Inkrementelles Skript **Version 29. September 2021**

Lineare Algebra 1

Goethe–Universität Frankfurt — Sommersemester 2020
für Bachelor Mathematik und L3-Mathematik

JAKOB STIX

Zusammenfassung. — Die Vorlesung Lineare Algebra 1 behandelt die Theorie der Vektorräume. Es geht um lineare Abbildungen, Matrizen, Determinanten, lineare Gleichungssysteme, Eigenwerte und Eigenräume, Diagonalisierbarkeit, symmetrische Bilinearformen sowie um Normalformen. Die Beschäftigung mit Vektoräumen vermittelt einen Einstieg in die abstrakte algebraische Art mathematisch zu denken.

Inhaltsverzeichnis

1. Einführung

Lineare Algebra bietet einen Einstieg in die universitäre Mathematik abstrakter Strukturen. Die meisten Schwierigkeiten resultieren nicht aus den Begriffen und ihren Beziehungen untereinander, den Sätzen, Theoremen und Lemmata, sondern aus der Tatsache, daß das Fach an der Schwelle von der Schule zur Universität gelehrt wird und die Studierenden sich an vielerlei Neues gewöhnen müssen. Ich sage bewußt gewöhnen, denn mit ausreichend viel Zeit und Durchhaltevermögen werden Sie Erfolg haben. Seien Sie jedenfalls eines versichert: die Konzepte und Sätze der linearen Algebra werden in Ihrem weiteren Studium ganz selbstverständlich verwendet werden, und irgendwann werden Sie es auch nicht mehr merken, ganz wie Sie die Erinnerung daran vergessen haben, als Sie beim Lesen aus einzelnen Buchstaben ganze Wörter geformt haben.

Obwohl die lineare Algebra als Theorie für viele Dinge der Physik und anderer Anwendungen zu einfach ist, zum Beispiel um periodische Vorgänge und Gleichgewichtszustände zu beschreiben, so gibt es doch trotzdem eine Fülle von versteckten Anwendungen. Einige davon wollen wir sehr kurz streifen, um ein wenig Werbung für das Fach zu machen. Keine der Anwendungen werden wir im Detail behandeln (Sie lernen ja das Lesen an sich, Bücher müssen Sie schon selbst zur Hand nehmen), aber Sie lernen im Prinzip die Sprache, in der die Erklärungen geschrieben sind.

1.1. **Bildkompression.** Zur Vereinfachung betrachten wir ein digitales Schwarzweißbild. Dies besteht nochmals vereinfachend aus einem rechteckigen Raster aus Bildpunkten, den Pixeln, von denen jedes entweder schwarz oder weiß sein kann. Diese Information codieren wir durch eine 0 für Weiß und eine 1 für Schwarz, und zwar für jedes Pixel. Ein Bild besteht also im Prinzip aus einer langen Folge von 0 oder 1, was man am besten als Vektor[1] in einem Vektorraum über dem Körper $\mathbb{F}_2 = \{0, 1\}$ auffaßt. Die Dimension ist dabei die Anzahl der Pixel und ziemlich groß.

Läßt man im Bestreben, das Bild mit weniger Information darzustellen, ein paar Pixelwerte weg, dann wird das Bild unscharf oder eine Stelle verschwindet ganz. Das geht also nicht.

Um den Speicherbedarf zu reduzieren, bietet es sich an, eine Basistransformation zu machen, so daß zwar dasselbe Bild (Vektor) beschrieben wird aber die Koordinaten nun über das ganze Bild verschmiert eine Bedeutung haben. Wenn man nun weiß, welche Koordinaten in unserer Wahrnehmung des Bildes keine so große Rolle spielen, dann kann man diese Koordinaten getrost vergessen und hat immer noch (fast) das gleiche Bild. So funktioniert im Prinzip Datenkompression.

1.2. **Komplexe Zahlen.** Beim Lösen quadratischer Gleichungen

$$X^2 + pX + q = 0$$

mit reellen Zahlen $p, q \in \mathbb{R}$ als Koeffizienten fehlen Quadratwurzeln aus negativen Zahlen. Es ist unmittelbar klar, daß es ausreicht, aus -1 eine Wurzel zu ziehen, denn für $a > 0$ bestehen wir auf $\sqrt{-a} = \sqrt{a} \cdot \sqrt{-1}$. Aber für alle $x \in \mathbb{R}$ gilt: $x^2 \geq 0$. Also geht das nicht — in $\mathbb{R}$! Die komplexen Zahlen werden formal am besten durch eine Körper-Struktur auf einem zweidimensionalen $\mathbb{R}$-Vektorraum erklärt. Hier hilft die lineare Algebra bei der Konstruktion von ansonsten imaginären Zahlen und hilft dabei ein psychologisches Hindernis zu überwinden.

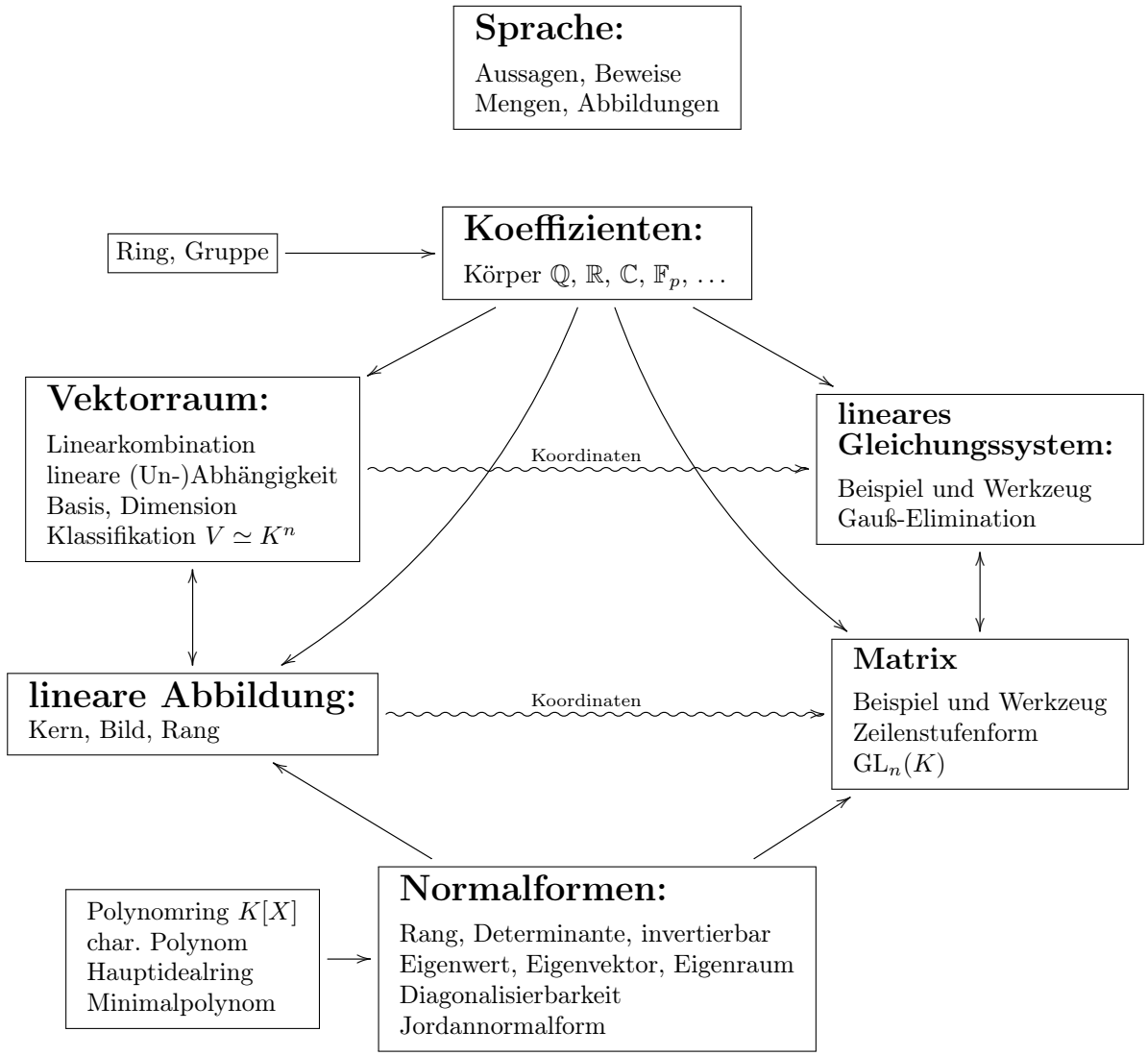
Nun, das ist ein bisschen gelogen, denn man muß ja zuerst zu der Einsicht kommen, daß die abstrakten Konstruktionen der linearen Algebra als *konkret* zu betrachten sind. Aber das ist nur ein Gewöhnungseffekt, den Sie desto schneller erreichen, je mehr Sie sich mit der linearen Algebra beschäftigen.

[1] Ich verwende bewußt bereits die Sprache der linearen Algebra, die noch einzuführen sein wird, damit Sie später, sofern Sie die Einleitung nochmals lesen, einen Bezug herstellen können.

1.3. **Methode der kleinsten Quadrate zur besten Approximation.** Will man durch eine Punktwolke in der Ebene, die das experimentelle Ergebnis eines postulierten linearen Zusammenhangs darstellt, eine beste Näherungsgerade finden, so gibt es die Methode der kleinsten Quadrate. Diese stammt von Gauß und wurde 1801 von ihm erfolgreich zur der Berechnung der Position des Zwergplaneten Ceres eingesetzt, dessen Spur die Astronomen verloren hatten.

Die Methode der kleinsten Quadrate ist reine lineare Algebra, obwohl der Name quadratisches verspricht. Es ist eine Besonderheit quadratischer Formen durch symmetrische Matrizen beschrieben zu werden, so daß die lineare Algebra hier auch ein gewaltiges Stück mitreden kann.

1.4. **Eine grobe Skizze der Landschaft *Lineare Algebra 1*.**



Literatur

[Be14] Albrecht Beutelspacher, *Lineare Algebra*, Springer, 2014, xiv+368 Seiten.
[Bo14] Siegfried Bosch, *Lineare Algebra*, Springer, 2014, x+385 Seiten.

Teil 1. Objekte

2. Grundbegriffe und Sprache

2.1. **Aussagen und Wahrheitstafeln.** In der Mathematik unterscheidet man die formale **syntaktische** Sprache von der **Semantik**, der Bedeutung. Eine Aussage ist zuerst einmal eine gültiger formaler Ausdruck (keine Rechtschreib- oder Grammatikfehler). Durch die Interpretation des Ausdrucks in einem Modell bekommt die Aussage (A) einen **Wahrheitsgehalt:** die Aussage ist **wahr (W)** oder **falsch (F)**.

Beispiel 2.1. Aussagen sind zum Beispiel:

- „$5 > 3$“
- „Es gibt keine natürlichen Zahlen a, b mit $\pi = \frac{a}{b}$.“
- „Die Erde ist eine Scheibe.“

Die ersten beiden Aussagen sind wahr, die dritte ist falsch. Keine Aussagen sind zum Beispiel:

- „Ist heute Donnerstag?“
- „$x = 3$.“

Die Gleichung $x = 3$ hat zwar keine Schreibfehler, aber ihr Wahrheitsgehalt ist nicht festgelegt, weil x nicht spezifiziert ist. Man kann daraus eine Aussage machen, indem man zuerst den Wert von x angibt:

- „Sei $x = 5$. Es gilt $x = 3$.“

Das ist dann eine falsche Aussage, aber immerhin eine Aussage.

2.1.1. *Logische Operationen.* Aussagen kann man mit logischen Operationen verbinden und erhält dadurch neue Aussagen. Deren Wahrheitsgehalt hängt auf festgelegte Art und Weise vom Wahrheitsgehalt der Aussagen ab, aus denen sich die neue Aussage zusammensetzt.

(1) Die **Negation** (oder Verneinung) einer Aussage A ist die Aussage $\neg A$.
(2) Das logische **und** zweier Aussagen A, B ist genau dann wahr, wenn beide Aussagen A und B wahr sind. Man findet die Notation $A \wedge B$.
(3) Das logische **oder** zweier Aussagen A, B ist genau dann wahr, wenn eine der beide Aussagen A oder B wahr ist. Gemeint ist hier mindestens eine, also „eine“ im Sinne von „es gibt eine“. Wenn mehr als eine Teilaussage wahr ist, dann ist das auch gut. Man findet die Notation $A \vee B$.

    Vorsicht: das mathematische oder ist nicht dasselbe wie das sich gegenseitig ausschließende entweder oder. Umgangssprachlich wird leider oft auf das „entweder“ verzichtet.

Am einfachsten erklärt sich das durch eine **Wahrheitstafel**:

A	B	$\neg A$	A und B	A oder B	(A und $\neg B$) oder ($\neg A$ und B)
W	W	F	W	W	F
W	F	F	F	W	W
F	W	W	F	W	W
F	F	W	F	F	F

Abbildung 1. Wahrheitstafel für Negation, und und oder, sowie entweder oder

Bemerkung 2.2. Aus der Wahrheitstafel lesen wir sofort die folgenden Regeln ab:

$$\begin{aligned}
\neg\neg A &\quad \text{hat den gleichen Wahrheitsgehalt wie} \quad A,\\
\neg(A \text{ und } B) &\quad \text{hat den gleichen Wahrheitsgehalt wie} \quad \neg A \text{ oder } \neg B,\\
\neg(A \text{ oder } B) &\quad \text{hat den gleichen Wahrheitsgehalt wie} \quad \neg A \text{ und } \neg B,\\
A \text{ und } (B \text{ oder } C) &\quad \text{hat den gleichen Wahrheitsgehalt wie} \quad (A \text{ und } B) \text{ oder } (A \text{ und } C),\\
A \text{ oder } (B \text{ und } C) &\quad \text{hat den gleichen Wahrheitsgehalt wie} \quad (A \text{ oder } B) \text{ und } (A \text{ oder } C).
\end{aligned}$$

2.1.2. *Logisches Argumentieren.* Auch das mathematische Argumentieren mit Aussagen läßt sich formalisieren. Dazu muß man festlegen, wann eine Argumentation, also eine Beweisführung, aus einer Aussage A eine Aussage B ableitet. Eine solche Argumentation wird **Implikation** genannt und mit $A \implies B$ notiert. In Abhängigkeit der Wahrheitswerte von A und B hat $A \implies B$ die folgenden Wahrheitswerte; wahr bedeutet, daß die Argumentation gültig, d.h. fehlerfrei, war, und falsch bedeutet, daß im Beweis ein Fehler unterlaufen ist.

Unter der **Äquivalenz** von Aussagen $A \iff B$ verstehen wir $(A \implies B)$ und $(B \implies A)$.

A	B	$A \implies B$	$\neg B \implies \neg A$	$\neg(A \text{ und } \neg B)$	$A \iff B$
W	W	W	W	W	W
W	F	F	F	F	F
F	W	W	W	W	F
F	F	W	W	W	W

Abbildung 2. Wahrheitstafel für Implikation und Äquivalenz

Bemerkung 2.3. (1) Die Implikation $A \implies B$ ist genau dann fehlerhaft (falsch), wenn man aus einer wahren Aussage eine falsche Aussage ableitet.

(2) Aus der Wahrheit von $A \implies B$ kann man nicht folgern, daß B wahr ist. Die Aussage

$$\text{„}5 = 3 \implies 5^2 = 3^2\text{“}$$

ist wahr. Nur wenn auch die Voraussetzung A wahr ist, folgt aus dem Beweis $A \implies B$ die Wahrheit von B.

(3) Das Symbol $A \iff B$ darf nur genau dann verwendet werden, wenn die Aussagen A und B den gleichen Wahrheitswert haben: beide wahr oder beide falsch.

(4) Die Implikation $A \implies B$ interpretieren wir als „wenn A, dann B“. Das erklärt die obige Wahrheitstafel. Wenn A falsch ist, dann haben wir durch „wenn A, dann B“ nichts versprochen, denn A ist ja nicht wahr. Wir können ungeachtet des Wahrheitswerts von B behaupten, daß B aus A folgt. Aus falschen Aussagen kann man alles folgern!

2.1.3. *Quantoren.* Manchmal möchte man eine Aussage A für eine Liste von Situationen gemeinsam betrachten. In diesem Fall kann A von einem Parameter abhängen, wir schreiben $A(x)$ mit Parameter x, und erst eine Aussage werden, wenn man diesen Parameter festlegt. Zum Beispiel

$$A(x) = \text{ das Quadrat der ganzen Zahl } x \text{ läßt bei Division durch 4 den Rest 1.}$$

Wenn man dann ausdrücken will, daß alle Aussagen gleichzeitig gelten, also ohne Ausnahme, dann nutzt man einen **Quantor**, den **Allquantor** $\forall$, genauer

$$\forall x \text{ aus der Liste der erlaubten Paramter}: \quad A(x).$$

Das Symbol $\forall$ wird „**für alle**“ gelesen. Diese neue Aussage ist genau dann wahr, wenn ohne Ausnahme für alle erlaubten x die spezifische Aussage $A(x)$ wahr ist. Ein Beispiel mit obigem $A(x)$:

$$\forall x \text{ ungerade ganze Zahl gilt: } A(x)$$

ist wahr, denn ein ungerades x hat die Form $x = 2k + 1$ und

$$(2k+1)^2 = 4k^2 + 4k + 1 = 4 \cdot k(k+1) + 1$$

läßt den 4er-Rest 1.

Was ist nun die Negation einer $\forall$-Aussage? Wie beim logischen und wird eine $\forall$-Aussage falsch, sobald es nur ein Gegenbeispiel gibt. Die Existenz eines geeigneten Parameters wird mit dem **Existenzquantor** $\exists$ beschrieben, der „**es existiert**“ gelesen wird. Beispielsweise wird die Wahrheit von

$$\exists x \text{ ganze Zahl mit } \neg A(x)$$

durch die Angabe eines Beispiels als wahr erkannt: etwa $x = 2$ (hat $x^2 = 4$ und den 4-er Rest 0).

Der Existenzquantor wird manchmal auch etwas nachlässig als „es gibt ein“ gelesen, was korrekterweise mit „es gibt mindestens ein“ interpretiert werden muss (nicht genau ein).

Bemerkung 2.4. Die Quantoren $\forall$ und $\exists$ sind keine Abkürzungen. Sie müssen stets als Indikator über die Qualität einer Aussage in Bezug auf eine Parameterwahl benuutzt werden.

Bemerkung 2.5. Es gelten die folgenden Regeln (die ... sind der Platzhalter für die Beschreibung welche x gemeint sind):

$$\neg\forall x \ldots A(x) \quad \text{hat den gleichen Wahrheitsgehalt wie} \quad \exists x \ldots \neg A(x),$$
$$\neg\exists x \ldots A(x) \quad \text{hat den gleichen Wahrheitsgehalt wie} \quad \forall x \ldots \neg A(x).$$

Die erste Regel drückt aus, daß eine $\forall$-Aussage widerlegt wird durch die Angabe bereits eines Gegenbeispiels (und so soll man das beim mathematischen Beweisen auch halten!). Eine Diskussion darüber, wie oft die Aussage trotzdem richtig ist, hat keine Bedeutung.

Eine $\exists$-Aussage zu widerlegen, ist deutlich anstrengender, denn hierbei muß man sich davon überzeugen, daß in allen Fällen das Gegenteil gilt.

Ein Beispiel: die Aussage

$$\forall \text{ Tage seit der Endeckung Amerikas gilt: der Tag hat 24h}$$

ist falsch, denn es gibt ab und zu Tage, an denen aus astronomischen Gründen eine Schaltsekunde eingefügt oder ausgelassen werden muß(die Erde eiert ein wenig um die Sonne).

Die mathematische Theorie der Begriffe „Aussage“ und „Beweis“ werden in der Disziplin **Logik** fundiert behandelt. Wir belassen es bei einem „naiven“ und intuitiven Umgang.

2.2. **Beweise und Beweisstrategien.** Es folgt eine Auswahl von Beweisstrategien.

2.2.1. *Direkter Beweis.* Ein direkter Beweis argumentiert vorwärts. Ausgehend von einer wahren Aussage A und einem wahren Beweis $A \Longrightarrow B$ wird auf die Wahrheit von B geschlossen:

$$\text{wenn } A \text{ und } (A \Longrightarrow B) \text{ wahr sind, dann ist auch } B \text{ wahr.}$$

Umgangssprachlich: die Voraussetzung A trifft zu, die Argumentation $A \Longrightarrow B$ ist nicht fehlerhaft, also gilt B. Es folgt ein Beispiel.

Proposition 2.6 (Goldener Schnitt)**.** *Sei x eine reelle Zahl mit $x^2 = x + 1$. Dann gilt*

$$x = \frac{1+\sqrt{5}}{2} \text{ oder } x = \frac{1-\sqrt{5}}{2}$$

Beweis. Aus $x^2 = x + 1$ folgt durch Umstellen und quadratisches Ergänzen $(x - 1/2)^2 = 5/4$. Weil Quadrate reller Zahlen nur dann gleich sind, wenn sie sich höchstens um das Vorzeichen unterscheiden, gilt $x - 1/2 = \sqrt{5/4}$ oder $x - 1/2 = -\sqrt{5/4}$. Daraus folgt unmittelbar die Behauptung. □

2.2.2. *Indirekter Beweis.* Der indirekte Beweis argumentiert gewissermaßen rückwärts. Wir nutzen aus, daß laut Tabelle 2 gilt

$$\text{„}A \implies B\text{“ genau dann wenn „}\neg B \implies \neg A\text{“.}$$

Die Implikation $\neg B \implies \neg A$ nennt man die **Kontraposition** zu $A \implies B$. Es folgt ein Beispiel.

Proposition 2.7. *Seien x, y reelle Zahlen und $xy \neq 0$. Dann gilt $x \neq 0$ und $y \neq 0$.*

Beweis. Anstelle von $A \implies B$ mit $A = \text{„}xy \neq 0\text{“}$ und $B = \text{„}x \neq 0$ und $y \neq 0\text{“}$ zeigen wir die Kontraposition. Zu zeigen ist also

$$\neg B = (x = 0 \text{ oder } y = 0) \implies \neg A = (xy = 0),$$

und das ist trivial, weil multiplizieren mit 0 immer den Wert 0 ergibt. □

2.2.3. *Beweis durch Widerspruch.* Der Beweis durch Widerspruch ist eine Stärke der Mathematik, um die uns die anderen Wissenschaften beneiden. Das funktioniert nur, weil wir an die Widerspruchsfreiheit der Mathematik glauben (die sich nicht beweisen läßt!). Angenommen, wir wollen die Aussage A zeigen. Dann nehmen wir das Gegenteil als wahr an: „Angenommen $\neg A$ ist wahr . . .“, und schauen, was daraus folgt. Wenn es uns gelingt,

$$\neg A \implies B$$

eine Aussage B zu folgern, von der wir wissen, daß B falsch ist, dann haben wir gewonnen. Die Widerspruchsfreiheit der Mathematik duldet das nicht, also muß $\neg A$ falsch sein, bzw. A wahr. Mit anderen Worten, der Beweis $\neg A \implies B$ ist korrekt, also auch die Kontraposition $\neg B \implies A$. Nun ist B falsch, also $\neg B$ wahr, und damit wegen der korrekten Implikation auch A wahr.

Bemerkung 2.8. Der Methode des Widerspruchsbeweises wohnt eine besondere Fehleranfälligkeit inne. Ein Fehler in einem direkten Beweis führt vom Weg ab und typischerweise nicht zu dem angestrebten Resultat. Man merkt sofort, daß der Beweis falsch ist. Bei einem Widerspruchsbeweis führt ein Fehler in der Regel zu einer falschen Aussage. Aber das ist genau das gewünschte Ergebnis! Und schon hat man sich vertan.

Hier ist das mit Recht berühmteste Beispiel eines Widerspruchsbeweises. Dazu erinnern wir an den Begriff der Primzahl.

Definition 2.9. Eine **Primzahl** ist eine natürliche Zahl $p \geq 2$, die nur durch 1 und sich selbst teilbar ist, d.h. wenn $p = ab$ mit natürlichen Zahlen a und b, dann gilt $a = 1$ oder $b = 1$.

Theorem 2.10 (Euklid)**.** *Es gibt unendlich viele Primzahlen.*

Beweis. **Angenommen, die Behauptung ist falsch**: es gäbe nur endlich viele Primzahlen. Dann kann man diese auf eine endliche Liste schreiben, sagen wir $p_1, p_2, \ldots, p_n$ sei die Liste aller Primzahlen. Weil $p = 2$ eine Primzahl ist, ist diese Liste nicht leer. Wir betrachten nun die natürliche Zahl $N = P + 1$ für das Produkt

$$P = p_1 \cdot p_2 \cdot \ldots \cdot p_n.$$

Dies ist eine Zahl $N \geq 2$ und hat daher einen Primteiler (das muß man auch noch beweisen), sagen wir ℓ. Weil unsere Liste der Primzahlen vollständig ist, teilt ℓ das Produkt P. Daher läßt N bei Division durch ℓ den Rest 1, denn die Division von P durch ℓ geht auf und $N = P + 1$.

Das ist der gesuchte Widerspruch, denn ℓ kann nicht gleichzeitig ein Primteiler von N sein und N bei Division durch ℓ den Rest 1 lassen. □

2.2.4. *Vollständige Induktion.* Die vollständige Induktion kommt oft ins Spiel, wenn man eine Folge von Aussagen A_n mit $n = 1, 2, 3 \ldots$ beweisen möchte. Man beweist dann zwei Dinge:

(1) **Induktionsanfang (Induktionsverankerung):** Die Aussage A_1 ist richtig.
(2) **Induktionsschritt:** es gilt für alle n die Implikation $A_n \Longrightarrow A_{n+1}$.
Hierbei formuliert man die Aussage A_n als **Induktionsvoraussetzung**, die man als wahr annimmt, und zeigt auf Grundlage dieser Voraussetzung die Aussage A_{n+1}.

Die vollständige Induktion schließt aus der Wahrheit von A_1 der Reihe nach auf die Wahrheit von A_2, dann von A_3, ... Es ist eine Eigenschaft der natürlichen Zahlen, daß damit dann auch jede Aussage A_n irgendwann erreicht wird und dann bewiesen ist.

Definition 2.11. Für eine natürliche Zahl $n \in \mathbb{N}$ ist die **Fakultät** definiert als

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n.$$

Es gilt per Konvention $0! = 1$ (denn das leere Produkt ist 1), damit gilt nämlich für alle $n \geq 1$

$$n! = n \cdot (n-1)!$$

als rekursive Definition (siehe Bemerkung 3.2) der Fakultät.

Proposition 2.12. *Sei $n \geq 1$ eine natürliche Zahl. Es gibt genau $n!$ verschiedene Möglichkeiten um n verschiedene Dinge in eine Reihenfolge zu bringen.*

Beweis. Es spielt offensichtlich keine Rolle, welche n verschiedenen Dinge wir sortieren. Nehmen wir also die Zahlen $1, 2, \ldots, n$. Die verschiedenen möglichen Reihenfolgen unterscheiden wir nun durch die Zahl, welche als erstes kommt. Für jedes i mit $1 \leq i \leq n$ passiert das genau so oft, wie man die restlichen $n-1$ vielen Dinge in eine Reihenfolge bringen kann (diese wird dann hinter das führende i sortiert. Nach **Induktionsvoraussetzung** geht das $(n-1)!$ mal. Und davon haben wir n viele, weil es für i genau n viele Möglichkeiten gibt. Insgesamt macht das

$$n \cdot (n-1)! = n!$$

viele Möglichkeiten. Das ist die Behauptung für n. Damit ist der **Induktionsschritt** vollzogen.

Aber ohne den **Induktionsanfang** zeigt das gar nichts. Der Induktionsanfang spricht für $n = 1$ und behauptet, daß es $1! = 1$ Möglichkeit gibt. Das stimmt auch, und damit ist der Induktionsbeweis abgeschlossen. □

Bemerkung 2.13. Das Beweisverfahren der vollständigen Induktion kann man modifizieren.

(1) Zum Beispiel kann die Induktion bei einer Aussage A_0 beginnen. Allgemeiner noch kann man eine Aussage A_n für alle $n \geq n_0$ haben und beginnt die Induktion mit der Induktionsverankerung bei der Aussage A_{n_0}.
(2) Es kann sein, daß der Beweis des Induktionsschritts leichter ist, wenn man eine Aussage überspringt. Dann zeigt man als Induktionsschritt $A_n \Longrightarrow A_{n+2}$. Als Induktionsanfang benötigt man dann die Aussagen A_0 und A_1. Im Grunde hat man dann zwei vollständige Induktionen: eine über die geraden Zahlen und eine über die ungeraden Zahlen.
(3) Weiterhin kann der Induktionsschritt mehr als die vorherige Aussage benötigen, etwa die letzten zwei. Es gelingt zum Beispiel nicht der Beweis $A_n \Longrightarrow A_{n+1}$ sondern „nur“ der Beweis von A_n und $A_{n+1} \Longrightarrow A_{n+2}$. In diesem Fall muß man den Induktionsanfang auch durch entsprechend die ersten zwei Aussagen nachweisen. Noch besser ist es in der Regel, die Aussagen A_n durch die Aussagen

$$A'_n = A_0 \text{ und } A_1 \text{ und } A_2 \text{ und } \ldots \text{ und } A_n$$

zu ersetzen. Suggestiv würde ich dafür sogar $A_{\leq n}$ schreiben. Beim Beweis nach vollständiger Induktion macht man ja genau das, man beweist der Reihe nach. Und dann kann man zum Beweis von A_{n+1} ja bereits auf die Wahrheit aller Aussagen A_i mit $0 \leq i \leq n$ zurückgreifen.

Beispiel 2.14. Die Fibonacci-Folge ist rekursiv definiert durch

$$a_{n+2} = a_n + a_{n+1} \qquad \text{für alle } n \geq 0$$

und die Anfangswerte $a_0 = 0$ und $a_1 = 1$. Die Folge beginnt

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

Sei $\varphi = \frac{1}{2}(1 + \sqrt{-5})$ die Lösung von $\varphi^2 = \varphi + 1$ aus Proposition 2.6. Dann gilt die geschlossene Formel für die Fibonacci-Folge:

$$a_n = \frac{1}{\sqrt{5}}\big(\varphi^n - (-\varphi)^{-n}\big).$$

Dies beweisen wir nun per vollständiger Induktion.

Im Induktionsschritt beweisen wir die Aussage für $n+2$ aus den Aussagen für n und $n+1$. Wir rechnen dazu

$$\begin{aligned} a_{n+2} = a_n + a_{n+1} &= \frac{1}{\sqrt{5}}\big(\varphi^n - (-\varphi)^{-n}\big) + \frac{1}{\sqrt{5}}\big(\varphi^{n+1} - (-\varphi)^{-(n+1)}\big) \\ &= \frac{1}{\sqrt{5}}\big(\varphi^n(\varphi+1) - (-\varphi)^{-(n+2)}((-\varphi)^2 + (-\varphi))\big) \\ &= \frac{1}{\sqrt{5}}\big(\varphi^{n+2} - (-\varphi)^{-(n+2)}\big), \end{aligned}$$

weil

$$\varphi + 1 = \varphi^2 \quad \text{ und } \quad (-\varphi)^2 + (-\varphi) = \varphi^2 - \varphi = 1.$$

Wenn man den Induktionsschritt zuerst erledigt, dann kann man sich danach Rechenschaft darüber ablegen, welche Verankerung notwendig ist. In diesem Fall gilt die Rechnung ab $n = 0$ und liefert die gesuchte Formel beginnend mit a_2 und Benutzung der Formel für die vorherigen beiden Folgenglieder. Als Verankerung brauchen wir demnach die Formel für a_0 und a_1. Es gilt

$$\frac{1}{\sqrt{5}}\big(\varphi^0 - (-\varphi)^{-0}\big) = \frac{1}{\sqrt{5}}\big(1 - 1\big) = 0 = a_0,$$

und, mittels $\varphi^2 - \varphi = 1$ sowie $2\varphi - 1 = \sqrt{5}$,

$$\frac{1}{\sqrt{5}}\big(\varphi^1 - (-\varphi)^{-1}\big) = \frac{\varphi^2 + 1}{\varphi\sqrt{5}} = \frac{2\varphi^2 - \varphi}{\varphi\sqrt{5}} = \frac{2\varphi - 1}{\sqrt{5}} = 1 = a_1,$$

und das beendet den Induktionsbeweis.

2.2.5. *Beweis von Äquivalenzen.* Wenn man von zwei Aussagen A und B deren Äquivalenz zu zeigen hat, dann ist es meistens einfacher, die Aufgabe in zwei Teile aufzuteilen. Es gilt nämlich

$$(A \iff B) \quad \text{hat den gleichen Wahrheitsgehalt wie} \quad (A \implies B) \text{ und } (B \implies A).$$

Oft gibt es dann die „einfache Richtung“, die man fast geschenkt bekommt, und die Schwierigkeiten verstecken sich in der anderen Richtung des Implikationsschlusses.

2.2.6. *Beweis durch Rückwärtsarbeiten.* Manchmal sind Aussagen A und B von denen man $A \Longrightarrow B$ zeigen soll, sehr weit voneinander entfernt. Man sieht nicht, was das eine mit dem anderen zu tun haben soll. Dann gilt es vermutlich eine Reihe $C_1, \ldots, C_n$ von Zwischenaussagen zu finden und zu zeigen, daß

$$A \Longrightarrow C_1 \Longrightarrow C_2 \ldots \Longrightarrow C_{n-1} \Longrightarrow C_n \Longrightarrow B.$$

Wenn der so zu bestreitende Beweisweg zu lange ist, dann weiß man nicht, in welche Richtung man losargumentieren soll. In diesem Fall ist es hilfreich auch zu raten, wie der Beweis am Ende verlaufen soll. Man sucht also neben dem ersten Schritt $A \Longrightarrow C_1$ auch den letzten $C_n \Longrightarrow B$, und so weiter. So arbeitet man sich von zwei Seiten aufeinander zu. Es folgt ein Beispiel.

Proposition 2.15. *Für zwei nichtnegative reele Zahlen $x, y \in \mathbb{R}$, $x, y \geq 0$ gilt die Ungleichung vom* ***geometrischen Mittel*** *gegen das* ***arithmetische Mittel****:*

$$\sqrt{x \cdot y} \leq \frac{x+y}{2}.$$

Beweis. Weil $x, y \geq 0$ positiv sind, folgt die Behauptung aus der Ungleichung

$$4xy \leq (x+y)^2,$$

die durch Mulitplikation mit 2 und anschließendem Quadrieren der Behauptung entsteht. Diese Aussage ist durch Ausmultiplizieren, Umstellen, sowie der Binomischen Formel äquivalent zu

$$0 \leq (x-y)^2.$$

und das ist wahr. □

Bemerkung 2.16. Im Beweis wurde die Voraussetzung $x, y \geq 0$ gebraucht, damit der erste Rückwärtsschritt gültig ist. Dann wird solange weiter rückwärts gearbeitet (sogar mit Äquivalenzschritten, so daß der Wert des Rückwärtsarbeiten nicht klar hervortritt), bis eine bekanntermaßen wahre Aussage entsteht. Dann ist man fertig.

2.2.7. *Beweis mittels Extremalprinzip.* Das Extremalprinzip ist eine mächtige Leitidee, wenn man mathematische Objekte mit einer besonderen Eigenschaft „$\mathscr{E}$“ konstruieren möchte. Unter allen möglichen Objekten $\mathscr{X}$ sucht man dazu ein Objekt $\mathscr{X}_{\text{ext}}$, in welchem ein Merkmal $\mathscr{M}$ extremal (minimal oder maximal) ausfällt. Die Existenz eines solchen extremalen Objekts $\mathscr{X}_{\text{ext}}$ muß man sicherstellen, und das ist eine oft übersehene kritische Stelle in einem Beweis mittels Extremalprinzip.

Hat man aber erst einmal eine extremale Wahl getroffen, dann zeigt man $\mathscr{E}$ für $\mathscr{X}_{\text{ext}}$ auf die folgende Art durch einen Widerspruchsbeweis. Aus der Annahme, daß $\mathscr{E}$ nicht gilt, zeigt man für eine kleine Variation $\mathscr{X}'$ von $\mathscr{X}_{\text{ext}}$, daß das Merkmal $\mathscr{M}$ bei $\mathscr{X}'$ extremaler als bei $\mathscr{X}_{\text{ext}}$ ausfällt, und das geht ja bekanntlich nicht (denn extremal läßt sich nicht steigern).

Der Beweis per Extremalprinzip hängt entscheidend von der Auswahl eines guten Merkmals $\mathscr{M}$ ab, einer Art Bewertungsfunktion. Man gewinnt durch die Auswahl eines bezüglich $\mathscr{M}$ extremalen Objekts, daß man nicht mehr mit Objekten ohne Unterschied arbeitet und $\mathscr{E}$ versucht bei einem dieser Objekte nachzuweisen, sondern man hat die extremale Ausprägung von $\mathscr{M}$, mit der man arbeiten kann.

Eine Variante des Extremalprinzips zeigt eine Behauptung dadurch, dass man die Menge A der Objekte betrachtet, welche der Behauptung widersprechen, die Menge der Ausnahmen. Ziel ist es, A als leer nachzuweisen. Aus dieser Menge A betrachtet man nun nach dem Extremalprinzip gemäß eines gut gewählten Merkmals extremale Elemente. Vorsicht: deren Existenz muß begründet werden. Durch Variation konstruiert man nun ein Element von A, in welchem dieses Merkmal extremaler ausfällt, ein Widerspruch zur extremalen Wahl. Und damit auch ein Widerspruch dazu, daß A überhaupt Elemente enthält. Es folgt ein Beispiel.

Proposition 2.17. *Für keine natürliche Zahl $n > 0$ ist $n\sqrt{2}$ eine ganze Zahl.*

Beweis. Sei A die Menge der Gegenbeispiele zur Behauptung, also die Menge aller natürlichen Zahlen n, für die $n\sqrt{2}$ eine ganze Zahl ist. Angenommen, A ist nicht leer. Dann gibt es in A wie in jeder nichtleeren Menge natürlicher Zahlen eine kleinste natürliche Zahl: ein kleinstes Gegenbeispiel. Dies sei $a \in A$, also ist

$$a\sqrt{2} \text{ ganz, und für alle natürlichen Zahlen } k < a \text{ ist } k\sqrt{2} \text{ nicht ganz.}$$

Mit $a\sqrt{2}$ ist auch $k = a\sqrt{2} - a$ eine natürliche Zahl. Des weitern gilt $1 \leq k < a$, weil $\sqrt{2} - 1 < 1$ und daher

$$k = a(\sqrt{2} - 1) < a.$$

Dann ist

$$k\sqrt{2} = a(\sqrt{2} - 1)\sqrt{2} = 2a - a\sqrt{2}$$

einerseits keine ganze Zahl, weil $k < a$ und wegen der Extremalität von $a \in A$ daher $k \notin A$. Andererseits handelt es sich um eine ganze Zahl wie die rechte Zahl als Differenz der ganzen Zahlen $2a$ und $a\sqrt{2}$, denn $a \in A$, zeigt. Dies ist der gesuchte Widerspruch. Folglich ist A leer und die Aussage gezeigt. □

2.3. **Mengen.** Eine **Menge** besteht aus **Elementen**. Ist a ein Element der Menge M, so schreiben wir

$$a \in M.$$

Wenn a kein Element von M ist, so schreiben wir $a \notin M$. Wir notieren/definieren eine Menge gerne auf die folgende Art und Weise:

$$M = \{a \ ; \ \text{Beschreibung, welche } a \text{ Elemente von } M \text{ sind}\}.$$

Das Semikolon, manchmal auch : oder |, nach a bedeutet, daß danach die charakterisierende Eigenschaft kommt, welche die zu betrachtenden a auswählt. Man liest dieses Zeichen als „für die gilt“.

Beispiel 2.18. Das (mathematisch) einfachste Beispiel einer Menge ist die **leere Menge**, für die es mehrere Notationen gibt, etwa

$$\emptyset = \{\ \},$$

und die dadurch definiert ist, daß sie keine Elemente enthält. Für alle a gilt also

$$a \notin \emptyset.$$

Beispiel 2.19. Grundlegende Zahlbereiche setzen wir in dieser Vorlesung als bekannt voraus. Dies sind die folgenden Mengen von „Zahlen“:

(1) die **natürlichen Zahlen (ohne** 0**)**

$$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\},$$

(2) die **natürlichen Zahlen (mit** 0**)**

$$\mathbb{N}_0 = \{0, 1, 2, 3, 4, 5, \ldots\},$$

(3) die **ganzen Zahlen**

$$\mathbb{Z} = \{\ldots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \ldots\},$$

(4) die **rationalen Zahlen** $\mathbb{Q}$

$$\mathbb{Q} = \{\frac{a}{b} \ ; a \in \mathbb{Z}, b \in \mathbb{N}, \text{ der Bruch } a/b \text{ sei gekürzt}\},$$

(5) und die **reellen Zahlen** $\mathbb{R}$ (die Beschreibung gehört in die Analysis).

Wir leugnen nicht, daß es bei der Konstruktion dieser Zahlbereiche etwas zu lernen gibt. Wir skizzieren die axiomatische Einführung der natürlichen Zahlen, der ganzen Zahlen und der rationalen Zahlen, in Abschnitt 3.1 und verschieben Details in den Bereich der Übungsaufgaben oder insbesondere für die reellen Zahlen auf eine andere Vorlesung.

Definition 2.20.

(1) Zwei Mengen sind **gleich**, wenn sie dieselben Elemente enthalten. Sind X und Y Mengen, so gilt

$$X = Y \quad \Longleftrightarrow \quad (\forall a \in X \text{ gilt: } a \in Y) \text{ und } (\forall a \in Y \text{ gilt: } a \in X).$$

(2) Eine **Teilmenge** einer Menge X ist eine Menge U, so daß alle Elemente von U auch Elemente von X sind:

$$a \in U \Longrightarrow a \in X.$$

Wir schreiben das als $U \subseteq X$ oder auch $X \supseteq U$. Manchmal ist auch die Notation $U \subset X$ (oder $X \supset U$) gebräuchlich. Diese wollen wir aber für eine **echte** Teilmenge benutzen, das ist eine Teilmenge von X, die nicht ganz X ist:

$$U \subset X \quad \Longleftrightarrow \quad U \subseteq X \text{ und } U \neq X.$$

Wenn U keine Teilmenge von X ist, dann schreiben wir das $U \not\subseteq X$.

Beispiel 2.21. Es folgen einige Beispiele für Teilmengen:

(1) Für jede Menge X gilt $\emptyset \subseteq X$ und $X \subseteq X$.

(2) Für die Zahlbereiche aus Beispiel 2.19 gelten die folgenden Teilmengenbeziehungen:

$$\mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z}.$$

Wie man auch die gewohnten Teilmengenbeziehungen

$$\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$$

verstehen kann, obwohl die Elemente aus $\mathbb{Z}$ ganze Zahlen, die Elemente aus $\mathbb{Q}$ gekürzte Brüche und die Elemente aus $\mathbb{R}$ Äquivalenzklassen (siehe Abschnitt 2.7) von Cauchyfolgen (siehe Analysis) sind, diskutieren wir hier nicht.

(3) Mengen können durchaus auch Elemente von (anderen) Mengen sein. Die Menge $\{\emptyset\}$ bezeichnet die Menge, die als einziges Element die leere Menge enthält. Vorsicht: die Menge $\{\emptyset\}$ enthält ein Element und ist daher nicht leer, auch wenn das enthaltene Element selbst die leere Menge ist. Ein Beispiel für eine echte Teilmenge ist

$$\{\emptyset\} \subset \{\emptyset, \{\emptyset\}\}.$$

Jetzt ist es nicht schwer, weitere Beispiele von Mengen zu konstruieren:

$$\{\emptyset, \{\emptyset\}\}, \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \quad \ldots$$

(4) Für Mengen X und Y gilt offenbar

$$X = Y \quad \Longleftrightarrow \quad X \subseteq Y \text{ und } Y \subseteq X.$$

Diese unscheinbare Umformulierung benutzt man oft, um die Gleichheit von Mengen nachzuweisen. Die Aussage „Gleichheit“ wird dabei in zwei Teilaussagen des Typs „Teilmenge“ zerlegt und somit der Nachweis in zwei Schritte in der Hoffnung, daß jeder einzelne Schritt einfacher zu beweisen sein möge als beide Schritte gleichzeitig.

Definition 2.22. Sei X eine Menge, und seien $U, V \subseteq X$ Teilmengen.

(1) Die **Vereinigung** von U und V ist die Menge

$$U \cup V = \{a \in X \ ; \ a \in U \text{ oder } a \in V\}.$$

(2) Der **Schnitt** von U und V ist die Menge
$$U \cap V = \{a \in X \ ; \ a \in U \text{ und } a \in V\}.$$
(3) Das **Komplement** von U in X ist die Menge
$$X \setminus U = \{a \in X \ ; \ a \notin U\}.$$
Es gibt keine einheitliche Notation für das Komplement. Die folgenden Bezeichnungen für das Komplement kommen vor: $U^c = X \setminus U = X - U = \complement U$.
(4) Die **Differenz** von U und V ist die Menge
$$U \setminus V = \{a \in U \ ; \ a \notin V\}.$$
(5) Die **symmetrische Differenz** von U und V ist die Menge
$$U \triangle V = (U \setminus V) \cup (V \setminus U).$$
(6) Die Teilmengen U und V heißen **disjunkt**, wenn
$$U \cap V = \emptyset.$$
Die Mengen $U \cap V$, $U \cup V$, $X \setminus U$, $U \setminus V$ und $U \triangle V$ sind Teilmengen von X.

Beispiel 2.23. Sei X eine Menge. Sind U, V Teilmengen von X, dann gilt
$$U \setminus V = U \cap (X \setminus V) = U \setminus (U \cap V).$$

Bemerkung 2.24. Seien $U, V \subseteq X$ Teilmengen.
(1) In der Regel gilt
$$U \setminus V \neq V \setminus U.$$
Gleichheit gilt genau dann, wenn $U = V$ gilt.
(2) Es gilt
$$U = U \setminus V \iff V \cap U = \emptyset \iff V = V \setminus U,$$
das heißt, wenn U und V disjunkt sind.
(3) Es gilt
$$U \triangle V = (U \cup V) \setminus (U \cap V).$$

Bemerkung 2.25. Wir betreiben in dieser Vorlesung naive Mengenlehre, aber das führt zu Problemen. Bertrand Russell und Ernst Zermelo haben auf das folgende Paradoxon hingewiesen, das **Russellsche Antinomie** genannt wird. Wie verhält es sich mit der Menge
$$M = \{a \ ; \ a \notin a\}?$$
Das ist die Menge, deren Elemente alle Mengen sind, die sich selbst nicht enthalten. Wir haben bereits gesehen, daß Mengen durchaus auch wieder Elemente von Mengen sein können. Warum also nicht auch von sich selbst?

Fall 1: $M \notin M$. Wenn M sich nicht selbst enthält, dann erfüllt es die definierende Eigenschaft der Elemente von M und damit gilt $M \in M$. Das geht aber nicht, denn wir haben ja angenommen, daß $M \notin M$, und beides gleichzeitig geht nicht.

Fall 2: $M \in M$. Wenn M in M enthalten ist, dann erfüllt M die charakterisierende Eigenschaft seiner Elemente, ist also eine Menge, die sich nicht selbst enthält: $M \notin M$. Aber das ist auch ein Widerspruch.

Weil aber eine der beiden Aussagen $M \in M$ oder $M \notin M$ zutreffen muß, haben wir ein Problem.

Wir lösen die Russellsche Antinomie in dieser Vorlesung nicht auf, versichern aber, daß es eine befriedigende Lösung gibt, z.B. durch ZFC. Das Kürzel steht für ein nach Zermelo und Frenkel benanntes Axiomensystem für Mengen, welches durch das **Auswahlaxiom** (axiom of **C**hoice) ergänzt wird. Wir gehen bei Gelegenheit weiter auf letzteres ein, während wir für axiomatische Mengenlehre und Grundlagen der Mathematik auf die entsprechende Literatur verweisen.

2.4. **Tupel und Produkte von Mengen.** In einer Menge sind die Elemente ungeordnet. Bei einem Tupel ist das anders.

Definition 2.26.
(1) Ein **Tupel** von Elementen aus einer Menge X und einer Menge Y ist ein geordnetes Paar bestehend aus einem Element von X und einem Element von Y. Die Notation ist
$$(a, b)$$
für das Tupel aus den Elementen $a \in X$ und $b \in Y$.
(2) Das **Produkt** der Mengen X und Y ist die Menge
$$X \times Y = \{(a, b) \ ; \ a \in X \text{ und } b \in Y\}$$
aller Tupel von Elementen aus X und Y.

Beispiel 2.27. Sei X eine Menge. Die **Diagonale** von X ist die Teilmenge $\Delta \subseteq X \times X$ definiert durch
$$\Delta = \{(a, a) \ ; \ a \in X\}.$$
Weiter gilt
$$X \times X \setminus \Delta = \{(a, b) \ ; \ a, b \in X, a \neq b\}.$$

Definition 2.28. Eine Familie von Mengen besteht aus einer Menge I von Indizes und zu jedem $i \in I$ eine Menge X_i.
(1) Ein **I-Tupel** aus Elementen der Mengen X_i, $i \in I$ ist eine Kollektion von Elementen $a_i \in X_i$ für alle $i \in I$. Die Notation für so ein I-Tupel ist $(a_i)_{i \in I}$.
(2) Das **Produkt** der Mengen X_i aus der Familie von Mengen mit Indexmenge I ist
$$\prod_{i \in I} X_i = \{(a_i)_{i \in I} \ ; \ a_i \in X_i \text{ für alle } i \in I\}$$
die Menge aller I-Tupel aus der Familie.

Der Fall $I = \{1, 2\}$ führt zum vorher eingeführten Begriff des Tupels und des Produkts. Im Falle von $I = \{1, 2, \ldots, n\}$ schreiben wir für das Tupel
$$(a_1, \ldots, a_n)$$
mit $a_i \in X_i$ und für das Produkt
$$\prod_{i=1}^{n} X_i = X_1 \times \ldots \times X_n = \{(a_1, \ldots, a_n) \ ; \ a_i \in X_i \text{ für alle } i = 1, \ldots, n\}.$$

Beispiel 2.29. Seien n, m natürliche Zahlen. Eine **$n \times m$-Matrix** mit Einträgen aus der Menge X ist ein mit $\{1, \ldots, n\} \times \{1, \ldots, m\}$-indiziertes Tupel
$$(a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$$
von Elementen $a_{ij} \in X$. Eine solche Matrix wird zweidimensional notiert als
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$
Man spricht von Matrizen mit den Abmessungen n **Zeilen** und m **Spalten**. Die Menge aller $n \times m$-Matrizen aus X bezeichnen wir mit
$$\mathrm{M}_{n \times m}(X).$$

Definition 2.30. Seien $U_i \subseteq X$ Teilmengen für alle $i \in I$, also eine Familie von Teilmengen. Dann setzen wir

(1) Schnitt:
$$\bigcap_{i\in I} U_i = \{a \in X \ ; \ a \in U_i \text{ für alle } i \in I\},$$

(2) Vereinigung:
$$\bigcup_{i\in I} U_i = \{a \in X \ ; \ \text{ es gibt ein } i \in I \text{ mit } a \in U_i\}.$$

Der Fall $I = \{1, 2\}$ führt zum vorher eingeführten Begriff des Schnitts und der Vereinigung. Im Falle von $I = \{1, 2, \ldots, n\}$ schreiben wir auch

$$\bigcap_{i=1}^{n} U_i = U_1 \cap \ldots \cap U_n = \{a \in X \ ; \ a \in U_i \text{ für alle } i = 1, \ldots, n\},$$
$$\bigcup_{i=1}^{n} U_i = U_1 \cup \ldots \cup U_n = \{a \in X \ ; \ \text{ es gibt ein } 1 \leq i \leq n \text{ mit } a \in U_i\}.$$

2.5. **Abbildungen von Mengen.** Nachdem wir Mengen behandelt haben, wollen wir Mengen miteinander in Beziehung setzen können. Das wird uns noch mehrmals mit komplizierteren Strukturen passieren.

Definition 2.31. Eine **Abbildung** von einer Menge X in eine Menge Y ist eine Vorschrift, die jedem Element $a \in X$ ein eindeutiges Element $b \in Y$ zuordnet. Ist f die Abbildung, so schreiben wir
$$f\colon X \to Y$$
und $f(a) \in Y$ für das Element, das f dem Element $a \in X$ zuordnet. Manchmal schreiben wir auch kurz $a \mapsto f(a)$. Wir nennen X den **Definitionsbereich** und Y den **Wertebereich** der Abbildung $f\colon X \to Y$.

Die Menge der Abbildungen von der Menge X nach der Menge Y bezeichnen wir mit
$$\mathrm{Abb}(X, Y) = \{f \ ; \ \ f\colon X \to Y \text{ ist eine Abbildung von Mengen}\}.$$

Beispiel 2.32. (1) Sei X eine Menge und Y eine Menge mit einem Element $* \in Y$. Dann gibt es die konstante Abbildung
$$c\colon X \to Y$$
die durch $c(x) = *$ für alle $x \in X$ definiert ist.

(2) Für jede Menge X gibt es die Abbildung $\mathrm{id}_X\colon X \to X$, die für alle $a \in X$ durch
$$\mathrm{id}_X(a) = a$$
definiert ist. Diese Abbildung wird **Identität** (auf der Menge X) genannt.

(3) Sei $U \subseteq X$ eine Teilmenge. Dann definieren wir die **Inklusion**(oder **Inklusionssabbildung**) durch
$$i\colon U \to X$$
$$i(a) = a.$$
Ein Spezialfall hiervon ist die Inklusion $i\colon \emptyset \to X$, der leeren Menge als Teilmenge $\emptyset \subseteq X$. Dies ist die einzige Abbildung $\emptyset \to X$, die es gibt.

(4) Seien $a_0, a_1, \ldots, a_n$ reelle Zahlen. Dann definert
$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$
eine Abbildung $f\colon \mathbb{R} \to \mathbb{R}$, eine **Polynomfunktion**. Über Polynome wird später noch zu reden sein.

(5) Sei X_i mit $i \in I$ eine Familie von Mengen. Ein I-Tupel von Elementen der Familie ist eigentlich eine Abbildung
$$a\colon I \to \bigcup_{i\in I} X_i$$
in die Vereinigung $X := \bigcup_{i\in I} X_i$ mit der Eigenschaft, daß für alle $i \in I$
$$a(i) \in X_i.$$
Das I-Tupel besteht dann aus
$$(a(i))_{i\in I} \in \prod_{i\in I} X_i.$$
Die Vereinigung X ist eine abstrakte Vereinigung, denn a priori sind die X_i ja nicht Teilmengen einer gegebenen Menge. Das gibt es nach Axiomatik der Mengenlehre, und man kann dann $X_i \subseteq X$ als Teilmenge begreifen.

(6) Man kann jedem Ort auf der Erdoberfläche seinen Längen- und Breitengrad zuordnen. Nur an den Polen hat man ein Problem. Die Nord- und Südpol haben keinen Längengrad. Diese Zuordnungsvorschrift liefert daher keine Abbildung.

(7) Wir wachsen mit der Vorstellung von der Zeit auf, daß es eine Abbildung Zeit gibt, die der physikalischen Welt (modern die Raumzeit) eine Zeitkoordinate gibt. Die Relativitätstheorie lehrt, daß es eine solche überall definierte Abbildung nicht gibt. Zeit hängt vom Beobachter ab. Die Frage, ob wenigstens jeder Beobachter für sich eine konsistente überall gültige Art der Zeitmessung haben kann, ist eine schwierige Frage an die globale topologische Gestalt und die semi-Riemannsche Geometrie der Massenverteilung des Universums.

Definition 2.33. Sei $f\colon X \to Y$ eine Abbildung von Mengen, und seien $U \subseteq X$, $V \subseteq Y$ Teilmengen.

(1) Das **Bild** von U unter f ist die Teilmenge von Y
$$f(U) := \{b \in Y \ ; \ \exists a \in U \text{ mit } b = f(a)\}.$$
Als Bild von f bezeichnet man die Teilmenge $f(X) \subseteq Y$.

(2) Das **Urbild** von V unter f ist die Teilmenge von X
$$f^{-1}(V) := \{a \in X \ ; \ f(a) \in V\}.$$
Besteht V nur aus einem Element $v \in V$, also $V = \{v\}$, so schreibt man kürzer
$$f^{-1}(v) := f^{-1}(\{v\}) = \{a \in X \ ; \ f(a) = v\}.$$

Das auftretende Symbol f^{-1} gehört nicht zu einer Abbildung $f^{-1}\colon Y \to X$ in die umgekehrte Richtung.

Die Definition 2.31 einer Abbildung enthält die undefinierte Terminologie *Vorschrift* und *zuordnen*. Die korrekte Definition einer Abbildung ist die folgende.

Definition 2.34. Eine **Abbildung** von einer Menge X nach einer Menge Y ist eine Teilmenge $\Gamma \subseteq X \times Y$, so daß für alle $a \in X$ genau ein $b \in Y$ existiert mit $(a, b) \in \Gamma$.

Wir schreiben trotzdem eine Abbildung f von X nach Y als $f\colon X \to Y$ und bezeichnen die Menge Γ (manchmal genauer Γ_f) als den **Graphen** von f.

Bemerkung 2.35. Den Graph Γ_f zu einer Abbildung $f\colon X \to Y$ aus dem ersten Abbildungsbegriff bekommt man als
$$\Gamma_f = \{(a, b) \in X \times Y \ ; \ b = f(a)\}.$$
Die Forderung der Eindeutigkeit des b zu einem a mit $(a, b) \in \Gamma_f$ besagt, daß f jedem a ein eindeutiges $b \in Y$ zuordnet. Die Forderung, daß jedes a in einem Tupel $(a, b) \in \Gamma_f$ auftritt, bedeutet, daß f auf ganz X definiert ist. Man sagt dann, daß f **wohldefiniert** ist.

Beispiel 2.36.

(1) Der Graph der Identität $\mathrm{id}_X \colon X \to X$ ist nichts anderes als die Diagonale
$$\Delta \subseteq X \times X.$$

(2) Sei $f \colon X \to Y$ eine Abbildung. Dann betrachten wir die Abbildung
$$F \colon X \to X \times Y,$$
die für alle $a \in X$ durch
$$F(a) = (a, f(a))$$
definiert ist. Dann ist das Bild von F nichts anderes als der Graph von f.

Definition 2.37. Die **Potenzmenge** einer Menge X ist die Menge $\mathscr{P}(X)$ aller Teilmengen:
$$\mathscr{P}(X) = \{U \ ; \ U \subseteq X\}.$$

Beispiel 2.38. Sei $f \colon X \to Y$ eine Abbildung von Mengen. Dann definieren das Bild unter f
$$\begin{aligned} f \colon \mathscr{P}(X) &\to \mathscr{P}(Y) \\ U &\mapsto f(U) \end{aligned}$$
und das Urbild
$$\begin{aligned} f^{-1} \colon \mathscr{P}(Y) &\to \mathscr{P}(X) \\ V &\mapsto f^{-1}(V) \end{aligned}$$
zwei Abbildungen von Mengen, die wir suggestiv wieder mit f bzw. f^{-1} bezeichnen.

Beispiel 2.39. Sei X eine Menge. Dann definieren wir zu einer Teilmenge $U \subseteq X$ die charakteristische Funktion $\mathbf{1}_U$ als Abbildung
$$\mathbf{1}_U \colon X \to \{0, 1\}$$
für alle $a \in X$ durch
$$\mathbf{1}_U(a) = \begin{cases} 1 & \text{wenn } a \in U, \\ 0 & \text{wenn } a \notin U. \end{cases}$$
Damit können wir nun die folgende Abbildung definieren:
$$\begin{aligned} \mathscr{P}(X) &\to \mathrm{Abb}(X, \{0, 1\}) \\ U &\mapsto \mathbf{1}_U \end{aligned}$$
Dies ist eine Bijektion. Den Nachweis überlassen wir als Übungsaufgabe.

2.6. **Eigenschaften von Abbildungen.**

Definition 2.40. Sei $f \colon X \to Y$ eine Abbildung von Mengen.

(1) Wir sagen f ist **injektiv**, wenn jedes Element von Y höchstens ein Mal als Bild vorkommt:
$$\forall a_1, a_2 \in X : \quad f(a_1) = f(a_2) \implies a_1 = a_2.$$
Eine injektive Abbildung wird **Injektion** genannt.

(2) Wir sagen f ist **surjektiv**, wenn jedes Element von Y mindestens ein Mal als Bild vorkommt:
$$\forall b \in Y : \exists a \in X : f(a) = b,$$
oder kürzer $f(X) = Y$. Eine surjektive Abbildung wird **Surjektion** genannt.

(3) Wir sagen f ist **bijektiv**, wenn jedes Element von Y genau ein Mal als Bild vorkommt:
$$\forall b \in Y : \exists^! a \in X : f(a) = b.$$

(Die Notation $\exists^!$ bedeutet „**genau ein**“.) Eine bijektive Abbildung wird **Bijektion** genannt.

Wir benutzen spezielle Pfeile, wenn wir betonen wollen, daß eine Abbildung

$$\text{injektiv ist: } \hookrightarrow,$$
$$\text{surjektiv ist: } \twoheadrightarrow,$$
$$\text{bijektiv ist: } \xrightarrow{\sim}.$$

Beispiel 2.41.

(1) Die Inklusion $i\colon U \to X$ einer Teilmenge $U \subseteq X$ ist eine injektive Abbildung mit Bild U.
(2) Für zwei Mengen X, Y ist die Projektion $\mathrm{pr}_1\colon X \times Y \to X$ auf den ersten Faktor definiert durch
$$\mathrm{pr}_1((a,b)) = a$$
für alle $a \in X$ und $b \in Y$. Wenn $Y \neq \emptyset$ gilt, dann ist die Projektion surjektiv. Analog gibt es die Projektion $\mathrm{pr}_2\colon X \times Y \to Y$ auf den zweiten Faktor.
(3) Die Multiplikationsabbildung $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ definiert durch $(x,y) \mapsto xy$ ist surjektiv.
(4) Die Abbildung $\mathbb{R} \to \mathbb{R}$ definiert durch $x \mapsto x^2$ ist weder injektiv noch surjektiv.
(5) Für jede Menge X ist die Identität id_X bijektiv.

Proposition 2.42. *bijektiv $\iff$ injektiv und surjektiv.*

Beweis. Das folgt sofort aus der Definition der Begriffe, denn

„genau ein“ $\iff$ „mindestens ein“ und „höchstens ein“. □

Definition 2.43. Die **Komposition** zweier Abbildungen von Mengen f und g ist definiert, wenn der Definitionsbereich von f mit dem Wertebereich von g übereinstimmt, etwa $f\colon Y \to Z$ und $g\colon X \to Y$. Die Komposition von f und g ist dann die Abbildung

$$f \circ g\colon X \to Z$$

definiert für alle $a \in X$ durch

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

Man nennt dann f und g komponierbar und spricht $f \circ g$ als „f **nach** g“, denn man wendet erst g und da**nach** f an.

Beispiel 2.44. Sei $f\colon X \to Y$ eine Abbildung von Mengen. Dann ist

$$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f.$$

Proposition 2.45. *Die Komposition von Abbildungen ist **assoziativ**: sind f, g und h komponierbar, etwa als*

$$X \xrightarrow{h} Y \xrightarrow{g} Z \xrightarrow{f} W,$$

dann sind f und $g \circ h$ sowie h und $f \circ g$ komponierbar und es gilt

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

Beweis. Die Behauptung über die Komponierbarkeit ist offensichtlich. Für alle $a \in X$ gilt überdies

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

Bemerkung 2.46. Die Komposition von Abbildungen ist nicht notwendig **kommutativ**, in der Regel

$$f \circ g \neq g \circ f,$$

was man schon allein daran sieht, daß die Komponierbarkeit von f mit g nichts mit der Komponierbarkeit von g mit f zu tun hat.

Definition 2.47. Die **Einschränkung** einer Abbildung $f\colon X \to Y$ auf eine Teilmenge $U \subseteq X$ ist die Abbildung

$$f|_U\colon U \to Y$$
$$f|_U(a) = f(a).$$

Mit der Inklusion $i\colon U \to X$ läßt sich die Einschränkung $f|_U$ als $f|_U = f \circ i$ schreiben.

Proposition 2.48. *Seien $f\colon Y \to Z$ und $g\colon X \to Y$ Abbildungen von Mengen.*
(1) Wenn f und g injektiv sind, dann ist auch $f \circ g$ injektiv.
(2) Wenn $f \circ g$ injektiv ist, dann ist auch g injektiv.
(3) Wenn f und g surjektiv sind, dann ist auch $f \circ g$ surjektiv.
(4) Wenn $f \circ g$ surjektiv ist, dann ist auch f surjektiv.

Beweis. (1) Wenn $a_1, a_2 \in X$ mit $f \circ g(a_1) = f \circ g(a_2)$, dann ist $f(g(a_1)) = f(g(a_2))$. Weil f injektiv ist, folgt $g(a_1) = g(a_2)$, und, weil g injektiv ist, auch $a_1 = a_2$. Damit ist $f \circ g$ injektiv.

(2) Wenn $a_1, a_2 \in X$ mit $g(a_1) = g(a_2)$, dann ist $f(g(a_1)) = f(g(a_2))$, also $f \circ g(a_1) = f \circ g(a_2)$. Weil $f \circ g$ injektiv ist, folgt $a_1 = a_2$. Damit ist g injektiv.

(3) Sei $c \in Z$ beliebig. Weil f surjektiv ist, gibt es ein $b \in Y$ mit $f(b) = c$. Weil g surjektiv ist gibt es ein $a \in X$ mit $g(a) = b$. Dann gilt

$$f \circ g(a) = f(g(a)) = f(b) = c.$$

Weil c beliebig war, folgt $f \circ g$ surjektiv.

(4) Sei $c \in Z$ beliebig. Da $f \circ g$ surjektiv ist, gibt es ein $a \in X$ mit $f \circ g(a) = c$. Dann ist mit $b = g(a)$

$$f(b) = f(g(a)) = f \circ g(a) = c,$$

somit f surjektiv, denn c war ja beliebig gewählt. □

Korollar 2.49. *Seien $f\colon Y \to Z$ und $g\colon X \to Y$ Abbildungen von Mengen. Wenn f und g bijektiv sind, dann ist auch $f \circ g$ bijektiv.*

Beweis. Das folgt sofort aus Proposition 2.42, denn $f \circ g$ ist nach Proposition 2.48 injektiv und surjektiv. □

Definition 2.50. Sei n eine natürliche Zahl. Eine Menge X hat die **Mächtigkeit** n, wenn es eine Bijektion

$$X \xrightarrow{\sim} \{1, 2, \ldots, n\}$$

gibt. Falls $n = 0$ so meinen wir eine Bijektion $X \to \emptyset$. Wir schreiben dann

$$|X| = n.$$

Bemerkung 2.51. Die subtile Frage, daß die Mächtigkeit einer endlichen Menge wohldefiniert ist, also aus einer Bijektion

$$\{1, 2, \ldots, n\} \xrightarrow{\sim} \{1, 2, \ldots, m\}$$

auf $n = m$ geschlossen werden kann, überlassen wir als Übungsaufgabe (modifizieren Sie die Bijektion, so daß n auf m abgebildet wird, schränken Sie dann auf das Komplement ein und verwenden Sie vollständige Induktion).

2.7. **Äquivalenzrelationen.** Wenn die Elemente einer Menge miteinander in Beziehung stehen, dann drückt man dies durch eine Relation aus. Es gibt verschiedene Arten von Relationen, die für uns wichtigste formalisiert Eigenschaften der Gleichheit.

Definition 2.52. Eine **(binäre) Relation** auf einer Menge X ist eine Teilmenge
$$R \subseteq X \times X.$$

Bemerkung 2.53. Eine Relation R hebt eine Teilmenge der Tupel (a, b) mit $a, b \in X$ hervor. Dies sind die Paare, die bezüglich R „zueinander in Relation stehen".

Beispiel 2.54. $X = \{\text{Schere, Stein, Papier}\}$ und
$$R = \{(a, b)\ ;\ \ a \text{ gewinnt gegen } b \text{ beim Spiel „Schere, Stein, Papier"}\}.$$

Definition 2.55. Eine **Äquivalenzrelation** auf einer Menge X ist eine Relation R auf X mit den folgenden Eigenschaften, die sich am besten in der suggestive Notation $\sim$ für die Relation definiert durch
$$a \sim b :\iff (a, b) \in R$$
formulieren lassen.

(i) **reflexiv:** für alle $x \in X$ gilt $x \sim x$.
(ii) **symmetrisch:** für alle $x, y \in X$ gilt: $x \sim y \iff y \sim x$.
(iii) **transitiv:** für alle $x, y, z \in X$ gilt: $x \sim y$ und $y \sim z$, dann auch $x \sim z$.

Notation 2.56. Eine Äquivalenzrelation ist eine Relation, welche gewisse Eigenschaften der Gleichheit formalisiert. Die Schreibweise $a \sim b$ anstelle von $(a, b) \in R$ erinnert daran. Andere übliche Symbole für eine Äquivalenzrelation sind $\equiv, \simeq, \cong, \ldots$.

Beispiel 2.57. (1) Die Relation „Schere, Stein, Papier" aus Beispiel 2.54 ist weder reflexiv noch symmetrisch noch transitiv.

(2) Gleichheit von Elementen von X ist eine Äquivalenzrelation. Diese Relation wird durch die Diagonale $\Delta \subseteq X \times X$ dargestellt und $a = b \iff (a, b) \in \Delta$ geschrieben.

(3) Sei n eine ganze Zahl. Eine wichtige Äquivalenzrelation auf der Menge der ganzen Zahlen
$$\mathbb{Z} = \{\ldots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \ldots\}$$
ist die Relation **kongruent modulo** n definiert durch
$$a \equiv b \pmod{n} :\iff \exists x \in \mathbb{Z} : a - b = n \cdot x.$$

- Reflexiv: $a - a = n \cdot 0$.
- Symmetrisch: wenn $a \equiv b \pmod{n}$, dann gibt es $x \in \mathbb{Z}$ mit $a - b = nx$, somit auch $b - a = n(-x)$ und damit $b \equiv a \pmod{n}$.
- Transitiv: wenn $a \equiv b \pmod{n}$ und $b \equiv c \pmod{n}$, dann gibt es $x, y \in \mathbb{Z}$ mit $a - b = nx$ und $b - c = ny$. Daraus folgt
$$a - c = (a - b) + (b - c) = nx + ny = n(x + y),$$
und weil $x + y \in \mathbb{Z}$, folgt $a \equiv c \pmod{n}$.

Bemerkung 2.58. Ein subtiler Fehler ist in der folgenden Argumentation versteckt. Behauptung: aus den Eigenschaften symmetrisch und transitiv einer Relation $\sim$ folgt bereits die Eigenschaft reflexiv. Denn für ein x und irgendein y mit $x \sim y$ folgt aus Symmetrie $y \sim x$. Transitivität zeigt aus $x \sim y$ und $y \sim x$ bereits $x \sim x$, fertig.

Der Fehler besteht darin, ohne Begründung für seine Existenz das Element y aus dem Hut gezaubert zu haben. Wenn es ein solches y nicht gibt, dann funktioniert das Argument nicht. Aus diesem Beispiel sollte man lernen, daß die Existenz von Elementen mit gewissen Eigenschaften stets zu begründen ist.

Definition 2.59. Sei $\sim$ eine Äquivalenzrelation auf der Menge X. Die **Äquivalenzklasse** zu einem $a \in X$ bezüglich $\sim$ ist die Teilmenge

$$[a] = \{b \in X \ ; \ a \sim b\} \subseteq X.$$

Beispiel 2.60. Auf der Menge der Schüler einer Schule definieren wir die Relation „gehen in die gleiche Klasse“. Das ist offenbar eine Äquivalenzrelation und die zugehörigen Äquivalenzklassen sind die Schulklassen dieser Schule.

Bemerkung 2.61 (Ringschluß). In der folgenden Proposition beweisen wir die Äquivalenz von mehreren Bedingungen per **Ringschluß**. Anstelle aller 12 Implikation $(x) \implies (y)$ für $x, y \in a, b, c, d$, $x \neq y$, zeigen wir nur die folgenden 4 Implikationen.

$$\begin{array}{ccc} (a) & \Longrightarrow & (b) \\ \Uparrow & & \Downarrow \\ (d) & \Longleftarrow & (c) \end{array}$$

Das ist ökonomischer und beweist auch alle Implikationen!

Proposition 2.62. *Sei $\sim$ eine Äquivalenzrelation auf der Menge X. Dann gilt:*

(1) *Für alle $a \in X$ ist $a \in [a]$.*

(2) *Für alle $a, b \in X$ sind äquivalent:*

- *(a)* $a \sim b$,
- *(b)* $b \in [a]$,
- *(c)* $[a] \cap [b] \neq \emptyset$,
- *(d)* $[a] = [b]$.

Beweis. (1) Zu zeigen ist $a \sim a$, und das gilt, weil $\sim$ reflexiv ist.

(2) Die Aussagen (a) und (b) sind äquivalent per Definiton von $[a]$. Wir zeigen nun

$$(b) \implies (c) \implies (d) \implies (a).$$

Nach dem Prinzip des Ringschluß sind wir dann fertig.

$(b) \implies (c)$: Aus (1) wissen wir auch $b \in [b]$, also gilt $b \in [a] \cap [b]$ und damit ist der Schnitt nicht leer.

$(c) \implies (d)$: Sei $[a] \cap [b] \neq \emptyset$. Dann gibt es ein $c \in [a] \cap [b]$. Per Definition einer Äquivalenzklasse gilt dann

$$a \sim c \quad \text{und} \quad b \sim c.$$

Wir zeigen nun $[a] \subseteq [b]$. Das analoge symmetrische Argument, in dem a und b die Rollen tauschen, zeigt dann $[b] \subseteq [a]$ und damit die Behauptung (d).

Sei $x \in [a]$, also $a \sim x$. Dann gelten $b \sim c$, und $c \sim a$ (wegen Symmetrie aus $a \sim c$), und $a \sim x$. Zweimaliges Anwenden von Transitivität macht daraus $b \sim x$. Damit gilt $x \in [b]$. Dies zeigt $[a] \subseteq [b]$.

$(d) \implies (a)$: Aus (1) wissen wir, daß $b \in [b]$ gilt. Wegen $[a] = [b]$ folgt $b \in [a]$, und das bedeutet per Definitiion $a \sim b$. □

Korollar 2.63. *Sei $\sim$ eine Äquivalenzrelation auf der Menge X. Dann gilt: Zwei Äquivalenzklassen sind gleich oder disjunkt: Für alle $a, b \in X$ gilt*

$$\textit{entweder} \quad [a] = [b] \quad \textit{oder} \quad [a] \cap [b] = \emptyset.$$

Beweis. Das folgt sofort aus Proposition 2.62 (2), genauer $(c) \iff (d)$. □

Definition 2.64. Sei X eine Menge mit einer Äquivalenzrelation $\sim$.

(1) Wir bezeichnen die **Menge der Äquivalenzklassen** von $\sim$ mit

$$X/_\sim := \{A \subseteq X \ ; \ \exists a \in X \text{ mit } A = [a]\}.$$

Dies ist eine Teilmenge der Potenzmenge von X.

(2) Die **Quotientenabbildung** der Relation ist die Abbildung

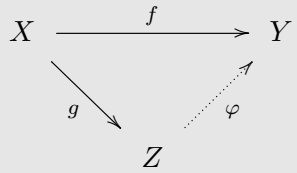
$$\begin{aligned} [\,]\colon X &\to X/_\sim \\ a &\mapsto [a], \end{aligned}$$

die jedem Element seine Äquivalenzklasse zuordnet. Die Quotientenabbildung ist offensichtlich surjektiv.

(3) Jedes Element einer Äquivalenzklasse heißt **Repräsentant** (dieser Äquivalenzklasse. Die Repräsentanten von $[x]$ sind genau die y mit $x \sim y$, also die Elemente des Urbilds von $[\,]^{-1}([x])$.

(4) Eine Teilmenge $T \subseteq X$, die zu jeder Äquivalenzklasse von $\sim$ genau einen Repräsentanten enthält, nennt man ein **(vollständiges) Repräsentantensystem** (von $\sim$ auf X).

