenoperationen.*

Beweis. (D6): Ist $A_{1,\sigma(1)}A_{2,\sigma(2)} \cdots A_{n,\sigma(n)} \neq 0$, dann ist $\sigma(i) \geq i$ für alle $1 \leq i \leq n$. Da σ eine Bijektion ist, geht dies nur für $\sigma = \mathrm{Id}$, mit Vorzeichen $+1$.

(D7): Sei $B \in M_r(k)$ und $D \in M_s(k)$, mit $r + s = n$. Es ist $A_{ij} = 0$ falls $i > r$ und $j \leq r$. Ist $A_{1,\sigma(1)}A_{2,\sigma(2)} \cdots A_{n,\sigma(n)} \neq 0$, dann $\sigma(i) > r$ für alle $i > r$, und deshalb $\sigma(i) \leq r$ für alle $i \leq r$. Das heißt, $\sigma = \sigma'\sigma''$ für Permutationen

$\sigma' \in \mathrm{Sym}(\{1,\dots,r\})$ und $\sigma'' \in \mathrm{Sym}(\{r+1,\dots,n\})$. Außerdem ist $A_{ij} = B_{ij}$ für $1 \le i,j \le r$, weshalb $A_{i\sigma'(i)} = B_{i\sigma'(i)}$ für $i \le r$; und $A_{r+i,r+j} = D_{ij}$ für $1 \le i,j \le s$, weshalb $A_{i\sigma''(i)} = D_{i-r,\tau(i-r)}$ für $1 \le i \le s$, wobei $\tau \in \mathrm{Sym}(\{1,\dots,s\})$ durch $\tau(i) + r = \sigma''(i+r)$ definiert wird. Also

$$\det(A) = \sum_{\sigma' \in S_r} \sum_{\sigma'' \in S_s} \varepsilon(\sigma')\varepsilon(\tau) B_{1\sigma'(1)} \cdots B_{r\sigma'(r)} D_{1\tau(1)} \cdots D_{s\tau(s)} = \det(B)\det(D)\,.$$

(D8): Durch Zeilenoperationen bringt der Gauß-Algorithmus jede Matrix auf Zeilenstufenform. (D5) beschreibt die Auswirkung dieser Operationen auf $\det(A)$, insbesondere bleibt $\det(A)$ entweder $= 0$ oder $\neq 0$. Ist $\mathrm{Rang}(A) < n$, so hat die Zeilenstufenform mindestens eine Nullzeile, also $\det(A) = 0$ wegen (D4). Matrizen in Zeilenstufenform haben obere Dreiecksgestalt. Ist $\mathrm{Rang}(A) = n$, so gibt es n Pivotstellen, die alle auf der Hauptdiagonale liegen – d.h. jede (i,i)-Stelle ist eine Pivotstelle. Wegen (D6) ist $\det(A)$ das Produkt der Pivot-Einträge, die ja alle $\neq 0$ sind. Also $\det(A) \neq 0$.

(D9): Es ist $\mathrm{Rang}(A) = \dim \mathrm{Bild}(L_A)$ und $L_{AB} = L_A \circ L_B$, also $\mathrm{Bild}(L_{AB}) \subseteq \mathrm{Bild}(L_A)$. Deshalb: ist $\mathrm{Rang}(A) < n$, dann $\mathrm{Rang}(AB) < n$, also $\det(A)$, $\det(AB)$ sind beide $= 0$ wegen (D8). Ist $\mathrm{Rang}(A) = n$, so gibt es nach Lemma 5.7 Elementarmatrizen $H_1,\dots,H_r \in M_n(k)$ mit $A = H_r H_{r-1} \cdots H_1$. Eine Matrix H heißt eine Elementarmatrix, wenn Linksmultiplikation $C \mapsto HC$ eine elementare Zeilenoperation bewirkt, vgl. Lemma 5.3. Mit $C = E_n$ sieht man, dass jede Elementarmatrix entsteht, indem man die entsprechende Zeilenoperation auf der Einheitsmatrix E_n anwendet. Nun, $\det(E_n) = 1$, und (D5) besagt, dass jede Zeilenoperation $\det(C)$ mit einem Skalar multipliziert, die nur von der Zeilenoperation abhängig ist, nicht von C selbst. Das heißt, $\det(HC) = \det(H)\det(C)$ gilt für Elementarmatrizen H. Wendet man diese Gleichung wiederholt an, so erhält man $\det(AC) = \det(H_r)\det(H_{r-1})\cdots\det(H_1)\det(C)$. Mit $C = E_n$ erhält man $\det(A) = \det(H_r)\det(H_{r-1})\cdots\det(H_1)$. Mit $C = B$ erhält man also $\det(AB) = \det(A)\det(B)$.

(D10) Es ist $\det(A^T) = \sum_{\sigma \in S_n} \varepsilon(\sigma) A_{\sigma(1),1} A_{\sigma(2),2} \cdots A_{\sigma(n),n}$. Ordnet man die Reihenfolge in den Produkten um, so erhält man

$$\begin{aligned}\det(A^T) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) A_{1,\sigma^{-1}(1)} A_{2,\sigma^{-1}(2)} \cdots A_{n,\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}\,.\end{aligned}$$

Dies ist aber $\det(A)$, denn $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$, wegen $\varepsilon(\sigma)\varepsilon(\sigma^{-1}) = \varepsilon(\mathrm{Id}) = 1$.

(D11) Es ist $\det(A) = \det(A^T)$. Spalten bzw. Spaltenoperationen für A entsprechen Zeilen bzw. Zeilenoperationen für A^T. ∎

Berechnung der Determinante Sei A eine Matrix. Durch elementare Zeilenoperationen können wir A auf Zeilenstufenform bringen. Zeilenstufenformen

haben obere Dreiecksgestalt, also können wir wegen (D6) die Determinante der Zeilenstufenform berechnen. Dann erlaubt uns (D5), die Determinante der ursprünglichen Matrix zu berechnen.

Beispiel Wir berechnen die Determinante $|A|$ der Matrix $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 4 & 6 & 1 \\ 8 & 3 & 2 & -1 \\ 0 & -1 & 3 & 2 \end{pmatrix}$. Es ist

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 4 & 6 & 1 \\ 8 & 3 & 2 & -1 \\ 0 & -1 & 3 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 8 & 3 & 2 & -1 \\ 5 & 4 & 6 & 1 \end{vmatrix} \qquad \text{(Zeilen 2 und 4 vertauscht)}$$

$$= - \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & -13 & -22 & -33 \\ 0 & -6 & -9 & -19 \end{vmatrix} \qquad \begin{pmatrix} \text{1. Zeile 8mal von der 3. und} \\ \text{5mal von der 4. abgezogen} \end{pmatrix}$$

$$= - \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & -61 & -59 \\ 0 & 0 & -27 & -31 \end{vmatrix} \qquad \begin{pmatrix} \text{2. Zeile 13mal von der 3. und} \\ \text{6mal von der 4. abgezogen} \end{pmatrix}$$

$$= - \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & -61 & -59 \\ 0 & 0 & 0 & -\frac{298}{61} \end{vmatrix} \qquad \begin{pmatrix} \text{3. Zeile } \frac{27}{61}\text{mal von der 4. ab-} \\ \text{gezogen} \end{pmatrix}$$

$$= -1 \cdot (-1) \cdot (-65) \cdot \left(-\frac{298}{65} \right) = 298 \qquad \text{(D6)}.$$

Laplace-Entwicklung

Satz 6.5 *Sei k ein Körper und $A \in M_n(k)$ eine quadratische Matrix. Für $i, j \in \{1, 2, \ldots, n\}$ werden wir mit $A(i,j)$ die Matrix in $M_{n-1}(k)$ bezeichnen, die entsteht, wenn man die $i$te Zeile und $j$te Spalte von A wegstreicht. Es ist dann:*

a) (Laplace-Entwicklung nach der $i$ten Zeile)
Für jedes $1 \leq i \leq n$ gilt $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A(i,j))$.

b) (Laplace-Entwicklung nach der $j$ten Spalte)
Für jedes $1 \leq j \leq n$ gilt $\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A(i,j))$.

Beweis. Zu b): Nach Satz 6.3 reicht es nachzuweisen, dass die rechte Seite der Gleichung die Bedingungen (D1), (D2) und (D3) erfüllt. (D1), d.h. Linearität in der $i_0$ten Zeile: Für $i = i_0$ hängt $A_{i_0 j}$ linear von der $i_0$ten Zeile ab, und $A(i_0, j)$ hängt gar nicht von dieser Zeile ab. Für $i \neq i_0$ ist es umgekehrt: A_{ij} bzw. $\det A(i,j)$ hängt gar nicht bzw. linear von der $i_0$ten Zeile ab. (D3) Es ist $A_{ij} =$

δ_{ij}, für $i \neq j$ ist dies 0. Also nur der Summand $i = j$ zählt. Dieser Summand ist $(-1)^{2j} \det A(j,j) = 1$, denn $A(j,j) = E_{n-1}$. (D2) Angenommen Zeilen a, b sind gleich, mit $a < b$. Für $i \neq a, b$ ist $\det A(i,j) = 0$ wegen (D2). Also

$$\text{Rechte Seite} = (-1)^{a+j} A_{aj} \det A(a,j) + (-1)^{b+j} A_{bj} \det A(b,j)\,.$$

Es ist $A_{bj} = A_{aj}$. Die $a$te Zeile $v \in k^{n-1}$ von $A(b,j)$ ist die wiederholte Zeile, mit der $j$ten Eintrag gestrichen. Die Matrix $A(a,j)$ ist gleich, außer man hat v solange nach unten an die anderen Zeilen vorbeigeschoben, bis v zur $(b-1)$ten Zeile wird. Hierfür wird v sukzessiv mit jeder der $b-1-a$ dazwischen liegenden Zeilen vertauscht, also $\det A(b,j) = (-1)^{b-1-a} \det A(a,j)$. Also

$$\text{Rechte Seite} = (-1)^{a+j} A_{aj} \det A(a,j) + (-1)^{b+j} A_{aj} (-1)^{b-a+1} \det A(a,j) = 0\,.$$

a) Folgt aus Teil b) und (D10). ∎

Die Determinante eines Endomorphismus

Definition Sei V ein k-Vektorraum. Ein *Endomorphismus* von V ist eine lineare Abbildung $f\colon V \to V$.

