Albert-Ludwigs-Universität Freiburg

Sommersemester 2016

Lineare Algebra

Skript zum Brückenkurs

Dozent:
Maximilian Gerhards

Version vom 11. August 2016
Maximilian Gerhards

Inhaltsverzeichnis

1 Mengen mit Verknüpfung

In der Algebra geht es im Wesentlichen um Mengen mit Verknüpfung. Eine Verknüpfung $*$ ist dabei eine Abbildung, die je zwei Elementen a und b der Menge ein drittes, geschrieben $a * b$, zuordnet. Die Definition ist sehr weit gefasst; Verknüpfungen sind euch schon oft begegnet. Schon aus der ersten Klasse bekannt ist natürlich die Addition, die aus zwei natürlichen Zahlen die Summe $a + b$ bildet. Später wurde die Addition auch auf anderen Mengen, bis hin zu der Menge der reellen Zahlen, definiert. Das ist auch gleich eines der wichtigsten Beispiele. Eine andere Verknüpfung ist aber zum Beispiel auch die Verkettung oder Komposition von Abbildungen, die zwei Abbildungen f und g von einer Menge in sich selbst zu der neuen Abbildung $x \mapsto f(g(x))$ verknüpft.
Man sieht, dass sich die beiden Verknüpfungen in einer wesentlichen Eigenschaft unterscheiden: Bei der Addition natürlicher oder reeller Zahlen spielt die Reihenfolge der beiden Elemente a und b keine Rolle, bei der Verkettung dagegen ist die Reihenfolge der Funktionen sehr wohl entscheidend, wie man am Beispiel der Funktionen $f : \mathbb{R} \to \mathbb{R}, x \mapsto x + 2$ und $g \colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ leicht sieht: $f \circ g$ bildet 0 auf 2 ab, $g \circ f$ dagegen auf 4.

Definition 1.1. (a) Eine *Verknüpfung* $*$ auf einer Menge X ist eine Abbildung

$$\begin{array}{rcl} X \times X & \to & X \\ (x, y) & \mapsto & x * y \end{array}$$

Man nennt das Paar $(X, *)$ aus der Menge und der Verknüpfung *Menge mit Verknüpfung* oder auch *Magma*.

(b) Die Verknüpfung $*$ heißt *assoziativ*, wenn

$$\forall x, y, z \in X \colon (x * y) * z = x * (y * z).$$

(c) Die Verknüpfung $*$ heißt *kommutativ*, wenn

$$\forall x, y \in X \colon x * y = y * x.$$

Beispiel 1.2. Die Verknüpfung

$$\begin{array}{rcl} \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x, y) & \to & x + y \end{array}$$

ist assoziativ und kommutativ. Dasselbe gilt für die Verknüpfung

$$\begin{array}{rcl} \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x, y) & \to & xy. \end{array}$$

Bemerkung 1.3. Wichtig ist, dass wir für *alle* Paare von Elementen aus der Menge ein Ergebnis der Verknüpfung definiert haben. $(\mathbb{R}, :)$ ist *keine* Menge mit Verknüpfung, $(\mathbb{R} \setminus \{0\}, :)$ dagegen schon.

Da Verknüpfungen so verschieden sein können, werden wir im folgenden meist spezielle Mengen mit Verknüpfungen betrachten.

Definition 1.4. Eine Menge mit Verknüpfung heißt *Halbgruppe*, wenn die Verknüpfung assoziativ ist.

Lemma 1.5. *Sei $(X, *)$ eine Halbgruppe.*
Definieren wir für $x_1, \ldots, x_n \in X$

$$x_1 * x_2 * \ldots * x_n := (\ldots (x_1 * x_2) * \ldots) * x_n,$$

so gilt für alle $m, n \in \mathbb{N} \setminus \{0\}$:

$$\begin{aligned} &\forall x_1, \ldots x_m, y_1 \ldots y_n \in X: \\ &\quad (x_1 * \ldots * x_m) * (y_1 * \ldots y_n) \;=\; x_1 * \ldots * x_m * y_1 * \ldots * y_n. \end{aligned}$$

Induktiv können wir also in einem beliebig geklammerten Ausdruck alle Klammern entfernen.

Beweis. Wir zeigen das über vollständige Induktion nach n:
Induktionsanfang $n = 1$:

$$\forall x_1, y_1 \ldots y_n \in X: \quad (x_1 * \ldots * x_m) * y_1 \overset{\text{Definition}}{=} x_1 * \ldots * x_n * y_1.$$

Induktionsannahme: Für ein $n \in \mathbb{N} \setminus \{0\}$ gilt:

$$\begin{aligned} &\forall x_1, \ldots x_m, y_1 \ldots y_n \in X: \\ &\quad (x_1 * \ldots * x_m) * (y_1 * \ldots * y_n) \;=\; x_1 * \ldots * x_m * y_1 * \ldots * y_n. \end{aligned}$$

Induktionsschritt „$n \to n + 1$":

$$\begin{aligned} &\forall x_1, \ldots x_m, y_1 \ldots y_{n+1} \in X: \\ &\quad (x_1 * \ldots * x_m) * (y_1 * \ldots y_{n+1}) \\ &\qquad \overset{\text{Definition}}{=} \quad (x_1 * \ldots * x_m) * ((y_1 * \ldots * y_n) * y_{n+1}) \\ &\qquad \overset{\text{Assoziativität}}{=} \quad ((x_1 * \ldots * x_m) * (y_1 * \ldots * y_n)) * y_{n+1} \\ &\qquad \overset{\text{Induktionsannahme}}{=} \quad (x_1 * \ldots * x_m * y_1 * \ldots * y_n) * y_{n+1} \\ &\qquad \overset{\text{Definition}}{=} \quad x_1 * \ldots * x_m * y_1 * \ldots * y_n * y_{n+1} \end{aligned}$$

□

Definition 1.6. Ist $(X, *)$ eine Menge mit Verknüpfung, so heißt $e \in X$ *neutrales Element* von $(X, *)$ oder bezüglich $*$, wenn

$$\forall x \in X: \quad e * x = x * e = x.$$

Satz 1.7. *Sei $(X, *)$ eine Menge mit Verknüpfung. Dann gibt es höchstens ein neutrales Element bezüglich $*$.*

Beweis. Seien e und $\tilde{e}$ zwei neutrale Elemente. Dann gilt:

$$e \overset{\tilde{e}\text{ neutral}}{=} e * \tilde{e} \overset{e\text{ neutral}}{=} \tilde{e}$$

□

Definition 1.8. Wenn $(X, *)$ eine Halbgruppe ist und e neutrales Element bezüglich $*$, so nennt man das Tripel $(X, *, e)$ *Monoid*.

Definition 1.9. Sei $(X, *)$ eine Menge mit Verknüpfung und e neutrales Element bezüglich $*$. Sei weiter $x \in X$ gegeben. Man nennt ein Element $y \in Y$

(a) *Linksinverses* von x, wenn $y * x = e$

(b) *Rechtsinverses* von x, wenn $x * y = e$

(c) *Inverses* von x, wenn es sowohl Links- als auch Rechtsinverses ist. In diesem Fall schreibt man $y =: x^{-1}$ (oder bei Unklarheit über die gemeinte Verknüpfung x^{*-1}) und nennt x *invertierbar*.

Lemma 1.10. *Ist $*$ assoziativ, also $(X, *, e)$ ein Monoid, so gilt für zwei Elemente x und y aus X:*

(a) Ist l_x Linksinverses von x und l_y Linksinverses zu y, so ist $l_y * l_x$ Linksinverses zu $x * y$.

(b) Ist r_x Rechtsinverses von x und r_y Rechtsinverses zu y, so ist $r_y * r_x$ Rechtsinverses zu $x * y$.

(c) Ist x^{-1} Inverses von x und y^{-1} Inverses zu y, so ist $y^{-1} * x^{-1}$ Inverses zu $x * y$.

Beweis. (a) Es gilt wegen der Assoziativität

$$(l_y * l_x) * (x * y) = l_y * (l_x * x) * y = l_y * e * y = l_y * y = e.$$

(b) analog

(c) folgt sofort aus (a) und (b)

□

Bemerkung 1.11. Die Reihenfolge dreht sich also beim Invertieren um. Man nennt das auch die *Socke-Schuh-Regel*: Beim Ausziehen von Schuhen und Socken muss man die Reihenfolge des Anziehens umkehren. (Quelle: Wikipedia `https://de.wikipedia.org/wiki/Inverses_Element`)

Satz 1.12. *Ist $*$ assoziativ, also $(X, *, e)$ ein Monoid, so sind für jedes $a \in X$ äquivalent:*

(a) a hat mindestens ein Linksinverses und mindestens ein Rechtsinverses.

(b) a hat mindestens ein Inverses.

(c) a hat genau ein Inverses.

(d) a hat ein Inverses a^{-1} und a^{-1} hat als ein Inverses wieder a.

(e) a hat ein Linksinverses b und b hat selbst ein Linksinverses.

Beweis. (a) ⇒ (b) Sei l ein Linksinverses von a und r ein Rechtsinverses. Dann gilt bereits

$$l = l * e = l * (a * r) = (l * a) * r = e * r = r.$$

Damit ist l sowohl Linksinverses als auch Rechtsinverses, also Inverses von a.

(b) ⇒ (c) Seien b und $\tilde{b}$ Inverse von a. Dann ist b Linksinverses von a und $\tilde{b}$ Rechtsinverses von a und wie eben sieht man, dass die beiden Elemente gleich sind.

(c) $\Rightarrow$ (d) Da $a * a^{-1} = e$, ist a Linksinverses zu a^{-1}. Da $a^{-1} * a = e$, ist es auch Rechtsinverses zu a^{-1}.

(d) $\Rightarrow$ (e) trivial

(e) $\Rightarrow$ (a) Sei c das Linksinverse zu b. Dann gilt:

$$a * b = e * a * b = (c * b) * a * b = c * (b * a) * b = c * e * b = c * b = e.$$

Also ist b auch ein Rechtsinverses zu a.

□

Definition 1.13. Eine *Gruppe* ist ein Monoid $(G, *, e)$, in dem jedes Element ein Linksinverses hat.

Bemerkung 1.14. Ausgeschrieben muss eine Gruppe eine Menge G mit Verknüpfung $*$ sein, die folgende drei Gruppenaxiome erfüllt:

(G1) $*$ ist assoziativ.

(G2) Es gibt ein neutrales Element e bezüglich $*$.

(G3) Für jedes Element $a \in G$ gibt es ein Linksinverses in G.

Bemerkung 1.15. Da in einer Gruppe für jedes Element die Eigenschaft (e) aus Satz 1.12 erfüllt ist, gelten für jedes Element sämtliche der äquivalenten Eigenschaften des Satzes. Insbesondere wissen wir:

(a) Jedes Element a einer Gruppe hat genau ein Inverses a^{-1}.

(b) Für jedes Element a in der Gruppe gilt $(a^{-1})^{-1} = a$.

Lemma 1.16. *Sei $(X, *, e)$ ein Monoid. Dann bildet die Menge $X^\times := X^* := \{a \in X \mid a \text{ hat ein Inverses}\}$ der invertierbaren Elemente, zusammen mit der eingeschränkten Verknüpfung $X^\times \times X^\times \to X^\times, (a, b) \mapsto a * b$ und dem neutralen Element e eine Gruppe.*

Beweis. Zunächst müssen wir prüfen, ob die Verknüpfung auf $X^\times$ überhaupt wohldefiniert ist in dem Sinne, dass für beliebige zwei Elemente $a, b \in X^\times$ auch $a * b$ wieder in $X^\times$ liegt. Das folgt aber sofort aus Lemma 1.10 (c), denn wenn a ein Inverses a^{-1} und b ein Inverses b^{-1} hat, dann ist $b^{-1} * a^{-1}$ das Inverse von $a * b$.
Jetzt müssen wir die drei Gruppenaxiome nachprüfen:

(G1) klar

(G2) Es gilt: $e \in X^\times$, weil aus $e * e = e$ folgt, dass e das Inverse von e ist. Dass e neutrales Element auch von $X^\times$ (nicht nur von X) ist, ist dann klar.

(G3) Wir wissen bereits, dass a ein Linksinverses, sogar ein Inverses in X hat. Die Frage ist aber, ob es auch eines in $X^\times$ hat. Aus Satz 1.12 folgt aber, dass auch a^{-1} ein Inverses hat, nämlich a, also $a^{-1} \in X^\times$.

□

Notation 1.17. Ist $*$ eine assoziative, kommutative Verknüpfung auf einer Menge X, so schreiben wir auch oft stattdessen $+$. Gibt es ein neutrales Element, bezeichnen wir das dann als 0. Gibt es zu einem Element a ein Inverses, so notieren wir das mit $-a$. Statt $b + (-a)$ schreiben wir dann auch kürzer $b - a$.
Ist I eine endliche Menge und $I \to X$, $i \mapsto a_i$ eine Abbildung, so schreiben wir weiter $\sum_{i \in I} a_i$ für die Verknüpfung aller a_i mit $i \in I$. $I = \emptyset$ ist dabei nur dann erlaubt, wenn es ein neutrales Element gibt. In diesem Fall definieren wir $\sum_{i \in \emptyset} a_i = 0$.

Lemma 1.18. *Sei $(G, *, e)$ eine Gruppe. Dann gilt:*

$$\forall a, b \in G\colon \quad \exists! x \in G\colon x * a = b \quad \wedge \quad \exists! y \in G\colon a * y = b.$$

Beweis. Existenz: Setze $x := b * a^{-1}$ und $y := a^{-1} * b$, dann gilt $x * a = b * a^{-1} * a = b$ und $a * y = a * a^{-1} * b = b$.
Eindeutigkeit: Sei $x*a = b$ und $\tilde{x}*a = b$. Dann gilt $x = x*a*a^{-1} = b*a^{-1} = \tilde{x}*a*a^{-1} = \tilde{x}$.
Die zweite Aussage zeigt man analog. □

Bemerkung 1.19. Man kann diese Eigenschaft als „Sudoku-Prinzip“ von Gruppen bezeichnen: Wenn man eine Verknüpfungstafel für eine Gruppe ausfüllt, darf in jeder Zeile und in jeder Spalte jeder Wert b nur genau einmal vorkommen.

Korollar 1.20 (Kürzungsregel in Gruppen)**.** *Sei $(G, *, e)$ eine Gruppe.*
Aus $x * a = \tilde{x} * a$ folgt bereits $x = \tilde{x}$; ebenso folgt aus $a * y = a * \tilde{y}$ bereits $y = \tilde{y}$.
Insbesondere folgt aus $a * x = a$ bereits $x = e$ und aus $y * a = a$ bereits $y = e$.

Beweis. Das folgt sofort aus der Eindeutigkeitsaussage von Lemma 1.18. □

Zum Schluss noch ein wichtiger Untertyp von Gruppen:

Definition 1.21. Eine Gruppe, deren Verknüpfung kommutativ ist, heißt *abelsche Gruppe*.

Notation 1.22. Manchmal verwenden wir die Menge G stellvertretend für eine Gruppe $(G, *, e)$ und behalten im Hinterkopf, dass die Verknüpfung und das neutrale Element auch dazugehören.

2 Homomorphismen und Unterobjekte

Bis jetzt haben wir mathematische Objekte mit verschiedenem Komplexitätsgrad definiert, die alle Mengen mit gewissen Zusatzstrukturen waren. Natürlich können wir zwischen je zwei Mengen Abbildungen definieren. Uns interessieren aber vor allem solche Abbildungen, die auch mit den Zusatzstrukturen „verträglich“ sind.
Solch eine strukturerhaltende Abbildung nennen wir allgemein Homomorphismus. Die genauen Eigenschaften, die der Homomorphismus erfüllen muss, richten sich nach dem mathematischen Objekt, um die es geht.
Die zusätzliche Struktur einer Menge mit Verknüpfung gegenüber einer bloßen Menge ist eben die Verknüpfung. So erklärt sich die folgende Definition:

Definition 2.1. Seien $(X, *)$ und $(\tilde{X}, \star)$ zwei Mengen mit Verknüpfung.
Ein *Homomorphismus von Mengen mit Verknüpfung* oder *Magmahomomorphismus* von $(X, *)$ nach $(\tilde{X}, \star)$ ist eine Abbildung $f\colon X \to \tilde{X}$, die die folgende Eigenschaft erfüllt:

$$\forall x, y \in X\colon \quad f(x * y) = f(x) \star f(y).$$

Eine Halbgruppe hat gegenüber einer Menge mit Verknüpfung keine zusätzliche Struktur, nur eine besondere Eigenschaft. Deshalb gibt es keine besonderen *Halbgruppenhomomorphismen*. In der Literatur ist es aber üblich, Magmahomomorphismen zwischen zwei Halbgruppen als Halbgruppenhomomorphismen zu bezeichnen.
Die zusätzliche Struktur eines Monoids gegenüber einer Halbgruppe ist das neutrale Element. Deshalb die folgende Definition:

Definition 2.2. Seien $(M, *, e)$ und $(\tilde{M}, \star, \tilde{e})$ zwei Monoide.
Ein *Monoidhomomorphismus* von $(M, *, e)$ nach $(\tilde{M}, \star, \tilde{e})$ ist ein Halbgruppenhomomorphismus von $(M, *)$ nach $(\tilde{M}, \star)$, der zusätzlich die folgende Eigenschaft erfüllt:

$$f(e) = \tilde{e}.$$

Ob die Existenz von Inversen zu jedem Element eigentlich eine Zusatzstruktur der Gruppe gegenüber einem Monoid sind, darüber lässt sich streiten. Man kann auf jeden Fall die folgende Definition geben:

Definition 2.3. Seien $(G, *, e)$ und $(\tilde{G}, \star, \tilde{e})$ Gruppen.
Ein *Gruppenhomomorphismus* von $(G, *, e)$ nach $(\tilde{G}, \star, \tilde{e})$ ist ein Monoidhomomorphismus von $(G, *, e)$ nach $(\tilde{G}, \star, \tilde{e})$, für den gilt:

$$\forall a \in G\colon f(a^{*-1}) = (f(a))^{\star-1}.$$

Es spielt aber keine Rolle, ob man die Inversen als Zusatzstruktur auffasst, denn für jeden Monoidhomomorphismus gilt bereits automatisch:

Lemma 2.4. *Seien $(M, *, e)$ und $(\tilde{M}, \star, \tilde{e})$ zwei Monoide und $f\colon M \to \tilde{M}$ ein Monoidhomomorphismus. Sei $a \in M$ beliebig. Dann gilt:*

(a) Ist l ein Linksinverses zu a bezüglich $*$, so ist $f(l)$ ein Linksinverses zu $f(a)$ bezüglich $\star$.

(b) Ist r ein Rechtsinverses zu a bezüglich $*$, so ist $f(r)$ ein Rechtsinverses zu $f(a)$ bezüglich $\star$.

(c) Ist a^{-1} das Inverse zu a bezüglich $*$, so ist $f(a^{-1}) = (f(a))^{-1}$ das Inverse zu $f(a)$ bezüglich $\star$. Insbesondere ist jeder Monoidhomomorphismus zwischen zwei Gruppen ein Gruppenhomomorphismus.

Beweis. (a) Es gilt:

$$f(l) \star f(a) = f(l * a) = f(e) = \tilde{e}.$$

(b) analog

(c) folgt aus (a) und (b).