Definition 2.65. Seien $f : X \to Y$ und $g : X \to Z$ Abbildungen. Man sagt **„f faktorisiert über g“**, wenn es eine Abbildung $\varphi : Z \to Y$ gibt mit $f = \varphi \circ g$. Visuell bedeutet das die Existenz des gepunkteten Pfeils im **kommutativen Diagramm**

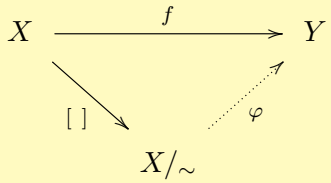


Das Diagramm von Abbildungen heißt dabei **kommutativ**, wenn bei Komposition der Abbildungen entlang eines Weges in Pfeilrichtung durch das Diagramm jeweils dieselbe Abbildung herauskommt, d.h. unabhängig vom gewählten Weg durch das Diagramm. Hier sind das nur zwei Wege und die Bedingung gerade $f = \varphi \circ g$.

Proposition 2.66 (Faktorisieren über die Quotientenabbildung)**.** *Sei X eine Menge mit einer Äquivalenzrelation $\sim$, und sei $f\colon X \to Y$ eine Abbildung.*

Die folgenden zwei Aussagen sind äquivalent.

(a) f faktorisiert über die Quotientenabbildung $[\,]$: Es gibt eine Abbildung $\varphi\colon X/_\sim \to Y$ mit $f = \varphi \circ [\,]$.



(b) f ist auf Äquivalenzklassen von $\sim$ konstant: $\forall\, x, x' \in X$ mit $x \sim x'$ gilt $f(x) = f(x')$.

Wenn die äquivalenten Aussagen (a) und (b) gelten, dann ist die Abbildung φ eindeutig festgelegt.

Beweis. Wir zeigen zuerst (a) $\Longrightarrow$ (b). Seien $x, x' \in X$ mit $x \sim x'$. Dann gilt $[x] = [x']$ nach Proposition 2.62, und somit

$$f(x) = \varphi([x]) = \varphi([x']) = f(x').$$

Dies zeigt Aussage (b).

Jetzt zeigen wir umgekehrt (b) $\Longrightarrow$ (a). Dazu müssen wir eine Abbildung

$$\varphi\colon X/_\sim \to Y$$

konstruieren. Jedes Element $c \in X/_\sim$ ist von der Form $[x]$ für ein $x \in X$. Wir definieren

$$\varphi(c) := f(x) \qquad \text{für } x \in X \text{ mit } c = [x].$$

Im Allgemeinen ist x nicht eindeutig. Damit φ wohldefiniert ist, müssen wir nachweisen, daß unabhängig von der Wahl von x bei unserer Vorschrift, die φ definiert, stets dasselbe Ergebnis herauskommt. Dazu vergleichen wir zwei Wahlen. Seien also $x, x' \in X$ Elemente mit $[x] = c = [x']$. Dann gilt nach Proposition 2.62 $x \sim x'$. Aus (b) folgt dann

$$f(x) = f(x').$$

Damit ist φ wohldefiniert. Nun weisen wir die Faktorisierungseigenschaft nach: zu $x \in X$ gilt

$$\varphi([x]) = f(x),$$

weil $x \in [x]$ eine der möglichen Wahlen von Elementen von X ist, mit der φ per Definition bestimmt werden kann. Die Faktorisierungseigenschaft war also in der Definition von φ bereits angelegt.

Zuletzt müssen wir noch zeigen, daß φ eindeutig ist, sofern φ existiert. Sei dazu $\psi\colon X/_\sim \to Y$ eine weitere Abbildung, welche die Eigenschaften von φ in (a) besitzt. Dann gilt für alle $c \in X/_\sim$, mit der Wahl von $x \in c$ also $c = [x]$, daß

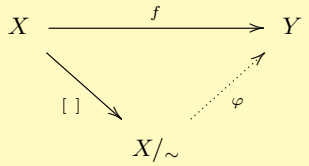
$$\varphi(c) = \varphi([x]) = f(x) = \psi([x]) = \psi(c).$$

Damit ist $\varphi = \psi$, und φ ist in der Tat eindeutig. □

Proposition 2.67 (Universelle Eigenschaft der Quotientenabbildung)**.** *Sei X eine Menge mit einer Äquivalenzrelation $\sim$. Dann hat die Quotientenabbildung $[\ ]\colon X \to X/_\sim$ die folgende universelle[a] Eigenschaft.*

(i) *Für alle $x, x' \in X$ gilt:*
$$x \sim x' \quad \Longrightarrow \quad [x] = [x'].$$

(ii) *Für alle Abbildungen $f : X \to Y$ gilt: wenn für alle $x, x' \in X$ mit $x \sim x'$ gilt, daß $f(x) = f(x')$ ist, dann faktorisiert f eindeutig durch $[\ ]$, d.h. es gibt eine eindeutige Abbildung $\varphi\colon X/_\sim \to Y$ mit $f = \varphi \circ [\ ]$.*



[a]Eine universelle Eigenschaft eines mathematischen Objekts charakterisiert diese Objekt unter den möglichen Kandidaten seiner Kategorie. Das verstehen Sie später.

Beweis. (i) wurde bereits in Proposition 2.62 gezeigt, und (ii) ist ein Teil von Proposition 2.66. □

Beispiel 2.68 (Modulare Arithmetik). In guten Fällen übertragen sich Strukturen auf die Menge der Äquivalenzklassen.

Sei $n \in \mathbb{Z}$. Die Menge der Äquivalenzklassen der Relation „kongruent modulo n“ auf $\mathbb{Z}$ wird mit

$$\mathbb{Z}/n\mathbb{Z}$$

bezeichnet (gelesen: $\mathbb{Z}$ modulo $n\mathbb{Z}$). Eine solche Äquivalenzklasse wird auch **Restklasse modulo** n genannt. Wir betrachten als Struktur die Addition und die Multiplikation der ganzen Zahlen

$\mathbb{Z}$. Wir definieren für alle $a, b \in \mathbb{Z}$ und die zugehörigen Restklassen $[a], [b] \in \mathbb{Z}/n\mathbb{Z}$ eine Addition und eine Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$ vermöge

$$[a] + [b] := [a + b],$$
$$[a] \cdot [b] := [a \cdot b].$$

Man sagt, Addition und Multiplikation werden auf Repräsentanten definiert. Vergleichen Sie dies mit der Definition der Abbildung φ im Beweis von Proposition 2.67.

Wir müssen nachweisen, daß diese Definition wohldefiniert ist! Dazu verfahren wir wie im Beweis von Proposition 2.67. Es ist zu zeigen, daß Addition und Multiplikation unabhängig von der Wahl der Repräsentanten der Restklassen ist: für alle $a, a', b, b' \in \mathbb{Z}$ mit $[a] = [a']$ und $[b] = [b']$ müssen wir $[a + b] = [a' + b']$ und $[a \cdot b] = [a' \cdot b']$ zeigen. Nach Voraussetzung gibt es $x, y \in \mathbb{Z}$ mit

$$a - a' = nx \qquad \text{und} \qquad b - b' = ny.$$

Dann gilt $[a + b] = [a' + b']$, weil

$$(a + b) - (a' + b') = (a - a') + (b - b') = nx + ny = n(x + y),$$

und es gilt $[ab] = [a'b']$, weil

$$ab - a'b' = ab - a'b + a'b - a'b' = (a - a')b + a'(b - b') = nxb + a'ny = n(xb + a'y).$$

Dies zeigt die Behauptung.

Die bekannten Rechengesetzte von $\mathbb{Z}$ übertragen sich sofort auf $\mathbb{Z}/n\mathbb{Z}$: assoziativ, kommutativ, distributiv, Null $[0]$, additives Inverses, Eins $[1]$. Die Details überlassen wir zur Übung. In der Sprache von Kapitel 3 haben wir $\mathbb{Z}/n\mathbb{Z}$ mit der Struktur eines Rings versehen.

Bemerkung 2.69. Sei $n \in \mathbb{N}$. Wir betonen ausdrücklich, daß $\mathbb{Z}/n\mathbb{Z}$ aus den n-vielen Restklassen

$$[0], [1], \ldots, [n - 1]$$

und nicht aus den Elementen

$$0, 1, \ldots, n - 1$$

besteht. Letzteres beschreibt ein vollständiges Repräsentatensystem: jede Restklasse enthält genau eine dieser Zahlen. Rechnungen in $\mathbb{Z}/n\mathbb{Z}$ lassen sich zwar mittels dieser Repräsentanten in $\mathbb{Z}$ durchführen, trotzdem sind die Elemente von $\mathbb{Z}/n\mathbb{Z}$ nicht gleich diesen Vertretern, sondern deren Klassen!

Übungsaufgaben zu §2

Übungsaufgabe 2.1. Alice bietet Bob die folgende Wette an: „Ich wette um 1€, daß ich Dir 10€ schenke, wenn Du mir 5€ schenkst."

Soll Bob die Wette annehmen?

Übungsaufgabe 2.2. Für welche Mengen X is die Diagonale $\Delta \subseteq X \times X$ keine echte Teilmenge?

Übungsaufgabe 2.3. Seien U, V und W Teilmengen von X. Zeigen Sie die folgenden Identitäten und Aussagen.

(1) $U \cap V \subseteq U$,
(2) $U \subseteq U \cup V$,
(3) $(U \cup V) \cup W = U \cup (V \cup W)$,
(4) $(U \cap V) \cap W = U \cap (V \cap W)$,
(5) $U \subseteq V$ und $V \subseteq W \implies U \subseteq W$,
(6) $U \subseteq V \implies U \cap W \subseteq V \cap W$,
(7) $U \subseteq V \implies X \setminus V \subseteq X \setminus U$,
(8) $X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V)$,
(9) $X \setminus (U \cup V) = (X \setminus U) \cap (X \setminus V)$,

(10) $X \setminus (X \setminus U) = U$,
(11) $(U \cap V) \cup W = (U \cup W) \cap (V \cup W)$,
(12) $(U \cup V) \cap W = (U \cap W) \cup (V \cap W)$.

Übungsaufgabe 2.4. Seien $f\colon Y \to Z$ und $g\colon X \to Y$ Abbildungen von Mengen. Finden Sie jeweils ein Gegenbeispiel für die folgende Aussage:

(1) Wenn $f \circ g$ injektiv ist, dann ist auch f injektiv.
(2) Wenn $f \circ g$ surjektiv ist, dann ist auch g surjektiv.

Übungsaufgabe 2.5. Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen von Mengen. Seien $A, B \subseteq X$ und $U, V \subseteq Y$ und $W \subseteq Z$ Teilmengen. Zeigen Sie die folgenden Aussagen.

(1) $A \subseteq B \implies f(A) \subseteq f(B)$
(2) $f(A \cup B) = f(A) \cup f(B)$,
(3) $f(A \cap B) \subseteq f(A) \cap f(B)$,
(4) finden Sie ein Gegenbeispiel zu $f(A \cap B) = f(A) \cap f(B)$,
(5) $f(X) \setminus f(A) \subseteq f(X \setminus A)$,
(6) $(g \circ f)(A) = g(f(A))$,
(7) $U \subseteq V \implies f^{-1}(U) \subseteq f^{-1}(V)$,
(8) $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$,
(9) $f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$,
(10) $f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V)$,
(11) $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$,
(12) $f(f^{-1}(U)) = U \cap f(X)$,
(13) $A \subseteq f^{-1}(f(A))$.

Übungsaufgabe 2.6. Seien $f\colon X \to Y$ und $g\colon X \to Y$ Abbildungen von Mengen.

(1) Sei $\varphi : Y \hookrightarrow Z$ eine injektive Abbildung. Zeigen Sie

$$\varphi \circ f = \varphi \circ g \quad \implies \quad f = g.$$

(2) Sei $\psi : Z \twoheadrightarrow X$ eine surjektive Abbildung. Zeigen Sie

$$f \circ \psi = g \circ \psi \quad \implies \quad f = g.$$

Übungsaufgabe 2.7 (Universelle Eigenschaft des Produkts). Seien X_1, X_2 Mengen und bezeichne $\mathrm{pr}_i\colon X_1 \times X_2 \to X_i$ die Projektion auf den i-ten Faktor, $i = 1, 2$. Zeigen Sie:

(1) Zu jeder Menge T und zwei Abbildungen $f_i\colon T \to X_i$ von Mengen, $i = 1, 2$, gibt es **genau eine** Abbildung von Mengen

$$F\colon T \to X_1 \times X_2$$

mit $f_i = \mathrm{pr}_i \circ F$ für $i = 1, 2$.

Diese Abbildung bezeichnen wir mit (f_1, f_2), und wir haben sie bei der Beschreibung des Graphen als Bild in einem Spezialfall bereits benutzt.

(2) Zu jeder Menge T ist

$$\begin{aligned} \mathrm{Abb}(T, X_1 \times X_2) &\to \mathrm{Abb}(T, X_1) \times \mathrm{Abb}(T, X_2) \\ F &\mapsto (\mathrm{pr}_1 \circ F, \mathrm{pr}_2 \circ F) \end{aligned}$$

eine Bijektion von Mengen. Schreiben Sie die inverse Abbildung hin und rechnen Sie nach, daß es sich tatsächlich um das Inverse handelt.

(3) Sei P eine Menge mit Abbildungen $\pi_i\colon P \to X_i$ für $i = 1, 2$, so daß die Abbildung

$$\begin{aligned} \mathrm{Abb}(T, P) &\to \mathrm{Abb}(T, X_1) \times \mathrm{Abb}(T, X_2) \\ F &\mapsto (\pi_1 \circ F, \pi_2 \circ F) \end{aligned}$$

für alle Mengen T eine Bijektion ist, dann gibt es eine Bijektion

$$\Phi\colon P \xrightarrow{\sim} X_1 \times X_2,$$

welche durch die Eigenschaft

$$\mathrm{pr}_i \circ \Phi = \pi_i$$

für $i = 1, 2$, eindeutig bestimmt wird.

Übungsaufgabe 2.8. Eine Partition einer Menge X ist eine Zerlegung $X = \bigcup_{i\in I} U_i$ mit paarweise disjunkten Teilmengen $U_i \subseteq X$: für $i \neq j$ gilt $U_i \cap U_j = \emptyset$.

(a) Zeigen Sie, daß durch

$$a \sim b :\iff \exists i \text{ mit } a, b \in U_i$$

eine Äquivalenzrelation auf X definiert wird.

(b) Zeigen Sie, daß jede Äquivalenzrelation in dieser Form durch eine Partition definiert werden kann. Die Partition ist im übrigen eindeutig. Welche ist es?

Übungsaufgabe 2.9. Sei $f\colon X \to Y$ eine Abbildung.

(a) Zeigen Sie, daß durch

$$a \sim b :\iff f(a) = f(b)$$

eine Äquivalenzrelation auf X definiert wird.

(b) Zeigen Sie, daß jede Äquivalenzrelation in dieser Form durch eine surjektive Abbildung $X \to Y$ für ein geeignetes Y definiert werden kann.

Übungsaufgabe 2.10. Sei $R \subseteq X \times X$ eine Relation auf X.

(a) Beschreiben Sie, was die Eigenschaften reflexiv und transitiv für die Teilmenge R bedeuten.

(b) Zu Abbildungen $f_1\colon A_1 \to B$ und $f_2\colon A_2 \to B$ bezeichnen wir mit

$$A_1 \times_B A_2 := A_1 \times_{f_1,B,f_2} A_2 := \{(a_1, a_2)\ ;\ f_1(a_1) = f_2(a_2)\}$$

das Faserprodukt der Mengen A_1, A_2 via f_1, f_2 über B.

Die Projektion $\mathrm{pr}_i\colon X \times X \to X$ auf den $i = 1, 2$-ten Faktor liefert ebenso bezeichnete Abbildungen $\mathrm{pr}_i\colon R \to X$. Es gibt dann eine Abbildung

$$c\colon R \times_{\mathrm{pr}_2,X,\mathrm{pr}_1} R \to X \times X$$

durch

$$c((x, y), (y, z)) = (x, z).$$

Beschreiben Sie, was die Eigenschaften transitiv für die Teilmenge R mittels der Abbildung c bedeutet.

Übungsaufgabe 2.11. Sei $R \subseteq X \times X$ eine Äquivalenzrelation auf X, und sei $f : X \to Y$ die Quotientenabbildung. Zeigen Sie, daß gilt:

$$R = X \times_{f,Y,f} X.$$

Übungsaufgabe 2.12. Zwei Mengen X und Y heißen **gleichmächtig**, wenn es eine Bijektion $X \xrightarrow{\sim} Y$ zwischen ihnen gibt. Vorsicht: das bedeutet nicht, daß es sich um endliche Mengen handelt. Sind die Mengen X und Y gleichmächtig, dann schreiben wir $|X| = |Y|$.

Zeigen Sie für drei Mengen X, Y und Z:

(i) $|X| = |X|$,

(ii) $|X| = |Y| \iff |Y| = |X|$,

(iii) $|X| = |Y|$ und $|Y| = |Z| \implies |X| = |Z|$.

Übungsaufgabe 2.13. Seien X und Y Mengen mit endlich vielen Elementen, und sei $f\colon X \to Y$ eine Abbildung. Zeigen Sie: wenn $|X| = |Y|$, dann sind äquivalent:

(a) f ist bijektiv.

(b) f ist injektiv.

(c) f ist surjektiv.

3. Gruppen, Ringe und Körper

3.1. **Zahlbereiche.** Grundlegende Zahlbereiche setzen wir in dieser Vorlesung zunächst als bekannt voraus. Wir skizzieren in diesem Abschnitt, wie man die Zahlbereiche $\mathbb{N}_0$, $\mathbb{Z}$ und $\mathbb{Q}$ mathematisch konstruiert.

3.1.1. *Die natürlichen Zahlen.*

Definition 3.1 (Peano Axiome)**.** Die **natürlichen Zahlen mit** 0 sind eine Menge, bezeichnet mit $\mathbb{N}_0$, die den **Peano**-Axiomen genügt:

(i) Es gibt ein Element $0 \in \mathbb{N}_0$, genannt **Null**.
(ii) Jedes $n \in \mathbb{N}_0$ hat einen **Nachfolger** in $\mathbb{N}_0$, den wir mit $n^* \in \mathbb{N}_0$ bezeichnen. Das ist eine Abbildung
$$\mathbb{N}_0 \to \mathbb{N}_0, \qquad n \mapsto n^*.$$
(iii) Die 0 hat keinen Vorgänger, d.h. 0 ist kein Nachfolger: $\forall n \in \mathbb{N}_0 : \ 0 \neq n^*$.
(iv) Jedes $n \in \mathbb{N}_0$, $n \neq 0$ hat einen eindeutigen Vorgänger, d.h. die Nachfolgerabbildung ist injektiv: wenn $m_1, m_2 \in \mathbb{N}_0$, dann folgt aus $m_1^* = m_2^*$ schon $m_1 = m_2$.
(v) Induktionsaxiom: gilt für eine Teilmenge $N \subseteq \mathbb{N}_0$
- $0 \in N$,
- wenn $n \in N$, dann ist auch $n^* \in N$,

dann gilt bereits $N = \mathbb{N}_0$.

Um natürliche Zahlen zu konstruieren, die den Peano-Axiomen genügen, muß man nun eine geeignete Menge $\mathbb{N}_0$ mit einer $0 \in \mathbb{N}_0$ und einem Begriff des Nachfolgers konstruieren. Man sagt, man konstruiert ein Modell der Axiome. Dazu beginnen wir mit der leeren Menge und bezeichnen sie mit
$$0 := \emptyset.$$
Dann definieren wir
$$0^* := 0 \cup \{0\} = \{\emptyset\}$$
und nennen dies 1. Dann konstruieren wir iteriert weiter, wenn wir eine Menge n konstruiert haben
$$n^* := n \cup \{n\}.$$
Schließlich setzen wir $\mathbb{N}_0$ als die Menge, welche alle Mengen $0, 1, 2, 3, \ldots$ enthält, konkret (aber unübersichtlich)
$$\mathbb{N}_0 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots\}$$
Damit haben wir alles beisammen und in der Tat gelten die Peano-Axiome.

Bemerkung 3.2. Dies ist eine Skizze, weil wir nichts bewiesen haben. Schlimmer noch: wir haben nicht diskutiert, was es heißt, iteriert etwas zu konstruieren, denn dieser Begriff scheint die natürlichen Zahlen und das Induktionsaxiom zu benutzen, womit sich die Katze in den Schwanz beißt. Das formal korrekt zu machen ist Aufgabe einer Vorlesung über Mengenlehre.

Bemerkung 3.3. Ich hoffe, es ist selbstverständlich, daß sich niemand die natürlichen Zahlen auf diese umständliche Weise vorstellen soll. Wir wissen eigentlich (und auf jeden Fall für das Lernziel dieser Vorlesung ausreichend), was die natürlichen Zahlen sind. Es ging nur darum zu zeigen, mit welchen Begrifflichkeiten man bei Fragen nach den Grundlagen der Mathematik umzugehen hat.

Bemerkung 3.4. Man kann ebenfalls rekursiv die Addition und Multiplikation auf den natürlichen Zahlen definieren:

(1) Die Addition (Summe) ist definiert durch $m + 0 = m$ und
$$m + n^* = (m + n)^*.$$
(2) Die Multiplikation (das Produkt) ist definiert durch $m \cdot 0 = 0$ und
$$m \cdot n^* = m \cdot n + m.$$

Sei $m \in \mathbb{N}_0$ fest und $N \subseteq \mathbb{N}_0$ die Menge der $n \in \mathbb{N}_0$, für die die obige Definition die Addition $n + m$ definiert. Dann liest man aus der Definition ab, daß $0 \in N$ und mit $n \in N$ wissen wir auch $m + n^*$, also $n^* \in N$. Das Induktionsaxiom zeigt damit $N = \mathbb{N}_0$ und somit ist $m + n$ für alle n (und alle m) definiert.

Genauso zeigt man, daß alle Produkte $m \cdot n$ definiert sind. Außerdem kann man leicht zeigen, daß die so definierte Addition und Multiplikation auf $\mathbb{N}_0$ die gewohnten Rechengesetze erfüllen.

Beispiel 3.5. Es gilt $n + 1 = n + 0^* = (n + 0)^* = n^*$.

Bemerkung 3.6. Das Induktionsaxiom begründet, warum der Beweis per vollständige Induktion funktioniert, und axiomatisiert so eine intuitiv verwendete Eigenschaft der natürlichen Zahlen.

Angenommen wir wollen eine Familie A_n von Aussagen für alle $n \in \mathbb{N}_0$ beweisen. Sei $N \subseteq \mathbb{N}_0$ die Menge der natürlichen Zahlen n, für die die Aussage A_n richtig ist. Wir zeigen nun

- **Induktionsanfang**: $0 \in N$,
- **Induktionsschritt**: wenn $n \in N$, dann auch $n + 1 \in N$,

und schließen aus dem **Peanoschen Induktionsaxiom** auf $N = \mathbb{N}_0$. Damit sind alle Aussagen A_n bewiesen!

3.1.2. *Die ganzen Zahlen.* Wir konstruieren nun die ganzen Zahlen aus den natürlichen Zahlen.

(1) Wir definieren auf der Menge $\mathbb{N}_0 \times \mathbb{N}_0$ die Relation

$$(a,b) \sim (c,d) : \Longleftrightarrow \; a+d = b+c.$$

Dies ist eine Äquivalenzrelation:

- Reflexiv: $(a,b) \sim (a,b)$ weil $a+b = b+a$.
- Symmetrisch: wenn $(a,b) \sim (c,d)$, dann $a+d = b+c$, also auch $c+b = d+a$ und das bedeutet $(c,d) \sim (a,b)$.
- Transitiv: wenn $(a,b) \sim (c,d)$ und $(c,d) \sim (e,f)$, dann gilt $a+d = b+c$ und $c+f = d+e$. Addieren wir diese Gleichungen, so erhalten wir

$$a+(d+c)+f = b+(c+d)+e.$$

Da auf $\mathbb{N}_0$ die Nachfolgerfunktion injektiv ist, zeigt man per Induktion, daß Addition von $(c+d)$ als Abbildung $\mathbb{N}_0 \to \mathbb{N}_0$, also $x \mapsto x+(c+d)$ injektiv ist. Wir schließen so auf

$$a+f = b+e,$$

und das bedeutet $(a,b) \sim (e,f)$.

(2) Die Menge der Äquivalenzklassen bezeichnen wir mit $\mathbb{Z}$. Die Zuordnung

$$\begin{aligned} \mathbb{N}_0 &\to \mathbb{Z} \\ n &\mapsto [(n,0)] \end{aligned}$$

ist injektiv (wenn $(n,0) \sim (m,0)$, dann ist $n = n+0 = m+0 = m$). Wir nehmen uns die Freiheit und bezeichnen $[(n,0)]$ wieder wie gewohnt mit n. Die Zuordnung

$$\begin{aligned} \mathbb{N}_0 &\to \mathbb{Z} \\ n &\mapsto [(0,n)] \end{aligned}$$

ist injektiv (wenn $(0,n) \sim (0,m)$, dann ist $m = 0+m = 0+n = n$). Wir bezeichnen $[(0,n)]$ wie gewohnt mit $-n$. Dann gilt für $n,m \in \mathbb{N}_0$

$$n = -m \Longrightarrow (n,0) \sim (0,m) \Longrightarrow n+m = 0+0 = 0,$$

und das geht nur für $n = m = 0$, weil 0 kein Nachfolger in $\mathbb{N}_0$ ist. Es gilt also $0 = -0$, aber $n \neq -m$ in allen anderen Fällen.

(3) Jedes Element von $\mathbb{Z}$ ist von der Form n oder $-n$ für ein eindeutiges $n \in \mathbb{N}_0$, wie man leicht sieht. Wir identifizieren $\mathbb{N}_0$ mit einer Teilmenge von $\mathbb{Z}$ durch $n \mapsto n$.

(4) Auf $\mathbb{Z}$ definieren wir durch

$$[(a,b)] + [(c,d)] := [(a+c, b+d)]$$

eine Addition.

- Diese Addition ist wohldefiniert: das geht wie im Beweis von Proposition 2.67 oder durch zweimalige Anwendung von Proposition 2.67. Übung!
- Für alle $n,m \in \mathbb{N}_0$ gilt

$$n+m = [(n,0)] + [(m,0)] = [(n+m,0)] = n+m.$$

Die Addition in $\mathbb{Z}$ setzt also die Addition von $\mathbb{N}_0$ fort.
- Die $0 = [(0,0)]$ ist das neutrale Element bezüglich der Addition. Das ist klar.
- Die Addition ist assoziativ. Das ist auch klar.
- Die Addition ist kommutativ. Das ist ebenso klar, weil es in $\mathbb{N}_0$ gilt.
- Jedes $[(a,b)] \in \mathbb{Z}$ hat ein Inverses bezüglich der Addition, nämlich $[(b,a)]$:

$$[(a,b)] + [(b,a)] = [(a+b, b+a)] = [(0,0)] = 0.$$

(5) Damit gibt es in $\mathbb{Z}$ auch Subtraktion als Additon mit dem Inversen. Wir rechnen

$$a-b = [(a,0)] - [(b,0)] = [(a,0)] + [(0,b)] = [(a,b)].$$

Wir schließen daraus, daß die Quotientenabbildung die folgende Gestalt hat:

$$\begin{aligned} \mathbb{N}_0 \times \mathbb{N}_0 &\twoheadrightarrow \mathbb{Z} \\ (a,b) &\mapsto a-b. \end{aligned}$$

(6) Auf ähnliche Weise setzt man durch

$$[(a,b)] \cdot [(c,d)] = [(ac+bd, bc+ad)]$$

die Multiplikation von $\mathbb{N}_0$ auf $\mathbb{Z}$ fort und zeigt, daß man in der Sprache von Kapitel 3 einen Ring $\mathbb{Z}$ erhält, den Ring der ganzen Zahlen.

(7) Zum Schluß setzen wir wieder die naive Brille auf und tun so, als ob man $\mathbb{N}_0 \subseteq \mathbb{Z}$ längst aus der Schule kennt. Mit diesem Verständnis definiert man auf $\mathbb{N}_0 \times \mathbb{N}_0$ die Relation

$$(a,b) \sim (c,d) :\iff a - b = c - d.$$

Hier basiert die Äquivalenzrelation auf Gleichheit der Differenz. Da ist es offensichtlich, daß wir eine Äquivalenzrelation haben und $\mathbb{N}_0 \times \mathbb{N}_0 \twoheadrightarrow \mathbb{Z}$ gegeben durch $(a,b) \mapsto a - b$ im Prinzip die Quotientenabbildung ist (bis auf Identifikation der Äquivalenzklassen $[(a,b)]$ mit dem gemeinsamen Wert $a-b$). Dies soll zeigen, daß die formale Definition nichts anderes als die bekannten ganzen Zahlen definiert und mit diesem Wissen die Konstruktion auch transparent wird.

3.1.3. *Die rationalen Zahlen.* Nun skizzieren wir die Konstruktion der rationalen Zahlen.

(1) Auf der Menge $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definieren wir die Relation

$$(a,s) \sim (b,t) :\iff at - bs = 0.$$

Überzeugen Sie sich, daß dies eine Äquivalenzrelation ist. Die Menge der Äquivalenzklassen bezeichnen wir mit $\mathbb{Q}$. Für die Äquivalenzklasse von (a,s) schreiben wir

$$\frac{a}{s} := [(a,s)].$$

(2) Dann gilt zum Beispiel für jedes $n, a, s \in \mathbb{Z}$, $n, s \neq 0$,

$$\frac{na}{ns} = [(na, ns)] = [(a,s)] = \frac{a}{s}.$$

(3) Mit den aus der Schule für die rationalen Zahlen bekannten Formeln definiert man Additon

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st}$$

und Multiplikation

$$\frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}.$$

Überzeugen Sie sich, daß diese Verknüpfungen wohldefiniert sind!

(4) Man zeigt leicht, daß in $\mathbb{Q}$ bezüglich $+$ und $\cdot$ die gewohnten Rechenregeln gelten. In der in der Sprache von Kapitel 3: $\mathbb{Q}$ ist ein Körper, der Körper der rationalen Zahlen. Übung!

3.2. Die inverse Abbildung.

Definition 3.7. Eine Abbildung heißt **inverse Abbildung** (oder **Inverses** oder **Umkehrabbildung**) zu einer Abbildung $f\colon X \to Y$ von Mengen, wenn es sich um eine Abbildung $g\colon Y \to X$ handelt mit

$$f \circ g = \mathrm{id}_Y \quad \text{und} \quad g \circ f = \mathrm{id}_X .$$

Gilt nur $f \circ g = \mathrm{id}_Y$, so nennt man g **Rechtsinverses** zu f. Gilt nur $g \circ f = \mathrm{id}_X$, so nennt man g **Linksinverses** zu f.

Als erstes überlegen wir uns, daß es höchstens ein Inverses zu einer gegebenen Abbildung geben kann.

Proposition 3.8 (Eindeutigkeit des Inversen)**.** *Falls eine Abbildung $f\colon X \to Y$ eine inverse Abbildung besitzt, dann ist dieses Inverse eindeutig.*

Beweis. Angenommen g und h wären Inverse zu f, dann gilt

$$g = g \circ \mathrm{id} = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id} \circ h = h. \qquad \square$$

Satz 3.9 (Kriterium für die Existenz eines Inversen)**.** *Sei $f\colon X \to Y$ eine Abbildung von Mengen. Dann gilt:*

$$f \text{ hat ein Inverses} \iff f \text{ ist bijektiv.}$$

Beweis. Angenommen, die Abbildung $g\colon Y \to X$ sei ein Inverses von f. Da die Identität jeder Menge eine bijektive Abbildung ist, folgt aus Proposition 2.48 (2) und $g \circ f = \mathrm{id}_X$ bereits f injektiv. Aus Proposition 2.48 (4) und $f \circ g = \mathrm{id}_Y$ folgt f surjektiv. Zusammen folgt aus Proposition 2.42, daß f bijektiv ist.

Nehmen wir nun umgekehrt an, daß f bijektiv ist, also injektiv und surjektiv. Wir konstruieren eine Abbildung $g : Y \to X$ wie folgt. Für ein beliebiges $b \in Y$ setzen wir

$$g(b) := a \qquad \text{für } a \in X \text{ mit } f(a) = b.$$

Unser Wunsch, ein Inverses zu f zu konstruieren, zwingt uns zu dieser Definition, denn dieses g erfüllt offensichtlich $f \circ g = \mathrm{id}_Y$ und $g \circ f = \mathrm{id}_X$. Aber nur, wenn dieses g wohldefiniert ist!

- Der Wert $g(b)$ ist nur definert für alle $b \in Y$, wenn wir zu jedem $b \in Y$ ein $a \in X$ mit $f(a) = b$ finden können. Das ist genau die Voraussetzung f surjektiv.
- Der Wert $g(b)$ muß eindeutig definiert sein: der Wert $g(b)$ darf nicht vom in der Definition nach Maßgabe von $f(a) = b$ benutzten a abhängen. Das ist genau die Voraussetzung f injektiv.
- Der Wert $g(b)$ muß aus X sein. Das ist klar.

Damit ist g wohldefiniert. Sei $a \in X$ beliebig. Dann gilt

$$g \circ f(a) = g(f(a)) = a,$$

denn für $b = f(a)$ ist a das Element von X, das in der Definition von $g(b)$ benutzt werden muß. Weiter gilt für ein beliebiges $b \in Y$ mit $a := g(b)$

$$f \circ g(b) = f(g(b)) = f(a) = b,$$

weil der Wert a von $g(b)$ ja durch die Bedingung $f(a) = b$ bestimmt wurde. □

Notation 3.10. Die typische Notation für die inverse Abbildung zu $f\colon X \to Y$ ist

$$f^{-1}\colon Y \to X.$$

Hier muß man aufpassen, die Notation nicht mit der Notation $f^{-1}(V)$ für das Urbild einer Teilmenge $V \subseteq V$ zu verwechseln. Das Urbild gibt es für *jede* Abbildung $f : Y \to X$, die inverse Abbildung nur für eine bijektive Abbildung, wie Satz 3.9 gezeigt hat.

Wenn eine Notation mehrfach verwendet wird, so spricht man manchmal von **Notationsmißbrauch**. Dieser kann gerechtfertigt sein, wenn dadurch die Notation weniger schwerfällig wird und ein Mißverständnis wenig wahrscheinlich ist. Im Fall von f^{-1} gibt es nur für bijektive Abbildungen zwei Interpretationen für $f^{-1}(v)$ für $v \in X$, und diese beiden Interpretationen stimmen überein! Das heilt den Notationsmißbrauch.

Korollar 3.11 (Bijektive Abbildungen haben bijektive Inverse)**.** *Sei $f\colon X \to Y$ eine bijektive Abbildung von Mengen. Dann ist die zu f inverse Abbildung $f^{-1}\colon Y \to X$ ebenfalls bijektiv, und zwar mit Inversem f:*

$$(f^{-1})^{-1} = f.$$

Beweis. Weil f^{-1} invers zu f ist, gelten $f \circ f^{-1} = \mathrm{id}$ und $f^{-1} \circ f = \mathrm{id}$. Wenn man dies aus Sicht von f^{-1} liest, folgt sofort, daß f das Inverse von f^{-1} ist („das“ und nicht mehr nur „ein“, weil es ja eindeutig ist). Satz 3.9 sagt dann, daß f^{-1} auch bijektiv ist. □

Proposition 3.12. *Sei $f\colon X \to Y$ eine Abbildung von Mengen.*

(1) *f hat ein Linksinverses $\Longrightarrow$ f ist injektiv.*
Wenn $X \neq \emptyset$, so gilt auch die Umkehrung: f ist injektiv $\Longrightarrow$ f hat ein Linksinverses.

(2) *f hat ein Rechtsinverses $\Longleftrightarrow$ f ist surjektiv.*

Beweis. (1) Sei λ ein Linksinverses von f, also $\lambda \circ f = \mathrm{id}_X$. Da id_X injektiv ist, folgt aus Proposition 2.48 (2) bereits f injektiv.

Für die umgekehrte Richtung nehmen wir nun an, daß f injektiv ist. Außerdem folgt aus $X \neq \emptyset$, daß wir ein Element $* \in X$ haben. Wir definieren ein Linksinverses $\lambda\colon Y \to X$ auf die folgende Art und Weise für $b \in Y$:

$$\lambda(b) = \begin{cases} a \text{ mit } b = f(a) & \text{falls } b \in f(X), \\ * & \text{falls } b \in Y \setminus f(X). \end{cases}$$

Jetzt müssen wir uns überlegen, daß die Abbildung λ wohldefiniert und ein Linksinverses von f ist.

- Es ist $\lambda(b)$ für jedes $b \in Y$ definiert, denn $Y = f(X) \cup (Y \setminus f(X))$.
- Aber ist der zugeordnete Wert eindeutig? Die Fallunterscheidung im Definitionsbereich durch $b \in f(X)$ und $b \in Y \setminus f(X)$ hat keine Überschneidungen, weil $f(X) \cap (Y \setminus f(X)) = \emptyset$. Das ist ok. Aber falls $b \in f(X)$ gilt, so haben wir ein Urbild $a \in X$ auszuwählen, und wer sagt, daß die Vorschrift für $\lambda(b)$ nicht von dieser Wahl abhängt? Nun, das tut sie nicht, denn f ist nach Voraussetztung injektiv und damit das $a \in X$ mit $f(a) = b$ eindeutig (sofern es existiert, was durch $b \in f(X)$ garantiert ist).
- Der definierte Wert $\lambda(b)$ liegt in jedem Fall in X.

Damit ist endlich $\lambda\colon Y \to X$ wohldefiniert. Daß λ ein Linksinverses zu f ist, folgt offensichtlich aus der Definition.

(2) Sei ρ ein Rechtsinverses von f, also $f \circ \rho = \mathrm{id}_X$. Da id_X surjektiv ist, folgt aus Proposition 2.48 (4) bereits f surjektiv. Das zeigt die eine Richtung.

Für die umgekehrte Richtung nehmen wir nun an, daß f surjektiv ist. Dann definieren wir ein Rechtsinverses $\rho\colon Y \to X$ zu f durch

$$\rho(b) = \text{ Wahl eines Elements von } f^{-1}(b)$$

für alle $b \in Y$. Die Abbildung ρ hängt von den getroffenen Wahlen ab, aber das stört uns nicht. Weil f surjektiv ist, ist für alle $b \in Y$ die Urbildmenge $f^{-1}(b)$ nicht leer. Also kann man für jedes b eine Wahl eines Urbilds treffen. Das so konstruierte ρ ist offensichtlich ein Rechtsinverses. □

Bemerkung 3.13. Der letzte Teil des Beweises von Proposition 3.12 hat ein ernstes Problem. Wir wissen zwar, daß wir für jedes $b \in Y$ ein $a \in f^{-1}(b)$ wählen können. Aber wer sagt uns, wie das gleichzeitig für alle b geht? Das **Auswahlaxiom**! In unserer naiven Mengenlehre sehen wir wahrscheinlich das Problem nicht, und das ist im Moment auch gut so.

Proposition 3.14. *Sei $f\colon X \to Y$ eine Abbildung von Mengen. Die folgenden Aussagen sind äquivalent.*

(a) *f ist bijektiv.*
(b) *f hat Links- und Rechtsinverses.*
(c) *f hat ein Inverses.*

Beweis. Wir wissen bereits (a) $\iff$ (c) aus Satz 3.9, und (c) $\implies$ (b) ist trivial, denn ein Inverses ist gleichzeitig Links- und Rechtsinverses. Wenn wir noch zeigen, daß (b) $\implies$ (a), dann sind wir fertig nach dem Prinzip des Ringschluß.

(b) $\implies$ (a): Aus Proposition 3.12 folgt, daß f injektiv und surjektiv ist, was f bijektiv nach Proposition 2.42 bedeutet. □

3.3. **Die symmetrische Gruppe.** Bevor wir lernen, was eine Gruppe ist, betrachten wir zunächst die symmetrische Gruppe als Beispiel.

Definition 3.15.

(1) Eine **Permutation** einer Menge X ist eine Bijektion $X \xrightarrow{\sim} X$. Die Menge aller Permutationen einer Menge X bezeichnen wir mit

$$\mathrm{Aut}(X) = \{\sigma\colon X \to X \ ; \ \sigma \text{ bijektiv}\}.$$

Die Notation „Aut“ kommt von **Automorphismengruppe**.

(2) Sei $n \in \mathbb{N}$. Die Menge der Permutationen der Menge $\{1, 2, \ldots, n\}$ bezeichnen wir mit

$$S_n = \mathrm{Aut}(\{1, 2, \ldots, n\})$$

und nennen sie die **Symmetrische Gruppe** auf n Elementen.

Bemerkung 3.16. Wir haben bereits einges über diese Menge der bijektiven Selbstabbildungen von X gelernt:

- Beispiel 2.41: $\mathrm{id}_X \in \mathrm{Aut}(X)$.
- Korollar 2.49: mit $\sigma, \tau \in \mathrm{Aut}(X)$ ist auch $\sigma \circ \tau \in \mathrm{Aut}(X)$.

- Proposition 2.45: die Komposition von Bijektionen aus $\mathrm{Aut}(X)$ ist assoziativ, denn das gilt ja ganz allgemein für die Komposition von Abbildungen von Mengen.
- Satz 3.9: jedes $\sigma \in \mathrm{Aut}(X)$ hat ein Inverses $\sigma^{-1}\colon X \to X$. Nach Proposition 3.8 ist σ^{-1} eindeutig und auch in $\mathrm{Aut}(X)$.

Bevor wir die hier angedeuteten algebraischen Eigenschaften von $\mathrm{Aut}(X)$ in Definition 3.21 im Gruppenbegriff abstrahieren und studieren, wenden wir uns dem konkreteren Spezialfall der symmetrischen Gruppe zu.

Wie stellt man ein Element von S_n dar und wie kann man die Komposition ausrechnen? Die erste Möglichkeit ist eine schlichte Wertetabelle. Ein $\sigma \in S_n$ wird beschrieben durch

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n-1) & \sigma(n) \end{pmatrix}$$

σ wirkt von oben nach unten: für $\sigma(i)$ schaut man in die Spalte, in der oben i steht, und findet unten den Wert $\sigma(i)$.

Beispiel 3.17. Wir nehmen $n = 3$. Das Symbol

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

beschreibt die Permutation mit $\sigma(1) = 2$, $\sigma(2) = 3$ und $\sigma(3) = 1$.

Verfahren 3.18. Seien zwei Permutationen aus S_n gegeben:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \tau(1) & \tau(2) & \tau(3) & \cdots & \tau(n) \end{pmatrix}.$$

Wir berechnen die Komposition $\sigma\tau := \sigma \circ \tau$ durch

- Umsortieren der ersten Permutation (die als zweites ausgeführt wird — das Umsortieren beschreibt als Wertetabelle offensichtlich die gleiche Permutation)

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix} = \begin{pmatrix} \tau(1) & \tau(2) & \tau(3) & \cdots & \tau(n) \\ \sigma(\tau(1)) & \sigma(\tau(2)) & \sigma(\tau(3)) & \cdots & \sigma(\tau(n)) \end{pmatrix}$$

- Übereinanderschreiben

$$\tau = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \tau(1) & \tau(2) & \tau(3) & \cdots & \tau(n) \end{pmatrix}$$
$$\sigma = \begin{pmatrix} \tau(1) & \tau(2) & \tau(3) & \cdots & \tau(n) \\ \sigma(\tau(1)) & \sigma(\tau(2)) & \sigma(\tau(3)) & \cdots & \sigma(\tau(n)) \end{pmatrix}$$

- und anschließendes Streichen der zwei mittleren Zeilen:

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(\tau(1)) & \sigma(\tau(2)) & \sigma(\tau(3)) & \cdots & \sigma(\tau(n)) \end{pmatrix}$$

Beispiel 3.19. Für $n \geq 3$ ist die Komposition in S_n nicht mehr kommutativ. Beispielsweise gilt in S_3 für

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

zum einen

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

und zum anderen

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

Man sieht insbesondere $\sigma\tau \neq \tau\sigma$.

Satz 3.20. *Die Mächtigkeit der symmetrischen Gruppe auf n Elementen ist*

$$|S_n| = n!.$$

Beweis. Jedes $\sigma \in S_n$ wird eindeutig durch die Reihenfolge $\sigma(1), \sigma(2), \ldots, \sigma(n)$ bestimmt, und jede solche Reihenfolge bestimmt ein eindeutiges σ. Die Anzahl der Elemente von S_n ist gleich der Anzahl der Anordnungen von $1, 2, \ldots, n$. Der Satz folgt somit direkt aus Proposition 2.12. □

3.4. **Gruppen.** Eine Gruppe ist ein algebraisches Konzept, welches das Beispiel der S_n oder allgemeiner $\mathrm{Aut}(X)$ für eine Menge X abstrakt faßt.

Definition 3.21. Eine **Gruppe** besteht aus einer Menge G und einer Abbildung

$$G \times G \to G$$
$$(a, b) \mapsto a * b,$$

Verknüpfung genannt, mit den folgenden Eigenschaften.

(i) Die Verknüpfung ist **assoziativ**: für alle $a, b, c \in G$ gilt
$$a * (b * c) = (a * b) * c.$$

(ii) Es gibt ein **neutrales Element** $e \in G$, so daß für alle $a \in G$ gilt
$$e * a = a = a * e.$$

(iii) Existenz von **Inversen**: für alle $a \in G$ gibt es ein $b \in G$ mit
$$a * b = b * a = e.$$

Notation 3.22. Wenn man die Verknüpfung $*$ einer Gruppe G betonen will, dann schreibt man auch $(G, *)$ für die Gruppe. Übliche Terminologie und Notation für die Verknüpfung von $a, b \in G$ sind $a + b$ Addition, $a \cdot b$ Multiplikation, oder ab Multiplikation.

Definition 3.23. Eine **kommutative Gruppe** (oder auch **abelsche Gruppe**) ist eine Gruppe, in der auch gilt:

(iv) Die Verknüpfung ist **kommutativ**: für alle $a, b \in G$ gilt
$$a * b = b * a.$$

Beispiel 3.24. (1) Sei X eine Menge. Dann ist
$$\mathrm{Aut}(X)$$
eine Gruppe bezüglich der Verknüpfung Komposition. Bemerkung 3.16 führt die nötigen Eigenschaften der Komposition auf. Das neutrale Element ist die Identität id_X, das Inverse ist die inverse Abbildung, und Kompositon ist wohldefiniert und assoziativ.

(2) Sei $n \geq 1$ eine natürliche Zahl. Die symmetrische Gruppe
$$S_n$$
ist eine Gruppe bezüglich Komposition. Diese Gruppe hat $n!$ viele Elemente. Die S_n ist kommutativ genau für $n \leq 2$.

(3) Die ganzen Zahlen
$$\mathbb{Z}$$
mit Addition als Verknüpfung bilden eine kommutative Gruppe. Das neutrale Element ist die 0, das Inverse zu $n \in \mathbb{Z}$ ist $-n$.

(4) Die kleinste Gruppe ist $G = \{e\}$ mit der einzig möglichen Verknüpfung
$$e * e = e.$$
Diese Gruppe nennt man die triviale Gruppe.

(5) Die Menge $\{1, -1\}$ von ganzen Zahlen mit der Verknüpfung definiert durch Multiplikation ganzer Zahlen ist eine Gruppe mit zwei Elementen. Das neutrale Element ist 1.
(6) Sei n eine natürliche Zahl. Die Menge

$$\mathbb{Z}/n\mathbb{Z}$$

der Restklassen modulo n ist mit der in Beispiel 2.68 definierten Addition als Verknüpfung eine abelsche Gruppe. Das neutrale Element ist $[0]$ und das Inverse zu $[a]$ ist $[-a]$.

Bemerkung 3.25. Man sollte der Versuchung widerstehen, eine (endliche) Gruppe durch ihre Verknüpfungstafel, also eine Tabelle, welche die Werte gh mit $g, h \in G$ angibt, verstehen zu wollen. Man sieht leicht, daß es genau eine Gruppe mit zwei Elementen $G = \{e, g\}$ gibt. Die Verknüpfungstafel muß lauten:

	e	g
e	e	g
g	g	e

Die dargestellte Information ist vollständig, aber auch vollständig nutzlos zum Verständnis der Gruppe.

Bemerkung 3.26. Manchmal hört man, daß zu den Eigenschaften einer Gruppe G gehört, daß G „abgeschlossen unter der Verknüpfung“ sein muß. Das ist Blödsinn, welcher der Vorstellung entspringt, das Ergebnis der Verknüpfung sei a priori irgendwo vorhanden, und die Frage, ob es in G liegt, sei sinnvoll.

Der Gehalt dieser Vorstellung betrifft den wahren Sachverhalt, daß eine Gruppe mehr ist als eine Menge: eine Menge mit Verknüpfung. Und diese Verknüpfung muß als Abbildung **wohldefiniert** sein. Man muß die Verknüpfung überhaupt erst definieren, und wenn sie wohldefiniert ist, dann hat man eine Verknüpfung mit Werten in G.

Proposition 3.27. *Das neutrale Element in einer Gruppe ist eindeutig.*

Beweis. Sei $(G, *)$ eine Gruppe und e, e' seien neutrale Elemente. Dann gilt

$$e = e * e' = e'.$$

□

Proposition 3.28. *Das inverse Element zu einem Gruppenelement ist eindeutig.*

Beweis. Sei $(G, *)$ eine Gruppe mit neutralem Element e, und seien $b, c \in G$ inverse Elemente zu a. Dann gilt

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

□

Lemma 3.29. *Seien $(G, *)$ eine Gruppe mit neutralem Element e. Seien $a, b, c \in G$. Dann gilt:*
(1) Aus $b * a = e$ folgt $b = a^{-1}$.
(2) Aus $a * c = e$ folgt $c = a^{-1}$.
In einer Gruppe sind Linksinverse (bzw. Rechtsinverse) bereits Inverse.

Beweis. Sei a^{-1} das inverse Element zu a. Dann folgt aus $b * a = e$ bereits

$$b = b * e = b * (a * a^{-1}) = (b * a) * a^{-1} = e * a^{-1} = a^{-1}.$$

Die andere Behauptung folgt analog. □

Proposition 3.30. *Seien $(G, *)$ eine Gruppe und $a, b \in G$. Dann ist*
$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

Beweis. Nach Lemma 3.29 reicht die folgende Rechnung:
$$(a * b) * (b^{-1} * a^{-1}) = a * b * (b^{-1} * a^{-1}) = a * (b * b^{-1}) * a^{-1} = a * e * a^{-1} = a * a^{-1} = e,$$
wobei e das neutrale Element in G ist. □

Bemerkung 3.31. (1) Den Beweis von Proposition 3.28 haben wir im Spezialfall $G = \mathrm{Aut}(X)$ bereits in Satz 3.9 geführt.
(2) Das eindeutige neutrale Element wird auch **Eins** oder **Einselement** genannt, oder **Null** oder **Nullelement**, je nachdem welche Notation man für die Verknüpfung wählt. Als Bezeichnung sind dann 0 oder 1 üblich.
(3) Das eindeutige Inverse Element zu einem $a \in G$ wird oft mit a^{-1} oder $-a$ bezeichnet, je nachdem welche Notation man für die Verknüpfung gewählt hat.

3.5. **Ringe.** Eine Gruppe formalisiert eine gewisse algebraische Struktur mit einer Verknüpfung. Die bekannte Arithmetik der rationalen/reellen Zahlen aus der Schule hat zwei Verknüpfungen: Addition und Multiplikation. Die zugehörige algebraische Struktur ist ein Ring.

Definition 3.32. Ein **Ring** ist eine Menge A zusammen mit zwei Verknüpfungen, einer **Addition**
$$\begin{aligned} +\colon A \times A &\to A \\ (a, b) &\mapsto a + b \end{aligned}$$
und einer **Multiplikation**
$$\begin{aligned} \cdot\colon A \times A &\to A \\ (a, b) &\mapsto a \cdot b \end{aligned}$$
mit den folgenden Eigenschaften.
(i) $(A, +)$ ist eine kommutative Gruppe.
(ii) Die Multiplikation ist assoziativ: für alle $a, b, c \in A$ gilt
$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$
(iii) Die Multiplikation ist kommutativ: für alle $a, b \in A$ gilt
$$a \cdot b = b \cdot a.$$
(iv) Es gibt ein neutrales Element der Multiplikation, genannt Eins und bezeichnet mit $1 \in A$:
$$1 \cdot a = a = a \cdot 1$$
für alle $a \in A$.
(v) Es gilt das **Distributivgesetz**: für alle $a, b, c \in A$ gilt
$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

Beispiel 3.33.
(1) Die ganzen Zahlen $\mathbb{Z}$ mit Addition und Multiplikation sind ein Ring.
(2) Sei $n \in \mathbb{N}$. Die Menge $\mathbb{Z}/n\mathbb{Z}$ der Restklassen modulo n ist ein Ring mit der in Beispiel 2.68 definitierten Addition und Multiplikation. Die Null ist $[0]$ und die Eins ist $[1]$.

Später lernen wir weitere Beispiele kennen, etwa den Polynomring.

Notation 3.34. (1) Wenn man die Verknüpfung $+$ und $\cdot$ bei einem Ring A betonen will, dann schreibt man auch $(A, +, \cdot)$ für den Ring.

(2) Das neutrale Element der Addition in einem Ring A bezeichnen wir mit 0. Dies hängt nur von der Gruppe $(A, +)$ ab. Wegen Proposition 3.27 ist 0 eindeutig.
(3) Das additive Inverse im Ring A, also das Inverse von $a \in A$ bezüglich der Addition, bezeichnen wir mit $-a$. Es ist nach Proposition 3.28 eindeutig. Außerdem kürzen wir ab:

$$a - b = a + (-b).$$

Es gelten die üblichen Rechenregeln für Vorzeichen. Zum Beispiel gilt für alle $a, b, c \in A$

$$a - (b - c) = (a - b) + c$$

(4) Es gilt die Konvention Punkt- vor Strichrechnung. Dies erspart wie gewöhnlich einige Klammern.
(5) Die Notation wird auch durch Weglassen des Multiplikationspunktes vereinfacht: $ab = a \cdot b$.

Bemerkung 3.35. In einem Ring A ist das multiplikative neutrale Element 1 eindeutig. Dies beweisen wir mit derselben Idee wie in Proposition 3.27: angenommen $1' \in A$ ist auch ein neutrales Element für die Multiplikation, dann gilt

$$1 = 1 \cdot 1' = 1'.$$

Das ist derselbe Beweis wie bei der Eindeutigkeit des neutralen Elements einer Gruppe. Dieser Beweis funktioniert, obwohl $(A, \cdot)$ keine Gruppe ist (es gibt für die $0 \in A$ kein Inverses, es sei denn $A = \{0\}$).

Proposition 3.36. *Sei A ein Ring und $a \in A$. Dann gilt*

$$0 \cdot a = 0 = a \cdot 0.$$

Beweis. Es gilt $0 + 0 = 0$, daher

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

Additiert man das Inverse $-(0 \cdot a)$, so ergibt sich die Behauptung durch

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

Dann gilt nämlich auch $a \cdot 0 = 0$, denn A ist kommutativ. □

Proposition 3.37. *Sei A ein Ring und $a, b \in A$. Dann gilt*
(1) $-(-a) = a$,
(2) $-(ab) = (-a)b = a(-b)$.

Beweis. (1) Wegen $a + (-a) = 0 = (-a) + a$ ist a das additive Inverse zu $-a$, also $-(-a) = a$. (Das haben wir schon mal gesehen!)
(2) Es gilt

$$ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$$

und

$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0 \cdot a = 0.$$

Weil Addition kommutativ ist, sind damit $a(-b)$ und $(-a)b$ das Inverse von ab. □

Korollar 3.38. *Sei A ein Ring und $a, b \in A$. Dann gilt*
(1) $ab = (-a)(-b)$,
(2) $-a = (-1)a$,
(3) $(-1) \cdot (-1) = 1$.

Beweis. (1) Aus Proposition 3.37 (1) und zweimal (2) folgt

$$ab = -(-(ab)) = -((-a)b) = (-a)(-b).$$

(2) Dies ist der Fall $b = 1$ in Proposition 3.37 (2):

$$-a = -(a \cdot 1) = a(-1) = (-1)a.$$

(3) Wir setzen $a = b = 1$ in (1):

$$1 = 1 \cdot 1 = (-1)(-1).$$ □

3.6. **Körper.** Um lineare Gleichungen lösen zu können, müssen wir durch Elemente $\neq 0$ dividieren können. Dieses Konzept formalisiert der Begriff des Körpers.

Definition 3.39. Ein **Körper** ist ein Ring K mit

(i) $0 \neq 1$.

(ii) Jedes $a \in K$, $a \neq 0$ hat bezüglich Multiplikation ein Inverses: es gibt $b \in K$ mit

$$a \cdot b = 1 = b \cdot a.$$

Beispiel 3.40. (1) Beispiele für Körper sind die bekannten Körper der rationalen Zahlen $\mathbb{Q}$, und die reellen Zahlen $\mathbb{R}$.

(2) Der Ring $\mathbb{Z}$ ist kein Körper.

Satz 3.41. *Sei K ein Körper und $a, b \in K$ verschieden von 0. Dann ist auch $ab \neq 0$.*

Beweis. Dies beweisen wir durch Widerspruch. Angenommen $ab = 0$. Nach Voraussetzung gibt es multiplikative Inverse $\alpha, \beta \in K$ mit $a\alpha = 1 = b\beta$. Aus Proposition 3.36 folgt

$$(ab)(\alpha\beta) = 0(\alpha\beta) = 0.$$

Andererseits gilt aber auch

$$(ab)(\alpha\beta) = (a\alpha)(b\beta) = 1 \cdot 1 = 1,$$

und das ist wegen $0 \neq 1$ ein Widerspruch. □

Korollar 3.42. *Sei K ein Körper. Die von 0 verschiedenen Elemente*

$$K^\times := K \setminus \{0\} = \{a \in K \ ; \ a \neq 0\}$$

sind mit der Einschränkung der Multiplikation eine kommutative Gruppe. Diese Gruppe heißt ***multiplikative Gruppe*** *des Körpers K.*

Beweis. Aus Satz 3.41 folgt, daß die Multiplikation eingeschränkt auf $K^\times$ eine Verknüpfung ist:

$$\begin{aligned} K^\times \times K^\times &\to K^\times \\ (a, b) &\mapsto ab. \end{aligned}$$

Assoziativität, Existenz des neutralen Elements (der Eins), Existenz des Inversen, Kommutativität sind Bestandteil der Körperaxiome.

Eine Kleinigkeit gibt es noch zu beachten. Das multiplikative Inverse a^{-1} zu einem $a \in K^\times$ gibt es nach Körperaxiomen zuerst einmal in K. Wir müssen beweisen, daß $a^{-1} \in K^\times$, also $a^{-1} \neq 0$. Dies beweisen wir durch Widerspruch. Angenommen $a^{-1} = 0$. Dann ist

$$0 = 0 \cdot a = a^{-1} \cdot a = 1,$$

im Widerspruch zu $0 \neq 1$ in einem Körper. □

Korollar 3.43. *In einem Körper K ist für jedes $0 \neq a \in K$ das multiplikative Inverse eindeutig. Wir schreiben a^{-1} für das multiplikative Inverse von a.*

Beweis. Das ist nun klar, weil das nach Proposition 3.28 in jeder Gruppe gilt. □

Beispiel 3.44. Es gibt genau einen Körper mit 2 Elementen. Dieser wird mit $\mathbb{F}_2$ bezeichnet und muß aus den Elementen

$$\mathbb{F}_2 = \{0, 1\}$$

bestehen. Die Addition erklärt sich von selbst (Übungsaufgabe!):

$+$	0	1
0	0	1
1	1	0

und die Multiplikation ebenfalls (Übungsaufgabe!):

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

Wie bereits erwähnt, sind Verknüpfungstafeln keine erhellende Methode, um eine algebraische Struktur zu erklären. Wir lernen in Kapitel 3.7 systematisch die Körper mit p Elementen, p eine Primzahl, kennen. Daher sparen wir uns die Mühe, die Körperaxiome nachzuprüfen. Es sind für Addition und Multiplikation jeweils assoziativ, kommutativ und für beide gemeinsam distributiv zu testen, was durch 32 einfache Rechnungen zu erledigen ist.

Wenn Sie sich in Euklids Lehre des Rechnens mit *gerade* (das ist die 0) und *ungerade* (das ist die 1) auskennen, dann glauben Sie die Körperaxiome sofort.

Bemerkung 3.45. Bei einem **Schiefkörper** verlangt man dieselben Axiome wie für einen Körper, nur daß auf die Kommutativität der Multiplikation verzichtet wird. Im Prinzip funktioniert die lineare Algebra auch mit Schiefkörpern. Als Beispiel für einen Schiefkörper, der kein Körper ist, konstruieren wir in Anhang B die Hamiltonschen Quaternionen.

Beispiel 3.46. Eine lineare Gleichung in einer Variablen X über einem Körper K ist eine Gleichung der Form

$$a_1X + b_1 = a_2X + b_2$$

mit $a_1, a_2, b_1, b_2 \in K$. Eine Lösung der linearen Gleichung in K ist ein $x \in K$ mit

$$a_1x + b_1 = a_2x + b_2.$$

Mit $a = a_1 - a_2$ und $b = b_2 - b_1$ hat offenbar

$$aX = b$$

die gleichen Lösungen in K. Daher beschränken wir uns auf lineare Gleichungen diese Typs. Offenbar gibt es genau dann eine Lösung (im Sinn von $\exists$),

- wenn $a = b = 0$, und dann löst jedes $x \in K$ die Gleichung, oder
- wenn $a \neq 0$, und dann löst nur $x = a^{-1}b$ die Gleichung.

Wenn wir später die allgemeine Aussage über die Lösbarkeit und die Lösungsmenge einer linearen Gleichung behandeln, soll zur Übung der einfache Fall mit nur einer Variablen und nur einer Gleichung damit verglichen werden.

Beispiel 3.47. Zwei lineare Gleichungen in den Variablen X, Y über einem Körper K sind ein System der Form

$$aX + bY = u \tag{I}$$
$$cX + dY = v \tag{II}$$

mit $a, b, c, d, u, v \in K$. Gesucht ist die Lösungsmenge derjenigen $(x, y) \in K \times K$, für die

$$ax + by = u \tag{I}$$
$$cx + dy = v \tag{II}$$

beide gelten. Die aus der Schule bekannte Kunst besteht darin, die Gleichungen I & II derart zu manipulieren, daß einerseits die Lösungsmenge erhalten bleibt (Äquivalenzumformung) und zweitens die Lösung dem neuen System unmittelbar ansehen werden kann.

Fall 1: $d \neq 0$. In diesem Fall können wir die Gleichung I mit d multiplizieren. Weil man diesen Schritt durch Multiplizieren mit d^{-1} revidieren kann, liegt eine Äquivalenzumformung vor:

$$adX + bdY = ud \qquad (\text{I'} = d\cdot\ \text{I})$$
$$cX + dY = v. \qquad (\text{II'} = \text{II})$$

Als nächstes ziehen wir von I' das b-fache der Gleichung II' ab. Diesen Schritt kann man ebenfalls revidieren, indem man das b-fache von II' wieder hinzuaddiert. Es liegt ebenfalls eine Äquivalenzumformung vor:

$$(ad - bc)X \qquad = ud - vb \qquad (\text{I''} = \text{I'} - b\cdot\ \text{II'})$$
$$cX + dY = v. \qquad (\text{II''} = \text{II'})$$

Wir definieren die Determinante des Gleichungssystems als

$$\Delta = ad - bc.$$

Wenn $\Delta \neq 0$ ist, dann können wir I'' nach X auflösen, das Ergebnis in II'' einsetzen und schließlich nach Y auflösen:

$$\begin{aligned} X &= (ud - vb)\Delta^{-1} \\ Y &= d^{-1}(v - cX) = d^{-1}(v - c(ud - vb)\Delta^{-1}) = d^{-1}(\Delta v - c(ud - vb))\Delta^{-1} \\ &= d^{-1}(adv - bcv - cud - cvb)\Delta^{-1} = d^{-1}(adv - cud)\Delta^{-1} = (av - cu)\cdot\Delta^{-1} \end{aligned}$$

Die Details, was bei $\Delta = 0$ passiert, und zum Fall $d = 0$ der Fallunterscheidung überlassen wir zur Übung. Es ging in diesem Beispiel weniger darum, die Lösungstheorie von linearen Gleichungssystemen zu erklären — das machen wir später mit der geeigneten Sprache — sondern die Auffassung zu unterstützen, daß man in einem beliebigen Körper genauso algebraisch rechnen kann, wie in den vertrauten Körpern.

3.7. **Endliche Körper von Primzahlordnung.** In diesem Abschnitt studieren wir, unter welcher Bedingung an $n \in \mathbb{N}$ der Ring der Restklassen $\mathbb{Z}/n\mathbb{Z}$ ein Körper ist.

Definition 3.48.

(1) Seien $a, n \in \mathbb{Z}$. Man sagt, a ist durch n **teilbar** und schreibt

$$n \mid a,$$

wenn es ein $x \in \mathbb{Z}$ gibt mit $a = nx$.

(2) Seien $a, b \in \mathbb{Z}$, nicht beide 0. Die größte ganze Zahl $d > 0$ mit $d \mid a$ und $d \mid b$ heißt **größter gemeinsamer Teiler** von a und b und wird mit

$$d = \mathrm{ggT}(a, b)$$

bezeichnet. Die Menge der gemeinsamen Teiler $\{t\ ;\ t \mid a \text{ und } t \mid b \text{ und } t > 0\}$ ist nicht leer, denn sie enthält 1, und ist nach oben beschränkt durch das Minimum $\min(|a|, |b|)$. Daher gibt es den größten gemeinsamen Teiler; dieser ist wohldefiniert.

Satz 3.49 (Lemma von Bézout)**.** *Seien $a, b \in \mathbb{Z}$ nicht beide 0 und $d = \mathrm{ggT}(a, b)$. Dann gibt es $x, y \in \mathbb{Z}$ mit*

$$d = ax + by.$$

Beweis. Wenn $t \mid a$ und $t \mid b$, so gibt es $u, v \in \mathbb{Z}$ mit $a = ut$ und $b = vt$, und damit

$$t \mid t(u \pm v) = a \pm b.$$

Angewandt auf a, b (mit $-$) und auf $a - b, b$ (mit $+$) zeigt dies

$$\{t \ ; \ t \mid a \text{ und } t \mid b \text{ und } t > 0\} = \{t \ ; \ t \mid a - b \text{ und } t \mid b \text{ und } t > 0\}.$$

Es folgt

$$\mathrm{ggT}(a, b) = \mathrm{ggT}(a - b, b).$$

Weil die positiven Teiler von a und $-a$ die gleichen sind, dürfen wir ohne Einschränkung annehmen, daß $a, b \geq 0$.

Jetzt argumentieren wir per Induktion nach $n = a + b$. Der Induktionsanfang $n = 1$ ist nur für $a = 1$ und $b = 0$ (oder umgekehrt) zu realisieren. Wir behandeln allgemein den Fall $b = 0$. Dann ist $a > 0$, und in diesem Fall ist $\mathrm{ggT}(a, 0) = a = ax + by$ mit $x = 1$ und $y = 0$.

Für den Induktionsschritt dürfen wir annehmen, daß der Satz bereits für Paare $a', b' \geq 0$ mit $a' + b' < n = a + b$ gilt. Nach eventuellem Vertauschen von a und b nehmen wir $a \geq b \geq 0$ an. Den Fall $b = 0$ haben wir bereits erledigt. Wenn $b > 0$ gilt, dann ist $a' = a - b \geq 0$ und $b' = b > 0$ mit Summe $a' + b' = a < a + b$. Demnach gibt es $u, v \in \mathbb{Z}$ mit

$$d = \mathrm{ggT}(a, b) = \mathrm{ggT}(a - b, b) = \mathrm{ggT}(a', b') = a'u + b'v = (a - b)u + bv = au + b(v - u).$$

Damit gilt der Satz auch für a, b, und zwar mit $x = u$ und $y = v - u$. □

Satz 3.50. *Sei $n \in \mathbb{N}$. Der Ring $\mathbb{Z}/n\mathbb{Z}$ ist ein Körper genau dann, wenn n eine Primzahl ist.*

Beweis. Wenn n keine Primzahl ist, dann gibt es $1 < a, b < n$ mit $ab = n$. Insbesondere ist $[a] \neq 0 \neq [b]$, aber

$$[a] \cdot [b] = [ab] = [n] = 0.$$

Das geht in einem Körper nach Satz 3.41 nicht.

Es bleibt zu zeigen, daß für eine Primzahl p tatsächlich $\mathbb{Z}/p\mathbb{Z}$ ein Körper ist. Weil $p > 1$, ist $0 = [0] \neq [1] = 1$ in $\mathbb{Z}/p\mathbb{Z}$. Wir müssen nun nur noch zu einem beliebigen $[a] \in \mathbb{Z}/p\mathbb{Z}$, mit $[a] \neq 0$, ein Inverses finden. Es gilt

$$\mathrm{ggT}(a, p) = 1,$$

weil p nur die Teiler 1 und p hat, und p wegen $[a] \neq 0$ kein Teiler von a ist. Nach dem Lemma von Bézout, Satz 3.49, gibt es $x, y \in \mathbb{Z}$ mit

$$1 = ax + py.$$

Das bedeutet

$$1 = [1] = [ax + py] = [a] \cdot [x] + [p] \cdot [y] = [a] \cdot [x] + 0 \cdot [y] = [a] \cdot [x].$$

Das gesuchte Inverse ist $[x]$ (wegen Kommutativität gilt auch $[x] \cdot [a] = 1$). □

Notation 3.51. Wir haben nun schlagartig unseren Beispielvorrat an Körpern um ein Beispiel zu jeder Primzahl p erweitert. Den Körper $\mathbb{Z}/p\mathbb{Z}$ mit p Elementen bezeichnen wir mit

$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}.$$

Der schon früher eingeführte Körper $\mathbb{F}_2$ stimmt mit diesem $\mathbb{F}_2$ überein.

Übungsaufgaben zu §3

Übungsaufgabe 3.1. Wir betrachten die folgenden Abbildungen von Mengen:

$$X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W.$$

Zeigen Sie, daß g eine Inverse Abbildung hat, wenn $\varphi = g \circ f$ und $\psi = h \circ g$ eine Inverse Abbildung besitzen. Geben Sie die Inverse Abbildung an!

Übungsaufgabe 3.2. Wir betrachten zu zwei Mengen X und Y die Abbildung $\tau\colon X \times Y \to Y \times X$ definiert für alle $a \in X$ und $b \in Y$ durch

$$\tau(a, b) = (b, a).$$

Sei $f\colon X \to Y$ eine bijektive Abbildung. Wie hängen der Graph von f, die Abbildung τ und der Graph der Umkehrabbildung f^{-1} miteinander zusammen?

Übungsaufgabe 3.3. Wir numerieren die Ecken des Quadrats im Uhrzeigersinn mit $1, 2, 3, 4$.

$$\begin{array}{ccc} 1 & — & 2 \\ | & & | \\ 4 & — & 3 \end{array}$$

(1) Eine Drehung um 90^o im Uhrzeigersinn induziert eine Bijektion der Ecken(nummern) und damit ein Element der S_4. Schreiben Sie dieses als Wertetabelle auf.
(2) Stellen Sie sich nun vor, das Quadrat sei starr. Schreiben Sie alle Permutationen der Ecken als Elemente der S_4 auf, die Sie durch Bewegungen des starren Quadrats hinbekommen.
(3) Wählen Sie nun zwei verschiedene Elemente (und nicht die Identität!) aus Ihrer Liste aus, berechnen Sie das Produkt der Permutationen in S_4 und vergleichen Sie diese mit der entsprechenden Hintereinanderausführung der Bewegungen des Quadrats.

Übungsaufgabe 3.4. Betrachten sie die folgende zyklische Permutation in S_n

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ n & 1 & 2 & \cdots & n-1 \end{pmatrix}$$

und berechnen Sie σ^r per Induktion nach r.

Visualisieren Sie diese Rechnung durch eine entsprechende Bewegung der geeignet markierten Ecken eines reglmäßigen n-Ecks.

Übungsaufgabe 3.5. Eine Permutation $\tau \in S_n$, die genau zwei Elemente vertauscht und die anderen festläßt, nennt man eine **Transposition**. Wenn τ die Elemente a, b vertauscht, schreiben wir $\tau = (a, b)$. Zeigen Sie, daß jedes $\sigma \in S_n$ eine Komposition von einer endlichen Zahl von Transpositionen der Form $(i, i+1)$ für $1 \leq i < n$ ist.

Übungsaufgabe 3.6. Sei A ein Ring mit $0 = 1$. Zeigen Sie, daß A genau ein Element hat.

Bemerkung: Diesen Ring nennt man den Nullring.

Übungsaufgabe 3.7. Die Quotientenabbildung $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ ist ein Ringhomomorphismus, und die definierte Addition und Multiplikation ist die einzige Addition und Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$, für die das gilt.

Übungsaufgabe 3.8. Zeigen Sie, daß ein Ring A mit $0 \neq 1$ genau dann ein Körper ist, wenn alle linearen Gleichungen

$$aX + b = c$$

mit $a, b, c \in A$ und $a \neq 0$ eine Lösung haben.

Übungsaufgabe 3.9. Wir bezeichnen mit $\mathbb{Q}(\sqrt{2})$ die Teilmenge von $\mathbb{R}$ gegeben durch

$$\mathbb{Q}(\sqrt{2}) = \{x \in \mathbb{R} \ ; \text{ es gibt } a, b \in \mathbb{Q} \text{ mit } x = a + b\sqrt{2}\}.$$

Zeigen Sie, daß $\mathbb{Q}(\sqrt{2})$ mit der von $\mathbb{R}$ geerbten Addition und Multiplikation ein Körper ist.

4. Vektorräume

4.1. **Beispiele von Vektorräumen.** Nun kommen wir zur Definition derjenigen algebraischen Struktur, von der diese Vorlesung hauptsächlich handelt.

Definition 4.1. Sei K ein Körper. Ein **Vektorraum** (genauer ein K-**Vektorraum**) ist eine Menge V mit zwei Verknüpfungen, einer **Addition**

$$\begin{aligned} +\colon\ & V \times V \to V \\ & (v, w) \mapsto v + w, \end{aligned}$$

einer **Skalarmultiplikation**

$$\begin{aligned} \cdot\colon\ & K \times V \to V \\ & (a, w) \mapsto a \cdot w, \end{aligned}$$

mit den folgenden Eigenschaften.

(i) $(V, +)$ ist eine kommutative Gruppe.

(ii) Die Skalarmultiplikation ist **assoziativ**: für alle $a, b \in K$ und $v \in V$ gilt

$$a \cdot (b \cdot v) = (a \cdot b) \cdot v.$$

(iii) Die Skalarmultiplikation ist **unitär**: für alle $v \in V$ und die Eins $1 \in K$ gilt

$$1 \cdot v = v.$$

(iv) Es gelten die **Distributivgesetze**: für alle $v, w \in V$ und $a, b \in K$ gilt

$$\begin{aligned} a \cdot (v + w) &= (a \cdot v) + (a \cdot w) \\ (a + b) \cdot v &= (a \cdot v) + (b \cdot v). \end{aligned}$$

Wenn man die Verknüpfung $+$ und $\cdot$ und den Körper der Skalare bei einem Vektorraum V betonen will, dann schreibt man auch $(V, K, +, \cdot)$ für den Vektorraum. Das werden wir nie tun, denn aus dem Kontext werden K sowie $+$ und $\cdot$ klar sein.

Bemerkung 4.2. Die Elemente eines Vektorraums nennt man **Vektoren**. Das ist nur ein Name! Die Elemente des Körpers K heißen bei einem K-Vektorraum die **Skalare**. Man kann Vektoren mit Elementen $a \in K$ skalieren: $v \in V$ wird zu $av := a \cdot v$.

Skalare stehen in der Notation immer links. Die Skalarmultiplikation ist schon aus formalen Gründen nicht symmetrisch, denn der Skalar ist aus K und der Vektor aus dem Vektorraum!

Beispiel 4.3. Sei K ein Körper. Zuerst beschreiben wir den Prototypen eines Vektorraums. Sei $n \in \mathbb{N}$. Der K-Vektorraum

$$K^n$$

ist als Menge $K^n = K \times \ldots \times K$ mit n Faktoren. Ein Vektor ist also ein n-Tupel mit Einträgen aus K. Diese werden vertikal geschrieben und dann **Spaltenvektoren** genannt. Die Addition ist komponentenweise:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

wobei $v_1, \ldots, v_n, w_1, \ldots, w_n \in K$ sind. Da $(K, +)$ eine kommutative Gruppe ist, erben die Spaltenvektoren diese Eigenschaft offensichtlich. Die 0 ist

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

und das Inverse ist

$$-\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}.$$

Die Skalarmultiplikation mit $a \in K$ geschieht auch komponentenweise:

$$a \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} av_1 \\ \vdots \\ av_n \end{pmatrix}.$$

Die Gültigkeit der Vektorraumaxiome für diese Addition und Skalarmultiplikation ist unmittelbar einsichtig.