Lemma 6.6 *Sei f ein Endomorphismus des n-dimensionalen Vektorraums V. Für jede Basis $B = b_1, \ldots, b_n$ von V nimmt* $\det({}_B\mathrm{M}_B(f))$ *den gleichen Wert. Diesen gemeinsamen Wert nennt man die Determinante* $\det(f)$ *des Endomorphismus f.*

Beweis. Sei $C = c_1, \ldots, c_n$ eine weitere Basis. Teil b) von Lemma 4.7 besagt: es ist ${}_C\mathrm{M}_D(g)\,{}_B\mathrm{M}_C(f) = {}_B\mathrm{M}_D(gf)$. Also ${}_C\mathrm{M}_C(f) = {}_B\mathrm{M}_C(\mathrm{Id})\,{}_B\mathrm{M}_B(f)\,{}_C\mathrm{M}_B(\mathrm{Id})$, und außerdem ist ${}_B\mathrm{M}_C(\mathrm{Id})\,{}_C\mathrm{M}_B(\mathrm{Id}) = {}_C\mathrm{M}_C(\mathrm{Id}) = E_n$. Nach der Produktregel (D9) ist also

$$\begin{aligned}\det {}_C\mathrm{M}_C(f) &= \det {}_B\mathrm{M}_C(\mathrm{Id}) \det {}_C\mathrm{M}_B(\mathrm{Id}) \det {}_B\mathrm{M}_B(f)\\ &= \det E_n \det {}_B\mathrm{M}_B(f) = \det {}_B\mathrm{M}_B(f)\,.\end{aligned}$$ ∎

7 Eigenwerte und Eigenvektoren

Definition Sei $\phi\colon V \to V$ ein Endomorphismus des k-Vektorraums ϕ. Gibt es einen Vektor $0 \neq v \in V$ und ein Skalar $\lambda \in k$ derart, dass $\phi(v) = \lambda v$ gilt, so heißt v ein *Eigenvektor* von ϕ mit *Eigenwert* λ.

Im wichtigen Fall $V = k^n$ lässt sich ϕ durch eine Matrix $A \in M_n(k)$ ausdrücken. Dann heißt $0 \neq v \in k^n$ ein Eigenvektor von A mit Eigenwert λ, falls $A \cdot v = \lambda v$ gilt.

Bemerkung Laut obiger Definition ist der Nullvektor kein Eigenvektor. Mit $v = 0$ gilt $\phi(v) = \lambda v$ für jeden Wert von λ. Dies bedeutet aber nicht, dass jeder Wert von λ ein Eigenwert ist.

Beispiele a) Zu den Eigenvektoren von $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} \in M_2(\mathbb{R})$ gehören unter anderen $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ mit Eigenwert 1 und $\left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right)$ mit Eigenwert -1: denn es ist

$$\begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ -3 \end{pmatrix}$$

b) Wegen

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & -1 & 1 \\ -1 & 1 & -5 \end{pmatrix} \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist $\begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix}$ ein Eigenvektor von $\begin{pmatrix} 1 & 2 & 2 \\ 0 & -1 & 1 \\ -1 & 1 & -5 \end{pmatrix}$ mit Eigenwert 0.

c) Die Zuordnung $f(x) \mapsto f''(x)$ (zweite Ableitung) ist ein Endomorphismus des $\mathbb{R}$-Vektorraums $C^\infty(\mathbb{R})$ aller beliebig oft stetig differenzierbarer Funktionen auf $\mathbb{R}$. Es ist $\sin''(x) = -\sin(x)$ und $\cos''(x) = -\cos(x)$. Deshalb sind die Sinus- und Kosinusfunktionen zwei Eigenvektoren der zweiten Ableitung, jeweils mit Eigenwert -1.

d) Die Zahl 2 ist kein Eigenwert der Matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 3 \end{smallmatrix}\right) \in M_2(\mathbb{R})$, denn wäre $v = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ ein Eigenvektor, so müsste

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \end{pmatrix}$$

gelten, d.h. $a + 2b = 2a$ und $3b = 2b$, woraus folgt $a = b = 0$. Aber der Nullvektor gilt nicht als Eigenvektor.

e) Als Element von $M_2(\mathbb{R})$ hat $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ keine Eigenwerte: ist $A\binom{x}{y} = \lambda\binom{x}{y}$, dann $-y = \lambda x$ und $x = \lambda y$, also $(\lambda^2+1)x = 0$, also $x = 0$ und $y = 0$. Arbeitet man stattdessen über $\mathbb{C}$, so ist $\binom{i}{1}$ ein Eigenvektor mit Eigenwert i.

Bezeichnung Ist $\lambda \in k$ ein Eigenwert des Endomorphismus ϕ von V, so nennt man

$$E_\lambda(\phi) := \{v \in V \mid \phi(v) = \lambda v\}$$

den *Eigenraum* von ϕ zum Eigenwert λ. Analog definiert man den Eigenraum $E_\lambda(A)$ einer Matrix $A \in M_n(k)$.

Lemma 7.1 *a) $\lambda \in k$ ist genau dann ein Eigenwert von $A \in M_n(k)$, wenn die Matrix $A - \lambda E_n$* singulär *ist, d.h. vom Rang $< n$.*

b) Der Eigenraum $E_\lambda(\phi)$ eines Endomorphismus ϕ von V ist ein Unterraum von V.

Beweis. a) Sei $B = A - \lambda E_n$. Für $v \in k^n$ ist $A \cdot v = \lambda v$ genau dann, wenn $B \cdot v = 0$. Die Eigenvektoren von A mit Eigenwert λ sind somit die Vektoren $\neq 0$ im Nullraum von B, d.h. im Kern von L_B. Es gibt also genau dann Eigenvektoren zu diesem Eigenwert, wenn $\dim \mathrm{Kern}(L_B) > 0$ ist. Laut Dimensionsformel ist dies genau dann der Fall, wenn $\dim \mathrm{Bild}(L_B) < n$ ist, d.h. wenn $\mathrm{Rang}(B) < n$ ist.

b) Auch wenn er kein Eigenvektor ist, liegt der Nullvektor doch im Eigenraum. Wie gerade gesehen, ist der Eigenraum $E_\lambda(\phi)$ der Kern des Endomorphismus $\phi - \lambda\,\mathrm{Id}$. Kerne sind Unterräume.

■

Das charakteristische Polynom

Eine häufige Aufgabe besteht darin, die Eigenwerte und Eigenvektoren einer vorgegebenen Matrix zu bestimmen. Hierfür benutzt man das sog. *charakteristische Polynom* der Matrix.

Definition Sei k ein Körper und $A \in M_n(k)$ eine quadratische Matrix. Das *charakteristische Polynom* $p_A(X)$ ist das Polynom $\det(XE_n - A)$ mit einer Variable X und mit Koeffizienten aus k.

Beispiel Für die Matrix $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} \in M_2(\mathbb{R})$ gilt

$$p_A(X) = \begin{vmatrix} X-2 & 1 \\ -3 & X+2 \end{vmatrix} = (X-2)(X+2) - (+1)(-3) = X^2 - 4 + 3 = X^2 - 1\,.$$

Satz 7.2 *Das charakteristische Polynom $p_A(X)$ ist normiert*[9] *und vom Grad n. Die Eigenwerte der Matrix A sind genau die Nullstellen des charakteristischen Polynoms. Hiervon gibt es höchstens n Stück.*

Für den Beweis des Satzes benötigen wir ein Ergebnis über Polynome.

Lemma 7.3 *Sei k ein Körper und $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_2X^2 + a_1X + a_0$ ein Polynom mit Koeffizienten aus k.*

a) Sei $\lambda \in k$. Es ist $f(\lambda) = 0$ genau dann, wenn das Polynom $f(X)$ durch $X - \lambda$ teilbar ist.

b) Ist n der Grad von $f \neq 0$, so hat f höchstens n verschiedene Nullstellen in k.

Beweis des Lemmas. Beide Teile werden per Induktion über den Grad n von f gezeigt. Das Nullpolynom $f(X) = 0$ hat Grad $-\infty$.

a) Ist $f(X) = (X-\lambda)g(X)$ für ein Polynom $g(X)$, so ist offenbar $f(\lambda) = 0$. Die andere Implikation wollen wir jetzt zeigen. Induktionsanfang $n = 0$: Unter den Polynomen $f(X) = a_0$ vom Grad höchstens 0 hat nur das Nullpolynom $f(X) = 0$ eine Nullstelle in λ. Das Nullpolynom ist durch $X - \lambda$ teilbar: $0 = (X - \lambda) \cdot 0$.

   Induktionsschritt: Es ist $f = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0$ vom Grad höchstens $n \geq 1$. Wir setzen voraus $f(\lambda) = 0$ und müssen zeigen, dass $f(X)$ durch $X-\lambda$ teilbar ist. Sei $f'(X) = f(X) - a_nX^{n-1}(X-\lambda)$, dann ist auch $f'(\lambda) = 0$. Aber $f'(X) = (a_{n-1}+\lambda a_n)X^{n-1}+a_{n-2}X^{n-2}+\cdots+a_1X+a_0$ hat Grad höchstens $n-1$. Nach der Induktionsannahme ist $f'(X)$ durch $X-\lambda$ teilbar: es gibt ein Polynom $g'(X)$ mit $f'(X) = (X-\lambda)g'(X)$. Dann ist $f(X) = (X-\lambda)g(X)$ für $g(X) = a_nX^{n-1} + g'(X)$.

b) Induktionsanfang $n = 0$: ein Polynom $f(X) = a_0$ mit $a_0 \neq 0$ ist konstant und hat keine Nullstellen. Induktionsschritt: Hat f vom Grad n eine Nullstelle in λ, so gibt es nach a) ein Polynom $g(X)$ mit $f(X) = (X - \lambda)g(X)$. Nun, der Grad von g muss $n-1$ sein, also hat g höchstens $n-1$ Nullstellen. Aber jede Nullstelle von $f(X)$ ist entweder eine Nullstelle von $g(X)$, oder eine Nullstelle von $X - \lambda$, d.h. λ. ∎

Beweis des Satzes. Sei C die Matrix $XE_n - A$. Normiert vom Grad n: Für $\sigma = \mathrm{Id}$ ist $\varepsilon(\sigma) = +1$ und $C_{11}C_{22}\cdots C_{nn} = (X - A_{11})(X - A_{22})\cdots(X - A_{nn}) = X^n + $ Terme vom Grad $\leq n-1$. Für alle weiteren Permutationen $\sigma \in S_n$ ist $\sigma(i) \neq i$ für mindestens zwei i, weshalb höchstens $n-2$ Faktoren von $C_{1\sigma(1)}\cdots C_{n\sigma(n)}$ der Art $X - A_{ii}$ sind, die restlichen sind $-A_{ij}$. Jeder weitere

[9] Das heißt, der führende Koeffizient beträgt 1.