□

Für Gruppen gilt sogar noch etwas Besseres:

Lemma 2.5. *Jeder Magmahomomorphismus zwischen zwei Gruppen ist bereits ein Gruppenhomomorphismus.*

Beweis. Seien $(G, *, e)$ und $(\tilde{G}, \star, \tilde{e})$ die Gruppen und $f: G \to \tilde{G}$ der Homomorphismus. Da wir bereits wissen, dass jeder Monoidhomomorphismus zwischen zwei Gruppen ein Gruppenhomomorphismus ist, müssen wir nur noch zeigen, dass $f(e) = \tilde{e}$.
Das folgt aber so: Sei $a \in G$ beliebig. Dann ist $f(a) \star f(e) = f(a * e) = f(a)$. Mit Korollar 1.20 muss dann $f(e)$ bereits $\tilde{e}$ sein. □

Bemerkung 2.6. Achtung! Es ist zwar jeder Magmahomomorphismus zwischen Gruppen ein Gruppenhomomorphismus, aber nicht jeder Magmahomomorphismus zwischen Monoiden ein Monoidhomomorphismus. Im Beweis haben wir das Sudoku-Prinzip benutzt, das eine Besonderheit von Gruppen ist und in Monoiden im Allgemeinen nicht gilt.

Bemerkung 2.7. Geht aus dem Kontext hervor, ob jetzt ein Magma-, Monoid-, Gruppen- oder noch ein anderer Homomorphismus gemeint ist, spricht man nur von Homomorphismen.

Gruppenhomomorphismen haben eine sehr nützliche Eigenschaft, die sie normalen Abbildungen voraus haben:

Lemma 2.8. *Seien $(G, *, e)$ und $(H, \star, \tilde{e})$ zwei Gruppen.*
Ein Gruppenhomomorphismus $\phi\colon G \to H$ ist genau dann injektiv, wenn

$$\phi^{-1}(\{\tilde{e}\}) = \{g \in G \mid \phi(g) = \tilde{e}\} = \{e\}.$$

Beweis. „$\Rightarrow$“:
Diese Richtung ist trivial: Wir wissen bereits, dass e unter ϕ auf $\tilde{e}$ abgebildet wird. Wenn ϕ injektiv ist, dann muss es das einzige Element von G sein.
„$\Leftarrow$“:
Seien $g, g' \in G$ mit $\phi(g) = \phi(g')$. Wir müssen zeigen, dass $g = g'$. Es gilt

$$\phi(g^{*-1} * g') = \phi(g^{*-1}) \star \phi(g') = \phi(g)^{\star-1}\phi(g') \overset{\phi(g)=\phi(g')}{=} \tilde{e}.$$

Da nach Voraussetzung e das einzige Element ist, das auf $\tilde{e}$ abgebildet wird, folgt $g^{*-1} * g' = e$, also $g = g * e = g * g^{*-1} * g' = g'$. □

Definition 2.9. Sei ϕ ein Gruppenhomomorphismus von $(G, *, e)$ nach $(H, \star, \tilde{e})$. Dann definiert man den *Kern* von ϕ als die Menge

$$\ker(\phi) := \phi^{-1}(\{\tilde{e}\}) = \{g \in G \mid \phi(g) = \tilde{e}\}.$$

Ihr wisst bereits, dass es Abbildungen $f\colon M \to N$ gibt, die man beidseitig umkehren kann, dass es also Abbildungen $g : N \to M$ und $h : N \to M$ gibt, sodass $f \circ g = \mathrm{id}_N$ und $h \circ f = \mathrm{id}_M$. Analog zur ersten Aussage von Satz 1.12 gilt dann $g = h$ und wir schreiben dafür f^{-1} und nennen das die *Umkehrabbildung* (Achtung: Das ist kein Inverses im Sinne unserer Definition, da bereits die Menge der Abbildungen zwischen M und N und zwischen N und M keine Menge mit Verknüpfung bilden - man kann ja nicht zwei Abbildungen von M nach N verknüpfen, wenn N keine Teilmenge von M ist). Solche Abbildungen nennt man dann bekanntlich *bijektiv*.
Das Analogon für Mengen mit Zusatzstruktur ist der Isomorphismus:

Definition 2.10. Ein *Isomorphismus* zwischen zwei Mengen X und Y mit Zusatzstruktur ist ein Homomorphismus $f : X \to Y$, der bijektiv ist und dessen Umkehrabbildung f^{-1} selbst wieder ein Homomorphismus ist.

Bemerkung 2.11. Hier muss natürlich der Homomorphismus der zur Zusatzstruktur passende und die Umkehrabbildung dieselbe Art von Homomorphismus sein. Ein Gruppenisomorphismus ist also zum Beispiel ein bijektiver Gruppenhomomorphismus, dessen Umkehrabbildung wieder ein Gruppenhomomorphismus ist.

Satz 2.12. *Jeder bijektive Magmahomomorphismus ist bereits ein Magmaisomorphismus. Entsprechendes gilt für Monoidhomomorphismen und Gruppenhomomorphismen.*

Beweis. Sei $f\colon X \to Y$ der Magmahomomorphismus.
Wir müssen zeigen, dass $f^{-1}(a \star b) = f^{-1}(a) * f^{-1}(b)$ für alle $a, b \in Y$.
Es gilt aber

$$\begin{aligned} f^{-1}(a \star b) &\overset{\text{Definition } f^{-1}}{=} f^{-1}\left(f(f^{-1}(a)) \star f(f^{-1}(b))\right) \\ &\overset{f \text{ Homomorphismus}}{=} f^{-1}\left(f(f^{-1}(a) * f^{-1}(b))\right) \\ &\overset{\text{Definition } f^{-1}}{=} f^{-1}(a) * f^{-1}(b). \end{aligned}$$

Den Fall der Gruppenhomomorphismen haben wir damit schon abgehandelt. Für Monoidhomomorphismen brauchen wir noch $f^{-1}(\tilde{e}) = e$, aber das ist ja nach Definition der Umkehrabbildung klar. □

Bemerkung 2.13. Man könnte sich jetzt fragen, warum man nicht gleich definiert, dass ein Isomorphismus ein bijektiver Homomorphismus ist, wenn die zusätzliche Eigenschaft, dass auch die Umkehrabbildung ein Homomorphismus ist, ohnehin von selbst folgt. Das tun manche auch.
Es gibt aber auch Zusatzstrukturen, deren Erhalt von einer Abbildung nicht automatisch auf die Umkehrabbildung übertragen wird. Spätestens dort muss man dann einen Unterschied machen zwischen bijektiven Homomorphismen und Isomorphismen. Deshalb finde ich es sinnvoll, den Unterschied gleich einzuführen, auch wenn er hier noch keine Anwendung findet.

Bemerkung 2.14. Ein Homomorphismus von einem Objekt in sich selbst wird *Endomorphismus* genannt. Ein Endomorphismus, der gleichzeitig Isomorphismus ist, heißt *Automorphismus.*

Bemerkung 2.15. Mathematische Objekte gleicher Art und die dazu passenden Homomorphismen bilden eine sogenannte *Kategorie.*
Beispiele für Kategorien sind also die Kategorie der Magmata mit den Magmahomomorphismen, die Kategorie der Monoide mit den Monoidhomomorphismen oder die Kategorie der Gruppen mit den Gruppenhomomorphismen.

Oft kommt es vor, dass ein Objekt aus einer Kategorie, z. B. eine Gruppe, in einem zweiten Objekt der gleichen Kategorie enthalten ist, und zwar nicht einfach nur als Menge, sondern so, dass die Strukturen, z. B. die Verknüpfung, auch „die gleichen" sind. Wir kennen das von der Gruppe der ganzen Zahlen mit der Addition, die in der Gruppe der reellen Zahlen mit der Addition liegt. Das können wir in die folgende allgemeine Definition fassen:

Definition 2.16. Sei (X, S) eine Objekt aus einer Kategorie von Mengen mit Zusatzstrukturen, wobei X die Menge und S die Strukturen bezeichne.
Ein Objekt (Y, T) derselben Kategorie nennen wir *Unterobjekt (dieser Kategorie)* von X, wenn die zugrundeliegende Menge Y eine Teilmenge von X ist und die natürliche Einbettung $i\colon Y \to X, y \mapsto y$ ein Homomorphismus der Kategorie ist.

Bemerkung 2.17. Ist also $(M, *)$ ein Magma, so heißt also ein Magma $(U, \star)$ *Untermagma* von $(M, *)$, wenn $U \subseteq M$ und für alle Elemente $u, v \in U$ gilt: $u \star v = i(u \star v) = u * v$. Da ein Gruppenhomomorphismus nach Lemma 2.5 dasselbe ist wie ein Magmahomomorphismus zwischen zwei Gruppen, heißt also eine Gruppe $(H, \star, \tilde{e})$ *Untergruppe* einer Gruppe $(G, *, e)$, wenn $H \subset G$ und für alle $h, h' \in H$ gilt: $h \star h' = h * h'$. Da ein Gruppenhomomorphismus das neutrale Element auf das neutrale Element abbildet, ist dann automatisch $\tilde{e} = i(\tilde{e}) = e$. Entsprechend gilt für die Inversen: $h^{\star -1} = i(h^{\star -1}) = h^{* -1}$.

Bemerkung 2.18. Achtung! Es genügt nicht, dass nur die Mengen ineinander enthalten sind. Z.B. ist $(\mathbb{R}^*, \cdot, 1)$ keine Untergruppe von $(\mathbb{R}, +, 0)$, denn für zwei Zahlen $a, b \in \mathbb{R}^*$ gilt im Allgemeinen nicht $a \cdot b = a + b$ (Natürlich sieht man es auch daran, dass die beiden neutralen Elemente nicht übereinstimmen).

Wenn jetzt die Verknüpfung auf der Untergruppe ohnehin einfach die Einschränkung der Verknüpfung der ursprünglichen Gruppe auf die Elemente der Untergruppe ist und das neutrale Element und die Inversen die gleichen wie in der ursprünglichen Gruppe sind, dann genügt es ja offensichtlich, bei einer Untergruppe die Menge festzulegen, der Rest ist dann bereits klar.
Es stellt sich also die Frage, welche Teilmengen von Gruppen Untergruppen sind. Die Antwort gibt das folgende einfache Lemma:

Lemma 2.19. *Sei $(G, *, e)$ eine Gruppe. Sei weiter $H \subseteq G$ und $*_H \colon H \times H \to G, (a, b) \mapsto a * b$ die Einschränkung der Verknüpfung $*$ auf Elemente von H.*
Genau dann ist $(H, *_H, e)$ eine Untergruppe von $(G, *, e)$, wenn gilt:

(a) $H \neq \emptyset$

(b) $\forall a, b \in H\colon \quad a * b \in H$

(c) $\forall a \in H\colon \quad a^{*-1} \in H$

Beweis. „$\Rightarrow$“ Wenn H eine Untergruppe und damit insbesondere eine Gruppe ist, muss es natürlich mindestens das neutrale Element enthalten und $a * b = a *_H b$ muss wieder in H liegen nach Definition einer Verknüpfung. Außerdem hat jedes Element a ein Inverses $a^{*_H -1} \in H$. Wie wir oben festgestellt haben, ist $a^{*_H -1} = a^{*-1}$. Dieses liegt also in H.
„$\Leftarrow$“ Als erstes stellen wir fest, dass $*_H$ tatsächlich eine Verknüpfung auf H ist, da für alle $a, b \in H$ das Ergebnis $a *_H b = a * b$ nach (b) wirklich in H liegt.
Wir müssen die drei Gruppenaxiome prüfen und sicherstellen, dass für alle $a, b \in H$ gilt: $a *_H b = a * b$.
Letzteres gilt nach Definition, also nur noch die Gruppenaxiome:

(G1) Klar: $\forall a, b, c \in H\colon \quad (a *_H b) *_H c = (a * b) * c = a * (b * c) = a *_H (b *_H c)$

(G2) $e \in H$, denn es gilt: $H \neq \emptyset$, also gibt es ein Element $a \in H$. Mit (c) ist $a^{*-1} \in H$, also mit (b) auch $a^{*-1} * a = e$.
e ist neutrales Element von H, denn für alle $a \in H$ gilt: $a *_H e = a * e = a$ und $e *_H a = e * a = a$.

(G3) Jedes Element a hat ein Linksinverses $a^{*_H -1}$: Setze $a^{*_H -1} := a^{*-1}$. Dann gilt $a^{*_H -1} *_H a = a^{*-1} * a = e$.

□

3 Ringe und Körper

Wenn wir uns die reellen Zahlen vor Augen führen, dann erkennen wir, dass es auf ihnen sogar zwei assoziative Verknüpfungen, $+$ und $\cdot$, gibt, die über das Distributivgesetz miteinander verbunden sind. Das ist die Motivation für die folgende Definition:

Definition 3.1. Ein *Ring (mit Eins)* ist ein Tupel $(R, +_R, 0_R, \cdot_R, 1_R)$ bestehend aus einer Menge R mit zwei Verknüpfungen $+_R$ und $\cdot_R$, sodass gilt:

(R1) $(R, +_R, 0_R)$ ist eine abelsche Gruppe.

(R2) $(R, \cdot_R, 1_R)$ ist ein Monoid.

(R3) Es gelten die Distributivgesetze:

$$\forall a, b, c \in R: \quad (a +_R b) \cdot_R c = (a \cdot_R c) +_R (b \cdot_R c), \quad a \cdot_R (b +_R c) = (a \cdot_R b) +_R (a \cdot_R c).$$

Wir bezeichnen die Verknüpfung $+_R$ als die *(Ring-)Addition*, die Verknüpfung $\cdot_R$ als die *(Ring-)Multiplikation.*

Bemerkung 3.2. Wenn dadurch keine Missverständnisse entstehen, schreiben wir auch kürzer $+$, 0, $\cdot$ und 1 für $+_R$, 0_R, $\cdot_R$ und 1_R.

Notation 3.3. Wir benutzen in Ringen die aus der Schule bekannte Regel „Punkt vor Strich“, sodass wir also statt $(a \cdot b) + c$ auch kürzer $a \cdot b + c$ schreiben können.
Weiter lassen wir oft den Punkt der zweiten Verknüpfung ganz weg, sodass wir also statt $a \cdot b$ auch einfach ab schreiben können.

Lemma 3.4. *Sei $(R, +, 0, \cdot, 1)$ ein Ring. Dann gilt:*

(a) $\forall a \in R: \quad 0 \cdot a = 0 = a \cdot 0.$

(b) $\forall a \in R: \quad (-1) \cdot a = -a = a \cdot (-1)$

(c) „Plus mal Minus gibt Minus“: $\quad \forall a, b \in R: \quad a \cdot (-b) = -(a \cdot b)$
„Minus mal Plus gibt Minus“: $\quad \forall a, b \in R: \quad (-a) \cdot b = -(a \cdot b)$
„Minus mal Minus gibt Plus“: $\quad \forall a, b \in R: \quad (-a) \cdot (-b) = a \cdot b$

Beweis. (a) $0 \cdot a = 0 \cdot a + \underbrace{0 \cdot a - 0 \cdot a}_{=0} = (0 + 0) \cdot a - 0 \cdot a = 0 \cdot a - 0 \cdot a = 0.$

Oder kürzer mit der Kürzungsregel in der Gruppe $(R, +, 0)$:

$$0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a.$$

Die andere Aussage beweist man analog.

(b) Es gilt: $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1 + 1) \cdot a = 0$, also ist $(-1) \cdot a$ das eindeutige Inverse $-a$ von a.
Auch hier beweist man die zweite Aussage analog.

(c) Wir beweisen die dritte Aussage, die anderen beiden zeigt man analog, nur einfacher.
$(-a) \cdot (-b) = (a \cdot (-1)) \cdot ((-1) \cdot b) = a \cdot ((-1) \cdot (-1)) \cdot b = a \cdot (-(-1)) \cdot b = a \cdot 1 \cdot b = a \cdot b$

□

Bemerkung 3.5. Nach Lemma 1.16 ist die Menge R^* (oder $R^\times$) aller bezüglich der Multiplikation invertierbaren Elemente, zusammen mit der Einschränkung der Multiplikation auf diese Menge und der 1 eine Gruppe. Man nennt sie die *Einheitengruppe* des Rings und ihre Elemente die *Einheiten* des Rings.
Ist in dem Ring $1 \neq 0$, so ist $0 \notin R^*$, denn $a \cdot 0$ ist immer 0 und nie 1, also hat 0 kein multiplikatives Inverses.
Ist in dem Ring $1 = 0$, so besteht der Ring nur aus diesem einen Element, denn für jedes Element $a \in R$ gilt: $a = 1 \cdot a = 0 \cdot a = 0$.

Besteht ein Ring also aus mehr als einem Element, so kann niemals $(R, \cdot, 1)$ eine Gruppe sein. Das Beste, was möglich ist, ist dass 0 das einzige Element ohne multiplikatives Inverses ist, sodass also $(R\backslash\{0\}, \cdot, 1)$ eine Gruppe ist. Das ist die Motivation der folgenden Definition:

Definition 3.6. (a) Ein Ring $(R, +, 0, \cdot, 1)$, in dem $0 \neq 1$ und $R^* = R \setminus \{0\}$, heißt *Schiefkörper*.

(b) Ein *Körper* ist ein Schiefkörper, bei dem auch die Multiplikation kommutativ ist.

Natürlich gibt es auch Ringhomomorphismen und Körperhomomorphismen. Wie man sich leicht denken kann, sind diese durch die folgenden Eigenschaften charakterisiert:

Definition 3.7. Seien $(R, +_R, 0_R, \cdot_R, 1_R)$ und $(S, +_S, 0_S, \cdot_S, 1_S)$.
Ein *Ringhomomorphismus* von $(R, +_R, 0_R, \cdot_R, 1_R)$ nach $(S, +_S, 0_S, \cdot_S, 1_S)$ ist eine Abbildung $\phi\colon R \to S$, für die gilt:

(a) ϕ ist ein Gruppenhomomorphismus von $(R, +_R, 0_R)$ nach $(S, +_S, 0_S)$.

(b) ϕ ist ein Monoidhomomorphismus von $(R, \cdot_R, 1_R)$ nach $(S, \cdot_S, 1_S)$.

Bemerkung 3.8. Wegen Lemma 2.5 sind das die Eigenschaften, die man für einen Ringhomomorphismus prüfen muss:

(a) $\forall a, b \in R\colon \quad \phi(a +_R b) = \phi(a) +_S \phi(b) \quad \wedge \quad \phi(a \cdot_R b) = \phi(a) \cdot_S \phi(b)$

(b) $\phi_R(1_R) = 1_S$

Bemerkung 3.9. Da ein Körper gegenüber einem Ring keine Zusatzstruktur besitzt, ist jeder Ringhomomorphismus zwischen zwei Körpern ein Körperhomomorphismus.

Notation 3.10. Auch bei Ringen und Körpern schreiben wir meist nur die Menge hin und denken uns die Verknüpfungen und die beiden neutralen Elemente dazu.

Beispiel 3.11. Die wichtigsten Körper sind $\mathbb{Q}$, der Körper der rationalen Zahlen, $\mathbb{R}$, der Körper der reellen Zahlen, und $\mathbb{C}$, der Körper der komplexen Zahlen.

4 Vektorräume

Betrachten wir das homogene reelle lineare Gleichungssystem

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

so können wir, bevor wir die Lösungen berechnet haben, bereits zwei Dinge feststellen:

1. Wenn (x_1, x_2, x_3) eine Lösung des linearen Gleichungssystems ist und (x'_1, x'_2, x'_3) eine zweite, dann ist auch die „Summe“ der Lösungen, zu verstehen als das Tupel $(x_1 + x'_1, x_2 + x'_2, x_3 + x'_3)$, eine Lösung, denn $2(x_1 + x'_1) + 3(x_2 + x'_2) - (x_3 + x'_3) = 2x_1 + 3x_2 - x_3 + 2x'_1 + 3x'_2 - x'_3 = 0 + 0 = 0$ und entsprechend $(x_1 + x'_1) - 2(x_2 + x'_2) + (x_3 + x'_3) = 0$.
2. Mit (x_1, x_2, x_3) ist auch das skalierte Tupel $(\lambda x_1, \lambda x_2, \lambda x_3)$ für jedes beliebige $\lambda \in \mathbb{R}$ eine Lösung, denn $2(\lambda x_1) + 3(\lambda x_2) - (\lambda x_3) = \lambda(2x_1 + 3x_2 - x_3) = \lambda \cdot 0 = 0$ und entsprechend $(\lambda x_1) - 2(\lambda x_2) + (\lambda x_3) = 0$.