Beispiel 4.4. Der „kleinste“ K-Vektorraum ist der Nullraum. Dieser besteht nur aus einem Element 0 und stimmt mit K^0 in der richtigen Interpretation überein. Der K^0 besteht aus der Menge der leeren Tupel (mit der leeren Menge indizierte Tupel). Ein leeres Tupel ist eine Abbildung $\emptyset \to K$ und davon gibt es nur eine: die Inklusion der leeren Menge als Teilmenge von K. Denken Sie an den Graph der Abbildung als Teilmenge von $\emptyset \times K = \emptyset$.

Die zugrundeliegende abelsche Gruppe $(K^0, +)$ ist die triviale Gruppe. Die Skalarmultiplikation

$$K \times K^0 \to K^0$$

macht $a \cdot 0 = 0$, was sonst. Die Vektorraumaxiome sind evident.

Der Nullvektorraum zu einem Körper K und zu einem anderen Körper L sind genaugenommen 2 verschiedene algebraische Strukturen. Beides sind algebraische Strukturen auf der Menge $\{0\}$, aber die Skalare, mit denen man multiplizieren darf, sind verschieden.

Beispiel 4.5. Der Fall $n = 1$. Der K-Vektorraum K^1 ist vom Körper K als Menge nicht zu unterscheiden. Ein 1-Tupel aus K ist nichts anderes als ein Element aus K. Die Addition in K^1 ist die Körperaddition, die Skalarmultiplikation

$$K \times K^1 \to K^1$$

ist die Körpermultiplikation. Man muß aufpassen, daß einen diese Koinzidenz nicht verwirrt. Dies hat aber auch eine gute Seite. Man kann Körper mit der Theorie der Vektorräume studieren.

Beispiel 4.6. Der Spezialfall $K = \mathbb{R}$ und $n = 3$ führt zum Vektorraum $\mathbb{R}^3$, den wir gerne als Modell für den uns umgebenden physikalischen Raum begreifen. Der $\mathbb{R}^4$ hat zusätzlich eine Zeitkoordinate und beschreibt somit ein Modell der Raumzeit. Der $\mathbb{R}^2$ ist ein Modell einer euklidischen Ebene.

Proposition 4.7. *Sei V ein K-Vektorraum. Dann gilt für alle $v \in V$ und $\lambda \in K$:*
(1) $0 \cdot v = 0$,
(2) $\lambda \cdot 0 = 0$,
(3) $(-1) \cdot v = -v$,
(4) $\lambda \cdot v = 0 \iff \lambda = 0$ *oder* $v = 0$.

Beweis. (1) Es gilt in K, daß $0 + 0 = 0$, also

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

Zieht man von beiden Seiten $0 \cdot v$ ab, dann erhält man $0 \cdot v = 0$.

(2) Es gilt in V, daß $0 + 0 = 0$, also

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

Zieht man von beiden Seiten $\lambda \cdot 0$ ab, dann erhält man $\lambda \cdot 0 = 0$.

(3) Nach Lemma 3.29 reicht die folgende Rechnung:

$$v + (-1)v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = 0.$$

(4) Wenn $\lambda = 0$ oder $v = 0$, dann zeigen (1) und (2), daß $\lambda \cdot v = 0$.

Wir müssen daher nur noch die andere Richtung zeigen. Sei also $\lambda \cdot v = 0$. Wenn $\lambda = 0$, dann sind wir fertig. Andernfalls gibt es $\lambda^{-1} \in K$ und dann gilt

$$0 = \lambda^{-1} \cdot 0 = \lambda^{-1} \cdot (\lambda \cdot v) = (\lambda^{-1} \cdot \lambda) \cdot v = 1 \cdot v = v. \qquad \square$$

Beispiel 4.8. Auch in der Analysis treten Vektorräume natürlich auf.

(1) Der $\mathbb{R}$-Vektorraum der reellen Funktionen auf einem Intervall $[a, b] = \{x \in \mathbb{R} \ ; \ a \leq x \leq b\}$:

$$\mathrm{Abb}([a, b], \mathbb{R}) = \{f \colon [a, b] \to \mathbb{R} \ ; \ \text{Abbildung}\}$$

ist mit punktweiser Addition und Multiplikation ein $\mathbb{R}$-Vektorraum. Zu $f, g \in \mathrm{Abb}([a, b], \mathbb{R})$ und $\lambda \in \mathbb{R}$ definieren

$$(f + g)(x) := f(x) + g(x)$$
$$(\lambda f)(x) := \lambda \cdot f(x)$$

eine Addition und eine Skalarmultiplikation auf $\mathrm{Abb}([a, b], \mathbb{R})$. Der Nullvektor ist die Funktion $0(x) = 0$ für alle $x \in [a, b]$. Das additive Inverse zur Funktion f ist

$$(-f)(x) = -f(x).$$

Beispiel 4.9. Wir verallgemeinern nun sowohl den K^n als auch den $\mathrm{Abb}([a, b], \mathbb{R})$.

Sei K ein Körper und X eine Menge. Dann ist der Raum der Abbildungen von X nach K

$$\mathrm{Abb}(X, K) = \{f \colon X \to K \ ; \quad \text{Abbildung}\}$$

in natürlicher Weise ein K-Vektorraum durch:

(1) Addition: zu $f, g \in \mathrm{Abb}(X, K)$ definieren wir $f + g$ durch

$$(f + g)(x) = f(x) + g(x)$$

für alle $x \in X$.

(2) Skalarmultiplikation: zu $f \in \mathrm{Abb}(X, K)$ und $\lambda \in K$ definieren wir $\lambda \cdot f$ durch

$$(\lambda \cdot f)(x) = \lambda \cdot f(x)$$

für alle $x \in X$.

Wir sprechen die Abbildungen $f \colon X \to K$ auch als (K-wertige) Funktionen an. Man sagt, Addition und Skalarmultiplikation der K-wertigen Funktionen seien punktweise (für jeden Punkt = Element $x \in X$) definiert. Beachten Sie, daß nur im Bildbereich addiert und multipliziert werden muß. Daher macht es auch nichts, daß X nur eine Menge ist.

Die Vektorraumaxiome verifizieren sich wie von selbst. Der Nullvektor ist wieder die Funktion, die identisch 0 ist, das Inverse bezüglich der Addition zur Funktion f ist die Funktion $-f$ mit

$$(-f)(x) = -f(x),$$

und die Rechengesetze erbt $\mathrm{Abb}(X, K)$ von K, eben punktweise.

Beispiel 4.10. Wenn wir speziell $X = \mathbb{N}$ betrachten, dann ist eine Funktion $f \colon \mathbb{N} \to K$ nichts anderes als eine K-wertige Folge $a_1, a_2, a_3, \ldots$ mit $a_n \in K$ für alle $n \in \mathbb{N}$, und $\mathrm{Abb}(\mathbb{N}, K)$ ist der K-Vektorraum der Folgen.

Im Spezialfall $X = \{1, \ldots, n\}$ haben wir eine naheliegende Idenitfikaton

$$\mathrm{Abb}(X, K) = K^n.$$

Ein n-Tupel aus K ist nichts anderes als die Liste der Werte einer Funktion

$$f \colon \{1, \ldots, n\} \to K.$$

Addition und Skalarmultiplikation entsprechen sich unter dieser Identifikation. So eine Identifikation, welche die betrachtete Struktur erhält, studieren wir später genauer und nennen sie dann einen **Isomorphismus**.

Beispiel 4.11. Die Potenzmenge $\mathscr{P}(X)$ einer Menge X hat die folgende Struktur eines $\mathbb{F}_2$-Vektorraums. Die Summe zweier Teilmengen $U, V \subseteq X$ wird definiert durch die symmetrische Differenz

$$U + V := U \triangle V = (U \setminus V) \cup (V \setminus U).$$

Die Skalarmultiplikation mit $a \in \mathbb{F}_2 = \{0, 1\}$ wird definiert durch

$$1 \cdot U = U$$
$$0 \cdot U = \emptyset.$$

Die Verifikation der $\mathbb{F}_2$-Vektorraumstruktur ist eine Übung. Wir sehen das später eleganter durch den Vergleich mit einem anderen Vektorraum.

4.2. **Unterräume.** Wir betrachten nun lineare Geometrie in einem Vektorraum.

Definition 4.12. Sei V ein K-Vektorraum. Ein **Unterraum** (oder **Untervektorraum**) von V ist eine Teilmenge $U \subseteq V$ zusammen mit der aus den folgenden Bedingungen resultierenden Struktur als K-Vektorraum $(U, K, +, \cdot)$, und zwar:

(i) Die Addition auf U durch Einschränkung der Addition von V ist wohldefiniert:
$$+ := +|_{U \times U} \colon U \times U \to U$$
$$(x, y) \mapsto x + y.$$

(ii) Die Skalarmultiplikation auf U durch Einschränkung der Skalarmultiplikation von V ist wohldefiniert:
$$\cdot := \cdot|_{K \times U} \colon K \times U \to U$$
$$(a, x) \mapsto a \cdot x.$$

(iii) U ist ausgestattet mit $+$ aus (i) und $\cdot$ aus (ii) ein K-Vektorraum.

Satz 4.13 (Unterraumkriterium)**.** *Sei V ein K-Vektorraum und $U \subseteq V$. Die folgenden Aussagen sind äquivalent.*

(a) *U ein Unterraum.*

(b) *(i)* *für alle $x, y \in U$ gilt $x + y \in U$,*
(ii) *für alle $x \in U$ und $\lambda \in K$ gilt $\lambda x \in U$,*
(iii) *$U \neq \emptyset$.*

(c) *U ist nicht leer und für alle $x, y \in U$ und $\lambda \in K$ gilt $\lambda x + y \in U$.*

Beweis. Wir zeigen die Äquivalenz durch einen Ringschluß. Wenn U ein Unterraum in V ist, dann gilt (c), weil Skalarmultiplikation und Addition den Unterraum nicht verlassen, und weil $0 \in U$. Dies zeigt (a) $\implies$ (c). Die Implikation (c) $\implies$ (b) folgt durch die Spezialfälle $\lambda = 1$ für (i) und $y = 0$ für (ii).

(b) $\implies$ (a): Sei $U \subseteq V$ eine Teilmenge mit (i)-(iii). Wegen (i) ist

$$+ \colon U \times U \to U$$

wohldefiniert. Als Einschränkung von $+$ des Vektorraums V ist $+$ für U assoziativ und kommutativ.

Sei $u \in U$ beliebig: Existenz nach (iii) gesichert! Nach (ii) ist dann auch $-u = (-1)u \in U$ und somit nach (i)

$$0 = u + (-u) \in U.$$

Die 0 ist auch ein neutrales Element für $+$ in U, weil die Addition nur die Einschränkung der Addition von V ist.

Weil für alle $u \in U$ nach (ii) auch

$$-u = (-1) \cdot u \in U,$$

ist auch das Inverse zu $u \in U$ aus der Gruppe $(V, +)$, also a priori das Element $-u \in V$, in U gelegen. Weil $+$ von U nur die Einschränkung von $+$ für V ist, gilt auch als Rechnung in U:

$$u + (-u) = 0 = (-u) + u.$$

Damit hat jedes $u \in U$ ein Inverses in U bezüglich $+$. Somit ist $(U, +)$ eine kommutative Gruppe.

Nach (ii) ist die Skalarmultiplikation

$$\cdot \colon K \times U \to U$$

wohldefiniert. Alle für einen Vektorraum geforderten Rechengesetze gelten damit offensichtlich auch für U, da ja $+$ und $\cdot$ für U nur die Einschränkung der entsprechenden Verknüpfungen für V sind. □

Beispiel 4.14. Alle Beispiele folgen aus dem Unterraumkriterium Proposition 4.13.

(1) Für jeden K-Vektorraum V sind der Nullraum $0 = \{0\}$ und der Vektorraum V Unterräume von V.

(2) Bezeichnen wir mit

$$\mathrm{Pol}_{\mathbb{R}} \subseteq \mathrm{Abb}(\mathbb{R}, \mathbb{R})$$

die Teilmenge der reellen Polynomfunktionen, also der Funktionen $\mathbb{R} \to \mathbb{R}$, von der Form: es gibt $n \in \mathbb{N}_0$ und $a_i \in \mathbb{R}$ für $i = 0, \ldots, n$ mit

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \ldots + a_1 x + a_0$$

für alle $x \in \mathbb{R}$. Dann ist $\mathrm{Pol}_{\mathbb{R}}$ ein Unterraum des $\mathbb{R}$-Vektorraums $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$.

Definition 4.15. Seien K ein Körper und $n, m \in \mathbb{N}_0$.

(1) Eine **lineare Gleichung** in n-vielen Variablen $X_1, \ldots, X_n$ mit **Koeffizienten** $a_j \in K$ für $j = 1, \ldots, n$ ist eine Gleichung der Form

$$a_1 X_1 + \ldots + a_n X_n = 0.$$

Eine lineare Gleichung ist einfach ein homogenes Gleichungssystem mit $m = 1$ Zeilen im Sinne der folgenden Definition.

(2) Ein **homogenes lineares Gleichungssystem** in n-vielen Variablen $X_1, \ldots, X_n$ mit m-vielen Gleichungen und Koeffizienten $a_{ij} \in K$ für $i = 1, \ldots, m$ und $j = 1, \ldots, n$ ist ein System $\mathscr{S}$ von Gleichungen der Form

$$\mathscr{S} = \left\{ \begin{array}{ccccccccc} a_{11}X_1 & + & \ldots & + & \ldots & + & \ldots & + a_{1n}X_n & = & 0 \\ \vdots & & & & \vdots & & & \vdots & & \vdots \\ a_{i1}X_1 & + & \ldots & + a_{ij}X_j & + & \ldots & + a_{in}X_n & = & 0 \\ \vdots & & & & \vdots & & & \vdots & & \vdots \\ a_{m1}X_1 & + & \ldots & + & \ldots & + & \ldots & + a_{mn}X_n & = & 0 \end{array} \right.$$

(3) Eine **Lösung des Systems** $\mathscr{S}$ ist ein $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ mit

$$\begin{array}{ccccccccc} a_{11}x_1 & + & \dots & + & \dots & + & a_{1n}x_n & = & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1}x_1 & + & \dots & + a_{ij}x_j + & \dots & & + a_{in}x_n & = & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & \dots & + & a_{mn}x_n & = & 0 \end{array}$$

(4) Der **Lösungsraum** des Systems $\mathscr{S}$ ist die Teilmenge

$$\mathscr{L}(\mathscr{S}) := \{x \in K^n \ ; \ x \text{ ist Lösung von } \mathscr{S}\} \subseteq K^n.$$

Der Lösungsraum eines homogenen linearen Gleichungssystems $\mathscr{S}$ in n Variablen ist ein Unterraum $\mathscr{L}(\mathscr{S}) \subseteq K^n$.

Proposition 4.16. *Sei $\mathscr{S}$ ein homogenes lineares Gleichungssystem in n Variablen und m Gleichungen mit Koeffizienten aus dem Körper K. Dann ist der Lösungsraum $\mathscr{L}(\mathscr{S})$ ein Unterraum von K^n:*

(1) Die Summe zweier Lösungen von $\mathscr{S}$ ist wieder eine Lösung.
(2) Sei $a \in K$. Das a-fache einer Lösung von $\mathscr{S}$ ist wieder eine Lösung.

Beweis. Der Satz folgt im Prinzip auf triviale Weise aus dem Distributivgesetz. Wir zeigen, daß das Unterraumkriterium Proposition 4.13 erfüllt ist. Wir betrachten das jomogene lineare Gleichungssytem $\mathscr{S}$ mit der Notation aus Definition 4.15. Seien

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n$$

Lösungen von $\mathscr{S}$. Dann ist für alle $i = 1, \dots, m$ und alle $\lambda \in K$

$$\begin{aligned} a_{i1}(\lambda x_1 + y_1) + \dots + a_{in}(\lambda x_n + y_n) &= \lambda\big(a_{i1}x_1 + \dots + a_{in}x_n\big) + \big(a_{i1}y_1 + \dots + a_{in}y_n\big) \\ &= \lambda \cdot 0 + 0 = 0, \end{aligned}$$

und damit ist $\lambda x + y \in K^n$ auch eine Lösung. Offensichtlich ist $0 \in K^n$ eine Lösung von $\mathscr{S}$. Daher ist $\mathscr{L}(\mathscr{S})$ nicht leer und das Unterraumkriterium erfüllt.

Die Aussagen (1) und (2) folgen sofort, weil der Lösungsraum ein Unterraum ist. □

Bemerkung 4.17. Als eines der Ziele der Vorlesung Lineare Algebra wollen wir lineare Gleichungssysteme perfekt beherrschen. Wenn der Lösungsraum ein Vektorraum ist, dann müssen wir zuerst die Theorie der Vektorräume verstehen.

4.3. **Linearkombinationen.** Die allgemeinste Form von Ausdrücken, die man in einem Vektorraum berechnen kann, sind Linearkombinationen.

Notation 4.18. Die Summe von Elementen $a_1, \dots, a_n$ bezüglich einer assoziativen und kommutativen Addition $+$ (etwa in einem Körper oder einem Vektorraum oder einer kommutativen Gruppe) wird mit der Summennotation abgekürzt:

$$\sum_{i=1}^{n} a_i := a_1 + \dots + a_n = \Big(\big((a_1 + a_2) + \dots + a_{n-1}\big) + a_n\Big).$$

Assoziativität und Kommutativität sorgen dafür, daß es bei der Auswertung der Summe nicht auf die Reihenfolge und Klammerung ankommt. Dies beweist man leicht durch vollständige Induktion nach der Anzahl der Summanden.

Sei I eine endliche Menge mit $|I| = n$. Wir können auch über ein I-Tupel $(a_i)_{i\in I}$ von Elementen eines Körpers (oder Vektorraums oder einer kommutativen Gruppe) summieren. Für jede Wahl einer Aufzählung

$$I = \{i_1, \ldots, i_n\}$$

der Elemente von I setzen wir

$$\sum_{i\in I} a_i := \sum_{r=1}^{n} a_{i_r}.$$

Per Induktion nach der Mächtigkeit von I zeigt man, daß dies wohldefiniert ist, also nicht von der Reihenfolge abhängt, in der man die Elemente von I aufgelistet hat. Übung!

Bemerkung 4.19. Sei K ein Körper. Es gelten die folgenden Regeln für die Summennotation. Mit $a_i \in K$ für $i = 1, \ldots, n+m$ und $b \in K$ gilt

$$\sum_{i=1}^{n+m} a_i = \sum_{i=1}^{n} a_i + \sum_{i=n+1}^{n+m} a_i.$$

Weiter gelten die beiden Distributivgesetze

$$b \cdot \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} ba_i,$$
$$\Big(\sum_{i=1}^{n} a_i\Big) \cdot b = \sum_{i=1}^{n} a_i b.$$

Und die Summen über die gleichen Indizes kann man wie folgt addieren:

$$\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i = \sum_{i=1}^{n} (a_i + b_i).$$

Des weiteren kann man Doppelsummen über Elemente $a_{ij} \in K$ für $i = 1, \ldots n$ und $j = 1, \ldots, m$ wie folgt schreiben:

$$\sum_{i=1}^{n} \Big(\sum_{j=1}^{m} a_{ij}\Big) = \sum_{\substack{1\le i\le n\\ 1\le j\le m}} a_{ij} = \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} a_{ij}\Big).$$

Analoge Regeln gelten auch für Summen von Vektoren.

Definition 4.20. Sei V ein K-Vektorraum.

(1) Als **Linearkombination** von Vektoren $v_1, \ldots, v_n \in V$ mit **Koeffizienten** $a_i \in K$, $i = 1, \ldots, n$, bezeichnen wir eine Summe

$$a_1 v_1 + \ldots + a_n v_n = \sum_{i=1}^{n} a_i v_i.$$

Die Linearkombination ist **nichttrivial**, wenn mindestens ein Koeffizient $a_i \neq 0$ ist.

(2) Allgemeiner, sei $M \subseteq V$ eine Teilmenge. Als **Linearkombination** von Vektoren aus M bezeichnen wir eine Summe

$$a_1 v_1 + \ldots + a_n v_n = \sum_{i=1}^{n} a_i v_i$$

mit Koeffizienten $a_i \in K$ und Vektoren $v_i \in M$.

Bemerkung 4.21. Eine Linearkombination ist stets eine *endliche* Summe. Unendliche Summen, also Reihen, setzen einen Konvergenzbegriff voraus und gehören daher in die Analysis.

Beispiel 4.22. (1) Die einzige Linearkombination, die man aus der leeren Menge von Vektoren bilden kann, ist die **leere Summe**, deren Wert als

$$0 := \sum_{i\in\emptyset} a_i v_i$$

festgelegt wird. Dies ist die einzige vernünftige Konvention, die mit

$$\sum_{i\in I} a_i v_i + \sum_{i\in\emptyset} a_i v_i = \sum_{i\in I\cup\emptyset} a_i v_i = \sum_{i\in I} a_i v_i$$

kompatibel ist.

(2) Die Summe von zwei Linearkombinationen der $v_1, \ldots, v_n$ ist wieder eine:

$$\sum_{i=1}^n a_i v_i + \sum_{i=1}^n b_i v_i = \sum_{i=1}^n (a_i + b_i) v_i.$$

Die Skalierung einer Linearkombination der $v_1, \ldots, v_n$ ist wieder eine:

$$\lambda(\sum_{i=1}^n a_i v_i) = \sum_{i=1}^n (\lambda \cdot a_i) v_i.$$

Beispiel 4.23. Sei K ein Körper und $\mathscr{S}$ das homogene lineare Gleichungssystem

$$a_{i1}X_1 + \ldots + a_{in}X_n = 0$$

für $i = 1, \ldots, m$. Aus den Spalten der Koeffizienten vor den Variablen X_j machen wir Vektoren

$$v_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m.$$

Dann gilt für Vektoren $x \in K^n$:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathscr{L}(\mathscr{S}) \quad \iff \quad \sum_{j=1}^n x_j v_j = 0.$$

Das ist eine einfache Rechnung: $\sum_{j=1}^n x_j v_j = 0$ ist eine Rechnung in K^m. Ein Vektor ist genau dann gleich 0, wenn alle seine Einträge 0 sind. Der i-te Eintrag hier ist

$$\big(\sum_{j=1}^n x_j v_j\big)_i = \sum_{j=1}^n x_j a_{ij} = \sum_{j=1}^m a_{ij} x_j,$$

weil a_{ij} der i-te Eintrag von v_j ist.

Lösungen von $\mathscr{S}$ haben also etwas mit Koeffizienten von Linearkombinationen zu tun!

Definition 4.24. Seien V ein K-Vektorraum.

(1) Sei $M \subseteq V$ eine Teilmenge. Die **lineare Hülle** (oder der **Spann**) von M ist die Menge der Linearkombinationen aus Vektoren von M in V. Wir bezeichnen die lineare Hülle von M mit

$$\langle M\rangle_K = \left\{ \sum_{i=1}^n a_i v_i \;;\; n \in \mathbb{N}_0, a_i \in K, v_i \in M,\ i = 1, \ldots, n \right\} \subseteq V.$$

(2) Speziell für $M = \{v_1, \ldots, v_n\}$ gilt

$$\langle v_1, \ldots, v_n\rangle_K := \langle\{v_1, \ldots, v_n\}\rangle_K = \left\{ \sum_{i=1}^n a_i v_i \;;\; a_i \in K,\ i = 1, \ldots, n \right\}.$$

Beispiel 4.25. (1) Seien V ein K-Vektorraum und $v \in V$ ein Vektor. Dann ist die Menge der Vielfachen des Vektors v

$$\langle v \rangle_K = \{av \ ; \ a \in K\} \subseteq V$$

ein Unterraum. Das folgt aus Proposition 4.26.

(2) Es gilt stets $\{0\} = \langle \emptyset \rangle_K$, denn dies lineare Hülle enthält einzig die leere Linearkombination, deren Wert als 0 festgelegt wurde.

Proposition 4.26. *Seien V ein K-Vektorraum und $M \subseteq V$ eine Teilmenge.*

(1) Dann ist die lineare Hülle $\langle M \rangle_K$ ein Unterraum von V.

(2) Es gilt $M \subseteq \langle M \rangle_K$.

(3) Jeder Unterraum U, der M enthält, enthält $\langle M \rangle_K$:

$$M \subseteq U \qquad \Longrightarrow \qquad \langle M \rangle_K \subseteq U.$$

Beweis. (1) Wir weisen für $\langle M \rangle_K$ das Unterraumkriterium aus Satz 4.13 nach. Seien $x, y \in \langle M \rangle_K$ zwei Linearkombinationen. Durch eventuelles Ergänzen von Summanden $0 \cdot v_i$ können wir annehmen, daß x und y als Linearkombination derselben Vektoren $v_1, \ldots, v_n$ aus M dargestellt werden. Seien

$$x = \sum_{i=1}^{n} a_i v_i, \qquad y = \sum_{i=1}^{n} b_i v_i.$$

Damit ist für ein beliebiges $\lambda \in K$ die Summe

$$\lambda x + y = \lambda \cdot \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{n} b_i v_i = \sum_{i=1}^{n} (\lambda a_i + b_i) v_i$$

auch eine Linearkombination von Vektoren aus M.

Nun müssen wir noch zeigen, daß die lineare Hülle nicht leer ist. Aber die leere Summe ist in $\langle M \rangle_K$ enthalten, also $0 \in \langle M \rangle_K$.

(2) Jedes $v \in M$ ist wegen $v = 1 \cdot v$, eine Linearkombination aus M. Das zeigt $M \subseteq \langle M \rangle_K$.

(3) Sei U ein Unterraum von V, der M enthält. Addition und Skalarmultiplikation verlassen U nicht, weil U ein Unterraum ist. Also liegen alle Linearkombinationen von Vektoren aus M in U. Daher gilt $\langle M \rangle_K \subseteq U$. □

Korollar 4.27. *Sei V ein K-Vektorraum. Für Teilmengen $A, B \subseteq V$ gilt*

(1) $A \subseteq \langle B \rangle_K \iff \langle A \rangle_K \subseteq \langle B \rangle_K$,

(2) $A \subseteq B \implies \langle A \rangle_K \subseteq \langle B \rangle_K$.

Beweis. Aussage (1) folgt aus Proposition 4.26 (3) für $M = A$ und $U = \langle B \rangle_K$. Aussage (2) folgt wegen $B \subseteq \langle B \rangle_K$ aus (1). □

Proposition 4.28. *Sei V ein K-Vektorraum. Für Teilmengen $A, B \subseteq V$ gilt*

$$B \subseteq \langle A \rangle_K \qquad \Longleftrightarrow \qquad \langle A \rangle_K = \langle A \cup B \rangle_K.$$

Beweis. $\Longleftarrow$: Wegen $A \cup B \subseteq \langle A \cup B \rangle_K$ folgt aus der Voraussetzung $\langle A \rangle_K = \langle A \cup B \rangle_K$ die Inklusion

$$B \subseteq A \cup B \subseteq \langle A \cup B \rangle_K = \langle A \rangle_K.$$

$\Longrightarrow$: Wegen $A \subseteq A \cup B$ folgt aus Korollar 4.27 (2) die Inklusion $\langle A \rangle_K \subseteq \langle A \cup B \rangle_K$. Für die umgekehrte Inklusion nutzen wir $A \subseteq \langle A \rangle_K$, so daß mit der Voraussetzung $B \subseteq \langle A \rangle_K$ gilt:

$$A \cup B \subseteq \langle A \rangle_K.$$

Nach Korollar 4.27 (1) folgt dann $\langle A \cup B \rangle_K \subseteq \langle A \rangle_K$, und das war noch zu zeigen. □

Definition 4.29. Seien V ein K-Vektorraum.
(1)
(2) Sei $U \subseteq V$ ein Unterraum. Eine Teilmenge $M \subseteq U$ heißt **Erzeugendensystem** von U (oder nur **Erzeugendensystem**, wenn $U = V$), wenn
$$U = \langle M \rangle_K.$$
Man sagt, M **erzeugt** U **als** K**-Vektorraum**.
(3) Wir betrachten ein Tupel von Vektoren $(v_i)_{i \in I}$ als Erzeugendensystem, wenn die zugrundeliegende Menge $M = \{v_i \ ; \ i \in I\}$ ein Erzeugendensystem ist.

Definition 4.30. Seien K ein Körper, $n \in \mathbb{N}$ und $1 \leq i \leq n$. Der i-te **Standardbasisvektor** von K^n ist der Spaltenvektor
$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i,$$
dessen einzige von 0 verschiedene Koordinate eine 1 in der i-ten Zeile ist.

Beispiel 4.31. Die Menge $\{e_1, \ldots, e_n\}$ ist ein Erzeugendensystem für K^n,
$$K^n = \langle e_1, \ldots, e_n \rangle_K,$$
denn für alle $x_1, \ldots, x_n \in K$ gilt
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \ldots + x_n e_n.$$

4.4. **Konstruktionen mit Unterräumen.** Wir beschreiben nun zwei Konstruktionen, mittels derer man aus Unterräumen neue Unterräume machen kann: den Schnitt und die Summe von Unterräumen.

Lemma 4.32. *Sei V ein K-Vektorraum, und seien $U_i \subseteq V$, $i = 1, 2$ Unterräume von V. Dann ist der* ***Schnitt***
$$U_1 \cap U_2 := \{v \in V \ ; \ v \in U_i \text{ für } i = 1, 2\}$$
wieder ein Unterraum von V.

Beweis. Das folgt sofort aus dem Unterraumkriterium, Satz 4.13. Es gilt $0 \in U_i$ für $i = 1, 2$, also
$$0 \in U_1 \cap U_2$$
und der Schnitt ist nicht leer. Wenn $x, y \in U_1 \cap U_2$, und $\lambda \in K$, dann ist auch $\lambda x + y \in U_i$ für $i = 1, 2$, also
$$\lambda x + y \in U_1 \cap U_2.$$
□

Bemerkung 4.33. Analog gilt dies auch für eine Familie von Unterräumen $U_i \subseteq V$ für alle $i \in I$. Dann ist der **Schnitt**
$$\bigcap_{i \in I} U_i \subseteq V$$
wieder ein Unterraum.

Proposition 4.34. *Seien V ein K-Vektorraum und $M \subseteq V$ eine Teilmenge.*
Der Unterraum $\langle M\rangle_K$ ist bezüglich Inklusion der kleinste Unterraum von V, der M enthält: es gilt als Schnitt in V:

$$\langle M\rangle_K = \bigcap_{U \text{ Unterraum, } M\subseteq U} U.$$

Beweis. Aus Proposition 4.26 wissen wir bereits, daß $\langle M\rangle_K \subseteq U$ für alle Unterräume U, die M enthalten. Damit folgt

$$\langle M\rangle_K \subseteq \bigcap_{U \text{ Unterraum, } M\subseteq U} U.$$

Die umgekehrte Inklusion ist trivial, weil $\langle M\rangle_K$ selbst einer der Unterräume ist, die M enthalten und über die der Schnitt zu nehmen ist. □

Beispiel 4.35. Die Vereinigung von Unterräumen ist in der Regel kein Unterraum mehr. Betrachten wir in $\mathbb{R}^2$ die Unterräume

$$U_1 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle_{\mathbb{R}} = \{ \begin{pmatrix} x \\ 0 \end{pmatrix} \ ; \ x \in \mathbb{R}\}$$

und

$$U_2 = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle_{\mathbb{R}} = \{ \begin{pmatrix} 0 \\ y \end{pmatrix} \ ; \ y \in \mathbb{R}\},$$

dann ist

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in U_1 \cup U_2,$$

aber

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin U_1 \cup U_2.$$

Die Rolle der Vereinigung als kleinster Unterraum, der die beiden Unterräume enthält, übernimmt die Summe.

Definition 4.36. Seien V ein K-Vektorraum und $U_i \subseteq V$, $i = 1, 2$ Unterräume. Die **Summe** von U_1 und U_2 ist

$$U_1 + U_2 := \{x_1 + x_2 \ ; \ x_1 \in U_1 \text{ und } x_2 \in U_2\}.$$

Proposition 4.37. *Die Summe zweier Unterräume ist ein Unterraum.*

Beweis. Das folgt sofort aus dem Unterraumkriterium, Satz 4.13. Zunächst ist $0 \in U_i$, $i = 1, 2$, somit auch

$$0 = 0 + 0 \in U_1 + U_2,$$

was also nicht leer ist. Seien $x, y \in U_1 + U_2$. Dann gibt es $x_1, y_1 \in U_1$ und $x_2, y_2 \in U_2$ mit

$$x = x_1 + x_2 \text{ und } y = y_1 + y_2.$$

Daraus folgt dann für ein beliebiges $\lambda \in K$:

$$\lambda x + y = \lambda(x_1 + x_2) + (y_1 + y_2) = (\lambda x_1 + y_1) + (\lambda x_2 + y_2) \in U_1 + U_2.$$ □

Übungsaufgaben zu §4

Übungsaufgabe 4.1. Wofür brauchen Sie in Satz 4.13 die Bedingung $U \neq \emptyset$?

Übungsaufgabe 4.2. Wir betrachten den K-Vektorraum $\mathrm{Abb}(X, K)$ der K-wertigen Abbildungen auf einer Menge X. Der Träger einer Abbildung $f\colon X \to K$ ist die Menge $\{x \in X \;;\; f(x) \neq 0\}$.

Eine Abbildung $f\colon X \to K$ hat **endlichen Träger**, wenn $f(x) = 0$ für alle $x \in X$ bis auf endlich viele Ausnahmen.

Zeigen Sie, daß die Menge

$$\{f \in \mathrm{Abb}(X, K) \;;\; f \text{ hat endlichen Träger}\}$$

einen Untervektorraum von $\mathrm{Abb}(X, K)$ bilden. Zeigen Sie, daß dieser Unterraum von den Abbildungen $\mathbf{1}_y$ zu allen $y \in X$ aufgespannt wird:

$$\mathbf{1}_y(x) = \begin{cases} 1 & y = x \\ 0 & y \neq x. \end{cases}$$

Übungsaufgabe 4.3. Seien $v_1, \ldots, v_m \in K^n$. Zeigen Sie, daß für alle $\lambda \in K$ und $1 \leq \alpha, \beta \leq m$ mit $\alpha \neq \beta$

$$\langle v_1, \ldots, v_m \rangle_K = \langle v_1, \ldots, v_\alpha + \lambda v_\beta, \ldots, v_\beta, \ldots, v_m \rangle_K,$$

und wenn $\mu \in K^\times$, dann auch

$$\langle v_1, \ldots, v_m \rangle_K = \langle v_1, \ldots, \mu v_\alpha, \ldots, v_m \rangle_K.$$

Übungsaufgabe 4.4. Sei V ein K-Vektorraum und sei

$$U_1 \subseteq U_2 \subseteq U_3 \subseteq \ldots$$

eine aufsteigende Folge von Unterräumen von V. Zeigen Sie: die Vereinigung

$$U = \bigcup_{i \geq 1} U_i$$

ist ein Unterraum von V.

Übungsaufgabe 4.5. Sei V ein K-Vektorraum und seien U_1, U_2 Unterräume von V. Zeigen Sie

$$U_1 + U_2 = \langle U_1 \cup U_2 \rangle_K.$$

5. Basis und Dimension

5.1. **Lineare Unabhängigkeit.** Wir kommen nun zu einem zentralen Begriff der linearen Algebra.

Definition 5.1. Sei V ein K-Vektorraum.

(1) Ein Tupel $(v_1, \ldots, v_n)$ von Vektoren aus V heißt **linear abhängig**, wenn es eine nichttriviale Linearkombination der $v_1, \ldots, v_n$ gibt, die den Wert 0 ergibt, d.h. es gibt $a_i \in K$, $i = 1, \ldots, n$ und nicht alle $a_i = 0$ mit

$$a_1 v_1 + \ldots + a_n v_n = 0. \tag{5.1}$$

Andernfalls heißt $(v_1, \ldots, v_n)$ **linear unabhängig**, d.h. für alle $a_i \in K$, $i = 1, \ldots, n$ gilt

$$a_1 v_1 + \ldots + a_n v_n = 0 \quad \Longrightarrow \quad a_i = 0 \text{ für alle } i = 1, \ldots, n.$$

(2) Eine Menge $M \subseteq V$ von Vektoren heißt **linear abhängig**, wenn es eine nichttriviale Linearkombination von Vektoren aus M gibt, die den Wert 0 ergibt, d.h. es gibt paarweise verschiedene $v_1, \ldots, v_n \in M$ und Elemente $a_i \in K$, $i = 1, \ldots, n$ und nicht alle $a_i = 0$ mit

$$a_1 v_1 + \ldots + a_n v_n = 0.$$

Andernfalls heißt die Menge M **linear unabhängig**, d.h. für alle paarweise verschiedene $v_1, \ldots, v_n \in M$ und Elemente $a_i \in K$, $i = 1, \ldots, n$ gilt

$$a_1 v_1 + \ldots + a_n v_n = 0 \quad \Longrightarrow \quad a_i = 0 \text{ für alle } i = 1, \ldots, n.$$

Bemerkung 5.2. (1) Die triviale Linearkombination hat immer den Wert 0. Lineare Unabhängigkeit bedeutet also, daß nur die triviale Linearkombination den Wert 0 hat.

(2) Wenn man betonen möchte, daß die fraglichen Koeffizienten aus dem Körper K stammen, dann sagt man auch **K-linear abhängig** bzw. **K-linear unabhängig**.

(3) Eine Gleichung wie (5.1) nennt man eine **lineare Relation** der Vektoren $v_1, \ldots, v_n$.

(4) Allgemeine Methoden, mit denen man entscheiden kann, ob Vektoren linear unabhängig sind, lernen wir in Kapitel 9.5 kennen.

Beispiel 5.3. (1) Die Vektoren im $\mathbb{R}^3$

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad y = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \quad z = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$$

erfüllen $x - 2y + z = 0$. Sie sind linear abhängig.

(2) Sei V ein K-Vektorraum. Ein Vektor $v \in V$ ist linear unabhängig genau dann, wenn $v \neq 0$. Das folgt sofort aus Proposition 4.7. In der Tat ist für $v = 0$ die nichttriviale Linearkombination

$$1 \cdot 0 = 0,$$

somit 0 linear abhängig. Und bei $v \neq 0$ folgt aus $a \cdot v = 0$ bereits $a = 0$, somit ist v linear unabhängig.

(3) Sei $\mathscr{S}$ ein homogenes lineares Gleichungssystem und seien $v_1, \ldots, v_n \in K^m$ die zugehörigen Spaltenvektoren wie in Beispiel 4.23. Dann gibt es eine nichttriviale Lösung $0 \neq x \in K^n$ von $\mathscr{S}$ genau dann, wenn die Vektoren $v_1, \ldots, v_n$ linear abhängig sind.

Das folgt sofort aus Beispiel 4.23, denn die Lösungen $x \in \mathscr{L}(\mathscr{S})$ sind genau die Koeffizienten derjenigen Linearkombinationen

$$x_1 v_1 + \ldots + x_n v_n,$$

die den Wert 0 haben.

(4) Wenn für die Vektoren im $\mathbb{Q}^3$

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad y = \begin{pmatrix} 4 \\ 6 \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} 7 \\ 0 \\ 0 \end{pmatrix}$$

eine Linearkombination $ax + by + cz = 0$ für $a, b, c \in \mathbb{Q}$ gilt, dann übersetzt sich das in Umkehrung zu (2) zum linearen Gleichungssystem

$$\left\{ \begin{array}{rcrcrcl} a & + & 4b & + & 7c & = & 0 \\ 2a & + & 6b & & & = & 0 \\ 3a & & & & & = & 0 \end{array} \right.$$

und das hat ersichtlich nur die Lösungen $a = b = c = 0$. Also sind x, y, z linear unabhängig in $\mathbb{Q}^3$.

Bemerkung 5.4. Sei V ein K-Vektorraum.

(1) Lineare Unabhängigkeit ist eine Eigenschaft des ganzen Tupels. Es kommt auf jeden Vektor an. Die lineare Unabhängigkeit des Tupels $(v_1, \ldots, v_n)$ ist nicht dasselbe wie die lineare Unabhängigkeit der einzelnen Vektoren v_i. Letztere müssen nur von 0 verschieden sein, um linear unabhängig mit sich selbst zu sein. Für lineare Unabhängigkeit als Tupel darf es keine lineare Relation geben!

(2) Seien $v_1, \ldots, v_n$ Vektoren aus V. Daß das Tupel $(v_1, \ldots, v_n)$ linear unabhängig ist, ist nicht dasselbe wie die lineare Unabhängigkeit der zugrundeliegenden Menge $\{v_1, \ldots, v_n\}$. Wenn es $i \neq j$ mit $v_i = v_j$ gibt, dann ist das Tupel nicht linear unabhängig, wie die nichttriviale Linearkombination

$$0 = v_i - v_j = 1 \cdot v_i + (-1) v_j$$

zeigt. In der Menge aber kommt der Vektor $v_i = v_j$ nur einmal vor. Diese Linearkombination steht also nicht zur Verfügung! Die zugrundeliegende Menge kann trotzdem noch linear unabhängig sein.

(3) Oft sprechen wir ungenau davon, daß die Vektoren $v_1, \ldots, v_n \in V$ linear unabhängig sind und meinen dabei die lineare Unabhängigkeit des Tupels $(v_1, \ldots, v_n)$. Das ist nicht so schlimm, weil aus dem Kontext klar wird, was gemeint ist.

(4) Sei $M \subseteq V$ eine Teilmenge. Dann folgt sofort aus der Definition, daß M linear unabhängig ist genau dann, wenn für alle $n \in \mathbb{N}$ und alle paarweise verschiedenen $v_1, \ldots, v_n \in M$ das Tupel $(v_1, \ldots, v_n)$ linear unabhängig ist.

(5) Die leere Menge $\emptyset \subseteq V$ ist linear unabhängig, denn es gibt keine nichttriviale Linearkombination mit Vektoren aus $\emptyset$, welche die lineare Abhängigkeit bezeugen könnte.

Jetzt zeigen wir, daß wir in einer nichttrivialen Linearkombination nach einem der beteiligten Vektoren „auflösen“ können.

Proposition 5.5. *Sei V ein K-Vektorraum.*

(1) Vektoren $v_1, \ldots, v_n \in V$ sind genau dann linear abhängig, wenn einer der Vektoren als Linearkombination der anderen geschrieben werden kann, d.h. es gibt $1 \leq j \leq n$ und $a_i \in K$ für $i = 1, \ldots, n$ und $j \neq i$ mit

$$v_j = \sum_{i=1,\ i\neq j}^{n} a_i v_i.$$

(2) Eine Teilmenge $M \subseteq V$ ist genau dann linear abhängig, wenn es ein $v \in M$ gibt, das als Linearkombination der anderen geschrieben werden kann, d.h.

$$v \in \langle M \setminus \{v\} \rangle_K.$$

Beweis. Die Aussage (2) folgt sofort aus Aussage (1), die wir jetzt beweisen. Wir nehmen zunächst an, daß

$$v_j = \sum_{i=1,\ i\neq j}^{n} a_i v_i.$$

Dann gilt

$$v_j + \sum_{i=1,\ i\neq j}^{n} (-a_i) v_i = 0,$$

und weil der Koeffizient von v_j mit $1 \neq 0$ in K nichttrivial ist, handelt es sich um eine nichttriviale Linearkombination, welche 0 als Wert hat. Damit sind $v_1, \ldots, v_n$ linear abhängig.

Für die umgekehrte Richtung nehmen wir an, daß $v_1, \ldots, v_n$ linear abhängig sind. Dann gibt es also $a_i \in K$ für $i = 1, \ldots, n$ mit

$$a_1 v_1 + \ldots + a_n v_n = 0$$

und nicht alle $a_i = 0$. Sei j ein Index mit $a_j \neq 0$. Wir können dann nach v_j auflösen:

$$v_j = a_j^{-1}(a_j v_j) = a_j^{-1}\Big(- \sum_{i=1,\ i\neq j}^{n} a_i v_i \Big) = \sum_{i=1,\ i\neq j}^{n} (-a_j^{-1} a_i) v_i,$$

und somit ist v_j der Wert einer Linearkombination der anderen Vektoren. □

Beispiel 5.6. Sei V ein K-Vektorraum. Aus Proposition 5.5 folgt: Zwei Vektoren $v, w \in V$ sind linear abhängig $\iff$ $v \in \langle w \rangle_K$ oder $w \in \langle v \rangle_K$ und das bedeutet, daß v Vielfaches von w oder w Vielfaches von v ist.

Das ist besonders im $\mathbb{R}^2$ instruktiv, wenn $v, w \neq 0$ sind. Dann ist die lineare Hülle von v bzw. w eine Gerade in der Ebene und die Vektoren sind genau dann linear abhängig, wenn v und w dieselbe Gerade aufspannen.

5.2. **Eindeutige Darstellung.** Eine Teilmenge M eines K-Vektorraums V ist ein Erzeugendensystem, wenn sie „groß genug“ ist, so daß jedes $v \in V$ eine Linearkombination von Vektoren aus M ist. Der Begriff der linearen Unabhängigkeit geht in die umgekehrte Richtung: die Menge ist „klein genug“ im folgenden Sinne.

Proposition 5.7. *Sei V ein K-Vektorraum und seien $v_1, \dots, v_n \in V$. Dann sind äquivalent:*

(a) *Die $v_1, \dots, v_n$ sind linear unabhängig.*

(b) *Jedes $x \in \langle v_1, \dots, v_n \rangle_K$ hat eine eindeutige Darstellung als Linearkombination der Vektoren $v_1, \dots, v_n$.*

Beweis. (a) $\Longrightarrow$ (b): Sei $x \in \langle v_1, \dots, v_n \rangle_K$. Die Existenz einer Darstellung als Linearkombination folgt per Definition der linearen Hülle. Wir müssen die Eindeutigkeit zeigen. Dazu nehmen wir an wir hätten zwei, also $a_i, b_i \in K$ für $i = 1, \dots, n$ mit

$$\sum_{i=1}^{n} a_i v_i = x = \sum_{i=1}^{n} b_i v_i.$$

Ziehen wir beide Relationen voneinander ab, so erhalten wir

$$0 = x - x = \sum_{i=1}^{n} a_i v_i - \sum_{i=1}^{n} b_i v_i = \sum_{i=1}^{n} (a_i - b_i) v_i.$$

Weil die $v_1, \dots, v_n$ linear unabhängig sind, folgt $a_i - b_i = 0$ für alle $i = 1, \dots, n$. Das bedeutet aber $a_i = b_i$, es waren doch keine zwei verschiedenen Linearkombinationen.

(b) $\Longrightarrow$ (a): Die 0 ist der Wert der trivialen Linearkombination $0 = \sum_{i=1}^{n} 0 \cdot v_i$. Aufgrund der Eindeutigkeit sind die Werte aller nichttrivialen Linearkombinationen $\neq 0$. Damit sind $v_1, \dots, v_n$ per Definition linear unabhängig. □

Korollar 5.8. *Sei V ein K-Vektorraum und seien $v_1, \dots, v_n \in V$ linear unabhängige Vektoren. Dann folgt für Koeffizienten $a_i, b_i \in K$, für $i = 1, \dots, n$:*

$$\sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{n} b_i v_i \quad \Longrightarrow \quad a_i = b_i \text{ für alle } i = 1, \dots, n.$$

*Dies nennt man **Koeffizientenvergleich**.*

Beweis. Das folgt sofort aus Proposition 5.7. □

Wir kommen nun zu einem zentralen Begriff der Theorie der Vektorräume: der Basis. Eine Basis wird ein Tupel von Vektoren eines Vektorraums V mit sehr speziellen Eigenschaften sein. Weil es sein kann, daß es unendlich viele Vektoren in einer Basis braucht, müssen wir zuerst den Begriff der linearen Unabhängigkeit auf unendliche Tupel ausweiten. Dazu müssen wir den intuitiven Begriff, ein Tupel ist in einem anderen enthalten bzw. ein Tupel ist größer/kleiner als ein anderes, formal beschreiben.

Definition 5.9.

(1) Wir sagen, ein I-Tupel $(x_i)_{i \in I}$ **enthält** ein J-Tupel $(y_j)_{j \in J}$, oder gleichbedeutend das J-Tupel $(y_j)_{j \in J}$ **ist** in dem I-Tupel $(x_i)_{i \in I}$ **enthalten**, wenn $J \subseteq I$ eine Teilmenge ist und für alle $j \in J$ gilt: $x_j = y_j$.

Wir sagen **enthält echt** bzw. **ist echt enthalten**, wenn darüberhinaus $J \neq I$.

(2) Sei K ein Körper und sei V ein K-Vektorraum. Sei I eine Menge. Ein I-Tupel $(v_i)_{i \in I}$ von Vektoren aus V heißt **linear unabhängig**, wenn für jedes $J \subseteq I$ mit $|J|$ endlich das Tupel von Vektoren $(v_j)_{j \in J}$, das in $(v_i)_{i \in I}$ enthalten ist, linear unabhängig ist.

Wenn das I-Tupel $(v_i)_{i\in I}$ nicht linear unabhängig ist, dann nennen wir es **linear abhängig**. Für endliche Tupel stimmt diese neue Definition mit der alten überein, wie man sich leicht überlegt.

Beispiel 5.10. Sei $I \subseteq \{1,\dots,n\}$, dann ist $(x_i)_{i\in I}$ ein Tupel, das in $(x_1,\dots,x_n)$ enthalten ist; und sogar echt enthalten ist, wenn $|I| < n$. Ist $I = \{i_1,\dots,i_r\}$, dann ist $(x_i)_{i\in I} = (x_{i_1},\dots,x_{i_r})$.

Definition 5.11. Seien K ein Körper und V ein K-Vektorraum. Eine **Basis** von V ist ein Tupel von Vektoren aus V, das
(i) linear unabhängig und
(ii) ein Erzeugendensystem von V ist.
Die Basis heißt **endlich**, wenn das Tupel ein endliches Tupel $(v_1,\dots,v_n)$ ist.

Bemerkung 5.12. Ist $\mathscr{B} = (v_1,\dots,v_n)$ eine Basis, so schreiben wir oft nur $v_1,\dots,v_n$ für die Basis $\mathscr{B}$. Die Reihenfolge, auf die es bei Koordinaten und Matrizen in Bezug auf die Basis $\mathscr{B}$ ankommt, wird dann als implizit gegeben angesehen. In der Regel verursacht der sorglose Umgang mit dem Unterschied zwischen dem Tupel $(v_1,\dots,v_n)$ und $v_1,\dots,v_n$ keine Probleme.

Beispiel 5.13. (1) Den Körper $\mathbb{Q}[\sqrt{2}]$ aus Übungsaufgabe 3.9 kann man als $\mathbb{Q}$-Vektorraum auffassen. Die Elemente $1, \sqrt{2}$ bilden eine Basis von $\mathbb{Q}[\sqrt{2}]$ als $\mathbb{Q}$-Vektorraum.
(2) In Beispiel 4.31 haben wir bereits gesehen, daß die Standardbasisvektoren $e_1,\dots,e_n$ den K^n erzeugen. Die lineare Unabhängigkeit folgt aus der Rechnung:

$$0 = a_1e_1 + \ldots + a_ne_n = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

dann $a_i = 0$ für alle $i = 1,\dots,n$. Damit sind $e_1,\dots,e_n \in K^n$ eine Basis von K^n. Diese nennen wir die **Standardbasis** des K^n, und das erklärt den Namen Standardbasisvektor für die e_i.
(3) Im Nullvektorraum $K^0 = \{0\}$ ist die leere Menge $\emptyset$ eine Basis.

Korollar 5.14. *Sei V ein K-Vektorraum und seien $v_1,\dots,v_n \in V$. Dann sind äquivalent:*
(a) Die $v_1,\dots,v_n$ sind eine Basis von V.
(b) Jedes $x \in V$ hat eine eindeutige Darstellung als Linearkombination der $v_1,\dots,v_n$.

Beweis. (a) $\Longrightarrow$ (b): Bei einer Basis ist $V = \langle v_1,\dots,v_n\rangle_K$. Damit folgt diese Richtung des Korollars sofort aus Proposition 5.7.

(b) $\Longrightarrow$ (a): Wenn jedes $x \in V$ eine Linearkombination der $v_1,\dots,v_n$ ist, dann ist $v_1,\dots,v_n$ ein Erzeugendensystem. Die Eindeutigkeit der darstellenden Linearkombination zeigt mittels Proposition 5.7, daß die $v_1,\dots,v_n$ linear unabhängig sind. Dies zeigt (a). □

Eine Basis ist nicht zu groß, weil aus linear unabhängigen Vektoren bestehend, und nicht zu klein, weil ein Erzeugendensystem.

Definition 5.15 (Koordinatenn)**.** Korollar 5.14 besagt, daß für eine Basis $\mathscr{B} = (v_1,\dots,v_n)$ von V die Abbildung $K^n \to V$

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \sum_{i=1}^{n} a_iv_i$$

bijektiv ist (surjektiv wegen erzeugend, und injektiv wegen linear unabhängig). Wir können vermöge dieser Abbildung die Vektoren $w \in V$ über ihre **Koordinaten** aus K^n ansprechen:

Für $w \in V$ mit $w = \sum_{i=1}^{n} a_i v_i$ nennen wir a_i die **i-te Koordinate** von w bezüglich der Basis $\mathscr{B}$. Mehr zu Koordinaten für einen Vektorraum erfahren wir in Kapitel 7.1.

5.3. **Die Existenz einer Basis.** Wir zeigen nun, daß Tupel von Vektoren mit speziellen extremalen Eigenschaften eine Basis sind.

Proposition 5.16 (Minimale Erzeugendensysteme sind linear unabhängig)**.** *Seien K ein Körper und V ein K-Vektorraum. Ein Tupel $(v_i)_{i\in I}$ von Vektoren aus V, das*
(i) ein Erzeugendensystem von V ist, und so dass
(ii) jedes echt in $(v_i)_{i\in I}$ enthaltene Tupel kein Erzeugendensystem ist,
ist eine Basis von V.

Beweis. Wir müssen zeigen, daß das Tupel $(v_i)_{i\in I}$ linear unabhängig ist und beweisen dies durch Widerspruch. Angenommen $(v_i)_{i\in I}$ wäre linear abhängig. Dann gibt es nach Proposition 5.5 ein $i_0 \in I$ so daß

$$v_{i_0} \in \langle v_i \ ; \ i \in I, i \neq i_0 \rangle_K.$$

Mit Proposition 4.28 für $A = \{v_i \ ; \ i \in I, i \neq i_0\}$ und $B = \{v_{i_0}\} \subseteq \langle A \rangle_K$ schließen wir

$$\langle v_i, i \in I, i \neq i_0 \rangle_K = \langle A \rangle_K = \langle A \cup B \rangle_K = \langle v_i, i \in I \rangle_K = V$$

und das ist ein Widerspruch zu (ii). □

Proposition 5.17 (Maximale linear unabhängige Tupel sind Erzeugendensysteme)**.** *Seien K ein Körper und V ein K-Vektorraum. Ein Tupel $(v_i)_{i\in I}$ von Vektoren aus V, das*
(i) linear unabhängig ist und so dass
(ii) jedes Tupel, das $(v_i)_{i\in I}$ echt enthält, linear abhängig ist,
ist eine Basis von V.

Beweis. Wir müssen zeigen, daß das Tupel $(v_i)_{i\in I}$ ein Erzeugendensystem ist und beweisen dies durch Widerspruch. Angenommen es gibt

$$w \in V \setminus \langle v_i, i \in I \rangle_K,$$

dann ist nach Voraussetzung das um w erweiterte Tupel linear abhängig. Formal handelt es sich um das Tupel $(v_j)_{j\in J}$ mit $J = I \cup \{*\}$ mit einem Index $* \notin I$ und $v_* = w$.

Es gibt also eine nichttriviale Linearkombination mit dem Wert 0. In dieser muß $w = v_*$ vorkommen, denn die restlichen Vektoren sind ja linear unabhängig. Diese Relation sei

$$aw + \sum_{k=1}^{n} a_k w_k = 0$$

mit $a, a_k \in K$ und $w_1, \ldots, w_n$ aus dem Tupel $(v_i)_{i\in I}$. Dabei gilt $a \neq 0$. Dann folgt wie im Beweis von Proposition 5.5

$$w = \sum_{k=1}^{n} (-a^{-1} a_k) w_k \ \in \ \langle w_1, \ldots, w_n \rangle_K \subseteq \langle v_i, i \in I \rangle_K,$$

ein Widerspruch zur Wahl von w. □

Im folgenden Satz sind „Maximales linear unabhängiges Tupel“ und „minimales Erzeugendensystem“ wie in den Propositionen 5.16 und 5.17 zu verstehen.

Satz 5.18. *Seien K ein Körper, V ein K-Vektorraum und I eine Menge.*
Für ein I-Tupel $(v_i)_{i\in I}$ von Vektoren aus V sind äquivalent:
(a) $(v_i)_{i\in I}$ ist Basis.

(b) $(v_i)_{i\in I}$ *ist ein maximales linear unabhängiges Tupel.*
(c) $(v_i)_{i\in I}$ *ist ein minimales Erzeugendensystem.*

Beweis. (b) $\Longrightarrow$ (a) ist Propositionen 5.17 und (c) $\Longrightarrow$ (a) ist Propositionen 5.16.

(a) $\Longrightarrow$ (b): Wenn $(v_i)_{i\in I}$ eine Basis ist, dann ist es ein linear unabhängiges Tupel. Es bleibt zu zeigen, daß jedes echt größere Tupel, das $(v_i)_{i\in I}$ enthält, nicht mehr linear unabhängig sein kann.

Angenommen es kommt der Vektor $w \in V$ hinzu. Dann ist $w \in V = \langle v_i \ ; \ i \in I\rangle_K$ eine Linearkombination von Vektoren aus $(v_i)_{i\in I}$. Nach Proposition 5.5 ist dann das erweiterte Tupel linear abhängig, und das war zu zeigen.

(a) $\Longrightarrow$ (c): Wenn $(v_i)_{i\in I}$ eine Basis ist, dann ist die zugrundeliegende Menge ein Erzeugendensystem. Es bleibt zu zeigen, daß jedes echt kleinere Tupel, das in $(v_i)_{i\in I}$ enthalten ist, nicht mehr Erzeugendensystem ist.

In einem echt enthaltenen Tupel fehlt mindestens ein Vektor, sagen wir zu $i_0 \in I$ fehlt der Vektor v_{i_0}. Weil $(v_i)_{i\in I}$ linear unabhängig ist, folgt aus Proposition 5.5, daß

$$v_{i_0} \notin \langle v_i \ ; \ i \in I \setminus \{i_0\}\rangle_K,$$

und damit ist das kleinere Tupel nicht mehr Erzeugendensystem (v_{i_0} wird nicht als Linearkombination dargestellt). Dies war zu zeigen. □

Definition 5.19. Ein Vektorraum heißt **endlich erzeugt**, falls er eine endliche Teilmenge hat, die ein Erzeugendensystem ist.

Theorem 5.20 (Existenz von Basen)**.** *Sei K ein Körper.*
(1) Jeder K-Vektorraum hat eine Basis.
(2) Jeder endlich erzeugte K-Vektorraum hat eine Basis aus endlich vielen Vektoren.

Beweis. Wir zeigen nur (2) und verweisen für (1) auf Appendix C, weil diese Aussage das (wichtige) Lemma von Zorn aus der Mengenlehre braucht. Aussage (2) folgt sofort aus dem gleich folgenden Satz 5.22. □

Satz 5.21. *Sei V ein endlich erzeugter K-Vektorraum. Seien $v_1, \ldots, v_r \in V$ linear unabhängige Vektoren, und seien $w_1, \ldots, w_n \in V$, so daß $v_1, \ldots, v_r, w_1, \ldots, w_n$ ein Erzeugendensystem von V ist. Dann gibt es $1 \leq i_1 < \ldots < i_s \leq n$, so daß*

$$v_1, \ldots, v_r, w_{i_1}, \ldots, w_{i_s}$$

eine Basis von V ist.

Beweis. Wir betrachten die linear unabhängigen Tupel, die

- in $(v_1, \ldots, v_r, w_1, \ldots, w_n)$ enthalten sind, und
- die $(v_1, \ldots, v_r)$ enthalten.

Das ist eine nichtleere Menge von Tupeln, denn $(v_1, \ldots, v_r)$ erfüllt die Bedingungen. Wir wählen darunter ein Tupel maximaler Länge (dies gibt es, denn die Länge von Tupeln ist begrenzt durch $r + n$). Dies hat die Form

$$(v_1, \ldots, v_r, w_{i_1}, \ldots, w_{i_s})$$

wie in der Behauptung, und wir müssen zeigen, daß dies ein Erzeugendensystem ist. Aufgrund der Maximalität sind für alle $1 \leq j \leq n$ die Vektoren

$$w_j, v_1, \ldots, v_r, w_{i_1}, \ldots, w_{i_s}$$

linear abhängig (entweder durch Wiederholung $w_j = w_{i_\alpha}$, oder weil das Tupel dann nach Umordnung in $v_1, \dots, v_r, w_1, \dots, w_n$ enthalten und länger wäre). In einer nichttrivialen Linearkombination, die 0 darstellt, muß w_j mit Koeffizient $\neq 0$ vorkommen, weil sonst bereits $v_1, \dots, v_r, w_{i_1}, \dots, w_{i_s}$ linear abhängig wäre im Widerspruch zur Auswahl. Daher gilt wie im Beweis von Proposition 5.17

$$w_j \in \langle v_1, \dots, v_r, w_{i_1}, \dots, w_{i_s} \rangle_K,$$

und das eben für alle $j = 1, \dots, n$. Daraus schließen wir mit Korollar 4.27 (1)

$$V = \langle v_1, \dots, v_r, w_1, \dots, w_n \rangle_K \subseteq \langle v_1, \dots, v_r, w_{i_1}, \dots, w_{i_s} \rangle_K \subseteq V,$$

woraus folgt, daß $v_1, \dots, v_r, w_{i_1}, \dots, w_{i_s}$ ein Erzeugendensystem ist. □

Wir folgern zwei unmittelbare Korollare.

Satz 5.22 (Basisauswahlsatz). *Seien K ein Körper und V ein K-Vektorraum. Dann enthält jedes endliche Erzeugendensystem von V eine Basis.*

Beweis. Das ist Satz 5.21 im Fall $r = 0$. □

Satz 5.23 (Basisergänzungssatz). *Sei V ein endlich erzeugter K-Vektorraum. Dann läßt sich jedes linear unabhängige Tupel von Vektoren aus V zu einer Basis ergänzen.*

Beweis. Seien $v_1, \dots, v_r \in V$ linear unabhängig, und sei $w_1, \dots, w_n$ ein endliches Erzeugendensystem von V gemäß Voraussetzung. Die Aussage folgt nun sofort aus Satz 5.21. □

5.4. **Der Basissatz.** Wir legen nun den Grundstein für den so wichtigen Dimensionsbegriff.

Theorem 5.24 (Basissatz). *Je zwei Basen eines Vektorraums V sind gleich mächtig.*

Beweis. Wir beweisen das mit Satz 5.26, und zwar nur für Vektorräume, die ein endliches Erzeugendensystem haben.

Für den allgemeinen Fall benötigt man das Konzept der gleichen **Mächtigkeit von Mengen** der (unendlichen) Mengenlehre. Und weiter benötigt man das Lemma von Zorn, siehe Anhang C; diesen Fall kann man daher getrost auf eine Zeit verschieben, wenn man die Grundlagen der Mathematik bereits besser verinnerlicht hat. □

Zuvor beweisen wir den Austauschsatz, einer Variante des Austauschlemmas von Steinitz.

Satz 5.25 (Austauschsatz). *Sei V ein K-Vektorraum, und seien $\mathscr{B}$ und $\mathscr{C}$ Basen von V. Sei b ein Basisvektor aus $\mathscr{B}$. Dann gibt es eine Basisvektor c aus der Basis $\mathscr{C}$, so daß der Austausch von b durch c in der Basis $\mathscr{B}$ weiterhin eine Basis von V ist.*

Beweis. Wir bezeichnen mit $\mathscr{B}_b$ das um den Eintrag b verkleinerte Tupel $\mathscr{B}$. Weil $\mathscr{B}$ eine Basis ist, läßt sich jeder Vektor aus $\mathscr{C}$ als Linearkombination von Vektoren aus $\mathscr{B}$ schreiben. Wenn in keiner dieser Linearkombinationen der Vektor b mit Koeffizient $\neq 0$ auftritt, dann reichen bereits die Vektoren aus $\mathscr{B}_b$ und es gilt

$$V = \langle \mathscr{C} \rangle_K \subseteq \langle \mathscr{B}_b \rangle_K.$$

Das ist ein Widerspruch, denn $\mathscr{B}_b$ ist kein Erzeugendensystem: $\mathscr{B}$ ist als Basis ein minimales Erzeugendensystem nach Satz 5.18.

Sei $c \in \mathscr{C}$ ein solcher Vektor, in dessen Darstellung als Linearkombination von Vektoren aus $\mathscr{B}$ der Vektor b vorkommt: für den es paarweise verschiedenen Vektoren $b = b_1, b_2, \dots, b_n$ aus $\mathscr{B}$ und Koeffizienten $\gamma_1, \dots, \gamma_n \in K$ gibt mit $\gamma_1 \neq 0$ und

$$c = \gamma_1 b + \sum_{i=2}^{n} \gamma_i b_i. \tag{5.2}$$

Sei $\mathscr{B}'$ das Tupel, das aus $\mathscr{B}$ durch Austausch des Eintrags b mit dem Vektor c entsteht. Wir zeigen nun, daß $\mathscr{B}'$ eine Basis ist.

$\mathscr{B}'$ ist Erzeugendensystem: Wegen $\gamma_1 \neq 0$ kann man die Gleichung (5.2) nach b auflösen:

$$b = \gamma_1^{-1}c - \sum_{i=2}^{n} \gamma_1^{-1}\gamma_i b_i.$$

Daher gilt $b \in \langle \mathscr{B}' \rangle_K$. Mittels Proposition 4.28 und Korollar 4.27 (1) schließen wir

$$\langle \mathscr{B}' \rangle_K = \langle \mathscr{B}', b \rangle_K \supseteq \langle \mathscr{B} \rangle_K = V,$$

und somit ist $\mathscr{B}'$ ein Erzeugendensystem.

$\mathscr{B}'$ ist linear unabhängig: Angenommen, es gibt eine lineare Relation der Vektoren von $\mathscr{B}'$. Indem wir zur Not Summanden mit dem Koeffizienten 0 hinzunehmen, dürfen wir diese Relation schreiben als

$$0 = \mu c + \sum_{i=2}^{m} \lambda_i b_i$$

mit $m \geq n$ und Vektoren $b_2, \ldots, b_m$ aus $\mathscr{B}_b$, welche die in der Gleichung (5.2) vorkommenden ergänzen (aber nicht $b = b_1$), und $\mu, \lambda_2, \ldots, \lambda_m \in K$. Jetzt nutzen wir für c den Ausdruck (5.2) und sortieren um:

$$0 = \mu\bigl(\gamma_1 b + \sum_{i=2}^{n} \gamma_i b_i\bigr) + \sum_{i=2}^{m} \lambda_i b_i = \mu\gamma_1 b + \sum_{i=2}^{n} (\lambda_i + \mu\gamma_i) b_i + \sum_{i=n+1}^{m} \lambda_i b_i.$$

Dies ist nun eine Relation zwischen Basisvektoren aus $\mathscr{B}$, folglich eine triviale Linearkombination. Daher gilt

$$\begin{aligned} \mu\gamma_1 &= 0, \\ \lambda_i + \mu\gamma_i &= 0 \qquad \text{für } i = 2, \ldots, n, \\ \lambda_i &= 0 \qquad \text{für } i = n+1, \ldots, m. \end{aligned}$$

Aus $\gamma_1 \neq 0$ folgt zunächst $\mu = 0$ und dann auch $\lambda_i = 0$ für alle i. Damit ist nur die triviale Linearkombination von Vektoren aus $\mathscr{B}'$ gleich 0: das Tupel $\mathscr{B}'$ ist linear unabhängig. □

Satz 5.26. *Sei V ein K-Vektorraum, und seien $\mathscr{B}$ und $\mathscr{C}$ Basen von V. Dann ist $\mathscr{B}$ endlich genau dann, wenn $\mathscr{C}$ endlich ist, und in diesem Falle haben $\mathscr{B}$ und $\mathscr{C}$ gleich viele Vektoren.*

Beweis. Aus Symmetriegründen reicht es anzunehmen, daß $\mathscr{B}$ endlich ist, und dann zu beweisen, daß $\mathscr{C}$ ebenfalls endlich ist mit $|\mathscr{C}| = |\mathscr{B}|$.

Wir führen Induktion nach der Anzahl r der Vektoren aus $\mathscr{B}$, welche keine Vektoren aus $\mathscr{C}$ sind. Die Induktionsverankerung bei $r = 0$ bedeutet, daß $\mathscr{B}$ ein Teiltupel der Basis $\mathscr{C}$ und selbst eine Basis ist. Nach Satz 5.18 geht das nur, wenn $\mathscr{B}$ bis auf eventuelle Umordnung des Tupels mit $\mathscr{C}$ übereinstimmt. In diesem Fall ist $\mathscr{C}$ auch endlich und $|\mathscr{B}| = |\mathscr{C}|$.

Wenn $r > 0$ ist, dann gibt es in $\mathscr{B}$ einen Vektor b, der nicht in $\mathscr{C}$ vorkommt. Nach dem Austauschsatz, Satz 5.25 gibt es einen Vektor c aus $\mathscr{C}$, so daß für das Tupel $\mathscr{B}'$, welches aus $\mathscr{B}$ durch den Austausch von b mit c entsteht, das folgende gilt:

- $\mathscr{B}'$ ist ebenfalls eine Basis,
- mit gleich vielen Elementen: $|\mathscr{B}'| = |\mathscr{B}|$, und
- die Anzahl r' der Vektoren aus $\mathscr{B}'$, welche keine Vektoren aus $\mathscr{C}$ sind, ist um eins kleiner geworden: $r' = r - 1 < r$.

Per Induktionsannahme hat dann $\mathscr{B}'$ genausoviele Vektoren wie $\mathscr{C}$ und wegen

$$|\mathscr{B}| = |\mathscr{B}'| = |\mathscr{C}|$$

gilt das auch für $\mathscr{B}$ und $\mathscr{C}$. Damit ist der Induktionsschritt vollzogen. □

Bemerkung 5.27. Im Beweis von Satz 5.26 kommt das Prinzip der vollständigen Induktion zum Einsatz. Der Parameter r, über den die Induktion strukturiert ist, erweist sich im Nachhinein als nach oben beschränkt durch die allen Basen gemeinsame Mächtigkeit. Mehr Vektoren der ersten Basis als diese Anzahl können nicht in der zweiten Basis fehlen. Was ist mit der **Vollständigkeit** der Induktion passiert, die ja die Wahrheit einer Aussage A_r für alle $r \in \mathbb{N}_0$ beweist? Nun, die Aussage A_r ist von dem Typ

$$A_r = \text{„Wenn es in } \mathscr{B} \text{ genau } r \text{ Vektoren gibt mit } \ldots, \text{ dann gilt } \ldots\text{“}$$

Diese Aussage wird durch die Induktion tatsächlich für alle r bewiesen. Aber es handelt sich um eine „Wenn-Dann“-Aussage. Sie ist auch dann wahr, falls der Wenn-Teil gar nicht zutrifft wie hier spätestens ab $r > |\mathscr{B}|$.

Definition 5.28. Sei V ein K-Vektorraum. Die **Dimension** von V ist die Mächtigkeit einer Basis von V. Wir bezeichnen die Dimension von V mit

$$\dim V = \dim(V) = \dim_K(V),$$

letzteres, wenn wir betonen wollen, daß es sich um einen K-Vektorraum handelt.

Bemerkung 5.29. Die Dimension ist wohldefiniert, weil

- jeder Vektorraum eine Basis hat, Theorem 5.20, und
- weil je zwei Basen eines Vektorraums V gleich mächtig sind, Theorem 5.24.

Die Dimension muß nicht endlich sein. In diesem Fall ist die Dimension eine **Kardinalzahl** und kann erst nach einer Einführung in Mengenlehre formal behandelt werden. Wir begnügen uns hier damit, daß die Existenz einer endlichen Basis von V die Endlichkeit aller Basen von V nach sich zieht. In dem Sinne ist hier

$$\dim(V) \in \mathbb{N}_0 \cup \{\infty\}.$$

Beispiel 5.30. (1) Es gilt $\dim(K^n) = n$, denn $e_1, \ldots, e_n$ ist Basis von K^n.

(2) Eine Polynomfunktion $f(x) = a_n x^n + \ldots + a_1 x + a_0$ hat, sofern nicht alle $a_i = 0$ sind, einen **Grad**

$$\deg(f) := \max\{i \ ; \ a_i \neq 0\}.$$

Der Grad ist der größte auftretende Exponent d, so daß x^d mit Vorfaktor $a_d \neq 0$ auftritt. Wir müssen uns später noch darum kümmern, daß dies wohldefiniert ist, und was für das Nullpolynom für ein Grad festgelegt werden soll. Für den Moment reicht es, wenn wir unter

$$\mathrm{Pol}_{\mathbb{R},\leq d} = \left\{ f\colon \mathbb{R} \to \mathbb{R} \ ; \ \begin{array}{c} \text{es gibt } a_0, \ldots, a_d \in \mathbb{R} \text{ mit} \\ f(x) = a_d x^d + \ldots + a_1 x + a_0 \text{ für alle } x \in \mathbb{R} \end{array} \right\}$$

den $\mathbb{R}$-Vektorraum der Polynomfunktionen vom Grad $\leq d$ betrachten. Dann gilt

$$\dim(\mathrm{Pol}_{\mathbb{R},\leq d}) = d + 1$$

mit der Basis

$$1 = x^0, x, x^2, \ldots, x^d.$$

Klarerweise ist $1, x, x^2, \ldots, x^d$ ein Erzeugendensystem, denn Polynomfunktionen sind ja gerade als $\mathbb{R}$-Linearkombinationen der Potenzfunktionen definiert. Zu zeigen bleibt die lineare Unabhängigkeit. Wenn

$$f(x) = a_d x^d + \ldots + a_1 x + a_0 = 0$$

für alle $x \in \mathbb{R}$ und nicht alle $a_i = 0$, dann haben wir ein nichttriviales Polynom (noch zu definieren!) mit Koeffizienten aus $\mathbb{R}$ und unendlich vielen Nullstellen. Das geht nicht (noch zu zeigen!).

(3) Hat ein Vektorraum V die Dimension 0, so besteht eine Basis aus 0-vielen Vektoren. Damit ist das leere Tupel eine Basis. Von diesem wird der Nullraum als lineare Hülle erzeugt, folglich ist $V = \{0\}$. Damit ist $\dim(V) = 0$ äquivalent zu $V = \{0\}$.

Satz 5.31. *Sei $n \in \mathbb{N}_0$. Seien V ein K-Vektorraum der Dimension $\dim(V) = n$ und $v_1, \ldots, v_m \in V$.*

(1) Wenn $v_1, \ldots, v_m$ linear unabhängig sind, dann ist

$$m \leq n.$$

Bei Gleichheit $m = n$ ist $v_1, \ldots, v_m$ eine Basis.

(2) Wenn $v_1, \ldots, v_m$ ein Erzeugendensystem ist, dann ist

$$m \geq n.$$

Bei Gleichheit $m = n$ ist $v_1, \ldots, v_m$ eine Basis.

Beweis. (1) Nach dem Basisergänzungssatz, Satz 5.23, kann man $v_1, \ldots, v_m$ zu einer Basis

$$v_1, \ldots, v_m, w_1, \ldots, w_s$$

ergänzen. Da nach dem Basissatz, Theorem 5.24, alle Basen $n = \dim(V)$-viele Vektoren enthalten, muß $m \leq m + s = n$ sein. Bei Gleichheit $m = n$ folgt $s = 0$ und man mußte kein w_j ergänzen, somit ist $v_1, \ldots, v_m$ schon eine Basis.

(2) Nach dem Basisauswahlsatz, Satz 5.22, enthält jedes Erzeugendensystem eine Basis. Da nach dem Basissatz, Theorem 5.24, alle Basen $n = \dim(V)$-viele Vektoren enthalten, enthält $v_1, \ldots, v_m$ eine n-elementige Teilmenge, die eine Basis ist. Insbesondere ist $m \geq n$. Und bei Gleichheit $m = n$ ist diese Teilmenge, die eine Basis ist, gleich $v_1, \ldots, v_m$. □

Bemerkung 5.32. Satz 5.31 besagt informell gesprochen für Vektoren $v_1, \ldots, v_n$ die Äquivalenz der Eigenschaften:

(a) Basis: linear unabhängig und Erzeugendensystem.
(b) linear unabhängig, und die richtige Anzahl (die Dimension).
(c) Erzeugendensystem, und die richtige Anzahl (die Dimension).