Summand in $\det(C)$ hat also Grad höchstens $n-2$ als Polynom in X. Damit hat die Determinante Grad n, und die Koeffizienten von X^n und von X^{n-1} sind die Koeffizienten in $(X-A_{11})(X-A_{22})\cdots(X-A_{nn})$. Der Koeffizient von X^n ist also 1.

Wegen Lemma 7.1 Teil a) sind die Eigenwerte genau die Nullstellen des charakteristischen Polynoms. Wegen Lemma 7.3 Teil b) gibt es höchstens n Nullstellen. ∎

Bezeichnung Die *Spur* einer quadratischen Matrix $A \in M_n(k)$ ist per Definition die Summe der Diagonaleinträge:

$$\mathrm{Spur}(A) = \sum_{i=1}^{n} A_{ii}\,.$$

Lemma 7.4 *Sei $A \in M_n(k)$ eine quadratische Matrix. Der Koeffizient von X^{n-1} im charakteristischen Polynom $p_A(X)$ ist $-\mathrm{Spur}(A)$. Das Absolutglied von $p_A(X)$ beträgt $(-1)^n \det(A)$.*

Beweis. Nach dem Beweis von Satz 7.2 stimmt der Koeffizient von X^{n-1} in $p_A(X)$ mit seinem Koeffizient in $\prod_{i=1}^{n}(X-A_{ii})$ überein. Dieser Koeffizient ist $-(A_{11}+A_{22}+\cdots+A_{nn})$.

Das Absolutglied ist $p_A(0)$, d.h. $\det(-A)$. Nach Linearität in den Zeilen ist $\det(-A) = (-1)^n \det(A)$. ∎

Beispiel Die Matrix $\left(\begin{smallmatrix} 3 & 4 \\ 1 & 7 \end{smallmatrix}\right)$ hat Spur $3+7=10$ und Determinante $3\cdot 7 - 4\cdot 1 = 17$. Nach dem Lemma beträgt das charakteristische Polynom also $X^2-10X+17$. Tatsächlich ist

$$\begin{aligned}\begin{vmatrix} X-3 & -4 \\ -1 & X-7 \end{vmatrix} &= (X-3)(X-7)-(-4)\cdot(-1) \\ &= (X^2-10X+21)-4 = X^2-10X+17\,.\end{aligned}$$

Lemma 7.5 *Sei V ein n-dimensionaler k-Vektorraum, und sei ϕ ein Endomorphismus von V. Dann:*

a) Für jede Basis B von V hat die Matrix ${}_B\mathrm{M}_B(\phi)$ das gleiche charakteristische Polynom. Dieses Polynom nennt man das charakteristische Polynom $p_\phi(X)$ des Endomorphismus ϕ.

b) Die Eigenwerte von ϕ sind genau die Nullstellen von $p_\phi(X)$.

c) Für jede Basis B von V hat die Matrix ${}_B\mathrm{M}_B(\phi)$ die gleiche Spur. Diesen gemeinsamen Wert nennt man die Spur $\mathrm{Spur}(\phi)$ des Endomorphismus ϕ.

Beweis. a) Seien B, C zwei Basen von V. Sei $R = {}_B\mathrm{M}_B(\phi)$, $S = {}_C\mathrm{M}_C(\phi)$ und $T = {}_B\mathrm{M}_C(\mathrm{Id})$. Nach dem Beweis von Lemma 6.6 ist $S = TRT^{-1}$, weshalb $XE_n - S = T(XE_n - R)T^{-1}$. Aufgrund der Produktregel haben somit $XE_n - S$ und $XE_n - R$ die gleiche Determinante, d.h. R und S haben das gleiche charakteristische Polynom.

b) λ ist genau dann ein Eigenwert von ϕ, wenn es ein Eigenwert von ${}_B\mathrm{M}_B(\phi)$ ist. Also folgt das Ergebnis aus Satz 7.2.

c) Folgt aus Teil a) und Lemma 7.4.

∎

Summen von Eigenräume sind direkt

Lemma 7.6 *Sei $A \in M_n(k)$. Seien $\lambda_1, \ldots, \lambda_r$ einige paarweise verschiedene Eigenwerte von A. Dann ist die Summe $E_{\lambda_1}(A) + \cdots + E_{\lambda_r}(A)$ eine* direkte *Summe $E_{\lambda_1}(A) \oplus \cdots \oplus E_{\lambda_r}(A)$.*

Anders gesagt lässt sich jedes $v \in k^n$ auf höchstens eine Weise als $v = v_1 + \cdots + v_r$ schreiben mit $A \cdot v_i = \lambda_i v_i$ für jedes $1 \leq i \leq r$.

Beweis. Ist $v = v_1 + \cdots + v_r = w_1 + \cdots + w_r$ mit $A \cdot v_i = \lambda_i v_i$ und $A \cdot w_i = \lambda_i w_i$, dann

$$0 = u_1 + \cdots + u_r \qquad (*)$$

mit $A \cdot u_i = \lambda_i u_i$ für alle i, wobei $u_i = v_i - w_i$. Wir zeigen jetzt per Induktion über r, dass $u_i = 0$ gilt für alle i. Der Induktionsanfang (der Fall $r = 1$) ist klar, also kommen wir zum Induktionsschritt. Hierfür wenden wir $A - \lambda_r E_n$ auf beiden Seiten der Gleichung (*) an. Es ist $(A - \lambda_r E_n) \cdot u_i = (\lambda_i - \lambda_r)u_i$, also

$$0 = (\lambda_1 - \lambda_r)u_1 + \cdots + (\lambda_{r-1} - \lambda_r)u_{r-1}\,.$$

Nach der Induktionsannahme ist $(\lambda_i - \lambda_r)u_i = 0$ für alle $1 \leq i \leq r-1$. Da die λ_i paarweise verschieden sind, ist $\lambda_i - \lambda_r \neq 0$ für $i \leq r-1$, also $u_i = 0$ für alle $1 \leq i \leq r-1$. Aus (*) folgt jetzt $u_r = 0$. ∎

Diagonalisierbarkeit

Definition Eine Matrix $A \in M_n(k)$ heißt eine *Diagonalmatrix*, falls $A_{ij} = 0$ gilt für alle $i \neq j$, d.h. falls alle Einträge $\neq 0$ auf der Hauptdiagonale liegen.

Ein Endomorphismus ϕ eines k-Vektorraums V heißt *diagonalisierbar*, wenn es eine Basis B von V gibt derart, dass ${}_B\mathrm{M}_B(\phi)$ eine Diagonalmatrix ist.

Eine Matrix $A \in M_n(k)$ heißt diagonalisierbar, wenn es zu einer Diagonalmatrix *ähnlich* ist, d.h. wenn es eine invertierbare Matrix $T \in M_n(k)$ gibt derart, dass TAT^{-1} eine Diagonalmatrix ist.

Äquivalente Formulierung: Ein Endomorphismus oder eine Matrix heißt diagonalisierbar, wenn es eine Basis gibt, die aus lauter Eigenvektoren besteht.

Beispiele a) Die Matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ aus $M_2(\mathbb{R})$ ist diagonalisierbar. Einerseits gibt es eine Basis $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ für $\mathbb{R}^2$, die aus Eigenvektoren besteht (Eigenwert 1 bzw. 2). Andrerseits ist $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ invertierbar mit $T^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$, und $T^{-1}AT = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$, eine Diagonalmatrix.

b) Die Matrix $B = \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix}$ aus $M_2(\mathbb{R})$ ist nicht diagonalisierbar, denn das charakteristische Polynom ist $X^2 - 2X + 1 = (X-1)^2$, d.h. 1 ist der einzige Eigenwert – und der Eigenraum $E_1(B)$ ist eindimensional, mit Basis $\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$. Es gibt also keine Basis, die aus Eigenvektoren besteht.

Lemma 7.7 *Gibt es n verschiedene Eigenwerte der Matrix $A \in M_n(k)$, so ist A diagonalisierbar.*

Beweis. Seien $\lambda_1, \ldots, \lambda_n$ diese Eigenwerte. Jede Eigenraum $E_{\lambda_i}(A)$ hat eine Basis aus Eigenvektoren. Somit hat auch die Summe $U := E_{\lambda_1}(A) + \cdots E_{\lambda_n}(A)$ eine Basis aus Eigenvektoren.

Nach Lemma 7.6 ist U die direkte Summe dieser Eigenräume. Somit ist $\dim(U)$ die Summe der Dimensionen der Eigenräume – s. hierzu Übungsserie 7, Aufgabe 7. Jeder Eigenraum hat Dimension mindestens 1. Also $\dim(U) \geq n$, d.h. $U = k^n$ und es gibt für k^n eine Basis von Eigenvektoren. ∎

Beispiel Die Matrix $C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ hat charakteristisches Polynom $X^2 + 1$. Über $\mathbb{R}$ hat das Polynom keine Nullstellen, also gibt es keine Eigenwerte, und die Matrix ist nicht diagonalisierbar.

Dagegen ist die Matrix diagonalisierbar über $\mathbb{C}$, und zwar ähnlich zur Diagonalmatrix $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. Denn das charakteristisches Polynom hat Nullstellen $i, -i$. Jetzt wendet man Lemma 7.7 an.

8 Skalarprodukte und der Spektralsatz

Für Vektoren $v = (v_1, v_2, v_3)$ und (w_1, w_2, w_3) des $\mathbb{R}^3$ wird das Skalaprodukt $v \bullet w$ oder $\langle v, w\rangle$ gegeben durch

$$\langle v, w\rangle = v_1w_1 + v_2w_2 + v_3w_3\,.$$

Die Länge $\|v\|$ des Vektors v ist gegeben durch $\|v\| = \sqrt{\langle v, v\rangle}$, und es gilt

$$\langle v, w\rangle = \|v\| \cdot \|w\| \cdot \cos(\theta)\,,$$

wobei θ der Winkel zwischen den Vektoren v, w ist. Wir verallgemeinern jetzt den Begriff Skalarprodukt.

Definition Sei V ein $\mathbb{R}$-Vektorraum. Ein *Skalarprodukt* auf V ist eine Abbildung $\langle , \rangle \colon V \times V \to \mathbb{R}$, $(v, w) \mapsto \langle v, w\rangle$, die die folgenden Bedingungen erfüllt:

a) $\langle , \rangle$ ist *bilinear*, d.h. $\langle \lambda u + \mu v, w\rangle = \lambda\langle u, w\rangle + \mu\langle v, w\rangle$ sowie $\langle u, \lambda v + \mu w\rangle = \lambda\langle u, v\rangle + \mu\langle u, w\rangle$ für alle $u, v, w \in V$, $\lambda, \mu \in \mathbb{R}$.

b) $\langle , \rangle$ ist *symmetrisch*, d.h. $\langle v, w\rangle = \langle w, v\rangle$ für alle v, w.

c) $\langle , \rangle$ ist *positiv definit*, d.h. $\langle v, v\rangle > 0$ für alle $v \neq 0$.