An keiner Stelle haben wir dabei die genauen Koeffizienten des linearen Gleichungssystems, die Anzahl der Variablen oder der Gleichungen benutzt. Das Einzige, was wichtig war, war die Homogenität, also dass auf der rechten Seite Nullen standen.
Ganz ähnlich sieht es aus, wenn wir uns die folgende reelle Differentialgleichung ansehen:

$$f(x) + 2f'(x) - 2f''(x) = 0.$$

Auch hier können wir, ohne die Lösungen zu kennen, bereits feststellen:

1. Wenn f und g zwei Lösungen der Differentialgleichung sind, dann auch ihre Summe, denn $(f+g)(x) + 2(f+g)'(x) - 2(f+g)''(x) \overset{\text{Summenregel}}{=} f(x) + g(x) + 2(f'(x) + g'(x)) - 2(f''(x) + g''(x)) = f(x) + 2f'(x) - 2f''(x) + g(x) + 2g'(x) - 2g''(x) = 0 + 0 = 0$.
2. Mit f ist auch die mit einem konstanten Faktor multiplizierte Funktion λf für jedes $\lambda \in \mathbb{R}$ eine Lösung, denn $(\lambda f)(x) + 2(\lambda f)'(x) - 2(\lambda f)''(x) \overset{\text{Faktorregel}}{=} \lambda f(x) + 2(\lambda f'(x)) - 2(\lambda f''(x)) = \lambda(f(x) + 2f'(x) - 2f''(x)) = \lambda \cdot 0 = 0$.

Wichtig zu verstehen ist aber, dass in beiden Fällen zwar das Produkt einer Lösung mit einem konstanten Faktor wieder eine Lösung ergibt, aber das das Produkt zweier Lösungen im allgemeinen nicht. Im Fall des linearen Gleichungssystems stößt man auf das Problem, dass man eine Summe von Produkten im allgemeinen nicht in Faktoren zerlegen kann, im zweiten davon abgesehen bereits auf das Hindernis, dass die Ableitung eines Produkts mit der Produktregel recht komplizierte Terme liefert.
Diese Ähnlichkeit in der Struktur der Lösungsmengen ist die Motivation für die folgende Definition, die das Zentrum der linearen Algebra bildet:

Definition 4.1. Sei $(K, +_K, 0_K, \cdot_K, 1_K)$ ein Körper.
Ein *K-Vektorraum* ist ein Tupel $(V, +_V, 0_V, \cdot_{K,V})$ bestehend aus einer abelschen Gruppe $(V, +_V, 0_V)$ und einer Abbildung $\cdot_{K,V} \colon K \times V \to V$, sodass für alle $\lambda, \mu \in K, v, w \in V$ gilt:

(V1) $\lambda \cdot_{K,V} (v +_V w) = \lambda \cdot_{K,V} v +_V \lambda \cdot_{K,V} w$

(V2) $(\lambda +_K \mu) \cdot_{K,V} v = \lambda \cdot_{K,V} v +_V \mu \cdot_{K,V} v$

(V3) $\lambda \cdot_{K,V} (\mu \cdot_{K,V} v) = (\lambda \cdot_K \mu) \cdot_{K,V} v$

(V4) $1_K \cdot_{K,V} v = v$

Die Gruppenverknüpfung $+_V$ heißt dabei *Vektoraddition*, die Abbildung $\cdot_{K,V}$ *skalare Multiplikation* oder *Skalarmultiplikation*. Die Elemente von V heißen *Vektoren*, die von K *Skalare*. Das neutrale Element 0_V wird als *Nullvektor* bezeichnet.
Ist $K = \mathbb{R}$ der Körper der reellen Zahlen, so heißen $\mathbb{R}$-Vektorräume auch *reelle Vektorräume*. Ist $K = \mathbb{C}$ der Körper der komplexen Zahlen, so heißen $\mathbb{C}$-Vektorräume auch *komplexe Vektorräume*.

Notation 4.2. Auch hier werden wir im Allgemeinen kürzer $+$ und 0 für $+_V$ und 0_V schreiben und $\cdot_{K,V}$ ganz weglassen. Außerdem bezeichnen wir das Inverse eines Vektors v in der Gruppe $(V, +_V, 0_V)$ wie gewohnt mit $-v$. Es ist deshalb wichtig, bei jeder Aussage zu überlegen, an welcher Stelle die Addition und Multiplikation des Körpers und wo die Vektoraddition und die Skalarmultiplikation gemeint sind.

Bemerkung 4.3. Achtung! Wir können zwei Vektoren addieren und einen Vektor mit einem Skalar multiplizieren. Zwei Vektoren multiplizieren kann man im Allgemeinen nicht.

Beispiel 4.4. (a) Sei $(K, +_K, 0_K, \cdot_K, 1_K)$ ein Körper. Dann ist $(K, +_K, 0_K, \cdot_K)$ ein K-Vektorraum. Die Vektorraumaxiome $(V1)$ und $(V2)$ sind einfach die Distributivgesetze, $(V3)$ ist das Assoziativgesetz der Multiplikation und $(V4)$ die Eigenschaft der 1_K, neutrales Element der Multiplikation zu sein.

(b) Sei $K^n = \underbrace{K \times \cdots \times K}_{n} = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in K\}$. Wir definieren für zwei Elemente $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ aus K^n ihre Summe

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 +_K y_1, \ldots, x_n +_K y_n)$$

und für $\lambda \in K$

$$\lambda(x_1, \ldots, x_n) := (\lambda \cdot_K x_1, \ldots, \lambda \cdot_K x_n).$$

Mit diesen beiden Abbildungen und dem Nullvektor $(0, \ldots, 0)$ ist K^n ein K-Vektorraum:
(V1): Seien $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in K^n$ und $\lambda \in K$ beliebig. Dann gilt:

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

Die Vektorraumaxiome (V2), (V3) und (V4) zeigt man auf ähnliche Weise.

(c) Eine einelementige Menge $\{0_V\}$ mit der trivialen Vektoraddition $0_V +_V 0_V = 0_V$, dem Nullvektor 0_V und der trivialen Skalarmultiplikation $\lambda \cdot_{K,V} 0_V = 0_V$ für alle $\lambda \in K$ ist ein K-Vektorraum. Er wird als *Nullvektorraum* bezeichnet.

(d) Ist $(V, +_V, 0_V, \cdot_{K,V})$ ein K-Vektorraum und L ein Unterkörper von K, dann ist $(V, +_V, 0_V, \cdot_{L,V})$ mit der auf $L \times V$ eingeschränkten Skalarmultiplikation auch ein L-Vektorraum.

(e) Ist $(V, +_V, 0_V, \cdot_{K,V})$ ein K-Vektorraum und X ein beliebige Menge, dann ist die Menge $\{f : X \to V\}$ aller Abbildungen von X nach V mit der Vektoraddition $f + g : x \mapsto f(x) +_V g(x)$ und der Skalarmultiplikation $\lambda f : x \mapsto \lambda \cdot_{K,V} f(x)$ und dem Nullvektor $0 : x \mapsto 0_V$ ein K-Vektorraum:
(V1): Seien f und g Abbildungen von X nach V und $\lambda \in K$. Wir müssen zeigen, dass die Abbildungen $\lambda(f + g)$ und $\lambda f + \lambda g$ gleich sind, also auf jedem Element von X dasselbe bewirken. Sei also $x \in X$ beliebig. Dann haben wir

$$\begin{aligned}
(\lambda(f + g))(x) &= \lambda \cdot_{K,V} ((f + g)(x)) \\
&= \lambda \cdot_{K,V} (f(x) +_V g(x)) \\
&= \lambda \cdot_{K,V} f(x) +_V \lambda \cdot_{K,V} g(x) \\
&= (\lambda f)(x) +_V (\lambda g)(x) \\
&= (\lambda f + \lambda g)(x)
\end{aligned}$$

Auch hier zeigt man die Vektorraumaxiome (V2), (V3) und (V4) auf ähnliche Weise.

Lemma 4.5. *Sei* $(K, +_K, 0_K, \cdot_K, 1_K)$ *ein Körper und* $(V, +_V, 0_V, \cdot_{K,V})$ *ein* K*-Vektorraum. Dann gilt:*

(a) $\forall v \in V\colon \quad 0_K \cdot_{K,V} v = 0_V$

(b) $\forall \lambda \in K\colon \quad \lambda \cdot_{K,V} 0_V = 0_V$

(c) $\forall \lambda \in K, v \in V\colon \quad \lambda \cdot_{K,V} v = 0_V \quad \Rightarrow \quad (\lambda = 0_K \ \vee \ v = 0_V)$

(d) $\forall v \in V\colon (-1_K) \cdot_{K,V} v = -v$

Beweis. (a) Es gilt $0_K \cdot_{K,V} v = (0_K +_K 0_K) \cdot_{K,V} v = 0_K \cdot_{K,V} v +_V 0_K \cdot_{K,V} v$. Mit der Kürzungsregel in der Gruppe $(V, +_V, 0_V)$ folgt damit die Aussage.

(b) Analog

(c) Sei $\lambda \cdot_{K,V} v = 0$ und $\lambda \neq 0$. Dann hat λ ein multiplikatives Inverses λ^{-1} und es folgt

$$v \overset{\text{(V4)}}{=} 1_K \cdot_{K,V} v = (\lambda^{-1} \cdot_K \lambda) \cdot_{K,V} v \overset{\text{(V3)}}{=} \lambda^{-1} \cdot_{K,V} (\lambda \cdot_{K,V} v) = \lambda^{-1} \cdot_{K,V} 0_V \overset{\text{(b)}}{=} 0_V.$$

(d) Es gilt: $v + (-1_K) \cdot_{K,V} v \overset{\text{(V4)}}{=} 1_K \cdot_{K,V} v + (-1_K) \cdot_{K,V} v \overset{\text{(V2)}}{=} (1_K - 1_K) \cdot_{K,V} v = 0_K \cdot_{K,V} v \overset{\text{(a)}}{=} 0_V$.
Damit ist $(-1_K) \cdot_{K,V} v$ das eindeutige Inverse $-v$ von v in der Gruppe $(V, +_V, 0_V)$.

□

Jetzt, da wir Vektorräume definiert haben, ist es kein Kunststück, darauf zu kommen, was ein Vektorraumhomomorphismus ist.

Definition 4.6. Sei $(K, +_K, 0_K, \cdot_K, 1_K)$ ein Körper und seien $(V, +_V, 0_V, \cdot_{K,V})$ und $(W, +_W, 0_W, \cdot_{K,W})$ K-Vektorräume.
Ein K*-Vektorraumhomomorphismus* oder auch eine K*-lineare Abbildung* von $(V, +_V, 0_V, \cdot_{K,V})$ nach $(W, +_W, 0_W, \cdot_{K,W})$ ist eine Abbildung $\phi\colon V \to W$, für die gilt:

1. ϕ ist ein Gruppenhomomorphismus von $(V, +_V, 0_V)$ nach $(W, +_W, 0_W)$.
2. $\forall \lambda \in K, v \in V\colon \quad \phi(\lambda \cdot_{K,V} v) = \lambda \cdot_{K,W} \phi(v)$.

Bemerkung 4.7. Da jeder Magmahomomorphismus zwischen zwei Gruppen bereits ein Gruppenhomomorphismus ist, reichen die folgenden beiden Eigenschaften aus, um K-lineare Abbildungen zu charakterisieren:

(L1) $\forall v, w \in V\colon \quad \phi(v +_V w) = \phi(v) +_W \phi(w)$ („Additivität")

(L2) $\forall \lambda \in K, v \in V\colon \quad \phi(\lambda \cdot_{K,V} v) = \lambda \cdot_{K,W} \phi(v)$ („Homogenität")

Bemerkung 4.8. Achtung! Der Körper der Skalare muss bei beiden Vektorräumen derselbe sein.

Wie wir uns denken können, ist ein K*-Vektorraumisomorphismus* ein bijektiver K-Vektorraumhomomorphismus, dessen Umkehrabbildung wieder ein K-Vektorraumhomomorphismus ist. Unsere erste Frage sollte wieder sein: Ist jeder bijektive K-Vektorraumhomomorphismus bereits ein K-Vektorraumisomorphismus? Die Antwort ist „ja":

Lemma 4.9. *Sei K ein Körper und seien $(V, +_V, 0_V, \cdot_{K,V})$ und $(W, +_W, 0_W, \cdot_{K,W})$ zwei Vektorräume.*
Wenn ϕ ein bijektiver K-Vektorraumhomomorphismus von $(V, +_V, 0_V, \cdot_{K,V})$ nach $(W, +_W, 0_W, \cdot_{K,W})$ ist, so ist seine Umkehrabbildung ϕ^{-1} auch ein K-Vektorraumhomomorphismus.

Beweis. Die Eigenschaft (L1) müssen wir nicht mehr zeigen, da ϕ auch ein bijektiver Gruppenhomomorphismus ist.
Die Eigenschaft (L2) sehen wir so:

$$\begin{aligned}
\phi^{-1}(\lambda \cdot_{W,K} w) &\overset{\text{Definition } \phi^{-1}}{=} \phi^{-1}(\lambda \cdot_{W,K} \phi(\phi^{-1}(w))) \\
&\overset{\text{Homogenität von } \phi}{=} \phi^{-1}(\phi(\lambda \cdot_{V,K} \phi^{-1}(w))) \\
&\overset{\text{Definition } \phi^{-1}}{=} \lambda \cdot_{V,K} \phi^{-1}(w).
\end{aligned}$$

□

Lemma 4.10. *Sei K ein Körper und seien U, V und W K-Vektorräume.*
Wenn $f\colon U \to V$ und $g\colon V \to W$ K-Vektorraumhomomorphismen sind, dann ist auch die Komposition $g \circ f\colon U \to W$ ein K-Vektorraumhomomorphismus.

Beweis. (L1) Seien $x, y \in U$ beliebig. Dann gilt:

$$\begin{aligned}
(g \circ f)(x + y) &= g(f(x + y)) \\
&\overset{f \text{ linear}}{=} g(f(x) + f(y)) \\
&\overset{g \text{ linear}}{=} g(f(x) + g(f(y)) \\
&= (g \circ f)(x) + (g \circ f)(y)
\end{aligned}$$

(L2) Seien $x \in U$ und $\lambda \in K$ beliebig. Dann gilt:

$$\begin{aligned}
(g \circ f)(\lambda x) &= g(f(\lambda x)) \\
&\overset{f \text{ linear}}{=} g(\lambda f(x)) \\
&\overset{g \text{ linear}}{=} \lambda g(f(x)) \\
&= \lambda (g \circ f)(x)
\end{aligned}$$

□

Wenn wir wissen, was Vektorraumhomomorphismen sind, dann wissen wir auch, was ein Untervektorraum ist.

Definition 4.11. Sei K ein Körper und $(V, +_V, 0_V, \cdot_{K,V})$ ein K-Vektorraum.
Ein *Untervektorraum* von V ist ein K-Vektorraum $(W, +_W, 0_W, \cdot_{K,W})$, sodass $W \subseteq V$ und die Inklusionsabbildung $i\colon W \to V$ ein Vektorraumhomomorphismus ist.

Bemerkung 4.12. Wie bei Untergruppen bedeutet diese etwas abstrakte Definition, dass $+_W$ und $\cdot_{K,W}$ gerade die Einschränkungen der Abbildungen $+_V$ und $\cdot_{K,V}$ auf $W \times W$ bzw. $K \times W$ sind. Außerdem folgt sofort, dass $0_W = 0_V$.

Bemerkung 4.13. Ebenfalls wie bei Untergruppen gibt es eine einfach zu prüfende Charakterisierung von Untervektorräumen:
Eine Teilmenge $W \subseteq V$ wird zusammen mit der Einschränkung der Vektoraddition und der Skalarmultiplikation und dem Nullvektor genau dann ein Untervektorraum, wenn die folgenden drei Eigenschaften erfüllt sind:

(U1) $W \neq \emptyset$

(U2) $\forall x, y \in W\colon \quad x + y \in W$

(U3) $\forall x \in W, \lambda \in K\colon \quad \lambda x \in W$

Beispiel 4.14. (a) Im Vektorraum $\mathbb{R}^3$ ist die Menge der Lösungen (x_1, x_2, x_3) des linearen Gleichungssystems

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

ein Untervektorraum.
Die beiden Eigenschaften (U2) und (U3) haben wir bereits nachgeprüft. (U1) ist erfüllt, da $(0, 0, 0)$ offensichtlich eine Lösung ist.

(b) Im Vektorraum aller Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ ist die Menge der Lösungen der Differentialgleichung

$$f(x) + 2f'(x) - 2f''(x) = 0$$

ein Untervektorraum aus denselben Gründen.

(c) In jedem Vektorraum V sind die Menge $\{0_V\}$ und der ganze Vektorraum V Untervektorräume. Man nennt sie die *trivialen Unterräume*.

(d) Seien V und W zwei K-Vektorräume. Dann ist die Menge

$$\mathrm{Hom}_K(V, W) := \{f\colon V \to W \mid f \ K\text{-linear}\}$$

der Vektorraumhomomorphismen von V nach W ein Untervektorraum des Vektorraums aller Abbildungen von V nach W:

(U1) Die Abbildung $V \to W, v \mapsto 0_W$ ist offensichtlich linear.

(U2) Seien $f, g \in \mathrm{Hom}_K(V, W)$. Dann gilt für $f + g$:

$$\begin{aligned} (f+g)(x+y) &= f(x+y) + g(x+y) \\ &\overset{f \text{ und } g \text{ linear}}{=} f(x) + f(y) + g(x) + g(y) \\ &= f(x) + g(x) + f(y) + g(y) \\ &= (f+g)(x) + (f+g)(y), \end{aligned}$$

$$(f+g)(\lambda x) = f(\lambda x) + g(\lambda x) = \lambda f(x) + \lambda g(x) = \lambda(f(x) + g(x)) = \lambda(f+g)(x).$$

Also ist auch $f + g \in \mathrm{Hom}_K(V, W)$.

(U3) Sei $f \in \mathrm{Hom}_K(V, W)$, $\lambda \in K$. Dann gilt für λf:

$$\begin{aligned} (\lambda f)(x+y) &= \lambda f(x+y) \\ &\overset{f \text{ linear}}{=} \lambda(f(x) + f(y)) \\ &= \lambda f(x) + \lambda f(y) \\ &= (\lambda f)(x) + (\lambda f)(y), \end{aligned}$$

$$(\lambda f)(\mu x) = \lambda f(\mu x) \overset{f \text{ linear}}{=} \lambda\mu f(x) = \mu\lambda f(x) = \mu(\lambda f)(x).$$

Also ist auch $\lambda f \in \mathrm{Hom}_K(V, W)$.

Definition 4.15. Sei K ein Körper und V ein K-Vektorraum.
Eine *Linearkombination* von endlich vielen Vektoren $v_1, \ldots, v_r$ ist ein Ausdruck der Form $\lambda_1 v_1 + \ldots + \lambda_r v_r$ für $\lambda_1, \ldots, \lambda_r \in K$.
Ist $E \subseteq V$ eine Teilmenge, so nennen wir

$$\mathrm{span}_K(E) := \{v \in V \mid \exists e_1, \ldots, e_r \in E, \lambda_1, \ldots, \lambda_r \in K \colon v = \lambda_1 e_1 + \ldots + \lambda_r v_r\}$$

das *Erzeugnis* von E über K.

Notation 4.16. Eine andere Schreibweise für das Erzeugnis ist $\langle E \rangle$.

Lemma 4.17. *Sei K ein Körper, V ein K-Vektorraum und $E \subseteq V$ beliebig. Dann gilt:*

(a) Das Erzeugnis $\mathrm{span}_K(E)$ *von E ist ein Untervektorraum, der E enthält.*

(b) Ist W ein beliebiger Untervektorraum von V, der E enthält, dann gilt $\mathrm{span}_K(E) \subseteq W$.