Die Aussagen (b) und (c) sind jeweils leichter zu prüfen als Aussage (a) und daher praktische Kriterien für *Basis*.

5.5. **Dimension von Unterräumen.** Der Dimensionsbegrff erlaubt Aussagen über Unterräume.

Korollar 5.33. *Seien V ein endlich erzeugter K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann gilt*

$$\dim(U) \leq \dim(V),$$

insbesondere ist U auch endlich erzeugt. Bei Gleichheit $\dim(U) = \dim(V)$ folgt $U = V$.

Beweis. Vektoren aus U, die in U linear unabhängig sind, sind das auch in V. Daher besteht nach Satz 5.31 jedes Tupel linear unabhängiger Vektoren aus U aus höchstens $\dim(V)$-vielen Vektoren. Damit gibt es endliche Tupel $(v_1, \ldots, v_m)$ von Vektoren aus U, die linear unabhängig und in U maximal mit dieser Eigenschaft sind. Daher handelt es sich um eine Basis nach Proposition 5.17. Somit ist U endlich erzeugt. Ferner ist

$$\dim(U) = m \leq \dim(V).$$

Nehmen wir nun an, $\dim(U) = \dim(V)$. Sei $v_1, \ldots, v_m$ eine Basis von U. Dann sind die $v_1, \ldots, v_m$ aufgefaßt als Vektoren von V immer noch linear unabhängig und $\dim(V)$-viele. Nach Satz 5.31 (1) folgt, daß $v_1, \ldots, v_m$ auch eine Basis von V ist. Damit gilt

$$U = \langle v_1, \ldots, v_m \rangle_K = V.$$ □

Korollar 5.34. *Sind $U \subseteq U'$ Unterräume des K-Vektorraums V und ist U' endlichdimensional, dann gilt*

$$\dim(U) \leq \dim(U')$$

und bei Gleichheit $\dim(U) = \dim(U')$ *folgt* $U = U'$.

Beweis. Der Unterraum U ist ein Unterraum von U'. Darauf wenden wir Korollar 5.33 an. □

Beispiel 5.35. Die einzigen Unterräume von $\mathbb{R}^2$ sind 0, $\mathbb{R}^2$ und für $x \in \mathbb{R}^2$, $x \neq 0$ die Geraden

$$\langle x \rangle_K = \{\lambda x \ ; \ \lambda \in \mathbb{R}\}.$$

Ein Unterraum $U \subseteq \mathbb{R}^2$ hat $\dim(U) \leq \dim(\mathbb{R}^2) = 2$, also

$$\dim(U) = 0, 1, \text{ oder } 2.$$

Wenn $\dim(U) = 0$, dann folgt nach Korollar 5.34 auf $0 \subseteq U$ angewandt

$$U = 0.$$

Wenn $\dim(U) = 2$, dann folgt nach Korollar 5.34 auf $U \subseteq \mathbb{R}^2$ angewandt

$$U = \mathbb{R}^2.$$

Bleibt der Fall $U = 1$. Hier besteht eine Basis von U aus genau einem Vektor $x \in \mathbb{R}^2$. Da x eine Basis von U ist, muß $x \neq 0$ sein. Damit ist $U = \langle x \rangle_K$ wie behauptet.

Beispiel 5.36. Daß man Unterräume, die ineinander enthalten sind, so leicht über die Dimension unterscheiden kann, ist eine Spezialität endlichdimensionaler Vektorräume. Für ein Gegenbeispiel bei unendlich-dimensionalen Vektorräumen betrachten wir den Raum

$$V = \mathrm{Abb}(\mathbb{N}, K)$$

aller Folgen $(a_n)_{n\in\mathbb{N}}$. Dieser enthält den Unterraum

$$U = \{(a_n)_{n\in\mathbb{N}} \ ; \ a_n \in K \text{ für alle } n \geq 1 \text{ und } a_1 = 0\} \subseteq \mathrm{Abb}(\mathbb{N}, K) = V.$$

Es gilt $U \neq V$, obwohl

$$\dim(U) = \infty = \dim(V).$$

Die Dimensionen zweier Unterräume sind mit den Dimensionen ihres Schnitts und ihrer Summe durch eine einfache Gleichung verbunden.

Satz 5.37 (Dimensionsformel für Unterräume)**.** *Seien V ein K-Vektorraum und $U_1, U_2 \subseteq V$ endlich erzeugte Unterräume. Dann sind auch $U_1 + U_2$ und $U_1 \cap U_2$ endlichdimensional und es gilt*

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

Beweis. Als Unterraum $U_1 \cap U_2 \subseteq U_1$ ist $\dim(U_1 \cap U_2) \leq \dim(U_1)$ nach Korollar 5.33, also $U_1 \cap U_2$ auch endlich erzeugt.

Wir beginnen nun mit einer Basis $b_1, \ldots, b_s$ von $U_1 \cap U_2$. Nach dem Basisergänzungssatz, Satz 5.23, können wir dies zu einer Basis von U_1

$$b_1, \ldots, b_s, x_1, \ldots, x_n$$

und einer Basis von U_2

$$b_1, \ldots, b_s, y_1, \ldots, y_m$$

ergänzen. Per Definition ist

$$\begin{aligned} U_1 + U_2 &= \langle b_1, \ldots, b_s, x_1, \ldots, x_n \rangle_K + \langle b_1, \ldots, b_s, y_1, \ldots, y_m \rangle_K \\ &= \langle b_1, \ldots, b_s, x_1, \ldots, x_n, y_1, \ldots, y_m \rangle_K. \end{aligned}$$

Insbesondere ist $U_1 + U_2$ auch endlich erzeugt und damit von endlicher Dimension.

Wir zeigen nun, daß die Vektoren $b_1, \ldots, b_s, x_1, \ldots, x_n, y_1, \ldots, y_m$ linear unabhängig sind. Seien $\lambda_1, \ldots, \lambda_s, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \in K$ mit

$$\sum_{i=1}^{s} \lambda_i b_i + \sum_{i=1}^{n} \alpha_i x_i + \sum_{i=1}^{m} \beta_i y_i = 0.$$

Dann ist
$$v := \underbrace{\sum_{i=1}^{s} \lambda_i b_i + \sum_{i=1}^{n} \alpha_i x_i}_{\in U_1} = \underbrace{-\sum_{i=1}^{m} \beta_i y_i}_{\in U_2} \in U_1 \cap U_2.$$
Also ist v eine Linearkombination der $b_1, \dots, b_s$. Aufgrund der Eindeutigkeit der Darstellung bezüglich einer Basis, Korollar 5.14, folgt $\alpha_i = 0$ für alle $i = 1, \dots, n$ (Eindeutigkeit bezüglich Basis von U_1) und $\beta_i = 0$ für alle $i = 1, \dots, m$ (Eindeutigkeit bezüglich Basis von U_2). Somit bleibt
$$v = \sum_{i=1}^{s} \lambda_i b_i = 0,$$
und dann folgt $\lambda_i = 0$ für alle $i = 1, \dots, s$, denn $b_1, \dots, b_s$ ist linear unabhängig. Dies beendet den Beweis, daß
$$b_1, \dots, b_s, x_1, \dots, x_n, y_1, \dots, y_m$$
linear unabhängig und damit eine Basis von $U_1 + U_2$ ist.

Die Dimensionsformel folgt nun aus
$$\begin{aligned} \dim(U_1 + U_2) + \dim(U_1 \cap U_2) &= (s + n + m) + s \\ &= (s + n) + (s + m) = \dim(U_1) + \dim(U_2). \end{aligned}$$ □

Korollar 5.38. *Sei V ein endlichdimensionaler K-Vektorraum. Seien U_1, U_2 Unterräume von V mit*
$$\dim(U_1) + \dim(U_2) > \dim(V).$$
Dann gibt es $0 \neq x \in V$ mit $x \in U_1 \cap U_2$.

Beweis. Aus der Dimensionsformel und weil $U_1 + U_2$ ein Unterraum von V ist, folgt nach Voraussetzung
$$\dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 + U_2) > \dim(V) - \dim(U_1 + U_2) \geq 0.$$
Damit ist $\dim(U_1 \cap U_2) > 0$ und $U_1 \cap U_2$ nicht der Unterraum 0. □

5.6. **Direkte Summe und Komplemente.**

Definition 5.39. Sei V ein K-Vektorraum. Die Summe $U_1 + U_2$ zweier Unterräume U_1 und U_2 von V ist **direkt**, genauer eine **innere direkte Summe** mit der Notation
$$U_1 \oplus U_2 \subseteq V,$$
wenn für jeden Vektor $x \in U_1 + U_2$ die Darstellung als $x = y_1 + y_2$ mit $y_i \in U_i$ für $i = 1, 2$ eindeutig ist.

Proposition 5.40. *Sei V ein K-Vektorraum. Die Summe zweier Unterräume U_1 und U_2 von V ist **direkt** genau dann, wenn $U_1 \cap U_2 = 0$.*

Beweis. Angenommen $U_1 \cap U_2 = 0$. Wenn $y_1 + y_2 = z_1 + z_2$ mit $y_i, z_i \in U_i$ für $i = 1, 2$, dann gilt
$$y_1 - z_1 = z_2 - y_2 \in U_1 \cap U_2 = \{0\}.$$
Daraus folgt mit $y_1 = z_1$ und $y_2 = z_2$ die behauptete Eindeutigkeit.

Sei umgekehrt die Summe $U_1 \oplus U_2$ direkt. Jedes $x \in U_1 \cap U_2$ hat die Darstellung
$$x = y_1 + y_2 = z_1 + z_2 \in U_1 + U_2$$
mit $y_1 = x$, $z_1 = 0 \in U_1$ und $y_2 = 0$, $z_2 = x \in U_2$. Weil die Darstellung eindeutig ist, folgt $x = 0$, somit $U_1 \cap U_2 = 0$. □

Korollar 5.41. *Sind U_1 und U_2 Unterräume von V, deren Summe in V direkt ist, dann gilt*

$$\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2).$$

Für jede Basis $(v_1, \dots, v_s)$ von U_1 und $(w_1, \dots, w_t)$ von U_2 ist eine Basis von $U_1 \oplus U_2$ gegeben durch

$$(v_1, \dots, v_s, w_1, \dots, w_t).$$

Beweis. Die Dimensionsformel, Satz 5.37, zeigt wegen $U_1 \cap U_2 = 0$

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

Die Aussage über die Basis von $U_1 \oplus U_2$ folgt aus dem Beweis von Satz 5.37. □

Definition 5.42. Sei V ein K-Vektorraum. Es ist V die **direkte Summe** (genauer die **innere direkte Summe**) zweier Unterräume U_1 und U_2 von V, wenn die Summe von U_1 und U_2 direkt ist und $U_1 + U_2 = V$ gilt. Wir schreiben dann

$$V = U_1 \oplus U_2.$$

Beispiel 5.43. Wir betrachten im K^n mit der Standardbasis $e_1, \dots, e_n$ die Unterräume

$$U_i = \langle e_1, \dots, e_i \rangle_K,$$
$$W_i = \langle e_{i+1}, \dots, e_n \rangle_K.$$

für $i = 1, \dots, n$. Dann gilt für alle $i = 1, \dots, n$

$$K^n = U_i \oplus W_i.$$

Definition 5.44. Seien V ein K-Vektorraum und U ein Unterraum. Ein **Komplement** von U in V ist ein Unterraum $W \subseteq V$ mit

$$V = U \oplus W.$$

Satz 5.45. *Jeder Unterraum U in einem endlich erzeugten K-Vektorraum V hat ein Komplement. Jedes Komplement W von U hat die Dimension* $\dim(W) = \dim(V) - \dim(U)$.

Beweis. Als Unterraum von V ist U auch endlich erzeugt. Wir wählen eine Basis $b_1, \dots, b_s$ von U und ergänzen diese nach dem Basisergänzungssatz, Satz 5.23, zu einer Basis $b_1, \dots, b_s, c_1, \dots, c_t$ von V. Wir setzen nun

$$W = \langle c_1, \dots, c_t \rangle_K.$$

Dann gilt $V = U + W$ und nach der Dimensionsformel, Satz 5.37, auch

$$\dim(U \cap W) = \dim(U + W) - \dim(U) - \dim(W) = (s + t) - s - t = 0.$$

Also ist $U \cap W = 0$, und die Summe $V = U \oplus W$ ist direkt. Die Dimensionsformel, Satz 5.37, zeigt auch die Behauptung zur Dimension des Komplements. □

Bemerkung 5.46. Im Allgemeinen hat ein Unterraum viele verschiedene Komplemente. Zum Beispiel hat jeder eindimensionale Unterraum $L \subseteq \mathbb{R}^2$, eine Ursprungsgerade, jeden von L verschiedenen Unterraum $L' \subseteq \mathbb{R}^2$ der Dimension 1 als Komplement $L \oplus L' = \mathbb{R}^2$.

Definition 5.47. Seien V_1 und V_2 zwei K-Vektorräume. Die **direkte Summe** von V_1 und V_2 ist der folgende K-Vektorraum $V_1 \oplus V_2$. Als Menge ist

$$V_1 \oplus V_2 = V_1 \times V_2$$

das Produkt von Mengen. Ein $v \in V_1 \oplus V_2$ ist also ein Tupel $v = (x_1, x_2)$ mit $x_i \in V_i$ für $i = 1, 2$. Die Addition von Tupeln ist komponentenweise definiert durch

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2)$$

für alle $x_i, y_i \in V_i$, $i = 1, 2$. Die Skalarmultiplikation ist ebenfalls komponentenweise definiert: für $\lambda \in K$ durch

$$\lambda(x_1, x_2) := (\lambda x_1, \lambda x_2).$$

Die Vektorraumaxiome verifiziert man für $V_1 \oplus V_2$ ohne Probleme. Letztendlich erbt die direkte Summe die Struktur als Vektorraum von den **Summanden** V_1 und V_2.

Bemerkung 5.48. In der direkten Summe $V = V_1 \oplus V_2$ gibt es zwei Unterräume U_1, U_2 gegeben durch

$$U_1 = \{(x_1, x_2) \in V \ ; \ x_2 = 0\}, \qquad U_2 = \{(x_1, x_2) \in V \ ; \ x_1 = 0\}.$$

Offensichtlich ist $V = U_1 + U_2$ und $U_1 \cap U_2 = \{0\}$. Daher ist $V = U_1 \oplus U_2$ als (innere) direkte Summe.

Wenn man zusätzlich bedenkt, daß durch

$$V_1 \xrightarrow{\sim} U_1, \qquad x_1 \mapsto (x_1, 0), \qquad \text{und} \qquad V_2 \xrightarrow{\sim} U_2, \qquad x_2 \mapsto (0, x_2),$$

der Unterraum U_i mit V_i als Vektorraum identifiziert werden kann, so wird die parallele Bezeichnung und Notation für die (innere) direkte Summe verständlich und vertretbar.

Übungsaufgaben zu §5

Übungsaufgabe 5.1. Sei V ein K-Vektorraum. Seien $v, w \in V$ beide von 0 verschieden. Zeigen Sie die Äquivalenz der folgenden Aussagen.

(a) v, w sind linear abhängig.
(b) Es gibt $\lambda \in K$ mit $v = \lambda w$.
(c) Es gibt $\lambda \in K$ mit $w = \lambda v$.

Übungsaufgabe 5.2. Sei K ein Körper. Bestimmen Sie alle Basen des K-Vektorraums K^1.

Übungsaufgabe 5.3. Sei V ein K-Vektorraum. Zeigen Sie:

Wenn $(v_1, \ldots, v_n)$ linear unabhängige Vektoren aus V sind, dann ist auch für jede Permuation $\sigma \in S_n$ das mit σ permutierte Tupel

$$(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$$

linear unabhängig. Genauer gilt dann und nur dann.

Tipp: Nutzen Sie die inverse Permutation.

Bemerkung: Es kommt also bei „linear unabhängig“ nicht auf die Reihenfolge an. Dies ist eine Eigenschaft der zugehörigen Menge von Vektoren, sofern es im Tupel keine Wiederholungen gibt.

Übungsaufgabe 5.4. Sei V ein K-Vektorraum. Zeigen Sie: Wenn von $v_1, \ldots, v_n \in V$ zwei Vektoren gleich sind oder einer der Vektoren 0 ist, dann ist das Tupel $(v_1, \ldots, v_n)$ linear abhängig.

Übungsaufgabe 5.5 (Austauschlemma). Sei V ein K-Vektorraum mit einer Basis $v_1, \ldots, v_n$ und sei $0 \neq w \in V$ beliebig.

Zeigen Sie: Dann gibt es ein $1 \leq i_0 \leq n$, so daß der Austausch von v_{i_0} mit w, also das Tupel $w_1, \ldots, w_n$ mit

$$w_i = \begin{cases} v_i & i \neq i_0 \\ w & i = i_0 \end{cases}$$

wieder eine Basis von V ist.

Vorsicht: Das ist nicht die Aussage des Austauschsatzes!

Übungsaufgabe 5.6. Beweisen Sie Satz 5.26 alternativ mittels des Austauschlemmmas aus Aufgabe 5.5. Wie müssen Sie die angesetzte Induktion verändern?

Übungsaufgabe 5.7. Zeigen Sie, daß der Raum der Folgen

$$\mathrm{Abb}(\mathbb{N}, K)$$

die Dimension ∞ hat, d.h. kein endliches Erzeugendensystem besitzt.

Übungsaufgabe 5.8. Wir betrachten Beispiel 5.6 (3) erneut. Die Vektoren

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad y = \begin{pmatrix} 4 \\ 6 \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} 7 \\ 0 \\ 0 \end{pmatrix}$$

kann man auch als Vektoren in $\mathbb{R}^3$ auffassen. In $\mathbb{Q}^3$ waren sie linear unabhängig. Sind sie das auch in $\mathbb{R}^3$?

Was passiert, wenn man $0, 1, 2, 3, 4, 5, 6, 7, \ldots$ in $\mathbb{F}_2$ auffaßt (n als Ergebnis von $1 + 1 + \ldots + 1$ mit n Summanden)? Sind dann die Vektoren x, y, z linear unabhängig?

Übungsaufgabe 5.9. Überlegen Sie sich, wie der Beweis des Basissatzes, Theorem 5.24, funktioniert, wenn die Basis $\mathscr{B}$ aus dem leeren Tupel besteht. Dann gibt es ja keine Vektoren zum austauschen!

Übungsaufgabe 5.10. Sei $n \in \mathbb{N}$ und sei $V = \mathscr{P}(\{1, \ldots, n\})$ der $\mathbb{F}_2$-Vektorraum aus Beispiel 4.11 auf der Potenzmenge von $\{1, \ldots, n\}$ mit symmetrischer Differenz als Addition.

(a) Wir setzen für $i = 1, \ldots, n$

$$v_i = \{1, \ldots, i\} \in V.$$

Zeigen Sie, daß $(v_1, \ldots, v_n)$ eine Basis von V ist.

(b) Geben Sie eine Basis $(w_1, \ldots, w_n)$ von V an, für die

$$|w_1| + \ldots + |w_n|$$

den minimal möglichen Wert annimmt.

Übungsaufgabe 5.11. Sei V ein endlichdimensionaler K-Vektorraum und sei

$$U_1 \subseteq U_2 \subseteq U_3 \subseteq \ldots$$

eine aufsteigende Folge von Unterräumen von V. Zeigen Sie, daß es ein $n \in \mathbb{N}_0$ gibt, so daß

$$U_i = U_n$$

für alle $i \geq n$.

Übungsaufgabe 5.12. Sei V ein endlichdimensionaler K-Vektorraum. Für eine echt aufsteigende Folge

$$U_0 \subsetneq U_1 \subsetneq U_2 \subsetneq \ldots \subsetneq U_n$$

von Unterräumen von V nennen wir n die Länge. „Echt“ bezieht sich darauf, daß für alle $0 \leq i \leq n-1$ gilt $U_i \neq U_{i+1}$.

Zeigen Sie die folgende alternative Beschreibung der Dimension:

$$\dim(V) = \sup\{n \ ; \text{ es gibt } U_0 \subsetneq U_1 \subsetneq U_2 \subseteq \ldots \subsetneq U_n \text{ Unterräume in } V\}.$$

Übungsaufgabe 5.13. Zeigen Sie, daß ein K-Vektorraum V genau dann endliche Dimension hat, wenn V endlich erzeugt ist.

Übungsaufgabe 5.14. Sei V ein endlich erzeugter K-Vektorraum. Zeigen Sie, daß jeder Unterraum von V auch endlich erzeugt ist.

6. Die komplexen Zahlen

Der Körper der komplexen Zahlen $\mathbb{C}$ spielt eine wichtige Rolle in der Mathematik. Historisch gesehen hat es einige Überwindung gekostet, diesen Körper zu akzeptieren. Heute ist es unvorstellbar, wie man ohne $\mathbb{C}$ auskommen soll.

Definition 6.1. Der Körper der **komplexen Zahlen** besteht aus der Menge $\mathbb{C} := \mathbb{R} \times \mathbb{R}$, wobei wir die Notation

$$a + b \cdot i := (a, b) \in \mathbb{C}$$

benutzen, mit Addition $+\colon \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ und Multiplikation $\cdot\colon \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ definert durch: für alle $a + b \cdot i, c + d \cdot i \in \mathbb{C}$ gilt

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

Das Symbol i wird **imaginäre Einheit** genannt und identifiziert mit

$$i := 0 + 1 \cdot i = (0, 1) \in \mathbb{C}.$$

Ein Element $z \in \mathbb{C}$ hat also eine eindeutige Darstellung $z = a + b \cdot i$ mit **Realteil** $a = \Re(z) \in \mathbb{R}$ und **Imaginärteil** $b = \Im(z) \in \mathbb{R}$.

Bemerkung 6.2. Auf dem Weg zum Nachweis, daß es sich bei $\mathbb{C}$ in der Tat um einen Körper handelt, beginnen wir mit ein paar Beobachtungen.

(1) Wir können die Sprache der $\mathbb{R}$-Vektorräume benutzen aufgrund der Bijektion

$$\mathbb{C} \xrightarrow{\sim} \mathbb{R}^2, \qquad (a, b) \mapsto \begin{pmatrix} a \\ b \end{pmatrix}.$$

In diesem Sinn hat $\mathbb{C}$ eine Basis als $\mathbb{R}$-Vektorraum gegeben durch

$$1 = 1 + 0 \cdot i, \qquad i = 0 + 1 \cdot i.$$

Die Koordinaten bezüglich dieser Basis sind gerade Real- und Imaginärteil. Die Addition von $\mathbb{C}$ stimmt mittels der Bijektion zu $\mathbb{R}^2$ mit der additiven Gruppe $(\mathbb{R}^2, +)$ des Vektorraums $\mathbb{R}^2$ überein. Somit ist $(\mathbb{C}, +)$ eine kommutative Gruppe. Das ist ein Anfang.

(2) Die geometrische Interpretation der komplexen Zahlen als die Punkte der reellen euklidischen Ebene verwendet die reellen Koordinaten $(\Re(z), \Im(z))$ in cartesischen Koordinaten für die komplexe Zahl/den Punkt $z \in \mathbb{C} = \mathbb{R}^2$.

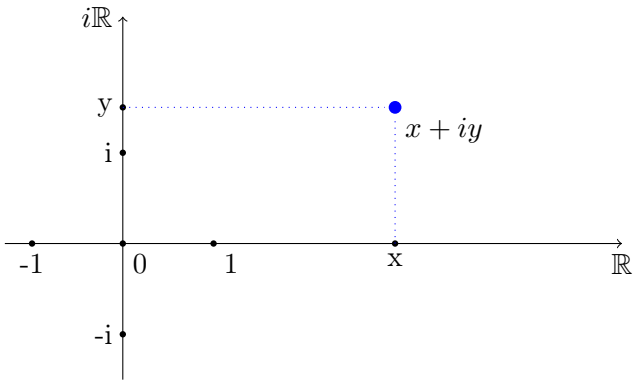


Abbildung 3. Die komplexen Zahlen $\mathbb{C}$ als Ebene $\mathbb{R}^2$.

(3) Die Notation einer komplexen Zahl $z = a + b \cdot i$ erklärt die Multiplikation. Wenn man mit i wie mit einer formalen Variablen rechnet, dann kommt fast dasselbe raus:

$$(a + b \cdot i) \cdot (c + d \cdot i) = ac + (ad + bc) \cdot i + bd \cdot i^2.$$

Es reicht, wenn wir zusätzlich benutzen, daß

$$i^2 = -1,$$

und wir sonst mit der Multiplikation aus $\mathbb{R}$ und dem Distributivgesetz arbeiten. In der Tat gilt in $\mathbb{C}$

$$i^2 = (0,1) \cdot (0,1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0)$$

und das wird die -1 des Körpers $\mathbb{C}$ sein.

(4) Wir können $\mathbb{R}$ vermöge der injektiven Abbildung $\mathbb{R} \hookrightarrow \mathbb{C}$

$$a \mapsto a + 0 \cdot i = (a, 0)$$

als Teilmenge von $\mathbb{C}$ auffassen. Diese Inklusion ist offensichtlich verträglich mit Addition und Multiplikation. Man sagt, daß $\mathbb{R}$ ein **Unterkörper** von $\mathbb{C}$ ist.

Wir werden im folgenden stets auf diese Weise $\mathbb{R} \subseteq \mathbb{C}$ als Teilmenge auffassen.

(5) Für alle $a \in \mathbb{R}$ und $z = x + y \cdot i \in \mathbb{C}$ gilt mit Addition und Multiplikation in $\mathbb{C}$

$$a \cdot z = (a + 0 \cdot i) \cdot (x + y \cdot i) = ax + ay \cdot i.$$

Dies ist die Skalarmultiplikation von $\mathbb{C}$ aufgefaßt als $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ wie oben. Das ist gut so, denn es bedeutet, daß wir beim Multiplizieren von reellen mit komplexen Zahlen nicht unterscheiden müssen, ob wir diese beide als komplexe Zahlen auffassen. oder ob es sich um die Skalarmultiplikation im $\mathbb{R}$-Vektorraum $\mathbb{C} \xrightarrow{\sim} \mathbb{R}^2$ handelt.

Definition 6.3. Die zu einer komplexen Zahl $z = a + b \cdot i \in \mathbb{C}$ **komplex konjugierte** Zahl ist

$$\bar{z} = a - b \cdot i.$$

Die **komplexe Konjugation** ist die Abbildung $z \mapsto \bar{z}$.

Lemma 6.4. *Für eine komplexe Zahl $z = a + b \cdot i$ gilt*

$$\bar{z} \cdot z = a^2 + b^2,$$

und dies liegt in der Teilmenge $\mathbb{R} \subseteq \mathbb{C}$.

Beweis. Das ist eine einfache Rechnung:

$$\bar{z} \cdot z = (a - b \cdot i)(a + b \cdot i) = (a^2 + b^2) + (ab + (-b)a) \cdot i = a^2 + b^2. \qquad \square$$

Satz 6.5. *Die komplexen Zahlen $(\mathbb{C}, +, \cdot)$ sind ein Körper.*

Beweis. Wir wissen bereits, $(\mathbb{C}, +)$ ist eine kommutative Gruppe. Daher wenden wir uns den Eigenschaften der Multiplikation zu. Die Multiplikation von $\mathbb{C}$ ist aus der Definition heraus offensichtlich kommutativ. Die Eins der Multiplikation in $\mathbb{C}$ ist

$$1 = 1 + 0 \cdot i,$$

denn für alle $a + b \cdot i \in \mathbb{C}$ gilt (Multiplikation mit $\mathbb{R}$ ist Skalarmultiplikation)

$$1(a + bi) = a + b \cdot i$$

und die andere Relation folgt aus kommutativ.

Assoziativ: Für beliebige Elemente $a+b\cdot i, x+y\cdot i, u+v\cdot i \in \mathbb{C}$ gilt

$$\begin{aligned}\big((a+b\cdot i)(x+y\cdot i)\big)(u+v\cdot i) &= \big((ax-by)+(ay+bx)\cdot i\big)(u+v\cdot i)\\ &= \big((ax-by)u-(ay+bx)v\big)+\big((ax-by)v+(ay+bx)u\big)\cdot i\\ &= \big(a(xu-yv)-b(yu+xv)\big)+\big(a(xv+yu)+b(xu-yv)\big)\cdot i\\ &= (a+b\cdot i)\big((xu-yv)+(yu+xv)\cdot i\big)\\ &= (a+b\cdot i)\big((x+y\cdot i)(u+v\cdot i)\big)\end{aligned}$$

Existenz des Inversen: Hier muß man wissen, was man tut. Zu jedem $0\neq z=a+b\cdot i\in\mathbb{C}$ brauchen wir ein multiplikatives Inverses. Als reelle Zahlen sind in jedem Fall $a^2\geq 0$ und $b^2\geq 0$, also $a^2+b^2\geq 0$. Nun sind nicht $a=b=0$, somit mindestens ein Quadrat echt >0 und deshalb

$$R := a^2+b^2 \neq 0.$$

Das Inverse bekommen wir nun als (Multiplikation mit $\mathbb{R}$ ist Skalarmultiplikation)

$$z^{-1} = R^{-1}\bar{z} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\cdot i,$$

denn nach Lemma 6.4 gilt

$$(R^{-1}\bar{z})z = R^{-1}(\bar{z}z) = R^{-1}R = 1,$$

wobei wir davon Gebrauch machen, auf dem Teil $\mathbb{R}\subset\mathbb{C}$ wie in $\mathbb{R}$ rechnen zu können.

Distributiv: Für beliebige Elemente $a+b\cdot i, x+y\cdot i, u+v\cdot i\in\mathbb{C}$ gilt:

$$\begin{aligned}(a+b\cdot i)\big((x+y\cdot i)+(u+v\cdot i)\big) &= (a+b\cdot i)\big((x+u)+(y+v)\cdot i\big)\\ &= a(x+u)-b(y+v)+\big(a(y+v)+b(x+u)\big)\cdot i\\ &= (ax-by+au-bv)+(ay+bx+av+bu)\cdot i\\ &= \big((ax-by)+(ay+bx)\cdot i\big)+\big((au-bv)+(av+bu)\cdot i\big)\\ &= (a+b\cdot i)(x+y\cdot i)+(a+b\cdot i)(u+v\cdot i).\end{aligned}$$

□

Beispiel 6.6. Anhand des Körpers der komplexen Zahlen können wir einige Begriffe wiederholen. Jede komplexe Zahl $z=a+b\cdot i\in\mathbb{C}$ (mit $a,b\in\mathbb{R}$) ist eine $\mathbb{R}$-Linearkombination von $1,i\in\mathbb{C}$:

$$z = a+b\cdot i = a\cdot 1 + b\cdot i.$$

Weil $i\in\mathbb{C}\notin\mathbb{R}$ gilt, sind 1 und i linear unabhängig, alternativ: aus $a+bi=0$ mit $a,b\in\mathbb{R}$ folgt $a=b=0$. Daher hat $\mathbb{C}$ die Dimension

$$\dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

Aber $\mathbb{C}$ ist als $\mathbb{C}^1$ auch ein $\mathbb{C}$-Vektorraum und $1,i$ sind $\mathbb{C}$-linear abhängig:

$$i\cdot 1 + (-1)\cdot i = 0.$$

Die Dimension als $\mathbb{C}$-Vektorraum ist natürlich

$$\dim_{\mathbb{C}}(\mathbb{C}) = 1,$$

weil hier $\mathbb{C}$ als $\mathbb{C}$-Vektorraum nichts anderes als $\mathbb{C}^1$ ist. Es ist also unbedingt wichtig, daß man für einen Vektorraum den Körper der Skalare festlegt.

Bemerkung 6.7. Eine komplexe Zahl hat auch eine Darstellung mittels Polarkoordinaten.

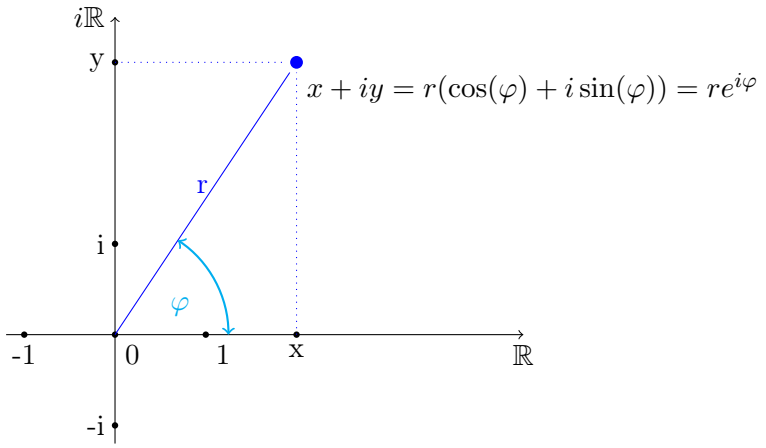


Abbildung 4. Polarkoordinaten für komplexen Zahlen.

Der Radius r berechnet sich aus $z = x + iy$ als

$$r = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} =: |z|.$$

In der Analysis lernt man mittels Potenzreihen, daß Exponentialfunktion und trigonometrische Funktionen Sinus und Cosinus auch für komplexe Argumente sinnvoll definiert sind. Aus den Potenzreihendarstellungen folgt insbesondere auch für $\varphi \in \mathbb{R}$ die Eulersche Formel

$$\cos(\varphi) + i\sin(\varphi) = e^{i\varphi},$$

und damit $z = re^{i\varphi}$. Jetzt wird die Geometrie hinter der Multiplikation komplexer Zahlen transparent. Die Abbildung $f_w(z) = wz$, also die Multiplikation mit $w = \rho e^{i\psi}$ führt zu

$$f_w(z) = wz = \rho e^{i\psi} \cdot re^{i\varphi} = (\rho r)e^{i(\varphi+\psi)}.$$

Das erklärt f_w als eine Drehstreckung mit Skalierungsfaktor $\rho = |w|$ und Drehwinkel ψ. Wenn man dies wieder in reellen Koordinaten ausdrückt, dann erhält man die Additionstheoreme für Sinus und Cosinus (der Einfachheit halber $r = \rho = 1$):

$$\begin{aligned}\cos(\varphi+\psi) + i\sin(\varphi+\psi) &= e^{i(\varphi+\psi)} \\ &= e^{i\varphi} \cdot e^{i\psi} = (\cos(\varphi) + i\sin(\varphi)) \cdot (\cos(\psi) + i\sin(\psi)) \\ &= (\cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi)) + i(\sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi))\end{aligned}$$

und Koeffizientenvergleich von Real- und Imaginärteil.

Bemerkung 6.8. Die komplexen Zahlen erlauben das Wurzelziehen auch aus negativen reellen Zahlen, was wegen

$$x \in \mathbb{R} \Longrightarrow x^2 \geq 0$$

mit reellen Zahlen allein nicht möglich ist. In der Tat, zu einer reellen Zahl $a < 0$ ist $z = \sqrt{|a|} \cdot i$ eine Wurzel, denn

$$z^2 = (\sqrt{|a|} \cdot i)^2 = (\sqrt{|a|})^2 \cdot i^2 = |a| \cdot (-1) = a.$$

Übungsaufgaben zu §6

Übungsaufgabe 6.1. Zeigen Sie:

(a) Für jedes $z \in \mathbb{C}$ gibt es ein $w \in \mathbb{C}$ mit $z = w^2$.

(b) Jede quadratische Gleichung

$$Z^2 + aZ + b = 0$$

mit Koeffizienten $a, b \in \mathbb{C}$ hat in $\mathbb{C}$ eine Lösung.

Teil 2. Morphismen

7. Lineare Abbildungen

Vektorräume sind der Hauptgegenstand der linearen Algebra. Nun wollen wir sehen, wie wir mehrere Vektorräume in Beziehung zueinander bringen können. Dazu betrachten wir Abbildungen, welche die vorhandene Struktur erhalten: Addition und Skalarmultiplikation.

Definition 7.1. Sei K ein Körper. Eine **lineare Abbildung** von K-Vektorräumen ist eine Abbildung

$$f : V \to W$$

von einem K-Vektorraum V nach einem K-Vektorraum W mit den folgenden Eigenschaften.

(i) f ist **additiv**: für alle $x, y \in V$ gilt

$$f(x + y) = f(x) + f(y).$$

(ii) f ist **homogen:** für alle $\lambda \in K$ und $x \in V$ gilt

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

Eine lineare Abbildung wird auch **Homomorphismus** genannt.

Bemerkung 7.2. Die Vektorräume, die in einer linearen Abbildung auftreten, haben denselben Körper K für ihre Skalarmultiplikation. Um den Körper zu betonen, bezüglich dessen die lineare Abbildung homogen ist, sagt man auch **K-lineare Abbildung**.

Lemma 7.3. *Sei $f : V \to W$ eine Abbildung von K-Vektorräumen. Dann sind äquivalent:*

(a) *f ist linear.*

(b) *Für alle $x, y \in V$ und alle $\lambda \in K$ gilt:*

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

Beweis. Wenn (a) gilt, also f linear ist, dann rechnen wir

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

Sei umgekehrt (b) erfüllt. Spezialisieren wir zunächst $\lambda = 1$, so folgt, f ist additiv:

$$f(x + y) = f(1 \cdot x + y) = 1 \cdot f(x) + f(y) = f(x) + f(y).$$

Als nächstes spezialisieren wir weiter zu $x = y = 0$ und erhalten aus additiv:

$$f(0) = f(0 + 0) = f(0) + f(0).$$

Addieren wir $-f(0)$ auf beiden Seiten, so folgt $f(0) = 0$. Mit dieser Information gehen wir zurück zu (b) und setzen $y = 0$:

$$f(\lambda x) = f(\lambda x + 0) = \lambda f(x) + f(0) = \lambda f(x).$$

Dies zeigt, daß f auch homogen ist. Dies zeigt (a). □

Proposition 7.4. *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Für alle $x, y \in V$ gilt:*

(1) $f(0) = 0$,

(2) $f(-x) = -f(x)$,

(3) $f(x - y) = f(x) - f(y)$.

Beweis. (1) haben wir gerade nebenbei bewiesen.

Die dritte Behauptung folgt durch

$$f(x - y) = f((-1)y + x) = (-1)f(y) + f(x) = f(x) - f(y)$$

und die zweite folgt durch Einsetzen von $x = 0$ wegen $f(0) = 0$. □

Bemerkung 7.5. Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Dann gilt für alle $n \in \mathbb{N}_0$ und $v_1, \ldots, v_n \in V$ und $a_1, \ldots, a_n \in K$

$$f(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i f(v_i).$$

Das folgt sofort per Induktion nach n mittels der Gleichung aus Lemma 7.3 (b). Der Induktionsschritt is die folgende Rechnung.

$$\begin{aligned} f(\sum_{i=1}^{n+1} a_i v_i) &= f\Big(a_{n+1}v_{n+1} + \sum_{i=1}^{n} a_i v_i\Big) = a_{n+1} f(v_{n+1}) + f(\sum_{i=1}^{n} a_i v_i) \\ &= a_{n+1} f(v_{n+1}) + \sum_{i=1}^{n} a_i f(v_i) = \sum_{i=1}^{n+1} a_i f(v_i). \end{aligned}$$

Proposition 7.6. *Die Komposition linearer Abbildungen ist wieder linear: sind $f : V \to W$ und $g : U \to V$ lineare Abbildungen von K-Vektorräumen, dann ist $h = f \circ g : U \to W$ linear.*

Beweis. Seien $x, y \in U$ und $\lambda \in K$. Dann gilt

$$h(\lambda x + y) = f(g(\lambda x + y)) = f(\lambda g(x) + g(y)) = \lambda f(g(x)) + f(g(y)) = \lambda h(x) + h(y).$$ □

7.1. **Beispiele und Eigenschaften linearer Abbildungen.** Das erste Beispiel beschreibt lineare Abbildungen mit Definitionsbereich K^n.

Beispiel 7.7. Sei V ein K-Vektorraum und seien $v_1, \ldots, v_n \in V$ beliebige Vektoren. Dann definiert

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \varphi(x) = x_1 v_1 + \ldots + x_n v_n$$

eine lineare Abbildung $f : K^n \to V$ mit $\varphi(e_i) = v_i$ für alle $i = 1, \ldots, n$. Wir benutzen das Kriterium aus Lemma 7.3. In der Tat gilt für alle $\lambda \in K$ und $x, y \in K^n$ mit Koordinaten x_i, y_i:

$$\begin{aligned} \varphi(\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}) &= \varphi(\begin{pmatrix} \lambda x_1 + y_1 \\ \vdots \\ \lambda x_n + y_n \end{pmatrix}) = \sum_{i=1}^{n} (\lambda x_i + y_i) v_i \\ &= \lambda \sum_{i=1}^{n} x_i v_i + \sum_{i=1}^{n} y_i v_i = \lambda \varphi(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) + \varphi(\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}). \end{aligned}$$

Das Bild von φ ist gerade die lineare Hülle $\langle v_1, \ldots, v_n \rangle_K$. Später folgt aus Satz 7.12, daß alle linearen Abbildungen $K^n \to V$ eindeutig durch das Tupel $(\varphi(e_1), \ldots, \varphi(e_n))$ von Vektoren aus V beschrieben werden.

Sei $\mathscr{B} = (v_1, \ldots, v_n)$ eine Basis des K-Vektorraums V. Dann gibt es nach Korollar 5.14 für jedes $x \in V$ eindeutig eine Linearkombination

$$x = x_1 v_1 + \ldots + x_n v_n.$$

Wir definieren mit diesen x_i

$$\kappa_{\mathscr{B}}(x) := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$$

und nennen die Einträge x_i des Vektors $\kappa_{\mathscr{B}}(x)$ die (i-te) **Koordinate** und $\kappa_{\mathscr{B}}(x)$ den **Koordinatenvektor** des Vektors x bezüglich der Basis $\mathscr{B}$.

Proposition 7.8. *Sei $\mathscr{B} = (v_1, \ldots, v_n)$ eine Basis des K-Vektorraums V. Dann ist die Koordinatenvektorabbildung*

$$\kappa_{\mathscr{B}} : V \to K^n$$

eine bijektive, lineare Abbildung.

Beweis. Wir benutzen das Kriterium aus Lemma 7.3. Wenn

$$x = x_1 v_1 + \ldots + x_n v_n,$$
$$y = y_1 v_1 + \ldots + y_n v_n,$$

und $\lambda \in K$ beliebig, dann ist

$$\lambda x + y = (\lambda x_1 + y_1)v_1 + \ldots + (\lambda x_n + y_n)v_n$$

und damit

$$\kappa_{\mathscr{B}}(\lambda x + y) = \begin{pmatrix} \lambda x_1 + y_1 \\ \vdots \\ \lambda x_n + y_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda \kappa_{\mathscr{B}}(x) + \kappa_{\mathscr{B}}(y).$$

Die Abbildung $\kappa_{\mathscr{B}}$ ist bijektiv, denn die inverse Abbildung $K^n \to V$ ist gegeben durch die Abbildung aus Beispiel 7.7

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \mapsto \quad x_1 v_1 + \ldots + x_n v_n. \qquad \square$$

Bemerkung 7.9. Wir möchten betonen, wie die Eigenschaften „linear unabhängig“ und „Erzeugendensystem“ einer Basis in Proposition 7.8 essentiell dafür sorgen, daß die Abbildung $\kappa_{\mathscr{B}} : V \to K^n$ wohldefiniert ist. Erstens erzeugt $\mathscr{B}$ den Vektorraum V, und daher hat jeder Vektor eine Darstellung als Linearkombination. Zweitens ist diese Darstellung eindeutig, weil $\mathscr{B}$ linear unabhängig ist.

Die Abbildung $\kappa_{\mathscr{B}}$ ist surjektiv, weil man jede Linearkombination $x_1 v_1 + \ldots + x_n v_n$ bilden kann. Und $\kappa_{\mathscr{B}}$ ist injektiv, weil man aus dem Spaltenvektor

$$\kappa_{\mathscr{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

den Vektor x als Linearkombination $x = x_1 v_1 + \ldots + x_n v_n$ zurückgewinnen kann. Injektiv und surjektiv bedeutet bijektiv, siehe Proposition 2.42.

Beispiel 7.10. (1) Jeder Vektorraum V hat eine ausgezeichnete lineare Abbildung mit Definitionsbereich und Wertebereich V: die **Identität**

$$\mathrm{id}_V : V \to V.$$

(2) Die **Inklusion** eines Unterraums $U \subseteq V$ ist eine lineare Abbildung

$$i : U \to V,$$

weil Addition und Skalarmultiplikation von U mit denen von V übereinstimmen.

(3) Sei V ein K-Vektorraum und $\lambda \in K$. Die **Streckung** mit dem Faktor λ ist die lineare Abbildung

$$\lambda\cdot : V \to V, \qquad v \mapsto \lambda v.$$

(4) Die Ableitung reeller Polynomfunktionen (vom Grad höchstens d) liefert eine lineare Abbildung:

$$\frac{d}{dX} : \operatorname{Pol}_{\mathbb{R},\leq d} \to \operatorname{Pol}_{\mathbb{R},\leq d}$$
$$f(X) \mapsto \frac{d}{dX} f(X) = f'(X).$$

Die Abbildung ist wohldefiniert, denn erstens besitzt eine Polynomfunktion $f(X)$ eine Ableitung $f'(X)$, und zweitens ist die Ableitung $f'(X)$ wieder eine Polynomfunktion. Dies entnimmt man der folgenden Formel (siehe Analysis/Oberstufe). Seien $a_0, \ldots, a_d \in \mathbb{R}$ mit

$$f(x) = a_d x^d + \ldots + a_2 x^2 + a_1 x + a_0$$

für alle $x \in \mathbb{R}$, dann ist

$$f'(x) = d a_d x^{d-1} + \ldots + 2a_2 x + a_1$$

wieder eine Polynomfunktion. Hat $f(X)$ höchstens den Grad d, dann hat $f'(X)$ höchstens den Grad $d-1$.

Die Ableitung ist linear: zu $\lambda \in \mathbb{R}$ und $f(X), g(X) \in \operatorname{Pol}_{\mathbb{R},\leq d}$ gilt für alle $x \in \mathbb{R}$:

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

Beispiel 7.11. Es mag kontraintuitiv sein, aber Drehungen sind lineare Abbildungen! Hier beschreiben wir Drehstreckungen mittels komplexer Zahlen.

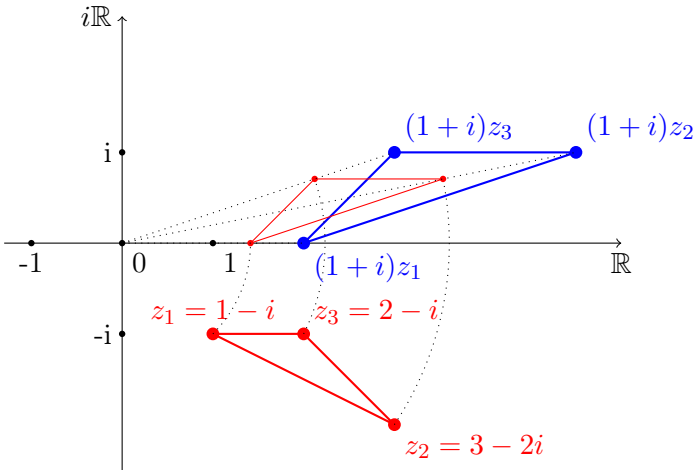


Abbildung 5. Multiplikation mit $w = 1 + i \in \mathbb{C}$ als Drehstreckung.

Sei $w \in \mathbb{C}$, und betrachten wir $\mathbb{C}$ als $\mathbb{R}$-Vektorraum. Dann definiert die Multiplikation mit w eine $\mathbb{R}$-lineare Abbildung

$$f_w : \mathbb{C} \to \mathbb{C}, \qquad f_w(z) = wz.$$

In der Tat gilt für $z_1, z_2 \in \mathbb{C}$ und $\lambda \in \mathbb{R}$:

$$f_w(\lambda z_1 + z_2) = w(\lambda z_1 + z_2) = \lambda w z_1 + w z_2 = \lambda f_w(z_1) + f_w(z_2).$$

Man rechnet leicht nach: für $z, w \in \mathbb{C}$ gilt $f_{zw} = f_z \circ f_w$ und

$$f_z = f_w \quad \Longrightarrow \quad f_z(1) = f_w(1) \quad \Longrightarrow \quad z = w.$$

Hier kann man nun das Assoziativgesetz der Multiplikation in $\mathbb{C}$ „umsonst“ aus der Assoziativität der Verknüpfung (linearer) Abbildungen bekommen. Zu $x, y, z \in \mathbb{C}$ gilt

$$f_{x(yz)} = f_x \circ f_{yz} = f_x \circ (f_y \circ f_z) = (f_x \circ f_y) \circ f_z = f_{xy} \circ f_z = f_{(xy)z}$$

also

$$x(yz) = (xy)z.$$

Die Multiplikationsabbildung f_w hat eine geometrische Interpretation als Drehstreckung von $\mathbb{C}$ aufgefaßt als $\mathbb{R}^2$. Der Spezialfall $w = i$ ist besonders hervorzuheben: die Multiplikation mit i beschreibt eine Drehung um 90^0 gegen den Uhrzeigersinn.

Satz 7.12. *Sei V ein K-Vektorraum mit der Basis $\mathscr{B} = (v_1, \ldots, v_n)$. Sei W ein K-Vektorraum.*

(1) Eine lineare Abbildung $f : V \to W$ ist eindeutig durch ihre Werte

$$f(v_1), \ldots, f(v_n)$$

auf der Basis $\mathscr{B}$ festgelegt: ist $g : V \to W$ eine lineare Abbildung mit

$$g(v_i) = f(v_i), \qquad \text{für alle } i = 1, \ldots, n,$$

dann gilt $g = f$.

(2) Zu jedem Tupel $(w_1, \ldots, w_n)$ von Vektoren aus W gibt es eine lineare Abbildung $f : V \to W$ mit

$$f(v_i) = w_i \qquad \text{für alle } i = 1, \ldots, n.$$

Beweis. (1) Seien $f, g : V \to W$ wie im Satz und $x \in V$ beliebig. Da $\mathscr{B}$ Basis ist, gibt es $a_i \in K$, $i = 1, \ldots, n$ mit

$$x = a_1 v_1 + \ldots + a_n v_n.$$

Dann gilt

$$f(x) = f(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i f(v_i) = \sum_{i=1}^{n} a_i g(v_i) = g(\sum_{i=1}^{n} a_i v_i) = g(x).$$

(2) Wir benutzen die Abbildung $\varphi : K^n \to W$ aus Beispiel 7.7 bezüglich der Vektoren $w_1, \ldots, w_n \in W$. Die gesuchte Abbildung $f : V \to W$ ist nun

$$f = \varphi \circ \kappa_{\mathscr{B}}.$$

Dieses f ist linear als Komposition linearer Abbildungen, Proposition 7.6, und für alle $i = 1, \ldots n$ gilt

$$f(v_i) = \varphi(\kappa_{\mathscr{B}}(v_i)) = \varphi(e_i) = w_i.$$

In dieser Rechnung ist $e_i \in K^n$ der i-te Standardbasisvektor. Damit ist $f : V \to W$ eine lineare Abbildung mit der gewünschten Eigenschaft. □

Beispiel 7.13. Der $\mathbb{R}$-Vektorraum $\mathbb{C}$ hat die $\mathbb{R}$-Basis $1, i$. Wir können also durch

$$1 \mapsto 1 \qquad i \mapsto -i$$

eine eindeutig bestimmte $\mathbb{R}$-lineare Abbildung $^- : \mathbb{C} \to \mathbb{C}$ definieren. Dies ist die **komplexe Konjugation**. Der Beweis aus dem Existenzsatz für lineare Abbildungen, Satz 7.12, führt auf die Formel

$$z = a + bi \mapsto \bar{z} = a - bi.$$

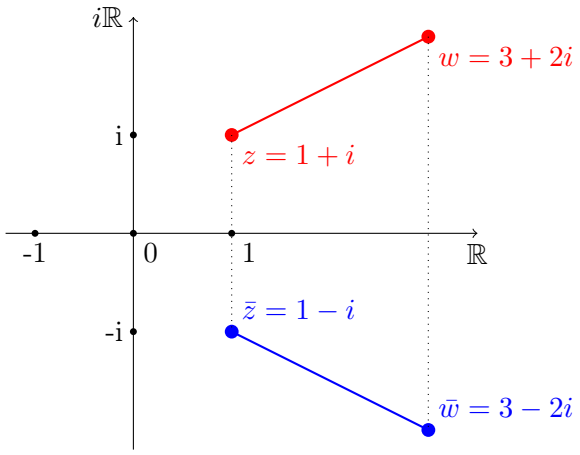


ABBILDUNG 6. Komplexe Konjugation als Spiegelung.

Die komplexe Konjugation ist geometrisch nichts anderes als die **Spiegelung** an der reellen Achse (x-Achse). Umgekehrt haben wir somit gelernt, daß die Spiegelung an der reellen Achse für $\mathbb{R}^2$ interpretiert als Ebene eine $\mathbb{R}$-lineare Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ liefert. In der Standardbasis $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ des $\mathbb{R}^2$ wird diese Spiegelung durch

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

eindeutig festgelegt und folgt daher allgemein der Formel

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y \end{pmatrix}.$$

Bemerkung 7.14. Man rechnet leicht nach, daß die komplexe Konjugation die Addition und Multiplikation auf $\mathbb{C}$ respektiert. Damit meinen wir: für alle $z, w \in \mathbb{C}$ gilt

$$\overline{z + w} = \bar{z} + \bar{w}$$
$$\overline{zw} = \bar{z}\bar{w}.$$

Man spricht davon, daß die Komplexe Konjugation einen Körperautomorphismus von $\mathbb{C}$ ist. Die Teilmenge der Fixpunkte der komplexen Konjugation ist

$$\{z \in \mathbb{Z} \ ; \ z = \bar{z}\} = \{z \in \mathbb{Z} \ ; \Im(z) = 0\} = \mathbb{R}$$

gerade der Unterkörper der reellen Zahlen.

7.2. **Kern, Bild und Rang.** Wir setzen unsere systematische Untersuchung linearer Abbildungen fort.

Definition 7.15. Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen.

(1) Der **Kern** von f ist

$$\ker(f) = \{x \in V \ ; \ f(x) = 0\} \subseteq V.$$

(2) Das **Bild** von f ist

$$\mathrm{im}(f) = f(V) = \{y \in W \ ; \ \exists x \in V : \ y = f(x)\} \subseteq W.$$

(3) Der **Rang** von f ist die Dimension des Bildes

$$\mathrm{rg}(f) = \dim_K(\mathrm{im}(f)).$$

Beispiel 7.16. Sei K ein Körper und

$$a_1X_1 + \ldots + a_nX_n = 0$$

eine lineare Gleichung in den Variablen $X_1, \ldots, X_n$. Der Lösungsraum ist der Kern der linearen Abbildung

$$f : K^n \to K, \quad f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = a_1x_1 + \ldots + a_nx_n.$$

Diese Beobachtung verallgemeinert sich auf lineare Gleichungssysteme und dern Lösungsraum.

Proposition 7.17. *Kern und Bild einer linearen Abbildung sind Unterräume.*

Beweis. Das ist ein Spezialfall der folgenden Proposition, und zwar für $U = 0$ bzw. $U = V$. □

Proposition 7.18. *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen.*

(1) Sei U ein Unterraum von W. Dann ist das Urbild

$$f^{-1}(U) = \{x \in V \ ; \ f(x) \in U\} \subseteq V$$

ein Unterraum.

(2) Sei U ein Unterraum von V. Dann ist das Bild

$$f(U) = \{y \in W \ ; \ \exists x \in U : \ y = f(x)\} \subseteq W$$

ein Unterraum.

Beweis. Wir prüfen das Unterraumkriterium, Satz 4.13.

(1) Wenn $x, y \in f^{-1}(U)$ und $\lambda \in K$, dann ist $\lambda x + y \in f^{-1}(U)$, weil

$$f(\lambda x + y) = \lambda f(x) + f(y) \in U.$$

Fehlt noch, daß $f^{-1}(U)$ nicht leer ist. Aber $f(0) = 0 \in U$ zeigt $0 \in f^{-1}(U)$.

(2) Wenn $y_1, y_2 \in f(U)$, dann gibt es $x_1, x_2 \in U$ mit $y_i = f(x_i)$ für $i = 1, 2$. Dann gilt für alle $\lambda \in K$

$$\lambda y_1 + y_2 = \lambda f(x_1) + f(x_2) = f(\lambda x_1 + x_2) \in f(U),$$

weil $\lambda x_1 + x_2 \in U$. Fehlt noch, daß $f(U)$ nicht leer ist. Aber $0 \in U$ zeigt $0 = f(0) \in f(U)$. □

Satz 7.19. *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen.*

(1) f ist injektiv $\iff \ker(f) = 0$.
(Diese 0 ist der Nullvektorraum als Unterraum von V.)

(2) f ist surjektiv $\iff \mathrm{im}(f) = W$.

Beweis. (1) Sei f injektiv und $x \in \ker(f)$. Dann gilt

$$f(x) = 0 = f(0)$$

und daher $x = 0$, weil f injektiv ist. Also ist $\ker(f) = \{0\}$ der Nullvektorraum.

Jetzt zeigen wir die umgekehrte Richtung. Sei $\ker(f) = 0$. Angenommen für $x, y \in V$ gilt $f(x) = f(y)$. Dann ist

$$f(x - y) = f(x) - f(y) = 0,$$

somit $x - y \in \ker(f)$. Weil $\ker(f) = 0$ folgt $x - y = 0$, also $x = y$. Dies zeigt, daß f injektiv ist.

Aussage (2) folgt sofort aus den Definitionen. □

Lemma 7.20. *Sei $f : V \to W$ eine lineare Abbildung und $v_1, \ldots, v_n$ ein Erzeugendensystem von V. Dann gilt*
$$f(V) = \langle f(v_1), \ldots, f(v_n)\rangle_K.$$

Beweis. Nach Voraussetzung gilt $V = \langle v_1, \ldots, v_n\rangle_K = \{\sum_{i=1}^n a_i v_i \ ; \ a_i \in K\}$. Daraus folgt
$$f(V) = \left\{ f(\sum_{i=1}^n a_i v_i) \ ; \ a_i \in K \right\} = \left\{ \sum_{i=1}^n a_i f(v_i) \ ; \ a_i \in K \right\} = \langle f(v_1), \ldots, f(v_n)\rangle_K. \qquad \square$$

Satz 7.21 (Kern–Bild–Dimensionsformel)**.** *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Es sind äquivalent*

(a) V ist endlichdimensional.

(b) Kern und Bild von f sind endlichdimensional.

Wenn $\dim(V) < \infty$, dann gilt die ***Kern-Bild-Dimensionsformel***
$$\dim(V) = \dim(\ker(f)) + \dim(\mathrm{im}(f)).$$
Damit gilt die ***Rangformel***
$$\mathrm{rg}(f) = \dim(V) - \dim(\ker(f)).$$

Beweis. (a) $\Longrightarrow$ (b): Sei V endlichdimensional. Dann ist $\ker(f)$ endlichdimensional als Unterraum von V. Sei $v_1, \ldots, v_n$ eine Basis von V. Dann ist $f(v_1), \ldots, f(v_n)$ nach Lemma 7.20 ein Erzeugendensystem von $\mathrm{im}(f) = f(V)$. Damit ist $f(V)$ endlich erzeugt. Nach dem Basisauswahlsatz, Satz 5.22, hat $f(V)$ eine endliche Basis: somit $\dim(\mathrm{im}(f)) < \infty$.

(b) $\Longrightarrow$ (a): Seien $\ker(f)$ und $\mathrm{im}(f)$ endlichdimensional. Wir wählen eine Basis $v_1, \ldots, v_r$ von $\mathrm{im}(f)$ und eine Basis $w_1, \ldots, w_s$ von $\ker(f)$. Sodann wählen wir Urbilder $v_i' \in V$ von v_i, also $f(v_i') = v_i$, für $i = 1, \ldots, r$. Wenn wir nun zeigen, daß
$$\mathscr{B} = (v_1', \ldots, v_r', w_1, \ldots, w_s)$$
eine Basis von V ist, dann haben wir sowohl $\dim(V) < \infty$ gezeigt, als auch die Dimensionsformel
$$\dim(V) = r + s = \dim(\ker(f)) + \dim(\mathrm{im}(f)).$$

$\mathscr{B}$ ist linear unabhängig: Seien $a_1, \ldots, a_r, b_1, \ldots, b_s \in K$ mit
$$a_1 v_1' + \ldots + a_r' v_r + b_1 w_1 + \ldots + b_s w_s = 0.$$
Darauf wenden wir f an und finden
$$\begin{aligned} 0 &= f(a_1 v_1' + \ldots + a_r v_r' + b_1 w_1 + \ldots + b_s w_s) \\ &= a_1 f(v_1') + \ldots + a_r f(v_r') + b_1 f(w_1) + \ldots + b_s f(w_s) \\ &= a_1 v_1 + \ldots + a_r v_r + b_1 \cdot 0 + \ldots + b_s \cdot 0 \\ &= a_1 v_1 + \ldots + a_r v_r. \end{aligned}$$
Weil $(v_1, \ldots, v_r)$ eine Basis von $\mathrm{im}(f)$ ist, muß diese Linearkombination trivial sein: $a_i = 0$ für alle $i = 1, \ldots, r$. Gehen wir damit in die Ausgangsrelation zurück, dann finden wir
$$0 = a_1 v_1' + \ldots + a_r v_r' + b_1 w_1 + \ldots + b_s w_s = b_1 w_1 + \ldots + b_s w_s. \tag{7.1}$$
Nun benutzen wir, daß $(w_1, \ldots, w_s)$ eine Basis von $\ker(f)$ ist. Die Linearkombination (7.1) muß auch trivial sein: $b_i = 0$ für alle $i = 1, \ldots, s$. Dies zeigt, daß $\mathscr{B}$ linear unabhängig ist.

$\mathscr{B}$ ist Erzeugendensystem: Wir müssen zu einem beliebigen $x \in V$ eine Darstellung als Linearkombination aus Vektoren von $\mathscr{B}$ finden. Zunächst bilden wir x ab auf $y = f(x) \in \mathrm{im}(f)$. Weil $(v_1, \ldots, vr)$ eine Basis von $\mathrm{im}(f)$ ist, gibt es $a_1, \ldots, a_r \in K$ mit
$$y = a_1 v_1 + \ldots + a_r v_r.$$

Jetzt setzen wir

$$x' = x - (a_1 v'_1 + \ldots + a_r v'_r).$$

Dann ist $x' \in \ker(f)$, denn

$$\begin{aligned} f(x') &= f\big(x - (a_1 v'_1 + \ldots + a_r v'_r)\big) \\ &= f(x) - f(a_1 v'_1 + \ldots + a_r v'_r) \\ &= f(x) - (a_1 f(v'_1) + \ldots + a_r f(v'_r)) \\ &= y - (a_1 v_1 + \ldots + a_r w_r) = 0. \end{aligned}$$

Dann gibt es $b_1, \ldots, b_s$ mit

$$x' = b_1 w_1 + \ldots + b_s w_s$$

und so

$$x = (x - x') + x' = (a_1 v'_1 + \ldots + a_r v'_r) + (b_1 w_1 + \ldots + b_s w_s) \in \langle \mathscr{B} \rangle_K.$$

Die Rangformel ist nun eine unmittelbare Konsequenz aus der Definition des Rangs und der Kern-Bild-Dimensionsformel. □

Bemerkung 7.22. Satz 7.21 lehrt uns zu denken, daß eine lineare Abbildung $f : V \to W$ den Vektorraum V in die Teile $\ker(f)$ und $\mathrm{im}(f)$ zerlegt.

Der folgende Satz ist eine überraschende Pointe über lineare Abbildungen. Für allgemeine Abbildungen haben injektiv und surjektiv ja recht wenig miteinander zu tun. Eine vergleichbare Aussage für endliche Mengen findet sich in Aufgabe 2.13.

Satz 7.23. *Seien V und W endlichdimensionale K-Vektorräume und $f : V \to W$ eine lineare Abbildung. Wenn $\dim(V) = \dim(W)$, dann sind äquivalent:*

(a) f ist injektiv,
(b) f ist surjektiv,
(c) f ist bijektiv.

Beweis. Nach Satz 7.19 gilt

$$f \text{ injektiv} \iff \ker(f) = 0.$$

Anwendung von Korollar 5.34 auf $0 \subseteq \ker(f)$ zeigt

$$\ker(f) = 0 \iff \dim(\ker(f)) = 0.$$

Satz 7.21 ergibt

$$\dim(\ker(f)) = 0 \iff \dim(V) = \dim(\mathrm{im}(f)).$$

Erneut Korollar 5.34, diesmal angewandt auf $\mathrm{im}(f) \subseteq W$, zeigt

$$\dim(W) = \dim(\mathrm{im}(f)) \iff \mathrm{im}(f) = W \iff f \text{ surjektiv}.$$

Weil nach Voraussetzung $\dim(V) = \dim(W)$, haben wir damit insgesamt (a) $\iff$ (b).

Weil außerdem (c) $\iff$ (a) und (b), folgt schon alles. □

Bemerkung 7.24. Die Voraussetzung von Satz 7.23 ist zum Beispiel gegeben, wenn $V = W$ ist, also $f : V \to V$ eine lineare Abbildung von einem endlich dimensionalen K-Vektorraum V nach demselben Vektorraum ist.

7.3. **Projektoren.** Die lineare Abbildung $P : \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch

$$P\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

projiziert vom dreidimensionalen Raum auf eine Koordinatenebene. Ein Projektor ist eine lineare Abbildung, die gleiches leistet.

Definition 7.25. Ein **Projektor** eines K-Vektorraums V ist ein Endomorphismus $P : V \to V$ mit der Eigenschaft
$$P \circ P = P.$$
Wir schreiben auch P^2 für $P \circ P$.

Proposition 7.26. *Sei $P : V \to V$ ein Projektor des K-Vektorraums V. Dann ist*
$$\ker(P) \oplus \operatorname{im}(P) = V.$$

Beweis. Zu jedem Vektor $x \in V$ gilt
$$P(x - P(x)) = P(x) - P^2(x) = P(x) - P(x) = 0.$$
Daher folgt $V = \ker(P) + \operatorname{im}(P)$ aus der Zerlegung
$$x = (x - P(x)) + P(x) \in \ker(P) + \operatorname{im}(P).$$
Die Summe ist direkt, weil für $x \in \ker(P) \cap \operatorname{im}(P)$ gilt: es gibt $y \in V$ mit $x = P(y)$ und dann
$$x = P(y) = P(P(y)) = P(x) = 0.$$ □

Proposition 7.27. *Sei $V = U \oplus W$ eine direkte Summe. Dann gibt es einen eindeutigen Projektor $P : V \to V$ mit*
(i) $\ker(P) = U$, *d.h.,* $P(y) = 0 \iff y \in U$.
(ii) $\operatorname{im}(P) = W$, *und genauer* $W = \{x \in V \ ; \ P(x) = x\}$,

Beweis. Die Eindeutigkeit ist klar, denn P ist auf den Summanden U und W, welche V erzeugen, durch (i) und (ii) vorgegeben.

Wir zeigen als nächstes, daß es eine lineare Abbildung $P : V \to V$ mit den geforderten Eigenschaften gibt, und erst im zweiten Schritt, daß dieses P ein Projektor ist. Zu $v \in V$ gibt es nach Voraussetzung eindeutig $v = x + y$ mit $x \in U$ und $y \in W$. Dann setzen wir
$$P(v) = y.$$
Zu $v' = x' + y'$ mit $x' \in U$ und $y' \in W$, also $P(v') = y'$, und $\lambda \in K$ beliebig ist die Zerlegung von $\lambda v + v'$ gegeben als
$$\lambda v + v' = \lambda(x + y) + (x' + y') = (\lambda x + x') + (\lambda y + y')$$
mit $\lambda x + x' \in U$ und $\lambda y + y' \in W$. Die Eindeutigkeit besagt, daß
$$P(\lambda v + v') = \lambda y + y' = \lambda P(v) + P(v').$$
Damit ist P linear.

Für $y \in W$ gilt per Definition $P(P(x)) = P(x)$, und für $y \in U$ gilt $P(P(y)) = P(0) = 0 = P(y)$. Also gilt $P^2 = P$ für alle Vektoren in U und in W, also auf einem Erzeugendensystem von V. Daher ist auch $P^2 = P$ ganz allgemein.

Zum Schluß bestimmen wir Kern und Bild von P. Sei wie oben $v = x + y$ mit $x \in U$ und $y \in W$. Weil $P(v) = y$ folgt
$$v \in \ker(P) \iff y = 0 \iff v = x \iff v \in U,$$
also $\ker(P) = U$. Das Bild von P ist offensichtlich in W enthalten, weil $y \in W$. Und zu $v \in W$ beliebig gehört die Zerlegung $v = x + y$ mit $y = v$ und $x = 0$, und dann ist $P(y) = y = v$. Dies zeigt $\operatorname{im}(P) = W$. □

Bemerkung 7.28. Proposition 7.27 erklärt, warum Projektoren so heißen. In einer direkten Summenzerlegung $V = U \oplus W$ gibt es einen Projektor $P : V \to V$, der $v = x + y$ mit $x \in U$ und $y \in W$, also v mit eindeutiger Darstellung durch das Tupel $(x, y) \in U \times W$, durch $P(v) = y$ auf den W-Anteil projiziert.

Beispiel 7.29. Der Projektor im Sinne von Proposition 7.27 zur direkten Summe $K^n = U_i \oplus W_i$ aus Beispiel 5.43 ist

$$P_i : K^n \to K^n, \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix} \mapsto P_i(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

7.4. **Isomorphismen und Klassifikation.** Als nächstes beschreiben wir, wann Vektorräume eine vergleichbare Gestalt haben.

Definition 7.30. Es besteht die folgende Terminologie für lineare Abbildungen.

(1) Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Man nennt f
- einen **Monomorphismus**, wenn f injektiv ist,
- einen **Epimorphismus**, wenn f surjektiv ist, und
- einen **Endomorphismus**, wenn $V = W$.

(2) Ein **Isomorphismus** von K-Vektorräumen V und W ist eine lineare Abbildung

$$f : V \to W,$$

für die es eine lineare Abbildung $g : W \to V$ gibt, die ein **Inverses** von f ist:

$$f \circ g = \mathrm{id}_W \quad \text{und} \quad g \circ f = \mathrm{id}_V .$$

Zwei K-Vektorräume V, W sind **isomorph**, wenn es einen Isomorphismus $f : V \to W$ gibt. Man schreibt dann

$$V \simeq W.$$

(3) Ein **Automorphismus** eines K-Vektorraums V ist ein bijektiver Endomorphismus, also ein Isomorphismus $f : V \to V$.

Beispiel 7.31. Sei V ein K-Vektorraum mit Basis $\mathscr{B}$ und Dimension $\dim(V) = n$. Dann ist die Koordinatenvektorabbildung

$$\kappa_{\mathscr{B}} : V \xrightarrow{\sim} K^n$$

ein Isomorphismus von K-Vektorräumen, der **Koordinatenisomorphismus** zu $\mathscr{B}$. Die inverse Abbildung $K^n \to V$ wurde in Proposition 7.8 konstruiert.

Proposition 7.32. *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Dann ist f ein Isomorphismus genau dann, wenn f bijektiv ist.*
Die zu f inverse Abbildung ist dann auch ein Isomorphismus.

Beweis. Wenn f ein Inverses hat, dann ist f bijektiv nach Satz 3.9. Dies zeigt die eine Richtung.

Nehmen wir nun an, f sei bijektiv. Da die inverse Abbildung von Mengen eindeutig ist, kommt für das Inverse $g : W \to V$ nur die Umkehrabbildung

$$f^{-1} : W \to V$$

in Frage, die nach Satz 3.9 exisiert, weil f bijektiv ist. Der Gehalt der Proposition versteckt sich nun in der leichten Beobachtung, daß dieses f^{-1} automatisch eine lineare Abbildung ist. Für alle $x_1, x_2 \in W$ gibt es $y_1, y_2 \in V$ mit $f(y_i) = x_i$ für $i = 1, 2$. Außerdem gilt dann $f^{-1}(x_i) = y_i$ für $i = 1, 2$. Es folgt für alle $\lambda \in K$

$$f^{-1}(\lambda x_1 + x_2) = f^{-1}(\lambda f(y_1) + f(y_2)) = f^{-1}(f(\lambda y_1 + y_2)) = \lambda y_1 + y_2 = \lambda f^{-1}(x_1) + f^{-1}(x_2).$$

Dieses f^{-1} ist dann auch eine bijektive lineare Abbildung, somit ein Isomorphismus. □

Beispiel 7.33. Sei X eine Menge. Wir betrachten die Potenzmenge $\mathscr{P}(X)$ wie in Beispiel 4.11 als $\mathbb{F}_2$-Vektorraum mit der symmetrischen Differenz als Addition. Ebenso betrachten wir $\mathrm{Abb}(X, \mathbb{F}_2)$ mit der üblichen $\mathbb{F}_2$-Vektorraumstruktur wie in Beispiel 4.9. Dann ist die Bijektion

$$\begin{aligned} \chi : \ \mathscr{P}(X) &\to \mathrm{Abb}(X, \mathbb{F}_2) \\ U &\mapsto \mathbf{1}_U \end{aligned}$$

ein Isomorphismus von $\mathbb{F}_2$-Vektorräumen. Dabei ist $\mathbf{1}_U$ die charakteristische Funktion von U:

$$\mathbf{1}_U(x) = \begin{cases} 1 & x \in U \\ 0 & x \notin U. \end{cases}$$

Die Umkehrabbildung ist

$$\begin{aligned} S : \ \mathrm{Abb}(X, \mathbb{F}_2) &\to \mathscr{P}(X) \\ f &\mapsto S(f) := \{x \in X \ ; \ f(x) \neq 0\}, \end{aligned}$$

denn man verifiziert leicht für eine beliebige Teilmenge $U \subseteq X$, also $U \in \mathscr{P}(X)$

$$S(\mathbf{1}_U) = U$$

und für eine beliebige Abbildung $f : X \to \mathbb{F}_2$

$$\mathbf{1}_{S(f)} = f.$$

Die Abbildung χ ist linear, weil für Teilmengen $A, B \subseteq X$ gilt

$$\chi(A) + \chi(B) = \mathbf{1}_A + \mathbf{1}_B = \mathbf{1}_{A \cup B} + \mathbf{1}_{A \cap B} = \mathbf{1}_{A \cup B} - \mathbf{1}_{A \cap B} = \mathbf{1}_{A \triangle B} = \chi(A \triangle B).$$

Beachten Sie dabei die Koeffizienten in $\mathbb{F}_2$! Und zum anderen für $\lambda = 0$ oder 1 durch Fallunterscheidung

$$\chi(\lambda A) = \mathbf{1}_{\lambda A} = \lambda \mathbf{1}_A = \lambda \chi(A).$$

Die einfachste Methode, $\mathscr{P}(X)$ mit der angegebenen Struktur als $\mathbb{F}_2$-Vektorraumstruktur zu erkennen, besteht im **Strukturtransport** der Vektorraumstruktur auf $\mathrm{Abb}(X, \mathbb{F}_2)$. Dazu macht man sich zuerst klar, daß die angegebene Abbildung $\chi : \mathscr{P}(X) \to \mathrm{Abb}(X, \mathbb{F}_2)$ bijektiv ist, und übersetzt dann Addition bzw. Skalarmultiplikation von Funktionen in Addition bzw. Skalarmultiplikation von Mengen. Man muß also unter anderem einsehen, daß für Teilmengen $A, B \subseteq X$ gilt

$$\chi^{-1}(\chi(A) + \chi(B)) = A \triangle B.$$

Proposition 7.34. *Sei K ein Körper.*

(1) Je zwei K-Vektorräume der gleichen Dimension sind isomorph.

(2) Sind V und W zwei K-Vektorräume und sind $\mathscr{B} = (v_1, \ldots, v_n)$ bzw. $\mathscr{C} = (w_1, \ldots, w_n)$ Basen von V bzw. von W, dann gibt es genau einen Isomorphismus $f : V \to W$, der $\mathscr{B}$ als Tupel auf $\mathscr{C}$ abbildet:

$$f(v_i) = w_i$$

für alle $i = 1, \ldots, n$.