Ein *euklidischer Raum* $(V, \langle , \rangle)$ besteht aus einem $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt $\langle , \rangle$ auf V.

Beispiele a) Zusammen mit dem sog. *Standardskalarprodukt*

$$\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)\rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

ist $\mathbb{R}^n$ ein euklidischer Raum.

b) Ein weiteres Skalarprodukt auf $\mathbb{R}^2$ ist $\langle (x, y), (x', y')\rangle := xx' + 2yy'$.

c) Dagegen ist $\langle (x, y), (x', y')\rangle := xx' - 2xy' - 2x'y + yy'$ kein Skalarprodukt: Bilinearität und Symmetrie sind schon gegeben, ferner ist $\langle e_1, e_1\rangle = \langle e_2, e_2\rangle = 1 > 0$. Aber $\langle (1, 1), (1, 1)\rangle = 1 - 2 - 2 + 1 = -2 < 0$, d.h. $\langle , \rangle$ ist nicht positiv definit.

d) In der speziellen Relitivitätstheorie setzt man $\|(x, y, z, t)\|^2 = x^2 + y^2 + z^2 - c^2t^2$. Auch dies ist nicht positiv definit.

Definition Sei V ein $\mathbb{C}$-Vektorraum. Ein *Skalarprodukt* auf V ist eine Abbildung $\langle , \rangle \colon V \times V \to \mathbb{C}$, $(v, w) \mapsto \langle v, w\rangle$, die die folgenden Bedingungen erfüllt:

a) $\langle , \rangle$ ist *sesquilinear*, d.h. $\langle u, v\rangle$ ist linear in u und *antilinear* in v, d.h. $\langle \lambda u + \mu v, w\rangle = \lambda\langle u, w\rangle + \mu\langle v, w\rangle$ sowie $\langle u, \lambda v + \mu w\rangle = \bar{\lambda}\langle u, v\rangle + \bar{\mu}\langle u, w\rangle$ für alle $u, v, w \in V$, $\lambda, \mu \in \mathbb{R}$.

b) Es ist $\langle v, w\rangle = \overline{\langle w, v\rangle}$ für alle v, w. Beachten Sie: $\langle v, v\rangle$ ist dann immer reel, denn $\langle v, v\rangle = \overline{\langle v, v\rangle}$. Sind a) und b) erfüllt, so heißt $\langle , \rangle$ *hermitesch*.

c) $\langle , \rangle$ ist *positiv definit*, d.h. die reelle Zahl $\langle v, v\rangle$ ist > 0 für alle $v \neq 0$.

Ein *unitärer Raum* $(V, \langle , \rangle)$ besteht aus einem $\mathbb{C}$-Vektorraum V zusammen mit einem Skalarprodukt $\langle , \rangle$ auf V.

Beispiele a) Zusammen mit dem sog. *Standardskalarprodukt*

$$\langle (z_1, z_2, \ldots, z_n), (w_1, w_2, \ldots, w_n)\rangle := z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n = \sum_{i=1}^{n} z_i\bar{w}_i$$

ist $\mathbb{C}^n$ ein euklidischer Raum.

b) Dagegen ist $\langle z, w\rangle = zw$ kein Skalarprodukt auf $\mathbb{C}$, denn $\langle 1, i\rangle = i \neq \overline{\langle i, 1\rangle} = \bar{i} = -i$.

c) Ist $(V, \langle , \rangle)$ ein unitärer Raum und $U \subseteq V$ ein Unterraum, so ist auch $(U, \langle , \rangle)$ ein unitärer Raum. Das gleiche gilt für euklidische Räume.

Bemerkung Ist $\langle u, v\rangle$ linear in u und symmetrisch, so ist es automatisch bilinear. Ist $\langle u, v\rangle$ linear in u und erfüllt $\langle u, v\rangle = \overline{\langle v, u\rangle}$, so ist es automatisch sesquilinear. Begründung für die zweite Aussage:

$$\begin{aligned}\langle u, \lambda v + \mu w\rangle = \overline{\langle \lambda v + \mu w, u\rangle} &= \overline{\lambda\langle v, u\rangle + \mu\langle w, u\rangle} \\ &= \bar{\lambda} \cdot \overline{\langle v, u\rangle} + \bar{\mu} \cdot \overline{\langle w, u\rangle} = \bar{\lambda}\langle u, v\rangle + \bar{\mu}\langle u, w\rangle\,.\end{aligned}$$

Orthonormalisierung

Definition Sei $(V, \langle , \rangle)$ ein euklidischer bzw. unitärer Raum.

a) Seien $u, v \in V$. Gilt $\langle u, v\rangle = 0$, so sagt man, dass u, v senkrecht aufeinander stehen. Bezeichnung: $u \perp v$. Beachten Sie: Da $\langle , \rangle$ symmetrisch bzw. hermitesch ist, gilt $\langle u, v\rangle = 0$ genau dann, wenn $\langle v, u\rangle = 0$ gilt. Das heißt, es ist $u \perp v$ genau dann, wenn $v \perp u$.

b) Ein System $v_1, \ldots, v_r$ von Vektoren in V heißt *orthogonal*, wenn $v_i \perp v_j$ gilt für alle $i \neq j$. Gilt außerdem $\langle v_i, v_i\rangle = 1$ für alle i, so heißt das System *orthonormal*.

Beispiele In $\mathbb{R}^3$ gelten u.a.

a) $(1,2,4) \perp (2,1,-1)$ und $(1,3,1) \not\perp (2,1,-1)$.

b) $(1,1,1), (-1,2,0), (0,0,0)$ ist ein orthogonales System.

c) $(1,0,0), (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ ist ein orthonormales System.

In $\mathbb{R}^n$ und in $\mathbb{C}^n$ ist die Standardbasis $e_1, e_2, \ldots, e_n$ eine orthonormales System.

Orthonormalisierungssatz *Jeder endlich dimensionaler euklidischer bzw. unitärer Raum besitzt mindestens eine Orthonormalbasis, d.h. eine Basis, die gleichzeitig ein orthonormales System ist. Jedes orthonormale System kann zu einer Basis fortgesetzt werden.*

Beweis. Sei $(V, \langle , \rangle)$ ein solcher Raum. Ob V euklidisch oder unitär ist, für jeden Vektor v und Skalar λ gilt $\langle \lambda v, \lambda v \rangle = |\lambda|^2 \langle v, v \rangle$.

Induktionsbeweis über $n = \dim V$. Sei $u_1, \ldots, u_n$ eine Basis von V. Sei $v_1 = \frac{1}{\sqrt{\langle u_1, u_1 \rangle}} u_1$. Dann $\langle v_1, v_1 \rangle = 1$. Für $2 \leq i \leq n$ sei $u_i' = u_i - \langle u_i, v_1 \rangle v_1$. Dann

$$\langle u_i', v_1 \rangle = \langle u_i - \langle u_i, v_1 \rangle v_1, v_1 \rangle = \langle u_i, v_1 \rangle - \langle u_i, v_1 \rangle \cdot \langle v_1, v_1 \rangle = 0 \,, \text{ da } \langle v_1, v_1 \rangle = 1.$$

Also $u_i' \perp v_1$ für alle $2 \leq i \leq n$. Außerdem ist $v_1, u_2', u_3', \ldots, u_n'$ eine Basis von V. Nach Induktionsannahme kann man aus $u_2', \ldots, u_n'$ eine Orthonormalbasis $v_2, \ldots, v_n$ für den $(n-1)$-dimensionalen Unterraum $U := \text{Spann}(u_2', \ldots, u_n')$ machen. Für $2 \leq i \leq n$ ist v_i eine Linearkombination von $u_2', \ldots, u_n'$, weshalb $\langle v_i, v_1 \rangle = 0$ ist – wegen Linearität von $\langle u, v \rangle$ in u. Induktionsanfang $n = 0$: die leere Menge ist eine Orthonormalbasis.

Letzter Teil: Jedes orthonormale System $u_1, \ldots, u_r$ ist linear unabhängig, denn: ist $\sum_{i=1}^r \lambda_i u_i = 0$, so nimmt man das Skalarprodukt mit u_j und erhält $\sum_{i=1}^r \lambda_i \langle u_i, u_j \rangle = 0$. Wegen Orthonormalität ist $\langle u_i, u_j \rangle = \delta_{ij}$, also $\lambda_j = 0$. Dies kann man für jedes j machen.

Man setzt dann das orthonormale System $u_1, \ldots, u_r$ zu einer Basis $u_1, \ldots, u_n$ von V fort und wendet dann das obige Verfahren an, um eine Orthonormalbasis $v_1, \ldots, v_n$ zu erhalten. Da $u_1, \ldots, u_r$ orthonormal sind, ist $v_1 = u_1$ und $u_i' = u_i$ für $2 \leq i \leq r$. Es folgt per Induktion, dass $v_i = u_i$ für $i \leq r$. ∎

Das Orthonormalisierungsverfahren nach Gram–Schmidt Die Beweismethode des Satzes ist konstruktiv. Hier ist eine leicht geänderte Fassung, die sich für konkrete Berechnungen eignet.

Input: Ein euklidischer bzw. unitärer Raum $(V, \langle , \rangle)$ und eine Basis $u_1, \ldots, u_n$ von V.

Output: Eine orthonormale Basis $v_1, \ldots, v_n$ von V. Sind $u_1, \ldots, u_r$ orthonormal, so ist $v_i = u_i$ für $i \leq r$.

Zuerst wird eine orthogonale Basis $w_1, \ldots, w_n$ berechnet, am Ende wird dies normalisiert.

1. Schritt: $w_1 = u_1$ setzen und $\langle w_1, w_1\rangle$ berechnen.

2. Schritt: $w_2 = u_2 - \frac{\langle u_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1$ setzen und $\langle w_2, w_2\rangle$ berechnen.

i. Schritt: Für $3 \leq i \leq n$ setzt man

$$w_i = u_i - \sum_{j=1}^{i-1} \frac{\langle u_i, w_j\rangle}{\langle w_j, w_j\rangle} w_j$$

und berechnet $\langle w_i, w_i\rangle$.