Zusammengefasst heißt das: $\mathrm{span}_K(E)$ *ist der kleinste Untervektorraum, der E enthält.*

Beweis. (a) Dass $E \in \mathrm{span}_K(E)$, ist klar, denn jedes Element $e \in E$ lässt sich als Linearkombination $1e$ aus Elementen (in diesem Fall einem einzigen) aus E schreiben.
Wir prüfen die drei charakterisierenden Eigenschaften für Untervektorräume aus Bemerkung 4.13:

(U1): Wenn $E \neq \emptyset$, dann ist es klar. Aber auch wenn $E = \emptyset$ ist, enthält $\mathrm{span}_K(E)$ noch den Nullvektor, wenn wir uns an die Konvention $\sum_{e_i \in \emptyset} e_i = 0$ in beliebigen abelschen Gruppen erinnern.

(U2): Seien $x, y \in \mathrm{span}_K(E)$ beliebig. Nach Definition des Erzeugnisses gibt es dann Vektoren $e_1, \ldots, e_r$ und $e_{r+1}, \ldots, e_s$ aus E (die auch nicht alle verschieden sein müssen) und Skalare $\lambda_1, \ldots, \lambda_r$ sowie $\lambda_{r+1}, \ldots, \lambda_s$, sodass $x = \sum_{i=1}^{r} \lambda_i e_i$ und $y = \sum_{i=r+1}^{s} \lambda_i e_i$. Dann ist aber $x + y = \sum_{i=1}^{s} \lambda_i e_i$ und damit in $\mathrm{span}_K(E)$.

(U3): Wieder gibt es Vektoren $e_1, \ldots, e_r \in E$ und Skalare $\lambda_1, \ldots, \lambda_r \in K$, sodass $x = \sum_{i=1}^{r} \lambda_i e_i$. Dann ist $\lambda x = \lambda(\sum_{i=1}^{r} \lambda_i e_i) = \sum_{i=1}^{r} \lambda(\lambda_i e_i) = \sum_{i=1}^{r} (\lambda\lambda_i) e_i$ und damit in $\mathrm{span}_K(E)$.

(b) Sei W ein Untervektorraum, der E enthält. Dann enthält W nach Eigenschaft $(U3)$ auch alle Elemente $\lambda_i e_i$ mit $\lambda_i \in K$ und $e_i \in E$ und nach Eigenschaft $(U2)$ induktiv auch alle Summen $\lambda_1 e_1 + \ldots + \lambda_r e_r$. Damit liegen aber alle Elemente aus $\mathrm{span}_K(E)$ in W.

$\square$

Definition 4.18. Sei K ein Körper, V ein K-Vektorraum und W ein Untervektorraum von V.
Ein *Erzeugendensystem* von W ist eine Teilmenge $E \subseteq W$ mit $\mathrm{span}_K(E) = W$. Hat W ein Erzeugendensystem, das nur endlich viele Elemente hat, so nennen wir W *endlich erzeugt.*

5 Basen und Dimension

Es ist leicht, für einen beliebigen Vektorraum ein Erzeugendensystem anzugeben, nämlich einfach den gesamten Vektorraum. Es ist aber klar, dass das „ineffizient“ ist in dem Sinne, als wir Vektoren hätten weglassen können, sodass die restlichen immer noch ein Erzeugendensystem bilden.
Wir hätten gerne eine Möglichkeit herauszufinden, ob ein gegebenes Erzeugendensystem minimal ist. Das wird uns auf den Begriff der Basis eines Vektorraums führen.

Definition 5.1. Seien I und A beliebige Mengen. Wir nennen eine Abbildung $I \to A, i \mapsto a_i$ auch *(durch I indizierte) Familie (in A)* und notieren sie als $(a_i)_{i\in I}$. I bezeichnen wir als *Indexmenge.*
Ist $I = \emptyset$, so heißt jede Familie $(a_i)_{i\in\emptyset}$ *leere Familie.*

Definition 5.2. Sei K ein Körper und V ein K-Vektorraum.
Eine Familie $(v_i)_{i\in I}$ von Vektoren aus V heißt *linear unabhängig*, wenn für jede endliche Teilmenge $\{i_1, \ldots, i_r\} \subseteq I$ paarweise verschiedener Indizes und beliebige $\lambda_1, \ldots, \lambda_r \in K$ gilt:
Wenn $\lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r} = 0$, dann gilt bereits $\lambda_1 = \ldots = \lambda_r = 0$.
Eine Familie $(v_i)_{i\in I}$ von Vektoren aus V, die nicht linear unabhängig ist, heißt *linear abhängig.*

Bemerkung 5.3. Achtung! Nur weil die Indizes $i_1, \ldots, i_r$ alle verschieden sein müssen, müssen das die Vektoren $v_{i_1}, \ldots, v_{i_r}$ noch lange nicht. Wenn aber zwei Vektoren v_i und v_j einer Familie gleich sind, dann ist die Familie automatisch linear abhängig, denn dann kann man den Nullvektor darstellen als $1 \cdot v_i + (-1) \cdot v_j$.

Lemma 5.4. *Sei K ein Körper, V ein K-Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren. Dann gilt:*

(a) Ist $I = \emptyset$, so ist die Familie $(v_i)_{i\in I}$ linear unabhängig.

(b) Ist $I = \{i\}$ eine einelementige Indexmenge, so ist die Familie $(v_i)_{i\in I}$ genau dann unabhängig, wenn $v_i \neq 0$.

(c) Falls $(v_i)_{i\in I}$ linear unabhängig und $J \subseteq I$, dann ist auch $(v_i)_{i\in J}$ linear unabhängig.

Beweis. (a) Wenn $I = \emptyset$, müssen wir für null Skalare etwas zeigen, also nichts.

(b) „$\Rightarrow$“: Wenn $v_i \neq 0$, dann folgt aus $\lambda_i v_i = 0$ bereits $\lambda_i = 0$ nach Lemma 4.5. Damit ist $(v_i)_{i\in I}$ linear unabhängig.
„$\Leftarrow$“: Wenn $v_i = 0$, dann ist $1 \cdot v_1 = 0$ und $1 \neq 0$, also muss die Familie (0) linear abhängig sein.

(c) Jede Linearkombination $\lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r}$ mit $i_k \in J$ ist auch eine Linearkombination mit $i_k \in I$. Da die Familie $(v_i)_{i\in I}$ linear unabhängig ist, müssen alle Skalare λ_i Null sein.

□

Lemma 5.5 (Alternative Charakterisierung von linear abhängigen Familien)**.** *Sei K ein Körper, V ein K-Vektorraum.*
Eine Familie $(v_i)_{i\in I}$ von Vektoren aus V ist genau dann linear abhängig, sich mindestens einer der Vektoren als Linearkombination der anderen darstellen lässt, wenn es also einen Index $i_0 \in I$ gibt und endlich viele andere Indizes $i_1, \ldots, i_r \in I$ sowie Skalare $\mu_1, \ldots, \mu_r$, sodass $v_{i_0} = \mu_1 v_{i_1} + \ldots + \mu_r v_{i_r}$.

Beweis. „$\Leftarrow$“: Wenn die Familie linear abhängig ist, dann gibt es also eine Darstellung $0 = \lambda_0 v_{i_0} + \ldots + \lambda_r v_{i_r}$, bei dem mindestens ein $\lambda_k \neq 0$, oBda $\lambda_0 \neq 0$. Dann gilt

$$\lambda_0 v_{i_0} = -\lambda_1 v_{i_1} - \ldots - \lambda_r v_{i_r}$$

und wegen $\lambda_0 \neq 0$ folgt

$$v_{i_0} = (-\lambda_0^{-1}\lambda_1)v_1 + \ldots + (-\lambda_0^{-1}\lambda_1)v_r.$$

Setze also $\mu_k := -\lambda_0^{-1}\lambda_k$.
„$\Rightarrow$“: Falls andersherum $v_{i_0} = \mu_{i_1} v_1 + \ldots + \mu_{i_r} v_r$, dann setze $\lambda_0 := 1$ und $\lambda_k := -\mu_k$ für $k \in \{1, \ldots, r\}$, und wir haben $\lambda_0 v_{i_0} + \lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r} = 0$, aber $\lambda_0 \neq 0$. □

Definition 5.6. Sei K ein Körper, V ein K-Vektorraum und W ein Untervektorraum. Wir nennen eine Familie $(v_i)_{i \in I}$ *Erzeugendensystem* von W, wenn die Menge $\{v_i \mid i \in I\}$ ein Erzeugendensystem von W ist.

Einen Zusammenhang zwischen Erzeugendensystemen und linear unabhängigen Familien erkennt man in der folgenden Aussage:

Satz 5.7. *Sei K ein Körper, V ein K-Vektorraum.*
Für eine Familie $(v_i)_{i \in I}$ von Vektoren $v_i \in V$ sind äquivalent:

(1) $(v_i)_{i \in I}$ ist ein unverkürzbares Erzeugendensystem von V, d.h. für jedes $i_0 \in I$ ist $(v_i)_{i \in I \setminus \{i_0\}}$ kein Erzeugendensystem von V.

(2) $(v_i)_{i \in I}$ ist ein linear unabhängiges Erzeugendensystem von V.

(3) $(v_i)_{i \in I}$ ist eine unverlängerbare linear unabhängige Familie, d.h. für jedes $w \in V$ ist $(w, (v_i)_{i \in I})$ linear abhängig.

Beweis. „(1) $\Rightarrow$ (2)“: Beweis durch Widerspruch:
Angenommen, $(v_i)_{i \in I}$ sei linear abhängig. Dann gibt es nach Lemma 5.5 einen Index $i_0 \in I$ und andere Indizes $i_1, \ldots, i_r \in I$ sowie Skalare $\mu_1, \ldots, \mu_r \in K$ sodass $v_{i_0} = \sum_{k=1}^{r} \mu_k v_{i_k}$, also $v_{i_0} \in \mathrm{span}_K(\{v_i \mid i \in I \setminus \{i_0\}\})$.
Dann ist also $E := \{v_i \mid i \in I\} \subseteq W := \mathrm{span}_K(\{v_i \mid i \in I \setminus \{i_0\})$ und damit nach Lemma 4.17 auch $\mathrm{span}_K(E) \subseteq W$. Da aber trivialerweise andersherum $\mathrm{span}_K(\{v_i \mid i \in I \setminus \{i_0\}) \subseteq \{v_i \mid i \in I\}$, gilt Gleichheit der beiden Erzeugnisse. Damit war aber $(v_i)_{i \in I}$ kein minimales Erzeugendensystem im Widerspruch zur Voraussetzung.
„(2) $\Rightarrow$ (1)“: Beweis durch Widerspruch:
Angenommen, wir könnten einen Vektor v_{i_0} der Familie weglassen und die Familie $(v_i)_{i \in I \setminus \{i_0\}}$ wäre immer noch ein Erzeugendensystem.
Dann könnten wir insbesondere den Vektor v_{i_0} als Linearkombination von Vektoren aus $(v_i)_{i \in I \setminus \{i_0\}}$ darstellen. Nach Lemma 5.5 wäre dann aber die Familie $(v_i)_{i \in I}$ linear abhängig im Widerspruch zur Voraussetzung.
„(2) $\Rightarrow$ (3)“: Dass $(v_i)_{i \in I}$ linear unabhängig ist, gilt nach Voraussetzung. Wir müssen also nur noch die Maximalität prüfen.
Sei dazu $w \in V$ beliebig. Da $(v_i)_{i \in I}$ ein Erzeugendensystem ist, gibt es eine Darstellung von w als Linearkombination endlich vieler v_i. Nach Lemma 5.5 ist dann die Familie $(w, (v_i)_{i \in I})$ linear abhängig.
„(3) $\Rightarrow$ (2)“: Wir müssen zeigen, dass sich jeder Vektor w als Linearkombination von Vektoren v_i darstellen lässt.
Wir wissen, dass $(w, (v_i)_{i \in I})$ linear abhängig ist, es gibt also eine Darstellung $0 = \lambda w +$

$\lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r}$ des Nullvektors als Linearkombination, sodass nicht alle Koeffizienten 0 sind. Tatsächlich muss $\lambda \neq 0$ sein, denn sonst hätten wir ja eine Darstellung $0 = \lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r}$, sodass nicht alle Koeffizienten 0 sind, im Widerspruch zur linearen Unabhängigkeit der Familie $(v_i)_{i \in I}$.
Wenn aber $\lambda \neq 0$, gehen wir vor wie im Beweis für Lemma 5.5:

$$w = \lambda^{-1}(-\lambda_1 v_1 - \ldots - \lambda_r v_r) = (-\lambda^{-1}\lambda_1)v_1 + \ldots + (-\lambda^{-1}\lambda_r)v_r.$$

Somit haben wir w als Linearkombination dargestellt. □

Definition 5.8. Sei K ein Körper, V ein K-Vektorraum.
Eine Familie $(v_i)_{i \in I}$, die die äquivalenten Eigenschaften des Satzes erfüllt, heißt *Basis* des Vektorraums V.

Wofür sind Basen gut? Eine erste Antwort liefert uns der folgende wichtige Satz:

Satz 5.9. *Sei K ein Körper, V ein K-Vektorraum.*
Sei $\mathcal{V} := (v_1, \ldots, v_n)$ eine Familie von Vektoren in V.

(a) Die Abbildung

$$\begin{array}{rrcl} \Phi\colon & K^n & \to & V \\ & (\lambda_1, \ldots, \lambda_n) & \mapsto & \lambda_1 v_1 + \ldots + \lambda_n v_n \end{array}$$

ist stets ein K-Vektorraumhomomorphismus.

(b) Genau dann ist $(v_1, \ldots, v_n)$ eine Basis von V, wenn $\Phi_{\mathcal{V}}$ ein Vektorraumisomorphismus ist.
Insbesondere heißt das, dass man jeden Vektor von V auf genau eine Weise als Linearkombination der Basisvektoren darstellen kann.

Beweis. (a) Φ ist ein Vektorraumhomomorphismus:
Seien $(\lambda_1, \ldots, \lambda_n)$ und $(\mu_1, \ldots, \mu_n)$ zwei Vektoren aus K^n. Dann gilt

$$\begin{aligned} \Phi((\lambda_1, \ldots, \lambda_n) + (\mu_1, \ldots, \mu_n)) &= \Phi((\lambda_1 + \mu_1, \ldots, \lambda_n + \mu_n)) \\ &= (\lambda_1 + \mu_1)v_1 + \ldots + (\lambda_n + \mu_n)v_n \\ &= (\lambda_1 v_1 + \ldots + \lambda_n v_n) + (\mu_1 v_1 + \ldots + \mu_n v_n) \\ &= \Phi((\lambda_1, \ldots, \lambda_n)) + \Phi((\mu_1, \ldots, \mu_n)) \end{aligned}$$

Ebenso zeigt man, dass $\Phi(\lambda(\lambda_1, \ldots, \lambda_n)) = \lambda\Phi((\lambda_1, \ldots, \lambda_n))$.

(b) Wir zeigen zunächst, dass Φ genau dann surjektiv ist, wenn $(v_1, \ldots, v_n)$ ein Erzeugendensystem ist. Dann zeigen wir, dass Φ genau dann injektiv ist, wenn $(v_1, \ldots, v_n)$ eine linear unabhängige Familie ist.

- $(v_1, \ldots, v_n)$ Erzeugendensystem $\Leftrightarrow \Phi$ surjektiv:
  Das ist genau die Definition eines Erzeugendensystems.
- $(v_1, \ldots, v_n)$ linear unabhängig $\Rightarrow \Phi$ injektiv:
  Sei $\Phi((\lambda_1, \ldots, \lambda_n)) = \Phi((\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n))$, also

  $$\lambda_1 v_1 + \ldots + \lambda_n v_n = \tilde{\lambda}_1 v_1 + \ldots + \tilde{\lambda}_n v_n.$$

  Dann folgt

  $$0 = (\lambda_1 v_1 + \ldots + \lambda_n v_n) - (\tilde{\lambda}_1 v_1 + \ldots + \tilde{\lambda}_n v_n) = (\lambda_1 - \tilde{\lambda}_1)v_1 + \ldots + (\lambda_n - \tilde{\lambda}_n)v_n$$

  ist eine Darstellung des Nullvektors als Linearkombination der v_i, also ist $\lambda_i - \tilde{\lambda}_i = 0$ für alle $i \in \{1, \ldots, n\}$, also $\Phi((\lambda_1, \ldots, \lambda_n)) = \Phi((\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n))$

- Φ injektiv $\Rightarrow (v_1, \ldots, v_n)$ linear unabhängig:
  Sei $0 = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Dann ist also $\Phi((\lambda_1, \ldots, \lambda_n)) = 0$. Wir kennen aber bereits ein Urbild der 0, nämlich $(0, \ldots, 0)$. Da jeder Vektor in V höchstens ein Urbild haben kann, folgt $\lambda_1 = \ldots = \lambda_n = 0$.

$\square$

Korollar 5.10. *Sei K ein Körper.*
Die Familie $\mathcal{E} := (e_1, \ldots, e_n)$ der Vektoren $e_j := (0, \ldots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \ldots, 0)$ ist eine Basis des Vektorraums K^n.

Beweis. Wir betrachten die Abbildung $\Phi_{\mathcal{E}}$. Es gilt offensichtlich $\Phi_{\mathcal{E}}(\lambda_1, \ldots, \lambda_n) = \lambda_1(1, 0, \ldots, 0) + \ldots + \lambda_n(0, \ldots, 0, 1) = (\lambda_1, \ldots, \lambda_n)$, also ist $\Phi_{\mathcal{E}} = \mathrm{id}_{K^n}$ und somit offensichtlich bijektiv, also ein Vektorraumisomorphismus. $\square$

Definition 5.11. Sei K ein Körper.
Die Familie $\mathcal{E} := (e_1, \ldots, e_n)$ der Vektoren $e_j := (0, \ldots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \ldots, 0)$ heißt *Standardbasis* des Vektorraums K^n.

Ebenso wichtig und nützlich ist die folgende Folgerung:

Korollar 5.12. *Sei K ein Körper und seien V und W zwei K-Vektorräume.*
Wenn $(v_1, \ldots, v_n)$ eine Basis ist und $w_1, \ldots, w_n \in W$ beliebig, dann gibt es genau eine lineare Abbildung $\phi\colon V \to W$ mit $\phi(v_1) = w_1, \ldots, \phi(v_n) = w_n$.
Jede lineare Abbildung $\phi\colon V \to W$ ist also eindeutig festgelegt durch die Bilder der Vektoren einer beliebigen Basis von V.

Beweis. Die Eindeutigkeit sieht man so: Sei $v \in V$ beliebig. Dann lässt sich v auf genau eine Weise als Linearkombination $\lambda_1 v_1 + \ldots + \lambda_n v_n$ der Basisvektoren darstellen.
Dann folgt aber:

$$\phi(v) = \phi(\lambda_1 v_1 + \ldots + \lambda_n v_n) \overset{\phi \text{ linear}}{=} \lambda_1 \phi(v_1) + \ldots + \lambda_n \phi(v_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n.$$

die Abbildung ist also bereits für jeden Vektor v durch die Werte $w_1, \ldots, w_n$ festgelegt.
Für die Existenz müssen wir nur zeigen, dass die Abbildung, die jedem Vektor $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ den Wert $\lambda_1 w_1 + \ldots + \lambda_n w_n$ zuordnet, tatsächlich linear ist.
Das ist aber genau die Hintereinanderausführung der Abbildung $\Phi_{\mathcal{V}}^{-1}$ für $\mathcal{V} := (v_1, \ldots, v_n)$ und der Abbildung $\Phi_{\mathcal{W}}$ für $\mathcal{W} := (w_1, \ldots, w_n)$. $\square$

Korollar 5.13. *Sei K ein Körper und V ein K-Vektorraum.*
Dann ist jede lineare Abbildung $\Phi : K^n \to V$ von der Form $\Phi = \Phi_{\mathcal{V}}$ für die Familie $\mathcal{V} := (\Phi(e_1), \ldots, \Phi(e_n))$ der Bilder der Standardbasisvektoren.