Beweis. Aussage (1) folgt sofort aus Aussage (2), der wir uns nun zuwenden. Die behauptete lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, n$ gibt es nach Satz 7.12. Des weiteren liefert der gleiche Satz auch eine lineare Abbildung $g : W \to V$ mit $g(w_i) = v_i$ für $i = 1, \ldots, n$. Die Komposition

$$g \circ f : V \to V$$

bildet v_i auf v_i ab für $i = 1, \ldots, n$. Dies tut auch die Identität id_V. Nach Satz 7.12 folgt

$$g \circ f = \mathrm{id}_V .$$

Analog gilt $f \circ g = \mathrm{id}_W$. Damit ist g ein Inverses zu f. Somit ist f bijektiv und ein Isomorphismus wie behauptet. □

Proposition 7.35. *Sei K ein Körper.*

(1) Zwei isomorphe K-Vektorräume haben die gleiche Dimension.

(2) *Ist $f : V \to W$ ein Isomorphismus von K-Vektorräumen, und sei $(v_1, \dots, v_n)$ ein Tupel aus V, dann gilt:*

$$(v_1, \dots, v_n) \text{ ist Basis von } V \iff (f(v_1), \dots, f(v_n)) \text{ ist Basis von } W.$$

Beweis. Aussage (1) folgt sofort aus Aussage (2).

(2) Sei $\mathscr{B} = (v_1, \dots, v_n)$ eine Basis. Dann folgt aus Lemma 7.20, daß $W = f(V)$ von $\mathscr{C} = (f(v_1), \dots, f(v_n))$ erzeugt wird. Damit $\mathscr{C}$ eine Basis ist, müssen noch zeigen, daß $\mathscr{C}$ linear unabhängig ist.

Wenn $\sum_{i=1}^n a_i f(v_i) = 0$ mit $a_i \in K$, dann ist auch

$$f(\sum_{i=1}^n a_i v_i) = \sum_{i=1}^n a_i f(v_i) = 0.$$

Weil f injektiv ist, folgt $\sum_{i=1}^n a_i v_i = 0$. Da $\mathscr{B}$ eine Basis ist, muß dann $a_i = 0$ für alle $i = 1, \dots, n$ sein. Also ist $(f(v_1), \dots, f(v_n))$ linear unabhängig.

Um von „$\mathscr{C}$ ist Basis“ auf „$\mathscr{B}$ ist Basis“ zu schließen, benutzen wir die Umkehrabbildung $f^{-1} : W \to V$. Wegen $f^{-1}(f(v_i)) = v_i$ für alle $i = 1, \dots, n$, folgt aus dem gerade Bewiesenen angewandt auf f^{-1} und das Tupel $\mathscr{C} = (f(v_1), \dots, f(v_n))$ die Behauptung. □

Satz 7.36 (Klassifikation endlichdimensionaler Vektorräume)**.** *Sei K ein Körper. Für jedes $n \in \mathbb{N}_0$ gibt es bis auf Isomorphie genau einen K-Vektorraum der Dimension n.*

Beweis. Wir tragen zusammen. *Existenz:* der Vektorraum K^n hat Dimension n. *Eindeutigkeit bis auf Isomorphie:* Proposition 7.34. □

Bemerkung 7.37. Isomorphe Vektorräume erlauben Fragen der linearen Algebra in dem einem Vektorraum durch einen Isomorphismus in den anderen zu transportieren. Zum Beispiel, ob ein Tupel von Vektoren linear unabhängig ist. Mit dem inversen Isomorphismus kann man die Antwort wieder zurücktransportieren. Dies ist sehr praktisch, da man zum Rechnen damit statt in einem beliebigen n-dimensionalen K-Vektorraum auch im K^n arbeiten kann, sofern man einen entsprechenden Isomorphismus $V \simeq K^n$ und sein Inverses zum Übersetzen kennt.

Beide Seiten haben ihre Berechtigung:

- allgemeine Vektorräume V sind zum begrifflichen, konzeptionellen Denken,
- der K^n ist zum Rechnen geeignet.

Übungsaufgaben zu §7

Übungsaufgabe 7.1. Zeigen Sie daß auf einer Menge von K-Vektorräumen der Isomorphiebegriff eine Äquivalenzrelation definiert, d.h., für K-Vektorräume U, V, W gilt:

(i) $V \simeq V$.
(ii) $U \simeq V \Longrightarrow V \simeq U$.
(iii) Wenn $U \simeq V$ und $V \simeq W$, dann auch $U \simeq W$.

Übungsaufgabe 7.2. Den $\mathbb{R}$-Vektorraum der Polynomfunktionen $\mathbb{R} \to \mathbb{R}$ vom Grad $\leq d$ bezeichnen wir mit

$$\mathrm{Pol}_{\mathbb{R},\leq d} = \{f : \mathbb{R} \to \mathbb{R} \ ; \ \exists a_i \in \mathbb{R}, i = 1, \dots, d \text{ mit } f(x) = \sum_{i=0}^d a_i x^i \quad \forall x \in \mathbb{R}\}.$$

Zeigen Sie, daß für alle $f \in \mathrm{Pol}_{\mathbb{R},\leq d}$ und alle $a \in \mathbb{R}$ gilt:

$$f(x + a) = \sum_{n=0}^d \frac{a^n}{n!} \partial_x^n f(x)$$

als Funktionen in $x \in \mathbb{R}$. Zeigen Sie, daß

$$f(x) \mapsto f(x+a)$$

und

$$f(x) \mapsto \sum_{n=0}^{d} \frac{a^n}{n!} \partial_x^n f(x)$$

lineare Abbildungen $\mathrm{Pol}_{\mathbb{R},\leq d} \to \mathrm{Pol}_{\mathbb{R},\leq d}$ sind und vergleichen Sie die Werte auf einer Basis.

Übungsaufgabe 7.3. Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen.

(1) Es sind äquivalent:
   (a) f ist surjektiv.
   (b) f bildet jede Basis von V auf ein Erzeugendensystem von W ab.
   (c) f bildet eine Basis von V auf ein Erzeugendensystem von W ab.

(2) Es sind äquivalent:
   (a) f ist injektiv.
   (b) f bildet jede Basis von V auf ein linear unabhängiges Tupel aus W ab.
   (c) f bildet eine Basis von V auf ein linear unabhängiges Tupel aus W ab.

(3) Es sind äquivalent:
   (a) f ist bijektiv.
   (b) f bildet jede Basis von V auf eine Basis von W ab.
   (c) f bildet eine Basis von V auf eine Basis von W ab.

Übungsaufgabe 7.4. Berechnen Sie für den K-Vektorraum $V = K^n$ die Koordinatenvektorabbildung

$$\kappa_{\mathscr{B}} : V \to K^n$$

für die Standardbasis $\mathscr{B} = (e_1, \ldots, e_n)$.

Übungsaufgabe 7.5. Seien V ein endlichdimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Wir kürzen die n-fache Komposition von f mit sich selbst durch

$$f^n = \underbrace{f \circ \ldots \circ f}_{n\text{-mal}}$$

ab. Dabei ist $f^0 = \mathrm{id}_V$ per Konvention. Zeigen Sie:

(a) Es gilt
$$V = \mathrm{im}(f^0) \supseteq \mathrm{im}(f^1) \supseteq \mathrm{im}(f^2) \supseteq \ldots.$$
Es gibt $n_0 \in \mathbb{N}_0$ mit $\mathrm{im}(f^n) = \mathrm{im}(f^{n_0})$ für alle $n \geq n_0$.

(b) Es gilt
$$0 = \ker(f^0) \subseteq \ker(f^1) \subseteq \ker(f^2) \subseteq \ldots$$
Es gibt $n_1 \in \mathbb{N}_0$ mit $\ker(f^n) = \ker(f^{n_1})$ für alle $n \geq n_1$.

Wir setzen $U = \mathrm{im}(f^{n_0})$ mit dem n_0 aus (a) und $N = \ker(f^{n_1})$ mit dem n_1 aus (b).

Zeigen Sie weiter: Es ist $V = U \oplus N$ eine innere direkte Summe und f respektiert die direkte Summenzerlegung, d.h. für $x \in U$ und $y \in N$ beliebig gilt wieder $f(x) \in U$ und $f(y) \in N$.

Übungsaufgabe 7.6 (Universelle Eigenschaft der (abstrakten) direkten Summe). Seien V_1 und V_2 zwei K-Vektorräume. Eine **(abstrakte) direkte Summe** von V_1 und V_2 ist ein K-Vektorraum V zusammen mit K-linearen Abbildungen

$$\iota_i : V_i \to V$$

für $i = 1, 2$, so daß zu jedem K-Vektorraum T und je zwei linearen Abbildungen $f_i : V_i \to T$, $i = 1, 2$ genau eine lineare Abbildung

$$F : V \to T$$

existiert mit

$$f_i = F \circ \iota_i$$

für $i = 1, 2$.

Zeigen Sie:

(a) Zu je zwei K-Vektorräumen V_1 und V_2 gibt es eine (abstrakte) direkte Summe.

(b) Je zwei (abstrakte) direkte Summen zu K-Vektorräumen V_1 und V_2 sind isomorph.

(c) Eine konkrete Konstruktion einer (abstrakten) direkten Summe ist wie folgt. Wir definieren zuerst

$$V_1 \oplus V_2 = V_1 \times V_2$$

als Menge. Dann defineren wir die Addition von Tupeln durch

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2)$$

für alle $x_i, y_i \in V_i$, $i = 1, 2$. Zuletzt setzen wir für $\lambda \in K$ die Skalarmultiplikation durch

$$\lambda(x_1, x_2) := (\lambda x_1, \lambda x_2).$$

(d) Sind U_1, U_2 Unterräume von V und ist ihre Summe $U_1 \oplus U_2 = U_1 + U_2 \subseteq V$ eine innere direkte Summe, dann gibt es einen eindeutigen Isomorphismus

$$U_1 \oplus U_2 \xrightarrow{\sim} U_1 + U_2,$$

der in der Konstruktion der direkten Summe als Vektorraum der Paare die Form

$$(x_1, x_2) \mapsto x_1 + x_2$$

hat.

Bemerkung: Teil (d) zeigt: innere direkte Summen sind auch abstrakte direkte Summen.

Übungsaufgabe 7.7 (Strukturtransport). Sei V eine Menge und sei $f : V \to K^n$ eine bijektive Abbildung von V mit der K^n zugrundeliegenden Menge. Zeigen Sie: Dann gibt es auf V eine K-Vektorraumstruktur, so daß f eine lineare Abbildung wird. Diese K-Vektorraumstruktur ist eindeutig.

Übungsaufgabe 7.8. Sei $P : V \to V$ ein Projektor. Zeigen Sie, daß die Abbildung $\pi = \mathrm{id}_V - P : V \to V$ definiert durch

$$\pi(x) = x - P(x)$$

für alle $x \in V$ auch ein Projektor ist. Zeigen Sie weiter, daß die Zerlegung $V = \ker(P) \oplus \mathrm{im}(P)$ und die Zerlegung $V = \ker(\pi) \oplus \mathrm{im}(\pi)$ nur die Reihenfolge der Summanden tauscht.

8. Matrizen und lineare Abbildungen

Wir haben die lineare Algebra und damit die Theorie der Vektorräume zunächst abstrakt aufgebaut. Um konkrete Fragen zu beantworten, braucht man eine Beschreibung in Koordinaten.

8.1. **Der Vektorraum der Matrizen.**

Definition 8.1. Sei K ein Körper und $n, m \in \mathbb{N}_0$. Eine $n \times m$**-Matrix** mit Koeffizienten aus K ist ein $\{1, \ldots, n\} \times \{1, \ldots, m\}$-Tupel von Elementen aus K. Eine Matrix A besteht damit aus Elementen

$$a_{ij} \in K \quad \text{für } i = 1, \ldots, n \text{ und } j = 1, \ldots, m$$

und wird in einem rechteckigen Schema angeordnet:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Eine $n \times m$-Matrix hat n-viele **Zeilen** und m-viele **Spalten**. Die Menge aller $n \times m$-Matrizen mit Koeffizienten aus K wird mit

$$\mathrm{M}_{n\times m}(K) = \{A \ ; \ A \text{ ist } n \times m\text{-Matrix mit Koeffizienten aus } K\}$$

bezeichnet. Im Spezialfall $n = m$ schreiben wir kurz

$$\mathrm{M}_n(K) := \mathrm{M}_{n\times n}(K).$$

Bemerkung 8.2. Wir vereinbaren die folgende Kurznotation für $n \times m$-Matrizen A als

$$A = (a_{ij})_{\substack{1\le i\le n\\ 1\le j\le m}},$$

und sogar einfach nur

$$A = (a_{ij}),$$

wenn der Bereich für die Indizes i, j aus dem Kontext klar ist.

Man beachte Zeile vor Spalte! In der Matrix $A = (a_{ij})$ ist i der Zeilenindex, der die Nummer der Zeile beschreibt (zählt vertikal von oben nach unten), und j der Spaltenindex, der die Nummer der Spalte beschreibt (zählt horizontal von links nach rechts).

Notation 8.3. Das Symbol **Kronecker–Delta** δ_{ij} für i, j mit $i = 1, \ldots, n$ und $j = 1, \ldots, m$ hat den Wert

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Man kann auch allgemeiner analog ein Kronecker-Delta δ_{ab} für Elemente $a, b \in X$ einer Menge X definieren:

$$\delta_{ab} = \begin{cases} 1 & a = b, \\ 0 & a \neq b. \end{cases}$$

Beispiel 8.4. (1) Die **Nullmatrix** 0 hat als Einträge an jeder Stelle $0 \in K$. Vorsicht: es gibt für jede Abmessung $n \times m$ eine eigene Nullmatrix.

(2) Sei $n \in \mathbb{N}_0$. Die **Einheitsmatrix** der Größe n ist die $n \times n$-Matrix

$$\mathbf{1}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij})_{\substack{1\le i\le n\\ 1\le j\le m}},$$

mit Einträgen 1, falls Zeilenindex = Spaltenindex (das nennt man die **Diagonale der Matrix**), und 0 sonst. Wenn die Abmessung $n \times n$ der Matrix aus dem Kontext klar ist, dann schreiben wir auch schon einmal $\mathbf{1}$ für $\mathbf{1}_n$.

(3) Hier ist ein Beispiel für eine 2×3-Matrix mit Koeffizienten aus $\mathbb{F}_p$:

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

(4) Ein **(Spalten-)Vektor** $x \in K^n$ mit Einträgen $x_1, \ldots, x_n$ ist eine $n\times 1$-Matrix mit denselben Einträgen

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

(5) Eine $1 \times m$-Matrix mit Einträgen $x_1, \ldots, x_m$ wird auch **Zeilenvektor** genannt und hat die Form

$$(x_1 \ \ldots \ x_m).$$

Ein Zeilenvektor unterscheidet sich von einem Tupel also durch das Fehlen der Kommata.

(6) Wir definieren zu $\alpha = 1, \ldots, n$ und $\beta = 1, \ldots, m$ eine $n \times m$-Matrix $e_{\alpha\beta}$, deren Eintrag an (i, j)-Position gegeben ist durch

$$(e_{\alpha\beta})_{ij} = \begin{cases} 1 & (\alpha, \beta) = (i, j) \\ 0 & \text{sonst} \end{cases} = \delta_{i\alpha} \cdot \delta_{j\beta}.$$

Die Matrix hat nur einen Eintrag $\neq 0$, und der ist eine 1 in der α-ten Zeile und β-ten Spalte.

$$e_{\alpha\beta} = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ \vdots & \cdots & 1 & \cdots & \vdots \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \leftarrow \alpha$$
$$\uparrow\beta$$

Notation 8.5. Sei K ein Körper und $n, m \in \mathbb{N}_0$. Die Spalten einer $n \times m$-Matrix sind Spaltenvektoren aus K^n, und zwar m-viele. Umgekehrt kann man aus $v_1, \ldots, v_m \in K^n$ mit Koordinaten

$$v_j = \begin{pmatrix} v_{1j} \\ \vdots \\ v_{nj} \end{pmatrix}$$

eine Matrix in **Spaltenvektorschreibweise** machen

$$[v_1, \ldots, v_m] := (v_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}} \in \mathrm{M}_{n \times m}(K),$$

deren j-te Spalte gerade v_j ist.

Definition 8.6. Sei K ein Körper und $n, m \in \mathbb{N}_0$. Auf $\mathrm{M}_{n \times m}(K)$ definieren wir für alle $A = (a_{ij}), B = (b_{ij}) \in \mathrm{M}_{n \times m}(K)$ die **Addition von Matrizen** als

$$A + B := (a_{ij} + b_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}} \tag{8.1}$$

und die **Skalarmultiplikation von Matrizen** mit $\lambda \in K$ als

$$\lambda A := (\lambda a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}}. \tag{8.2}$$

Beispiel 8.7. Es gilt in $\mathrm{M}_n(K)$

$$\mathbf{1}_n = e_{11} + e_{22} + \ldots + e_{nn},$$

wobei die Matrizen e_{ii} auch $n \times n$-Matrizen sein sollen.

Proposition 8.8. *Sei K ein Körper und $n, m \in \mathbb{N}_0$. Dann ist $\mathrm{M}_{n \times m}(K)$ mit der durch (8.1) und (8.2) definierten Addition und Skalarmultiplikation ein K-Vektorraum. Die Abbildung*

$$\mathrm{M}_{n \times m}(K) \to K^{nm}, \qquad (a_{ij}) \mapsto \begin{pmatrix} a_{11} \\ \vdots \\ a_{nm} \end{pmatrix},$$

in der die nm-vielen Einträge der Matrix übereinander als Spaltenvektor geschrieben werden, ist ein Isomorphismus von K-Vektorräumen.

Beweis. Das ist wirklich klar. □

Korollar 8.9. *Eine Basis von* $\mathrm{M}_{n\times m}(K)$ *ist gegeben durch die* nm*-vielen Vektoren*

$$e_{\alpha\beta} \in \mathrm{M}_{n\times m}(K) \quad \text{für } (\alpha,\beta) \in \{1,\dots,n\}\times\{1,\dots,m\}$$

und es gilt

$$\dim(\mathrm{M}_{n\times m}(K)) = nm.$$

Beweis. Das folgt aus Proposition 8.8 mittels Proposition 7.35: Das Tupel der $e_{\alpha\beta}$ ist das Bild der Standardbasis von K^{nm} mittels des inversen Isomorphismus zu dem aus Proposition 8.8. □

8.2. Transponieren und das Standardskalarprodukt.

Definition 8.10. Sei K ein Körper und $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Die **transponierte Matrix** zu A ist die Matrix

$$A^t = (a_{ji})_{\substack{1\le j\le m\\ 1\le i\le n}} \in \mathrm{M}_{m\times n}(K)$$

mit dem Eintrag a_{ji} and der Stelle (i,j). Dies ist die Spiegelung von A an der **Hauptdiagonale** entlang der Einträge $a_{11}, a_{22}, \dots, a_{ii}, \dots$.

Beispiel 8.11. (1) Die Abmessungen der Matrix vertauschen sich!

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

(2) $\mathbf{1}^t = \mathbf{1} \in \mathrm{M}_n(K)$.
(3) Für alle $A \in \mathrm{M}_{n\times m}(K)$ gilt: $(A^t)^t = A$.

Für das Transponieren von Matrizen gelten die folgenden Regeln.

Proposition 8.12. *Das Transponieren* $A \mapsto A^t$ *ist eine bijektive lineare Abbildung*

$$(\ )^t : \mathrm{M}_{n\times m}(K) \to \mathrm{M}_{m\times n}(K).$$

Die Inverse Abbildung ist das Transponieren mit vertauschten Abmessungen n *und* m.

(i) Für $A, B \in \mathrm{M}_{n\times m}(K)$ *gilt*

$$(A+B)^t = A^t + B^t.$$

(ii) Für $A \in \mathrm{M}_{n\times m}(K)$ *und* $\lambda \in K$ *gilt*

$$(\lambda A)^t = \lambda A^t.$$

Beweis. Das ist trivial. Die Basis aus Korollar 8.9 wird durch das Transponieren auf eine Permutation eben dieser Basis für den Raum der Matrizen mit den umgekehrten Abmessungen abgebildet: $e^t_{\alpha\beta} = e_{\beta\alpha}$. □

Definition 8.13. Sei K ein Körper und $n \in \mathbb{N}$. Das **Standardskalarprodukt** ist die folgende auf zwei Argumenten aus K^n definierte Abbildung:

$$-\bullet-\colon K^n \times K^n \to K$$

$$(x,y) \mapsto x \bullet y := \sum_{i=1}^{n} x_i y_i, \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Notation 8.14. Das Standardskalarprodukt hat viele Notationen. Die hier gewählte lehnt sich an den amerikanischen Sprachraum an, bei dem dieses Produkt auch das Dot-Produkt genannt wird, daher der dicke Dot $\bullet$.

Bemerkung 8.15. Das Standardskalarprodukt ist nur für $K = \mathbb{R}$ ein Skalarprodukt, für allgemeine Körper K hingegen ist es nur eine symmmetrische Bilinearform. Das ist ein Vorgriff auf das Thema Bilinearformen, das im Laufe der Vorlesungen Lineare Algebra 1+2 noch behandelt werden wird.

Proposition 8.16. *Das Standardskalarprodukt ist linear in beiden Argumenten: für alle $x, x', y, y' \in K^n$ und $\lambda \in K$ gilt*

(1) $x \bullet (y + y') = x \bullet y + x \bullet y'$,
(2) $(x + x') \bullet y = x \bullet y + x' \bullet y$,
(3) $\lambda(x \bullet y) = (\lambda x) \bullet y = x \bullet \lambda y$.

Das Standardskalarprodukt ist symmetrisch: für alle $x, y \in K^n$ gilt:

$$x \bullet y = y \bullet x.$$

Beweis. Das ist trivial nachzurechnen. Es folgt später aus der Formel für das Skalarprodukt mittels Matrixmultiplikation

$$x \bullet y = x^t y$$

aus den Eigenschaften derselben, Proposition 8.22. Daher sparen wir uns den Beweis hier. □

8.3. **Matrixmultiplikation.** Lineare Abbildungen lassen sich komponieren. Wir werden lineare Abbildungen durch Matrizen beschreiben, und daher sollte man auch Matrizen komponieren können. Dies führt zur Matrixmultiplikation.

Definition 8.17. Seien K ein Körper und $\ell, n, m \in \mathbb{N}_0$. Wir definieren **Matrixmultiplikation** für $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$ und $B = (b_{jk}) \in \mathrm{M}_{m\times \ell}(K)$ durch

$$A \cdot B := (\sum_{j=1}^{m} a_{ij} b_{jk})_{\substack{1\le i\le n \\ 1\le k\le \ell}} \in \mathrm{M}_{n\times \ell}(K).$$

Statt $A \cdot B$ schreiben wir auch AB.

Bemerkung 8.18. Matrixmultiplikation funktioniert nur, wenn die Matrizen aufeinander abgestimmte Abmessungen haben.

AB nur wenn **Spaltenanzahl** von A = **Zeilenanzahl** von B.

Bemerkung 8.19. Matrixmultiplikation ist keine Multiplikation der Form $M \times M \to M$ auf einer festen Menge M, sondern auf Matrizen passender Größe:

$$\begin{aligned} \mathrm{M}_{n\times m}(K) \times \mathrm{M}_{m\times \ell}(K) &\to \mathrm{M}_{n\times \ell}(K) \\ (A, B) &\mapsto AB \end{aligned}$$

Für den Eintrag von AB an der Stelle ik, also i-te Zeile und k-te Spalte, braucht man nur die i-te Zeile von A und die k-te Spalte von B:

$$(AB)_{ik} = ik\text{ - Eintrag von } \begin{pmatrix} \vdots & & \vdots \\ a_{i1} & \cdots a_{ij} \cdots & a_{im} \\ \vdots & & \vdots \end{pmatrix} \cdot \begin{pmatrix} \cdots & b_{1k} & \cdots \\ & \vdots & \\ & b_{jk} & \\ & \vdots & \\ \cdots & b_{mk} & \cdots \end{pmatrix}$$

$$= \begin{pmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{im} \end{pmatrix} \cdot \begin{pmatrix} b_{1k} \\ \vdots \\ b_{jk} \\ \vdots \\ b_{mk} \end{pmatrix} = \sum_{j=1}^{m} a_{ij} b_{jk}.$$

Schreiben wir $B = [w_1, \ldots, w_\ell]$ in Spaltenvektorschreibweise mit den Vektoren $w_k \in K^m$ in der k-ten Spalte, und ebenso in Spaltenvektorschreibweise $A^t = [v_1^t, \ldots, v_\ell^t]$ mit den Zeilenvektoren $v_i \in \mathrm{M}_{1\times m}(K)$ in der i-ten Zeile von A (diese wird beim Transponieren zur i-ten Spalte von A^t), dann wird obige Rechnung zu

$$(AB)_{ik} = v_i \cdot w_k = v_i^t \bullet w_k.$$

Das erste Produkt $\cdot$ ist Matrixmultiplikation, und das zweite Produkt $\bullet$ beschreibt das Standardskalarprodukt. Beachten Sie, daß die Anforderungen an die Abmessungen von A und B genau bedeuten, daß v_i^t und w_k im selben K^m liegen. Insgesamt ergibt sich die folgende kompakte Formel für die Matrixmultiplikation

$$AB = (v_i^t \bullet w_k)_{ik}.$$

Beispiel 8.20.

(1) Für alle $x, y \in K$ gilt in $\mathrm{M}_2(K)$:

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

(2) Für alle $x, y, a, b \in K$ gilt in $\mathrm{M}_2(K)$:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \cdot \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} = \begin{pmatrix} ax & 0 \\ 0 & by \end{pmatrix}.$$

(3) Im Allgemeinen ist Matrixmultiplikation nicht kommutativ. Das geht schon alleine formal nicht, wenn es sich nicht um quadratische Matrizen handelt, weil die Multiplikation in umgekehrter Reihenfolge nicht definiert ist oder andere Abmessungen hat. Besonders drastisch ist es bei $1 \times n$ und $n \times 1$, wie das folgende Beispiel illustriert

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} = \begin{pmatrix} \frac{n(n+1)}{2} \end{pmatrix} \qquad \in \mathrm{M}_1(K),$$

$$\begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2 & \cdots & 2 \\ \vdots & & & \vdots \\ n & n & \cdots & n \end{pmatrix} \qquad \in \mathrm{M}_n(K).$$

Aber auch im quadratischen Fall ist $AB = BA$ nicht der Normalfall:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Proposition 8.21. *Transponieren vertauscht die Reihenfolge bei Matrixmultiplikation: für $A \in \mathrm{M}_{n\times m}(K)$ und $B \in \mathrm{M}_{m\times r}(K)$ gilt*

$$(AB)^t = B^t A^t.$$

Beweis. Sei $A = (a_{ij})$ und $B = (b_{jk})$, dann ist der ki-Eintrag von $(AB)^t$ gerade

$$(AB)^t_{ki} = (AB)_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} = \sum_{j=1}^{m} b_{jk} a_{ij} = \sum_{j=1}^{m} (B^t)_{kj} (A^t)_{ji} = (B^t A^t)_{ki}. \qquad \square$$

Proposition 8.22. *Matrixmultiplikation ist assoziativ, distributiv und verträglich mit Skalarmultiplikation:*

(1) Für alle $A \in \mathrm{M}_{r\times n}(K)$ und $B \in \mathrm{M}_{n\times m}(K)$ und $C \in \mathrm{M}_{m\times s}(K)$ gilt

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

(2) Für alle $A \in \mathrm{M}_{n\times m}(K)$ und $B, C \in \mathrm{M}_{m\times s}(K)$ gilt

$$A \cdot (B + C) = A \cdot B + A \cdot C.$$

(3) Für alle $B, C \in \mathrm{M}_{n\times m}(K)$ und $A \in \mathrm{M}_{m\times s}(K)$ gilt

$$(B + C) \cdot A = B \cdot A + C \cdot A.$$

(4) Für alle $A \in \mathrm{M}_{r\times n}(K)$ und $B \in \mathrm{M}_{n\times m}(K)$ und $\lambda \in K$ gilt

$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$

Beweis. (1) Sei $A = (a_{ij}) \in \mathrm{M}_{r\times n}(K)$, $B = (b_{jk}) \in \mathrm{M}_{n\times m}(K)$ und $C = (c_{k\ell}) \in \mathrm{M}_{m\times s}(K)$. Dann ist der Eintrag in der i-ten Zeile und ℓ-ten Spalte wie folgt:

$$\begin{aligned} \big((AB)C\big)_{i\ell} &= \sum_{k=1}^{m} (AB)_{ik} c_{k\ell} = \sum_{k=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} b_{jk} \Big) c_{k\ell} = \sum_{k=1}^{m} \sum_{j=1}^{n} (a_{ij} b_{jk}) c_{k\ell} \\ &= \sum_{j=1}^{n} \sum_{k=1}^{m} a_{ij} (b_{jk} c_{k\ell}) = \sum_{j=1}^{n} a_{ij} \Big(\sum_{k=1}^{m} b_{jk} c_{k\ell} \Big) = \sum_{j=1}^{n} a_{ij} (BC)_{j\ell} = \big(A(BC)\big)_{i\ell}. \end{aligned}$$

(2) Sei $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$, $B = (b_{jk}) \in \mathrm{M}_{m\times s}(K)$ und $C = (c_{jk}) \in \mathrm{M}_{m\times s}(K)$. Dann ist der Eintrag in der i-ten Zeile und k-ten Spalte wie folgt:

$$\begin{aligned} \big(A(B+C)\big)_{ik} &= \sum_{j=1}^{m} a_{ij} (B+C)_{jk} = \sum_{j=1}^{m} a_{ij} (b_{jk} + c_{jk}) = \sum_{j=1}^{m} a_{ij} b_{jk} + a_{ij} c_{jk} \\ &= \sum_{j=1}^{m} a_{ij} b_{jk} + \sum_{j=1}^{m} a_{ij} c_{jk} = (AB)_{ik} + (AC)_{ik} = (AB + AC)_{ik}. \end{aligned}$$

(3) Das folgt formal aus (2) angewandt auf die Transponierten. Denn aus

$$A^t (B+C)^t = A^t (B^t + C^t) \overset{(2)}{=} A^t B^t + A^t C^t = (BA)^t + (CA)^t = (BA + CA)^t$$

folgt durch transponieren die Behauptung:

$$(B+C)A = \Big(A^t (B+C)^t \Big)^t = \Big((BA + CA)^t \Big)^t = BA + CA.$$

(4) Das folgt trivial aus den Definitionen von Matrix- und Skalarmultiplikation. $\square$

Wir betrachten einen wichtigen Spezialfall, indem wir einen Spaltenvektor aus K^n als $n \times 1$-Matrix auffassen.

$n \times m$-Matrix mal Vektor aus K^m ergibt einen Vektor in K^n.

Korollar 8.23. *Sei $A \in \mathrm{M}_{n\times m}$. Die **Linksmultiplikation mit** A*

$$\begin{aligned} L_A : K^m &\to K^n \\ x &\mapsto Ax \end{aligned}$$

ist eine lineare Abbildung.

Beweis. Wir fassen Spaltenvektoren $x, y \in K^m$ als $m \times 1$-Matrizen auf. Dann gilt nach Proposition 8.22 (2) und (4) für alle $\lambda \in K$

$$L_A(\lambda x + y) = A(\lambda x + y) = A(\lambda x) + Ay = \lambda(Ax) + Ay = \lambda L_A(x) + L_A(y). \qquad \square$$

Beispiel 8.24. Für $A = \begin{pmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \end{pmatrix} \in \mathrm{M}_{2\times 3}(\mathbb{R})$ ist $L_A : \mathbb{R}^3 \to \mathbb{R}^2$ gegeben durch

$$L_A(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + 3y + 5z \\ 7x + 11y + 13z \end{pmatrix}.$$

Sehen Sie die Koeffizienten von A in der Formel für L_A?

Definition 8.25. Der **Nullraum** (oder **Kern**) einer Matrix A ist der Kern der zugehörigen linearen Abbildung L_A:

$$\ker(A) := \ker(L_A).$$

Das folgende Lemma interpretiert die Spalten einer Matrix A und eines Matrixprodukts BA. Diese Formeln erleichtern manchmal die Berechnung des Matrixprodukts.

Lemma 8.26. *Seien $v_1, \ldots, v_m \in K^n$ die Spalten der Matrix $A = [v_1, \ldots, v_m] \in \mathrm{M}_{n\times m}(K)$.*

(1) Sei e_j der j-te Standardbasisvektor von K^m. Dann ist

$$Ae_j = v_j.$$

(2) Sei $B \in \mathrm{M}_{r\times n}(K)$. Dann ist

$$BA = [Bv_1, \ldots, Bv_m].$$

Beweis. (1) Die k-te Koordinate von e_j wird durch δ_{kj} beschrieben. Damit rechnet man dann den i-ten Eintrag aus zu

$$(Ae_j)_i = \sum_{k=1}^{m} a_{ik}\delta_{kj} = a_{ij} = (v_j)_i.$$

(2) Die j-te Spalte von BA ist nach (1) und Proposition 8.22 (1) gerade

$$(BA)e_j = B(Ae_j) = Bv_j. \qquad \square$$

Bemerkung 8.27. Die Rechnung von Lemma 8.26 (1) ist graphisch einfacher zu begreifen:

$$Ae_j = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{im} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{nj} \end{pmatrix}$$

Lemma 8.28. *Sei $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Dann ist der Eintrag an Position (i,j) gerade*

$$a_{ij} = e_i^t A e_j = e_i \bullet A e_j.$$

Beweis. Wir nutzen Lemma 8.26 und rechnen

$$e_i^t A e_j = e_i^t \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = \sum_{k=1}^{n} \delta_{ik} a_{kj} = a_{ij}.$$

Die Gleichung $e_i \bullet A e_j = e_i^t A e_j$ folgt aus der Definition des Standardskalarprodukts. □

8.4. **Koordinaten für lineare Abbildungen.** Wir beschreiben nun die fundamentale Beziehung zwischen linearen Abbildungen und Matrizen.

Definition 8.29. Sei V ein K-Vektorraum der endlichen Dimension $\dim(V) = m$ mit Basis $\mathscr{B} = (b_1, \ldots, b_m)$, und sei W ein K-Vektorraum der endlichen Dimension $\dim(W) = n$ mit Basis $\mathscr{C} = (c_1, \ldots, c_n)$. Sei

$$f : V \to W$$

eine lineare Abbildung. Wir definieren die **Darstellungsmatrix** von f bezüglich der Basen $\mathscr{B}$ und $\mathscr{C}$ als

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) := [\kappa_{\mathscr{C}}(f(b_1)), \kappa_{\mathscr{C}}(f(b_2)), \ldots, \kappa_{\mathscr{C}}(f(b_m))] \in \mathrm{M}_{n\times m}(K).$$

Die j-te Spalte von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$ enthält die Koordinaten von $f(b_j)$ bezüglich $\mathscr{C}$.

Konkret, wenn $a_{ij} \in K$ mit

$$f(b_j) = \sum_{i=1}^{n} a_{ij} c_i,$$

dann gilt

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) = (a_{ij})_{\substack{1\le i\le n \\ 1\le j\le m}}.$$

Bemerkung 8.30. Die Notation $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(-)$ ist so gewählt, daß stets Koordinaten in der Quelle bezüglich der oben indizierten Basis und im Ziel bezüglich der unten indizierte Basis zu betrachten sind. Es geht also von oben nach unten. Im Folgenden lernen wir einige Formeln kennen, bei denen mehrere von einer Basis abhängende „Faktoren“ nebeneinander stehen. Dabei wird stets, von rechts nach links gelesen, eine unten indizierte Basis im nächsten Faktor oben stehen. Damit kann man sich die Formeln leichter merken!

Beispiel 8.31. Sei $A \in \mathrm{M}_{n\times m}(K)$. Die Darstellungsmatrix zu $L_A : K^m \to K^n$ bezüglich der Standardbasen $\mathscr{E} = (e_1, \ldots, e_m)$ von K^m und $\mathscr{F} = (e_1, \ldots, e_n)$ von K^n berechnen wir durch ($\kappa_{\mathscr{F}}$ ist die Identität auf K^n)

$$\kappa_{\mathscr{F}}(L_A(e_j)) = L_A(e_j) = A e_j = \text{ } j\text{-te Spalte von } A$$

nach Lemma 8.26. Also gilt, wie könnte es auch anders sein,

$$\mathrm{M}_{\mathscr{F}}^{\mathscr{E}}(L_A) = [\kappa_{\mathscr{F}}(L_A(e_1)), \ldots, \kappa_{\mathscr{F}}(L_A(e_m))] = A.$$

Die Beschreibung eines Vektors als Spaltenvektor seiner Koordinaten bezüglich einer Basis und die Beschreibung einer linearen Abbildung durch ihre Darstellungsmatrix paßt zusammen. In Koordinatenschreibweise ist jede lineare Abbildung die Linksmultiplikation mit der Darstellungsmatrix.

Proposition 8.32. *Seien K ein Körper und V, W zwei K-Vektorräume. Sei $f : V \to W$ eine lineare Abbildung, und seien $\mathscr{B} = (b_1, \ldots, b_m)$ bzw. $\mathscr{C} = (c_1, \ldots, c_n)$ Basen von V bzw.*

von W. Dann gilt für alle $x \in V$ als Gleichung in K^n:

$$\kappa_{\mathscr{C}}(f(x)) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\kappa_{\mathscr{B}}(x). \tag{8.3}$$

Bemerkung 8.33. Die Gleichung (8.3) wird auch graphisch durch das folgende kommutative Diagramm (in der Notation von Proposition 8.32) ausgedrückt:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle \kappa_{\mathscr{B}}}\downarrow & & \downarrow{\scriptstyle \kappa_{\mathscr{C}}} \\ K^m & \xrightarrow[L_{\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)}]{} & K^n. \end{array}$$

Das Diagramm (von Vektorräumen und linearen Abbildungen) ist **kommutativ**, wenn zu je zwei Wegen im Diagramm mit selbem Start und selbem Ziel die Komposition der Abbildungen entlang des Weges dieselbe lineare Abbildung ergibt. Hier gibt es nur zwei Wege, rechts oder links um das Quadrat von links oben nach rechts unten. Die entsprechende Gleichheit

$$\kappa_{\mathscr{C}} \circ f = L_{\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)} \circ \kappa_{\mathscr{B}} \tag{8.4}$$

von Abbildungen $V \to K^n$ ist genau (8.3).

Beweis von Proposition 8.32. Wir zeigen (8.4). Da nach Korollar 8.23 Linksmultiplikation mit einer Matrix $A = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$ eine lineare Abbildung ist, reicht es nach Satz 7.12, die Gleichung (8.4) nur auf einer Basis von V nachzuweisen.

Wir nehmen natürlich die Basis $\mathscr{B}$. Für alle $j = 1, \ldots, m$ haben wir $\kappa_{\mathscr{B}}(b_j) = e_j$ mit dem j-ten Standardbasisvektor von K^m. Dann gilt nach Lemma 8.26 und der Definition der Darstellungsmatrix

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\kappa_{\mathscr{B}}(b_j) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)e_j = j\text{-te Spalte von } \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) = \kappa_{\mathscr{C}}(f(b_j)). \qquad \square$$

Beispiel 8.34. (1) Die Abbildung $f : K^3 \to K^2$ definiert durch

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x + 3y + 5z \\ 7x + 11y + 13z \end{pmatrix}$$

für alle $x, y, z \in K$ ist eine lineare Abbildung (Übung!). Wir wählen $\mathscr{B} = (e_1, e_2, e_3)$ und $\mathscr{C} = (\varepsilon_1, \varepsilon_2)$ jeweils die Standardbasis, die wir in K^2 zur besseren Unterscheidung als $\varepsilon_i = e_i$ schreiben. Dann ist

$$f(e_1) = \begin{pmatrix} 2 \\ 7 \end{pmatrix} = 2\varepsilon_1 + 7\varepsilon_2,$$
$$f(e_2) = \begin{pmatrix} 3 \\ 11 \end{pmatrix} = 3\varepsilon_1 + 11\varepsilon_2,$$
$$f(e_3) = \begin{pmatrix} 5 \\ 13 \end{pmatrix} = 5\varepsilon_1 + 13\varepsilon_2,$$

und deshalb

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) = \begin{pmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \end{pmatrix}.$$

In der Tat folgt

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

(2) Die Nullabbildung ist $f_0 : V \to W$ mit $f_0(v) = 0$ für alle $v \in V$. Seien V, W endlichdimensional. Egal welche Basen $\mathscr{B}$ von V und $\mathscr{C}$ von W man benutzt, es gilt stets

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f_0) = 0$$

die Nullmatrix.

(3) Sei V ein endlichdimensionaler K-Vektorraum mit Basis $\mathscr{B}$. Dann ist die Matrix zur Identität

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(\mathrm{id}_V) = \mathbf{1}_n$$

die Einheitsmatrix der Größe $\dim(V) = n$.

(4) Wir betrachten konkreter die Identität $\mathrm{id} : K^2 \to K^2$ auf K^2 und die Basen $\mathscr{B} = (\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix})$ und $\mathscr{C} = (\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix})$. Dann ist

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(\mathrm{id}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

und

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\mathrm{id}) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Die Darstellungsmatrix hängt also entscheidend von der Wahl der Basen $\mathscr{B}$ und $\mathscr{C}$ ab!

(5) Sei $f : V \to W$ ein Isomorphismus von endlichdimensionalen K-Vektorräumen. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V und $\mathscr{C} = (f(b_1), \ldots, f(b_n))$ deren Bild in W. Dann ist $\mathscr{C}$ eine Basis von W und

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) = \mathbf{1}_n.$$

Satz 8.35 (Matrixmultiplikation entspricht Komposition)**.** *Sei K ein Körper und seien U, V, W drei K-Vektorräume mit Basen $\mathscr{A} = (a_1, \ldots, a_\ell)$ von U, und $\mathscr{B} = (b_1, \ldots, b_m)$ von V, und $\mathscr{C} = (c_1, \ldots, c_n)$ von C. Seien $f : V \to W$ und $g : U \to V$ lineare Abbildungen. Dann gilt*

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{A}}(f \circ g) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\, \mathrm{M}_{\mathscr{B}}^{\mathscr{A}}(g).$$

Beweis. Es reicht, die Matrixidentität aus $\mathrm{M}_{n\times\ell}(K)$ spaltenweise nachzuprüfen. Für die j-te Spalte müssen wir mit dem Standardbasisvektor $e_j \in K^\ell$ multiplizieren. Dieser ist $\kappa_{\mathscr{A}}(a_j)$. Wegen (dreifach) Proposition 8.32 gilt nun

$$\begin{aligned}
\mathrm{M}_{\mathscr{C}}^{\mathscr{A}}(f \circ g)e_j &= \mathrm{M}_{\mathscr{C}}^{\mathscr{A}}(f \circ g)\kappa_{\mathscr{A}}(a_j) \\
&= \kappa_{\mathscr{C}}(f \circ g(a_j)) = \kappa_{\mathscr{C}}(f(g(a_j)) \\
&= \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\kappa_{\mathscr{B}}(g(a_j)) \\
&= \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)(\mathrm{M}_{\mathscr{B}}^{\mathscr{A}}(g)\kappa_{\mathscr{A}}(a_j)) = (\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\, \mathrm{M}_{\mathscr{B}}^{\mathscr{A}}(g))e_j.
\end{aligned}$$

□

Bemerkung 8.36. Satz 8.35 beinhaltet die finale Rechtfertigung für die zunächst willkürlich erscheinende Definition der Matrixmultiplikation.

Notation 8.37. Seien $n = r + s$ und $m = t + u$ mit $s, r, t, u \in \mathbb{N}_0$. Aus

$$A \in \mathrm{M}_{r\times t}(K),\ B \in \mathrm{M}_{r\times u}(K),\ C \in \mathrm{M}_{s\times t}(K),\ D \in \mathrm{M}_{s\times u}(K)$$

kann man eine Matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{M}_{n\times m}(K)$$

durch Aneinanderfügen der rechteckigen Teilmatrizen in der angegebenen Anordnung erzeugen. Dies ist eine einfache Version einer **Blockmatrix**.

Beispiel 8.38. Seien V und W zwei endlichdimensionale K-Vektorräume, und sei $f : V \to W$ eine lineare Abbildung. Wir betrachten Basen $(v_1, \dots, v_s)$ von $\mathrm{im}(f)$ und $w_1, \dots, w_r$ von $\ker(f)$ wie im Beweis der Kern–Bild–Dimensionsformel aus Satz 7.21. Wie ebenda ergänzen wir zu einer Basis

$$\mathscr{B} = (v'_1, \dots, v'_r, w_1, \dots, w_s)$$

von V durch Wahl von Urbildern $f(v'_j) = v_j$, $j = 1, \dots, r$. Darüberhinaus ergänzen wir nach dem Basisergänzungssatz die Basis von $\mathrm{im}(f)$ zu einer Basis

$$\mathscr{C} = (v_1, \dots, v_r, u_1, \dots, u_t)$$

von W. Bezüglich dieser Basen hat die Matrix von f eine besonders einfache Form als Blockmatrix mit Zeilenaufteilung $n = r + t$ und Spaltenaufteilung $m = r + s$ als

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{n\times m}(K).$$

In der Tat, die ersten r Basisvektoren v'_j werden zu

$$\kappa_{\mathscr{C}}(f(v'_j)) = \kappa_{\mathscr{C}} v_j) = e_j,$$

also zum j-ten Standardbasisvektor in K^n. Die weiteren s Spalten sind von der Form

$$\kappa_{\mathscr{C}}(f(w_j)) = \kappa_{\mathscr{C}}(0) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

denn $w_j \in \ker(f)$.

Bemerkung 8.39. Das Beispiel 8.38 ist teilweise irreführend. Die Matrix von f wird nur deshalb so einfach, weil man die Basen $\mathscr{B}$ und $\mathscr{C}$ sehr speziell wählt. Die Wahl hängt stark von f ab. Dies bedeutet nicht, daß lineare Abbildungen $f : V \to W$ allein durch die Werte r, s, t aus Beispiel 8.38 festgelegt sind, die ja allein über das Aussehen der Matrix

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$$

entscheiden, obwohl ja nach Satz 10.4 die Matrix eindeutig f beschreibt. Dies liegt am folgenden: in Satz 10.4 sind die Basen $\mathscr{B}$ des Definitionsbereichs V und $\mathscr{C}$ des Wertebereichs W fixiert. Es variiert dann $f : V \to W$ und wird dann bei festem $\mathscr{B}$ und $\mathscr{C}$ eindeutig durch $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$ beschrieben. Im Beispiel 8.38 hingegen suchen wir für jedes f die passenden Basen aus. Das ist etwas völlig anderes.

8.5. **Invertierbare Matrizen und Isomorphismen.**

Definition 8.40. Eine Matrix $A \in \mathrm{M}_n(K)$ heißt **invertierbar**, wenn es ein $B \in \mathrm{M}_n(K)$ gibt mit

$$AB = \mathbf{1}_n = BA.$$

Proposition 8.41 (Die **allgemeine lineare Gruppe**[a]). *Sei K ein Körper und $n \in \mathbb{N}_0$. Dann ist*

$$\mathrm{GL}_n(K) = \{A \in \mathrm{M}_n(K) \ ; \ A \text{ invertierbar}\}$$

mit Matrixmultiplikation als Verknüpfung eine Gruppe.

[a] Die Notation GL kommt aus dem Englischen: general linear group.

Beweis. Zuerst müssen wir zeigen, daß Matrixmultiplikation auf $\mathrm{GL}_n(K)$ eine Verknüpfung definiert. Sei $A, B \in \mathrm{GL}_n(K)$ und Inverse $C, D \in \mathrm{M}_n(K)$, die das bezeugen, also

$$AC = CA = \mathbf{1}_n = BD = DB.$$

Dann gilt

$$(DC)(AB) = D(CA)B = D\mathbf{1}_n B = DB = \mathbf{1}_n$$

und

$$(AB)(DC) = A(BD)C = A\mathbf{1}_n C = AC = \mathbf{1}_n.$$

Damit ist $AB \in \mathrm{GL}_n(K)$.

Assoziativ: Matrixmultiplikation ist assoziativ nach Proposition 8.22. Hier ist weiter nichts zu zeigen.

Existenz eines neutralen Elements: Das neutrale Element ist die Einheitsmatrix $\mathbf{1}_n$. Es gilt für alle $A \in \mathrm{M}_n(K)$

$$A\mathbf{1}_n = A = \mathbf{1}_n A,$$

was für $A = \mathbf{1}_n$ erst einmal zeigt, daß $\mathbf{1}_n \in \mathrm{GL}_n(K)$ ist, und zweitens $\mathbf{1}_n$ als neutrales Element ausweist.

Existenz des Inversen: Zu $A \in \mathrm{GL}_n(K)$ gibt es per Definition $B \in \mathrm{M}_n(K)$ mit $AB = \mathbf{1}_n = BA$. Damit ist zuerst $B \in \mathrm{GL}_n(K)$ auch invertierbar und dann auch das Inverse zu A. □

Satz 8.42. *Seien V, W zwei K-Vektorräume und $f : V \to W$ eine lineare Abbildung. Sei $\mathscr{B} = (b_1, \ldots, b_m)$ eine Basis von V, und sei $\mathscr{C} = (c_1, \ldots, c_n)$ eine Basis von W. Dann sind äquivalent:*

(a) Die Abbildung f ist ein Isomorphismus.

(b) Es ist $n = m$ und die Darstellungsmatrix $A = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$ ist eine invertierbare Matrix.

Falls (a) und (b) zutreffen, dann ist $A^{-1} = \mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(f^{-1})$ die Darstellungsmatrix zur inversen Abbildung $f^{-1} : W \to V$.

Beweis. $(a) \Longrightarrow (b)$: Wenn f ein Isomorphismus ist, dann haben V und W nach Proposition 7.35 die gleiche Dimension:

$$m = \dim(V) = \dim(W) = n.$$

Ferner gibt es die Inverse Abbildung $f^{-1} : W \to V$ nach Proposition 7.32. Wir setzen $B = \mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(f^{-1})$ und finden mit Satz 8.35

$$AB = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\,\mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(f^{-1}) = \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f \circ f^{-1}) = \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(\mathrm{id}_W) = \mathbf{1}_n.$$

Genauso sieht man wegen $f^{-1} \circ f = \mathrm{id}_V$ auch $BA = \mathbf{1}_n$. Somit ist A invertierbar.

$(b) \Longrightarrow (a)$: Sei nun umgekehrt $A \in \mathrm{GL}_n(K)$. Zum Inversen A^{-1} betrachten wir

$$g := \kappa_{\mathscr{B}}^{-1} \circ L_{A^{-1}} \circ \kappa_{\mathscr{C}} : W \xrightarrow{\sim} K^n \xrightarrow{A^{-1}\cdot} K^m \xrightarrow{\sim} V.$$

Wir berechnen für $j = 1, \ldots, n$ die j-te Spalte von $\mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(g)$ durch

$$\kappa_{\mathscr{B}}(g(c_j)) = \kappa_{\mathscr{B}}\big(\kappa_{\mathscr{B}}^{-1} \circ L_{A^{-1}} \circ \kappa_{\mathscr{C}}(c_j)\big) = L_{A^{-1}}(e_j) = A^{-1}e_j.$$

Daraus folgt

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(g) = A^{-1}.$$

Wir rechnen wieder mit Satz 8.35 dann

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f \circ g) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\,\mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(g) = AA^{-1} = \mathbf{1} = \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(\mathrm{id}_W),$$

also ist $f \circ g = \mathrm{id}_W$. Genauso sieht man $g \circ f = \mathrm{id}_V$ aus $A^{-1}A = \mathbf{1}$. Damit ist f ein Isomorphismus mit inverser Abbildung $f^{-1} = g$, und $A^{-1} = \mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(f^{-1})$ folgt aus der Konstruktion von g. □

Korollar 8.43. *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gilt*
$$A \in \mathrm{GL}_n(K) \iff L_A : K^n \to K^n \text{ ist ein Automorphismus.}$$

Beweis. Das ist wegen $A = \mathrm{M}_{\mathscr{E}}^{\mathscr{E}}(L_A)$ ein Spezialfall von Satz 8.42. □

Proposition 8.44. *Ist $A \in \mathrm{GL}_n(K)$, dann gilt $A^t \in \mathrm{GL}_n(K)$ und*
$$(A^t)^{-1} = (A^{-1})^t.$$

Beweis. Wir setzen $B = A^{-1}$ und finden
$$B^t A^t = (AB)^t = \mathbf{1}^t = \mathbf{1}$$
und genauso $A^t B^t = (BA)^t = \mathbf{1}^t = \mathbf{1}$. Das zeigt bereits alles. □

8.6. **Basiswechsel.** Koordinaten hängen von der gewählten Basis ab. Die Basis kann man wechseln, und dann ändern sich die Koordinaten, nicht aber der durch die Koordinaten beschriebene Vektor. Dieses Phänomen betrifft ebenso Koordinaten für lineare Abbildungen, die Darstellungsmatrizen.

Definition 8.45. Sei V ein endlichdimensionaler K-Vektorraum der Dimension $n = \dim(V)$. Die **Basiswechselmatrix** von einer Basis $\mathscr{B}$ in eine Basis $\mathscr{C}$ von V ist
$$S_{\mathscr{C}}^{\mathscr{B}} := \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\mathrm{id}_V) \in \mathrm{M}_n(K).$$

Proposition 8.46. *Sei $S_{\mathscr{C}}^{\mathscr{B}} = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\mathrm{id}_V)$ die Basiswechselmatrix zu Basen $\mathscr{B}$ und $\mathscr{C}$ des K-Vektorraums V. Dann gilt für alle $x \in V$:*
$$\kappa_{\mathscr{C}}(x) = S_{\mathscr{C}}^{\mathscr{B}} \cdot \kappa_{\mathscr{B}}(x).$$

Beweis. Das ist ein Spezialfall von Proposition 8.32. □

Bemerkung 8.47. Wenn wir eine lineare Abbildung f in die Linksmultiplikation mit einer Matrix A übersetzen, dann beschreibt $x \mapsto Ax$, wie sich der Vektor mit Koordinaten x sich unter f verändert. Basiswechsel zeigt die Flexibilität des Konzepts und die Abhängigkeit von der gewählten Basis. Die Basiswechselmatrix gehört zur Abbildung id_V, der durch x beschriebene Vektor ändert sich nicht! Was sich ändert, sind die Koordinaten der Beschreibung von x, weil sich die Basis ändert!

Proposition 8.48. *Seien V, W zwei endlichdimensionale K-Vektorräume. Seien $\mathscr{B}$, $\mathscr{B}'$ Basen von V und $\mathscr{C}$, $\mathscr{C}'$ Basen von W. Sei $f : V \to W$ eine lineare Abbildung und schreiben wir zur Abkürzung $A = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$ und $A' = \mathrm{M}_{\mathscr{C}'}^{\mathscr{B}'}(f)$ für die darstellenden Matrizen. Dann gilt*
$$A' = S_{\mathscr{C}'}^{\mathscr{C}} A S_{\mathscr{B}}^{\mathscr{B}'}.$$

Beweis. Das folgt sofort aus $f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$ und zweimal Satz 8.35. □

Beispiel 8.49. Oft will man im K^n von der Standardbasis $\mathscr{E} = (e_1, \ldots, e_n)$ zu einer anderen Basis $\mathscr{B} = (b_1, \ldots, b_n)$ wechseln. Die Basiswechselmatrix $S_{\mathscr{B}}^{\mathscr{E}}$ ist nicht direkt aus den Vektoren $b_1, \ldots, b_n$ abzulesen. Das ist besser mit dem Basiswechsel in der umgekehrten Richtung. Die entsprechende Basiswechselmatrix ist einfach
$$S_{\mathscr{E}}^{\mathscr{B}} = [b_1, \ldots, b_n] \in \mathrm{M}_n(K).$$
Wie bekommt man die umgekehrte Basiswechselmatrix $S_{\mathscr{B}}^{\mathscr{E}}$?

Proposition 8.50. *Sei V ein K-Vektorraum der Dimension n mit Basen $\mathscr{B}$ und $\mathscr{C}$. Dann ist die Basiswechselmatrix $S^{\mathscr{B}}_{\mathscr{C}} \in \mathrm{GL}_n(K)$ invertierbar mit Inversem*

$$(S^{\mathscr{B}}_{\mathscr{C}})^{-1} = S^{\mathscr{C}}_{\mathscr{B}}.$$

Beweis. Es gilt wegen Satz 8.35

$$S^{\mathscr{B}}_{\mathscr{C}} S^{\mathscr{C}}_{\mathscr{B}} = \mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(\mathrm{id}_V)\,\mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id}_V) = \mathrm{M}^{\mathscr{C}}_{\mathscr{C}}(\mathrm{id}_V \circ \mathrm{id}_V) = \mathrm{M}^{\mathscr{C}}_{\mathscr{C}}(\mathrm{id}_V) = \mathbf{1}_n.$$

Die Gleichung $S^{\mathscr{C}}_{\mathscr{B}} S^{\mathscr{B}}_{\mathscr{C}} = \mathbf{1}_n$ folgt aus dieser durch Vertauschen von $\mathscr{B}$ und $\mathscr{C}$ per Symmetrie. □

Bemerkung 8.51. Mit Proposition 8.50 haben wir nur zur Hälfte beantwortet, wie wir die umgekehrte Basiswechselmatrix $S^{\mathscr{E}}_{\mathscr{B}}$ bekommen. Aber wir haben es auf ein anderes fundamentales Problem zurückgeführt: wie finden wir das Inverse zu einer invertierbaren Matrix?

Proposition 8.50 hat eine Umkehrung.

Proposition 8.52. *Sei V ein K-Vektorraum der Dimension n und $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis. Zu jedem $S \in \mathrm{GL}_n(K)$ gibt es eine Basis $\mathscr{C}$ mit $S = S^{\mathscr{B}}_{\mathscr{C}}$.*

Beweis. Sei $\mathscr{B} = (b_1, \ldots, b_n)$. Dann ist $\mathscr{C} = (c_1, \ldots, c_n)$ gegeben durch

$$c_j = \kappa^{-1}_{\mathscr{B}}(S^{-1} e_j).$$

Zum einen ist $\mathscr{C}$ als Bild der Standardbasis unter dem Isomorphismus $\kappa^{-1}_{\mathscr{B}} \circ L_{S^{-1}}$ eine Basis von V. Und es ist $S^{\mathscr{C}}_{\mathscr{B}} = S^{-1}$, weil die j-te Spalte von $S^{\mathscr{C}}_{\mathscr{B}}$ den Vektor $\kappa_{\mathscr{B}}(c_j) = S^{-1} e_j$ enthält, und das ist die j-te Spalte von S^{-1}. Aus Proposition 8.50 folgt nun die Behauptung $S = S^{\mathscr{B}}_{\mathscr{C}}$. □

Beispiel 8.53. Wir formulieren explizit den folgenden Spezialfall der Basiswechselformel, Proposition 8.48. Sei $L_A : K^n \to K^n$ die Linksmultiplikation mit $A \in \mathrm{M}_n(K)$. Dann gilt $A = \mathrm{M}^{\mathscr{E}}_{\mathscr{E}}(L_A)$ bezüglich der Standardbasis $\mathscr{E} = (e_1, \ldots, e_n)$ von K^n. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine weitere Basis von K^n. Dann ist

$$S := S^{\mathscr{B}}_{\mathscr{E}} = [b_1, \ldots, b_n]$$

und für $B = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(L_A)$ gilt

$$B = S^{-1} A S.$$

Setzt man $T = S^{-1} = S^{\mathscr{E}}_{\mathscr{B}}$, dann ist $T^{-1} = (S^{-1})^{-1} = S$ und $B = T A T^{-1}$.

Übungsaufgaben zu §8

Übungsaufgabe 8.1. Sei $A \in \mathrm{M}_{n\times m}(K)$. Zeigen Sie, daß

$$A^t A$$

eine symmetrische Matrix ist.

Bemerkung: Eine (notwendigerweise) quadratische Matrix $B = (b_{ij})$ heißt **symmetrisch**, wenn für alle i, j gilt $b_{ij} = b_{ji}$.

Übungsaufgabe 8.2. Sei $w = \alpha + i\beta$ eine komplexe Zahl. Bestimmen Sie die Darstellungsmatrix für die Abbildung $f_w(z) = wz$ als Endomorphismus von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum bezüglich der Basis $(1, i)$. Für $w = e^{i\varphi}$ erhalten sie die Drehmatrix für die Drehung um den Winkel φ. Wie lautet diese?

Übungsaufgabe 8.3. Sei V ein endlichdimensionaler K-Vektorraum mit Basis $\mathscr{B}$. Zeigen Sie:

(1) Ein Endomorphismus $P : V \to V$ ist ein Projektor genau dann, wenn $A = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(P)$ die Gleichung

$$A^2 = A$$

erfüllt. (Es ist $A^2 = AA$.)

(2) Zeigen Sie, daß in einer geeigneten Basis die Matrix eines Projektors P die Form

$$\mathbf{1}_{r,n} = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_n(K)$$

als Blockmatrix mit Blöcken zur Aufteilung $n = r + (n-r)$ ist.

(3) Ein Endomorphismus $P : V \to V$ ist ein Projektor genau dann, wenn

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(P) = S\mathbf{1}_{r,n}S^{-1}$$

für eine geeignete Matrix $S \in \mathrm{GL}_n(K)$.

Übungsaufgabe 8.4. Bestimmen Sie eine Matrix für den Endomorphismus $X\frac{d}{dX}$ auf $\mathrm{Pol}_{\mathbb{R},\leq d}$, also die Abbildung

$$P(X) \mapsto X \cdot P'(X).$$

Übungsaufgabe 8.5. Sei V ein K-Vektorraum der Dimension n und seien $\mathscr{B}$ und $\mathscr{C}$ Basen von V. Zeigen Sie

$$\mathscr{B} = \mathscr{C} \quad \Longleftrightarrow \quad \kappa_{\mathscr{B}} = \kappa_{\mathscr{C}}.$$

Übungsaufgabe 8.6. Sei V ein K-Vektorraum mit endlicher Basis $\mathscr{B}$. Zu einem Vektor $v \in V$ definieren wir die lineare Abbildung $f_v : K^1 \to V$ durch

$$f_v(a) = av \qquad \text{für alle } a \in K.$$

Sei $\mathscr{E} = (1)$ die Standardbasis von K^1. Beschreiben Sie die Darstellungsmatrix $\mathrm{M}_{\mathscr{B}}^{\mathscr{E}}(f_v)$ mittels der Koordinatenabbildung $\kappa_{\mathscr{B}}$.

Übungsaufgabe 8.7. Zeigen Sie, daß die Formel

$$\kappa_{\mathscr{C}}(f(x)) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\kappa_{\mathscr{B}}(x)$$

aus Proposition 8.32 ein Spezialfall der Formel

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{A}}(f \circ g) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\,\mathrm{M}_{\mathscr{B}}^{\mathscr{A}}(g).$$

aus Satz 8.35 ist.

Übungsaufgabe 8.8. Sei V ein Vektorraum.

(1) Zeigen Sie, daß die Menge der Automorphismen

$$\mathrm{GL}(V) = \{f : V \to V \ ; \ \text{bijektiv und linear}\}$$

mit Komposition als Verknüpfung eine Gruppe ist.

(2) Finden Sie für den Fall $\dim(V) = n \in \mathbb{N}_0$ eine bijektive Abbildung

$$\varphi : \mathrm{GL}(V) \xrightarrow{\sim} \mathrm{GL}_n(K),$$

welche die Gruppenverknüpfungen respektiert (einen Isomorphismus von Gruppen):

$$\varphi(f \circ g) = \varphi(f)\varphi(g)$$

für alle $f, g \in \mathrm{GL}(V)$.

Übungsaufgabe 8.9. Sei K ein Körper und $n \in \mathbb{N}$. Zeigen Sie, daß es für jede Matrix $S \in \mathrm{GL}_n(K)$ eine Basis $\mathscr{B}$ von K^n gibt, so daß S die Basiswechselmatrix von der Standardbasis $\mathscr{E} = (e_1, \ldots, e_n)$ zur Basis $\mathscr{B}$ ist.

Übungsaufgabe 8.10. Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Zeigen Sie:

(a) Wenn es ein $B \in \mathrm{M}_n(K)$ mit $AB = \mathbf{1}$ gibt, dann ist $A \in \mathrm{GL}_n(K)$.

(b) Wenn es ein $C \in \mathrm{M}_n(K)$ mit $CA = \mathbf{1}$ gibt, dann ist $A \in \mathrm{GL}_n(K)$.

Übungsaufgabe 8.11. Sei K ein Körper und $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(K)$. Finden Sie ein Kriterium in Abhängigkeit von a, b, c, d dafür, daß $A \in GL_2(K)$, und finden Sie eine Formel für A^{-1}.

Anleitung: Machen Sie einen Ansatz $B = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$ und lösen Sie die Gleichungen in den Variablen X, Y, Z, W, die der Gleichung

$$AB = \mathbf{1}_2$$

entsprechen. Das Kriterium tritt als Bedingung für die Lösbarkeit auf. Zeigen Sie, daß für diese Lösungen auch $BA = \mathbf{1}_2$ gilt.

9. Rang und Zeilenstufenform

9.1. **Rang, Spaltenrang und Zeilenrang.** Wir erinnern an die Definition des Rangs einer linearen Abbildung f als Dimension des Bildes $\mathrm{rg}(f) = \dim(\mathrm{im}(f))$. Wir definieren nun auf drei Arten den Rang einer Matrix und zeigen anschließend, daß alle drei Werte übereinstimmen.

Definition 9.1. Seien K ein Körper und $A = [v_1, \ldots, v_m] \in \mathrm{M}_{n\times m}(K)$ eine Matrix in Spaltenvektorschreibweise.

(1) Der **Spaltenraum**[a] von A ist der von den Spalten aufgespannte Unterraum in K^n:
$$C(A) = \langle v_1, \ldots, v_m\rangle_K \subseteq K^n.$$

(2) Der **Rang** von A ist die Dimension des Spaltenraums von A:
$$\mathrm{rg}(A) = \dim C(A).$$

(3) Der **Spaltenrang** von A, ist
$$\text{s-rg}(A) = \text{ maximale Anzahl linear unabhängiger Spalten von } A.$$

(4) Zeilen von A fassen wir als Zeilenvektoren in $\mathrm{M}_{1\times m}(K)$ auf. Der **Zeilenrang** von A ist
$$\text{z-rg}(A) = \text{ maximale Anzahl linear unabhängiger Zeilen von } A.$$

Wir werden zeigen, daß stets $\mathrm{rg}(A) = \text{s-rg}(A) = \text{z-rg}(A)$ gilt, und werden daraufhin nur noch die Notation $\mathrm{rg}(A)$ benutzen. Die Notation für Zeilenrang und Spaltenrang ist nur temporär.

[a] Englisch *column space*.

Beispiel 9.2. Wir betrachten Matrizen in $\mathrm{M}_{2\times 3}(\mathbb{Q})$, und zwar

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 8 \end{pmatrix}.$$

Die Matrix A hat Spaltenrang 1, denn die Spalten sind Vielfache voneinander, aber nicht 0. Der von den Spalten aufgespannte Vektorraum ist daher auch von Dimension 1, daher $\mathrm{rg}(A) = 1$. Der Zeilenrang von A ist ebenso 1.

Die Matrix B hat linear unabhängige Zeilen, daher Zeilenrang 2. Die Spalten sind nicht linear unabhängig, denn die dritte Spalte ist die Summe der ersten und zweiten (in $\mathbb{Q}^2$ sind je drei Vektoren stets linear abhängig; wieso?). Der Spaltenrang ist 2, weil die ersten beiden Spalten eine Basis von $\mathbb{Q}^2$ sind. Daher ist der Spaltenraum $\mathbb{Q}^2$ und ebenfalls $\mathrm{rg}(B) = 2$.

Proposition 9.3. *Sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix und $L_A : K^m \to K^n$ die Linksmultiplikation mit A. Es gilt:*

(1) $\text{z-rg}(A) = \text{s-rg}(A^t)$.
(2) $\mathrm{rg}(A) = \text{s-rg}(A)$.
(3) $\mathrm{rg}(A) = \mathrm{rg}(L_A)$

Beweis. Aussage (1) folgt offensichtlich, weil Transponieren Zeilen und Spalten vertauscht und weil Transponieren K-linear ist: lineare Relationen der Zeilen entsprechen linearen Relationen der transponierten Zeilen, aka Spalten.

(2) Sei $A = [v_1, \dots, v_m]$ die Spaltenvektorschreibweise. Eine maximale linear unabhänigige Teilmenge der Spalten ist eine maximal linear unabhängige Teilmenge $\mathscr{B}$ des Erzeugendensystemes $v_1, \dots, v_m$ des Spaltenraums $C(A)$. Nach dem Beweis von Satz 5.21 ist $\mathscr{B}$ eine Basis von $C(A)$. Daher gilt

$$\text{z-rg}(A) = |\mathscr{B}| = \dim C(A) = \operatorname{rg}(A).$$

(3) Die Bilder $v_1 = Ae_1, \dots, v_m = Ae_m$ der Standardbasis $e_1, \dots, e_m$ unter der Abbildung L_A sind nach Lemma 7.20 ein Erzeugendensystem des Bildes $\operatorname{im}(L_A)$. Dies sind aber genau die Spalten von $A = [v_1, \dots, v_m]$. Daraus folgt:

$$C(A) = \langle v_1, \dots, v_m \rangle_K = \operatorname{im}(L_A),$$

und daher für die Dimensionen

$$\operatorname{rg}(A) = \dim C(A) = \dim(\operatorname{im}(L_A)) = \operatorname{rg}(L_A). \qquad \square$$

Die Begriffe Rang für Matrizen und für lineare Abbildungen sind in der bestmöglichen Weise kompatibel.

Proposition 9.4. *Seien V und W endlichdimensionale K-Vektorräume und sei $f : V \to W$ eine lineare Abbildung. Sei $\mathscr{B} = (b_1, \dots, b_m)$ eine Basis von V und $\mathscr{C}$ eine Basis von W. Dann ist die Einschränkung des Koordinatenisomorphismus $\kappa_{\mathscr{C}}$ ein Isomorphismus des Bildes von f mit dem Spaltenraum der Darstellungsmatrix:*

$$\kappa_{\mathscr{C}}|_{\operatorname{im}(f)} \colon \operatorname{im}(f) \xrightarrow{\sim} C(\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)).$$

Insbesondere gilt: der Rang von f ist gleich dem Rang der Darstellungsmatrix:

$$\operatorname{rg}(f) = \operatorname{rg}(\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)).$$

Beweis. Sei $n = \dim(W)$. Der Koordinatenisomorphismus $\kappa_{\mathscr{C}} : W \to K^n$ bildet $f(b_j)$ auf die j-te Spalte von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$, und damit

$$\operatorname{im}(f) = \langle f(b_1), \dots, f(b_m) \rangle_K \xrightarrow{\kappa_{\mathscr{C}}, \sim} C(\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f))$$

isomorph auf Spaltenraum $C(\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f))$ der Darstellungsmatrix ab. Daher sind die Dimensionen dieser Unterräume gleich, folglich per Definition auch die beiden Ränge. $\square$

Proposition 9.5. *Seien K ein Körper und $n \in \mathbb{N}$. Sei $A \in \mathrm{M}_{n\times m}(K)$ und $S \in \mathrm{GL}_n(K)$ sowie $T \in \mathrm{GL}_m(K)$. Dann gilt*

$$\operatorname{rg}(A) = \operatorname{rg}(SA) = \operatorname{rg}(AT) = \operatorname{rg}(SAT).$$

Beweis. Es reicht, $\operatorname{rg}(A) = \operatorname{rg}(SA)$ und $\operatorname{rg}(A) = \operatorname{rg}(AT)$ für alle A zu zeigen. Denn dann gilt angewandt auf A und auf SA auch

$$\operatorname{rg}(A) = \operatorname{rg}(SA) = \operatorname{rg}((SA)T).$$

Zur invertierbaren Matrix S gehört ein Isomorphismus $L_S : K^n \to K^n$. Wenn $A = [v_1, \dots, v_n]$, dann ist $SA = [Sv_1, \dots, Sv_n]$ und daher für die Spaltenräume

$$C(SA) = \langle Sv_1, \dots, Sv_n \rangle_K = L_S(\langle v_1, \dots, v_n \rangle_K) = L_S(C(A)).$$

Vergleich der Dimension liefert, weil L_S ein Isomorphismus ist,

$$\operatorname{rg}(A) = \dim C(A) = \dim L_S(C(A)) = \dim C(SA) = \operatorname{rg}(SA).$$

Zur invertierbaren Matrix T gehört ein Isomorphismus $L_T : K^m \to K^m$, also ist $L_T(K^m) = K^m$. Damit gilt auch

$$\operatorname{im}(L_{AT}) = L_{AT}(K^m) = L_A(L_T(K^m)) = L_A(K^m) = \operatorname{im}(L_A).$$

Daraus folgt sofort

$$\operatorname{rg}(AT) = \dim(\operatorname{im}(L_{AT})) = \dim(\operatorname{im}(L_A)) = \operatorname{rg}(A). \qquad \square$$

Bemerkung 9.6. Proposition 9.5 folgt sofort aus Proposition 9.4, wenn man sich überlegt, daß geeignete Basiswechsel die Darstellungsmatrix A von L_A in die Matrizen SA, AT und SAT transformieren. Diese Matrizen beschreiben bei geeigneter Interpretation alle die gleiche lineare Abbildung L_A, und haben wegen Proposition 9.4 daher alle den Rang $\operatorname{rg}(L_A)$.

Satz 9.7 (Zeilenrang gleich Spaltenrang)**.** *Sei K ein Körper und sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Dann gilt*

$$\operatorname{rg}(A^t) = \operatorname{rg}(A),$$

mit anderen Worten Zeilenrang ist gleich Spaltenrang.

Beweis. Die Behauptung $\operatorname{rg}(A^t) = \operatorname{rg}(A)$ ist wegen $\text{z-rg}(A) = \operatorname{rg}(A^t)$ und $\operatorname{rg}(A) = \text{s-rg}(A)$ aus Proposition 9.3 in der Tat äquivalent zur Aussage „Zeilenrang ist gleich Spaltenrang“.

Wir wollen Proposition 9.5 dazu benutzen, nach einer geeigneten Basistransformation den Rang von A direkt ablesen zu können. Dazu betrachten wir für die Linksmultiplikation $L_A : K^m \to K^n$ die in Beispiel 8.38 basierend auf dem Beweis der Kern-Bild-Dimensionsformel, Satz 7.21, gefundene einfache Darstellungsmatrix in Bezug auf geeignet angepaßte Basen $\mathscr{B}$ von K^m und $\mathscr{C}$ von K^n. Bezüglich dieser Basen hat die Matrix von L_A die Struktur einer Blockmatrix mit Zeilenaufteilung $n = r + t$ und Spaltenaufteilung $m = r + s$ als

$$B := \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(L_A) = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{n\times m}(K)$$

mit der $r \times r$-Einheitsmatrix $\mathbf{1}_r$. Seien $S = S_{\mathscr{E}}^{\mathscr{C}}$ und $T = S_{\mathscr{B}}^{\mathscr{E}}$ die Basiswechselmatrizen von der jeweiligen Standardbasis $\mathscr{E} = (e_1, e_2, \ldots)$ nach $\mathscr{B}$ bzw. von $\mathscr{C}$ in die Standardbasis. Dann gilt

$$A = \mathrm{M}_{\mathscr{E}}^{\mathscr{E}}(L_A)) = S\,\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(L_A)T = SBT.$$

Es gilt weiter $A^t = T^t B^t S^t$. Weil S und T, sowie ihre transponierten S^t und T^t invertierbar sind, folgt aus Proposition 9.5

$$\operatorname{rg}(A) = \operatorname{rg}(B) \qquad \text{und} \qquad \operatorname{rg}(A^t) = \operatorname{rg}(B^t).$$

Der Satz gilt damit für A genau dann, wenn der Satz für B gilt. Für B ist aber alles klar: die Spaltenräume sind $C(B) = \langle e_1, \ldots, e_r\rangle_K \subseteq K^n$ und $C(B^t) = \langle e_1, \ldots, e_r\rangle_K \subseteq K^m$, zwar als Unterräume in im allgemeinen verschiedenen Vektorräumen, aber dennoch mit

$$\operatorname{rg}(B) = \dim\langle e_1, \ldots, e_r\rangle_K = r = \dim\langle e_1, \ldots, e_r\rangle_K = \operatorname{rg}(B^t). \qquad \square$$

Bemerkung 9.8. Die Gleichheit von Zeilenrang und Spaltenrang ist im Sinne dieser Vorlesung eine nichttriviale Aussage und vielleicht auf den ersten Blick verblüffend, weil Zeilen und Spalten einer Matrix nur beinahe zufällig etwas miteinander zu tun haben. Nun, so zufällig ist der Zusammenhang nicht. Wir geben später in Bemerkung 9.52 eine Interpretation der Aussage im Kontext von linearen Gleichungssystemen.