Letzter Schritt: Für jedes $1 \leq i \leq n$ setzt man $v_i = \frac{1}{\sqrt{\langle w_i, w_i\rangle}} w_i$.

Bemerkungen a) Man kann auch v_i sofort nach Berechnung von w_i berechnen, dann setzt man $w_i = u_i - \sum_{j=1}^{i-1} \langle u_i, v_j\rangle v_j$. Einerseits eine einfachere Formel, anderseits muss man bereits sehr früh mit Quadratwurzeln rechen.

b) Ersetzt man $\langle u_i, w_j\rangle$ durch $\langle w_j, u_i\rangle$, so funktioniert die Methode im unitären Fall nicht. Im euklidischen Fall gibt es keine Unterschied.

c) Per Induktion sieht man, dass w_i in Spann$(u_1, \ldots, u_i)$ liegt, und dass die Vektoren $u_1, \ldots, u_i$ in Spann$(w_1, \ldots, w_i)$ liegen. Somit sind $w_1, \ldots, w_i$ linear unabhängig.

d) Nach Konstruktion von w_i gilt $\langle w_i, w_j\rangle = 0$ für alle $j < i$. Das System $w_1, \ldots, w_n$ ist also eine orthogonale Basis, und das Verfahren funktioniert.

Beispiel Sei $V \subseteq \mathbb{R}^4$ der Lösungsraum der Gleichung $x_1 + x_2 + x_3 + x_4 = 0$, ein euklidischer Raum bezüglich des Standardskalarprodukts auf $\mathbb{R}^4$. Wir konstruieren eine Orthonormalbasis von V.

Hierzu wählen wir eine Basis $u_1 = (1, -1, 0, 0)$, $u_2 = (1, 0, -1, 0)$, $u_3 = (1, 0, 0, -1)$ von V. Zuerst setzen wir $w_1 = u_1$, also $w_1 = (1, -1, 0, 0)$ und $\langle w_1, w_1\rangle = 2$. Jetzt setzen wir $w_2 = u_2 - \frac{\langle u_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1$, d.h. $w_2 = u_2 - \frac{1}{2} w_1 = \left(\frac{1}{2}, \frac{1}{2}, -1, 0\right)$. Um vorerst Brüche zu vermeiden, multiplizieren wir mit 2, also $w_2 = (1, 1, -2, 0)$ und $\langle w_2, w_2\rangle = 6$. Man beachte, dass $w_1 \perp w_2$ tatsächlich gilt.

Nun setzen wir

$$\begin{aligned} w_3 &= u_3 - \frac{\langle u_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle u_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 \\ &= u_3 - \frac{1}{2} w_1 - \frac{1}{6} w_2 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1\right), \end{aligned}$$

mit $\langle w_3, w_3\rangle = \frac{4}{3}$. Die Orthonormalbasis lautet also $v_1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0\right)$, $v_2 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, 0\right)$, $v_3 = \left(\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, -\frac{3}{2\sqrt{3}}\right)$.

Das orthogonale Komplement

Definition Sei $(V, \langle , \rangle)$ ein euklidischer bzw. unitärer Raum. Ist $T \subseteq V$ eine Teilmenge, so setzt man

$$T^\perp := \{u \in V \mid u \perp T, \text{ d.h. } u \perp v \text{ für jedes } v \in T\}\,.$$

Beispiel In $\mathbb{R}^3$ ist ${e_3}^\perp$ die x, y-Ebene durch den Ursprung. Ist $U \subseteq \mathbb{R}^3$ ein zweidimensionaler Unterraum, d.h. eine Ebene durch den Ursprung, so ist $U^\perp$ die Gerade durch den Urpsrung, die senkrecht auf U steht.

Lemma 8.1 *Sei $(V, \langle , \rangle)$ ein euklidischer bzw. unitärer Raum.*

a) Für jede Teilmenge $T \subseteq V$ ist $T^\perp \subseteq V$ ein Unterraum. Ist $v \in T \cap T^\perp$, dann $v = 0$.

b) Ist $U \subseteq V$ ein Unterraum, so ist $V = U \oplus U^\perp$ und $(U^\perp)^\perp = U$.

Beweis. Schreiben wir k für $\mathbb{R}$ bzw. $\mathbb{C}$.

a) Immer gilt $0 \perp T$. Sind $v, w \in T^\perp$, $\lambda, \mu \in k$ und $u \in T$, so ist

$$\langle \lambda v + \mu w, u\rangle = \lambda\langle v, u\rangle + \mu\langle w, u\rangle = \lambda \cdot 0 + \mu \cdot 0 = 0\,.$$

Also $\lambda v + \mu w \in T^\perp$, und $T^\perp$ ist ein Unterraum. Ist $v \in T \cap T^\perp$, dann $\langle v, v\rangle = 0$. Für alle $v \neq 0$ ist aber $\langle v, v\rangle > 0$.

b) Orthonormalisierungssatz: U hat eine Orthonormalbasis $v_1, \ldots, v_r$, und diese lässt sich zu einer Orthonormalbasis $v_1, \ldots, v_n$ von V fortsetzen. Für $j > r$ gilt $v_j \perp U$, denn v_j steht senkrecht auf jedem Element der Basis $v_1, \ldots, v_r$. Also $U' = \text{Spann}(v_{r+1}, \ldots, v_n)$ liegt in $U^\perp$. Wegen $U^\perp \cap U = \{0\}$ und $U \oplus U^\perp \subseteq V$ gilt $\dim U^\perp \leq n - r = \dim U'$. Also $U' = U^\perp$. Wendet man das gleiche Argument auf U' an, so sieht man, dass $(U')^\perp = U$. ∎

Selbstadjungierte Endomorphismen und der Spektralsatz

Definition Sei V ein euklidischer bzw. unitärer Raum, und sei $F\colon V \to V$ ein Endomorphismus von V. Gilt $\langle F(u), v\rangle = \langle u, F(v)\rangle$ für alle $u, v \in V$, so heißt F *selbstadjungiert*.

Wir wollen den folgenden Satz beweisen. Hierfür sind einige Vorarbeiten nötig.

Spektralsatz *Sei V ein endlich-dimensionaler euklidischer bzw. unitärer Raum, und sei F ein selbstadjungierter Endomorphismus von V. Dann: alle Eigenwerte von F sind reell, und es gibt eine Orthonormalbasis von V, die aus Eigenvektoren von F besteht.*

Zuerst drücken wir den Selbstadjungiert-Begriff in der Matrizensprache aus, für das Standardskalarprodukt.

Lemma 8.2 *a) Sei $A \in M_n(\mathbb{R})$ eine Matrix. Dann: Der Endomorphismus L_A des $\mathbb{R}^n$ ist genau dann selbstadjungiert, wenn A symmetrisch ist, d.h. $A^T = A$, d.h. $A_{ji} = A_{ij}$ für alle i, j.*

b) Sei $A \in M_n(\mathbb{C})$ eine Matrix. Dann: Der Endomorphismus L_A des $\mathbb{C}^n$ ist genau dann selbstadjungiert, wenn A hermitesch *ist, d.h. $A^T = \bar{A}$, d.h. $A_{ji} = \bar{A}_{ij}$ für alle i, j.*

Beweis. Wir beweisen den zweiten Teil; der Beweis des ersten Teils ist analog aber einfacher. Für $u = (\lambda_1, \ldots, \lambda_n)$ und $v = (\mu_1, \ldots, \mu_n)$ ist

$$\langle L_A(u), v\rangle = \sum_{i,j=1}^{n} A_{ij}\lambda_j\bar{\mu}_i$$

$$\langle u, L_A(v)\rangle = \sum_{i,j=1}^{n} \lambda_j\bar{A}_{ji}\bar{\mu}_i\,.$$

Da die Werte von λ_i, μ_j beliebig gewählt werden können, folgt: es ist genau dann $\langle u, L_A(v)\rangle = \langle L_A(u), v\rangle$ für alle u, v, wenn $A_{ij} = \bar{A}_{ji}$ für alle i, j. ∎

Für den Spektralsatz müssen wir insbesondere zeigen, dass F immer einen reellen Eigenwert hat, auch im unitären Fall. Im unitären Fall zitieren wir den Fundamentalsatz der Algebra; im euklidischen Fall müssen wir mittels Matrizen argumentieren.

Fundamentalsatz der Algebra *Jedes nichtkonstante Polynom mit komplexen Koeffizienten besitzt mindestens eine komplexe Nullstelle.*

Dieser wichtige Satz wird an dieser Stelle nicht bewiesen. Es gibt einige Beweise, aber alle setzen mehr Hilfsmittel voraus, als im 1. Semester zur Verfügung stehen.

Ein Korollar des Fundamentalsatzes ist die Tatsache, dass jede Matrix in $M_n(\mathbb{C})$ mindestens einen Eigenwert hat: denn das charakteristische Polynom muss eine Nullstelle haben.

Lemma 8.3 *Sei V ein endlich-dimensionaler euklidischer bzw. unitärer Raum. Jeder selbstadjunguerte Endomorphismus F von V hat mindestens einen Eigenwert, und alle Eigenwerte sind reell.*

Das gleiche gilt für symmetrische reelle Matrizen und für komplexe hermitesche Matrizen.

Beweis. Der unitäre Fall: Nach dem Fundamentalsatz der Algebra hat das charakteristische Polynom von F mindestens eine Nullstelle λ. Sei $v \neq 0$ ein Eigenvektor von F mit Eigenwert λ. Dann $\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle F(v), v\rangle$.

Einerseits ist $\langle F(v), v\rangle = \langle v, F(v)\rangle$, da F selbstadjungiert ist. Andererseits ist $\langle F(v), v\rangle = \overline{\langle v, F(v)\rangle}$, da $\langle , \rangle$ hermitesch ist. Folglich ist $\langle F(v), v\rangle$ und somit $\lambda\langle v, v\rangle$ reell. Da $\langle v, v\rangle$ reell und > 0 ist, folgt: λ ist reell.

Matrizen: Ist A eine hermitesche komplexe Matrix, so ist L_A nach Lemma 8.2 ein selbstadjungierter Endomorphismus des $\mathbb{C}^n$. Die Eigenwerte von L_A sind die von A, und umgekehrt. Die Aussage gilt also für A. Ist $A \in M_n(\mathbb{R})$ symmetrisch, so ist A gleichzeitig eine hermitesche Matrix in $M_n(\mathbb{C})$. Die Aussage gilt also für diese A.