Beweis. Setze $w_j := \Phi(e_j)$ für alle $j \in \{1, \ldots, n\}$.
Es gibt nach Korollar 5.12 genau eine Abbildung $\phi\colon K^n \to V$ mit $\phi(e_1) = w_1, \ldots, \phi(e_n) = w_n$, nämlich Φ.
Für die Abbildung $\Phi_{\mathcal{V}}$ gilt aber ebenfalls:

$$\Phi_{\mathcal{V}}(e_1) = \Phi_{\mathcal{V}}(1, 0, \ldots, 0) = 1w_1 + 0w_2 + \ldots + 0w_n = w_1$$

und allgemein

$$\forall j \in \{1, \ldots, n\}\colon \quad \Phi_{\mathcal{V}}(e_j) = w_j.$$

Damit folgt $\Phi = \Phi_{\mathcal{V}}$. $\square$

Die zweite Frage ist, wie man Basen finden kann. Hat überhaupt jeder Vektorraum eine Basis? Der folgende Satz beantwortet das zumindest schon einmal für alle endlich erzeugten Vektorräume.

Satz 5.14 (Basisauswahlsatz)**.** *Sei K ein Körper, V ein K-Vektorraum.*
Sei die endliche Familie $(v_1, \ldots, v_r)$ ein Erzeugendensystem von V. Dann gibt es Indizes $i_1, \ldots, i_n \in \{1, \ldots, r\}$, sodass die Familie $(v_{i_1}, \ldots, v_{i_n})$ eine Basis ist.

Beweis. Wir verkleinern das Erzeugendensystem Schritt für Schritt.
Wenn das Erzeugendensystem bereits linear unabhängig ist, so sind wir bereits fertig.
Wenn das Erzeugendensystem nicht linear unabhängig ist, gibt es nach Satz 5.7 einen Index $i_0 \in \{1, \ldots, r\}$, sodass bereits $\{v_1, \ldots, v_r\} \setminus \{v_{i_0}\}$ ein Erzeugendensystem ist.
Prüfe nach jedem Schritt, ob das Erzeugendensystem linear unabhängig ist, sonst verkleinere weiter. Spätestens wenn wir keinen Vektor mehr übrig haben, ist die leere Familie, die übrig bleibt, linear unabhängig. Da wir mit nur endlich vielen Vektoren angefangen haben, endet das Verfahren also irgendwann. □

Lemma 5.15 (Austauschlemma)**.** *Sei K ein Körper, V ein K-Vektorraum.*
Sei $(v_1, \ldots, v_n)$ eine Basis von V und $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ mit $\lambda_i \in K$ für $i \in \{1, \ldots, n\}$. Falls $j \in \{1, \ldots, n\}$ existiert mit $\lambda_j \neq 0$, dann ist auch $(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n\}$ eine Basis.

Beweis. Nach Umnummerieren können wir oBdA $j = 1$ annehmen. Zu zeigen ist also: $(w, v_2, \ldots, v_n)$ ist linear unabhängiges Erzeugendensystem.
$(w, v_1, \ldots, v_n)$ ist linear unabhängig:
Sei

$$\mu_1 w + \mu_2 v_2 + \ldots + \mu_n v_n = 0$$

eine Darstellung des Nullvektors als Linearkombination der Vektoren $w, v_2, \ldots, v_n$.
Dann ist also

$$0 = \mu_1(\lambda_1 v_1 + \ldots + \lambda_n v_n) + \mu_2 v_2 + \ldots + \mu_n v_n = \mu_1 \lambda_1 v_1 + \ldots + (\mu_1 \lambda_2 + \mu_2) v_2 + \ldots (\mu_1 \lambda_n + \mu_n) v_n$$

eine Darstellung des Nullvektors als Linearkombination der linear unabhängigen Vektoren $v_1, \ldots, v_n$, also

$$\mu_1 \lambda_1 = \mu_1 \lambda_2 + \mu_2 = \ldots \mu_1 \lambda_n + \mu_n = 0.$$

Da $\lambda_1 \neq 0$, muss $\mu_1 = 0$ sein. Dann folgt sofort, dass $\mu_2 = \ldots = \mu_n = 0$.
$(w, v_1, \ldots, v_n)$ ist Erzeugendensystem:
Da $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ mit $\lambda_1 \neq 0$, ist

$$v_1 = \lambda_1^{-1} w - \lambda_1^{-1} \lambda_2 v_2 - \ldots - \lambda_1^{-1} \lambda_n v_n.$$

Sei nun $v \in V$ beliebig. Da $(v_1, \ldots, v_n)$ ein Erzeugendensystem ist, gibt es eine Darstellung

$$v = \mu_1 v_1 + \ldots + \mu_n v_n$$

der Vektoren $v_1, \ldots, v_n$. Dann ist aber

$$\begin{aligned} v &= \mu_1(\lambda_1^{-1} w - \lambda_1^{-1} \lambda_2 v_2 - \ldots - \lambda_1^{-1} \lambda_n v_n) + \mu_2 v_2 + \ldots + \mu_n v_n \\ &= \mu_1 \lambda_1^{-1} w + (\mu_2 - \lambda_1^{-1} \lambda_2) v_2 + \ldots + (\mu_n - \lambda_1^{-1} \lambda_n) v_n \end{aligned}$$

eine Darstellung als Linearkombination der Vektoren $w, v_2, \ldots, v_n$. □

Insbesondere folgt sofort die folgende Aussage:

Korollar 5.16. *Sei K ein Körper und V ein K-Vektorraum. Ist $(v_1, \ldots, v_n)$ eine Basis von V und $\lambda \neq 0$, dann ist auch $(v_1, \ldots, v_{j-1}, \lambda v_j, v_{j+1}, \ldots, v_n)$ für beliebig gewähltes $j \in \{1, \ldots, n\}$ eine Basis von V.*

Satz 5.17 (Basisaustauschsatz von Steinitz)**.** *Sei K ein Körper, V ein K-Vektorraum. Ist $(v_1, \ldots, v_n)$ eine Basis von V und $(w_1, \ldots, w_m)$ eine linear unabhängige Familie von Vektoren in V, dann gilt:*

(a) $m \leq n$

(b) $(v_1, \ldots, v_n)$ kann so umnummeriert werden, dass $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$ eine Basis von V ist.

Beweis. Wir beweisen die Aussage durch Induktion über m: **Induktionsanfang** $m = 0$: Es ist nichts zu zeigen.
Induktionsannahme: Für jede linear unabhängige Familie $(w_1, \ldots, w_m)$ von Vektoren aus V ist $m \leq n$ und nach geeigneter Umnummerierung ist $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$ eine Basis.
Induktionsschritt „$m \to m + 1$": $(w_1, \ldots, w_m)$ sei eine linear unabhängige Familie von Vektoren in V. Dann ist $(w_1, \ldots, w_m)$ nach Lemma 5.4 linear unabhängig und wir können die Induktionsannahme darauf anwenden.
Wir zeigen jetzt beide Teilaussagen: $m + 1 \leq n$:
Beweis durch Widerspruch:
Angenommen, $m + 1 > n$. Da wir $m \leq n$ bereits wissen, folgt dann $m = n$. Dann ist also (ohne dass Ummummerieren noch nötig ist) $(w_1, \ldots, w_m)$ eine Basis und damit nach Satz 5.7 eine unverlängerbare linear unabhängige Familie im Widerspruch zur Voraussetzung, dass $(w_1, \ldots, w_{m+1})$ linear unabhängig ist.
Nach Umnummerieren der v_i ist $(w_1, \ldots, w_{m+1}, v_{m+2}, \ldots, v_n)$ eine Basis:
Wir wissen, dass nach geeigneter Umnummerierung der v_i $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$ eine Basis, also insbesondere ein Erzeugendensystem ist. Dann lässt sich w_{m+1} also als Linearkombination $\lambda_1 w_1 + \ldots, +\lambda_m w_m + \lambda_{m+1} v_{m+1} + \ldots + \lambda_n v_n$ darstellen.
Wäre $\lambda_{m+1} = \ldots = \lambda_n = 0$, so wäre w_{m+1} als Linearkombination der Vektoren $w_1, \ldots, w_m$ darstellbar im Widerspruch zur linearen Unabhängigkeit. Also muss für mindestens ein $j \in \{m + 1, \ldots, n\}$ $\lambda_j \neq 0$ gelten, oBdA $\lambda_{m+1} \neq 0$. Nach Lemma 5.15 können wir also v_{m+1} durch w_{m+1} ersetzen. □

Korollar 5.18. *Sei K ein Körper, V ein K-Vektorraum.*
Wenn V eine Basis mit n Elementen hat, dann hat jede Basis von V genau n Elemente.

Beweis. Seien $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ zwei Basen von V. Da beide Basen linear unabhängige Familien sind, gilt sowohl $m \leq n$ als auch $n \leq m$, also $m = n$. □

Definition 5.19. Sei K ein Körper, V ein K-Vektorraum.
Wenn V eine Basis mit n Elementen hat und damit jede Basis von V aus n Elementen besteht, nennen wir n die *Dimension* von V und schreiben $\dim_K(V) := n$. In diesem Fall sagen wir, V sei *endlichdimensional*.
Wenn V keine Basis mit endlich vielen Elementen hat, dann nennen wir V *unendlichdimensional* und schreiben $\dim_K(V) := \infty$.

Bemerkung 5.20. Interessant ist, dass auch jeder unendlichdimensionale Vektorraum eine Basis besitzt. Das ist aber für uns etwas zu schwer zu zeigen.

6 Matrizen

Definition 6.1. Eine $(m \times n)$-*Matrix* ist ein rechteckiges Schema

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

wobei die *Einträge* a_{ij} aus einer beliebigen Menge X stammen.
Wir schreiben abgekürzt $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ oder auch noch kürzer $A = (a_{ij})$.

$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ heißt i-te *Zeile* von A, $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ heißt j-te *Spalte* von A.

Die Menge aller $(m \times n)$-Matrizen mit Einträgen in einer Menge X bezeichnen wir mit $\mathrm{Mat}_X(m \times n)$.

Bemerkung 6.2. Im Folgenden seien die Einträge der Matrizen stets aus einem Körper K.

Definition 6.3. Gegeben zwei Matrizen $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ und $B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ & \vdots & \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$ aus $\mathrm{Mat}_K(m \times n)$ definieren wir ihre *Summe* $A + B := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ & \vdots & \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$.

Für $\lambda \in K$ definieren wir weiter $\lambda A := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ & \vdots & \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$.

Bemerkung 6.4. Man sieht leicht, dass $\mathrm{Mat}_K(m \times n)$ mit den so definierten Addition und Skalarmultiplikation ein K-Vektorraum ist. Der Nullvektor dieses Vektorraums ist die $(m \times n)$-Matrix, deren Einträge alle 0 sind.

Wenn die Vektorraumstruktur alles wäre, was wir von Matrizen erwarten, könnten wir die insgesamt mn Einträge der Matrix auch in ein Tupel schreiben. Zwei Matrizen können wir aber unter den geeigneten Voraussetzungen nicht nur addieren, sondern auch miteinander multiplizieren:

Definition 6.5. Sei $A = (a_{ij})$ eine $(l \times m)$-Matrix und $B = (b_{ij})$ eine $(m \times n)$-Matrix. Dann definieren wir das *Matrixprodukt* $AB := C = (c_{ij}) \in \mathrm{Mat}_K(l \times n)$ als die Matrix, deren Einträge sich berechnen als

$$c_{ij} := \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Bemerkung 6.6. Achtung! Zwei Matrizen kann man nur multiplizieren, wenn die erste genau so viele Spalten hat wie die zweite Zeilen.

Jetzt stellen sich natürlich ein paar Fragen: Welche Eigenschaften hat das Matrixprodukt? Wie verträgt es sich mit den anderen beiden Operationen? Und vor allem: Wie kommt jemand auf so eine seltsame Definition?
Die Antworten auf diese Fragen werden sich im Laufe des Kapitels präsentieren. Am Anfang steht dabei eine ziemlich trivial klingende Beobachtung:

Bemerkung 6.7. Man kann jeden Vektor $(x_1, \ldots, x_n)$ aus K^n auf zwei natürliche Weisen als Matrix auffassen:
Einmal als $(1 \times n)$-Matrix: $\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$ („Zeilenvektor")
Oder als $(n \times 1)$-Matrix: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ („Spaltenvektor")
Besonders die letztere Identifikation müssen wir uns unbedingt merken.

Lemma 6.8. *Sei $A = (a_{ij}) \in \mathrm{Mat}_K(m \times n)$. Dann ist die Abbildung*

$$\begin{array}{rcccl} F_A\colon & K^n & \to & K^m & \\ x = & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & Ax = & \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{pmatrix} \end{array}$$

ein Vektorraumhomomorphismus von K^n nach K^m.

Beweis. Wir stellen zunächst einmal fest, dass die Abbildung tatsächlich in K^m landet, da die Bildmatrix eine $(m \times 1)$-Matrix ist, die wir wieder als Vektor von K^m auffassen können. Dann prüfen wir die beiden Axiome für lineare Abbildungen nach:

(L1) Seien $x, y \in K^n$ beliebig. Es gilt:

$$\begin{aligned} F_A(x+y) &= F_A(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}) \\ &= F_A(\begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}) \\ &= \begin{pmatrix} \sum_{k=1}^n a_{1k}(x_k + y_k) \\ \vdots \\ \sum_{k=1}^n a_{mk}(x_k + y_k) \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k + \sum_{k=1}^n a_{1k}y_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k + \sum_{k=1}^n a_{mk}y_k \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^n a_{1k}y_k \\ \vdots \\ \sum_{k=1}^n a_{mk}y_k \end{pmatrix} \\ &= F_A(x) + F_A(y) \end{aligned}$$

(L2) analog

$\square$

Es kommt aber noch besser: Tatsächlich ist jede lineare Abbildung von K^n nach K^m von dieser Form:

Satz 6.9. *Die Abbildung*

$$\begin{array}{rrcl} \Psi: & \mathrm{Mat}_K(m\times n) & \to & \mathrm{Hom}_K(K^m, K^n) \\ & A & \mapsto & F_A \end{array}$$

ist ein K-Vektorraum-Isomorphismus.

Beweis. Wir zeigen zunächst, dass die Abbildung linear, also ein Vektorraumhomomorphismus, ist. Dann zeigen wir Surjektivität und Injektivität.

- Ψ ist linear:
  Seien $A, B \in \mathrm{Mat}_K(m\times n)$ beliebig. Wir wollen prüfen, dass $F_{A+B} = F_A + F_B$ ist. Dazu setzen wir in beide Seiten alle möglichen Elemente ein und prüfen, dass jedes Mal dasselbe herauskommt.
  Sei also $x = (x_1, \ldots, x_n) \in K^n$ beliebig. Dann ist

$$\begin{aligned} F_{A+B}(x) &= (A+B)x \\ &= \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ & \vdots & \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^n (a_{1k}+b_{1k})x_k \\ \vdots \\ \sum_{k=1}^n (a_{mk}+b_{mk})x_k \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^n b_{1k}x_k \\ \vdots \\ \sum_{k=1}^n b_{mk}x_k \end{pmatrix} \\ &= Ax + Bx \\ &= F_A(x) + F_B(x). \end{aligned}$$

  Die Homogenität, also $F_{\lambda A} = \lambda F_A$, zeigt man ebenso.

- Ψ ist surjektiv:
  Wir wissen aus Korollar 5.12, dass jede lineare Abbildung eindeutig durch die Bilder aller Basisvektoren gegeben ist.
  Sei $(e_1, \ldots, e_n)$ also die Standardbasis des K^n und seien ihre Bilder unter der Abbildung F bezeichnet mit $F(e_j) =: \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$.
  Betrachte dann die Matrix $A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$. Es gilt

$$F_A(e_1) = Ae_1 = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} = F(e_1)$$

und entsprechend $F_A(e_j) = F(e_j)$ für alle $j \in \{1, \ldots, n\}$. Also sind die Abbildungen F_A und F gleich.

- Ψ ist injektiv:
  Wir zeigen: Wenn wir zwei Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ nehmen, die verschieden sind, dann sind auch die linearen Abbildung F_A und F_B verschieden.
  Wenn die Matrizen verschieden sind, muss es also mindestens einen Eintrag $a_{ij} \neq b_{ij}$ geben, an dem sie sich unterscheiden. Betrachte dann das Bild des j-ten Standardvektors e_j unter den beiden Abbildungen:
  $$F_A(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad F_B(e_j) = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix}.$$
  Die beiden Bilder sind verschieden, nämlich mindestens im i-ten Eintrag. Also sind die Abbildungen F_A und F_B verschieden.

□

Bemerkung 6.10. Was wir aus dem Beweis mitnehmen sollten, ist die Umkehrabbildung, die wir zum Beweis der Surjektivität konstruiert haben:
Ist $F : K^n \to K^m$ eine lineare Abbildung, so ergibt sich die dazugehörige Matrix, indem man die Bilder der Standardbasisvektoren $e_1, \ldots, e_n$ hintereinander als Spalten in die Matrix einträgt.

Wir haben bereits gesehen, dass die Hintereinanderausführung zweier linearer Abbildungen wieder linear ist. Wir können uns also fragen, welche Matrix dazugehört, wenn die einzelnen Abbildungen die Matrizen A und B haben. Das wird durch das folgende Lemma beantwortet:

Lemma 6.11. *Seien $A \in \mathrm{Mat}_K(l \times m)$, $B \in \mathrm{Mat}_K(m \times n)$. Dann ist*
$$F_A \circ F_B = F_{AB}.$$

Beweis. Wir nutzen die Bemerkung 6.10, um die Matrix zu finden:
Sei e_j ein Standardbasisvektor des K^n. Dann gilt
$$(F_A \circ F_B)(e_j) = F_A(F_B(e_j)) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k} b_{kj} \\ \vdots \\ \sum_{k=1}^n a_{mk} b_{kj} \end{pmatrix},$$
was also unsere j-te Spalte der gesuchten Matrix ergibt. Das ist aber genau die j-te Spalte der Produktmatrix AB. □

Die beiden Resultate liefern uns eine Fülle von weiteren:

Bemerkung 6.12. Da wir wissen, dass die Komposition von Abbildungen assoziativ ist in dem Sinne, dass $(f \circ g) \circ h = f \circ (g \circ h)$ für alle solchen Abbildungen, die sich komponieren lassen, folgt sofort, dass auch $(AB)C = A(BC)$ für alle passenden Matrizen.
Weiter folgt aus der einfachen Tatsache $(f + g) \circ h = f \circ h + g \circ h$ für Abbildungen die Aussage $(A + B)C = AC + BC$.
Entsprechend gilt für lineare Abbildungen auch $f \circ (g + h) = f \circ g + f \circ h$, also $A(B + C) = AB + AC$.

Wenn wir $m = n$ wählen, sehen wir, dass die $(n \times n)$-Matrizen den Endomorphismen des K^n entsprechen.
Dort gibt es nun sogar noch zusätzliche Strukturen: Zum einen ein neutrales Element id_{K^n}. Da wir für diese Abbildung einfach die n Standardbasisvektoren nebeneinander in die Matrix schreiben müssen, entspricht ihr natürlich die Matrix $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$, die also das neutrale Element der Matrizenmultiplikation ist.