Ein anderer Zugang zum Verständnis der Gleichheit von Zeilenrang und Spaltenrang gelingt über einen konzeptionellen Beweis. Der eben geführte Beweis von Satz 9.7 basiert auf der Auswahl spezieller Basen, die A in eine Normalform transformiert, aus der die Gleichheit der Ränge trivial abgelesen werden kann. Einen zweiten konzeptionellen Beweis von Satz 9.7 mittels Dualität liefern wir am Ende von Kapitel 10.2.

Korollar 9.9 (Rang von Untermatrizen)**.** *Entsteht eine Matrix A' aus der Matrix A durch entfernen eines Teils der Spalten und eines Teils der Zeilen, dann gilt* $\mathrm{rg}(A') \leq \mathrm{rg}(A)$.

Beweis. Das folgt sofort aus Satz 9.7 sowie der Definition von Zeilen- und Spaltenrang in Definition 9.1. Entfernt man Zeilen, so vergrößert sich der Zeilenrang nicht, und der Spaltenrang vergrößert sich nicht, wenn man Spalten entfernt. □

Proposition 9.10. *Sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Dann gilt:*
(1) $\mathrm{rg}(A) \leq \min\{n, m\}$.
(2) $\mathrm{rg}(A) = 0 \iff A = 0$.

Beweis. (1) Der Rang ist die Dimension eines Unterraums von K^n, also $\mathrm{rg}(A) \leq n$. Außerdem wird dieser Unterraum von den m Spalten erzeugt und besitzt nach dem Basisauswahlsatz eine Basis bestehend aus einem Teil dieser Spalten. Damit ist $\mathrm{rg}(A) \leq m$.

(2) Es ist $\mathrm{rg}(A) = 0 \iff$ der Raum, der von den Spalten erzeugt wird, ist $\{0\} \subseteq K^n$. Das passiert genau dann, wenn sämtliche Spalten 0 sind, d.h. $A = 0$. □

Definition 9.11. Eine quadratische Matrix $A \in \mathrm{M}_n(K)$ heißt **regulär**, wenn $\mathrm{rg}(A) = n$. Das ist der maximal mögliche Wert für den Rang einer $n \times n$-Matrix.

Satz 9.12 (Maximaler Rang)**.** *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gilt*

$$A \text{ ist invertierbar} \iff \mathrm{rg}(A) = n, \text{ d.h. } A \text{ ist regulär.}$$

Beweis. Nach Korollar 8.43 ist A invertierbar, genau dann wenn die Linksmultiplikation $L_A : K^n \to K^n$ ein Isomorphismus ist. Nach Satz 7.23 ist L_A ein Isomorphismus, genau dann wenn L_A surjektiv ist. Das Bild von L_A ist gerade der Spaltenraum $C(A) \subseteq K^n$. Daher ist L_A surjektiv, genau dann wenn $C(A) = K^n$ ist. Und das ist nach Korollar 5.33 äquivalent zu $\dim(C(A)) = \dim K^n$. Das übersetzt sich per Definition in $\mathrm{rg}(A) = n$. □

9.2. **Elementare Zeilenoperationen.** Unsere nächste Aufgabe besteht darin, den Rang einer Matrix algorithmisch zu bestimmen. Das gleiche Verfahren löst auch lineare Gleichungssysteme.

Definition 9.13. Seien K ein Körper und $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Unter einer **elementaren Zeilenoperation** versteht man eine der folgenden Transformationen:

(1) **Addition des Vielfachen** einer Zeile zu einer anderen: Seien $1 \leq \alpha, \beta \leq n$ mit $\alpha \neq \beta$ und $\lambda \in K$. Die Addition des λ-fachen der β-ten Zeile zur α-ten Zeile macht aus $A = (a_{ij})$ die Matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \vdots & \vdots \\ a_{\alpha 1} + \lambda a_{\beta 1} & \cdots & a_{\alpha m} + \lambda a_{\beta m} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \leftarrow \alpha.$$

(2) **Skalieren einer Zeile**: Seien $1 \leq \alpha \leq n$ und $\mu \in K^\times$. Das μ-fache der α-ten Zeile macht aus $A = (a_{ij})$ die Matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \vdots & \vdots \\ \mu a_{\alpha 1} & \cdots & \mu a_{\alpha m} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \leftarrow \alpha.$$

(3) **Vertauschen zweier Zeilen:** Seien $1 \leq \alpha, \beta \leq n$ mit $\alpha \neq \beta$. Die Vertauschung der α-ten mit der β-ten Zeile macht aus $A = (a_{ij})$ die Matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \vdots & \vdots \\ a_{\beta 1} & \cdots & a_{\beta m} \\ \vdots & \vdots & \vdots \\ a_{\alpha 1} & \cdots & a_{\alpha m} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \alpha \\ \\ \leftarrow \beta. \\ \\ \\ \end{matrix}$$

Die Abmessungen der Matrix bleiben bei elementaren Zeilenoperationen erhalten.

Definition 9.14. Seien K ein Körper und $n \in \mathbb{N}$. Wir nennen die folgenden Matrizen in $\mathrm{M}_n(K)$ **Elementarmatrizen**:

(1) Zu $\lambda \in K$ und $1 \leq \alpha, \beta \leq n$ und $\alpha \neq \beta$ definieren wir

$$E_{\alpha\beta}(\lambda) = \mathbf{1}_n + \lambda \cdot e_{\alpha\beta} = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \lambda & & \vdots \\ \vdots & \cdots & \ddots & \cdots & \vdots \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix} \begin{matrix} \\ \leftarrow \alpha \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} \\ \\ \\ \\ \uparrow\beta \end{matrix}$$

(2) Zu $\mu \in K^\times$ und $1 \leq \alpha \leq n$ definieren wir

$$E_{\alpha\alpha}(\mu) = \mathbf{1}_n + (\mu - 1) \cdot e_{\alpha\alpha} = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ \vdots & \cdots & \mu & \cdots & \vdots \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \alpha. \\ \\ \\ \end{matrix}$$
$$\begin{matrix} \\ \\ \uparrow\alpha \\ \\ \\ \end{matrix}$$

Die elementaren Zeilenoperation werden durch Multiplikation von links mit den Elementarmatrizen realisiert.

Satz 9.15 (Elementaroperationen 1+2)**.** *Seien K ein Körper und $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Dann gilt:*

(1) ***Addition des Vielfachen einer Zeile zu einer anderen****: Seien $1 \leq \alpha, \beta \leq n$ mit $\alpha \neq \beta$ und $\lambda \in K$. Die Addition des λ-fachen der β-ten Zeile zur α-ten Zeile macht*

aus $A = (a_{ij})$ die Matrix

$$E_{\alpha\beta}(\lambda)A.$$

(2) ***Skalieren einer Zeile:*** *Seien $1 \leq \alpha \leq n$ und $\mu \in K^\times$. Das μ-fache der α-ten Zeile macht aus $A = (a_{ij})$ die Matrix*

$$E_{\alpha\alpha}(\mu)A.$$

Beweis. Das folgt jeweils aus einer einfachen Matrixmultiplikation. □

Lemma 9.16. *Für Elementarmatrizen in $\mathrm{M}_n(K)$ gelten die folgenden Rechenregeln.*

(1) *Für $1 \leq \alpha, \beta \leq n$, $\alpha \neq \beta$ und $\lambda_1, \lambda_2 \in K$:*

$$E_{\alpha\beta}(\lambda_1)E_{\alpha\beta}(\lambda_2) = E_{\alpha\beta}(\lambda_1 + \lambda_2).$$

(2) *Für $1 \leq \alpha \leq n$ und $\mu_1, \mu_2 \in K^\times$:*

$$E_{\alpha\alpha}(\mu_1)E_{\alpha\alpha}(\mu_2) = E_{\alpha\alpha}(\mu_1\mu_2).$$

Beweis. Am einfachsten betrachtet man die zugehörigen linearen Abbildungen $K^n \to K^n$ durch Matrixmultiplikation. Dort verfolgt man die Vektoren e_j der Standardbasis mittels Satz 9.15. Alle e_j mit $j \neq \beta$ bleiben e_j. Der Vektor e_β wird bei (1) abgebildet zu

$$e_\beta \mapsto e_\beta + \lambda_2 e_\alpha \mapsto e_\beta + (\lambda_1 + \lambda_2)e_\alpha.$$

Bei (2) erhalten wir

$$e_\alpha \mapsto \mu_2 e_\alpha \mapsto \mu_1\mu_2 e_\alpha,$$

und alle e_j mit $j \neq \alpha$ bleiben e_j.

Alternativ kann man die Formeln auch einfach nachrechnen. Die nötigen Rechnungen sind analog zu Beispiel 8.20 (1), (2). Der dort behandelte 2×2-Fall spielt sich hier im Unterraum $\langle e_\alpha, e_\beta\rangle_K$ ab bei (1), und bei (2) sogar nur im eindimensionalen Unterraum $\langle e_\alpha\rangle_K$. □

Lemma 9.17. *Elementarmatrizen in $\mathrm{M}_n(K)$ sind invertierbar.*

(1) *Für $1 \leq \alpha, \beta \leq n$, $\alpha \neq \beta$ und $\lambda \in K$:*

$$E_{\alpha\beta}(\lambda)^{-1} = E_{\alpha\beta}(-\lambda).$$

(2) *Für $1 \leq \alpha \leq n$ und $\mu \in K^\times$:*

$$E_{\alpha\alpha}(\mu)^{-1} = E_{\alpha\alpha}(\mu^{-1}).$$

Beweis. Das folgt sofort aus Lemma 9.16 mittels $E_{\alpha\beta}(0) = \mathbf{1}_n$ für $\alpha \neq \beta$ und $E_{\alpha\alpha}(1) = \mathbf{1}_n$. □

Definition 9.18. Sei K ein Körper und $n \in \mathbb{N}$. Seien $1 \leq \alpha, \beta \leq n$ mit $\alpha \neq \beta$. Die **Permutationsmatrix** der Vertauschung von α und β ist definiert als die Matrix

$$P_{\alpha\beta} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \alpha \\ \\ \leftarrow \beta. \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & \uparrow\alpha & & \uparrow\beta & & \end{matrix}$$

Die Matrix $P_{\alpha\beta}$ weicht nur in den vier Positionen (i, j) mit $\{i, j\} = \{\alpha, \beta\}$ in der angezeigten Weise von der Einheitsmatrix ab.

Lemma 9.19. *Sei K ein Körper und $n \in \mathbb{N}$. Seien $1 \leq \alpha, \beta \leq n$ mit $\alpha \neq \beta$. Es gilt*

$$P_{\alpha\beta} = E_{\beta\beta}(-1)E_{\alpha\beta}(1)E_{\beta\alpha}(-1)E_{\alpha\beta}(1) \in \mathrm{M}_n(K).$$

Beweis. Wir nutzen aus, daß für Matrizen $A = B$ gilt genau dann, wenn $Av = Bv$ für alle Vektoren v gilt. Weil sich in den Zeilen $i \neq \alpha, \beta$ nichts tut, notieren wir nur die Zeilen $i = \alpha, \beta$ (wenn man so möchte, behandeln wir den Fall $n = 2$):

$$E_{\beta\beta}(-1)E_{\alpha\beta}(1)E_{\beta\alpha}(-1)E_{\alpha\beta}(1)\begin{pmatrix} x_\alpha \\ x_\beta \\ \vdots \end{pmatrix} = E_{\beta\beta}(-1)E_{\alpha\beta}(1)E_{\beta\alpha}(-1)\begin{pmatrix} x_\alpha + x_\beta \\ x_\beta \\ \vdots \end{pmatrix}$$

$$= E_{\beta\beta}(-1)E_{\alpha\beta}(1)\begin{pmatrix} x_\alpha + x_\beta \\ -x_\alpha \\ \vdots \end{pmatrix} = E_{\beta\beta}(-1)\begin{pmatrix} x_\beta \\ -x_\alpha \\ \vdots \end{pmatrix} = \begin{pmatrix} x_\beta \\ x_\alpha \\ \vdots \end{pmatrix} = P_{\alpha\beta}\begin{pmatrix} x_\alpha \\ x_\beta \\ \vdots \end{pmatrix}. \qquad \square$$

Satz 9.20 (Elementaroperation 3)**.** *Seien K ein Körper und $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Seien $1 \leq \alpha, \beta \leq n$ mit $\alpha \neq \beta$. Die Vertauschung der α-ten mit der β-ten Zeile macht aus $A = (a_{ij})$ die Matrix*

$$P_{\alpha\beta}A.$$

Beweis. Das folgt aus einer einfachen Matrixmultiplikation, die bereits im Beweis von Lemma 9.19 enthalten ist. $\square$

Korollar 9.21. *Elementare Zeilenoperationen ändern den Rang nicht.*

Beweis. Elementare Zeilenoperationen sind Linksmultiplikationen mit (Produkten von) Elementarmatrizen, und diese sind invertierbar, Lemma 9.17. Nach Proposition 9.5 ändert dies den Rang nicht. $\square$

9.3. **Zeilenstufenform mittels Gaußschem Eliminationsverfahren.**

Definition 9.22. Eine quadratische Matrix $A = (a_{ij}) \in \mathrm{M}_n(K)$ ist eine **obere Dreiecksmatrix** (bzw. **untere Dreiecksmatrix**), wenn für alle $i > j$ (bzw. alle $i < j$ gilt $a_{ij} = 0$.

obere Dreiecksmatrix: $\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * \end{pmatrix}$, untere Dreiecksmatrix: $\begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & * \end{pmatrix}$.

Die Einträge mit Zeilenindex = Spaltenindex heißen **Diagonaleinträge**.

Proposition 9.23. *Sei*

$$A = (a_{ij}) = \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \alpha_r \end{pmatrix} \in \mathrm{M}_n(K)$$

eine obere Dreiecksmatrix mit von 0 verschiedenen Diagonaleinträgen

$$\alpha_i = a_{ii} \neq 0 \text{ für } i = 1, \ldots, n.$$

Dann hat A maximalen Rang: $\operatorname{rg}(A) = n$. *Insbesondere ist A invertierbar.*

Beweis. Wir müssen zeigen, daß die Spalten von A linear unabhängig sind. Sei $A = [v_1, \ldots, v_n]$ in Spaltenschreibweise. Dann gilt für $x = (x_1 \ldots x_n)^t \in K^n$

$$Ax = x_1 v_1 + \ldots + x_n v_n.$$

Eine Linearkombination der Spalten mit Koeffizienten $x_1, \ldots, x_n \in K$ ergibt daher 0 genau dann, wenn

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

Dieses Gleichungssystem für $x_1, \ldots, x_n \in K$ läßt sich rückwärts beginnend mit x_n und dann fallend im Index eindeutig lösen: die Gleichung der n-ten Zeile ist:

$$\alpha_n x_n = 0 \Longrightarrow x_n = 0.$$

Und wenn bereits $x_{s+1}, \ldots, x_n = 0$, dann folgt wegen $\alpha_s \neq 0$ aus der Gleichung der s-ten Zeile auch

$$0 = \alpha_s x_s + \sum_{j=s+1}^{n} a_{sj} x_j = \alpha_s x_s \Longrightarrow x_s = 0.$$

Wir erhalten nur die eine Lösung $x_i = 0$ für $i = 1, \ldots, n$. Daher sind die Spalten von A linear unabhängig.

Daß A invertierbar ist, folgt nun aus Satz 9.12. □

Definition 9.24. Eine Matrix $R = (r_{ij}) \in \mathrm{M}_{n\times m}(K)$ hat **Zeilenstufenform**, wenn es eine ganze Zahl r mit

$$0 \leq r \leq n$$

und eine streng monoton wachsende Folge natürlicher Zahlen ρ_i mit

$$1 \leq \rho_1 < \rho_2 < \ldots < \rho_r \leq m$$

gibt mit den folgenden Eigenschaften:

(i) Die **Pivoteinträge** $\alpha_i = r_{i\rho_i}$ an den Positionen (i, ρ_i) sind für alle $i = 1, \ldots, r$ von Null verschieden:
$$\alpha_i \neq 0.$$

(ii) Für $i = 1, \ldots, r$ sind in der i-ten Zeile links vom Pivoteintrag sämtliche Einträge 0: für alle $1 \leq j < \rho_i$ ist
$$r_{ij} = 0.$$

(iii) Die Zeilen mit Zeilenindizes $i = r+1, \ldots, n$ enthalten nur den Eintrag 0 (Nullzeilen).

Graphisch bedeutet das für R die folgende Gestalt mit $\alpha_1, \ldots, \alpha_r \in K^\times$.

$$R = \begin{pmatrix}
0 & \cdots & 0 & \alpha_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & \alpha_2 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & & & & & \cdots & 0 & \ddots & & & & & & * \\
0 & \cdots & & & & & \cdots & 0 & \alpha_r & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & & & & & & \cdots & 0 & 0 & \cdots & & \cdots & 0 \\
\vdots & & & & & & & & \vdots & \vdots & & & & \vdots \\
0 & \cdots & & & & & & \cdots & 0 & 0 & \cdots & & \cdots & 0
\end{pmatrix} \tag{9.1}$$

$\uparrow\rho_1 \qquad \uparrow\rho_2 \qquad \uparrow\rho_r$

Die Nullspalten links und die Nullzeilen unten, sowie die unbekannten $*$-Einträge rechts oben können, müssen aber nicht auftreten.

Die **Pivotspalten** der Matrix in Zeilenstufenform sind die Spalten der Indizes $\rho_1, \ldots, \rho_r$. Dies sind die Spalten mit den Pivoteinträgen $\alpha_1, \ldots, \alpha_r$.

Beispiel 9.25. Eine obere Dreiecksmatrix in $\mathrm{M}_n(K)$ hat Zeilenstufenform, wenn die Diagonaleinträge ungleich 0 sind. In diesem Fall ist $r = n$ sowie $\rho_i = i$: jede Spalte eine Pivotspalte der Matrix. Ein konkretes Beispiel ist die Elementarmatrix $E_{\alpha\beta}(\lambda)$ für $\alpha < \beta$.

Wenn nicht alle Diagonaleinträge der oberen Dreiecksmatrix von 0 verschieden sind, dann kann die Matrix Zeilenstufenform haben, muß aber nicht: zum Beispiel hat die obere Dreiecksmatrix

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

keine Zeilenstufenform.

Satz 9.26. *Sei $R \in \mathrm{M}_{n\times m}(K)$ eine Matrix in Zeilenstufenform. Dann hat R den Rang*

$$\mathrm{rg}(R) = \textit{Anzahl der Pivotspalten von } R,$$

d.h. in der Form (9.1) ist $\mathrm{rg}(R) = r$. Eine Basis für den Spaltenraum $C(R)$ ist durch die Pivotspalten gegeben.

Beweis. Weil alle Spalten von R in $\langle e_1, \ldots, e_r\rangle_K \subseteq K^n$ enthalten sind, haben wir

$$\mathrm{rg}(R) \leq \dim\langle e_1, \ldots, e_r\rangle_K = r. \tag{9.2}$$

Streichen wir alle anderen Spalten und die Nullzeilen, so entsteht mit

$$R' = \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \alpha_r \end{pmatrix}$$

eine obere Dreiecksmatrix mit von 0 verschiedenen Diagonaleinträgen. Proposition 9.23 zeigt daher

$$\mathrm{rg}(R') = r.$$

Mit Korollar 9.9 und (9.2) folgt

$$r = \mathrm{rg}(R') \leq \mathrm{rg}(R) \leq r,$$

also die Behauptung $\mathrm{rg}(R) = r$.

Wir zeigen nun, daß die Pivotspalten von R linear unabhängig sind. Seien $v_1, \ldots, v_r$ die Pivotspalten von R, und sei $R' = [v'_1, \ldots, v'_r]$ in Spaltenschreibweise. Dann entstehen die Spalten v'_j von R' aus den Pivotspalten v_j durch Weglassen der 0-Zeilen mit Index $i = r+1, \ldots, n$. Daher folgt für $x_1, \ldots, x_r \in K$:

$$x_1v_1 + \ldots + x_rv_r = 0 \iff x_1v'_1 + \ldots + x_rv'_r = 0, \tag{9.3}$$

denn die ausgelassenen Zeilen sind sowieso 0. Nun zeigt Proposition 9.23, genauer, daß die Spalten von R' linear unabhängig sind. Daraus folgt nun mit (9.3), daß auch die Pivotspalten von R linear unabhängig sind.

Weil die Pivotspalten linear unabhängig und von der richtigen Anzahl sind (nämlich $\mathrm{rg}(R)$-viele), handelt es sich bei den Pivotspalten nach Satz 5.31 um eine Basis von $C(R)$. □

Algorithmus 9.27 (Gaußsches Eliminationsverfahren)**.** Das **Gaußsche Eliminationsverfahren** bietet einen Algorithmus, um den Rang einer Matrix ablesen und lineare Gleichungssysteme systematisch lösen zu können.

Wir starten mit einer Matrix $A = A_0 \in \mathrm{M}_{n\times m}(K)$. Im s-ten Schritt produziert der Algorithmus eine Matrix $A_s \in \mathrm{M}_{n\times m}(K)$ der Form

$$A_s = \begin{pmatrix}
0 & \cdots & 0 & \alpha_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & \alpha_2 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & & & & & \cdots & 0 & \ddots & * & & & & & * \\
0 & \cdots & & & & & \cdots & 0 & \alpha_s & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & & & & & & \cdots & 0 & \multicolumn{5}{c}{\multirow{3}{*}{B_s}} \\
\vdots & & & & & & & & \vdots & & & & & \\
0 & \cdots & & & & & & \cdots & 0 & & & & &
\end{pmatrix}$$

mit $\alpha_1, \ldots, \alpha_s \in K^\times$ und einer Matrix B_s mit $n - s$ Zeilen und m_s-vielen Spalten für gewisses $m_s \in \mathbb{N}_0$. Zu Anfang ist $s = 0$ und $B_0 = A_0$.

Der Algorithmus arbeitet nun pro Durchgang die folgenden Schritte ab:

I Sei $B_s = [b_1, \ldots, b_{m_s}]$. Wenn $B_s = 0$ die Nullmatrix ist (oder wenn keine Zeilen oder Spalten mehr übrig sind), dann stoppt der Algorithmus.

Wenn $B_s \neq 0$, dann betrachten wir die Spalte mit dem kleinsten Spaltenindex von B_s, welche als Spalte von 0 verschieden ist. In dieser Spalte wählen wir eine Zeile mit von 0 verschiedenem Eintrag. Dies sei in der Matrix A_s der Eintrag an der Position (i_s, j_s), den wir nun mit α_{s+1} bezeichnen. Per Wahl ist $\alpha_{s+1} \neq 0$.

Sind mehrere Einträge von 0 verschieden, so kann man manchmal geschickt wählen und so den Rechenaufwand minimieren.

II Wir tauschen nun die i_s-te Zeile von A_s mit der $s+1$-ten Zeile.

$$A_s \rightsquigarrow A'_s = \begin{pmatrix}
0 & \cdots & 0 & \alpha_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & \alpha_2 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & & & & & \cdots & 0 & \ddots & & & & & & * \\
0 & \cdots & & & & & \cdots & 0 & \alpha_s & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & & & & & & & \cdots & 0 & \alpha_{s+1} & * & \cdots & * \\
\vdots & & & & & & & & & \vdots & * & * & \cdots & * \\
\vdots & & & & & & & & & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & & & & & & & \cdots & 0 & * & * & \cdots & *
\end{pmatrix}$$

III Nun addieren wir ein jeweils geeignetes Vielfaches der $s+1$-ten Zeile von $A'_s = (a'_{ij})$ zu den Zeilen mit Index $s + 2 \leq i \leq n$ hinzu. Das Vielfache zur i-ten Zeile ist

$$\lambda_{i,s} = -\frac{a'_{ij_s}}{\alpha_{s+1}}.$$

Dadurch entstehen unter dem Eintrag α_{s+1} an der Position $(s+1, j_s)$ lauter Nullen, denn

$$a'_{ij_s} + \lambda_{i,s} a'_{s+1\ j_s} = a'_{ij_s} - \frac{a'_{ij_s}}{\alpha_{s+1}} \cdot \alpha_{s+1} = 0,$$

also:

$$A'_s \rightsquigarrow A_{s+1} = \begin{pmatrix} 0 & \cdots & 0 & \alpha_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \alpha_2 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & & & & & \cdots & 0 & \ddots & & & & & & * \\ 0 & \cdots & & & & & \cdots & 0 & \alpha_s & * & \cdots & \cdots & \cdots & * \\ 0 & \cdots & & & & & & & \cdots & 0 & \alpha_{s+1} & * & \cdots & * \\ \vdots & & & & & & & & & \vdots & 0 & * & \cdots & * \\ \vdots & & & & & & & & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & & & & & & & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}$$

Danach geht es weiter mit Schritt $\boxed{\text{I}}$.

Nach Satz 9.15 und Satz 9.20 sind elementare Zeilenoperationen Linksmultiplikation mit Produkten von Elementarmatrizen. Jeder Schritt im Gauß-Eliminationsverfahren entspricht der Multiplikation von links mit Elementarmatrizen: $A_{s+1} = E_s A_s$ mit in der Notation von oben

$$E_s := E_{ns}(\lambda_{n,s}) E_{(n-1)s}(\lambda_{(n-1),s}) \dots E_{(s+1)s}(\lambda_{s+1,s}) P_{s+1,i_s}.$$

Hierbei ist zu beachten, daß die Permutationsmatrix P_{s+1,i_s} nach Lemma 9.19 ebenfalls ein Produkt von Elementarmatrizen ist.

Theorem 9.28 (Zeilenstufenform aus dem **Gaußschen Eliminationsverfahren**). *Seien K ein Körper und $n, m \in \mathbb{N}_0$. Sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix.*

(1) Das Gaußsche Eliminationsverfahren endet nach höchstens n Durchgängen, und zwar mit einer Matrix R in Zeilenstufenform.

(2) Sei S die Matrix, die aus $\mathbf{1}_n \in \mathrm{M}_n(K)$ entsteht, indem man dieselben elementaren Zeilenoperationen in derselben Reihenfolge anwendet, die R aus A im Gaußschen Eliminationsverfahren ergeben haben. Dann ist
   (i) $R = SA$,
   (ii) S ist ein Produkt von Elementarmatrizen, und
   (iii) $S \in \mathrm{GL}_n(K)$.

(3) Es gilt:
$$\mathrm{rg}(A) = \textit{Anzahl der Pivotspalten von } R.$$

(4) Die Spalten von A mit den Spaltenindizes der Pivotspalten von R sind eine Basis des Spaltenraums $C(A)$.

Beweis. (1) Das ist klar. Im Schritt A_s hat man in den ersten s Zeilen bereits Zeilenstufenform erreicht. Das Verfahren bricht erst ab, wenn in den weiteren Spalten nur 0 steht oder keine weitere Zeile mehr vorhanden ist. Die Anzahl der Schritte ist durch die Anzahl n der Zeilen begrenzt.

(2) In der Notation von Algorithmus 9.27 gilt $A_{s+1} = E_s A_s$ mit einem Produkt von Elementarmatrizen E_s. Sei F das Produkt

$$F = E_{r-1} \dots E_2 E_1 E_0.$$

Dann gilt (per Induktion nach s)

$$FA = E_{r-1} \dots E_2 E_1 E_0 A_0 = E_{r-1} \dots E_{s+1} E_s A_s = E_{r-1} \dots E_{s+1} A_{s+1} = \dots = A_r = R.$$

Dieselben Operationen angewandt auf $\mathbf{1}_n$ liefert nach Definition von S

$$S = F\mathbf{1}_n = F.$$

Es folgt $SA = R$. Elementarmatrizen sind nach Lemma 9.17 invertierbar. Also ist S als Produkt von invertierbaren Matrizen auch invertierbar, d.h. $S \in \mathrm{GL}_n(K)$.

(3) Nach Proposition 9.5 und Satz 9.26 ist wegen $S \in \mathrm{GL}_n(K)$

$$\mathrm{rg}(A) = \mathrm{rg}(SA) = \mathrm{rg}(R) = r.$$

(4) Satz 9.26 liefert genauer, daß die Pivotspalten von R eine Basis des Spaltenraums $C(R)$ sind. Sei $A = [v_1, \ldots, v_m]$ mit Spalten $v_j \in K^n$. Dann ist $R = [Sv_1, \ldots, Sv_m]$, also

$$C(R) = L_S(C(A)).$$

Linksmultiplikation mit S ist ein Isomorphismus und das Urbild der Pivotspalten sind gerade die Spalten von A mit den Pivotspaltenindizes. Diese sind somit eine Basis von $C(A)$. □

Korollar 9.29. *Die Anzahl der Pivotspalten, die das Gaußsche Eliminationsverfahren liefert, ist unabhängig von den auftretenden Wahlen.*

Beweis. In jedem Fall ist diese Anzahl gleich dem Rang der anfänglichen Matrix. □

Algorithmus 9.30 (Gauß–Jordan–Verfahren). Sei $A = [v_1, \ldots, v_m]$. Man bestimmt S aus Theorem 9.28 am besten dadurch, daß man mit der erweiterten Matrix (geschrieben als Blockmatrix)

$$\begin{pmatrix} A & \mathbf{1}_n \end{pmatrix} = [v_1, \ldots, v_m, e_1, \ldots, e_n]$$

arbeitet und abbricht, sobald der Teil A in Zeilenstufenform gebracht ist. Dann hat man, in der Notation von Theorem 9.28 mit $F = S$ von links multipliziert, so daß

$$S \cdot \begin{pmatrix} A & \mathbf{1}_n \end{pmatrix} = \begin{pmatrix} SA & S \end{pmatrix} = \begin{pmatrix} R & S \end{pmatrix}.$$

Man kann die Zeilenstufenform weiter verbessern zur reduzierten Zeilenstufenform, aus der man dann eine Basis des Kerns einer Matrix ablesen kann.

Definition 9.31. Man sagt, daß eine Matrix R **reduzierte Zeilenstufenform** hat, wenn R Zeilenstufenform hat mit Pivotspaltenindizes $\rho_1 < \ldots < \rho_r$ und die j-te Pivotspalte für alle $1 \leq j \leq r$ (die Spalte mit Index ρ_j) gerade der j-te Standardbasisvektor e_j ist. Das bedeutet, daß die Einträge in den Stufen 1 und alle anderen Einträge über dem Stufeneintrag (so wie auch unterhalb) gleich 0 sind.

$$R = \begin{pmatrix}
0 & \cdots & 0 & 1 & * & \cdots & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & * \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1 & * & 0 & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & & & & & \cdots & 0 & \ddots & \vdots & & & & & * \\
0 & \cdots & & & & & \cdots & 0 & 1 & * & \cdots & \cdots & \cdots & * \\
0 & \cdots & & & & & & \cdots & 0 & 0 & \cdots & & \cdots & 0 \\
\vdots & & & & & & & & \vdots & \vdots & & & & \vdots \\
0 & \cdots & & & & & & \cdots & 0 & 0 & \cdots & & \cdots & 0
\end{pmatrix} \qquad (9.4)$$

$\uparrow\rho_1 \qquad \uparrow\rho_2 \qquad \uparrow\rho_r$

Algorithmus 9.32 (Reduzierte Zeilenstufenform)**.** Man kann das Gauß-sche Eliminationsverfahren leicht so erweitern, daß eine reduzierte Zeilenstufenform berechnet wird.

- Nachdem in der s-ten Pivotspalte mit Spaltenindex ρ_s der von 0 verschiedene Eintrag α_s in die s-te Zeile getauscht wurde, skaliert man mittels $E_{ss}(\alpha_s^{-1})$ zunächst die s-te Zeile und damit den Eintrag α_s in der s-ten Stufe auf 1.
- Insbesondere muß man in den folgenden elementaren Zeilenoperationen zur Erzeugung der 0-Einträge der s-ten Pivotspalte nicht mehr durch den Pivoteintrag α_s teilen, denn durch den Skalierungsschritt wird $\alpha_s = 1$.
- Anschließend nutzt man weitere Zeilenoperationen $E_{is}(-a_{i\rho_s})$, für $i \neq s$ (und nicht nur für $i > s$ wie im Eliminationsverfahren zur Bestimmung der allgemeinen Zeilenstufenform), um die zusätzlichen Nullen in Zeile i an der Stelle (i, ρ_s) zu erzeugen.

Verfahren 9.33. Aus einer reduzierten Zeilenstufenform $R = (r_{ij})$ für die Matrix A kann man wie folgt eine Basis für den Nullraum $\ker(A)$ ablesen. Seien $\rho_1 < \ldots < \rho_r$ die Pivotspaltenindizes und

$$J = \{1, \ldots, n\} \setminus \{\rho_1, \ldots, \rho_r\}$$

die Menge der Nichtpivotspaltenindizes. Zu $j \in J$ lesen wir aus dem j-ten Spaltenvektor von R den Vektor $w_j \in K^n$ definiert durch

$$(w_j)_i := i\text{-te Zeile von } w_j := \begin{cases} r_{sj} & i = \rho_s, \\ -1 & i = j, \\ 0 & i \in J,\ i \neq j. \end{cases}$$

ab. Geometrisch formuliert notiert man die Einträge der j-ten Spalte von R in die Zeilen mit dem entsprechenden Pivotspaltenindex, fügt in der j-ten Zeile eine -1 ein und ergänzt in allen Zeilen mit Index $k \in J \setminus \{j\}$ eine Null.

Satz 9.34. *Die w_j mit $j \in J$ aus Verfahren 9.33 bilden eine Basis von* $\ker(A)$.

Beweis. Wegen $R = SA$ und $S \in \mathrm{GL}_n(K)$ gilt für alle $x \in K^n$:

$$Ax = 0 \iff SAx = 0 \iff Rx = 0,$$

ergo $\ker(A) = \ker(R)$. Daher müssen wir zeigen, daß die w_j für $j \in J$ eine Basis von $\ker(R)$ sind.

Wir zeigen zunächst $w_j \in \ker(R)$. Die i-te Zeile von Rw_j ist

$$\begin{aligned}(Rw_j)_i &= \sum_{k=1}^{n} r_{ik} \cdot (w_j)_k = r_{ij} \cdot (w_j)_j + \sum_{k \notin J} r_{ik} \cdot (w_j)_k = r_{ij} \cdot (w_j)_j + \sum_{s=1}^{r} r_{i\rho_s} \cdot (w_j)_{\rho_s} \\ &= -r_{ij} + \sum_{s=1}^{r} r_{i\rho_s} r_{sj} = -r_{ij} + \sum_{s=1}^{r} \delta_{is} r_{sj} = -r_{ij} + r_{ij} = 0.\end{aligned}$$

Dies zeigt $Rw_j = 0$, also $w_j \in \ker(R)$ für alle $j \in J$.

Sei $W = [w_j\ ;\ j \in J]$ die Matrix mit Spalten w_j für $j \in J$. Die Matrix W hat $|J| = n - r$ Spalten. Streicht man in W die Zeilen mit Zeilenindex $i \in \{\rho_1, \ldots, \rho_r\}$ aus der Menge der Pivotspaltenindizes, dann entsteht mit $-\mathbf{1}_{n-r}$, das Negative zur Einheitsmatix mit der Abmessung $|J| = n - r$. Daher gilt nach Korollar 9.9

$$n - r = \mathrm{rg}(-\mathbf{1}_{n-r}) \leq \mathrm{rg}(W) \leq \text{ Anzahl der Spalten von } W \ = n - r.$$

Daher ist $\mathrm{rg}(W) = n - r$ und die Spalten von W sind linear unabhängig. Aus der Kern-Bild-Dimensionsformel, Satz 7.21, folgt

$$\dim\big(\ker(R)\big) = n - \mathrm{rg}(R) = n - r.$$

Damit sind die Spalten von W ein linear unabhängiges System von $\dim\big(\ker(R)\big)$-vielen Vektoren aus $\ker(R)$, und demnach eine Basis von $\ker(R)$. □

Beispiel 9.35. Die folgende Matrix hat bereits reduzierte Zeilenstufenform:

$$R = \begin{pmatrix} 1 & 2 & 0 & 0 & 3 & 11 \\ 0 & 0 & 1 & 0 & 5 & 13 \\ 0 & 0 & 0 & 1 & 7 & 17 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Pivotspalten haben die Indizes 1, 3 und 4. Die Nichtpivotspaltenindizes sind $J = \{2, 5, 6\}$. Die Basis (w_2, w_5, w_6) von $\ker(R)$ lesen wir ab als

$$w_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w_5 = \begin{pmatrix} 3 \\ 0 \\ 5 \\ 7 \\ -1 \\ 0 \end{pmatrix}, \quad w_6 = \begin{pmatrix} 11 \\ 0 \\ 13 \\ 17 \\ 0 \\ -1 \end{pmatrix}.$$

Algorithmus 9.36 (Inverse Matrix)**.** Sei $A \in \mathrm{M}_n(K)$ eine invertierbare quadratische Matrix. Nach Satz 9.12 erkennt man das daran, daß A maximalen Rang hat:

$$\mathrm{rg}(A) = n.$$

Die zu A inverse Matrix A^{-1} bestimmt man durch Kombination des erweiterten Gauß-sche Eliminationsverfahren Algorithmus 9.32 mit dem Gauß–Jordan–Verfahren Algorithmus 9.30. Wir erweitern $A = [v_1, \ldots, v_n]$ zu

$$\begin{pmatrix} A & \mathbf{1}_n \end{pmatrix} = [v_1, \ldots, v_n, e_1, \ldots, e_n]$$

und lassen das erweiterten Gauß-sche Eliminationsverfahren laufen. Dies multipliziert von links mit einer invertierbaren Matrix S und bricht ab mit einer Matrix, die im vorderen Teil die reduzierte Zeilenstufenform von A ist. Nach Voraussetzung $\mathrm{rg}(A) = n$ ist die Einheitsmatrix $\mathbf{1}_n$ die reduzierte Zeilenstufenform von A. Also erreicht der Algorithmus bei Abbruch die Matrix

$$S \cdot \begin{pmatrix} A & \mathbf{1}_n \end{pmatrix} = \begin{pmatrix} SA & S \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n & S \end{pmatrix}.$$

Also gilt $SA = \mathbf{1}_n$. Weil A invertierbar ist, folgt (das ist Lemma 3.29 in $\mathrm{GL}_n(K)$)

$$A^{-1} = \mathbf{1}_n A^{-1} = (SA)A^{-1} = S(AA^{-1}) = S.$$

Die gesuchte zu A inverse Matrix ist demnach S und befindet sich bei Abbruch des Algorithmus im hinteren $n \times n$-Block.

Korollar 9.37. *Jede invertierbare Matrix A ist ein Produkt von Elementarmatrizen.*

Beweis. Die Matrix S in Algorithmus 9.36 ist ein Produkt aus Elementarmatrizen. Also ist $A^{-1} = S$ ein Produkt von Elementarmatrizen. Wenn wir dieses Argument für $A^{-1} \in \mathrm{GL}_n(K)$ verwenden, finden wir daß $A = (A^{-1})^{-1}$ ein Produkt von Elementarmatrizen ist. □

9.4. **Wie entscheidet man ...** Das Gaußsche Eliminationsverfahren angewandt auf eine Matrix

$$A = [v_1, \ldots, v_m] \in \mathrm{M}_{n \times m}(K)$$

erlaubt es, die folgenden Fragen algorithmisch zu bewältigen (vorausgesetzt, man kann im Körper K algorithmisch rechnen):

- Der Rang $\mathrm{rg}(A)$ von A ist die Anzahl der Pivotspalten der zugehörigen Zeilenstufenform R.
- Eine Basis des Spaltenraums $C(A) = \langle v_1, \ldots, v_m \rangle_K$ ist gegeben durch die Spalten v_j, die zu den Pivotspaltenindizes j von R gehören.
- Die Dimension des Spaltenraums $C(A)$ ist $\mathrm{rg}(A) = \dim(C(A))$.

Verfahren 9.38. Wie findet man eine Basis für die lineare Hülle von Vektoren $v_1, \ldots, v_m$?

Man bildet die Matrix $A = [v_1, \ldots, v_m]$ und bestimmt eine Basis des Spaltenraums.

Verfahren 9.39. Sei $v_1, \ldots, v_m$ ein Erzeugendensystem für einen Unterraum von K^n. Der Basisauswahlsatz, Satz 5.22, behauptet die Existenz einer Basis bestehend (als Tupel) aus einer Teilmenge von $v_1, \ldots, v_m$. Wie findet man eine solche Teilmenge.

Man bildet die Matrix $A = [v_1, \ldots, v_m]$ und bestimmt eine Basis des Spaltenraums.

Verfahren 9.40. Erzeugen $v_1, \ldots, v_m \in K^n$ den Vektorraum K^n?

Wenn $m < n$, dann sicher nicht, und wir fangen auch gar nicht erst mit Rechnen an! Ansonsten bilden wir die Matrix $A = [v_1, \ldots, v_m]$ und nutzen wir, daß gilt:

$$\operatorname{rg}(A) = n \iff \dim(C(A)) = \dim(K^n) \iff v_1, \ldots, v_m \text{ erzeugen } K^n.$$

Verfahren 9.41. Sind die Vektoren $v_1, \ldots, v_m \in K^n$ linear unabhängig?

Wenn $m > n$, dann sicher nicht, und wir fangen auch gar nicht erst mit Rechnen an! Ansonsten bilden wir die Matrix $A = [v_1, \ldots, v_m]$ und nutzen wir, daß gilt:

$$\operatorname{rg}(A) = m \iff (v_1, \ldots, v_m) \text{ sind linear unabhängig.}$$

Verfahren 9.42. Liegt $v \in K^n$ in der linearen Hülle von $v_1, \ldots, v_m \in K^n$?

Wir betrachten die erweiterte Matrix $B = [A, v] = [v_1, \ldots, v_m, v]$. Da

$$\langle v_1, \ldots, v_m \rangle_K \subseteq \langle v_1, \ldots, v_m, v \rangle_K$$

gilt nach Korollar 5.34

$$v \in \langle v_1, \ldots, v_m \rangle_K \iff \langle v_1, \ldots, v_m \rangle_K = \langle v_1, \ldots, v_m, v \rangle_K \iff \operatorname{rg}(A) = \operatorname{rg}(B).$$

Verfahren 9.43. Angenommen die Vektoren $v_1, \ldots, v_s \in K^n$ sind linear unabhängig. Wie ergänzt man zu einer Basis? Der Basisergänzungssatz 5.23 garantiert ja, daß dies möglich ist.

Man braucht ein Erzeugendensystem von K^n, dazu nehmen wir die Standardbasis. Sei $A = [v_1, \ldots, v_s]$ und $B = [v_1, \ldots, v_s, e_1, \ldots, e_n]$. Das Gaußsche Eliminationsverfahren für A und B unterscheidet sich einzig dadurch, daß die Matrix von B mehr Spalten hat. In den Spalten mit Index $1, \ldots, s$ passiert genau das gleiche. Weil die $v_1, \ldots, v_s$ linear unabhängig sind, ist $\operatorname{rg}(A) = s$. Daher sind alle Spalten aus A Pivotspalten auch in B. Die vom Algorithmus ausgewählte Basis der Pivotspalten von B enthält daher die Vektoren $v_1, \ldots, v_s$. Das ist die gesuchte Basisergänzung.

Verfahren 9.44. Stellt die Matrix $A \in \mathrm{M}_{n\times m}(K)$ eine injektive, eine surjektive oder eine bijektive lineare Abbildung dar?

Die Abbildung $L_A : K^m \to K^n$ ist

- injektiv: wenn $\ker(L_A) = 0$, also nach der Kern/Bild-Dimensionsformel, wenn $\dim(\operatorname{im}(L_A)) = m$, also wenn
$$\operatorname{rg}(A) = m.$$
- surjektiv: wenn $\operatorname{im}(L_A) = K^n$, also wenn
$$\operatorname{rg}(A) = n.$$
- bijektiv: wenn sowohl injektiv als auch surjektiv, also wenn
$$m = \operatorname{rg}(A) = n,$$
siehe dazu auch Satz 9.12.

9.5. **Lineare Gleichungssysteme als Matrixgleichungen.** In Definition 4.15 haben wir bereits definiert, was ein homogenes lineares Gleichungssystem ist.

Definition 9.45. Seien K ein Körper und $n, m \in \mathbb{N}_0$. Ein **inhomogenes lineares Gleichungssystem** in m-vielen **Variablen** $X_1, \ldots, X_m$ mit n-vielen Gleichungen und Koeffizienten $a_{ij} \in K$ für $i = 1, \ldots, n$ und $j = 1, \ldots, m$ und **Inhomogenitäten** $b_1, \ldots, b_n \in K$ ist

ein System $\mathscr{S}$ von Gleichungen der Form

$$\mathscr{S} = \left\{ \begin{array}{ccccccccc} a_{11}X_1 & + & \dots & + & \dots & + & \dots & + a_{1m}X_m & = & b_1 \\ \vdots & & & & \vdots & & & \vdots & & \vdots \\ a_{i1}X_1 & + & \dots & + a_{ij}X_j & + & \dots & + a_{im}X_m & = & b_i \\ \vdots & & & & \vdots & & & \vdots & & \vdots \\ a_{n1}X_1 & + & \dots & + & \dots & + & \dots & + a_{nm}X_m & = & b_n \end{array} \right.$$

Eine **Lösung des Systems** $\mathscr{S}$ ist ein $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in K^m$ mit

$$\begin{array}{ccccccccc} a_{11}x_1 & + & \dots & + & \dots & + & \dots & + a_{1m}x_m & = & b_1 \\ \vdots & & & & \vdots & & & \vdots & & \vdots \\ a_{i1}x_1 & + & \dots & + a_{ij}x_j & + & \dots & + a_{im}x_m & = & b_i \\ \vdots & & & & \vdots & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \dots & + & \dots & + & \dots & + a_{nm}x_m & = & b_n \end{array}$$

Der **Lösungsraum** des Systems $\mathscr{S}$ ist die Teilmenge

$$\mathscr{L}(\mathscr{S}) := \{x \in K^m \ ; \ x \text{ ist Lösung von } \mathscr{S}\}.$$

Bemerkung 9.46. Ein lineares Gleichungssystem erlaubt eine kompakte Darstellung in **Matrixschreibweise:**

(1) Die Koeffizienten a_{ij} eines (inhomogenen) linearen Gleichunssystem kann man als Einträge einer Matrix auffassen:

$$A = (a_{ij}) \in \mathrm{M}_{n\times m}(K).$$

- Die Spaltenanzahl m ist gleich der Variablenanzahl.
- Die Zeilenanzahl n ist gleich der Anzahl der linearen Gleichungen im System.

(2) Die Inhomogenitäten $b_1, \dots, b_n$ sind die Komponenten eines Spaltenvektors

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in K^n.$$

(3) Wir bilden weiter aus den Variablen $X_1, \dots, X_m$ einen Vektor von Variablen

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix},$$

aber das ist nur eine Notation und kein Vektor in K^m.

Die Matrixschreibweise des linearen Gleichungssystems lautet nun

$$AX = b,$$

denn ein $x \in K^m$ ist Lösung genau dann, wenn

$$Ax = b.$$

Notation 9.47. Wir schreiben für den Lösungsraum des Systems $AX = b$ einfach

$$\mathscr{L}(AX = b).$$

Bemerkung 9.48. (1) Ein homogenes lineares Gleichungssystem ist nichts anderes als ein inhomogenes lineares Gleichungssystem $AX = b$ mit $b = 0$.

(2) Der Lösungsraum von $AX = b$ hat mit der linearen Abbildung $L_A : K^m \to K^n$ zu tun:

$$\mathscr{L}(AX = b) = L_A^{-1}(b).$$

(3) Wenn $A \in GL_n(K)$ invertierbar ist, dann hat für $b \in K^n$ das Gleichungssystem

$$AX = b$$

nur die Lösung $x = A^{-1}b$.

Definition 9.49. Seien K ein Körper, $n, m \in \mathbb{N}_0$, weiter $A \in \mathrm{M}_{n\times m}(K)$ und $b \in K^n$. Sei

$$AX = b$$

ein inhomogenes lineares Gleichungssystem.

(1) Das **zugehörige homogene lineare Gleichungssystem** ist

$$AX = 0.$$

(2) Sei $A = [v_1, \ldots, v_m]$. Die **erweiterte Matrix** des inhomogenen linearen Gleichungssystems ist

$$B = [A, b] := [v_1, \ldots, v_m, b],$$

also die Matrix aus den Spalten von A verlängert um die Spalte b.

Satz 9.50. *Sei K ein Körper, $A \in \mathrm{M}_{n\times m}(K)$ und $b \in K^n$. Wir betrachten das inhomogene lineare Gleichungssystem*

$$AX = b.$$

(1) Der Lösungsraum des zugehörigen homogenen linearen Gleichungssystems $AX = 0$ ist der Kern der linearen Abbildung $L_A(x) = Ax$:

$$\mathscr{L}(AX = 0) = \ker(A) = \ker(L_A).$$

Dies ist ein Unterraum der Dimension

$$\dim(\mathscr{L}(AX = 0)) = m - \mathrm{rg}(A).$$

(2) Der Lösungsraum des inhomogenen linearen Gleichungssystems $AX = b$ ist genau eine der beiden Alternativen:

- *entweder leer: $\mathscr{L}(AX = b) = \emptyset$, oder*
- *zu einer beliebigen Lösung $x_0 \in \mathscr{L}(AX = b)$ gilt*

$$\mathscr{L}(AX = b) = x_0 + \ker(A) := \{x \in K^m \ ; \ \exists y \in \ker(A) \text{ mit } x = x_0 + y\}.$$

(3) Kriterium für die Existenz von Lösungen: es gibt eine Lösung für $AX = b$ genau dann, wenn

$$\mathrm{rg}(A) = \mathrm{rg}([A, b]),$$

d.h. wenn die Matrix A und die erweiterte Matrix $B = [A, b]$ den gleichen Rang haben.

Beweis. (1) Es gilt

$$x \in \mathscr{L}(AX = 0) \iff Ax = 0 \iff L_A(x) = 0.$$

Der Kern $\ker(L_A)$ ist aber bekanntlich ein Unterraum, Proposition 7.17, und seine Dimension bestimmt sich nach Satz 7.21 und Proposition 9.3(3) zu

$$\dim(\mathscr{L}(AX = 0)) = \dim(\ker(L_A)) = m - \dim(\mathrm{im}(L_A)) = m - \mathrm{rg}(A).$$

(2) Angenommen $\mathscr{L}(AX = b)$ ist nicht leer. Sei $x_0 \in \mathscr{L}(AX = b)$ eine Lösung. Für jedes $y \in \ker(A)$ ist $x = x_0 + y$ eine Lösung wegen:

$$Ax = A(x_0 + y) = Ax_0 + Ay = b + 0 = b.$$

Umgekehrt, wenn x eine weitere Lösung ist, dann ist $y = x - x_0$ in $\ker(A)$ wegen

$$Ay = A(x - x_0) = Ax - Ax_0 = b - b = 0.$$

(3) Sei $A = [v_1, \ldots, v_m]$ die Spaltenvektorschreibweise von A. Dann gilt

$$\mathscr{L}(AX = b) \neq \emptyset \iff \exists \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in K^m \text{ mit } x_1v_1 + \ldots + x_mv_m = b \iff b \in \langle v_1, \ldots, v_m \rangle_K$$

und das ist nach Proposition 4.28 äquivalent zu

$$\langle v_1, \ldots, v_m \rangle_K = \langle v_1, \ldots, v_m, b \rangle_K.$$

Da der linke Unterraum in jedem Fall im rechten enthalten ist, kann man nach die Gleichheit aus der Gleichheit der Dimensionen ablesen, ergo ist dies per Definition des Rangs äquivalent zu

$$\operatorname{rg}(A) = \operatorname{rg}([A, b]). \qquad \square$$

Bemerkung 9.51. (1) Man sagt: jede Lösung eines inhomogenen linearen Gleichungssystems ist von der Form

spezielle Lösung + Lösung des zug. homogenen linearen Gleichungssystems.

Dabei ist die spezielle Lösung keinesfalls speziell, sondern einfach nur eine Lösung, die man zur Beschreibung aller Lösungen ausgewählt hat. Jede Lösung von $AX = b$ ist als „spezielle Lösung“ brauchbar.

(2) Sei $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$ eine Matrix, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in K^n$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in K^m$.

Wir haben nun drei äquivalente Perspektiven:

(a) Zeilenweise: x ist Lösung des lineares Gleichungssystems:

$$\sum_{j=1}^{m} a_{ij}x_j = b_i \qquad \text{für } i = 1, \ldots, n.$$

(b) Matrixschreibweise: x löst die Matrixgleichung:

$$AX = b.$$

(c) Spaltenweise: b ist Linearkombination der Spalten mit den Koordinaten von x als Koeffizienten:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots + x_m \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Jede Perspektive hat ihre Bedeutungen.

(a) Zeilenweise: Gleichungssysteme möchte man in Anwendungen lösen. Die Gleichungen beschreiben den Unterraum $\ker(A)$. Man kann sich fragen, welche und wieviele der Gleichungen bereits ausreichen, um $\ker(A)$ zu beschreiben. Wieviele der Gleichungen sind linear unabhängig? Das wird durch den Zeilenrang beantwortet, siehe Satz 9.7.

(b) Matrixschreibweise: Strukturaussagen und Normalformen für Matrizen beeinflussen Matrixgleichungen. Ein weiterer Vorteil besteht in der bestechenden Kürze der Notation. Man kann so schnell auch ein großes lineares Gleichungssystem ansprechen und darüber nachdenken, weil man nicht von den Details der einzelnen Koeffizienten und der besonders fehleranfälligen Indizes abgelenkt wird.

(c) Spaltenweise: Dieser Standpunkt betont die Geometrie des Spaltenraums $C(A)$, der von den Spalten erzeugt wird, und damit das Lösbarkeitskriterium $b \in C(A)$.

Bemerkung 9.52. Wir betrachten das lineare Gleichungssystem $AX = 0$ mit $A \in \mathrm{M}_{n\times m}(K)$. Der Zeilenrang von A gibt an, wieviele linear unabhängige Gleichungen sich in den n Gleichungen von A verbergen. Intuitiv geht die Dimension des Lösungsraumes mit jeder weiteren, unabhängig von den anderen zu erfüllenden Gleichung um 1 runter. Wir erwarten also

$$\dim(\ker(A)) = m - \text{z-rg}(A).$$

Andererseits folgt aus der Kern–Bild–Dimensionsformel, Satz 7.21,

$$\dim(\ker(A)) = m - \mathrm{rg}(A) = m - \text{s-rg}(A).$$

Unsere Intuition von Gleichungen und Dimension besagt somit $\text{z-rg}(A) = \text{s-rg}(A)$. Und das gilt nach Satz 9.7.

Korollar 9.53. *Sei $AX = 0$ ein homogenes Gleichungssystem mit m Variablen und n Gleichungen. Wenn $m > n$, dann gibt es sicher eine nichttriviale, d.h. eine von 0 verschiedene Lösung $x \in K^m$.*

Beweis. Es gibt eine nichttriviale Lösung genau dann, wenn $\ker(A) \neq 0$. Aber

$$\dim(\ker(A)) = m - \mathrm{rg}(A) \geq m - n > 0,$$

also ist $\ker(A) \neq 0$. □

Verfahren 9.54. Wir besprechen nun ein Verfahren, mit dem lineare Gleichungssysteme gelöst werden können. Im Verfahren 9.33 haben wir bereits beschrieben, wie man aus der **reduzierten** Zeilenstufenform eine Basis des Lösungsraums eines homogenen linearen Gleichungssystems bestimmt. Dies ist das bevorzugte Verfahren. Hier besprechen wir nun, wie man mit der Methode des Rückwärtseinsetzens bereits aus der Zeilenstufenform (nicht notwendig reduziert) den Lösungsraum beschreibenn kann.

Sei $A \in \mathrm{M}_{n\times m}(K)$ und R die vom Gaußschen Eliminationsverfahren aus A erzeugte Matrix in Zeilenstufenform. Sei S das Produkt von Elementarmatrizen mit $R = SA$. Weil S invertiertbar ist, folgt

$$Ax = 0 \iff S(Ax) = 0 \iff Rx = 0,$$

und damit

$$\mathscr{L}(AX = 0) = \ker(A) = \ker(R) = \mathscr{L}(RX = 0).$$

Das Gaußsche Eliminationsverfahren transformiert also ein beliebiges Gleichungssystem in ein äquivalentes[2] Gleichungssystem in Zeilenstufenform. Derart kann man es nun leicht durch **Rückwärtseinsetzen** lösen. Man geht fallend in den Indizes vor und unterscheidet Pivotspalten von den Nicht-Pivotspalten. Sei $J \subseteq \{1, \ldots, m\}$ die Menge der Nicht-Pivotspaltenindizes.

- Wenn $j \in J$ ein Index ist, der zu einer Nicht-Pivotspalte gehört, dann wird
$$x_j = t_j$$
ein **Parameter** $t_j \in K$ des Lösungsraums.
- Wenn $j = \rho_s$ der Index einer Pivotspalte ist, etwa wie in (9.1) zur s-ten Zeile, dann müssen wir die s-te Zeile nach x_{ρ_s} auflösen, was wegen $\alpha_s = r_{s\rho_s} \in K^\times$ geht:
$$x_{\rho_s} = -\alpha_s^{-1} \sum_{k=\rho_s+1}^{m} r_{sk} x_k.$$

[2] Gleichungssysteme heißen äquivalent, wenn sie den gleichen Lösungsraum haben.

Hier substituiert man stets für x_k mit $k > \rho_s$ die bereits erhaltene Linearkombination in den t_k mit den Indizes k der Nicht-Pivotspalten. So gewinnt man eine Darstellung der Form

$$x_{\rho_s} = \sum_{j=\rho_s+1, j\in J}^{m} \xi_{\rho_s j} t_j$$

für gewisse $\xi_{\rho_s j} \in K$. Hierbei sind stets $j > \rho_s$ und $j \in J$.

- Wir ergänzen für $i, j \in J$

$$\xi_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

und $\xi_{\rho_s j} = 0$ für $j < \rho_s$ und $j \in J$. Dies ergibt für $j \in J$ die Vektoren

$$\xi_j = \begin{pmatrix} \xi_{1j} \\ \vdots \\ \xi_{mj} \end{pmatrix} \in K^m.$$

Schreiben wir zur Abkürzung $\underline{t} = (t_j)_{j\in J}$ für das Tupel der Parameter zu einer Lösung, so hat die Lösung zum Parameter $\underline{t}$ die Form

$$x(\underline{t}) = \sum_{j\in J} t_j \xi_j.$$

Der Lösungsraum wird daher von den Vektoren ξ_j aufgespannt. Wegen

$$\dim(\ker(A)) = m - \operatorname{rg}(A) = m - \text{ Anzahl Pivotspalten } = \text{ Anzahl Nichtpivotspalten} = |J|$$

ist das die richtige Anzahl von Vektoren für eine Basis des Nullraums von A. Diese Basis bekommt man auch durch die folgende spezielle Wahl der Parameter: man setze alle Parameter $t_k = 0$, $k \neq j$ und $t_j = 1$. Dann entsteht ξ_j.

Möchte man nun das inhomogene Gleichungssystem $AX = b$ lösen, so arbeitet man mit dem Ergebnis der Gauß-Elimination von A angewandt auf die erweiterte Matrtix $B = [A, b]$, das ist

$$[R, Sb]$$

mit S wie oben das Produkt der Elementarmatrizen aus der Gauß-Elimination mit $R = SA$. Schreiben wir

$$Sb = c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

dann haben wir beim Rückwärtseinsetzen nur in den Auflösungsschritten für x_{ρ_s} aus der s-ten Zeile das folgende zu verändern:

$$x_{\rho_s} = \alpha_s^{-1}\Big(c_s - \sum_{k=\rho_s+1}^{m} r_{sk} x_k\Big),$$

woraus nach Substitution der x_j mit $j > \rho_s$ jetzt

$$x_{\rho_s} = \delta_{\rho_s} + \sum_{j=\rho_s+1, j\in J}^{m} \xi_{\rho_s j} t_j$$

für gewisse zusätzliche $\delta_{\rho_s} \in K$ folgt. Ergänzen wir $\delta_j = 0$ für die $j \in J$, welche nicht zu Pivotspalten gehören, so haben wir

$$\delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix} \in K^m$$

und die Lösung zum Parameter $\underline{t}$ wird nun

$$x(\underline{t}) = \delta + \sum_{j \in J} t_j \xi_j.$$

Bei δ handelt es sich um eine spezielle Lösung, die in gewisser Weise von der Zeilenstufenform ausgesucht wird.

Beispiel 9.55. Wir betrachten das lineare Gleichungssystem mit Koeffizienten aus $K = \mathbb{Q}$ und Variablen $X_1, \ldots, X_5$:

$$\begin{array}{rcrcrcrcrcl} X_1 & - & 2X_2 & & & - & X_4 & + & X_5 & = & b_1 \\ & & & & & & & & X_5 & = & b_2 \\ & & & & X_3 & + & 3X_4 & - & X_5 & = & b_3 \\ 2X_1 & - & 4X_2 & + & X_3 & + & X_4 & & & = & b_4 \end{array}$$

Die erweiterte Matrix ist:

$$[A, b] = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 1 & b_1 \\ 0 & 0 & 0 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 3 & -1 & b_3 \\ 2 & -4 & 1 & 1 & 0 & b_4 \end{array}\right),$$

wobei wir durch den vertikalen Strich die Spalte der Inhomogenitäten abgrenzen.

Nun lassen wir den Gauß-Algorithmus laufen:

(1) Das -2-fache der Zeile 1 zur Zeile 4 addieren:

$$E_{41}(-2)[A, b] = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 1 & b_1 \\ 0 & 0 & 0 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 3 & -1 & b_3 \\ 0 & 0 & 1 & 3 & -2 & -2b_1 + b_4 \end{array}\right)$$

(2) Zeile 3 und Zeile 2 tauschen:

$$P_{23}E_{41}(-2)[A, b] = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 1 & b_1 \\ 0 & 0 & 1 & 3 & -1 & b_3 \\ 0 & 0 & 0 & 0 & 1 & b_2 \\ 0 & 0 & 1 & 3 & -2 & -2b_1 + b_4 \end{array}\right)$$

(3) Das -1-fache der Zeile 2 zur Zeile 4 addieren:

$$E_{42}(-1)P_{23}E_{41}(-2)[A, b] = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 1 & b_1 \\ 0 & 0 & 1 & 3 & -1 & b_3 \\ 0 & 0 & 0 & 0 & 1 & b_2 \\ 0 & 0 & 0 & 0 & -1 & -2b_1 - b_3 + b_4 \end{array}\right)$$

(4) Das 1-fache der Zeile 3 zur Zeile 4 addieren:

$$E_{43}(1)E_{42}(-1)P_{23}E_{41}(-2)[A, b] = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 1 & b_1 \\ 0 & 0 & 1 & 3 & -1 & b_3 \\ 0 & 0 & 0 & 0 & 1 & b_2 \\ 0 & 0 & 0 & 0 & 0 & -2b_1 + b_2 - b_3 + b_4 \end{array}\right)$$

Damit ist Zeilenstufenform für A erreicht.

Jetzt lesen wir ab:

- Der Rang von A ist 3.

- Die Spalten mit den Indizes 1, 3 und 5 erzeugen den Spaltenraum von A:
$$C(A) = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \rangle_K.$$
- Der Rang der erweiterten Matrix $[A, b]$ ist gleich $\mathrm{rg}(A)$ genau dann, wenn
$$-2b_1 + b_2 - b_3 + b_4 = 0.$$
- Wenn $-2b_1 + b_2 - b_3 + b_4 = 0$, dann lösen wir rückwärts nach den X_i auf. Die Nichtpivotspalten, die mit Indizes 2 und 4, liefern Parameter t_2, $t_4 \in K$ für die Lösung.
$$\begin{aligned} x_5 &= b_2 \\ x_4 &= t_4 \\ x_3 &= b_3 - 3x_4 + x_5 = b_3 - 3t_4 + b_2 \\ x_2 &= t_2 \\ x_1 &= b_1 + 2x_2 + x_4 - x_5 = b_1 + 2t_2 + t_4 - b_2. \end{aligned}$$
In Spaltenvektorschreibweise bekommt man
$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = t_2 \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_4 \cdot \begin{pmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} b_1 - b_2 \\ 0 \\ b_3 + b_2 \\ 0 \\ b_2 \end{pmatrix}$$

Eine spezielle Lösung bekommt man durch Wahl von Werten für t_2 und t_4, die bequemste Wahl ist oft $t_2 = t_4 = 0$. Damit ist
$$x_0 = \begin{pmatrix} b_1 - b_2 \\ 0 \\ b_3 + b_2 \\ 0 \\ b_2 \end{pmatrix} = \delta$$
eine spezielle Lösung von $AX = b$. Der Nullraum von A, also der Lösungsraum im Fall $b = 0$, hat als Basis die Koeffizientenvektoren in der Parameterdarstellung der allgemeinen Lösung:
$$\ker(A) = \langle \xi_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ \xi_4 = \begin{pmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} \rangle_K.$$

9.6. **Die *LR*-Zerlegung.** Die Matrizen $P_{\alpha,\beta}$, welche für die Zeilenvertauschung benutzt werden, lassen sich wie folgt verallgemeinern.

Definition 9.56. Sei K ein Körper und $\sigma \in S_n$ eine Permutation. Die **Permutationsmatrix** zu σ ist die Matrix $P_\sigma = (P_{\sigma,ij}) \in \mathrm{M}_n(K)$ mit Einträgen
$$(P_\sigma)_{ij} = \delta_{i,\sigma(j)},$$
also j-ter Spalte $P_\sigma e_j = e_{\sigma(j)}$ und damit
$$P_\sigma = [e_{\sigma(1)}, \ldots, e_{\sigma(n)}].$$

Beispiel 9.57. (1)
$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \Longrightarrow P_\sigma \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} c \\ a \\ b \end{pmatrix}.$$

(2)
$$P_\sigma \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\delta_{i,\sigma(j)})_{\substack{1\le i\le n \\ 1\le j\le m}} (x_i)_{1\le i\le n} = (\sum_{k=1}^{n} \delta_{i,\sigma(k)} x_k)_{1\le i\le n} = \begin{pmatrix} x_{\sigma^{-1}(1)} \\ \vdots \\ x_{\sigma^{-1}(n)} \end{pmatrix}$$
Linksmultiplikation mit P_σ permutiert die Koordinaten so, wie es die Permutation σ vorgibt. In Zeile i von $P_\sigma x$ steht die Koordinate $x_{\sigma^{-1}(i)}$, die von σ von Position $\sigma^{-1}(i)$ an die Position $\sigma(\sigma^{-1}(i)) = i$ gebracht wird.

Proposition 9.58. *Für $\sigma, \tau \in S_n$ gilt $P_{\sigma\tau} = P_\sigma P_\tau$.*

Beweis. Nachrechnen. □

Das algorithmische Resultat des Gaußschen Eliminationsverfahrens führt zu einer strukturellen Aussage.

Satz 9.59 (LR-Zerlegung)**.** *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gibt es eine Permutationsmatrix P, eine obere Drieecksmatrix R und eine untere Dreiecksmatrix L mit*
$$PA = LR.$$
Wenn im Eliminationsverfahren nach Gauß keine Zeilenvertauschungen nötig sind, dann gibt es die Zerlegung $A = LR$ mit $P = \mathbf{1}_n$.

Beweis. Eine Zeilenvertauschung kommutiert mit der Addition des Vielfachen einer Zeile zu einer anderen, sofern man die Indizes der beteiligten Zeilen anpaßt. Das bedeutet, daß man alle Zeilenvertauschungen, die im Gauß-Verfahren auftreten bereits vorweg am Anfang zusammenfassen kann, siehe Proposition 9.58. Das gibt eine Permutationsmatrix P und für PA läuft der Algorithmus jetzt ohne Zeilenvertauschungen.

Die auftretenden elementaren Zeilenoperationen multiplizieren von links mit Elementarmatrizen $E_{\alpha\beta}(\lambda)$, die untere Dreiecksmatrizen sind ($\alpha > \beta$). Das Produkt sei S und $R = SPA$ die entsehende Zeilenstufenform, eine obere Dreiecksmatrix. Die Inversen sind Elementarmatrizen derselben Form, also auch untere Dreiecksmatrizen $E_{\alpha\beta}(-\lambda)$. Das Podukt von unteren Dreiecksmatrizen ist wieder eine, also ist $L = S^{-1}$ eine untere Dreiecksmatrix, mit der gilt
$$LR = PA.$$ □

Bemerkung 9.60. Sei $A \in \mathrm{M}_n(K)$. Das Lösen des Gleichungssystems $AX = b$ benötigt einen Rechenaufwand von $\approx n^3$ Operationen. Es gilt aber bei $PA = LR$ wie in Satz 9.59, daß
$$Ax = b \quad \Longleftrightarrow \quad LRx = P^{-1}b.$$
Das berechnen von $c = P^{-1}b$ kostet $\approx n$ Operationen. Das Lösen von $LRx = c$ teilen wir auf in $Ly = c$ und $Rx = y$. Das berechnen von y aus c durch Vorwärtseinsetzen und das anschließende Lösen von $Rx = y$ durch Rückwärtseinsetzen kostet $\approx n^2$ viele Operationen. Wenn man viele Gleichungen $AX = b$ mit derselben Matrix A und variierendem b zu Lösen hat, dann ist der gestufte Prozeß über die LR-Zerlegung ökonomischer bezüglich des Rechenaufwandes.

Übungsaufgaben zu §9

Übungsaufgabe 9.1. Sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix vom Rang $\mathrm{rg}(A) = 1$. Zeigen Sie, daß es Vektoren $v \in K^n$ und $w \in K^m$ gibt mit
$$A = v \cdot w^t.$$

Übungsaufgabe 9.2. Seien $n \in \mathbb{N}_0$, $1 \le \alpha, \beta \le n$ mit $\alpha \ne \beta$. Zeigen Sie:

(a) Für $\lambda_1, \lambda_2 \in K$ gilt
$$E_{\alpha\beta}(\lambda_1) \cdot E_{\alpha\beta}(\lambda_2) = E_{\alpha\beta}(\lambda_1 + \lambda_2).$$

(b) Für $\mu_1, \mu_2 \in K^\times$ gilt
$$E_{\alpha\alpha}(\mu_1) \cdot E_{\alpha\alpha}(\mu_2) = E_{\alpha\alpha}(\mu_1 \cdot \mu_2).$$

Übungsaufgabe 9.3 (Elementare Spaltenoperationen). Sei K ein Körper und $A \in \mathrm{M}_{n\times m}$. Berechnen Sie für $1 \le \alpha, \beta \le m$, $\alpha \ne \beta$ und $\lambda \in K$, $\mu \in K^\times$
$$AE_{\alpha\beta}(\lambda) \quad \text{und} \quad AE_{\alpha\alpha}(\mu).$$
Formulieren Sie, was elementare Spaltenoperationen mit Matrizen machen sollen.

Übungsaufgabe 9.4. Welche Form einer Matrix erreicht das Gaußsche Eliminationsverfahren, wenn man anstelle von Zeilenoperationen Spaltenoperationen verwendet?

Übungsaufgabe 9.5. (a) Nehmen Sie an, die Matrix $R \in \mathrm{M}_{n\times m}$ habe Zeilenstufenform wie in (9.1). Beschreiben Sie, wie man mit Spaltenoperationen, also durch Multiplikation mit Elementarmatrizen von rechts (das Produkt all dieser sei T) die folgende Form erhält:

$$RT = \begin{pmatrix} \alpha_1 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \alpha_r & & & \\ & & & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & & \cdots & 0 & 0 \end{pmatrix}$$

Dabei sind die Spalten rechts und Zeilen unten aus lauter Nullen optional.

(b) Schließen Sie, daß es für eine beliebige Matrix $A \in \mathrm{M}_{n\times m}(K)$ Produkte von Elementarmatrizen $S \in \mathrm{M}_n(K)$ und $T \in \mathrm{M}_m(K)$ gibt mit

$$SAT = \begin{pmatrix} \alpha_1 & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \alpha_r & & & \\ & & & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & & \cdots & 0 & 0 \end{pmatrix} =: D$$

(c) Sei nun $A \in \mathrm{GL}_n(K)$ invertierbar. Schließen Sie, daß dann D auch invertierbar ist und daher $n = r$. Schreiben Sie das Inverse von D hin und zeigen Sie

$$A^{-1} = TD^{-1}S.$$

Bemerkung: Auf diese Weise kann man mittels Gauß-Elimination, zuerst für Zeilen, dann für Spalten, das Inverse einer invertierbaren Matrix bestimmen.

Übungsaufgabe 9.6. Sei $A \in \mathrm{M}_n(K)$. Zeigen Sie die Äquivalenz der folgenden Aussagen:

(a) Die Gleichung $AX = b$ hat für alle $b \in K^n$ mindestens eine Lösung.
(b) Die Gleichung $AX = 0$ hat höchstens eine Lösung.

Wenn diese Aussagen zutreffen, dann hat $AX = b$ für jedes $b \in K^n$ genau eine Lösung, es ist $A \in \mathrm{GL}_n(K)$, und die eindeutige Lösung ist

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

10. Vektorräume linearer Abbildungen

10.1. **Lineare Abbildungen als Vektoren.** Anstatt eine einzelne lineare Abbildung $V \to W$ zu verstehen, versuchen wir, alle linearen Abbildungen $V \to W$ gemeinsam zu verstehen.

Notation 10.1. Seien V und W zwei K-Vektorräume. Die Menge aller linearen Abbildungen von V nach W bezeichnen wir mit

$$\mathrm{Hom}_K(V, W) = \{f : V \to W \ ; \ \text{linear}\}.$$

Die Abkürzung „Hom“ steht für **Homomorphismus**, bei Vektorräumen eine alternative Terminologie für *lineare Abbildung*.

Wir definieren die Addition linearer Abbildungen wie folgt. Zu $f, g \in \mathrm{Hom}_K(V,W)$ definieren wir durch

$$(f+g)(x) = f(x) + g(x) \tag{10.1}$$

für alle $x \in V$ eine Abbildung

$$f+g : V \to W.$$

Dieses $h := f+g$ ist ein lineare Abbildung: für $x, y \in V$ und $\lambda \in K$ gilt

$$\begin{aligned} h(\lambda x + y) &= f(\lambda x + y) + g(\lambda x + y) = \lambda f(x) + f(y) + \lambda g(x) + g(y) \\ &= \lambda\big(f(x) + g(x)\big) + \big(f(y) + g(y)\big) = \lambda h(x) + h(y). \end{aligned}$$

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

Als nächstes definieren wir die Skalarmultiplikation für K-lineare Abbildungen mit einem Skalar aus K. Zu $a \in K$ und $f \in \mathrm{Hom}_K(V,W)$ definieren wir durch

$$(af)(x) = af(x). \tag{10.2}$$

für alle $x \in V$ eine Abbildung

$$af : V \to W.$$

Dieses $h := af$ ist eine lineare Abbildung: für $x, y \in V$ und $\lambda \in K$ gilt

$$h(\lambda x + y) = af(\lambda x + y) = a\big(\lambda f(x) + f(y)\big) = \lambda(af(x)) + af(y) = \lambda h(x) + h(y).$$

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

Proposition 10.2. *Seien V und W zwei K-Vektorräume. Dann ist $\mathrm{Hom}_K(V,W)$ mit der durch (10.1) und (10.2) definierten Addition und Skalarmultiplikation ein K-Vektorraum.*

Beweis. Die Vektorraumaxiome kann man zur Übung selbst nachrechnen. Wir beweisen das gleich (im endlichdimensionalen Fall) als Konsequenz der Beschreibung durch Matrizen. □

Bemerkung 10.3. Weil $\mathrm{Hom}_K(V,W)$ selbst ein K-Vektorraum ist, wird damit die Theorie der Vektorräume nutzbar, um lineare Abbildungen und damit wiederum Vektorräume zu studieren!

Es stellt sich die Frage nach einer Basis von $\mathrm{Hom}_K(V,W)$ und dem zugehörigen Koordinatenisomorphismus zu einem K^r für ein geeignetes r. Die Antwort auf diese zweite Frage wird durch Matrizen geleistet. Die Einträge der Darstellungsmatrix sind Koordinaten für linearer Abbildungen $V \to W$, also für Vektoren des Vektorraums $\mathrm{Hom}_K(V,W)$. Eine Basis erhält man durch Transport einer Basis von $\mathrm{M}_{n\times m}(K)$.

Satz 10.4. *Sei V ein K-Vektorraum mit Basis $\mathscr{B} = (b_1, \ldots, b_m)$, und sei W ein K-Vektorraum mit Basis $\mathscr{C} = (c_1, \ldots, c_n)$. Dann ist*

$$\begin{aligned} \mathrm{M}_{\mathscr{C}}^{\mathscr{B}} : \ \mathrm{Hom}_K(V,W) &\to \mathrm{M}_{n\times m}(K) \\ f &\mapsto \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) \end{aligned}$$

ein Isomorphismus von K-Vektorräumen.