Euklidischer Fall: Nach dem Orthonormalisierungssatz hat V eine Orthonormalbasis B. Sei A die Matrix ${}_B\mathrm{M}_B(F)$ von F bezüglich dieser Basis. Bezüglich dieser Basis ist das Skalarprodukt auf V das Standardskalarprodukt auf $\mathbb{R}^n$, und F ist L_A. Nach dem ersten Teil von Lemma 8.2 ist A also symmetrisch. Die Eigenwerte von A sind die Eigenwerte von F, und umgekehrt. ∎

Lemma 8.4 *Sei F ein selbstadjungierter Endomorphismus des euklidischen bzw. unitären Vektorraums V. Sei $U \subseteq V$ ein Unterraum mit der Eigenschaft, dass $F(U) \subseteq U$. Dann $F(U^\perp) \subseteq U^\perp$.*

Beweis. Sei $u \in U$, $v \in U^\perp$. Wir müssen zeigen: es ist $\langle u, F(v)\rangle = 0$. Tatsächlich gilt $\langle u, F(v)\rangle = \langle F(u), v\rangle = 0$, denn F ist selbstadjungiert, $F(u) \in U$ und $v \in U^\perp$. ∎

Beweis des Spektralsatzes. Induktion über $n = \dim(V)$. Nach Lemma 8.3 hat F einen reellen Eigenwert λ. Sei $v \neq 0$ ein Eigenvektor mit diesem Eigenwert. Sei $v_1 = \frac{1}{\sqrt{\langle v,v\rangle}} v$. Dann v_1 ist ein Eigenvektor mit Eigenwert λ, und es ist $\langle v_1, v_2\rangle = 1$.

Sei $U = \mathrm{Spann}(v_1)$. Wegen $F(v_1) = \lambda v_1$ gilt $F(U) \subseteq U$. Wegen Lemma 8.4 ist also $F(U^\perp) \subseteq U^\perp$. Wegen Lemma 8.2 Teil b) ist $V = U \oplus U^\perp$. Nach Induktionsannahme hat $U^\perp$ eine Orthonormalbasis $v_2, \ldots, v_n$, die aus Eigenvektoren mit reellen Eigenwerten besteht. Als ist $v_1, \ldots, v_n$ eine Orthonormalbasis von V, die aus Eigenvektoren mit reellen Eigenwerten besteht. ∎

Rechenverfahren zur Bestimmung einer solchen Basis Man geht in drei Schritten vor:

a) Charakteristisches Polynom von F berechnen, Nullstellen bestimmen: diese sind die Eigenwerte von F.

b) Zu jedem Eigenwert λ eine Basis des Eigenraums $E_\lambda(F)$ berechnen (Gauß-Verfahren).

c) Für jeden Eigenraum eine Orthonormalbasis bestimmen (Gram–Schmidt). Diese Basen zusammenlegen.

Um zu sehen, dass dieses Rechenverfahren funktioniert, benötigen wir ein weiteres Lemma:

Lemma 8.5 *Sei F ein selbstadjungierter Endomorphismus des euklidischen bzw. unitären Raums V. Seien u, v zwei Eigenvektoren von F mit unterschiedlichen Eigenwerten. In diesem Fall gilt $u \perp v$.*

Beweis. Sei λ bzw. μ der Eigenwert von u bzw. v. Nach Lemma 8.3 sind $\lambda, \mu \in \mathbb{R}$. Also

$$\mu\langle u, v\rangle = \langle u, \mu v\rangle = \langle u, F(v)\rangle = \langle F(u), v\rangle = \langle \lambda u, v\rangle = \lambda\langle u, v\rangle\,.$$

Wegen $\lambda \neq \mu$ folgt also $\langle u, v\rangle = 0$. ∎

Orthogonale und unitäre Endomorphismen

Sei V ein euklidischer bzw. unitärer Raum. Wegen des Vergleichs mit dem herkömmlichen Fall $\mathbb{R}^3$ erklärt man die Länge $\|v\|$ eines Vektors $v \in V$ durch $\|v\| := \sqrt{\langle v, v\rangle}$. Im euklidischen Fall erklärt man außerdem den Winkel $\theta \in [0, \pi]$ zwischen zwei Vektoren $v, w \neq 0$ durch $\cos(\theta) = \frac{\langle v,w\rangle}{\|v\|\cdot\|w\|}$.

Definition a) Ein Endomorphismus F eines euklidischen Raums V heißt *orthogonal*, wenn $\|F(v)\| = \|v\|$ gilt für jedes $v \in V$. Ein Endomorphismus F eines unitären Raums V heißt *unitär*, wenn $\|F(v)\| = \|v\|$ gilt für jedes $v \in V$.

Das heißt, ein Endomorphismus ist orthogonal bzw. unitär, wenn er Längen erhält.

b) Eine Matrix $A \in M_n(\mathbb{R})$ heißt *orthogonal*, falls $A \cdot A^T = A^T \cdot A = E_n$ gilt. Eine Matrix $A \in M_n(\mathbb{C})$ heißt *unitär*, falls $A \cdot \overline{A^T} = \overline{A^T} \cdot A = E_n$ gilt.

Lemma 8.6 *a) Sei F ein Endomorphismus des euklidischen bzw. unitären Vektorraums V. Dann: F ist genau dann orthogonal bzw. unitär, wenn $\langle F(v), F(w)\rangle = \langle v, w\rangle$ gilt für alle $v, w \in V$.*

b) Sei $A \in M_n(\mathbb{R})$ bzw. $M_n(\mathbb{C})$. Dann: der Endomorphismus L_A des $\mathbb{R}^n$ bzw. des $\mathbb{C}^n$ ist genau dann orthogonal bzw. unitär (bzgl. des Standardskalarprodukts), wenn die Matrix A orthogonal bzw. unitär ist.

c) Um zu prüfen, dass $A \in M_n(\mathbb{R})$ orthogonal ist, reicht es, eine der Gleichungen $A \cdot A^T = E_n$, $A^T \cdot A = E_n$ nachzuweisen. Für unitäre Matrizen gilt die entsprechende Aussage.

Beweis. a) Wir müssen zeigen: ist $\langle F(v), F(v)\rangle = \langle v, v\rangle$ für alle v, dann $\langle F(v), F(w)\rangle = \langle v, w\rangle$ für alle v, w. Euklidischer Fall: Folgt aus der Polarisierungs-Formel:

$$\langle v+w, v+w\rangle - \langle v, v\rangle - \langle w, w\rangle = 2\langle v, w\rangle\,,$$

die man mittels Bilinearität und Symmetrie nachweist. Unitärer Fall: Diesmal lautet die Polarisierungs-Formel

$$\langle v+w, v+w\rangle - \langle v, v\rangle - \langle w, w\rangle = 2\,\mathrm{Re}\langle v, w\rangle\,.$$

Somit ist $\mathrm{Re}\langle F(v), F(w)\rangle = \mathrm{Re}\langle v, w\rangle$ für alle v, w. Das gleiche gilt für den imaginären Teil, denn $\mathrm{Re}\langle v, iw\rangle = \mathrm{Im}\langle v, w\rangle$.

b), c): Wir behandeln den unitären Fall, der orthogonale Fall ist analog aber einfacher. L_A ist genau dann unitär, wenn $\langle A\cdot v, A\cdot w\rangle = \langle v, w\rangle$ für alle $v = (z_1, \ldots, z_n)$ und $w = (w_1, \ldots, w_n) \in \mathbb{C}^n$. Das heißt, wenn

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n A_{ij} z_j \overline{A_{ik} w_k} = \sum_{i=1}^n z_i \overline{w_i}\,, \quad \text{d.h.}$$
$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n A_{ij} \overline{A_{ik}} z_j \overline{w_k} = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \delta_{ij}\delta_{ik} z_j \overline{w_k}\,.$$

Da z_j, w_k beliebige Werte annehmen dürfen, folgt:

$$\sum_{i=1}^n A_{ij}\overline{A_{ik}} = \sum_{i=1}^n \delta_{ij}\delta_{ik}\,, \quad \text{d.h.} \quad \sum_{i=1}^n A_{ij}\overline{A_{ik}} = \delta_{jk}\,.$$

Dies ist die Aussage $A^T \cdot \overline{A} = E_n$. Das komplex Konjugierte dieser Gleichung lautet $\overline{A^T}\cdot A = E_n$. Für b) reicht es also, dass wir c) zeigen.

Aus $\overline{A^T}\cdot A = E_n$ folgt $\det(\overline{A^T})\det(A) = 1$ wegen der Produktregel für Determinanten. Somit ist $\det(A) \neq 0$, also A hat Rang n und ist invertierbar. Multipliziert man die Gleichung $\overline{A^T}\cdot A = E_n$ von rechts mit A^{-1}, so erhält man $\overline{A^T} = A^{-1}$. Wir sind dann fertig, denn $A\cdot A^{-1} = A^{-1}\cdot A = E_n$. Analog folgt $\overline{A^T} = A^{-1}$ aus $A\cdot\overline{A^T} = E_n$. ∎

Bemerkung Ist $A \in M_n(\mathbb{R})$ orthogonal, so folgt $\det(A)^2 = 1$ aus $A\cdot A^T = E_n$. Also $\det(A) = \pm 1$. Die Matrizen

$$B_+ = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad B_- = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

sind orthogonal, mit $\det(B_+) = +1$, $\det(B_-) = -1$.

Ist $A \in M_n(\mathbb{C})$ unitär, so folgt $|\det(A)| = 1$ aus $A\cdot\overline{A^T} = E_n$. Für jedes $z \in \mathbb{C}$ mit $|z| = 1$ ist die matrix $B_z = \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}$ unitär mit $\det(B_z) = z$.