Definition 6.13. Die Matrix

$$E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathrm{Mat}_K(n \times n)$$

heißt *Einheitsmatrix.*

Mit einem neutralen Element und der Assoziativität der Matrixmultiplikation aus Bemerkung 6.12 bilden die $(n \times n)$-Matrizen ein Monoid. Sind sie vielleicht sogar eine Gruppe?
Die Antwort ist nein, denn offensichtlich ist das Produkt der Nullmatrix mit jeder weiteren Matrix wieder die Nullmatrix. Damit kann sie kein Inverses haben. Die Frage ist nun: Gibt es außer der Nullmatrix noch weitere Matrizen, die kein Inverses haben, und wenn ja, wie findet man das für eine gegebene Matrix heraus? Das wird das Thema des letzten Kapitels.
Jetzt wollen wir aber noch eine weitere Überlegung zu Ende führen:
Bisher haben wir nur die Abbildungen zwischen den speziellen Vektorräumen K^n und K^m angesehen. Wie sieht es mit denen zwischen allgemeinen endlichdimensionalen Vektorräumen aus?

Bemerkung 6.14. Wir erinnern uns an Satz 5.9: Gegeben ein beliebiger endlichdimensionaler Vektorraum V mit Basis $\mathcal{A} := (v_1, \ldots, v_n)$ gibt es einen Vektorraumisomorphismus

$$\begin{array}{rrcl} \Phi_{\mathcal{A}}\colon & K^n & \to & V \\ & (\lambda_1, \ldots, \lambda_n) & \mapsto & \lambda_1 v_1 + \ldots + \lambda_n v_n. \end{array}$$

Damit können wir jede lineare Abbildung F von einem K-Vektorraum V mit Basis $\mathcal{A} := (v_1, \ldots, v_n)$ in einen anderen Vektorraum W mit Basis $\mathcal{B} := (w_1, \ldots, w_m)$ zu einer Abbildung von K^n nach K^m umwandeln, indem wir erst von K^n über $\Phi_{\mathcal{A}}$ nach V abbilden, dann mit F nach W und schließlich von W mit der Umkehrabbildung $\Phi_{\mathcal{B}}^{-1}$ nach K^m gehen. Das die Hintereinanderausführung dreien linearer Abbildungen wieder linear ist, ergibt das eine lineare Abbildung von K^n nach K^m, zu der es eine Matrix gibt.
Diese Matrix nennen wir die *darstellende Matrix* zu der Abbildung F bezüglich der gegebenen Basen $\mathcal{A}$ und $\mathcal{B}$ und notieren sie

$$M(\mathcal{B}, F, \mathcal{A}).$$

Die Basis des Zielraums steht dabei an erster Stelle.

7 Multilinearformen

Bevor wir mit der Invertierbarkeit von Matrizen beschäftigen, erarbeiten wir uns Konzepte aus einem anderen Bereich:
Bisher haben wir lineare Abbildungen zwischen K-Vektorräumen V und W angesehen. Eine Möglichkeit, das Konzept zu erweitern, besteht darin, Abbildungen von einem Produkt von Vektorräumen $V_1 \times \cdots \times V_n$ nach W anzusehen, die ganze Tupel von n Vektoren $v_1 \in V_1, \ldots, v_n \in V_n$ nehmen und nach W abbilden, sodass die Abbildung in jeder Komponente linear ist, wenn man alle anderen Komponenten festhält.

Definition 7.1. Sei K ein Körper und seien $V_1, \ldots, V_n$ sowie W K-Vektorräume.
Eine Abbildung $\Phi\colon V_1 \times \cdots \times V_n \to W$ heißt *multilinear*, wenn die Abbildung

$$\begin{aligned} V_i &\to W \\ x &\mapsto \Phi(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) \end{aligned}$$

für jede Wahl von $i \in \{1, \ldots, n\}$ und $v_1 \in V_1, \ldots, v_{i-1} \in V_{i-1}, v_{i+1} \in V_{i+1}, \ldots, v_n \in V_n$ linear ist.
Im Fall von $n = 2$ spricht man von einer *bilinearen Abbildung.*

Bemerkung 7.2. Die Menge aller multilinearen Abbildungen von $V_1 \times \ldots \times V_n$ nach W bildet offensichtlich einen Untervektorraum aller Abbildungen von $V_1 \times \ldots \times V_n$ nach W. Da wir nur die Linearität in jeder Komponente prüfen müssen, können wir im Prinzip den Beweis des Falls $n = 1$, also des Falles linearer Abbildungen, übernehmen.

Ein spezieller Fall ist, wenn erstens $V_1 = \ldots = V_n$ alle derselbe Vektorraum V sind und zweitens W der Grundkörper K (der immer auch ein K-Vektorraum ist).

Definition 7.3. Sei K ein Körper und V ein K-Vektorraum.
Eine multilineare Abbildung $\Phi\colon V^n \to K$ heißt *Multilinearform.*
Im Fall $n = 2$ spricht man von einer *Bilinearform.*
Den Vektorraum der Multilinearformen auf V wird mit $\mathcal{J}^n_K(V)$ bezeichnet.

Definition 7.4. Sei K ein Körper und V ein K-Vektorraum.
Eine Multilinearform $\Phi\colon V^n \to K$ heißt

(a) *symmetrisch*, wenn für alle Vektoren $v_1, \ldots, v_n \in V$ und alle $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ gilt:
$$\Phi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = \Phi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n),$$

(b) *antisymmetrisch*, wenn für alle Vektoren $v_1, \ldots, v_n \in V$ und alle $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ gilt:
$$\Phi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\Phi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n),$$

(c) *alternierend*, wenn für alle Vektoren $v_1, \ldots, v_n \in V$ und alle $i \in \{1, \ldots, n\}$ gilt:
$$\Phi(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) = 0.$$

Bemerkung 7.5. Eine Bilinearform ist also symmetrisch, wenn $\forall v, w \in V : B(v, w) = B(w, v)$, antisymmetrisch, wenn $\forall v, w \in V : B(v, w) = -B(w, v)$, und alternierend, wenn $\forall v \in V : B(v, v) = 0$.

Lemma 7.6. *Sei K ein Körper und V ein K-Vektorraum. Jede alternierende Multilinearform $\Phi\colon V^n \to K$ ist auch antisymmetrisch.*

Beweis. Seien $v_1, \ldots, v_n \in V$ beliebig und $i, j \in \{1, \ldots, n\}$ verschieden. Dann gilt:

$$\begin{aligned}
0 &\overset{\Phi \text{ alternierend}}{=} \Phi(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) \\
&\overset{\substack{\text{Linearität im} \\ i\text{-ten Argument}}}{=} \Phi(v_1, \ldots, v_i, \ldots, v_i + v_j, \ldots, v_n) + \Phi(v_1, \ldots, v_j, \ldots, v_i + v_j, \ldots, v_n) \\
&\overset{\substack{\text{Linearität im} \\ j\text{-ten Argument}}}{=} \Phi(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + \Phi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) \\
&\quad + \Phi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + \Phi(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) \\
&\overset{\Phi \text{ alternierend}}{=} \Phi(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + \Phi(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)
\end{aligned}$$

□

Ihr habt das Wort „bilinear“ schon einmal gehört: Im Kontext mehrdimensionaler Analysis. Auf dem $\mathbb{R}^n$ habt ihr das *Standardskalarprodukt*

$$\begin{aligned}
\langle \cdot, \cdot \rangle \colon \quad \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\
((x_1, \ldots, x_n), (y_1, \ldots, y_n)) &\mapsto x_1 y_1 + \ldots + x_n y_n.
\end{aligned}$$

kennengelernt.
Jetzt, aus dem Blickwinkel der linearen Algebra ist sollte diese Summe irgendwie vertraut vorkommen – das ist doch das Matrixprodukt der Zeilenmatrix $\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$ mit der Spaltenmatrix $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$!
Damit ist die Bilinearität der Abbildung bewiesen, denn für Matrizen gilt sowohl $(A + B)C = AB + AC$ und $(\lambda A)B = (\lambda A)B$, also Linearität in der ersten Komponente, als auch $A(B + C) = AB + AC$ und $A(\lambda B) = \lambda AB$, also Linearität in der zweiten Komponente.
Mehr noch, damit haben wir eine Idee für eine ganze Klasse von Bilinearformen auf einem allgemeinen K^n:

Beispiel 7.7. Sei K ein Körper.
Für jede quadratische Matrix $A = (a_{ij}) \in \mathrm{Mat}_K(n \times n)$ definiert die Abbildung

$$\begin{aligned}
\langle \cdot, \cdot \rangle_A \colon \quad \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\
((x_1, \ldots, x_n), (y_1, \ldots, y_n)) &\mapsto \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.
\end{aligned}$$

eine Bilinearform. Wir schreiben auch $x^\top A y := \langle x, y \rangle_A$ und meinen damit, dass wir den Vektor y als Spaltenvektor, x als Zeilenvektor auffassen.

Tatsächlich sind alle Bilinearformen auf dem Vektorraum K^n von dieser Gestalt. Mehr noch, wir können sie mit den $(n \times n)$-Matrizen identifizieren:

Satz 7.8. *Sei K ein Körper.*
Die Abbildung

$$\begin{array}{rrcl} \Psi: & \mathrm{Mat}_K(n\times n) & \to & \mathcal{J}^2_K(K^n) \\ & A & \mapsto & \langle\cdot,\cdot\rangle_A \end{array}$$

ist ein Vektorraum-Isomorphismus.

Beweis. Es gilt für alle x, y: $x^\top(A+B)y = x^\top Ay + x^\top By$ und $x^\top(\lambda A)y = \lambda x^\top Ay$ nach den Regeln der Matrizenmultiplikation, also ist die Abbildung ein Vektorraumhomomorphismus.
Für die Surjektivität sei B eine beliebige Bilinearform auf K^n. Betrachte dann die Matrix $A := \begin{pmatrix} B(e_1,e_1) & \dots & B(e_1,e_n) \\ \vdots & & \vdots \\ B(e_n,e_1) & \dots & B(e_n,e_n) \end{pmatrix}$, wobei $(e_1,\dots,e_n)$ die Standardbasis des K^n ist.
Es gilt für jedes Paar von Vektoren $x = (x_1,\dots,x_n) = x_1e_1 + \dots + x_ne_n$ und $y = (y_1,\dots,y_n) = y_1e_1 + \dots + y_ne_n$:

$$\begin{aligned} & \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} B(e_1,e_1) & \dots & B(e_1,e_n) \\ \vdots & & \vdots \\ B(e_n,e_1) & \dots & B(e_n,e_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ = & \begin{pmatrix} \sum_{k=1}^n x_k B(e_k,e_1) & \cdots & \sum_{k=1}^n x_k B(e_k,e_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ = & \begin{pmatrix} B(\sum_{k=1}^n x_k e_k, e_1) & \cdots & B(\sum_{k=1}^n x_k e_k, e_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ = & \begin{pmatrix} B(x,e_1) & \cdots & B(x,e_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ = & \sum_{l=1}^n B(x,e_l)y_l \\ = & B(x, \sum_{l=1}^n y_l e_l) \\ = & B(x,y) \end{aligned}$$

Damit ist $B = \langle\cdot,\cdot\rangle_A = \Psi(A)$ für die angegebene Matrix A.
Für die Injektivität betrachte den Wert der Bilinearform $\langle e_i, e_j\rangle_A$ für zwei Standardbasisvektoren e_i und e_j.

Es ist offensichtlich

$$\begin{aligned}
&\begin{pmatrix} 0 & \cdots & 0 & 1 & 0\cdots & 0 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
= &\begin{pmatrix} 0 & \cdots & 0 & 1 & 0\cdots & 0 \end{pmatrix} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \\
= &\; a_{ij}.
\end{aligned}$$

Stimmen für zwei Matrizen die Bilinearformen überein, dann insbesondere für alle Paare von Standardbasisvektoren. Dann sind aber alle Einträge gleich und die Matrizen identisch.

Bemerkung 7.9. Was wir aus diesem Beweis mitnehmen sollten, ist die Vorschrift, wie man aus einer Bilinearform eine Matrix bildet: Man trägt in den Eintrag der i-ten Zeile und j-ten Spalte den Wert der Bilinearform an der Stelle (e_i, e_j) ein.

□

Bemerkung 7.10. Erinnern wir uns noch einmal an Satz 5.9: Für jeden endlichdimensionalen Vektorraum V mit Basis $\mathcal{A} := (v_1, \ldots, v_n)$ gibt es einen Vektorraumisomorphismus

$$\begin{aligned}
\Phi_{\mathcal{A}}\colon \quad K^n &\to V \\
(\lambda_1, \ldots, \lambda_n) &\mapsto \lambda_1 v_1 + \ldots + \lambda_n v_n.
\end{aligned}$$

Diesen können wir nutzen, um auch Bilinearformen auf beliebigen endlichdimensionalen Vektorräumen eine Matrix zuzuordnen.
Ist nämlich $B\colon V \times V \to K$ eine Bilinearform, dann ist auch

$$\begin{aligned}
\tilde{B}_{\mathcal{A}}\colon \quad K^n \times K^n &\to K \\
(x, y) &\mapsto B(\Phi_{\mathcal{A}}(y), \Phi_{\mathcal{A}}(y))
\end{aligned}$$

eine Bilinearform: Für alle $x, x' \in K^n$ und alle $y \in K^n$ gilt nämlich wegen der Linearität von Φ:

$$\begin{aligned}
\tilde{B}_{\mathcal{A}}(x + x', y) &= B(\Phi(x + x'), \Phi(y)) = B(\Phi(x) + \Phi(x'), \Phi(y)) \\
&= B(\Phi(x), \Phi(y)) + B(\Phi(x'), \Phi(y)) = \tilde{B}_{\mathcal{A}}(x, y) + \tilde{B}_{\mathcal{A}}(x', y).
\end{aligned}$$

Die Homogenität im ersten Argument und die Additivität und Homogenität im zweiten zeigt man analog.
Da $\tilde{B}_{\mathcal{A}}$ eine Bilinearform auf K^n ist, gibt es dazu eine Matrix; diese nennen wir die *darstellende Matrix* zu der Bilinearform B bezüglich der gegebenen Basis $\mathcal{A}$ und schreiben

$$M_{\mathcal{A}}(B).$$

Wie man sieht, wenn man die Definition der darstellenden Matrix ausschreibt, ist der Eintrag in der i-ten Zeile und j-ten Spalte der Wert $B(v_i, v_j)$. Außerdem folgt aus der Definition für zwei beliebige Vektoren $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ und $w = \mu_1 v_1 + \ldots + \mu_n v_n$:

$$B(v, w) = \tilde{B}_{\mathcal{A}}(\Phi^{-1}(v), \Phi^{-1}(w)) = \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix} M_{\mathcal{A}}(B) \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Bemerkung 7.11. Ist V endlichdimensional, so können wir die Bilinearform bezüglich einer beliebig vorgegebenen Basis $\mathcal{A} = (v_1, \ldots, v_n)$ durch eine Matrix darstellen. Eine Bilinearform B ist genau dann symmetrisch, wenn die zugehörige Matrix $M_\mathcal{A}(B) = (a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}$ symmetrisch ist, also $a_{ij} = a_{ji}$ für alle $i, j \in \{1, \ldots, n\}$:
Das das notwendig ist, ist klar, denn $a_{ij} = B(v_i, v_j) \overset{\text{Symmetrie}}{=} B(v_j, v_i) = a_{ji}$.
Hinreichend hinwiederum ist es, weil dann für je zwei Vektoren $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ und $w = \mu_1 v_1 + \ldots + \mu_n v_n$ gilt:

$$B(v, w) = \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix} M_\mathcal{A}(B) \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \sum_{k=1}^{n}\sum_{l=1}^{n} \lambda_k a_{kl} \mu_l = \sum_{k=1}^{n}\sum_{l=1}^{n} \lambda_k a_{lk} \mu_l$$

$$= \sum_{l=1}^{n}\sum_{k=1}^{n} \mu_l a_{lk} \lambda_k = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} M_\mathcal{A}(B) \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix} = B(w, v).$$

Während wir Symmetrie für beliebige Multilinearformen und einen beliebigen Grundkörper definieren konnten, benötigen wir für die nächste Definition die Ordnungsrelation der reellen Zahlen und sie gilt auch nur für bilineare Abbildungen:

Definition 7.12. Sei V ein $\mathbb{R}$-Vektorraum.

(a) Eine Bilinearform $\Phi\colon V^2 \to \mathbb{R}$ heißt *positiv semidefinit*, wenn für jeden Vektor $v \in V$ gilt:
$$\Phi(v, v) \geq 0.$$

(b) Eine Bilinearform $\Phi\colon V^2 \to \mathbb{R}$ heißt *positiv definit*, wenn für jeden Vektor $v \in V \setminus \{0\}$ gilt:
$$\Phi(v, v) > 0.$$

Entsprechend sind die Begriffe *negativ semidefinit* und *negativ definit* definiert.
Eine Bilinearform, die weder positiv noch negativ semidefinit ist, heißt *indefinit*.

Bemerkung 7.13. Es gibt auch Möglichkeiten, die Definitheit einer Bilinearform von ihrer darstellenden Matrix abzulesen, das erfordert aber mehr Kenntnisse.

Definition 7.14. Sei V ein $\mathbb{R}$-Vektorraum.
Ein *Skalarprodukt* auf V ist eine positiv definite, symmetrische Bilinearform $\Phi\colon V \times V \to \mathbb{R}$.

Bemerkung 7.15. Achtung, nicht verwechseln mit der Skalarmultiplikation:
Die Skalarmultiplikation macht aus einem Skalar und einem Vektor einen Vektor, ein Skalarprodukt aus zwei Vektoren einen Skalar.

Bemerkung 7.16. Das wichtigste Beispiel für ein Skalarprodukt habt ihr bereits kennengelernt: Das *euklidische* oder auch *Standardskalarprodukt* des $\mathbb{R}^n$:

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \ldots + x_n y_n$$

Da die zugehörige Matrix offensichtlich die Einheitsmatrix ist, die symmetrisch ist, ist das Standardskalarprodukt symmetrisch. Für die positive Definitheit beachten wir, dass $\langle (x_1, \ldots, x_n), (x_1, \ldots, x_n) \rangle = x_1^2 + \ldots + x_n^2$ und eine Summe von Quadraten reeller Zahlen immer größer als Null ist, solange eine der Zahlen nicht Null ist.
Damit handelt es sich tatsächlich um ein Skalarprodukt nach unserer Definition.

8 Invertierbarkeit und Determinante

Wie wir bereits herausgefunden haben, bilden die $(n \times n)$-Matrizen über einem beliebigen Körper mit der Einheitsmatrix ein Monoid. Wie in jedem Monoid nennen wir eine Matrix A invertierbar, wenn es eine zweite Matrix, die wir dann A^{-1} nennen, gibt, sodass $AA^{-1} = E_n = A^{-1}A$. Nach Lemma 1.16 bilden dann die invertierbaren Matrizen eine Gruppe.

Definition 8.1. Sei K ein Körper.
Die Gruppe der invertierbaren $(n \times n)$-Matrizen notieren wir

$$GL_n(K) := \mathrm{Mat}_K(n \times n)^*$$

und nennen sie die *allgemeine lineare Gruppe der $(n \times n)$-Matrizen.*

Bemerkung 8.2. Im Prinzip kann auch das Produkt AB zweier nicht quadratischer Matrizen $A \in \mathrm{Mat}_K(n \times m)$ und $B \in \mathrm{Mat}_K(m \times n)$ die Einheitsmatrix E_n ergeben. Dann kann aber niemals auch $BA = E_m$ gelten, wie man durch Dimensionsargumente, die wir nicht behandelt haben, sehen kann. In jedem Fall werden wir uns damit im folgenden nicht beschäftigen.