Beweis. Wenn wir die Übung aus dem Beweis von Proposition 10.2 nicht gemacht haben, dann wissen wir noch gar nicht, daß $\mathrm{Hom}_K(V,W)$ ein K-Vektorraum ist. Wenn wir aber zeigen, daß

(i) $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}$ bijektiv ist und

(ii) mit Addition verträglich ist: für alle $f, g \in \mathrm{Hom}_K(V,W)$ gilt

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f+g) = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) + \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(g)$$

(iii) und mit Skalarmultiplikation verträglich ist: für alle $\lambda \in K$ und $f \in \mathrm{Hom}_K(V,W)$ gilt

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\lambda f) = \lambda\, \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f),$$

dann haben wir die Vektorraumaxiome für $\mathrm{Hom}_K(V,W)$ dadurch nachgewiesen, dass diese in $\mathrm{M}_{n\times m}(K)$ gelten und aufgrund der mit Addition und Skalarmultiplikation verträglichen Bijektion in $\mathrm{M}_{n\times m}(K)$ verifiziert werden können.

(i) Die Koordinatenvektorabbildung $\kappa_{\mathscr{C}}$ ist bijektiv. Also bestimmen die Spalten von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$ die Bilder $f(b_1),\dots,f(b_m)$ der Basisvektoren der Basis $\mathscr{B}$. Nach Satz 7.12 bestimmt dies die lineare Abbildung f eindeutig. Damit ist $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}$ injektiv.

Zeigen wir nun die Surjektivität. Dazu sei eine Matrix $A\in\mathrm{M}_{n\times m}(K)$ gegeben. Diese schreiben wir als

$$A=[y_1,\dots,y_m]\in\mathrm{M}_{n\times m}(K)$$

mit m-vielen Spaltenvektoren $y_j\in K^n$. Die Koordinatenvektorabbildung $\kappa_{\mathscr{C}}$ ist bijektiv, also gehören dazu Vektoren $w_1,\dots,w_m\in W$ mit

$$y_j=\kappa_{\mathscr{C}}(w_j)\quad\text{für alle } j=1,\dots,m.$$

Nach Satz 7.12 gibt es eine lineare Abbildung $f:V\to W$ mit

$$f(b_j)=w_j\quad\text{für alle } j=1,\dots,m.$$

Dann gilt per Definition von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}$ gerade

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)=[\kappa_{\mathscr{C}}(f(b_1)),\dots,\kappa_{\mathscr{C}}(f(b_m))]=[\kappa_{\mathscr{C}}(w_1),\dots,\kappa_{\mathscr{C}}(w_m)]=[y_1,\dots,y_m]=A.$$

Damit ist $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}$ surjektiv.

(ii) Dies ist eine Rechnung: die j-te Spalte von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f+g)$ ist

$$\kappa_{\mathscr{C}}((f+g)(b_j))=\kappa_{\mathscr{C}}(f(b_j)+g(b_j))=\kappa_{\mathscr{C}}(f(b_j))+\kappa_{\mathscr{C}}(g(b_j))$$

und das ist die j-te Spalte von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)+\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(g)$.

(iii) Dies ist auch eine Rechnung: die j-te Spalte von $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\lambda f)$ ist

$$\kappa_{\mathscr{C}}((\lambda f)(b_j))=\kappa_{\mathscr{C}}(\lambda(f(b_j)))=\lambda\kappa_{\mathscr{C}}(f(b_j))$$

und das ist die j-te Spalte von $\lambda\,\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)$. □

Korollar 10.5. *Für endlichdimensionale K-Vektorräume V und W gilt*

$$\dim\big(\mathrm{Hom}_K(V,W)\big)=\dim(V)\cdot\dim(W).$$

Beweis. Das folgt sofort aus Satz 10.4 und Korollar 8.9. □

Korollar 10.6. *Sei K ein Körper, und seien $n,m\in\mathbb{N}_0$. Die Zuordnung $A\mapsto L_A$ definiert einen Isomorphismus von K-Vektorräumen*

$$L\ :\ \mathrm{M}_{n\times m}(K)\xrightarrow{\sim}\mathrm{Hom}_K(K^m,K^n),$$

die invers zu $\mathrm{M}_{\mathscr{F}}^{\mathscr{E}}$ ist, wenn $\mathscr{E}$ Standardbasis von K^m und $\mathscr{F}$ Standardbasis von K^n sind.

Beweis. Weil $\mathrm{M}_{\mathscr{F}}^{\mathscr{E}}(-)$ ein Isomorphismus von Vektorräumen ist, reicht es aus, wenn wir zeigen, daß L die Inverse Abbildung ist. Damit ist L dann automatisch linear (was man auch direkt nachweisen kann).

Aus Beispiel 8.31 haben wir für alle Matrizen A

$$\mathrm{M}_{\mathscr{F}}^{\mathscr{E}}(L_A)=A.$$

Das zeigt $\mathrm{M}_{\mathscr{F}}^{\mathscr{E}}(-)\circ L=\mathrm{id}$.

Der Koordinatenisomorphismus bezüglich der Standardbasis des K^n ist die Identität. Also gilt für alle linearen Abbildungen $f:K^m\to K^n$ nach Proposition 8.32

$$L_{M_{\mathscr{F}}^{\mathscr{E}}(f)}=f.$$

Das zeigt $L \circ \mathrm{M}_{\mathscr{F}}^{\mathscr{E}}(-) = \mathrm{id}$. □

10.2. **Dualisieren.** Der Dualraum ist ein Spezialfall des Vektorraums der linearen Abbildungen, und damit sind wir dem Dualraum bereits begegnet.

Definition 10.7. Seien K ein Körper und V ein K-Vektorraum.

(1) Eine **Linearform** auf V ist eine lineare Abbildung
$$\pi : V \to K.$$
Hier bezeichnen wir den Vektorraum K^1 mit K.

(2) Der **Dualraum** von V ist der Vektorraum aller Linearformen auf V:
$$V^* = \mathrm{Hom}_K(V, K) = \{\pi : V \to K \ ; \ \text{ linear}\}.$$

Beispiel 10.8. Eine Linearform auf dem $\mathbb{R}^3$ hat die Gestalt $\pi : \mathbb{R}^3 \to \mathbb{R}$ mit zum Beispiel
$$\pi(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = 2x + 3y + 5z.$$
Die Formel hat eine „Form“, in der die Koordinaten ausschließlich „linear“ kombiniert vorkommen.

Beispiel 10.9. Zu $v \in K^n$ betrachten wir
$$v \bullet - : K^n \to K, \qquad x \mapsto v \bullet x.$$
Diese Abbildung ist nach Proposition 8.16 eine Linearform auf K^n.

Bemerkung 10.10. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V. Eine Linearform $\pi : V \to K$ ist nach Satz 7.12 eindeutig durch das Tupel der Werte
$$\pi(b_1), \ldots, \pi(b_n) \in K$$
festgelegt, und jede Wahl von Werten aus K gehört zu einer Linearform. Das 1-Tupel $\mathscr{E} = (1)$ ist die Standardbasis des $K^1 = K$. Dann ist
$$\mathrm{M}_{\mathscr{E}}^{\mathscr{B}}(\pi) = (\pi(b_1), \ldots, \pi(b_n)) \in \mathrm{M}_{1\times n}(K)$$
die Koordinatenbeschreibung der Linearform. Wir haben einen Isomorphismus aus Satz 10.4
$$\mathrm{M}_{\mathscr{E}}^{\mathscr{B}} : V^* \xrightarrow{\sim} \mathrm{M}_{1\times n}(K).$$
Das Inverse dieses Isomorphismus transportiert die Standardbasis der Zeilenvektoren
$$e_j^t = (0, \ldots, 0, \underset{\uparrow j}{1}, 0, \ldots, 0) \in \mathrm{M}_{1\times n}(K)$$
zu einer Basis $\mathscr{B}^* = (b_1^*, \ldots, b_n^*)$ von V^*. Dazu die folgende Proposition und dann Definition.

Proposition 10.11. *Sei V ein K-Vektorraum mit Basis $\mathscr{B} = (b_1, \ldots, b_n)$. Durch die Vorgabe der Werte auf $\mathscr{B}$ werden Linearformen $b_j^* : V \to K$ definiert durch*
$$b_j^*(b_i) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$
Dann ist $\mathscr{B}^* = (b_1^*, \ldots, b_n^*)$ eine Basis von V^*.

Beweis. Sei $\pi \in V^*$ beliebig. Dann gilt
$$\pi = \pi(b_1)b_1^* + \ldots + \pi(b_n)b_n^*$$

nach Satz 7.12, weil beide Seiten Linearformen sind und auf der Basis $\mathscr{B}$ den gleichen Wert annehmen: in der Tat gilt für alle $i = 1, \ldots, n$

$$\big(\pi(b_1)b_1^* + \ldots + \pi(b_n)b_n^*\big)(b_i) = \sum_{j=1}^n \pi(b_j)b_j^*(b_i) = \sum_{j=1}^n \pi(b_j)\delta_{ij} = \pi(b_i).$$

Dies zeigt, daß $\mathscr{B}^*$ ein Erzeugendensystem ist.

Wir zeigen nun, daß $\mathscr{B}^*$ auch linear unabhängig ist. Seien dazu $a_i \in K$, $i = 1, \ldots, n$ mit $\sum_{i=1}^n a_i b_i^* = 0$. Einsetzen von b_j liefert

$$0 = \Big(\sum_{i=1}^n a_i b_i^*\Big)(b_j) = \sum_{i=1}^n a_i b_i^*(b_j) = \sum_{i=1}^n a_i\delta_{ij} = a_j.$$

Daher handelt es sich um die triviale Linearkombination. □

Definition 10.12. Die in Proposition 10.11 konstruierte Basis $\mathscr{B}^*$ von V^* heißt die zu $\mathscr{B}$ duale Basis.

Vorsicht: Die duale Basis dualisiert nicht einzelne Vektoren! Die Linearform $b_j^* \in V^*$ hängt nicht nur von b_j ab, sondern benötigt auch die b_i für $i \neq j$.

Korollar 10.13. *Sei V ein endlichdimensionaler K-Vektorraum. Dann gilt*

$$\dim(V^*) = \dim(V).$$

Beweis. Sei $\mathscr{B}$ eine Basis von V und $\mathscr{B}^*$ die dazu duale Basis von V^*. Dann

$$\dim(V^*) = |\mathscr{B}^*| = |\mathscr{B}| = \dim(V).$$ □

Die duale Basis $\mathscr{B}^*$ beschreibt die Koordinatenabbildung bezüglich $\mathscr{B}$:

Proposition 10.14. *Sei V ein Vektorraum mit Basis $\mathscr{B} = (b_1, \ldots, b_n)$, und sei $\mathscr{B}^* = (b_1^*, \ldots, b_n^*)$ die zu $\mathscr{B}$ duale Basis von V^*. Dann gilt für alle $x \in V$ und alle $\pi \in V^*$:*

$$x = b_1^*(x)b_1 + \ldots + b_n^*(x)b_n, \tag{10.3}$$

$$\pi = \pi(b_1)b_1^* + \ldots + \pi(b_n)b_n^*. \tag{10.4}$$

Insbesondere gilt für die Koordinaten bezüglich $\mathscr{B}$ bzw. $\mathscr{B}^*$:

$$\kappa_{\mathscr{B}}(x) = \begin{pmatrix} b_1^*(x) \\ \vdots \\ b_n^*(x) \end{pmatrix}, \qquad \kappa_{\mathscr{B}^*}(\pi) = \begin{pmatrix} \pi(b_1) \\ \vdots \\ \pi(b_n) \end{pmatrix}.$$

Beweis. Sei $x \in V$ beliebig. Weil $\mathscr{B}$ eine Basis ist, gibt es $x_i \in K$ mit $x = \sum_{i=1}^n x_i b_i$. Wir berechnen nun

$$b_j^*(x) = b_j^*(\sum_{i=1}^n x_i b_i) = \sum_{i=1}^n x_i b_j^*(b_i) = \sum_{i=1}^n x_i\delta_{ij} = x_j.$$

Daher gilt für alle $x \in V$

$$x = b_1^*(x)b_1 + \ldots + b_n^*(x)b_n,$$

und das zeigt auch die Behauptung für $\kappa_{\mathscr{B}}(x)$ nach der Definition von $\kappa_{\mathscr{B}}$.

In gewissem Sinne dual dazu gibt es für $\pi \in V^*$, weil $\mathscr{B}^*$ eine Basis ist, Elemente $a_i \in K$ mit $\pi = \sum_{i=1}^n a_i b_i^*$. Die Rechnung

$$\pi(b_j) = \Big(\sum_{i=1}^n a_i b_i^*\Big)(b_j) = \sum_{i=1}^n a_i b_i^*(b_j) = \sum_{i=1}^n a_i\delta_{ij} = a_j$$

zeigt $\pi = \pi(b_1)b_1^* + \ldots + \pi(b_n)b_n^*$, und daraus folgen die restlichen Behauptungen. □

Korollar 10.15. *Sei V ein Vektorraum mit Basis $\mathscr{B} = (b_1, \dots, b_n)$, und sei $\mathscr{B}^* = (b_1^*, \dots, b_n^*)$ die zu $\mathscr{B}$ duale Basis von V^*. Dann gilt für alle $x \in V$ und alle $\pi \in V^*$ mit dem Standardskalarprodukt $- \bullet -$ auf K^n:*

$$\pi(v) = \kappa_{\mathscr{B}^*}(\pi) \bullet \kappa_{\mathscr{B}}(v).$$

Beweis. Aus Proposition 10.14 folgt

$$\begin{aligned}\pi(v) &= \Big(\sum_{i=1}^n \pi(b_i) b_i^*\Big)\Big(\sum_{j=1}^n b_j^*(x) b_j\Big) \\ &= \sum_{i=1}^n \sum_{j=1}^n \pi(b_i) b_j(x) b_i^*(b_j) = \sum_{i=1}^n \sum_{j=1}^n \pi(b_i) b_j(x) \delta_{ij} \\ &= \sum_{i=1}^n \pi(b_i) b_i(x) = \kappa_{\mathscr{B}^*}(\pi) \bullet \kappa_{\mathscr{B}}(v).\end{aligned}$$

□

Nun möchten wir auch lineare Abbildungen dualisieren.

Definition 10.16. Sei $f : V \to W$ eine lineare Abbildung. Die Abbildung

$$\begin{aligned} f^* : W^* &\to V^* \\ \pi &\mapsto f^*(\pi) = \pi \circ f \end{aligned}$$

heißt die zu f **duale (lineare) Abbildung**. $f^*(\pi) = \pi \circ f : V \to K$ ist als Komposition linearer Abbildungen wieder linear, also in V^*, und damit f^* als Abbildung wohldefiniert.

Proposition 10.17. *Die duale Abbildung ist linear.*

Beweis. Wir nutzen die Notation aus der Definition. Seien $\pi, \rho \in W^*$ und $\lambda \in K$. Dann ist für alle $x \in V$

$$f^*(\lambda\pi + \rho)(x) = (\lambda\pi + \rho)(f(x)) = \lambda\pi(f(x)) + \rho(f(x)) = \lambda\big(f^*(\pi)\big)(x) + \big(f^*(\rho)\big)(x),$$

also

$$f^*(\lambda\pi + \rho) = \lambda f^*(\pi) + f^*(\rho).$$

□

Proposition 10.18. *Dualisieren ist ein **kontravarianter Funktor**, das heißt:*

(1) Komposition: seien $f : V \to W$ und $g : U \to V$ lineare Abbildungen. Dann gilt

$$(f \circ g)^* = g^* \circ f^*$$

als lineare Abbildungen $W^* \to V^* \to U^*$.

(2) Identität: die duale Abbildung zur Identität auf V ist die Identität auf V^*.

Beweis. (1) Für $\pi \in W^*$ gilt

$$(f \circ g)^*(\pi) = \pi \circ (f \circ g) = (\pi \circ f) \circ g = (f^*(\pi)) \circ g = g^*(f^*\pi) = (g^* \circ f^*)(\pi).$$

(2) Für alle $\pi \in V^*$ gilt

$$(\mathrm{id}_V)^*(\pi) = \pi \circ \mathrm{id}_V = \pi.$$

□

Satz 10.19. *Seien V, W zwei endlichdimensionale K-Vektorräume mit Basis $\mathscr{B}$ von V und Basis $\mathscr{C}$ von W. Sei $f : V \to W$ eine lineare Abbildung. Dann gilt*

$$\mathrm{M}_{\mathscr{B}^*}^{\mathscr{C}^*}(f^*) = \big(\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\big)^t.$$

Beweis. Sei $\mathscr{B} = (b_1, \ldots, b_m)$ die Basis von V und $\mathscr{C} = (c_1, \ldots, c_n)$ die Basis von W. Wir berechnen für beide Matrizen den Eintrag an Position (i, j):

$$\begin{aligned}
\mathrm{M}_{\mathscr{B}^*}^{\mathscr{C}^*}(f^*)_{ij} &= e_i \bullet \mathrm{M}_{\mathscr{B}^*}^{\mathscr{C}^*}(f^*) e_j && \text{(Lemma 8.28)} \\
&= \kappa_{\mathscr{B}}(b_i) \bullet \kappa_{\mathscr{B}^*}(f^*(c_j^*)) = \kappa_{\mathscr{B}^*}(f^*(c_j^*)) \bullet \kappa_{\mathscr{B}}(b_i) \\
&= \big(f^*(c_j^*)\big)(b_i) = (c_j^* \circ f)(b_i) = c_j^*(f(b_i)) && \text{(Korollar 10.15)} \\
&= \kappa_{\mathscr{C}^*}(c_j^*) \bullet \kappa_{\mathscr{C}}\big(f(b_i)\big) && \text{(Korollar 10.15)} \\
&= e_j \bullet \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) e_i = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)_{ji} = \big(\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f)\big)_{ij}^t. && \text{(Lemma 8.28)}
\end{aligned}$$

Damit ist alles bewiesen. □

Satz 10.20. *Sei $f : V \to W$ eine lineare Abbildung endlichdimensionaler K-Vektorräume.*
(1) Wenn f surjektiv ist, dann ist f^* injektiv.
(2) Wenn f injektiv ist, dann ist f^* surjektiv.

Beweis. (1) Sei $\pi \in W^*$ im Kern von f^*. Das heißt, daß für alle $x \in V$ gilt

$$0 = f^*(\pi)(x) = \pi(f(x)).$$

Weil f surjektiv ist, nimmt $f(x)$ jeden Wert in W an. Daher folgt $\pi = 0$. Dies zeigt f^* injektiv.

(2) Sei $\mathscr{B} = (v_1, \ldots, v_n)$ eine Basis von V. Weil f injektiv ist, sind die $f(v_1), \ldots, f(v_n)$ linear unabhängig in W. Wir ergänzen mittels des Basisergänzungssatzes zu einer Basis

$$(f(v_1), \ldots, f(v_n), w_1, \ldots, w_s)$$

von W. Sei nun $\pi \in V^*$ beliebig. Wir definieren eine Linearform $\rho \in W^*$ durch

$$\begin{aligned}
\rho(f(v_i)) &= \pi(v_i) && \text{für } i = 1, \ldots, n, \\
\rho(w_j) &= 0 && \text{für } j = 1, \ldots, s.
\end{aligned}$$

Dann gilt $f^*(\rho) = \pi$ durch Vergleich der Werte auf der Basis $\mathscr{B}$. Also ist f^* surjektiv. □

Satz 10.21. *Sei $f : V \to W$ eine lineare Abbildung endlichdimensionaler K-Vektorräume. Dann ist*

$$\dim(\mathrm{im}(f)) = \dim(\mathrm{im}(f^*)).$$

Beweis. Die Abbildung f faktorisiert über die Inklusion des Bildes $i : \mathrm{im}(f) \hookrightarrow W$

$$f : V \xrightarrow{q} \mathrm{im}(f) \xrightarrow{i} W.$$

Dabei ist $q(x) = f(x)$ mit Wertebereich $\mathrm{im}(f)$. Die Abbildung i ist injektiv, und die Abbildung q ist surjektiv. Nun dualisieren wir (Proposition 10.18 besagt $f^* = q^* \circ i^*$)

$$f^* : W^* \xrightarrow{i^*} \mathrm{im}(f)^* \xrightarrow{q^*} V^*.$$

Nach Satz 10.20 ist q^* injektiv und i^* surjektiv. Daher ist

$$\mathrm{im}(f)^* = i^*(W^*) \simeq q^*(i^*(W^*)) = (i \circ q)^*(W^*) = f^*(W^*) = \mathrm{im}(f^*).$$

Aus Korollar 10.13 folgt sofort

$$\dim(\mathrm{im}(f^*)) = \dim(\mathrm{im}(f)^*) = \dim(\mathrm{im}(f)).$$ □

Zweiter Beweis von Satz 9.7. Es ist A^t die Darstellungsmatrix von $(L_A)^*$ bezüglich der dualen Basis, also nach Satz 10.21 und zweimal Proposition 9.4

$$\mathrm{rg}(A^t) = \dim(\mathrm{im}(L_A^*)) = \dim(\mathrm{im}(L_A)) = \mathrm{rg}(A).$$ □

In Linearformen kann man Vektoren einsetzen und erhält so Linearformen auf dem Dualraum.

Definition 10.22. Sei V ein K-Vektorraum. Zu $v \in V$ gehört die **Auswertung** in v

$$\mathrm{ev}_v : V^* \to K, \quad \mathrm{ev}_v(\pi) = \pi(v).$$

Lemma 10.23. *Die Abbildung* ev_v *ist eine Linearform auf* V^*.

Beweis. Zu $\lambda \in K$ und $\pi, \rho \in V^*$ rechnen wir

$$\mathrm{ev}_v(\lambda\pi + \rho) = (\lambda\pi + \rho)(v) = \lambda\pi(v) + \rho(v) = \lambda \mathrm{ev}_v(\pi) + \mathrm{ev}_v(\rho). \qquad \square$$

Satz 10.24. *Sei* V *ein Vektorraum.*

(1) Die Zuordnung

$$\begin{aligned} \mathrm{ev} : V &\to (V^*)^* \\ x &\mapsto \mathrm{ev}_x = (\pi \mapsto \pi(x)) \end{aligned}$$

ist eine lineare Abbildung.

(2) Wenn V *endlichdimensional ist, dann ist* $\mathrm{ev} : V \to (V^*)^*$ *ein Isomorphismus.*

Beweis. (1) Seien $x, y \in V$ und $\lambda \in K$ beliebig. Dann ist für alle $\pi \in V^*$

$$\mathrm{ev}_{\lambda x+y}(\pi) = \pi(\lambda x + y) = \lambda\pi(x) + \pi(y) = \lambda \mathrm{ev}_x(\pi) + \mathrm{ev}_y(\pi) = (\lambda \mathrm{ev}_x + \mathrm{ev}_y)(\pi).$$

Folglich gilt

$$\mathrm{ev}_{\lambda x+y} = \lambda \mathrm{ev}_x + \mathrm{ev}_y,$$

und ev ist linear.

(2) Sei $x \in V$, $x \neq 0$. Dann kann man x zu einer Basis $\mathscr{B} = (b_1 = x, b_2, \ldots, b_n)$ von V ergänzen. Für die Linearform $b_1^* \in V^*$ gilt dann $\mathrm{ev}_x(b_1^*) = b_1^*(x) = 1$. Somit liegt x nicht in $\ker(\mathrm{ev})$, denn dann wäre $\mathrm{ev}_x = 0$ als Linearform auf V^*. Wegen $\ker(\mathrm{ev}) = 0$ ist ev injektiv.

Zweimaliges Anwenden von Korollar 10.13 ergibt $\dim(V) = \dim(V^*) = \dim((V^*)^*)$. Damit ist ev eine injektive lineare Abbildung zwischen K-Vektorräumen gleicher Dimension und somit nach Satz 7.23 bijektiv. Mit Proposition 7.32 folgt, daß ev ein Isomorphismus ist. $\square$

Übungsaufgaben zu §10

Übungsaufgabe 10.1. Seien V und W zwei endlichdimensionale K-Vektorräume mit Basis $\mathscr{B} = (b_1, \ldots, b_m)$ von V und Basis $\mathscr{C} = (c_1, \ldots, c_n)$ von W. Beschreiben Sie die linearen Abbildungen $f_{\alpha\beta} : V \to W$, die mittels des Isomorphismus aus Satz 10.4

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}} : \mathrm{Hom}_K(V, W) \to \mathrm{M}_{n\times m}(K)$$

der Basis der Matrizen $e_{\alpha\beta}$ für $\alpha = 1, \ldots, n$ und $\beta = 1, \ldots, m$ von $\mathrm{M}_{n\times m}(K)$ entspricht.

Übungsaufgabe 10.2 (Natürlichkeit). Zeigen Sie, daß die Abbildung ev aus Satz 10.24 natürlich ist. Zur Unterscheidung bezeichnen wir ev für den Vektorraum V mit $\varphi_V : V \to (V^*)^*$. Dann bedeutet „natürlich“das folgende: für alle linearen Abbildungen $f : V \to W$ ist das folgende Diagramm kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{\varphi_V} & (V^*)^* \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle (f^*)^*} \\ W & \xrightarrow{\varphi_W} & (W^*)^* \end{array}$$

Das bedeutet: für alle $x \in V$ gilt

$$(f^*)^*(\varphi_V(x)) = \varphi_W(f(x)).$$

Teil 3. Determinanten, Eigenwerte und Diagonalisierbarkeit

11. Determinanten

11.1. Das Vorzeichen einer Permutation.

Definition 11.1. Sei $n \in \mathbb{N}$. Ein Element τ von S_n heißt **Transposition**, wenn τ genau zwei Elemente von $\{1, \ldots, n\}$ vertauscht und den Rest fest läßt. Es gibt also zu $1 \leq a, b \leq n$ mit $a \neq b$ die Transposition $\tau = (a, b) \in S_n$, mit

$$\tau(i) = \begin{cases} a & i = b, \\ b & i = a, \\ i & i \neq a, b. \end{cases} = \begin{pmatrix} 1 & 2 & \ldots & a & \ldots & b & \ldots & n-1 & n \\ 1 & 2 & \ldots & b & \ldots & a & \ldots & n-1 & n \end{pmatrix}$$

Satz 11.2. *Jedes Element von S_n ist ein Produkt von Transpositionen.*

Beweis. Wir führen Induktion nach n. Der Fall $n = 1$ ist trivial. Nehmen wir also an, daß die Aussage für $n - 1$ bereits bewiesen ist.

Sei $\sigma \in S_n$ beliebig, und sei $\sigma(n) = a$. Falls $a \neq n$, so betrachten wir $\tau = (a, n)\sigma$. Dann ist

$$\tau(n) = (a, n)(\sigma(n)) = n.$$

Weil $(a, n)\tau = (a, n)(a, n)\sigma = \sigma$, so reicht es anstelle von σ die Permutation τ als Produkt von Transpositionen zu schreiben. Dieses τ nennen wir wieder σ und dürfen nun annehmen, daß $\sigma(n) = n$ gilt. In diesem Fall ist σ eingeschränkt auf $\{1, \ldots, n-1\}$ ein Element $\sigma_0 \in S_{n-1}$. Per Induktionsannahme gibt es (a_i, b_i), $i = 1, \ldots, r$, mit

$$\sigma_0 = (a_1, b_1) \cdots (a_r, b_r) \in S_{n-1}.$$

Dann gilt auch $\sigma = (a_1, b_1) \cdots (a_r, b_r)$ in S_n, denn für die Komposition der Transpositionen spielt es keine Rolle, ob wir auch noch das Element n permutieren, denn das bleibt bei den (a_i, b_i) sowieso fix. □

Definition 11.3. Das **Vorzeichen** (oder **Signum**) einer Permutation $\sigma \in S_n$ ist definiert als

$$\operatorname{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Beispiel 11.4. Sei $n \geq 2$. Wir berechnen das Signum der Transposition $(1, 2)$. Die Faktoren zu $3 \leq i < j$ sind 1. Die restlichen Faktoren gehören zu $i = 1$ und $2 \leq j$ sowie $i = 2$ und $3 \leq j$, so daß

$$\operatorname{sign}((1,2)) = \Big(\frac{1-2}{2-1} \cdot \prod_{3 \leq j \leq n} \frac{j-2}{j-1}\Big) \cdot \prod_{3 \leq j \leq n} \frac{j-1}{j-2} = \frac{1-2}{2-1} = -1.$$

Das Signum der Transposition $(1, 2)$ ist -1, folglich ist $(1, 2)$ ungerade.

Definition 11.5. Ein **Fehlstand** der Permutation $\sigma \in S_n$ ist ein Paar (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$. Die **Inversionszahl** von σ ist die Anzahl der Fehlstände:

$$\operatorname{inv}(\sigma) := |\{(i, j)\ ;\ 1 \leq i < j \leq n \text{ und } \sigma(i) > \sigma(j)\}|.$$

Proposition 11.6. *Für $\sigma \in S_n$ gilt*

$$\operatorname{sign}(\sigma) = (-1)^{\operatorname{inv}(\sigma)}.$$

Beweis. In der Definierenden Formel für $\operatorname{sign}(\sigma)$ treten, bis auf ein Vorzeichen, im Zähler die gleichen Differenzen wie im Nenner auf. Im Nenner sind die Differenzen stets positiv. Negative

Zähler gibt es genau so viele, wie es Paare $i < j$ mit $\sigma(i) > \sigma(j)$ gibt. Also enthält $\operatorname{sign}(\sigma)$ für jeden Fehlstand ein -1. □

Korollar 11.7. *Das Signum einer Permutation ist* 1 *oder* -1.

Beweis. Das folgt sofort aus Proposition 11.6. Wir beweisen es ein zweites Mal direkt aus der definierenden Formel. Per Definition ist zunächst $\operatorname{sign}(\sigma) \in \mathbb{Q}$. Um $\operatorname{sign}(\sigma) \in \{\pm 1\}$ zu erkennen, berechnen wir das Quadrat:

$$\begin{aligned}\operatorname{sign}(\sigma)^2 &= \prod_{1\le i<j\le n} \left(\frac{\sigma(j)-\sigma(i)}{j-i}\right)^2 = \prod_{1\le i<j\le n} \frac{(\sigma(i)-\sigma(j))(\sigma(j)-\sigma(i))}{(i-j)(j-i)} \\ &= \prod_{i\neq j} \frac{\sigma(i)-\sigma(j)}{i-j} = \frac{\prod_{i\neq j}(\sigma(i)-\sigma(j))}{\prod_{i\neq j}(i-j)} = 1.\end{aligned}$$

Zähler und Nenner im letzten Bruch unterscheiden sich nur in der Reihenfolge der Faktoren. □

Definition 11.8. Eine Permutation $\sigma \in S_n$ heißt **gerade**, wenn $\operatorname{sign}(\sigma) = 1$, und **ungerade**, wenn $\operatorname{sign}(\sigma) = -1$.

Korollar 11.9. *Jede Transposition ist eine ungerade Permutation.*

Beweis. Sei $1 \le a < b \le n$. Wir zählen die Fehlstände der Transposition $\tau = (a, b) \in S_n$. Ein Fehlstand kann nur entstehen, wenn i oder j in $\{a, b\}$ ist, damit von der Transposition auf das andere getauscht wird und dadurch ein Fehlstand entsteht: der andere Index muß im Intervall $[a, b]$ liegen.

(1) $i = a$ und $a < j < b$,
(2) $a < i < b$ und $j = b$,
(3) $i = a$ und $j = b$.

Damit gilt $\operatorname{inv}((a,b)) = 2(b-a-1)+1 \equiv 1 \pmod 2$, folglich $\operatorname{sign}((a,b)) = (-1)^{\operatorname{inv}(\sigma)} = -1$. □

Satz 11.10. *Für* $\sigma, \tau \in S_n$ *gilt*

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau).$$

Beweis. Per Definition ist das Signum ein Produkt über alle $1 \le i < j \le n$. Weil

$$\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j},$$

kommt es nicht auf $i < j$ oder $j < i$ an, sondern nur, daß jede zweielementige Teilmenge $\{i, j\} \subseteq \{1, \dots, n\}$ nur einmal einen Faktor beiträgt. Weil für jedes $\tau \in S_n$ auch $\{\tau(i), \tau(j)\}$ durch alle zweielementigen Teilmengen läuft, ist klar, daß

$$\begin{aligned}\operatorname{sign}(\sigma) &= \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i} = \prod_{1\le i<j\le n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \\ &= \prod_{1\le i<j\le n} \left(\frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \cdot \frac{j-i}{\tau(j)-\tau(i)}\right) \\ &= \prod_{1\le i<j\le n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \cdot \left(\prod_{1\le i<j\le n} \frac{\tau(j)-\tau(i)}{j-i}\right)^{-1} = \operatorname{sign}(\sigma\tau)\cdot\operatorname{sign}(\tau)^{-1}.\end{aligned}$$

Hieraus folgt sofort die Behauptung $\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau)$. □

Definition 11.11. Ein **Gruppenhomomorphismus** von einer Gruppe G nach einer Gruppe H ist eine Abbildung $\varphi : G \to H$, so daß für alle $x, y \in G$ gilt
$$\varphi(xy) = \varphi(x)\varphi(y).$$

Beispiel 11.12. Die Gruppe der Vorzeichen $\{\pm 1\} = \{1, -1\}$ mit Multiplikation wie in $\{\pm 1\} \subseteq \mathbb{R}^\times$ ist eine Gruppe mit 2 Elementen. Das Signum ist nach Satz 11.10 ein Gruppenhomomorphismus
$$\operatorname{sign} : S_n \to \{\pm 1\}.$$
Nach Beispiel 11.4 ist $\operatorname{sign}((1,2)) = -1$. Demnach ist das Signum für $n \geq 2$ surjektiv.

Korollar 11.13. *Wenn $\sigma \in S_n$ sich als Produkt von r Transpositionen schreiben läßt, dann ist*
$$\operatorname{sign}(\sigma) = (-1)^r.$$

Beweis. Das folgt sofort aus der Homomorphie aus Satz 11.10 und aus Korollar 11.9. □

Lemma 11.14. *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Es bezeichne 1 das neutrale Element sowohl von G als auch[a] von H. Dann gilt:*
(1) $\varphi(1) = 1$.
(2) Für alle $x \in G$ gilt: $\varphi(x^{-1}) = \varphi(x)^{-1}$.

[a] Wenn man eine Notation absichtlich ein wenig ungenau verwendet spricht man von Notationsmißbrauch. Dieser wird in der Regel dadurch gerechtfertigt, daß die Notation einfacher lesbar wird, und das mögliche Mißverständnis nicht die naheliegende Interpretation der Notation darstellt. Die Formeln werden somit richtig verstanden, bevor die fehlerhafte Notation auffällt.

Beweis. (1) Es gilt $\varphi(1) = \varphi(1 \cdot 1) = \varphi(1) \cdot \varphi(1)$. Multiplikation mit $\varphi(1)^{-1}$ liefert $1 = \varphi(1)$.
(2) Aus $x^{-1}x = 1$ und (1) folgt
$$1 = \varphi(1) = \varphi(x^{-1}x) = \varphi(x^{-1})\varphi(x),$$
und das bedeutet gerade $\varphi(x)^{-1} = \varphi(x^{-1})$. □

Bemerkung 11.15. Aus Lemma 11.14 erhalten wir einen neuen Beweis für Korollar 11.9: alle Transpositionen sind ungerade. Sei $\tau = (a, b)$ eine beliebige Transposition in S_n. Als Übungsaufgabe überlege man sich, daß es ein $\sigma \in S_n$ gibt mit $\sigma(1) = a$ und $\sigma(2) = b$. Dann gilt
$$\sigma(1,2)\sigma^{-1} = (a, b)$$
und daher mittels Satz 11.10, Lemma 11.14 und Beispiel 11.4
$$\operatorname{sign}(\tau) = \operatorname{sign}(\sigma(1,2)\sigma^{-1}) = \operatorname{sign}(\sigma)\operatorname{sign}((1,2))\operatorname{sign}(\sigma)^{-1} = \operatorname{sign}((1,2)) = -1.$$

Definition 11.16. Ein Charakter auf einer Gruppe G mit Werten in einem Körper K ist ein Gruppenhomomorphismus
$$\chi : G \to K^\times.$$

Beispiel 11.17. Das Signum definiert für jeden Körper K mit den Interpretationen von $1, -1 \in K$ einen Charakter
$$\operatorname{sign} : S_n \to K^\times.$$
Das Bild von sign in K besteht für $n \geq 2$ aus $\{1, -1\} \subseteq K^\times$. Allerdings gibt es Körper, etwa $\mathbb{F}_2$, für die $1 = -1$. Das passiert genau dann, wenn $2 = 0$ in K gilt. Man spricht dann davon, daß der Körper die Charakteristik 2 hat. Aber selbst in diesem Fall, wenn man also die Werte 1 und -1 in K nicht mehr unterscheiden kann, gelten die im folgenden entwickelten Formeln zur Determinante mit dieser Interpretation des Signum-Charakters.

Beispiel 11.18. Ein Rubic's Cube 3er-Würfel hat 12 Kantenwürfel mit jeweils 2 Kacheln. Die Drehungen des Würfels haben einen Effekt durch Permutieren dieser 24 Kacheln. Jede Drehung führt dabei zwei 4-er Ringtausche durch, man spricht jeweils von einem 4-Zykel. Das Signum von einem 4-Zykel ist -1, und das Signum von zwei solchen 4-Zyklen ist dann das Produkt, also 1. Man kann also in dieser S_{24}, der Permutationsgruppe der Kantenkacheln, nur gerade Permutationen ausführen. Daher kann es auch nicht sein, daß man beim Lösen des Zauberwürfels am Ende vor der Aufgabe steht, daß nur ein Kachelpaar, also ein Kantenwürfel, gekippt werden mußẆenn Sie einen solchen Rubic's Cube beobachten, dann haben Sie bewiesen, daß dieses Exemplar falsch zusammengebaut worden ist.

11.2. **Multilineare Abbildungen.** Auch in der linearen Algebra treten nichtlineare Funktionen auf, zum Beispiel die Determinante.

Definition 11.19. Sei K ein Körper und $n \in \mathbb{N}$. Wir definieren **die Determinante**

$$\det : \mathrm{M}_n(K) \to K$$

durch die **Leibniz–Formel**

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}. \tag{11.1}$$

Hier ist $\operatorname{sign} : S_n \to \{\pm 1\} \subseteq K^\times$ der Charakter, welcher durch das Signum gegeben ist.

Beispiel 11.20. Im Fall $n = 2$ liefert die Leibniz–Formel (11.1) die Determinante

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \operatorname{sign}(\mathrm{id})ad + \operatorname{sign}((1,2))bc = ad - bc.$$

Im Fall $n = 3$ liefert die Leibniz–Formel (11.1) eine als **Sarrus'sche Regel** bekannte Formel. Dazu ergänzt man die Spalten 1 und 2 erneut als vierte und fünfte Spalte:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix}$$

In der entstandenenen 3×5-Matrix gibt es drei nach rechts unten gerichtete Diagonalen,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix}$$

die man jeweils aufmultipliziert und mit dem Vorzeichen $+$ versieht, sowie drei nach links unten gerichtete Diagonalen,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix}$$

die jeweils aufmultipliziert und mit dem Vorzeichen $-$ versehen werden. Die Summe dieser Terme ergibt die Determinante der Matrix (a_{ij}) nach der Leibniz–Formel.

Bemerkung 11.21. Schreibt man die Matrix $A = [v_1, \ldots, v_n] \in \mathrm{M}_n(K)$ in Spaltenschreibweise, so kann man die Determinante als eine Funktion

$$\det(v_1, \ldots, v_n) := \det([v_1, \ldots, v_n]) \in K$$

auffassen, in die man n Vektoren aus V einsetzt:

$$\det : \underbrace{K^n \times \ldots \times K^n}_{n\text{-mal}} \to K.$$

Bemerkung 11.22. Zur Bedeutung der Determinante in der Geometrie betrachten wir den Fall $K = \mathbb{R}$. Für reelle Matrizen $A \in \mathrm{M}_n(\mathbb{R})$ werden meßbare Teilmengen $X \subseteq \mathbb{R}^n$ durch die lineare Abbildung $L_A : \mathbb{R}^n \to \mathbb{R}^n$ derart abgebildet, daß sich das Volumen mit dem Betrag der Determinante multipliziert:

$$\mathrm{vol}(L_A(X)) = |\det(A)| \cdot \mathrm{vol}(X).$$

Das Vorzeichen von $\det(A)$ gibt an, ob L_A die Orientierung erhält oder wechselt. Orientierung formalisiert den Unterschied zwischen einer rechten und einer linken Hand.

Definition 11.23. Sei K ein Körper, und seien $V_1, \ldots, V_n$ und W Vektorräume über K. Eine Abbildung

$$d : V_1 \times \ldots \times V_n \to W$$

(1) ist **multilinear**, wennn für alle $s = 1, \ldots, n$ und alle $v_1, \ldots, v_{s-1}, v_{s+1}, \ldots, v_n$ mit $v_i \in V_i$ durch

$$x \mapsto f(x) = d(v_1, \ldots, v_{s-1}, x, v_{s+1}, \ldots, v_n)$$

für $x \in V_s$ eine lineare Abbildung $f : V_s \to W$ definiert wird (d ist linear im s-ten Argument bei fixierten restlichen Argumenten).

Wenn $W = K$ und $V = V_1 = \ldots = V_n$, dann nennt man die multilineare Abbildung eine **Multilinearform** auf V.

(2) Sei nun $V = V_1 = \ldots = V_n$. Dann ist d **alternierend**, wenn für alle $1 \leq a < b \leq n$ und alle $v_1, \ldots, v_n \in V$ mit $v_a = v_b$ gilt

$$d(v_1, \ldots, v_n) = 0.$$

Beispiel 11.24. Multilineare Abbildungen werden systematisch in der Vorlesung *Lineare Algebra 2* studiert. Hier kommen im Vorgriff ein paar Beispiele, die wir jetzt schon leicht beschreiben können und die bereits eine Rolle gespielt haben.

(1) Das Standardskalarprodukt ist eine multilineare Abbildung

$$K^n \times K^n \to K, \qquad (x, y) \mapsto x \bullet y.$$

In diesem Fall gibt es zwei Argumente und man spricht von einer **bilinearen** Abbildung. Weil das Ziel der Vektorraum $K = K^1$ ist, spricht man genauer von einer **Bilinearform** auf dem Vektorraum K^n. Allgemeiner: eine Bilinearform auf einem K-Vektorraum V ist eine bilineare Abbildung

$$V \times V \to K.$$

(2) Sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Dann definiert für $x \in K^n$ und $y \in K^m$

$$(x, y)_A := x^t A y = x \bullet Ay$$

eine bilineare Abbildung

$$(-,-)_A : K^n \times K^m \to K.$$

Man kann sich leicht überlegen, daß jede bilineare Abbildung $V \times W \to K$ nach Wahl von Koordinaten diese Form hat. Zum Beispiel gehört das Standardskalarprodukt auf K^n zur Einheitsmatrix $\mathbf{1}_n \in \mathrm{M}_n(K)$ Insbesondere kann man das Wissen über Matrizen erneut einsetzen, wenn es darum geht, allgemeine bilineare Abbildungen zu verstehen.

(3) Sei V ein K-Velktorraum und V^* sein Dualraum. Die Auswertung liefert eine Bilinearform

$$V^* \times V \to K, \quad (\pi, x) \mapsto \pi(x).$$

Sei $\mathscr{B}$ eine Basis von V und $\mathscr{B}^*$ die zugehörige duale Basis. Dann gilt

$$\pi(x) = \kappa_{\mathscr{B}^*}(\pi) \bullet \kappa_{\mathscr{B}}(x)$$

nach Korollar 10.15. Bezüglich $\mathscr{B}$ und $\mathscr{B}^*$ ist die Auswertung in Koordinaten nichts anderes als das Standardskalarprodukt.

(4) Wir sehen gleich, daß die Determinante eine multilineare Abbildung ist mit $V_1 = \ldots = V_n = K^n$ und $W = K$.

Proposition 11.25. *Die Abbildung* det *ist multilinear in den Spalten: Sei* $1 \leq s \leq n$*, und seien* $v_1, \ldots, v_{s-1}, v_{s+1}, \ldots, v_n$ *in* K^n *fest gewählt. Dann ist die Abbildung*

$$f : K^n \to K, \quad f(x) = \det(v_1, \ldots, v_{s-1}, x, v_{s+1}, \ldots, v_n)$$

eine K*-lineare Abbildung.*

Beweis. Sei $A = [v_1, \ldots, v_n] = (a_{ij})$ mit $v_s = x$ beliebig. In der Tat, wenn wir die Spalten $\neq s$ fixieren und für jedes $k = 1, \ldots, n$ setzen

$$c_k = \sum_{\sigma \in S_n \text{ mit } \sigma(k)=s} \operatorname{sign}(\sigma) \prod_{i=1,\ i\neq k}^{n} a_{i\sigma(i)},$$

dann hängt c_k nicht von der s-ten Spalte

$$x = \begin{pmatrix} a_{1s} \\ \vdots \\ a_{ns} \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ab. Als Funktion der s-ten Spalte x ist

$$\begin{aligned} f(x) = \det((a_{ij})) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \\ &= \sum_{k=1}^{n} \Big(\sum_{\sigma \in S_n \text{ mit } \sigma(k)=s} \operatorname{sign}(\sigma) \prod_{i=1,\ i\neq k}^{n} a_{i\sigma(i)} \Big) a_{ks} = \sum_{k=1}^{n} c_k a_{ks} = \sum_{k=1}^{n} c_k x_k \end{aligned}$$

offensichlich linear. □

Proposition 11.26. *Die Abbildung* det *ist alternierend.*

Beweis. Seien die Spalten mit den Indizes $\alpha \neq \beta$ identisch, also mit der Transposition $\tau = (\alpha, \beta)$ gilt für alle i, j

$$a_{ij} = a_{i\tau(j)}. \tag{11.2}$$

Wir betrachten auf S_n die folgende Äquivalenzrelation[3]

$$\sigma \sim \sigma' \iff \sigma = \sigma' \text{ oder } \sigma = \tau\sigma'.$$

Dies ist offensichtlich reflexiv. Wegen $\tau\tau = \mathrm{id}$ und dem daraus resultierenden

$$\sigma' = \tau\sigma \iff \sigma = \tau\sigma'$$

ist $\sim$ auch symmetrisch. Ebenso durch Fallunterscheidung zeigt man leicht die Transitivität. Die Äquivalenzklassen bestehen jeweils aus 2 Elementen σ und $\tau\sigma$, die nicht gleich sind(!):

$$\sigma = \tau\sigma \implies \mathrm{id} = \tau \implies \text{Widerspruch!}$$

Wir sortieren die Summe über alle $\sigma \in S_n$ in der Formel (11.1) nun so, daß wir zunächst über die Permutationen einer Äquivalenzklasse summieren. Wir bringen also den Term für σ und den Term für $\tau\sigma$ zusammen. Das Konzept der Äquivalenzklassen hilft einzusehen, daß bei diesem

[3] Das ist die Relation mit Nebenklassen nach der Untergruppe $\langle\tau\rangle = \{\mathrm{id}, \tau\} \subseteq S_n$ als Äquivalenzklassen, also ein Spezialfall einer allgemeinen Konstruktion.

Umsortieren jeder Summand nur einmal benötigt wird (das ist die paarweise Disjunktheit der Äquivalenzklassen). Also

$$\begin{aligned}\sum_{\pi\in\{\sigma,\tau\sigma\}} \operatorname{sign}(\pi)\prod_{i=1}^n a_{i\pi(i)} &= \operatorname{sign}(\sigma)\prod_{i=1}^n a_{i\sigma(i)} + \operatorname{sign}(\tau\sigma)\prod_{i=1}^n a_{i\tau\sigma(i)}\\ &= \operatorname{sign}(\sigma)\prod_{i=1}^n a_{i\sigma(i)} + \operatorname{sign}(\tau\sigma)\prod_{i=1}^n a_{i\sigma(i)} \qquad (11.2)\\ &= (\operatorname{sign}(\sigma)+\operatorname{sign}(\tau)\operatorname{sign}(\sigma))\cdot\prod_{i=1}^n a_{i\sigma(i)} = 0, \qquad (\text{Satz } 11.10)\end{aligned}$$

weil $\operatorname{sign}(\tau) = -1$. Somit summieren in $\det(A)$ die Beiträge zu σ und $\tau\sigma$ zu 0, folglich $\det(A) = 0$ und det ist in der Tat alternierend. □

11.3. **Determinantenfunktionen.** Wir studieren nun die Determinante systematisch.

Definition 11.27. Sei K ein Körper, und sei V ein K-Vektorraum der Dimension $n \in \mathbb{N}_0$. Eine **Determinantenfunktion** auf V ist eine multilineare und alternierende Abbildung

$$d : \underbrace{V \times \ldots \times V}_{n\text{-mal}} \to K,$$

(die Anzahl der Faktoren V ist gleich $\dim(V)$).

Bemerkung 11.28. Eine Determinantenfunktion auf K^n darf man als eine Abbildung

$$d : \mathrm{M}_n(K) \to K, \qquad d([v_1,\ldots,v_n]) = d(v_1,\ldots,v_n)$$

interpretieren, indem man die n-vielen Spaltenvektoren zu einer $n\times n$ Matrix zusammenfaßt. Aus den definierenden Eigenschaften werden dann in Bezug auf Matrizen die folgenden Eigenschaften:

(i) d ist **multilinear** in den Spalten.
(ii) d ist **alternierend:** $d(A) = 0$, falls $A \in \mathrm{M}_n(K)$ zwei identischen Spalten hat.

Proposition 11.23 und Proposition 11.26 zeigen, daß durch (11.1) eine Determinantenfunktion det auf K^n definiert ist.

Satz 11.29. *Sei V ein K-Vektorraum der Dimension n. Eine Determinantenfunktionen d auf V skaliert sich unter Permutation der Argumente mit dem Charakter* $\operatorname{sign} : S_n \to K^\times$*, d.h. für alle $\sigma \in S_n$ und $v_1,\ldots,v_n \in V$ gilt*

$$d(v_{\sigma(1)},\ldots,v_{\sigma(n)}) = \operatorname{sign}(\sigma)\cdot d(v_1,\ldots,v_n).$$

Beweis. Der Beweis hat drei Schritte.

Schritt 1: Wir beweisen den Satz zunächst für eine Transposition $\tau = (a,b)$ mit $a < b$. Für feste $v_i \in V$ für $i \neq a, b$ kürzen wir ab: für $v, w \in V$ sei

$$f(v,w) = d(v_1,\ldots,v_{a-1},v,v_{a+1},\ldots,v_{b-1},w,v_{b+1},\ldots,v_n).$$

Dann gilt wegen multilinear und alternierend

$$\begin{aligned}0 = f(v+w,v+w) &= f(v,v+w) + f(w,v+w)\\ &= f(v,v) + f(v,w) + f(w,v) + f(w,w)\\ &= f(v,w) + f(w,v).\end{aligned}$$

Also folgt wie behauptet:

$$f(v,w) = -f(w,v) = \operatorname{sign}(\tau)\cdot f(w,v).$$

Schritt 2: Jetzt nehmen wir an, daß der Satz bereits für Permutationen σ und $\pi \in S_n$ und **alle** $v_1, \ldots, v_n \in V$ bewiesen ist, und zeigen dann die Gültigkeit auch für die Komposition $\sigma\pi$. Dazu setzen wir $w_i = v_{\sigma(i)}$ für $i = 1, \ldots, n$ und rechnen:

$$\begin{aligned} d(v_{\sigma\pi(1)}, \ldots, v_{\sigma\pi(n)}) &= d(v_{\sigma(\pi(1))}, \ldots, v_{\sigma(\pi(n))}) = d(w_{\pi(1)}, \ldots, w_{\pi(n)}) \\ &= \operatorname{sign}(\pi) \cdot d(w_1, \ldots, w_n) = \operatorname{sign}(\pi) \cdot d(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \\ &= \operatorname{sign}(\pi) \cdot \operatorname{sign}(\sigma) \cdot d(v_1, \ldots, v_n) \\ &= \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\pi) \cdot d(v_1, \ldots, v_n) = \operatorname{sign}(\sigma\pi) \cdot d(v_1, \ldots, v_n). \end{aligned}$$

Schritt 3: Nun sei $\sigma \in S_n$ beliebig. Nach Satz 11.2 ist σ ein Produkt von Transpositionen, sagen wir mit r Faktoren. Wir zeigen nun per Induktion nach der Anzahl r der benötigten Transpositionen die Aussage des Satzes für σ. Wenn man $r = 0$ Transpositionen braucht, dann ist $\sigma = \mathrm{id}$ und der Satz trivial. Ansonsten läßt sich σ als $\sigma = \tau\pi$ schreiben mit einer Transposition τ und einer Permutation π, die ein Produkt von weniger als r Transpositionen ist. Für τ und π gilt der Satz dann wegen Schritt 1 oder der Induktionsannahme. Wegen Schritt 2 folgt der Satz dann auch für σ. □

Bemerkung 11.30. Aus der Aussage von Satz 11.29 kann man für eine multilineare Abbildung d auf alternierend schließen, sofern man in K durch 2 teilen darf. Der Körper K darf also nicht Charakteristik 2 haben, d.h. bei $2 = 0$ in K gilt dieser Schluß nicht. Dazu schauen wir uns den entscheidenden Schritt der Rechnung mit $f(v, w)$ im Beweis von Satz 11.29 genauer an. Alternierend besagt $f(v, v) = 0$ für alle v, während die Aussage von Satz 11.29

$$f(v, w) = -f(w, v)$$

liefert. Setzen wir $w = v$ und bringen beide Terme auf eine Seite, so finden wir $2f(v, v) = 0$. Wenn $2 \in K^\times$, so folgt $f(v, v)$ wie gewünscht.

Definition 11.31. Sei V ein endlichdimensionaler K-Vektorraum. Die Menge der Determinantenfunktionen auf V bezeichnen[a] wir mit

$$\det(V^*).$$

[a] Die Notation für den Raum der Determinantenfunktionen ist so gewählt, weil das in das allgemeine Bezeichnungsschema paßt. Das ist nur ein Name und, daß darin die Notation für den Dualraum vorkommt, soll uns nicht stören.

Wir erinnern an Beispiel 4.9: für eine beliebige Menge X ist die Menge der Abbildungen

$$\mathrm{Abb}(X, K) = \{f : X \to K \ ; \quad \text{Abbildung}\}$$

ein K-Vektorraum. Elemente in $\mathrm{Abb}(X, K)$ kann man als K-Wertige Funktionen auf X ansprechen. Funktionen werden addiert, indem die Funktionswerte addiert werden. Und die Skalarmultiplikation einer Funktion geschieht durch die werteweise Multiplikation mit dem Skalar.

Proposition 11.32. *Sei V ein endlichdimensionaler K-Vektorraum der Dimension n. Dann bilden die Determinantenfunktionen einen K-Unterraum*

$$\det(V^*) \subseteq \mathrm{Abb}(V^{\times n}, K)$$

des K-Vektorraums der Abbildunngen auf der Menge $V^{\times n} := \underbrace{V \times \ldots \times V}_{n\text{-mal}}$.

Beweis. Die definierenden Eigenschaften multilinear und alternierend von Determinantenfunktionen bleiben bei werteweiser Addition und Skalarmultiplikation erhalten. Damit folgt die Aussage aus dem Unterraumkriterium, denn $\det(V^*)$ enthält die konstante Abbildung mit Wert 0 als triviales Beispiel einer Determinantenfunktion. □

Satz 11.33. *Sei K ein Körper. Sei V ein endlichdimensionaler K-Vektorraum. Dann hat der Raum der Determinantenfunktionen auf V die Dimension* 1*:*

$$\dim(\det(V^*)) = 1.$$

Beweis. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V und d eine Determinantenfunktion. Wir zeigen zunächst, daß der Wert

$$d(b_1, \ldots, b_n)$$

die Determinantenfunktion d eindeutig festlegt. Wenn alle $v_i \in \{b_1, \ldots, b_n\}$, dann gibt es entweder eine Permutation σ mit $v_i = b_{\sigma(i)}$, und dann gilt:

$$d(v_1, \ldots, v_n) = d(b_{\sigma(1)}, \ldots, b_{\sigma(n)}) = \operatorname{sign}(\sigma) \cdot d(b_1, \ldots, b_n).$$

Oder aber zwei der v_i sind gleich, woraus $d(v_1, \ldots, v_n) = 0$ folgt. Damit ist d auf Tupeln aus Basisvektoren von $\mathscr{B}$ eindeutig durch $d(b_1, \ldots, b_n)$ festgelegt.

Für allgemeine Tupel $v_1, \ldots, v_n$ arbeiten wir per Induktion über die Anzahl der Vektoren, die nicht aus $\mathscr{B}$ sind:

$$r = |\{i \ ; \ v_i \notin \{b_1, \ldots, b_n\}\}|$$

Den Induktionsanfang $r = 0$ haben wir bereits erledigt, denn dann sind alle v_i Basisvektoren der Basis $\mathscr{B}$. Wenn $r \geq 1$, dann gibt es ein j mit v_j ist kein Basisvektor der Basis $\mathscr{B}$. Dieses v_j läßt sich aber als Linearkombination

$$v_j = \sum_{i=1}^{n} x_i b_i$$

für gewisse $x_i \in K$ ausdrücken. Dann gilt

$$d(v_1, \ldots, v_n) = \sum_{i=1}^{n} x_i \cdot d(v_1, \ldots, \underset{\uparrow j}{b_i}, \ldots, v_n)$$

und weil auf der rechten Seite v_j durch b_i als Argument von d ersetzt wurde, sind die Werte von d auf der rechten Seite bereits per Induktionsannahme bekannt.

Damit ist gezeigt, daß die Auswertung

$$\det(V^*) \to K, \qquad d \mapsto d(b_1, \ldots, b_n)$$

injektiv ist. Per Definition der Vektorraumstruktur auf dem Raum der Determinantenfunktionen handelt es sich um eine lineare Abbildung. Damit ist

$$\dim(\det(V^*)) \leq \dim(K) = 1.$$

Es bleibt nun zu zeigen, daß es eine von 0 verschiedene Determinantenfunktion gibt. Dafür reicht der Fall $V = K^n$ (der uns im Grunde hauptsächlich interessiert), denn mittels eines Isomorphismus $f : V \xrightarrow{\sim} K^n$ und einer Determinantenfunktion $d \neq 0$ auf K^n bekommt man durch

$$(f^*d)(v_1, \ldots, v_n) = d(f(v_1), \ldots, f(v_n))$$

eine Determinantenfunktion f^*d auf V.

Definition 11.19 liefert eine Determinantenfunktion det auf K^n. Es bleibt nur zu zeigen, daß dies eine nichttriviale Determinantenfunktion ist. Dazu werten wir (11.1) auf der Einheitsmatrix aus, also für $a_{ij} = \delta_{ij}$. Dann bleibt einzig der Summand zu $\sigma = \mathrm{id}$ bestehen, alle anderen Produkte enthalten einen Faktor 0. Somit ist

$$\det(\mathbf{1}_n) = \operatorname{sign}(\mathrm{id}) \prod_{i=1}^{n} \delta_{ii} = 1. \qquad \square$$

Korollar 11.34. *Sei K ein Körper und $n \in \mathbb{N}$. Die Determinante*

$$\det : \mathrm{M}_n(K) \to K$$

ist die eindeutige Determinantenfunktion auf K^n, die auf der Standardbasis $(e_1, \dots, e_n)$ den Wert 1 annimmt, also

$$\det(e_1, \dots, e_n) = \det(\mathbf{1}_n) = 1.$$

Beweis. Satz 11.33 und genauer der Beweis zeigt, daß es zu einer Basis $\mathscr{B} = (b_1, \dots, b_n)$ eines Vektorraums V genau eine Determinantenfunktion d auf V mit dem Wert $d(b_1, \dots, b_n) = 1$. □

11.4. **Eigenschaften der Determinante.** Wir illustrieren die Macht von Satz 11.33 durch den folgenden eleganten Beweis des wichtigen Determinanten-Multiplikationssatzes, der mit der expliziten Formel (11.1) nur mit einer Index-Schlacht zu haben ist.

Satz 11.35 (Determinanten-Multiplikationssatz)**.** *Seien K ein Körper und $n \in \mathbb{N}$. Dann gilt für alle $A, B \in \mathrm{M}_n(K)$*

$$\det(AB) = \det(A) \cdot \det(B).$$

Beweis. Die Abbildung $\delta : \mathrm{M}_n(K) \to K$ gegeben durch

$$\delta(X) := \det(AX)$$

ist eine Determinantenfunktion. In der Tat, wenn in Spaltenschreibweise $X = [x_1, \dots, x_n]$, dann ist

$$\delta(X) = \det([Ax_1, \dots, Ax_n]).$$

Weil Multiplikation mit A linear ist, ist $\delta(X)$ multilinear in den Spalten. Und alternierend ist $\delta(X)$ offensichtlich auch. Damit gibt es nach Satz 11.33 ein $c \in K$, so daß für alle $X \in \mathrm{M}_n(K)$:

$$\delta(X) = c \det(X).$$

Auswertung bei der Einheitsmatrix $X = \mathbf{1}_n$ berechnet c zu

$$c = c \det(\mathbf{1}_n) = \delta(\mathbf{1}_n) = \det(A\mathbf{1}_n) = \det(A).$$

Ausgewertet in B erhalten wir

$$\det(AB) = \delta(B) = c \det(B) = \det(A) \cdot \det(B).$$ □

Proposition 11.36. *Sei V ein K-Vektorraum der Dimension n, und seien $v_1, \dots, v_n \in V$.*

(1) Wenn $v_j = 0$, dann ist

$$\det(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_n) = 0.$$

(2) Wenn $v_1, \dots, v_n$ linear abhängig sind, dann ist

$$\det(v_1, \dots, v_n) = 0.$$

Beweis. (1) Das folgt sofort, weil det als Funktion des j-ten Eintrags linear ist und als solche 0 auf 0 abbildet.

(2) Wenn $v_1, \dots, v_n$ linear abhängig sind, dann gibt es ein j und $x_i \in K$ für $i \neq j$ mit $v_j = \sum_{i \neq j} x_i v_i$. Dann gilt

$$\begin{aligned} \det(v_1, \dots, v_n) &= \det(v_1, \dots, v_{j-1}, \sum_{i \neq j} x_i v_i, v_{j+1}, \dots, v_n) \\ &= \sum_{i \neq j} x_i \det(v_1, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) = 0, \end{aligned}$$

weil in jedem Summanden der Vektor v_i zweimal als Argument in det auftritt. □

Satz 11.37. *Sei $A \in \mathrm{M}_n(K)$. Dann gilt*

$$A \text{ invertierbar} \iff \det(A) \neq 0.$$

Mit anderen Worten

$$\mathrm{GL}_n(K) = \{A \in \mathrm{M}_n(K) \,;\, \det(A) \in K^\times\}.$$

Beweis. Wenn A invertierbar ist, dann gibt es $B \in \mathrm{M}_n(K)$ mit $AB = \mathbf{1}_n$. Satz 11.35 liefert dann

$$\det(A) \cdot \det(B) = \det(AB) = \det(\mathbf{1}_n) = 1.$$

Folglich ist $\det(A) \neq 0$.

Für die umgekehrte Richtung zeigen wir das logisch äquivalente

$$A \text{ nicht invertierbar} \implies \det(A) = 0.$$

Wenn A nicht invertierbar ist, dann ist $\mathrm{rg}(A) < n$. Also sind die Spalten von A linear abhängig. Damit folgt $\det(A) = 0$ direkt aus Proposition 11.36. □

Korollar 11.38. *Sei K ein Körper und $n \in \mathbb{N}$. Dann ist*

$$\det : \mathrm{GL}_n(K) \to K^\times$$

ein Gruppenhomomorphismus. Insbesondere gilt für alle $A \in \mathrm{GL}_n(K)$

$$\det(A^{-1}) = \det(A)^{-1}.$$

Beweis. Aus Satz 11.37 folgt, daß $\det(A)$ für $A \in \mathrm{GL}_n(K)$ Werte in $K^\times$ annimmt, also det im Sinne des Korollars wohldefiniert ist. Die Gruppenhomomorphismuseigenschaft folgt aus dem Determinantenmultiplikationssatz, Satz 11.35.

Die Aussage über $\det(A^{-1})$ ist eine allgemeine Aussage über Gruppenhomomorphismen, siehe Lemma 11.14. □

Der Begriff der Determinante überträgt sich von Matrizen über die Darstellungsmatrix auf Endomorphismen.

Definition 11.39. Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums von endlicher Dimension. Die **Determinante** von f ist

$$\det(f) = \det\left(\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)\right) \in K.$$

Dabei ist $\mathscr{B}$ eine Basis von V und $\det(f)$ ist unabhängig von der Wahl der Basis $\mathscr{B}$.

Beweis. Wir müssen zeigen, daß $\det(f)$ nicht von der Wahl der Basis $\mathscr{B}$ abhängt. Sei

$$A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$$

eine Darstellungsmatrix und bezüglich der Basis $\mathscr{B}'$

$$A' = \mathrm{M}_{\mathscr{B}'}^{\mathscr{B}'}(f)$$

eine weitere. Mit der Basiswechselmatrix

$$S = S_{\mathscr{B}'}^{\mathscr{B}} = \mathrm{M}_{\mathscr{B}'}^{\mathscr{B}}(\mathrm{id}_V)$$

und der umgekehrten Basiswechselmatrix, siehe Proposition 8.50,

$$S^{-1} = S_{\mathscr{B}}^{\mathscr{B}'} = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}'}(\mathrm{id}_V)$$

gilt, siehe Proposition 8.48,

$$A' = SAS^{-1}.$$

Daraus folgt mit dem Determinantenmultiplikationssatz, Satz 11.35,

$$\det(A') = \det(SAS^{-1}) = \det(S)\det(A)\det(S^{-1}) = \det(S)\det(A)\det(S)^{-1} = \det(A)$$

die Unabhängigkeit der Definition von $\det(f)$ von der Wahl der Basis. □

Bemerkung 11.40. Für die Berechnung der Determinante eines Endomorphismus $f : V \to V$ als Determinante der Darstellungsmatrix ist es essentiell, daß man in Definitionsbereich und Wertebereich, beides V, die **gleiche Basis** benutzt.

Proposition 11.41. *Für alle $A \in \mathrm{M}_n(K)$ gilt*

$$\det(A^t) = \det(A).$$

Beweis. Wir nutzen, daß das Signum nur Werte ± 1 annimmt und somit

$$\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)^{-1} = \operatorname{sign}(\sigma)$$

für alle $\sigma \in S_n$. Außerdem ist der Übergang zum Inversen $\sigma \mapsto \sigma^{-1}$ eine Bijektion auf S_n, so daß man statt über $\sigma \in S_n$ auch über die entsprechenden Inversen σ^{-1} summieren kann.

Damit folgt dann für $A = (a_{ij})$ und $(A^t)_{ij} = a_{ji}$ nach der Leibniz–Formel (11.1)

$$\begin{aligned}\det(A^t) &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^{n} (A^t)_{j\sigma(j)} = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^{n} a_{\sigma(j)j} \\ &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma^{-1}) \prod_{i=1}^{n} a_{i\sigma^{-1}(i)} = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \det(A).\end{aligned}$$ □

Korollar 11.42. *Sei K ein Körper und $n \in \mathbb{N}$.*

(1) Die Determinante det *auf $\mathrm{M}_n(K)$ ist multilinear und alternierend als Funktion in den Zeilen.*

(2) Wenn $A \in \mathrm{M}_n(K)$ linear abhängige Zeilen hat, dann ist $\det(A) = 0$.

(3) Wenn $A \in \mathrm{M}_n(K)$ eine Nullzeile hat, dann ist $\det(A) = 0$.

Beweis. Dies folgt sofort aus $\det(A) = \det(A^t)$ nach Proposition 11.41 und den entsprechende Aussagen bezüglich der Spalten, denn das Transponieren macht Zeilen zu Spalten.

Alternativ kann man für (2) argumentieren, daß linear abhängige Zeilen $\operatorname{rg}(A) < n$ bedeuten. Damit gibt es dann auch linear abhängige Spalten und Proposition 11.36 zeigt $\det(A) = 0$.

Aussage (3) folgt alternativ auch aus der Leibniz-Formel für die Determinante, denn in jedem Summand steckt ein Faktor aus der Nullzeile. □

Beispiel 11.43. Wir berechnen einige nützliche Determinanten.

(1) Sei $A = (a_{ij}) \in \mathrm{M}_n(K)$ eine obere Dreiecksmatrix, und seien $a_{ii} = \alpha_i$ die Diagonaleinträge.

$$\det(A) = \det \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \alpha_n \end{pmatrix} = \prod_{i=1}^{n} \alpha_i.$$

In der Tat fallen in der Leibniz–Formel alle Terme zu Permutationen σ weg, in denen für ein i gilt $i > \sigma(i)$, denn solche sorgen für einen Faktor 0. Die verbleibenden Permutationen σ haben für alle $i = 1, \ldots, n$ die Eigenschaft $\sigma(i) \geq i$. Das erfüllt nur $\sigma = \mathrm{id}$. Der entsprechende Summand in der Leibniz–Formel ist das Produkt über die Diagonalterme.

(2) Für alle $\alpha \neq \beta$ und $\lambda \in K$ gilt für die entsprechende Elementarmatrix

$$\det\big(E_{\alpha\beta}(\lambda)\big) = 1.$$

Wenn $\alpha < \beta$, dann sind wir im Fall einer oberen Dreiecksmatrix und dort speziell mit nur 1 auf der Diagonalen. Im Fall $\alpha > \beta$ kann man analog schließen, oder man kann sich Proposition 11.41 zu Nutze machen und das offensichtliche

$$E_{\alpha\beta}(\lambda) = E_{\beta\alpha}(\lambda)^t.$$

(3) Für $1 \leq \alpha \leq n$ und $\mu \in K^\times$ gilt

$$\det(E_{\alpha\alpha}(\mu)) = \mu.$$

Das ist nur ein Spezialfall von (1).

(4) Für alle $\alpha \neq \beta$ gilt

$$\det\big(P_{\alpha\beta}\big) = \det\Big(E_{\beta\beta}(-1)E_{\alpha\beta}(1)E_{\beta\alpha}(-1)E_{\alpha\beta}(1)\Big) = -1 \cdot 1 \cdot 1 \cdot 1 = -1$$

mit Hilfe von (2) und (3) und dem Determinantenmultiplikationssatz, Satz 11.35.

Bemerkung 11.44. Zur Permutation $\sigma \in S_n$ definieren wir eine Permutationsmatrix $P_\sigma \in \mathrm{M}_n(K)$ als

$$P_\sigma = [e_{\sigma(1)}, \dots, e_{\sigma(n)}],$$

also als die Matrix, deren Linksmultiplikation die Standardbasis $\mathscr{E} = (e_1, \dots, e_n)$ wie σ permutiert: $P_\sigma e_j = e_{\sigma(j)}$. Es folgt für $\sigma, \pi \in S_n$ und $j = 1, \dots, n$:

$$P_\sigma P_\pi e_j = P_\sigma e_{\pi(j)} = e_{\sigma\pi(j)} = P_{\sigma\pi} e_j,$$

also $P_\sigma P_\pi = P_{\sigma\pi}$. Weiter gilt:

$$\det(P_\sigma) = \operatorname{sign}(\sigma),$$

denn die Formel stimmt für Transpositionen, weil $P_{\alpha\beta}$ die Permutationsmatrix zur Transposition (α, β) ist. Außerdem sind beide Seiten multiplikativ, und jede Permutation ist Produkt von Transpositionen.

Die Linksmultiplikation mit P_σ vertauscht die Zeilen einer Matrix wie es die Permutation σ angibt. Es reicht dies bei einem Spaltenvektornachzurechnen und dort gilt

$$P_\sigma \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P_\sigma(\sum_{i=1}^n x_i e_i) = \sum_{i=1}^n x_i e_{\sigma(i)} = \sum_{i=1}^n x_{\sigma^{-1}(i)} e_i = \begin{pmatrix} x_{\sigma^{-1}(1)} \\ \vdots \\ x_{\sigma^{-1}(n)} \end{pmatrix}.$$

Man darf sich nicht am σ^{-1} in der Formel stören: das bedeutet, daß in der Eintrag x_i aus der i-ten Zeile von x in $P_\sigma x$ nun in der $\sigma(i)$-ten Zeile auftritt, weil $\sigma^{-1}(\sigma(i)) = i$ ist.

Im allgemeinen Fall ist die Leibniz–Formel nicht geeignet, die Determinante einer Matrix zu berechnen. Gerade wenn man sowieso mit dem Gaußschen Eliminationsverfahren eine Zeilenstufenform berechnet hat, dann kommt man mit dem folgenden Satz viel besser an die Determinante.

Proposition 11.45. *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix, sei R die aus dem Gaußschen Eliminationsverfahren berechnete Zeilenstufenform. Sei t die Anzahl der im Verfahren nötigen Zeilenvertauschungen und seien α_i bei $1 \leq i \leq r = \mathrm{rg}(A)$ die auftretenden Einträge in der i-ten Zeile und i-ten Pivotspalte. Dann ist*

$$\det(A) = \begin{cases} (-1)^t \cdot \prod_{i=1}^n \alpha_i & \text{wenn } r = n \\ 0 & \text{wenn } r < n. \end{cases}$$

Beweis. Aus dem Gaußschen Eliminationsverfahren, siehe Theorem 9.28, bekommt man eine Matrix S, die ein Produkt von t Matrizen der Form $P_{\alpha\beta}$ und ansonsten der Form $E_{\alpha\beta}(\lambda)$ mit $\alpha \neq \beta$ ist. Ferner gilt $R = SA$. Nach dem Determinantenmultiplikationssatz, Satz 11.35, und Beispiel 11.43 folgt

$$\det(R) = \det(S) \cdot \det(A) = (-1)^t \det(A).$$

Wenn $r = \operatorname{rg}(A) < n$, dann hat R eine Nullzeile und $\det(R) = 0$ nach Korollar 11.42. Wenn $r = n$, dann ist R eine obere Dreiecksmatrix wie in Beispiel 11.43 (1) und $\det(R) = \prod_{i=1}^{n} \alpha_i$. Daraus folgt die Behauptung. □

Bemerkung 11.46. Man merke sich:

- Elementare Zeilenumformungen, bei denen man das Vielfache einer Zeile zu einer anderen Zeile addiert, ändern die Determinante nicht.
- Eine Zeilenvertauschung multipliziert die Determinante mit -1.
- Das Skalieren einer Zeile mit einem Faktor μ multipliziert die Determinante entsprechend um denselben Faktor μ.

Durch Transponieren erhält man die entsprechenden elementaren Spaltenoperationen, bezüglich derer die Determinante sich dann genauso verhält.

Die **Cramersche Regel** löst lineare Gleichungssysteme durch eine explizite Formel, sofern die Voraussetzungen erfüllt sind.

Proposition 11.47 (Cramersche Regel)**.** *Seien K ein Körper, $A = [v_1, \dots, v_n] \in \mathrm{M}_n(K)$ und $b \in K^n$. Sei*

$$A_j = [v_1, \dots, v_{j-1}, \underset{\uparrow j}{b}, v_{j+1}, \dots, v_n]$$

die Matrix A mit j-ter Spalte ersetzt durch b.

Wenn $\det(A) \neq 0$, dann hat $AX = b$ nur eine Lösung, und zwar für alle $1 \leq j \leq n$ gegeben durch

$$x_j = \frac{\det(A_j)}{\det(A)}.$$

Beweis. Wegen $\det(A) \neq 0$ ist A invertierbar nach Satz 11.37, und $AX = b$ hat eine eindeutige Lösung. Diese Lösung sei $x \in K^n$. Das bedeutet

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{und} \quad b = x_1 v_1 + \ldots + x_n v_n.$$

Einsetzen in die j-te Spalte von A_j liefert

$$\begin{aligned} \det(A_j) &= \det([v_1, \dots, v_{j-1}, \sum_{i=1}^{n} x_i v_i, v_{j+1}, \dots, v_n]) \\ &= \sum_{i=1}^{n} x_i \det([v_1, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n]) \\ &= x_j \det([v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n]) = x_j \det(A), \end{aligned}$$

weil in allen anderen Summanden zwei identische Spalten mit Wert v_i im Argument von det stehen. Wenn nun $\det(A) \neq 0$, dann kann man dies nach x_j auflösen. □

Bemerkung 11.48. Die Cramersche Regel liefert eine Lösungsformel für ein lineares Gleichungssystem. Insbesondere kann man jede Variable unabhängig von den anderen Variablen bestimmen. Das ist nützlich, wenn man nur an dem Wert einer der Variablen interessiert ist.