Definition Die *orthogonale Gruppe* $O(n)$, die *spezielle orthogonale Gruppe* $SO(n)$, die *unitäre Gruppe* $U(n)$ und die *spezielle unitäre Gruppe* $SU(n)$ werden wie folgt definiert:

$$\begin{aligned} O(n) &= \{A \in M_n(\mathbb{R}) \mid A \text{ ist orthogonal}\} \\ SO(n) &= \{A \in O(n) \mid \det(A) = 1\} \\ U(n) &= \{A \in M_n(\mathbb{C}) \mid A \text{ ist unitär}\} \\ SU(n) &= \{A \in U(n) \mid \det(A) = 1\}\,. \end{aligned}$$

Hilfssatz 7 *Diese vier Gruppen sind tatsächlich Gruppen, bezüglich Matrixmultiplikation.*

Beweis. Matrixmultiplikation ist bekanntlich assoziativ. Die Einheitsmatrix liegt in allen vier Gruppen und ist das jeweilige neutrale Element. Was wir noch zeigen müssen ist: ist G eine dieser vier Gruppen, und sind A, B Elemente aus G, so liegen auch AB und A^{-1} in G. Ist A unitär, dann $A^{-1} = \overline{A^T}$ und $\overline{(A^{-1})^T} = A$. Wegen $\overline{A^T}A = E_n$ ist also auch A^{-1} unitär. Sind A, B unitär, so ist $\overline{(AB)^T} = \overline{B^T} \cdot \overline{A^T}$ und deshalb $AB \cdot \overline{(AB)^T} = AB \cdot \overline{B^T} \cdot \overline{A^T} = AE_n\overline{A^T} = E_n$, d.h. AB ist unitär. Der orthogonale Fall ist analog. Ist $\det(A) = \det(B) = 1$, dann $\det(A^{-1}) = \det(AB) = 1$ aufgrund der Produktregel. ∎

Satz 8.7 (Die Hauptachsentransformation) *a) Ist $A \in M_n(\mathbb{R})$ symmetrisch, so gibt es ein $S \in SO(n)$ derart, dass $S^{-1}AS$ Diagonalgestalt hat.*

b) Ist $A \in M_n(\mathbb{C})$ hermitesch, so gibt es ein $S \in SU(n)$ derart, dass $S^{-1}AS$ Diagonalgestalt hat. Außerdem sind alle Einträge dieser Diagonalmatrix reell.

Beweis. Der komplexe Fall: L_A ist nach Lemma 8.2 ein selbstadjungierter Endomorphismus des $\mathbb{C}^n$. Nach dem Spektralsatz gibt es also eine Orthonormalbasis $b_1, \ldots, b_n$ des $\mathbb{C}^n$ derart, dass jedes b_i ein Eigenvektor von A ist. Sei T die Matrix, deren $i$te Spalte der Vektor $b_i \in \mathbb{C}^n$ ist. Da die Spalten von T orthonormal sind, liegt T in $U(n)$. Da die Spalten von T Eigenvektoren von A sind, ist $T^{-1}AT$ eine Diagonalmatrix. Sei $z = \det(T)$, dann $|z| = 1$. Ersetzt man den ersten Basisvektor b_1 durch $\frac{1}{z}b_1$, so liegt eine Orthonormalbasis weiterhin vor. Sei S die entsprechende Matrix, d.h. die erste Spalte von S ist $\frac{1}{z}$ mal die erste Spalte von T, und für $j \geq 2$ sind die $j$ten Spalten von S und T gleich. Da auch die Spalten von S orthonormal sind und Eigenvektoren sind, ist $S \in U(n)$, und $S^{-1}AS$ ist diagonal. Nach Konstruktion von S ist außerdem $\det(S) = 1$, d.h. $S \in SU(n)$. Der reeller Fall ist analog.

Letzter Teil: Da A hermitesch und S unitär ist, ist auch $S^{-1}AS$ hermitesch. Die Diagonaleinträge einer hermiteschen Matrix sind reell. ∎

Lemma 8.8 *a) $SO(2)$ besteht aus allen Matrizen der Art $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ mit $\theta \in \mathbb{R}$. Diese Matrix entspricht eine Drehung um 0 durch θ gegen den Uhrzeigersinn. Ist $\sin\theta \neq 0$, so hat die Matrix keine Eigenvektoren.*

b) Ist $A \in O(2) \setminus SO(2)$, dann gibt es ein $\theta \in \mathbb{R}$ mit $A = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$. Diese Matrix enspricht eine Spiegelung in der Gerade durch 0 mit Richtung $(\cos\frac{\theta}{2}, \sin\frac{\theta}{2})$. Sie hat die Eigenwerte 1 und -1.

c) Ist $A \in O(3)$, so hat es mindestens einen Eigenwert. Nach einem Basiswechsel erhält A den Gestalt $\begin{pmatrix}\pm 1 & 0\\ 0 & B\end{pmatrix}$, wobei B eine Matrix aus $O(2)$ ist.

Beweis. a) Tatsächlich ist die Matrix $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ orthogonal, und die Determinante beträgt 1. Liegt $A = \left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) \in SO(2)$, so ist $\left(\begin{smallmatrix} a & c\\ b & d\end{smallmatrix}\right)\cdot\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0\\ 0 & 1\end{smallmatrix}\right)$ und $ad - bc = 1$, also

$$a^2 + c^2 = 1 \qquad b^2 + d^2 = 1 \qquad ab + cd = 0 \qquad ad - bc = 1$$

Aufgrund der ersten beiden Gleichungen gibt es Zahlen θ, ϕ mit $a = \cos\theta$, $c = \sin\theta$, $b = \cos\phi$, $d = \sin\phi$. Wegen $ad - bc = 1$ ist $\sin(\phi - \theta) = 1$, also $\phi = \theta + \frac{\pi}{2}$. Deshalb ist $b = -\sin\theta$, $d = \cos\theta$.

Das charakteristische Polynom der Matrix ist $X^2 - 2\cos\theta X + 1 = (X - \cos\theta)^2 + \sin^2\theta$. Für $\sin\theta \neq 0$, so verschwindet das Polynom nirgendwo.

b) Aus $A \in O(2) \setminus SO(2)$ folgt $\det(A) = -1$. Sei $B = A \cdot \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. Dann $A = B \cdot \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, und $B \in SO(2)$. Nach Teil a) gibt es also ein θ mit $A = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$. Dann ist $(\cos\frac{\theta}{2}, \sin\frac{\theta}{2})$ ein Eigenvektor mit Eigenwert 1, und $(-\sin\frac{\theta}{2}, \cos\frac{\theta}{2})$ ist ein Eigenvektor mit Eigenwert -1. Diese Eigenvektoren stehen senkrecht aufeinander, und A entspricht eine Spiegelung in der Gerade durch 0 in der Richtung des Eigenvektors mit Eigenwert $+1$.

c) Das charakteristische Polynom hat Grad 3 und ist somit ungerade. Somit gibt es mindestens eine Nullstelle. Als Nullstellen kommen nur $+1$ und -1 in Betracht. Sei b_1, b_2, b_3 eine Orthonormalbasis des $\mathbb{R}^3$ derart, dass b_1 ein Eigenvektor von A ist. Nach Wechsel zu dieser Basis hat A die angekündigte Gestalt. ∎

9 Affine Unterräume und Affine Abbildungen

Geraden und Ebenen im $\mathbb{R}^3$ sind nur dann Untervektorräume, wenn sie den 0 enthalten. Dagegen schließt der Begriff „affiner Unterraum“ alle Geraden und Ebenen mit ein.

Zwei gleichwertige Wege, eine Ebene E anzugeben, sind

- Ein Punkt $P \in E$ sowie zwei Richtungen v, w in E.
- Dreipunktform: Drei Punkte P, Q, R auf E, die nicht auf einer Gerade liegen.

Im ersten Fall ist

$$E = \{Q \mid \overrightarrow{OQ} = \overrightarrow{OP} + \mu v + \nu w,\ \mu, \nu \in \mathbb{R}\}\,.$$

In der Sprache der Linearen Algebra sagt man, dass E die *Nebenklasse* $E = \overrightarrow{OP} + U$ des zweidimensionalen Untervektorraums $U = \text{Spann}(v, w)$ ist.

Vom ersten zum zweiten Fall kommt man, indem man Q, R durch $\overrightarrow{OQ} = \overrightarrow{OP} + v$, $\overrightarrow{OR} = \overrightarrow{OP} + w$ definiert. Vom zweiten zum ersten Fall kommt man, indem man $v = \overrightarrow{PQ}$, $w = \overrightarrow{PR}$ setzt. Somit liegt S genau dann auf E, wenn es Skalaren μ, ν gibt mit

$$\begin{aligned}
&\overrightarrow{OS} = \overrightarrow{OP} + \mu\overrightarrow{PQ} + \nu\overrightarrow{PR}\ \text{, d.h.}\\
&\overrightarrow{OS} = (1-\mu-\nu)\overrightarrow{OP} + \mu\overrightarrow{OQ} + \nu\overrightarrow{OR}\ \text{, d.h.}\\
&\overrightarrow{OS} = \lambda\overrightarrow{OP} + \mu\overrightarrow{PQ} + \nu\overrightarrow{PR} \quad \text{mit } \lambda+\mu+\nu = 1.
\end{aligned}$$

Bemerkung In diesem Kapitel ist es einfacher, sich auf Körper k zu beschränken, die $1 + 1 \neq 0$ erfüllen. Häufig schreibt man diese Bedingung als „$\text{char}(k) \neq 2$“. Bisher kennen wir nur einen Körper, wo $1 + 1 = 0$ gilt: $\mathbb{F}_2$. Meistens arbeiten wir mit $\mathbb{R}$ oder mit $\mathbb{C}$, dort gilt auf jedem Fall $1 + 1 \neq 0$.

1. Definition Sei k ein Körper mit $1 + 1 \neq 0$, und sei V ein k-Vektorraum. Eine Teilmenge $A \subseteq V$ heißt ein *affiner Unterraum*, wenn gelten:

a) Sind $a, b \in A$ und $\lambda, \mu \in k$ mit $\lambda + \mu = 1$, so liegt auch $\lambda a + \mu b$ in A.

b) $A \neq \emptyset$

Vorsicht: Für *manche* Zwecke ist es sinnvoll, die leere Menge als einen -1-dimensionalen affinen Unterraum zu betrachten.

Beispiel Jede Ebene und jede Gerade im $\mathbb{R}^3$ ist ein affiner Unterraum.

Ist $C \in M(m \times n, k)$ eine Matrix und $b \in k^m$ ein Vektor derart, dass das Gleichungssystem $C \cdot v = b$ lösbar ist, so ist der Lösungsraum $L := \text{LR}(C, b) = \{v \in k^n \mid C \cdot v = b\}$ ein affiner Unterraum des k^n: denn ist $C \cdot u = C \cdot v = b$ und $\lambda + \mu = 1$, so ist $C \cdot (\lambda u + \mu v) = \lambda C \cdot u + \mu C \cdot v = \lambda b + \mu b = b$.

Lemma 9.1 *Sei V ein k-Vektorraum, wobei $1+1 \neq 0$ in k. Sei $A \subseteq V$ ein affiner Unterraum. Seien $a_1, \ldots, a_n$ Elemente von A, und seien $\lambda_1, \ldots, \lambda_n \in k$ Skalare mit $\sum_{i=1}^n \lambda_i = 1$. Dann liegt die Summe $\sum_{i=1}^n \lambda_i a_i$ in A.*

Beweis. Induktion über n. Für $n = 1$ ist nichts zu beweisen, der Fall $n = 2$ folgt aus der Definition. Sei also $n \geq 3$. Ist irgendein $\lambda_i = 0$, so können wir einen Term streichen, das Ergebnis folgt aus der Induktionsannahme. Seien also alle $\lambda_i \neq 0$.