Die Fragen, die uns nun beschäftigen, sind: Wie prüft man, ob eine gegebene $(n\times n)$-Matrix invertierbar ist? Und wie berechnet man dann ihr Inverses?
Eine erste Antwort liefert uns Satz 6.9:

Lemma 8.3. *Eine Matrix $A \in \mathrm{Mat}_K(n \times n)$ ist genau dann invertierbar, wenn die entsprechende lineare Abbildung $F_A \colon K^n \to K^n$ invertierbar ist, also eine Umkehrabbildung hat.*

Jetzt erinnern wir uns an Satz 5.9: Genau dann ist die lineare Abbildung

$$\begin{array}{rrcl} \Phi_{\mathcal{V}}\colon & K^n & \to & V \\ & (\lambda_1, \ldots, \lambda_n) & \mapsto & \lambda_1 v_1 + \ldots + \lambda_n v_n \end{array}$$

für die Familie $\mathcal{V} = (v_1, \ldots, v_n)$ ein Vektorraumisomorphismus, also invertierbar, wenn die Familie eine Basis von V bildet.
In unserem Fall haben wir $V = K^n$ und wollen die Abbildung F_A auf Invertierbarkeit überprüfen. Wir wissen aus Korollar 5.13, dass $F_A = \Phi_{\mathcal{V}}$ für die Familie $\mathcal{V} := (F_A(e_1), \ldots, F_A(e_n))$ der Bilder der Standardbasisvektoren. Wir haben also die folgende Aussage:

Lemma 8.4. *Eine Matrix A ist genau dann invertierbar, wenn die Bilder der Standardbasisvektoren unter der Abbildung F_A eine Basis von K^n bilden.*

Nach Bemerkung 6.10 sind die Bilder der Standardbasisvektoren aber genau die Spalten der Matrix. Deshalb können wir die Aussage auch etwas umformulieren:

Korollar 8.5. *Eine Matrix A ist genau dann invertierbar, wenn ihre Spaltenvektoren eine Basis von K^n bilden.*

Diese Erkenntnis ist der Schlüssel zu zwei Verfahren, um die Invertierbarkeit einer Matrix zu prüfen. Die erste Idee ist es, eine Kenngröße zu entwickeln, die man von einer Matrix ablesen – das heißt aus ihren Einträgen ausrechnen – kann und Werte aus einer bestimmten Menge, dem „Annahmebereich“, immer dann annimmt, wenn die Matrix invertierbar ist,

und Werte aus einer anderen Menge, dem „Ablehnungsbereich“, immer dann, wenn sie es nicht ist. Da diese Kenngröße bestimmt, also determiniert, wann die Matrix invertierbar ist, nennt man sie die *Determinante*.
Um zu sehen, wie wir diese Determinante sinnvoll definieren, können uns gleich einmal ein paar Dinge klarmachen:

- Da die Entscheidung, ob die Matrix invertierbar ist, aus den Spalten gelesen werden kann, sollte die Determinante eine Abbildung aus dem Raum $\underbrace{K^n \times \ldots \times K^n}_{n\text{-mal}}$ sein.

- Der einfachste Fall, dass die Spaltenvektoren keine Basis bilden, ist, wenn einer von ihnen der Nullvektor ist. Wir möchten also, dass die Determinante dann im Ablehnungsbereich landet.
  Genauer wissen wir aus Korollar 5.16: Wenn $(v_1, \ldots, v_n)$ eine Basis ist, dann ist $(v_1, \ldots, v_{j-1}, \lambda v_j, v_{j-1}, \ldots, v_n)$ genau dann eine Basis, wenn $\lambda \neq 0$.
  Das bringt uns auf die folgende Idee: Wir könnten als Ablehnungsbereich die Menge $\{0_K\}$ und als Annahmebereich $K \setminus \{0_K\}$ wählen und eine Abbildung $\det\colon K^n \times \ldots \times K^n \to K$ mit der folgenden Eigenschaft suchen:

  $$\forall j \in \{1, \ldots, n\}, \lambda \in K\colon \quad \det(v_1, \ldots, \lambda v_j, \ldots, v_n) = \lambda \det(v_1, \ldots, v_j, \ldots v_n).$$

  Wenn dann nämlich $(v_1, \ldots, v_j, \ldots v_n)$ eine Basis ist, also $\det(v_1, \ldots, v_j, \ldots v_n) \neq 0$, dann bleibt $\det(v_1, \ldots, \lambda v_j, \ldots, v_n) \neq 0$ genau dann, wenn $\lambda \neq 0$, also genau dann, wenn $(v_1, \ldots, \lambda v_j, \ldots, v_n)$ eine Basis ist.

- Wenn wir eine Basis $(v_1, \ldots, v_n)$ von K^n haben und einen der Vektoren v_j durch einen anderen Vektor $w = \lambda_1 v_1 + \ldots + \lambda_n v_n$ ersetzen, dann geschieht das Folgende:
  Wenn $\lambda_j \neq 0$, so ist nach dem Austauschlemma 5.15 auch $(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n)$ eine Basis.
  Wenn dagegen $\lambda_j \neq 0$, sodass wir w als Linearkombination der anderen Vektoren $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$ schreiben können, so ist nach Lemma 5.5 die Familie $(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n)$ linear abhängig, also keine Basis.
  Daraus folgt: Wenn die Familie $(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n)$ linear unabhängig ist – sich also zu einer Basis $(v_1, \ldots, v_n)$ ergänzen lässt – und w ein beliebiger Vektor asus K^n ist, so entscheidet sich die Frage, ob $(v_1, \ldots, w, \ldots, v_n)$ einen Basis ist, an dem Koeffizienten λ_j in der Basisdarstellung von w. Dieser Koeffizient ist aber linear: Wenn λ_j der Koeffizient in der Basisdarstellung von w ist und λ'_j der in der Basisdarstellung von w', dann ist $\lambda_j + \lambda'_j$ der in der Basisdarstellung von $w + w'$.
  Damit wünschen wir für unsere Determinantenabbildung also weiter:

  $$\begin{aligned}&\forall j \in \{1, \ldots, n\}, \forall w, w' \in K^n\colon\\ &\qquad \det(v_1, \ldots, v_{j-1}, w + w', v_{j+1}, \ldots, v_n)\\ &\qquad = \det(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n) + \det(v_1, \ldots, v_{j-1}, w', v_{j+1}, \ldots, v_n).\end{aligned}$$

  Für den Fall, dass $(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n)$ linear abhängig ist, gibt es keine Möglichkeit mehr, die Familie zu einer Basis zu ergänzen, sodass alle Terme einfach 0 sind und die Gleichheit also auch dann eine sinnvolle Forderung ist.
  Zusammen mit dem zweiten Punkt erhalten wir also, dass wir als Determinante gerne eine multilineare Abbildung hätten.

- Der zweiteinfachste Fall, dass eine Familie $(v_1, \dots, v_n)$ von Vektoren keine Basis ist, ist, wenn zwei von ihnen gleich sind. Wir suchen also eine Abbildung, für die weiter gilt:
$$\forall i \in \{1, \dots, n\}: \quad \det(v_1, \dots, v_i, \dots, v_i, \dots, v_n) = 0.$$
Die Determinante sollte also alternierend sein.

- Die Standardbasis $\{e_1, \dots, e_n\}$ ist eine Basis. Wir möchten also
$$\det(e_1, \dots, e_n) \neq 0.$$

Insgesamt suchen wir also als Determinante eine Abbildung $\det\colon K^n \times \ldots \times K^n \to K$, die multilinear und alternierend ist und der Standardbasis nicht den Wert 0 zuweist.
Jetzt stellen sich zwei Fragen:
Erstens: Gibt es eine solche Abbildung überhaupt?
Zweitens: Erfüllt jede solche Abbildung tatsächlich unsere Forderung, dass sie genau dann 0 wird, wenn die Vektoren keine Basis bilden? Die Frage erscheint vielleicht seltsam, weil wir uns ja so viel Mühe gegeben haben, genau diese Eigenschaften herauszulesen, aber es könnte ja zum Beispiel sein, dass wir noch eine wichtige Eigenschaft übersehen haben, die die Abbildung auch bräuchte.
Die erste Frage beantworten wir, indem wir eine solche Abbildung konkret angeben. Dafür brauchen wir ein wenig Rückgriff auf die Gruppentheorie:

Lemma 8.6. *Sei $n \in \mathbb{N} \setminus \{0\}$. Dann bilden die Permutationen von n Elementen, also die bijektiven Abbildungen $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ mit der Hintereinanderausführung eine Gruppe.*

Beweis. (G1) Die Verknüpfung $\circ$ ist assoziativ, denn für drei Abbildungen ρ, σ und τ und ein beliebiges $j \in \{1, \dots, n\}$ gilt:
$$((\rho \circ \sigma) \circ \tau)(j) = (\rho \circ \sigma)(\tau(j)) = \rho(\sigma(\tau(j))) = \rho((\sigma \circ \tau)(j)) = (\rho \circ (\sigma \circ \tau))(j).$$

(G2) Die Identität $\mathrm{id}_{\{1,\dots,n\}}$ ist offensichtlich neutrales Element.

(G3) Da alle Permutationen σ nach Voraussetzung bijektiv sind, hat jede eine inverse Abbildung σ^{-1} mit $\sigma^{-1} \circ \sigma = \mathrm{id}_{\{1,\dots,n\}}$ und diese ist natürlich selbst bijektiv.

□

Bemerkung 8.7. Diese Gruppe wird die *symmetrische Gruppe S_n* genannt.

Definition 8.8. Sei $\sigma \in S_n$ eine Permutation.

(a) Ein *Fehlstand* von σ ist ein Paar $(i, j) \in \{1, \dots n\}^2$ mit $i < j$, aber $\sigma(i) > \sigma(j)$.

(b) Das *Signum* von σ ist
$$\mathrm{sgn}(\sigma) := (-1)^{|\{(i,j) \in \{1,\dots,n\}^2 | i<j, \sigma(i)>\sigma(j)\}|},$$
also 1, wenn die Zahl der Fehlstände von σ gerade ist, und -1, wenn sie ungerade ist.

Definition 8.9. Eine Permutation τ, die zwei Zahlen vertauscht und alle anderen festhält, nennen wir *Transposition.*

Bemerkung 8.10. Die Fehlstände der Transposition, die i und j mit $i < j$ vertauscht, sind genau das Paar (i, j) und alle der Form (i, k) oder (k, j) mit $i < k < j$. Damit ist ihr Signum $(-1)^{1+2(j-i-1)} = -1$.

Lemma 8.11. sgn *ist ein Gruppenhomomorphismus von* S_n *nach* $(\mathbb{R}, \cdot, 1)$*, also*

$$\forall \sigma, \tau \in S_n\colon \quad \operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau) \cdot \operatorname{sgn}(\sigma).$$

Beweis. Wir definieren zunächst ein *geordnetes Paar* von $\{1, \dots, n\}$ als ein Paar $(i, j) \in \{1, \dots, n\}^2$ mit $i < j$. Die Menge aller geordneten Paare bezeichnen von $\{1, \dots, n\}$ bezeichnen wir mit P_n.
Wir stellen zunächst eine Vorüberlegung zu den geordneten Paaren von $\{1, \dots, n\}$ an: Für jedes $(i, j) \in P_n$ und jede Permutation σ gilt, dass entweder $(\sigma(i), \sigma(j))$ oder $(\sigma(j), \sigma(i))$ wieder ein geordnetes Paar ist, das wir $\tilde{\sigma}(i, j)$ nennen. So erhalten wir eine Abbildung $\tilde{\sigma}$ von der Menge P_n auf sich selbst.
Diese Abbildung ist offensichtlich surjektiv, denn für jedes geordnete Paar $(k, l) \in \{1, \dots, n\}^2$ gibt es ein $\sigma^{-1}(k)$ und ein $\sigma^{-1}(l)$. Ist $\sigma^{-1}(k) < \sigma^{-1}(l)$, so ist $(\sigma^{-1}(k), \sigma^{-1}(l))$ ein geordnetes Paar mit $\tilde{\sigma}(\sigma^{-1}(k), \sigma^{-1}(l)) = (k, l)$. Ist dagegen $\sigma^{-1}(k) > \sigma^{-1}(l)$, dann ist $(\sigma^{-1}(l), \sigma^{-1}(k))$ ein geordnetes Paar mit $\tilde{\sigma}(\sigma^{-1}(l), \sigma^{-1}(k)) = (k, l)$. Da P_n endlich ist, ist die Abbildung $\tilde{\sigma}$ sogar bijektiv. Wenn wir die Paare $\tilde{\sigma}(i, j)$ für $i < j$ der Reihe nach ansehen, erhalten wir also alle geordneten Paare genau einmal.
Weiter definieren wir für jede Permutation σ die Funktion

$$\begin{array}{rccl} f_\sigma\colon & P_n & \to & \{0, 1\} \\ & (i, j) & \mapsto & \begin{cases} 1 & \text{für } (i, j) \text{ Fehlstand} \\ 0 & \text{sonst,} \end{cases} \end{array}$$

die jür jedes geordnete Paar prüft, ob es ein Fehlstand ist oder nicht. Klarerweise gilt

$$\operatorname{sgn}(\sigma) = \prod_{(i,j) \in P_n} (-1)^{f_\sigma(i,j)}.$$

Jetzt prüfen wir für jedes geordnete Paar $(i, j) \in \{1, \dots, n\}^2$, ob es ein Fehlstand von $\tau \circ \sigma$ ist oder nicht. Es gibt die folgenden Fälle:
Ist (i, j) kein Fehlstand von σ, so gilt $\sigma(i) < \sigma(j)$. In diesem Fall ist $\tilde{\sigma}(i, j) = (\sigma(i), \sigma(j))$. Genau dann, wenn $\tilde{\sigma}(i, j)$ ein Fehlstand von τ ist, ist $(\tau \circ \sigma)(i) < (\tau \circ \sigma)(j)$ und (i, j) ein Fehlstand von $\tau \circ \sigma$. Also gilt in diesem Fall

$$f_{\tau \circ \sigma}(i, j) = f_\tau(\tilde{\sigma}(i, j)) = f_\tau(\tilde{\sigma}(i, j)) + \underbrace{f_\sigma(i, j)}_{=0}.$$

Ist dagegen (i, j) ein Fehlstand von σ, so ist $\tilde{\sigma}(i, j) = (\sigma(j), \sigma(i))$. Ist nun dieses geordnete Paar kein Fehlstand von τ, so ist $\tau(\sigma(j)) < \tau(\sigma(i)$ und (i, j) ein Fehlstand von $\tau \circ \sigma$. Auch in diesem Fall gilt also

$$f_{\tau \circ \sigma}(i, j) = \underbrace{f_\tau(\tilde{\sigma}(i, j))}_{=0} + \underbrace{f_\sigma(i, j)}_{=1}.$$

Ist schließlich (i, j) ein Fehlstand von σ und $\tilde{\sigma}(i, j)$ ein Fehlstand von τ, so ist $\tau(\sigma(j)) > \tau(\sigma(i))$ und (i, j) kein Fehlstand von $\tau \circ \sigma$. In diesem Fall gilt also

$$f_{\tau \circ \sigma}(i, j) = \underbrace{f_\tau(\tilde{\sigma}(i, j))}_{=1} + \underbrace{f_\sigma(i, j)}_{=1} - 2.$$

In allen drei Fällen gilt aber

$$(-1)^{f_{\tau\circ\sigma}(i,j)} = (-1)^{f_\tau(\tilde{\sigma}(i,j))+f_\sigma(i,j)} = (-1)^{f_\tau(\tilde{\sigma}(i,j))} \cdot (-1)^{f_\sigma(i,j)}.$$

Insgesamt erhalten wir

$$\begin{aligned}
\operatorname{sgn}(\tau\circ\sigma) \quad &= \quad \prod_{(i,j)\in P_n} (-1)^{f_{\tau\circ\sigma}(i,j)} \\
&= \quad \prod_{(i,j)\in P_n} \left((-1)^{f_\tau(\tilde{\sigma}(i,j))} \cdot (-1)^{f_\sigma(i,j)}\right) \\
&= \quad \prod_{(i,j)\in P_n} (-1)^{f_\tau(\tilde{\sigma}(i,j))} \cdot \prod_{(i,j)\in P_n} (-1)^{f_\sigma(i,j)} \\
&\overset{\text{Vorüberlegung}}{=} \quad \prod_{(i,j)\in P_n} (-1)^{f_\tau(i,j)} \cdot \prod_{(i,j)\in P_n} (-1)^{f_\sigma(i,j)} \\
&= \quad \operatorname{sgn}(\tau)\cdot\operatorname{sgn}(\sigma).
\end{aligned}$$

□

Korollar 8.12. *Für jede Transposition τ ist die Abbildung*

$$\begin{aligned}
\{\sigma\in S_n \mid \operatorname{sgn}(\sigma)=1\} \quad &\to \quad \{\rho\in S_n \mid \operatorname{sgn}(\rho)=-1\} \\
\sigma \quad &\mapsto \quad \sigma\circ\tau
\end{aligned}$$

eine Bijektion.

Beweis. Wegen $\operatorname{sgn}(\sigma\circ\tau) = \underbrace{\operatorname{sgn}(\sigma)}_{1}\cdot\underbrace{\operatorname{sgn}(\tau)}_{-1}$ ist die Abbildung wohldefiniert. Aus dem gleichen Grund ist $\operatorname{sgn}(\tau^{-1}) = \operatorname{sgn}(\tau)^{-1} = (-1)^{-1} = -1$ (tatsächlich ist $\tau^{-1}=\tau$), sodass auch die Abbildung

$$\begin{aligned}
\{\rho\in S_n \mid \operatorname{sgn}(\rho)=-1\} \quad &\to \quad \{\sigma\in S_n \mid \operatorname{sgn}(\sigma)=1\} \\
\rho \quad &\mapsto \quad \rho\circ\tau^{-1}
\end{aligned}$$

wohldefiniert ist. Diese Abbildung ist aber offensichtlich die Umkehrabbildung, denn $\sigma\circ\tau\circ\tau^{-1} = \sigma\circ\mathrm{id}_{\{1,\ldots,n\}} = \sigma$ für alle σ und andersherum $\rho\circ\tau^{-1}\circ\tau = \rho\circ\mathrm{id}_{\{1,\ldots,n\}} = \rho$ für alle ρ. □

Jetzt haben wir alle Begriffe und Hilfsmittel zur Hand, um die Determinantenfunktion zu definieren:

Satz 8.13. *Sei K ein Körper und $n\in\mathbb{N}\setminus\{0\}$. Dann ist die Abbildung*

$$\begin{aligned}
\det\colon \quad K^n\times\ldots\times K^n \quad &\to \quad K \\
\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1}\end{pmatrix},\ldots,\begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn}\end{pmatrix}\right) \quad &\mapsto \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1}\cdots a_{\sigma(n)n}
\end{aligned}$$

multilinear, alternierend und erfüllt $\det(e_1,\ldots,e_n)=1$.

Beweis. • det ist multilinear:
In jedem Summanden wird aus jedem Vektor genau ein Eintrag verwendet, nämlich aus dem j-ten Vektor an der j-ten Stelle.
Damit gilt

$$\begin{aligned}
&\det(\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\dots,\begin{pmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{pmatrix}+\begin{pmatrix}b_{1j}\\ \vdots\\ b_{nj}\end{pmatrix},\dots,\begin{pmatrix}a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix})\\
&= \det(\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\dots,\begin{pmatrix}a_{1j}+b_{ij}\\ \vdots\\ a_{nj}+b_{nj}\end{pmatrix},\dots,\begin{pmatrix}a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix})\\
&= \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{\sigma(1)1}\cdots(a_{\sigma(j)j}+b_{\sigma(j)j})\cdots a_{\sigma(n)n}\\
&= \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(j)j}\cdots a_{\sigma(n)n}+\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{\sigma(1)1}\cdots b_{\sigma(j)j}\cdots a_{\sigma(n)n}\\
&= \det(\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\dots,\begin{pmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{pmatrix},\dots,\begin{pmatrix}a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix})+\det(\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\dots,\begin{pmatrix}b_{1j}\\ \vdots\\ b_{nj}\end{pmatrix},\dots,\begin{pmatrix}a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix}).
\end{aligned}$$

Die Homogenität sieht man ebenso ein.

- det ist alternierend:
Wir nutzen aus, dass wir S_n disjunkt in die beiden Mengen $\{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$ und $\{\rho \in S_n \mid \operatorname{sgn}(\rho) = -1\} = \{\sigma \circ \tau \mid \operatorname{sgn}(\sigma) = 1\}$ für eine beliebige Transposition τ zerlegen können.
Damit gilt nämlich:

$$\begin{aligned}
\det(a_1,\dots,a_n) &= \sum_{\substack{\sigma\in S_n\\ \operatorname{sgn}(\sigma)=1}}\operatorname{sgn}(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n}+\sum_{\substack{\sigma\in S_n\\ \operatorname{sgn}(\sigma)=1}}\operatorname{sgn}(\sigma\circ\tau)a_{(\sigma\circ\tau)(1)1}\cdots a_{(\sigma\circ\tau)(n)n}\\
&= \sum_{\substack{\sigma\in S_n\\ \operatorname{sgn}(\sigma)=1}}\left(a_{\sigma(1)1}\cdots a_{\sigma(n)n}-a_{(\sigma\circ\tau)(1)1}\cdots a_{(\sigma\circ\tau)(n)n}\right)
\end{aligned}$$

Gilt nun $\begin{pmatrix}a_{1i}\\ \vdots\\ a_{ni}\end{pmatrix} = \begin{pmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{pmatrix}$, so erhalten wir für die Transposition τ, die i und j vertauscht, und jedes $\sigma \in S_n$ mit $\operatorname{sgn}(\sigma) = 1$: $a_{(\sigma\circ\tau)(i)i} = a_{\sigma(j)i} = a_{\sigma(j)j}$, $a_{(\sigma\circ\tau)(j)j} = a_{\sigma(i)j} = a_{\sigma(i)i}$ und $a_{(\sigma\circ\tau)(k)k} = a_{\sigma(k)k}$ für alle $k \notin \{i, j\}$, zusammen also

$$a_{(\sigma\circ\tau)(1)1}\cdots a_{(\sigma\circ\tau)(n)n} = a_{\sigma(1)1}\cdots a_{\sigma(n)n}.$$

Damit sind alle Summanden 0.

- $\det(e_1,\dots,e_n) = 1$:
Ein Eintrag a_{ij} ist in diesem Fall nur dann 1 und nicht 0, wenn $i = j$. Das Produkt $a_{\sigma(1)1}\cdots a_{\sigma(n)n}$ ist also nur dann 1 und nicht 0, wenn $\sigma(1) = 1,\dots,\sigma(n) = n$, also $\sigma = \mathrm{id}_{\{}1,\dots,n\}$. In diesem Fall ist $\operatorname{sgn}(\sigma) = 1$, woraus die Behauptung folgt.

□

Bemerkung 8.14. Da man jede Matrix offensichtlich in ihre Spalten zerlegen kann, kann man mit derselben Formel auch eine Abbildung

$$\begin{aligned} \mathrm{Mat}_K(n \times n) &\to K \\ \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} &\mapsto \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \end{aligned}$$

definieren, die wir auch mit det bezeichnen. Diese Abbildung ist aber nicht mehr multilinear, da sie nur noch ein einziges Argument aufnehmen kann, und damit auch nicht alternierend.
Linear ist sie für $n > 1$ auch nicht, denn

$$\det(\lambda \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}) = \det(\begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{n1} & \dots & \lambda a_{nn} \end{pmatrix}) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)(\lambda a_{\sigma(1)1}) \cdots (\lambda a_{\sigma(n)n})$$

$$= \lambda^n \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \lambda^n \det(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}).$$

Trotzdem wird man in der Literatur meist diese Definition der Determinantenfunktion finden (auch wir werden ihren Wert später kennen lernen). Um die Eigenschaften der Determinantenfunktion auf $K^n \times \ldots \times K^n$ auszudrücken, sagt man dann „die Determinantenfunktion ist multilinear und alternierend in den Spalten“.

Nachdem wir nun eine Abbildung gefunden haben, die multilinear und alternierend ist und der Standardbasis nicht den Wert 0 zuordnet, wollen wir uns überlegen, ob diese Abbildung tatsächlich zuverlässig zwischen Basen und Nichtbasen, also invertierbaren und nicht invertierbaren Matrizen, unterscheiden kann.

Lemma 8.15. *Sei K ein Körper, $n \in \mathbb{N} \setminus \{0\}$ und V ein n-dimensionaler K-Vektorraum. Bilden die Vektoren $v_1, \ldots, v_n \in V$ keine Basis von V, so gilt* $\det(v_1, \ldots, v_n) = 0$.

Beweis. Da V n-dimensional ist, kann die Familie $(v_1, \ldots, v_n)$ nur dann keine Basis sein, wenn sie linear abhängig ist. Dann gibt es aber einen Vektor v_i, der sich als Linearkombination $\sum_{j \neq i} \lambda_j v_j$ der anderen schreiben lässt. Dann gilt aber

$$\det(v_1, \ldots, v_i, \ldots, v_n) = \det(v_1, \ldots, \sum_{j \neq i} \lambda_j v_j, \ldots, v_n) = \sum_{j \neq i} \lambda_j \det(v_1, \ldots, v_j, \ldots, v_n).$$

Weil det alternierend ist und in jeder der Determinanten am Ende der Zeile ein Eintrag doppelt vorkommt, ist jede davon 0. □

An sich haben wir auch das Handwerkszeug, um andersherum zu beweisen, dass die Determinante jeder Basis ungleich 0 ist. Das wäre aber zum jetzigen Zeitpunkt umständlich. Ebenso möglich, aber noch einmal umständlicher, wäre es, mit der Determinante allein die Inverse einer Matrix zu berechnen, wenn es sie gibt.
Stattdessen begeben wir uns auf einen Weg, auf dem wir beides – zuzüglich einem einfachen Berechnungsalgorithmus für die Determinante einer Matrix – bequem erreichen.
Wir beginnen damit, dass wir uns drei (zum Teil schon bekannte) Dinge klarmachen:

Bemerkung 8.16. Sei $(v_1, \dots, v_n)$ eine Basis von V.

(a) Für alle $i, j \in \{1, \dots, n\}$ mit $i \neq j$ ist auch das Tupel $(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n)$, bei dem der i-te und j-te Vektor getauscht sind, eine Basis von V. Wir nennen diese Operation, aus einem Tupel von Vektoren ein neues zu machen, *elementare Spaltenumformung des Typs I.*

(b) Ist $\lambda \neq 0$, dann ist für alle $i \in \{1, \dots, n\}$ auch das Tupel $(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_n)$, bei dem der i-te Vektor um den Faktor λ gestreckt ist, eine Basis von V. Diese Operation nennen wir *elementare Spaltenumformung des Typs II.*

(c) Für alle $i, j \in \{1, \dots, n\}$ mit $i \neq j$ und alle $\lambda \in K$ ist auch das Tupel $(v_1, \dots, v_{j-1}, v_j + \lambda v_i, v_{j+1}, \dots, v_n)$, bei dem auf den j-ten Vektor das λ-Fache des i-ten Vektors addiert ist, eine Basis von V. Diese Operation nennen wir *elementare Spaltenumformung des Typs III.*

Wir werden jetzt ein Verfahren konstruieren, mit dem man jede Basis allein durch elementare Spaltenumformungen dieser drei Typen erst in eine Basis der Form

$$\begin{pmatrix} b_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{pmatrix}$$

mit $b_{11} \neq 0, \dots, b_{nn} \neq 0$ und dann in die Standardbasis umformen kann.
Wir gehen dabei so vor: Angenommen, wir haben es bereits geschafft, die Basis durch elementare Spaltenumformungen in eine Basis der Gestalt

$$\underbrace{\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{a_1}, \dots, \underbrace{\begin{pmatrix} a_{1m} \\ \vdots \\ a_{mm} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{a_m}, \underbrace{\begin{pmatrix} b_{1(m+1)} \\ \vdots \\ b_{m(m+1)} \\ b_{(m+1)(m+1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{b_{m+1}}, \dots, \underbrace{\begin{pmatrix} b_{1n} \\ \\ \vdots \\ \\ \vdots \\ \\ b_{nn} \end{pmatrix}}_{b_n}$$

mit $b_{(m+1)(m+1)} \neq 0, \dots, b_{nn} \neq 0$ für ein $m \in \{1, \dots, n\}$ zu bringen. Dann können wir diese mit weiteren elementaren Spaltenumformungen in eine Basis der entsprechenden Gestalt für ein um 1 kleineres m umformen:
Zunächst einmal muss es unter den ersten m Basisvektoren mindestens einen geben, dessen m-ter Eintrag ungleich 0 ist. Andernfalls hätten wir nämlich m Vektoren in dem $m-1$-dimensionalen Untervektorraum $\{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_m = \dots = x_n = 0\} = \operatorname{span}_K(e_1, \dots, e_{m-1})$, die nicht linear unabhängig sein könnten – im Widerspruch zur Annahme, dass es sich um eine Basis handelt.
Wenn dieser Vektor bereits der m-te ist, nenne diesen b_m, definiere also $b_{1j} := a_{1j}$ für alle $j \in \{1, \dots, n\}$; ist es der i-te mit $i \neq m$, so tausche den i-ten und den m-ten Vektor mittels der entsprechenden elementaren Spaltenumformung des Typs I und nenne den ehemaligen

Vektor a_i um in b_m und den ehemaligen Vektor a_m um in a_i. In beiden Fällen haben wir $b_{mm} \neq 0$ nach Konstruktion.
Jetzt addiere der Reihe nach auf alle Vektoren a_j für $j \in \{1, \ldots, m-1\}$ das $-\frac{a_{mj}}{b_{mm}}$-Fache des Vektors b_m mittels der entsprechenden elementaren Spaltenumformungen des Typs III. Wenn wir die entstehenden Vektoren wieder a_j nennen, gilt nach Konstruktion offenbar $a_{mj} = 0$ und wir haben eine Basis des gewünschten Typs.
Wenn wir dieses Verfahren immer weiter durchführen, erhalten wir irgendwann eine Basis der Gestalt

$$\begin{pmatrix} b_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{pmatrix}$$

mit $b_{11} \neq 0$, …, $b_{nn} \neq 0$. Nun wandeln wir diese Basis in die Standardbasis um, indem wir so vorgehen:
Angenommen, wir haben bereits die Basis $e_1, \ldots, e_m, b_{m+1}, \ldots, b_n$ für ein $m \in \{1, \ldots, n-1\}$ erreicht. Dann addiere auf b_{m+1} der Reihe nach das $-b_{j(m+1)}$-Fache des Vektors e_j für alle $j \in \{1, \ldots, m\}$ mittels der entsprechenden elementaren Spaltenumformung des Typs III, um einen Vektor zu erhalten, dessen einziger von 0 verschiedener Eintrag b_{mm} ist. Diesen Vektor streckt man schließlich mittels der entsprechenden elemantaren Spaltenumformung des Typs II um den Faktor $\frac{1}{b_{mm}}$, um den m-ten Standardbasisvektor e_m zu erhalten.
Insgesamt haben wir also ein Verfahren, den *Gauß-Algorithmus für Spalten*, entwickelt, mit dem man eine beliebige Basis durch elementare Spaltenumformungen in die Standardbasis umformen kann.
Damit können wir erst einmal das versprochene Resultat zur Determinante beweisen. Wir verfolgen dazu, was mit der Determinante geschieht, wenn wir elementare Spaltenumformungen durchführen:

Bemerkung 8.17. (a) Jede elementare Spaltenumformung des Typs I vertauscht zwei Spalten. Wegen der Antisymmetrie von det gilt

$$\det(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m) = -\det(v_1, \ldots, v_n).$$

(b) Jede elementare Spaltenumformung des Typs II ersetzt einen Vektor durch das λ-Fache. Wegen der Multilinearität von det gilt

$$\det(v_1, \ldots, \lambda v_i, \ldots, v_n) = \lambda \det(v_1, \ldots, v_n).$$

(c) Für jede elementare Spaltenumformung des Typs III gilt, da det multilinear und alternierend ist,

$$\det(v_1, \ldots, v_j + \lambda v_i, \ldots, v_n) = \det(v_1, \ldots, v_n) + \lambda \det(v_1, \ldots, v_i, \ldots, v_n) = \det(v_1, \ldots, v_n).$$

Lemma 8.18. *Sei K ein Körper und $n \in \mathbb{N} \setminus \{0\}$.*
Für jede Basis $(v_1, \ldots, v_n)$ von K^n gilt $\det(v_1, \ldots, v_n) \neq 0$.

Beweis. Beweis durch Widerspruch: Angenommen, es gäbe eine Basis $(v_1, \ldots, v_n)$ von K^n mit $\det(v_1, \ldots, v_n) = 0$. Mittels des Gaußalgorithmus' kann diese Basis nun in die Standardbasis überführt werden.
Durch elementare Spaltenumformungen wird die Determinante aber lediglich mit Faktoren -1 oder λ multipliziert. In jedem Fall heißt das: Wenn für die ursprüngliche Basis $(v_1, \ldots, v_n)$ die Determinante 0 wäre, dann bliebe sie dies den ganzen Rest des Verfahrens. Am Schluss kommt aber die Standardbasis mit Determinante 1 heraus – Widerspruch. □

Damit ist aber der Wert des Gauß-Algorithmus' noch lange nicht ausgeschöpft: Betrachten wir ein Tupel $\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}\right)$ nämlich wieder als Matrix $A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, so entspricht

- der elementaren Spaltenumformung des Typs I, die den i-ten und j-ten Vektor vertauscht, die Multiplikation der Matrix A mit der Matrix

$$P_{ij} := \begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 &&&& 1 &&& \\ &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ &&& 1 &&&& 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix},$$

  die Einsen nur in den kk-ten Einträgen für $k \in \{1, \dots, n\} \setminus \{i, j\}$, dem ij-ten und dem ji-ten Eintrag hat und sonst überall Nullen,

- der elementare Spaltenumformung des Typs II, die den i-ten Vektor mit dem Faktor λ multipliziert, der Multiplikation der Matrix A mit der Matrix

$$M_i(\lambda) := \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix},$$

  die außer einem λ in ii-ten Eintrag auf der Hauptdiagonalen Einsen und sonst überall Nullen hat,

- und der elementaren Spaltenumformung des Typs III, die auf den j-ten Vektor das λ-Fache des i-ten Vektors addiert, der Multiplikation der Matrix A mit der Matrix

$$Q_{ij}(\lambda) := \begin{pmatrix} 1 &&&& \\ & \ddots & \lambda && \\ && \ddots && \\ &&& \ddots & \\ &&&& 1 \end{pmatrix},$$

  die in der i-ten Zeile und j-ten Spalte ein λ, auf der Hauptdiagonalen Einsen und sonst nur Nullen hat.

Das bedeutet aber, dass wir mit dem Gaußalgorithmus im Grunde den folgenden Satz bewiesen haben:

Satz 8.19. *Sei K ein Körper und $n \in \mathbb{N} \setminus \{0\}$.*
Genau dann ist A eine invertierbare Matrix, wenn es Matrizen

$$\begin{aligned} S_1, \ldots, S_N \quad \in \quad & \{P_{ij} \mid i,j \in \{1,\ldots,n\}, i \neq j\} \\ & \cup \{M_i(\lambda) \mid i \in \{1,\ldots,n\}, \lambda \in K \setminus \{0\}\} \\ & \cup \{Q_{ij}(\lambda) \mid i,j \in \{1,\ldots,n\}, i \neq j, \lambda \in K\} \end{aligned}$$

gibt mit

$$AS_1 \cdots S_N = E_n.$$

In diesem Fall ist $A^{-1} = S_1 \cdots S_N$.

Beweis. „$\Rightarrow$“:
Das ist einfach die Umformulierung des Gaußalgorithmus' in die Sprache der Matrizen.
„$\Leftarrow$“:
Wenn $AS_1 \cdots S_N = E_n$, dann ist per Definition A invertierbar mit Inverser $S_1 \cdots S_N$. □

Aus dem Satz erhalten wir auch gleich eine Konstruktionsvorschrift für die Inverse einer Matrix A: Es gilt ja $S_1 \cdots S_N = E_n S_1 \cdots S_N$, deshalb erhält man die Inverse, indem man alle elementaren Spaltenoperationen, die man braucht, um aus A die Einheitsmatrix zu machen, in derselben Reihenfolge auf die Einheitsmatrix anwendet.
Und noch eine Erkenntnis können wir mit dem Gaußalgorithmus gewinnen: Schreiben wir die Rechenregeln von Bemerkung 8.17 in Matrixschreibweise noch einmal aus, so erhalten wir die Aussagen

$$\det(AP_{ij}) = -\det(A) \quad \det(AM_i(\lambda)) = \lambda \det(A) \quad \det(AQ_{ij}(\lambda)) = \det(A)$$

für alle Matrizen $A \in \mathrm{Mat}_K(n \times n)$.
Wenn wir nun für A speziell die Einheitsmatrix E_n einsetzen, dann erhalten wir

$$\det(P_{ij}) = -1 \quad \det(M_i(\lambda)) = \lambda \quad \det(Q_{ij}(\lambda) = 1.$$

Dadurch können wir die Aussagen aber noch einmal umformulieren:

$$\det(AP_{ij}) = \det(A)\det(P_{ij})$$

$$\det(AM_i(\lambda)) = \det(A)\det(M_i(\lambda))$$

$$\det(AQ_{ij}(\lambda)) = \det(A)\det(Q_{ij}(\lambda)).$$

Dieses Resultat können wir jetzt verallgemeinern:

Satz 8.20 (Multiplikationsregel für Determinanten)**.** *Sei K ein Körper, $n \in \mathbb{N} \setminus \{0\}$.*
Für alle $A, B \in \mathrm{Mat}_K(n \times n)$ gilt:

$$\det(BA) = \det(B)\det(A).$$

Die Determinante ist also ein Halbgruppenhomomorphismus von $(\mathrm{Mat}_K(n \times n), \cdot)$ nach $(K, \cdot)$.

Beweis. Wir unterscheiden zwei Fälle:
A nicht invertierbar:
Dann ist auch BA nicht invertierbar, denn sonst würde für das Inverse $(BA)^{-1}$ gelten: $((BA)^{-1}B)A = (BA)^{-1}(BA) = E_n$, also wäre A doch invertierbar mit Inversem $(BA)^{-1}B$. Damit lautet die Formel aber $0 = \det(B) \cdot 0$ und ist offensichtlich korrekt.
A invertierbar:
Dann existieren nach Satz 8.19 Matrizen

$$\begin{aligned} S_1, \dots, S_N \in \; & \{P_{ij} \mid i,j \in \{1,\dots,n\}, i \neq j\} \\ & \cup \{M_i(\lambda) \mid i \in \{1,\dots,n\}, \lambda \in K \setminus \{0\}\} \\ & \cup \{Q_{ji}(\lambda) \mid i,j \in \{1,\dots,n\}, i \neq j, \lambda \in K\} \end{aligned}$$

mit $E_n = AS_1 \cdots S_N$. Es folgt also

$$1 = \det(E_n) = \det(AS_1 \cdots S_N) = \det(A)\det(S_1)\cdots\det(S_N).$$

Mit denselben Matrizen gilt aber auch $B = BE_n = BAS_1 \cdots S_N$, sodass wir $\det(B) = \det(BA)\det(S_1)\cdots\det(S_N)$ erhalten. Multiplizieren wir beide Seiten mit $\det(A)$ und benutzen rechts die Kommutativität der Multiplikation in K, erhalten wir

$$\det(B)\det(A) = \det(BA)\underbrace{\det(S_1)\cdots\det(S_N)\det(A)}_{=1} = \det(BA).$$

□

Bemerkung 8.21. In der Literatur findet man den Gaußalgorithmus meist mit Zeilenumformungen statt Spaltenumformungen formuliert. In dieser Form entsteht er in natürlicher Weise als Verfahren, um lineare Gleichungssysteme zu lösen.
Für die Berechnung der Inversen nach Satz 8.19 eignen sich beide Verfahren gleich gut. Für den Zusammenhang mit der Determinante ist allerdings die hier definierte Variante ungleich nützlicher, deshalb habe ich mich für diese entschieden.