Gilt $\lambda_1 + \lambda_2 = \lambda_1 + \lambda_3 = 0$, so ist $\lambda_2 + \lambda_3 = -2\lambda_1 \neq 0$, denn $\lambda_1 \neq 0$ und $1 + 1 \neq 0$. Nachdem wir zur Not umnummeriert haben, können wir also davon ausgehen, dass $\lambda_1 + \lambda_2 \neq 0$. Dann $\sum_{i=1}^n \lambda_i a_i = \mu b + \sum_{i=3}^n \lambda_i a_i$, wobei $\mu = \lambda_1 + \lambda_2$ und $b = \frac{\lambda_1}{\mu} a_1 \frac{\lambda_2}{\mu} a_2$. Also $b \in A$, $\mu + \sum_{i=3}^n \lambda_i = 1$, und $\sum_{i=1}^n \lambda_i a_i \in A$. ∎

Lemma 9.2 *Sei A ein affiner Unterraum des k-Vektorraums V, wobei $1+1 \neq 0$ im Körper k. Sei $U := \{b - a \mid a, b \in A\}$, eine Teilmenge von V. Dann:*

a) U ist ein Untervektorraum von V.

b) Für jedes $a \in A$ ist A die Nebenklasse $a + U$, d.h. $A = \{a + u \mid u \in U\}$.

c) Ist umgekehrt $W \subseteq V$ ein Untervektorraum und $b \in V$ ein Vektor, so ist die Nebenklasse $B := b + W = \{b + w \mid w \in W\}$ ein affiner Unterraum des V.

Definition Man nennt U den zu A assoziierten Untervektorraum des V, und definiert die Dimension von A durch $\dim(A) = \dim(U)$.

Beweis. Für $a \in A$ setzen wir $U_a = \{b - a \mid b \in A\}$. Es ist dann $A = a + U_a$, und $U_a \subseteq U$. Für a) und b) reicht es, zu zeigen, dass U_a ein Untervektorraum ist, und dass $U \subseteq U_a$ gilt.

Wegen $a+0 = a \in A$ ist $0 \in U_a$. Seien $u, v \in U_a$ und $\lambda, \mu \in k$ mit $\lambda+\mu = 1$. Sei $b = a+u$, $c = a+v$. Nach Definition von U_a liegen b, c in A. Dann $a + \lambda u + \mu v = \lambda(u + a) + \mu(v + a) = \lambda b + \mu c \in A$. Somit ist $\lambda u + \mu v \in U_a$, und U_a ist ein Untervektorraum.

Ist $u \in U$, so gibt es $b, c \in A$ mit $u = c - b$. Dann $a + u = a - b + c$. Die Summe der Koeffizienten ist $1 + (-1) + 1 = 1$, also liegt $a + u$ in A, weshalb $u \in U_a$ gilt.

Teil c): Es ist $b = b + 0 \in b + W$, also $b + W \neq \emptyset$. Sind $a_1, \ldots, a_n \in b + W$, so gibt es $w_1, \ldots, w_n \in W$ mit $a_i = b + w_i$. Seien $\lambda_1, \ldots, \lambda_n \in k$ mit $\sum_{i=1}^n \lambda_i = 1$. Dann

$$\sum_{i=1}^n \lambda_i a_i = \sum_{i=1}^n \lambda_i (b + w_i) = b + \sum_{i=1}^n \lambda_i w_i \,.$$

Also $\sum_{i=1}^n \lambda_i a_i \in A$, denn $\sum_{i=1}^n \lambda_i w_i \in W$. ∎

2. Definition Sei k ein beliebiger Körper, V ein k-Vektorraum und $A \subseteq V$ eine Teilmenge. Gibt es ein $a \in V$ und einen Untervektorraum $U \subseteq V$ derart, dass $A = a + U$ gilt, so heißt A ein affiner Unterraum von V.

Der Beweis von Lemma 9.2 Teil c) zeigt, dass Lemma 9.1 jetzt für alle Körper k gilt. Die ersten Teile des Lemmas zeigen, dass die beiden Definitionen im Fall $1 + 1 \neq 0$ überseinstimmen.

Lemma 9.3 *Sei k ein beliebiger Körper. Sei $C \cdot x = b$ ein lösbares lineares Gleichungssystem, mit $C \in M(m \times n, k)$ und $b \in k^m$. Sei $A \subseteq k^n$ der Lösungsraum $A = \mathrm{LR}(C, b)$. Dann: A ist ein affiner Unterraum, und es ist*

$$\dim(A) + \mathrm{Rang}(C) = n\,.$$

Umgekehrt lässt sich jeder affine Unterraum des k^n als einen solchen Lösungsraum realisieren.

Beweis. Sei $U \subseteq k^n$ der Nullraum $U = \mathrm{LR}(C, 0)$ der Matrix C. Dann $U = \mathrm{Kern}(L_C)$, und nach der Dimensionsformel gilt $\dim(U) + \dim \mathrm{Bild}(L_C) = n$, d.h. $\dim(U) = n - \mathrm{Rang}(C)$. Sei $a \in k^n$ eine Lösung von $C \cdot x = b$, d.h. $a \in A$. Für jedes $u \in U$ ist $C \cdot (a + u) = C \cdot a + C \cdot u = b + 0 = b$, also $a + U \subseteq A$. Ist dagegen $a' \in A$, so ist $C \cdot (a' - a) = 0$, weshalb $a' \in U$. Also $A = a + U$ und deshalb $\dim(A) = \dim(U)$.

Letzter Teil: Sei $A = a + U$ ein affiner Unterraum des k^n. Nach Korollar 4.5 gibt es eine lineare Abbildung $f\colon k^n \to k^n$ mit $U = \mathrm{Kern}(f)$. Nach Lemma 4.6 Teil a) gibt es eine Matrix $C \in M_n(k)$ mit $f = L_C$. Somit ist U der Nullraum von f, und A ist der Lösungsraum des Gleichungssystems $C \cdot x = f(a)$. ∎

Affine Abbildungen

Definition Sei k ein Körper. Sei A bzw. B ein affiner Unterraum V bzw. W. Eine Abbildung $f\colon A \to B$ heißt eine *affine Abbildung*, wenn für alle $n \geq 1$, alle $a_1, \ldots, a_n \in A$ und alle Skalare $\lambda_1, \ldots, \lambda_n \in k$ mit $\sum_{i=1}^n \lambda_i = 1$ gilt

$$f\left(\sum_{i=1}^n \lambda_i a_i\right) = \sum_{i=1}^n \lambda_i f(a_i)\,.$$

Lemma 9.4 *a) Sei $f\colon A \to B$ eine affine Abbildung. Sei $T \subseteq V$ bzw. $U \subseteq W$ der zu A bzw. B assoziierte Untervektorraum. Dann gibt es genau eine lineare Abbildung $g\colon T \to U$ derart, dass für ein (und sogar für jedes) $a \in A$ gilt $f(a + t) = f(a) + g(t)$ für alle $t \in T$.*

b) Umgekehrt ist $f\colon A \to B$, $f(a+t) = b+g(t)$ eine affine Abbildung für jedes $b \in B$ und für jede lineare Abbildung $g\colon T \to U$.

c) *Gilt* $1+1 \neq 0$ *in* k*, so ist jede Abbildung* $g\colon A \to B$ *affin, die* $g(\lambda a+\lambda' a') = \lambda g(a)+\lambda' g(a')$ *erfüllt für alle* $a, a' \in A$ *und für alle* $\lambda, \lambda' \in k$ *mit* $\lambda+\lambda' = 1$*.*

Beweis. a) Eindeutigkeit von g: ist $f(a+t) = f(a) + g(t)$, und ist $a' \in A$, dann ist $f(a') + g(t) = f(a') - f(a) + f(a+t)$. Da f affin ist, gilt $f(a') - f(a) + f(a+t) = f(a' - a + (a+t))$, d.h. $f(a') + g(t) = f(a'+t)$. Somit ist g eindeutig, und $f(a+t) = f(a) + g(t)$ gilt für alle $a \in A$, sofern es für mindestens ein a gilt.

Existenz: Man wähle ein $a \in A$ und definiere $g\colon T \to W$ durch $g(t) = f(a+t) - f(a)$. Da $f(a+t), f(a)$ beide in B liegen, liegt $g(t)$ in U. Einerseits ist $f(a+\lambda s+\mu t) = f(a) + g(\lambda s + \mu t)$. Andrerseits ist

$$\begin{aligned} f(a+\lambda s+\mu t) &= f(\lambda(a+s) + \mu(a+t) + (1-\lambda-\mu)a) \\ &= \lambda f(a+s) + \mu f(a+t) + (1-\lambda-\mu) f(a) \\ &= \lambda f(a) + \lambda g(s) + \mu f(a) + \mu g(t) + (1-\lambda-\mu) f(a) \\ &= f(a) + \lambda g(s) + \mu g(t)\,. \end{aligned}$$

b) Betrachten wir n Elemente $a+t_1, \ldots, a+t_n$ aus A. Ist $\sum_{i=1}^n \lambda_i = 1$, so ist

$$\begin{aligned} f\left(\sum_{i=1}^n \lambda_i (a+t_i)\right) = f\left(a + \sum_{i=1}^n \lambda_i t_i\right) &= b + g\left(\sum_{i=1}^n \lambda_i t_i\right) = b + \sum_{i=1}^n \lambda_i g(t_i) \\ &= \sum_{i=1}^n \lambda_i (b + g(t_i)) = \sum_{i=1}^n \lambda_i f(a+t_i)\,. \end{aligned}$$

c) Man benutzt die Beweismethode aus Lemma 9.1. ∎

Beispiel Wir betrachten k^n als einen affinen Unterraum von sich selbst und betrachten affine Abbildungen $f\colon k^n \to k^n$. Nach Lemma 9.4 gibt es zu jeder solchen affinen Abbildung f eine lineare Abbildung $g\colon k^n \to k^n$ mit $f(v) = f(0) + g(v)$.

Die affinen Abbildungen $f\colon k^n \to k^n$ sind also alle Abbildungen $f(v) = a + A \cdot v$, wobei $a \in k^n$ und $A \in M_n(k)$. Jede affine Abbildung lässt sich faktorisieren als $f = h \circ g$, wobei g der lineare Endomorphismus $g(v) = A \cdot v$ ist, und h die Verschiebung $h(v) = a + v$ ist.