Andererseits ist die Berechnung von Determinanten aufwendig. Billiger ist es, vor allem wenn bei selber Matrix A bezüglich mehrerer Inhomogenitäten b das lineare Gleichungssystem $AX = b$ gelöst werden soll, wenn man ein für alle Mal eine Zeilenstufenform $R = SA$ bestimmt und dann mit der transformierten Inhomogenität $c = Sb$ durch Rückwärtseinsetzen $RX = c$ löst.

11.5. **Determinante und Cofaktoren.**

Bemerkenswerterweise kann man mittels der Determinante eine Formel für die inverse Matrix angeben, sofern die Matrix invertierbar ist. Dazu brauchen wir zunächst den Laplace'schen Entwicklungssatz, der ein rekursives Verfahren zur Berechnung der Determinante angibt. Dieses ist für manche **dünn** besetzte Matrizen (viele Einträge sind 0) geeignet.

Notation 11.49. Seien $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$ und $1 \leq r \leq n$, $1 \leq s \leq m$. Zu Zeilenindizes $1 \leq i_1 < \ldots < i_r \leq n$ und Spaltenindizes $1 \leq j_1 < \ldots < j_s \leq m$, oder Teilmengen $I = \{i_1, \dots, i_r\} \subseteq \{1, \dots, n\}$ und $J = \{j_1, \dots, j_s\} \subseteq \{1, \dots, m\}$, sei

$$A_{I,J} = (a_{i_\alpha j_\beta})_{1\leq\alpha\leq r, 1\leq\beta\leq s} \in \mathrm{M}_{r\times s}$$

die Matrix mit r Zeilen und s Spalten, die aus A entsteht, wenn man nur die Zeilen mit Index aus I und nur Spalten mit Index aus J berücksichtigt.

Wenn $I = \{i = 1, \dots, n \ ; \ i \neq \alpha\}$ und $J = \{j = 1, \dots, m \ ; \ j \neq \beta\}$, dann kürzen wir ab:

$$A_{\widehat{\alpha\beta}} = A_{I,J}.$$

Definition 11.50. Sei $A = (a_{ij}) \in \mathrm{M}_n(K)$. Dann definieren wir den i,j-ten **Cofaktor** als

$$(-1)^{i+j} \det(A_{\widehat{i,j}})$$

und die **Adjunkte** zu A als die Matrix

$$(A^{\#})_{ij} = (-1)^{j+i} \det(A_{\widehat{j,i}}),$$

also die transponierte Matrix zur Matrix der Cofaktoren.

Satz 11.51. *Sei $A = (a_{ij}) \in \mathrm{M}_n(K)$. Dann gilt*

$$A \cdot A^{\#} = \det(A) \cdot \mathbf{1} = A^{\#} \cdot A.$$

Insbesondere ist für invertierbare A

$$A^{-1} = \frac{1}{\det(A)} \cdot A^{\#}. \tag{11.3}$$

Bemerkung 11.52. (1) Es ist besonders zu betonen, daß die Einträge der inversen Matrix A^{-1} bis auf den gemeinsamen Nenner $\det(A)$ polynomial in den Einträgen von A sind. Der Nenner $\det(A)$ ist überdies auch polynomial.

(2) Speziell erhält man für 2×2-Matrizen:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beweis von Satz 11.51. Wenn wir die s-te Zeile von A mit einer Kopie der t-ten Zeile überschreiben und das Resultat $B = (b_{ij})$ nennen, dann folgt

$$\det(A) \cdot \delta_{ts} = \det(B).$$

Weil die s-te Zeile ausgeblendet wird, ändert sich für die s,k-ten Cofaktoren für $k = 1, \dots, n$ nichts:

$$(-1)^{s+k} \det(A_{\widehat{s,k}}) = (-1)^{s+k} \det(B_{\widehat{s,k}}).$$

Aus Laplace-Entwicklung von $\det(B)$ nach der s-ten Zeile, siehe der folgender Satz 11.53, folgt dann

$$\begin{aligned} \det(A) \cdot \delta_{ts} = \det(B) &= \sum_{k=1}^{n} b_{sk} (-1)^{s+k} \det(B_{\widehat{s,k}}) \\ &= \sum_{k=1}^{n} a_{tk} (-1)^{s+k} \det(A_{\widehat{s,k}}) = \sum_{k=1}^{n} a_{tk} A^{\#}_{ks} = (A \cdot A^{\#})_{ts}. \end{aligned}$$

Daraus folgt

$$A \cdot A^{\#} = \det(A) \cdot \mathbf{1}.$$

Die Formel $A^{\#} \cdot A = \det(A) \cdot \mathbf{1}$ folgt entsprechend durch Überschreiben der t-ten Spalte mit einer Kopie der s-ten Spalte und Entwicklung nach der t-ten Spalte nach Satz 11.53 (oder durch Anwendung auf A^t der ersten Gleichung und anschließendem Transponieren). □

Satz 11.53 (Laplace'scher Entwicklungssatz)**.** *Sei $A = (a_{ij}) \in \mathrm{M}_n(K)$. Dann gilt:*

(1) Entwicklung nach der i-ten Zeile:

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(A_{\widehat{ik}}),$$

(2) Entwicklung nach der j-ten Spalte:

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

Beweis. Wegen $\det(A) = \det(A^t)$ sind die Aussagen (1) und (2) zueinander äquivalent. Wir beweisen daher nur (2). In Proposition 11.25 haben wir die Multilinearität der Leibniz–Formel nachgewiesen durch

$$\det((a_{ij})) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{k=1}^{n} c_k a_{kj}. \tag{11.4}$$

Wir fixieren nun j und k und wollen c_k bestimmen. Die Beziehung zwischen den Matrixeinträgen von A und von $A_{\widehat{kj}}$ wird durch die folgenden Permutationen ausdrückbar. Sei $\tau_{\ell,n} \in S_n$ die Permutation, welche ℓ nach n schickt und alle $i > \ell$ um eins vermindert:

$$\begin{aligned} \tau_{\ell,n} &= \begin{pmatrix} 1 & \cdots & \ell-1 & \ell & \ell+1 & \cdots & n-2 & n-1 & n \\ 1 & \cdots & \ell-1 & n & \ell & \cdots & n-3 & n-2 & n-1 \end{pmatrix} \\ &= (n, n-1)(n-1, n-2) \cdots (\ell+1, \ell). \end{aligned} \tag{11.5}$$

Für $\sigma \in S_n$ mit $\sigma(k) = j$ ist

$$\hat{\sigma} = \tau_{j,n} \sigma \tau_{k,n}^{-1}$$

eine Permutation mit $\hat{\sigma}(n) = n$, die wir als eine Permutation von $\{1, \ldots, n-1\}$ und damit als Element von S_{n-1} auffassen können. Die Abbildung $\sigma \mapsto \hat{\sigma}$ ist eine Bijektion

$$\{\sigma \in S_n \ ; \ \sigma(k) = j\} \xrightarrow{\sim} S_{n-1}.$$

Die inverse Abbildung ordnet einem Element $\pi \in S_{n-1}$ nach Fortsetzung durch $\tilde{\pi}(i) = \pi(i)$ für $i < n$ und $\tilde{\pi}(n) = n$ die Permutation $\tau_{j,n}^{-1} \tilde{\pi} \tau_{k,n}$ zu.

Das Signum von $\hat{\sigma}$ hängt nicht davon ab, ob wir $\hat{\sigma} \in S_n$ oder nach Einschränkung auf $\{1, \ldots, n-1\}$ als Element von S_{n-1} auffassen. Schreiben wir $\hat{\sigma} \in S_{n-1}$ als Produkt von Transpositionen, dann gilt dieselbe Formel auch in S_n. Das Signum wird in beiden Fällen dadurch festgelegt, ob wir eine gerade oder ungerade Anzahl von Transpositionen im Produkt haben. Weil $\tau_{\ell,n}$ als Produkt von $n - \ell$ Transpositionen geschrieben werden kann, siehe (11.5), gilt

$$\operatorname{sign}(\hat{\sigma}) = \operatorname{sign}(\tau_{j,n}) \cdot \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau_{k,n}^{-1}) = (-1)^{(n-j)+(n-k)} \cdot \operatorname{sign}(\sigma) = (-1)^{j+k} \cdot \operatorname{sign}(\sigma).$$

Es gilt nun für $i \neq k$ und $\ell \neq j$:

$$(A_{\widehat{kj}})_{\tau_{k,n}(i), \tau_{j,n}(\ell)} = a_{i\ell},$$

und wegen $\ell = \sigma(i)$ genau dann wenn $\tau_{j,n}(\ell) = \hat{\sigma}(\tau_{k,n}(i))$ folgt für alle $\sigma \in S_n$ mit $\sigma(k) = j$:

$$\prod_{\substack{i=1 \\ i \neq k}}^{n} a_{i\sigma(i)} = \prod_{\substack{i=1 \\ i \neq k}}^{n} (A_{\widehat{kj}})_{\tau_{k,n}(i), \tau_{j,n}(\ell)} = \prod_{i=1}^{n-1} (A_{\widehat{kj}})_{i\hat{\sigma}(i)}.$$

Daraus folgt in (11.4), und mit der Leibniz–Formel angewandt auf $A_{\widehat{kj}}$:

$$c_k = \sum_{\substack{\sigma \in S_n \\ \text{mit } \sigma(k)=j}} \operatorname{sign}(\sigma) \prod_{\substack{i=1 \\ i \neq k}}^{n} a_{i\sigma(i)} = (-1)^{j+k} \sum_{\hat{\sigma} \in S_{n-1}} \operatorname{sign}(\hat{\sigma}) \prod_{i=1}^{n-1} (A_{\widehat{kj}})_{i\hat{\sigma}(i)} = (-1)^{j+k} \det(A_{\widehat{kj}}). \qquad \square$$

Bemerkung 11.54. Man erhält die Formeln des Laplace'schen Entwicklungssatzes auch über die eindeutige Charakterisierung der Determinante $\det(-)$ als Determinantenfunktion (multilinear und alternierend in den Spalten) mit Wert $\det(\mathbf{1}_n) = 1$ auf der Einheitsmatrix. Wie oben hat man nur eine der beiden Formeln zu beweisen, etwa (1). Dazu betrachtet man die rechte Seite der Formel (1) als Funktion der Matrix A. Die rechte Seite der Formel in (1) ist multilinear in den Spalten.

- Additiv in der s-ten Spalte: Sei $B = (b_{ij}) \in \mathrm{M}_n(K)$ mit $b_{ij} = a_{ij}$ für alle i und alle $j \neq s$. Ferner sei $C = (c_{ij}) \in \mathrm{M}_n(K)$ mit

$$c_{ij} = \begin{cases} a_{ij} & j \neq s \\ a_{ij} + b_{ij} & j = s. \end{cases}$$

Dann ist $C_{\widehat{ik}} = A_{\widehat{ik}} = B_{\widehat{ik}}$, falls $k = s$. Und für $k \neq s$ unterscheiden sich $A_{\widehat{ik}}$ und $B_{\widehat{ik}}$ nur in der ehemals (vor Streichen der k-ten Spalte) s-ten Spalte und $C_{\widehat{ik}}$ entsteht durch Addition dieser entsprechenden Spalten (und sonst wie $A_{\widehat{ik}}$ und $B_{\widehat{ik}}$). Somit gilt

$$\begin{aligned}\sum_{k=1}^{n}(-1)^{i+k}c_{ik}\det(C_{\widehat{ik}}) &= (-1)^{i+s}c_{is}\det(C_{\widehat{is}}) + \sum_{k=1,k\neq s}^{n}(-1)^{i+k}c_{ik}\det(C_{\widehat{ik}})\\ &= (-1)^{i+s}(a_{is}+b_{is})\det(C_{\widehat{is}}) + \sum_{k=1,k\neq s}^{n}(-1)^{i+k}c_{ik}\big(\det(A_{\widehat{ik}})+\det(B_{\widehat{ik}})\big)\\ &= \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det(A_{\widehat{ik}}) + \sum_{k=1}^{n}(-1)^{i+k}b_{ik}\det(B_{\widehat{ik}}).\end{aligned}$$

- Homogen in der s-ten Spalte: Sei $\lambda \in K$ und $B = (b_{ij})$ entstehe aus A durch Multiplikation der s-ten Spalte mit λ. Gleiches gilt für alle $k \neq s$ dann für $B_{\widehat{ik}}$ in Bezug auf $A_{\widehat{ik}}$. Für $k = s$ hingegen haben wir $A_{\widehat{is}} = B_{\widehat{is}}$.
  Dann gilt

$$\begin{aligned}\sum_{k=1}^{n}(-1)^{i+k}b_{ik}\det(B_{\widehat{ik}}) &= (-1)^{i+s}b_{is}\det(B_{\widehat{is}}) + \sum_{k=1,k\neq s}^{n}(-1)^{i+k}b_{ik}\det(B_{\widehat{ik}})\\ &= (-1)^{i+s}\lambda a_{is}\det(A_{\widehat{is}}) + \sum_{k=1,k\neq s}^{n}(-1)^{i+k}a_{ik}\big(\lambda\cdot\det(A_{\widehat{ik}})\big)\\ &= \lambda\cdot\Big(\sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det(A_{\widehat{ik}})\Big).\end{aligned}$$

Außerdem ist die rechte Seite der Formel in (1) alternierend in den Spalten. Seien die Spalten von A mit den Indizes $r < s$ gleich. Dann hat für alle $k \notin \{r, s\}$ die Matrix $A_{\widehat{ik}}$ auch zwei gleiche Spalten, somit verschwindende Determinante. Es entsteht $A_{\widehat{ir}}$ aus $A_{\widehat{is}}$ durch Vertauschen der Spalten mit der Permutation

$$\sigma = \begin{pmatrix}1 & \cdots & r-1 & r & r+1 & \cdots & s-2 & s-1 & s & \cdots & n-1\\ 1 & \cdots & r-1 & s-1 & r & \cdots & s-3 & s-2 & s & \cdots & n-1\end{pmatrix},$$

das ist das Produkt der Transpositionen

$$\sigma = (s-2, s-1)\circ\ldots\circ(j-1, j)\circ\ldots\circ(r, r+1),$$

(die r-te Spalte wandert jedesmal einen Index nach hinten, bis sie bei der Position $s-1$ angelangt ist; das ist das richtige Ziel, denn in $A_{\widehat{ir}}$ fehlt ja die r-te Spalte, so daß die ehemals s-te Spalte von A nun die $s-1$-te Spalte ist). Das Vorzeichen dieser Permutation ist

$$\operatorname{sign}(\sigma) = (-1)^{s-1-r}.$$

Daraus ergibt sich nun

$$\begin{aligned}\sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det(A_{\widehat{ik}}) &= (-1)^{i+r}a_{ir}\det(A_{\widehat{ir}}) + (-1)^{i+s}a_{is}\det(A_{\widehat{is}})\\ &= (-1)^{i+r}a_{is}\cdot\operatorname{sign}(\sigma)\cdot\det(A_{\widehat{is}}) + (-1)^{i+s}a_{is}\det(A_{\widehat{is}})\\ &= \big((-1)^{r-s}\cdot\operatorname{sign}(\sigma)+1\big)\cdot(-1)^{i+s}a_{is}\det(A_{\widehat{is}})\\ &= \big((-1)^{r-s}\cdot(-1)^{s-1-r}+1\big)\cdot(-1)^{i+s}a_{is}\det(A_{\widehat{is}}) = 0.\end{aligned}$$

Es bleibt, die Formel für die Einheitsmatrix $A = \mathbf{1}$ nachzuweisen. Wenn gestrichener Zeilenindex i nicht mit dem gestrichenen Spaltenindex j übereinstimmt, entsteht eine Matrix mit einer 0-Spalte in der i-ten Spalte. Diese hat Determinante 0. Damit ist der Nachweis im Fall der Einheitsmatrix eine leichte Übung.

Bemerkung 11.55. Die Vorzeichen $(-1)^{i+j}$, welche in der Laplace-Entwicklung der Determinante auftreten, kann man sich geometrisch leicht merken: sie bilden ein Schachbrettmuster:

$$\begin{pmatrix}+ & - & + & - & + & \cdots\\ - & + & - & + & - & \cdots\\ + & - & + & - & & \\ - & + & - & & \ddots & \ddots\\ + & - & & \ddots & \ddots & \\ \vdots & \vdots & & \ddots & & \end{pmatrix}$$

Beispiel 11.56. Wir berechnen durch Entwicklung nach der ersten Zeile:

$$\det\begin{pmatrix}2&3&7\\1&1&0\\0&3&0\end{pmatrix} = 2\cdot\det\begin{pmatrix}1&0\\3&0\end{pmatrix} - 3\cdot\det\begin{pmatrix}1&0\\0&0\end{pmatrix} + 7\cdot\det\begin{pmatrix}1&1\\0&3\end{pmatrix}$$
$$= 2\cdot 0 - 3\cdot 0 + 7\cdot 3 = 21.$$

Bei genauerer Betrachtung wäre aufgrund der vielen Nullen in der letzten Spalte eine Entwicklung nach dieser geschickter (die Summanden $\pm 0\cdot\det(\ )$ lassen wir gleich weg):

$$\det\begin{pmatrix}2&3&7\\1&1&0\\0&3&0\end{pmatrix} = 7\cdot\det\begin{pmatrix}1&1\\0&3\end{pmatrix} = 7\cdot 3 = 21.$$

Definition 11.57. Sei $1 \le r \le \min\{n,m\}$ und $A \in \mathrm{M}_{n\times m}(K)$. Ein $r\times r$**-Minor** der Matrix A ist

$$\det A_{I,J}$$

für eine Wahl von Teilmengen $I \subseteq \{1,\dots,n\}$ von $r = |I|$-vielen Zeilen und $J \subseteq \{1,\dots,m\}$ von $r = |J|$-vielen Spalten.

Eine Matrix hat also in der Regel mehr als einen $r\times r$-Minor.

Beispiel 11.58. Die Einträge der Adjunkten zu $A \in \mathrm{M}_n(K)$ sind bis auf das Vorzeichen genau die $(n-1)\times(n-1)$-Minoren der Matrix A.

Satz 11.59. *Sei $A \in \mathrm{M}_{n\times m}(K)$ eine Matrix. Dann gilt*

$$\mathrm{rg}(A) = \max\{s\ ;\ \ A \text{ hat einen } s\times s\text{-Minor } \neq 0\}$$

Beweis. Schritt 1: In Korollar 9.9 wurde bewiesen, daß für alle $I \subseteq \{1,\dots,n\}$ und $J =\subseteq \{1,\dots,m\}$

$$\mathrm{rg}(A_{I,J}) \le \mathrm{rg}(A).$$

Schritt 2: Wenn A einen $s\times s$-Minor $\neq 0$ hat, dann gibt es I und J wie oben mit $\det(A_{I,J}) \neq 0$. Nach Satz 11.37 ist $A_{I,J}$ ist regulär: $\mathrm{rg}(A_{I,J}) = s$. Dies zeigt

$$\mathrm{rg}(A) \ge \max\{s\ ;\ \ A \text{ hat einen } s\times s\text{-Minor } \neq 0\}.$$

Schritt 3: Jetzt wählen wir die Spalten und Zeilen sorgfältig aus. Sei $\mathrm{rg}(A) = r$. Dann gibt es Spalten $J = \{j_1 < \ldots < j_r\} \subseteq \{1,\dots,m\}$ von A, die linear unabhängig sind. Sei B die Matrix dieser Spalten mit Indizes $j \in J$. Dann ist $\mathrm{rg}(B^t) = \mathrm{rg}(B) = \mathrm{rg}(A) = r$. In B^t wählen wir wieder Spalten $I = \{i_1 < \ldots < i_r\}$, die linear unabhängig sind. Der Rang bleibt wieder r und erneutes Transponieren liefert wie oben $A_{I,J}$. Damit ist

$$r = \mathrm{rg}(A) = \mathrm{rg}(A_{I,J}).$$

Somit ist $A_{I,J}$ regulär, und nach Satz 11.37 gilt $\det(A_{I,J}) \neq 0$. Damit hat A einen $r\times r$-Minor $\neq 0$, wobei $r = \mathrm{rg}(A)$. Dies zeigt die umgekehrte Ungleichung. □

Übungsaufgaben zu §11

Übungsaufgabe 11.1. Zeigen Sie, daß für einen Körper K die Zuordnung

$$K \to \mathrm{GL}_2(K)$$
$$x \mapsto \begin{pmatrix}1&x\\0&1\end{pmatrix}$$

einen Gruppenhomomorphismus von der additiven Gruppe $(K,+)$ nach $\mathrm{GL}_2(K)$ beschreibt.

Übungsaufgabe 11.2. Sei K ein Körper. Mit komponentenweiser Multiplikation ist

$$K^\times \times K^\times$$

eine Gruppe. Zeigen Sie, daß für einen Körper K die Zuordnung

$$K^\times \times K^\times \to \mathrm{GL}_2(K)$$
$$(x,y) \mapsto \begin{pmatrix}x&\\&y\end{pmatrix}$$

einen Gruppenhomomorphismus beschreibt.

Übungsaufgabe 11.3. Zwei Gruppen G und H heißen isomorph, wenn es einen Gruppenisomorphismus $G \to H$ gibt, d.h. einen bijektiven Gruppenhomomorphismus.

Zeigen Sie: Es gibt bis auf Isomorphie genau eine Gruppe mit zwei Elementen.

Übungsaufgabe 11.4. Sei $n \in \mathbb{N}_0$. Zeigen Sie, daß die Menge der geraden Permutationen

$$A_n := \{\sigma \in S_n \ ; \ \sigma \text{ gerade}\}$$

eine Gruppe bezüglich der Komposition von Permutationen in S_n ist. Zeigen Sie weiter, daß A_n für $n \geq 2$ genau $n!/2$-viele Elemente hat.

Bemerkung: Diese Gruppe heißt die **alternierende Gruppe** (auf n Elementen).

Übungsaufgabe 11.5. Die **spezielle lineare Gruppe** der Dimension n über dem Körper K ist

$$\mathrm{SL}_n(K) = \{A \in \mathrm{GL}_n(K) \ ; \ \det(A) = 1\}.$$

Zeigen Sie, daß $\mathrm{SL}_n(K)$ mit Matrixmultiplikation eine Gruppe ist.

Übungsaufgabe 11.6. Betrachten Sie die Abbildung $q : \mathrm{M}_2(\mathbb{F}_2) \to \mathbb{F}_2$ definiert durch

$$q(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ab + cd.$$

Zeigen Sie, daß q multilinear in den Spalten und antisymmetrisch ist. Dabei bedeutet **antisymmetrisch**, daß für alle Vektoren v, w gilt $q([v, w]) = -q([w, v])$, also die Reaktion der Form mit einem Vorzeichen auf das Vertauschen der Spalten ist wie bei Satz 11.29.

Zeigen Sie weiter, daß q keine Determinantenfunktion auf $(\mathbb{F}_2)^2$ ist.

Übungsaufgabe 11.7 (Vandermonde–Determinante). Seien $a_1, \ldots, a_n \in K$. Zeigen Sie

$$\det \begin{pmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{(n-1)} \\ 1 & a_2 & a_2^2 & \ldots & a_2^{(n-1)} \\ \vdots & \vdots & & & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{(n-1)} \end{pmatrix} = \prod_{i<j} (a_j - a_i).$$

Übungsaufgabe 11.8 (Alternative Definition der Determinante eines Endomorphismus). Sei V ein K-Vektorraum mit $\dim(V) = n$ und sei $f : V \to V$ ein Endomorphismus. Sei $\delta \neq 0$ eine Determinantenfunktion auf V. Dann ist $f^*(\delta)$ definiert auf $v_1, \ldots, v_n \in V$ durch

$$f^*(\delta)(v_1, \ldots, v_n) = \delta(f(v_1), \ldots, f(v_n))$$

eine Determinantenfunktion. Wir definieren nun $\det(f) \in K$ als dasjenige Element, für das

$$f^*(\delta) = \det(f) \cdot \delta \in \det(V^*) \tag{11.6}$$

gilt. Zeigen Sie, daß diese Definition wohldefiniert ist (es existiert ein eindeutiger solcher Skalar und dieser Skalar ist unabhängig von der Wahl von $\delta \neq 0$).

Zeigen Sie für einen weiteren Endomorphismus $g : V \to V$, daß

$$\det(f \circ g) = \det(f) \cdot \det(g),$$

indem Sie die Definition der Determinante mittels (11.6) nutzen.

Übungsaufgabe 11.9. Sei $A = [v_1, \ldots, v_n] \in \mathrm{M}_n(K)$ eine quadratische Matrix. Zeigen Sie, daß die folgenden Aussagen äquivalent sind.

(a) $(v_1, \ldots, v_n)$ ist Basis des K^n.
(a_1) $(v_1, \ldots, v_n)$ ist linear unabhängig.
(a_2) $(v_1, \ldots, v_n)$ ist Erzeugendensystem des K^n.

(b) $AX = b$ hat für alle $b \in K^n$ eine eindeutige Lösung.
(b_1) $AX = 0$ hat nur die Lösung $x = 0$.

(b_2) $AX = b$ hat für alle $b \in K^n$ eine Lösung.

(c) Die lineare Abbildung $L_A : K^n \to K^n$ ist bijektiv.
(c_1) Die lineare Abbildung $L_A : K^n \to K^n$ ist injektiv.
(c_2) Die lineare Abbildung $L_A : K^n \to K^n$ ist surjektiv.

(d) Die lineare Abbildung $L_A : K^n \to K^n$ ist ein Isomorphismus.
(d') A ist invertierbar: $A \in \mathrm{GL}_n(K)$, d.h. es gibt $B \in \mathrm{M}_n(K)$ mit $AB = \mathbf{1} = BA$.

(e) $\det(A) \neq 0$.

(f) Das Gaußsche Eliminationsverfahren liefert als Zeilenstufenform eine obere Dreiecksmatrix mit Diagonaleinträgen aus $K^\times$.

(g) A ist regulär: $\mathrm{rg}(A) = n$.

12. Eigenwerte und Eigenvektoren

12.1. **Spektrum und Eigenräume.** Im Allgemeinen ist es nicht leicht, sich eine lineare Abbildung geometrisch zu veranschaulichen. Ein besonders einfaches Beispiel ist jedoch die Streckung. Seien V ein K-Vektorraum und $\lambda \in K$. Dann ist die Streckung von V um den Faktor λ die lineare Abbildung

$$\lambda \cdot \mathrm{id}_V : V \to V,$$

die jedes $v \in V$ auf $\lambda v \in V$ abbildet. Bei $K = \mathbb{R}$ und $V = \mathbb{R}^2$ interpretiert als Ebene kommt hier die Streckung um den Faktor λ mit Streckzentrum im Ursprung $0 \in \mathbb{R}^2$ heraus.

Wir isolieren nun für einen beliebigen Endomorphismus $f : V \to V$ Teile von V, auf denen sich f wie eine Streckung benimmt.

Definition 12.1. Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V.

(1) Ein **Eigenvektor** von f ist ein Vektor $v \in V$, $v \neq 0$, so daß es ein $\lambda \in K$ gibt mit
$$f(v) = \lambda \cdot v.$$
(2) Ein **Eigenwert** von f ist ein $\lambda \in K$, so daß es einen Eigenvektor $v \in V$ von f gibt mit
$$f(v) = \lambda \cdot v.$$
(3) Das **Spektrum** von f ist die Menge seiner Eigenwerte.

Bemerkung 12.2. (1) Zu einem Eigenvektor gehört immer ein eindeutiger Eigenwert. In der Tat, wenn $v \in V$ ein Eigenvektor ist und $\mu, \lambda \in K$ mit

$$\mu v = f(v) = \lambda v,$$

dann folgt $(\lambda - \mu)v = 0$, und weil $v \neq 0$, folgt dann $\lambda = \mu$.

(2) Die Eigenwertgleichung $f(v) = \lambda v$ wird auch von $v = 0$ und dann jedem beliebigen $\lambda \in K$ gelöst. Diese triviale Lösung $v = 0$ ist per Definition **kein Eigenvektor** und führt auch nicht dazu, daß jedes $\lambda \in K$ ein Eigenwert ist.

(3) Die Begriffe Eigenvektor und Eigenwert eines Endomorphismus machen keine Voraussetzungen an die Dimension des Vektorraums. Diese Dimension darf auch unendlich sein.

Definition 12.3. Sei K ein Körper. Eine **Diagonalmatrix** in $\mathrm{M}_n(K)$ ist eine Matrix $D = (d_{ij})$ mit $d_{ij} = 0$ für alle $i \neq j$:

$$D = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_n \end{pmatrix}.$$

Hier ist $d_{ii} = \alpha_i$. Wir schreiben kurz

$$D = \operatorname{diag}(\alpha_1, \dots, \alpha_n).$$

Proposition 12.4. *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums, und sei $\mathscr{B} = (b_1, \dots, b_n)$ eine Basis von V. Dann sind äquivalent:*

(a) Die Darstellungsmatrix $\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ ist eine Diagonalmatrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

(b) Die Basis $\mathscr{B}$ besteht aus Eigenvektoren von f, und zwar $f(b_j) = \lambda_j b_j$ für $j = 1, \dots, n$.

Beweis. Für die Darstellungsmatrix $\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ müssen die Vektoren $f(b_j)$ in der Basis $\mathscr{B}$ ausgedrückt werden. Die Eigenwertgleichung $f(b_j) = \lambda_j b_j$ bedeutet gerade wegen $\kappa_{\mathscr{B}}(b_j) = e_j$, daß in der j-ten Spalte der Vektor $\lambda_j e_j$ steht. □

Lineare Abbildungen lassen sich durch Matrizen und Vektoren durch Spaltenvektoren darstellen. Die Begriffe Eigenvektor und Eigenwert haben entsprechend passende Bedeutungen.

Definition 12.5. Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix.

(1) Ein **Eigenvektor** von A ist ein Vektor $x \in K^n$, $x \neq 0$, so daß es ein $\lambda \in K$ gibt mit
$$Ax = \lambda \cdot x.$$

(2) Ein **Eigenwert** von A ist ein $\lambda \in K$, so daß es einen Eigenvektor $x \in K^n$ von A gibt mit
$$Ax = \lambda \cdot x.$$

(3) Das **Spektrum** von A ist die Menge seiner Eigenwerte.

Bemerkung 12.6. Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit Basis $\mathscr{B}$ und Dimension n. Sei A die Darstellungsmatrix bezüglich $\mathscr{B}$ sowohl in Definitions- wie Wertebereich von f, also

$$A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f).$$

Mit der Koordinatenabbildung $\kappa_{\mathscr{B}} : V \xrightarrow{\sim} K^n$ gilt für $v \in V$ und $x = \kappa_{\mathscr{B}}(v) \in K^n$:

$$v \neq 0 \iff x \neq 0$$

und für $\lambda \in K$ auch, siehe Proposition 8.32,

$$f(v) = \lambda v \iff Ax = \lambda x.$$

Daher ist v ein Eigenvektor von f genau dann, wenn x ein Eigenvektor von A ist. Die Eigenwerte und damit das Spektrum sind dann die gleichen.

Beispiel 12.7. (1) Eine Diagonalmatrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in \mathrm{M}_n(K)$ hat offensichtlich Eigenvektoren $e_i \in K^n$, für $i = 1, \dots n$, denn

$$De_i = \lambda_i \cdot e_i.$$

Somit sind $\lambda_1, \dots, \lambda_n$ schon einmal Eigenwerte von D. Wir sehen gleich, daß dies alle Eigenwerte von D sind.

(2) Betrachten wir dazu nun den Spezialfall $K = \mathbb{R}$ und $n = 3$. Dann vermittelt die Matrix

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

eine lineare Abbildung $\mathbb{R}^3 \to \mathbb{R}^3$, und zwar

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 2y \\ 3z \end{pmatrix}.$$

Die Koordinaten werden also mit einem individuellen Streckfaktor gestreckt. Das entspricht den drei Richtungen der Eigenvektoren und den entsprechenden Eigenwerten als Streckfaktor.

Wir kommen nun zur fundamentalen Eigenschaft von Eigenwerten, über die sie letztlich berechenbar werden.

Satz 12.8. *Sei $A \in \mathrm{M}_n(K)$. Dann gilt:*

$$\lambda \in K \text{ ist Eigenwert von } A \iff \det(\lambda \cdot \mathbf{1} - A) = 0.$$

Beweis. Das Element $\lambda \in K$ ist Eigenwert zu A per Definition, wenn es $x \in K^n$, $x \neq 0$ gibt mit $Ax = \lambda x$. Dies ist äquivalent zu

$$Ax = \lambda x \iff Ax = \lambda \cdot \mathbf{1}x \iff 0 = (\lambda \cdot \mathbf{1} - A)x = 0,$$

somit dazu, daß die Matrix $\lambda \cdot \mathbf{1} - A$ einen nichttrivialen Kern hat. Das wiederum ist äquivalent nach Satz 7.23 zu

$$\lambda \cdot \mathbf{1} - A \text{ ist nicht invertierbar}$$

und dies wiederum nach Satz 11.37 zu $\det(\lambda \cdot \mathbf{1} - A) = 0$. □

Um die analoge Aussage für Eigenwerte von Endomorphismen aufstellen zu können, nutzen wir den Begriff der Determinante eines Endomorphismus, siehe Definition 11.39.

Korollar 12.9. *Sei K ein Körper. Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums. Dann sind äquivalent:*

(a) $\lambda \in K$ *ist Eigenwert von* f.
(b) $\det(\lambda \cdot \mathrm{id}_V - f) = 0$.

Beweis. Sei $\mathscr{B}$ eine Basis von V und $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ die Darstellungsmatrix von f bezüglich $\mathscr{B}$. Dann ist λ Eigenwert von f genau dann, wenn λ Eigenwert von A ist. Und gemäß Satz 12.8 ist dies äquivalent zu $\det(\lambda \cdot \mathbf{1} - A) = 0$. Nun folgt alles aus der Rechnung:

$$\det(\lambda \cdot \mathrm{id}_V - f) = \det\left(\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(\lambda \cdot \mathrm{id}_V - f)\right) = \det\left(\lambda \cdot \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(\mathrm{id}_V) - \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)\right) = \det(\lambda \cdot \mathbf{1} - A). \quad \square$$

Definition 12.10. Seien K ein Körper und $\lambda \in K$.

(1) Sei $f : V \to V$ ein Endomorphismus des K-Vektorraums V. Wir setzen

$$V_\lambda = \ker(\lambda \cdot \mathrm{id}_V - f).$$

Ist λ ein Eigenwert von f, dann heißt V_λ der **Eigenraum** von f zum Eigenwert λ.

(2) Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Wir setzen

$$V_\lambda = \ker(\lambda \cdot \mathbf{1} - A).$$

Ist λ ein Eigenwert von A, dann heißt V_λ der **Eigenraum** von A zum Eigenwert λ.

Proposition 12.11. *Seien K ein Körper und $\lambda \in K$.*

(a) Sei $f : V \to V$ ein Endomorphismus des K-Vektorraums V.
  (1) V_λ ist ein Unterraum von V.
  (2) $V_\lambda \neq 0 \iff \lambda$ ist Eigenwert von f.
  (3) Wenn λ ein Eigenwert von f ist, dann gilt

$$V_\lambda = \{0\} \cup \{v \in V \ ; \ v \text{ ist Eigenvektor von } f \text{ zum Eigenwert } \lambda\}.$$

(b) Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix.
  (1) V_λ ist ein Unterraum von K^n.
  (2) $V_\lambda \neq 0 \iff \lambda$ ist Eigenwert von A.

(3) *Wenn λ ein Eigenwert von A ist, dann gilt*
$$V_\lambda = \{0\} \cup \{x \in K^n \ ; \ x \text{ ist Eigenvektor von } A \text{ zum Eigenwert } \lambda\}.$$

Beweis. (1) Ein Kern ist ein Unterraum. Aussage (2) haben wir im Beweis von Satz 12.8 und seinem Analogon Korollar 12.9 bereits behandelt: λ ist genau dann Eigenwert, wenn $\lambda \cdot \mathbf{1} - A$ (bzw. $\lambda \cdot \mathrm{id}_V - f$) einen nicht-trivialen Kern hat. Dieser Kern ist V_λ.

(3) Ein $v \in V$ ist Eigenvektor zum Eigenwert λ genau dann, wenn $v \neq 0$ und
$$(\lambda \cdot \mathbf{1} - f)(v) = 0,$$
also genau wenn wenn $v \in V_\lambda \setminus \{0\}$. Die Variante für die Matrix A geht genauso. □

Bemerkung 12.12. (1) Sei λ ein Eigenwert des Endomorphismus $f : V \to V$. Der Eigenraum $V_\lambda \subseteq V$ ist genau der Teil von V, auf dem f wie die Streckung um den Faktor λ wirkt.

(2) Wenn λ ein Eigenwert von $A \in \mathrm{M}_n(K)$ ist, dann bestimmt man einen zugehörigen Eigenvektor, indem man V_λ als Kern/Nullraum von $\lambda\mathbf{1} - A$ bestimmt.

(3) Für Eigenvektoren zu Endomorphismen $f : V \to V$ geht man analog vor: zuerst bestimmt man eine Darstellungsmatrix A bezüglich einer Basis $\mathscr{B}$ von V, bestimmt dafür den entsprechenden Eigenraum und transformiert zum Schluß wieder mittels der Inversen der Koordinatenabbildung $\kappa_{\mathscr{B}} : V \to K^n$ zum entsprechenden Unterraum von V zurück.

12.2. **Das charakteristische Polynom.** Polynome werden erst in Kapitel 17 systematisch behandelt. Wir nutzen das charakteristische Polynom, um die Beschäftigung mit Polynomen zu motivieren.

Wenn wir λ als Unbestimmte, als Variable X betrachten, dann müssen wir zur Bestimmung der Eigenwerte von $A = (a_{ij}) \in \mathrm{M}_n(K)$ die Gleichung
$$\det(X \cdot \mathbf{1} - A) = 0$$
lösen mit einer Variablen X, die Werte in K annehmen kann. Die linke Seite der Gleichung ist ein Ausdruck der Form
$$\begin{aligned}\det(X \cdot \mathbf{1} - A) = \det((\delta_{ij}X - a_{ij})_{ij}) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} (\delta_{i\sigma(i)}X - a_{i\sigma(i)}) \\ &= X^n + c_{n-1}X^{n-1} + \ldots + c_2X^2 + c_1X + c_0,\end{aligned}$$
für gewisse $c_i \in K$, $i = 0, \ldots, n-1$. Dabei haben wir die rekursiv definierte Potenzschreibweise für die Variable X benutzt: für $n \geq 0$ sei
$$X^n := \begin{cases} 1 & n = 0 \\ X \cdot X^{n-1} & n \geq 1 \end{cases} = \underbrace{X \cdot \ldots \cdot X}_{n-\text{mal}}$$
Der Koeffizient vor X^n ist 1, denn dieser Teil des Ausdrucks kommt nur aus dem Summanden, der zu $\sigma = \mathrm{id}$ gehört:
$$\prod_{i=1}^{n} (\delta_{ii}X - a_{ii}) = \prod_{i=1}^{n} (X - a_{ii}) = X^n - (\sum_{i=1}^{n} a_{ii})X^{n-1} + \ldots.$$

Definition 12.13. Ein **Polynom (in einer Variablen X)** mit Koeffizienten im Körper K besteht aus Koeffizienten $a_i \in K$, $i \in \mathbb{N}_0$, so daß es ein $n \in \mathbb{N}_0$ gibt mit
$$a_i = 0 \text{ für alle } i > n.$$
Das zugehörige Polynom $P = P(X)$ wird dann notiert als
$$P = P(X) = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n = \sum_{i=0}^{n} a_iX^i,$$

wobei n so groß gewählt wird, daß $a_i = 0$ für alle $i > n$. Welche n man dabei wählen darf, hängt vom Polynom $P(X)$ ab. Wenn $a_i = 0$ ist, dann darf der Term $a_i X^i$ in der Notation weggelassen werden. Insbesondere sind, mit n und $P(X)$ wie oben, für alle $m \geq n$ durch

$$\sum_{i=0}^{n} a_i X^i = a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n + 0 \cdot X^{n+1} + \ldots + 0 \cdot X^m = \sum_{i=0}^{m} a_i X^i$$

verschiedene Darstellungen desselben Polynoms $P(X)$ gegeben.

Die Menge aller Polynome mit Koeffizienten aus K bezeichnen wir mit $K[X]$. Formal präzise (aber leider sehr unintuitiv) besteht $K[X]$ einfach aus der Menge der Folgen $(a_i)_{i \in \mathbb{N}_0}$ von Elementen $a_i \in K$, die für hinreichend große i zur Nullfolge $a_i = 0$ werden.

Der **Grad** eines Polynoms $P(X) = \sum_{i=0}^{n} a_i X^i$ ist

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

Bemerkung 12.14. Auf $K[X]$ ist eine Addition, eine Multiplikation und eine Skalarmultiplikation mit Skalaren aus K erklärt, die $K[X]$ zu einem Ring und einem K-Vektorraum machen, den **Polynomring** mit Koeffizienten aus K.

Zu einem $x_0 \in K$ und einem Polynom $P(X) = \sum_{i=0}^{n} a_i X^i$ gibt es die **Auswertung** in x_0:

$$P(x_0) := \sum_{i=0}^{n} a_i (x_0)^i \in K,$$

wobei hier in K ebenfalls die Potenzschreibweise benutzt wurde. Die Abbildung

$$K[X] \to K, \qquad P \mapsto P(x_0)$$

ist ein K-linearer Ringhomomorphismus. Bei $P \in K[X]$ mit $P(x_0) = 0$ spricht man davon, daß x_0 eine **Nullstelle** von P aus K ist.

Definition 12.15. Sei K ein Körper. Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Das **charakteristische Polynom** von A ist

$$\chi_A(X) = \det(X \cdot \mathbf{1} - A) \in K[X].$$

Analog definieren wir für einen Endomorphismus $f : V \to V$ eines endlichdimensionalen K-Vektorraums das **charakteristische Polynom** als

$$\chi_f(X) = \det(X \cdot \mathrm{id}_V - f) \in K[X].$$

Proposition 12.16. *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums V der Dimension n. Sei $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ die Darstellungsmatrix von f bezüglich der Basis $\mathscr{B}$ von V. Dann gilt*

$$\chi_f(X) = \chi_A(X).$$

Beweis. Das folgt direkt aus

$$\begin{aligned} \chi_f(X) &= \det(X \cdot \mathrm{id}_V - f) = \det\left(\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(X \cdot \mathrm{id}_V - f)\right) \\ &= \det\left(X \cdot \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(\mathrm{id}_V) - \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)\right) = \det(X \cdot \mathbf{1}_n - A) = \chi_A(X). \end{aligned}$$

□

Proposition 12.17. *Sei K ein Körper. Ähnliche Matrizen haben das gleiche charakteristische Polynom: für alle $A \in \mathrm{M}_n(K)$ und alle $S \in \mathrm{GL}_n(K)$ gilt*

$$\chi_{SAS^{-1}}(X) = \chi_A(X).$$

Beweis. Die Matrizen SAS^{-1} und A sind Darstellungsmatrizen von $L_A : K^n \to K^n$ bezüglich einer Basis (genauer aber unerheblich: diejenige aus den Spalten von S^{-1} für SAS^{-1} und bezüglich der Standardbasis für A). Daraus folgt

$$\chi_{SAS^{-1}}(X) = \chi_{L_A}(X) = \chi_A(X). \qquad \square$$

Die wichtige Rolle des charakteristischen Polynoms in Bezug auf Eigenwerte unterstreicht der folgende Satz.

Satz 12.18. *Seien K ein Körper, $A \in \mathrm{M}_n(K)$ eine quadratische Matrix und $f : V \to V$ ein Endomorphismus des endlichdimensionalen K-Vektorraums V.*
(1) Die Eigenwerte von A sind genau die Nullstellen von $\chi_A(X)$ aus K.
(2) Die Eigenwerte von f sind genau die Nullstellen von $\chi_f(X)$ aus K.

Beweis. Wegen der Gleichungen

$$\chi_A(\lambda) = \det(\lambda \cdot \mathbf{1}_n - A) \quad \text{und} \quad \chi_f(\lambda) = \det(\lambda \cdot \mathrm{id}_V - f)$$

folgt die Behauptung sofort aus Satz 12.8 und Korollar 12.9. $\square$

Wenn $K = \mathbb{R}$ und das Polynom höchstens quadratische ist, lernt man die Lösungstheorie in der Schule. Für systematische Antworten auf die Frage, was man im Allgemeinen sagen kann, müssen wir auf Kapitel 17 warten.

Proposition 12.19. *Sei $P(X) \in K[X]$ ein Polynom und $\alpha \in K$. Dann gibt es ein Polynom $Q(X) \in K[X]$ mit*

$$P(X) = (X - \alpha) \cdot Q(X) + P(\alpha).$$

Beweis. Man rechnet sofort nach, daß für $Q_n(X) = X^{n-1} + \alpha X^{n-2} + \ldots + \alpha^{n-2}X + \alpha^{n-1}$

$$(X - \alpha) \cdot Q_n(X) = X^n - \alpha^n.$$

Daraus folgt für $P(X) = \sum_{i=0}^d a_i X^i$

$$P(X) - P(\alpha) = \sum_{i=0}^d a_i(X^i - \alpha^i) = \sum_{i=0}^d a_i(X - \alpha) \cdot Q_i(X) = (X - \alpha) \cdot \sum_{i=0}^d a_i \cdot Q_i(X),$$

und die Behauptung gilt mit $Q(X) = \sum_{i=0}^d a_i \cdot Q_i(X)$. $\square$

Korollar 12.20. *Sei $P(X) \in K[X]$ ein Polynom und $\alpha \in K$. Dann sind äquivalent:*
(a) α ist Nullstelle von $P(X)$.
(b) $X - \alpha$ ist ein Faktor von $P(X)$, d.h. es gibt $Q(X) \in K[X]$ mit $P(X) = (X - \alpha) \cdot Q(X)$.

Beweis. Wenn $P(\alpha) = 0$ ist, dann folgt die Existenz einer Faktorisierung $P(X) = (X - \alpha) \cdot Q(X)$ sofort aus Proposition 12.19.

Wenn umgekehrt $P(X) = (X - \alpha) \cdot Q(X)$ für ein $Q \in K[X]$ gilt, dann ist

$$P(\alpha) = (\alpha - \alpha) \cdot Q(\alpha) = 0 \qquad \square$$

Nullstellen eines Polynoms erkennt man daher an Faktoren der Form $X - \alpha$, die zur Nullstelle α gehören.

Beispiel 12.21. Eine Diagonalmatrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathrm{M}_n(K)$ hat genau die Eigenwerte $\lambda_1, \ldots, \lambda_n$. Das charakteristische Polynom ist

$$\chi_D(X) = \det(X \cdot \mathbf{1}_n - D) = \det(\mathrm{diag}(X - \lambda_1, \ldots, X - \lambda_n)) = \prod_{i=1}^n (X - \lambda_i).$$

Damit ist für alle $i = 1, \ldots, n$ der Faktor $X - \lambda_i$ ein Teiler von $\chi_D(X)$, somit λ_i eine Nullstelle und damit ein Eigenwert von D. Das wußten wir schon, denn $De_i = \lambda_i e_i$. Aber für μ verschieden von den $\lambda_1, \ldots, \lambda_n$ gilt

$$\chi_D(\mu) = \prod_{i=1}^{n} (\mu - \lambda_i) \neq 0,$$

also ist so ein μ kein Eigenwert von D. Das Spektrum von D ist daher $\lambda_1, \ldots, \lambda_n$.

12.3. **Diagonalisierbarkeit.** Mit diesem Abschnitt steigen wir in die Suche nach Normalformen von Matrizen und Endomorphismen ein. Die Leitfrage nimmt ein Endomorphismus $f : V \to V$ eines K-Vektorraums und fragt, wie man eine zu f passende Basis $\mathscr{B}$ von V findet, in der die Darstellungsmatrix

$$A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$$

eine möglichst einfache Gestalt hat? Eine Übersicht darüber, welche Matrizen grunndsätzlich zur Auswahl stehen, liefert der Begriff der Ähnlichkeit von Matrizen.

Definition 12.22. Sei K ein Körper und $n \in \mathbb{N}$. Zwei Matrizen $A, B \in \mathrm{M}_n(K)$ heißen **ähnlich**, Notation $A \sim B$, wenn

$$A \sim B :\iff \text{ es gibt } S \in \mathrm{GL}_n(K) \text{ mit } A = SBS^{-1}.$$

Proposition 12.23. *Ähnlichkeit von Matrizen ist eine Äquivalenzrelation auf* $\mathrm{M}_n(K)$.

Beweis. Das ist eine leichte Übungsaufgabe. Seien $A, B, C \in \mathrm{M}_n(K)$ beliebig.

- *Reflexiv*: Das folgt mit $S = \mathbf{1}$ aus $A = \mathbf{1}A\mathbf{1}^{-1}$ wegen $\mathbf{1}^{-1} = \mathbf{1}$.
- *Symmetrisch*: Wenn $A \sim B$, dann gibt es $S \in \mathrm{GL}_n(K)$ mit $B = SAS^{-1}$. Setzen wir $T = S^{-1}$, dann ist $T^{-1} = S$ und $T \in \mathrm{GL}_n(K)$. Dann folgt $B \sim A$ aus
$$TBT^{-1} = T(SAS^{-1})T^{-1} = S^{-1}SAS^{-1}S = A.$$
- *Transitiv*: Wenn $A \sim B$ und $B \sim C$, dann gibt es $S, T \in \mathrm{GL}_n(K)$ mit $B = SAS^{-1}$ und $C = TBT^{-1}$. Setzen wir $U = TS$, dann ist $U^{-1} = S^{-1}T^{-1}$ und $U \in \mathrm{GL}_n(K)$. Dann folgt $A \sim C$ aus
$$C = TBT^{-1} = T(SAS^{-1})T^{-1} = UAU^{-1}.$$ □

Proposition 12.24. *Sei V ein K-Vektorraum der Dimension n und $f : V \to V$ ein Endomorphismus. Sei $\mathscr{B}$ eine Basis von V und $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ die zugehörige Darstellungsmatrix. Sei $B \in \mathrm{M}_n(K)$ eine weitere Matrix. Dann gilt*

$$B \sim A \iff \text{es gibt eine Basis } \mathscr{C} \text{ von } V \text{ mit } B = \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f).$$

Insbesondere ist die Menge der Darstellungsmatrizen zu f bezüglich aller möglicher Wahlen (die gleiche für Quelle und Ziel) von V ist eine Ähnlichkeitsklasse von Matrizen.

Beweis. Wenn $B = \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f)$, dann folgt $B \sim A$ aus der Basiswechselformel Proposition 8.48 und Proposition 8.50.

Wenn $B = SAS^{-1}$ mit $S \in \mathrm{GL}_n(K)$, dann müssen wir eine Basis $\mathscr{C}$ finden, so daß $S^{-1} = S_{\mathscr{B}}^{\mathscr{C}}$ gilt. Das leistet Proposition 8.52. □

Gemäß Proposition 12.24 müssen wir die Frage beantworten, wie eine zu A ähnlichen Matrix

$$SAS^{-1}$$

mit besonders guten Eigenschaften aussehen kann.

Definition 12.25. Sei K ein Körper.

(1) Eine Matrix $A \in \mathrm{M}_n(K)$ heißt **diagonalisierbar**, wenn es ein $S \in \mathrm{GL}_n(K)$ und eine Diagonalmatrix D gibt mit
$$A = SDS^{-1}.$$
(2) Ein Endomorphismus $f : V \to V$ eines endlichdimensionalen K-Vektorraums V heißt **diagonalisierbar** (oder **split halbeinfach**), wenn es eine Darstellungsmatrix von f gibt, die eine Diagonalmatrix ist.

Bemerkung 12.26. Die Gleichung $A = SDS^{-1}$ aus der Definition der Diagonalisierbarkeit kann man auch „umdrehen“. Es ist $T = S^{-1} \in \mathrm{GL}_n(K)$ und
$$TAT^{-1} = T(SDS^{-1})T^{-1} = (TS)D(TS)^{-1} = D.$$
In dieser Betrachtung wird die diagonalisierbare Matrix A durch Konjugation mit T zur Diagonalmatrix $D = TAT^{-1}$. Wenn wir geometrisch denken, dann entspricht A einem Endomorphismus, den wir durch Übergang zur dazu ähnlichen Matrix TAT^{-1} in einer anderen Basis ausgedrückt haben, nun mit einer Diagonalmatrix und damit viel einfacher.

Proposition 12.27. *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums. Es sind äquivalent:*
(a) f ist diagonalisierbar.
(b) Es gibt eine Basis $\mathscr{B}$ von V, so daß die Darstellungsmatrix $\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ diagonalisierbar ist.
(c) Für jede Basis $\mathscr{B}$ von V ist die Darstellungsmatrix $\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ diagonalisierbar.

Beweis. (a) $\implies$ (c): Wenn f diagonalisierbar ist, dann gibt es per Definition eine Basis $\mathscr{C}$, bezüglich der die Darstellungsmatrix $D = \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f)$ eine Diagonalmatrix ist. Alle anderen Darstellungsmatrizen sind dann diagonalisierbar wegen
$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = S\,\mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f)S^{-1} = SDS^{-1}$$
mit $S = S_{\mathscr{B}}^{\mathscr{C}}$. Das zeigt (c).

Die Implikation (c) $\implies$ (b) ist trivial.

(b) $\implies$ (a): Sei $\mathscr{B}$ eine Basis, so daß $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ diagonalisierbar ist. Dann gibt es eine Diagonalmatrix $D \sim A$. Nach Proposition 12.24 bilden die Darstellungsmatrizen von f bezüglich allen möglichen Wahlen einer Basis von V genau eine Ähnlichkeitsklasse. Also ist auch D eine Darstellungsmatrix von f, somit ist f diagonalisierbar. □

Proposition 12.28. *Sei $A \in \mathrm{M}_n(K)$ gegeben, und sei $S = [v_1, \ldots, v_n] \in \mathrm{M}_n(K)$ sowie $\lambda_1, \ldots, \lambda_n \in K$ gesucht. Die folgenden Aussagen beschreiben äquivalente Forderungen an die v_j und λ_j:*
(a) S ist invertierbar und $A = SDS^{-1}$ ist diagonalisierbar mit der Diagonalmatrix
$$D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$
(b) Die Spalten $v_1, \ldots, v_n$ von S sind eine Basis von K^n aus Eigenvektoren von A, und zwar ist v_j ein Eigenvektor zum Eigenwert λ_j für $j = 1, \ldots, n$.

Beweis. (a) $\implies$ (b): Die Spalten einer invertierbaren Matrix bilden immer eine Basis. Die j-te Spalte von S ist $v_j = Se_j$, und
$$Av_j = (SDS^{-1})(Se_j) = SD(S^{-1}S)e_j = SDe_j = S(\lambda_j e_j) = \lambda_j(Se_j) = \lambda_j v_j.$$
(b) $\implies$ (a): Bilden die Spalten einer Matrix S eine Basis, dann ist $S \in GL_n(K)$. Wegen $v_j = Se_j$, folgt aus
$$(SDS^{-1})v_j = SD(S^{-1}Se_j) = SDe_j = S(\lambda_j e_j) = \lambda_j(Se_j) = \lambda_j v_j = Av_j,$$

daß Linksmultiplikation mit A und SDS^{-1} auf einer Basis übereinstimmen und daher $A = SDS^{-1}$ gilt. □

Bemerkung 12.29. Propositon 12.28(b) beschreibt, wie man zu diagonalisierbarem $A \in \mathrm{M}_n(K)$ ein $S \in \mathrm{GL}_n(K)$ und eine Diagonalmatrix D finden kann mit $A = SDS^{-1}$:

- Es wird eine Basis aus Eigenvektoren gesucht, als Spalten von S.
- Die zugehörigen Eigenwerte erscheinen im zugehörigen Diagonaleintrag von D.

Proposition 12.28 besagt darüberhinaus, daß man das auch genauso machen muß.

12.4. **Matrixpotenzen.** In Anwendungen kommt es vor, daß die Zustände eines „Systems“ als Vektoren eines Vektorraums V aufgefaßt werden können. Zustandsänderungen sind dann durch eine Abbildung $f : V \to V$ beschrieben, von der wir annehmen wollen, daß sie linear ist. Angenommen, die Abbildung f beschreibt, wie sich das System in einer Zeiteinheit verändert. Wie verhält sich das System dann nach 2 oder r Zeiteinheiten? Dazu müssen wir $f^2 = f \circ f$ und $f^r = f \circ \ldots \circ f$ (mit r Faktoren) verstehen. Und wie verhält sich das System im Limes $r \to \infty$? Diese Fragen führen zu Matrixpotenzen: Iteriert man eine lineare Abbildung r-mal, dann wird die Darstellungsmatrix zur r-ten Potenz erhoben.

Definition 12.30. Sei $A \in \mathrm{M}_n(K)$ und $r \in \mathbb{N}$, dann definieren wir

$$A^r = \underbrace{A \cdot \ldots \cdot A}_{r\text{-mal}} \in \mathrm{M}_n(K).$$

Außerdem setzen wir $A^0 = \mathbf{1}$.

Lemma 12.31. *Es gelten die folgenden üblichen Potenzgesetze:*
(1) $A^{r+s} = A^r \cdot A^s$,
(2) $(A^r)^s = A^{rs}$.

Beweis. Das ist nur Abzählen der Faktoren A. □

Bemerkung 12.32. Wenn $A, B \in \mathrm{M}_n(K)$ nicht miteinander kommutieren, also $AB \neq BA$, dann ist in der Regel

$$(AB)^r \neq A^r B^r.$$

Proposition 12.33. *Seien $A \in \mathrm{M}_n(K)$ und $S \in \mathrm{GL}_n(K)$. Dann gilt für alle $r \in \mathbb{N}_0$:*

$$(SAS^{-1})^r = SA^rS^{-1}.$$

Beweis. Allgemein gilt für eine weitere Matrix $B \in \mathrm{M}_n(K)$

$$(SAS^{-1})(SBS^{-1}) = SA(S^{-1}S)BS^{-1} = SA \cdot \mathbf{1} \cdot BS^{-1} = S(AB)S^{-1},$$

so daß die Behauptung leicht per Induktion nach r folgt. □

Als Korollar halten wir fest, daß sich Potenzen von diagonalisierbaren Matrizen besonders leicht berechnen lassen.

Korollar 12.34. *Sei $A = SDS^{-1} \in \mathrm{M}_n(K)$ eine diagonalisierbare Matrix mit seiner Diagonalisierung $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ und $S \in \mathrm{GL}_n(K)$. Dann ist*

$$A^r = S \cdot \mathrm{diag}((\lambda_1)^r, \ldots, (\lambda_n)^r) \cdot S^{-1}.$$

Beweis. Das folgt aus $D^r = \mathrm{diag}((\lambda_1)^r, \ldots, (\lambda_n)^r)$, siehe Beispiel 8.20 für den Fall $n = 2$. □

Beispiel 12.35. Wir erhalten eine Formel für die Fibonacci-Folge, die rekursiv durch

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_{n+2} &= a_{n+1} + a_n \qquad \text{für alle } n \geq 0 \end{aligned}$$

definiert ist. Die Folge beginnt mit

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, \ldots$$

Mit der Matrix

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

gilt dann

$$A \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} a_{n+1} + a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix}$$

also für alle $n \geq 1$

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = A^n \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Die Eigenwertgleichung $\det(\lambda \mathbf{1} - A) = 0$ ist

$$\det \begin{pmatrix} \lambda - 1 & -1 \\ -1 & \lambda \end{pmatrix} = \lambda^2 - \lambda - 1 = 0$$

mit Nullstellen

$$\varphi = \frac{1 + \sqrt{5}}{2} \text{ und } 1 - \varphi = \frac{1 - \sqrt{5}}{2}.$$

Somit ist A diagonalisierbar mit den beiden Eigenwerten φ und $1 - \varphi$ als Diagonaleinträgen der Diagonalmatrix. Konkret

$$A = S \begin{pmatrix} \varphi & 0 \\ 0 & 1 - \varphi \end{pmatrix} S^{-1}, \qquad S = \begin{pmatrix} \varphi & 1 - \varphi \\ 1 & 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$$

(die Spalten von S sind Wahlen von Eigenvektoren zu den jeweiligen Eigenwerten).

Nach der Formel für die inverse Matrix bei 2×2-Matrizen folgt

$$S^{-1} = \frac{1}{\det(S)} \begin{pmatrix} 1 & \varphi - 1 \\ -1 & \varphi \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & \varphi - 1 \\ -1 & \varphi \end{pmatrix}.$$

Daraus berechnet sich

$$\begin{aligned} A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= S \begin{pmatrix} \varphi^n & 0 \\ 0 & (1-\varphi)^n \end{pmatrix} S^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi & 1 - \varphi \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & (1-\varphi)^n \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi & 1 - \varphi \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n \\ -(1-\varphi)^n \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{n+1} - (1-\varphi)^{n+1} \\ \varphi^n - (1-\varphi)^n \end{pmatrix} \end{aligned}$$

und folglich die geschlossene Formel für die Fibonacci–Folge

$$a_n = \frac{1}{\sqrt{5}} \Big(\varphi^n - (1-\varphi)^n \Big) = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n.$$

12.5. **Die Eigenraumzerlegung.** Ein Eigenvektor v einer linearen Abbildung $f : V \to V$ ist Eigenvektor zu nur einem einzigen Eigenwert. Sind $\lambda \neq \mu$ zwei Eigenwerte und $v \in V_\lambda \cap V_\mu$ liegt in beiden Eigenräumen V_λ und V_μ, dann folgt $v = 0$. Also ist der Schnitt der Nullvektorraum

$$V_\lambda \cap V_\mu = 0.$$

Eine Verallgemeinerung (auf mehr Eigenwerte) und Ausweitung dieses Fakts ist die Aussage des folgenden Satzes.

Satz 12.36. *Seien $\lambda_1, \ldots, \lambda_r \in K$ paarweise verschiedene Eigenwerte einer Matrix $A \in \mathrm{M}_n(K)$ (bzw. einer linearen Abbildung $f : V \to V$). Seien $v_i \in V$ (mit $V = K^n$ im Matrixfall) Eigenvektoren zum Eigenwert λ_i, $i = 1, \ldots, r$. Dann sind die Vektoren $v_1, \ldots, v_r$ linear unabhängig.*

Beweis. Der Matrixfall übersetzt sich mit $f = L_A$, der Linksmultiplikation mit A, sofort in den Fall der linearen Abbildung. Wir benutzen daher die Notation in Bezug auf $f : V \to V$.

Angenommen, der Satz wäre falsch. Dann gibt es eine nichttriviale Linearkombination

$$0 = \sum_{i=1}^{r} a_i v_i$$

mit $a_i \in K$, nicht alle 0. Unter diesen Relationen gibt es eine kürzeste, d.h., mit den wenigsten Koeffizienten $\neq 0$. Durch Verzicht auf die Eigenvektoren v_i mit $a_i = 0$ dürfen wir annehmen, daß alle $a_i \neq 0$ sind. Jetzt wenden wir f an und rechnen

$$0 = f(0) - \lambda_1 \cdot 0 = f(\sum_{i=1}^{r} a_i v_i) - \lambda_1 \cdot \sum_{i=1}^{r} a_i v_i = \sum_{i=1}^{r} a_i \cdot (f(v_i) - \lambda_1 v_i) = \sum_{i=1}^{r} a_i(\lambda_i - \lambda_1)v_i.$$

Dies ist eine neue Linearkombination mit Koeffizienten $a_i(\lambda_1 - \lambda_i)$. Da dies für $i = 1$ gleich 0 ist, ist die neue Linearkombination kürzer im obigen Sinne. Das ist ein Widerspruch. Es muß sich also um eine triviale Linearkombination handeln: für alle i ist $a_i(\lambda_1 - \lambda_i) = 0$. Weil die λ_i paarweise verschieden sind und $a_i \neq 0$ angenommen werden konnte, muß $r = 0$ oder $r = 1$ sein. Dieser Fall ist aber klar, weil die leere Menge von Vektoren linear unabhängig ist ($r = 0$), und weil $v_1 \neq 0$ als Eigenvektor gilt und damit v_1 linear unabhängig ist ($r = 1$). □

Wir verallgemeinern nun die innere direkte Summe von Definition 5.39 auf mehr als zwei Summanden.

Definition 12.37. Sei V ein K-Vektorraum. Die Summe

$$U_1 + \ldots + U_r$$

von r-vielen Unterräumen $U_1, \ldots, U_r$ von V ist **direkt**, genauer eine **innere direkte Summe** mit der Notation

$$U_1 \oplus \ldots \oplus U_r = \bigoplus_{i=1}^{r} U_i \subseteq V,$$

wenn für jeden Vektor $x \in U_1 + \ldots + U_r$ die Darstellung als $x = y_1 + \ldots + y_r$ mit $y_i \in U_i$ für $i = 1, \ldots, r$ eindeutig ist. Man sagt, daß V die **direkte Summe** der $U_1, \ldots, U_r$ ist, wenn

$$V = U_1 \oplus \ldots \oplus U_r = \bigoplus_{i=1}^{r} U_i.$$

Beispiel 12.38. Sei V ein K-Vektorraum. Sei $(b_1, \ldots, b_n)$ ein linear unabhängiges Tupel von Vektoren aus V. Die Summe der Unterräume

$$U_i = \langle b_i \rangle_K =: K \cdot b_i$$

ist eine innere direkte Summe aufgrund der linearen Unabhängigkeit der $b_1, \ldots, b_n$. Genauer gilt:

$$\mathscr{B} = (b_1, \ldots, b_n) \text{ ist Basis} \quad \iff \quad V = \bigoplus_{i=1}^{n} K \cdot b_i.$$

Proposition 12.39 (Kriterium für direkte Summe)**.** *Seien V ein K-Vektorraum und $U_1, \ldots, U_r$ Unterräume von V. Dann sind äquivalent:*

(a) *Die Summe ist direkt: $U_1 \oplus \ldots \oplus U_r \subseteq V$.*

(b) *Für alle $i = 1, \ldots, r$ gilt*

$$U_i \cap \sum_{j=1,\ j\neq i}^{r} U_j = \{0\}.$$

Beweis. (a) $\implies$ (b): Ein Vektor $x \in U_i \cap \sum_{j=1,\ j\neq i}^{r} U_j$ hat zwei Darstellungen

$$x = y_1 + \ldots + y_r = z_1 + \ldots + z_r$$

mit $y_j, z_j \in U_j$ für alle $j = 1, \ldots, r$. Die eine Darstellung hat $y_i = x$ und $y_j = 0$ für $j \neq i$, und die andere hat $z_i = 0$. Die Eindeutigkeit erzwingt $x = y_i = z_i = 0$.

(b) $\implies$ (a): Angenommen, es gibt $x \in U_1 + \ldots + U_r$ mit zwei Darstellungen

$$y_1' + \ldots + y_r' = x = y_1'' + \ldots + y_r''$$

und $y_i', y_i'' \in U_i$. Mit der Differenz $z_i = y_i' - y_i'' \in U_i$ gilt dann

$$0 = z_1 + \ldots + z_r.$$

Umgestellt erhalten wir

$$z_i = \sum_{j\neq i} (-z_j) \in U_i \cap \sum_{j\neq i}^{r} U_j = \{0\}.$$

Daher ist $z_i = 0$ und damit $y_i' = y_i''$ für alle $i = 1, \ldots, r$. Die Darstellung ist also eindeutig und die Summe direkt. □

Proposition 12.40. *Sei V ein K-Vektorraum und $U_1, \ldots, U_r$ seien Unterräume von V, deren Summe in V direkt ist. Dann ist*

$$\dim(U_1 \oplus \ldots \oplus U_r) = \sum_{i=1}^{r} \dim(U_i).$$

Beweis. Das folgt analog wie der Fall mit zwei Summanden aus Korollar 5.41. Wenn $\mathscr{B}_i$ eine Basis von U_i ist, dann ist das Tupel von Vektoren von V, das durch Hintereinanderhängen der Tupel $\mathscr{B}_i$ entsteht, eine Basis der direkten Summe. □

Satz 12.41 (Eigenraumzerlegung)**.** *Seien $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte einer Matrix A (oder eines Endomorphismus $f : V \to V$). Dann ist die Summe der Eigenräume V_{λ_i} mit $i = 1, \ldots, r$ direkt:*

$$\bigoplus_{i=1}^{r} V_{\lambda_i} \subseteq V \tag{12.1}$$

(mit $V = K^n$ im Fall der Eigenräume der Matrix A).

Beweis. Wir beweisen dies per Induktion über die Anzahl r der λ_i. Für $r = 0$ und $r = 1$ ist nichts zu beweisen.

Sei nun die Aussage richtig für Anzahlen $< r$. Nehmen wir an, daß die Summe der Eigenräume nicht direkt ist. Dann gibt es nach Proposition 12.39 ein $v_1 \in V_{\lambda_1}, \dots v_r \in V_{\lambda_r}$ und ein i mit

$$v_i = \sum_{j \neq i} v_j \neq 0.$$

Dies Widerspricht der linearen Unabhängigkeit von Eigenvektoren aus Satz 12.36, angewandt auf die Menge der v_i, welche ungleich 0 und damit Eigenvektoren zum Eigenwert λ_i sind. □

Bemerkung 12.42. Satz 12.41 besagt informell, daß Eigenräume zu paarweise verschiedenen Eigenwerten zueinander maximal linear unabhängig sind.

Korollar 12.43. *Seien $\lambda_1, \dots, \lambda_r$ paarweise verschiedene Eigenwerte einer Matrix A (oder eines Endomorphismus $f : V \to V$). Dann ist*

$$\sum_{i=1}^{r} \dim(V_{\lambda_i}) \leq \dim(V)$$

(mit $V = K^n$ im Fall der Eigenräume der Matrix A). Insbesondere ist die Anzahl der Eigenwerte durch

$$|\{\lambda \ ; \ \lambda \text{ ist Eigenwert von } A \text{ (oder } f)\}| \leq \dim(V)$$

nach oben beschränkt.

Beweis. Unmittelbar klar aus Satz 12.41, weil $\dim(V_\lambda) \geq 1$ für jeden Eigenwert λ. □

Als naheliegende Frage drängt sich nun auf: Wann gilt in (12.1) Gleichheit? Für welche Endomorphismen beschreibt die Summe der Eigenräume alles?

Beispiel 12.44. Wir schauen uns zunächst ein Beispiel an, das zeigt, daß in (12.1) nicht immer Gleicheit herrschen muß, und zwar die Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{M}_2(K).$$

Ein $\lambda \in K$ ist Eigenwert, wenn $\det(\lambda \mathbf{1} - A) = (\lambda - 1)^2 = 0$ ist. Es gibt also nur den Eigenwert $\lambda = 1$. Der zugehörige Eigenraum V_1 enthält e_1, ist aber nicht 2-dimensional, weil sonst $V_1 = K^2$ und $A = \mathbf{1}$, was ersichtlich nicht stimmt. Also ist

$$V_1 = K \cdot e_1 = \langle e_1 \rangle_K$$

und in (12.1) gilt nicht Gleichheit.

Satz 12.45 (Diagonalisierbarkeit)**.** *Sei $A \in \mathrm{M}_n(K)$, und seien $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A. Dann sind äquivalent:*

(a) *A ist diagonalisierbar.*
(b) *K^n hat eine Basis aus Eigenvektoren von A.*
(c) *$V_{\lambda_1} \oplus \dots \oplus V_{\lambda_r} = K^n$.*
(d) *$\sum_{i=1}^{r} \dim(V_{\lambda_i}) = n$.*

Beweis. (a) $\iff$ (b): Proposition 12.28.

(b) $\implies$ (c): Sei $v_1, \dots, v_n$ eine Basis aus Eigenvektoren von A. Dann gilt

$$K^n = \langle v_1, \dots, v_n \rangle_K \subseteq V_{\lambda_1} \oplus \dots \oplus V_{\lambda_r} \subseteq K^n.$$

Folglich beschreibt die innere direkte Summe $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_r}$ den gesamten K^n.

(c) $\Longrightarrow$ (b): Wir wählen als Basis von K^n eine Basis, die aus der Aneinanderreihung von Basen der Eigenräume V_{λ_i} besteht, siehe der Beweis von Proposition 12.40.

(c) $\iff$ (d): Nach Satz 12.41 ist die Summe der Eigenräume $U := V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r}$ direkt. Seine Dimension ist demnach $\dim(U) = \sum_{i=1}^r \dim(V_{\lambda_i})$. Aus Korollar 5.33 folgt die Äquivalenz von $U = K^n$, also (c), mit $\dim(U) = \dim(K^n) = n$, und das ist (d). □

Korollar 12.46. *Sei $A = SDS^{-1}$ mit $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ und $S \in \mathrm{GL}_n(K)$. Dann ist für alle $\lambda \in K$ der Eigenraum V_λ von A zum Eigenwert λ genau*

$$V_\lambda = \langle Se_j \ ; \ j \text{ mit } \lambda_j = \lambda \rangle_K$$

und diese Spalten Se_j von S bilden eine Basis von V_λ.

Beweis. Es gilt sicherlich $V_\lambda(S) := \langle Se_j \ ; \ j \text{ mit } \lambda_j = \lambda \rangle_K \subseteq V_\lambda$ und die betrachteten Spalten sind als Spalten einer invertierbaren Matrix S linear unabhängig. Weiter folgt

$$n = \dim(C(S)) = \sum_\lambda \dim(V_\lambda(S)) \leq \sum_\lambda \dim(V_\lambda) = n.$$

Also herrscht Gleichheit der Dimension $\dim(V_\lambda(S)) = \dim(V_\lambda)$, und daher $V_\lambda(S) = V_\lambda$. □

Korollar 12.47 (Diagonalisierbarkeit von Endomorphismen)**.** *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums V, und seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A. Dann sind äquivalent:*

(a) *f ist diagonalisierbar.*
(b) *V hat eine Basis aus Eigenvektoren von f.*
(c) *$V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r} = V$.*
(d) *$\sum_{i=1}^r \dim(V_{\lambda_i}) = \dim(V)$.*

Beweis. Das ist die Übersetzung von Satz 12.45 für Endomorphismen. □

Übungsaufgaben zu §12

Übungsaufgabe 12.1. Sei V ein K-Vektorraum $\neq 0$. Bestimmen sie Eigenwerte und Eigenvektoren für die Identität $\mathrm{id}_V : V \to V$ und die Nullabbildung $0 : V \to V$.

Übungsaufgabe 12.2. Seien K ein Körper und $a, x, b \in K$. Zeigen Sie, daß die Matrix

$$A = \begin{pmatrix} a & x \\ 0 & b \end{pmatrix}$$

genau dann diagonalisierbar ist, wenn $a \neq b$ oder $x = 0$ gilt.

Übungsaufgabe 12.3. Seien K ein Körper, $a, x, b \in K$ und

$$A = \begin{pmatrix} a & x \\ 0 & b \end{pmatrix}.$$

Bestimmen Sie eine Formel für A^n.

Übungsaufgabe 12.4. Sei $A \in \mathrm{M}_n(K)$ eine diagonalisierbare Matrix.

(a) Beschreiben Sie alle Diagonalmatrizen D, für die es ein $S \in \mathrm{GL}_n(K)$ gibt mit

$$A = SDS^{-1}.$$

(b) Beantworten Sie diesselbe Frage für die möglichen $S \in \mathrm{GL}_n(K)$, welche A diagonalisieren.

Übungsaufgabe 12.5. Sei $A \in \mathrm{M}_n(K)$. Zeigen Sie:

(1) $A \in \mathrm{GL}_n(K) \iff 0$ ist kein Eigenwert von A.

(2) Sei $A \in \mathrm{GL}_n(K)$ und $\lambda \in K$ ein Eigenwert. Dann ist λ^{-1} ein Eigenwert von A^{-1}.

Übungsaufgabe 12.6. Seien $A, B \in \mathrm{GL}_n(K)$ invertierbare Matrizen mit $AB \neq BA$. Zeigen Sie, daß dann

$$(AB)^2 \neq A^2B^2$$

gilt, und finden Sie für $K = \mathbb{R}$ und $K = \mathbb{F}_2$ sowie $n = 2$ jeweils ein Beispiel.

Übungsaufgabe 12.7. Sei $A \in \mathrm{M}_n(K)$. Zeigen Sie, daß A und A^t die gleichen Eigenwerte haben, indem Sie

$$\chi_A(X) = \chi_{A^t}(X)$$

beweisen.

13. Die Spur

Definition 13.1. Sei K ein Körper.

(1) Die **Spur** einer quadratischen Matrix $A = (a_{ij}) \in \mathrm{M}_n(K)$ ist die Summe der Diagonalelemente

$$\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}.$$

(2) Die **Spur** eines Endomorphismus $f : V \to V$ eines endlichdimensionalen K-Vektorraums V ist bezüglich jeder Basis $\mathscr{B}$ von V die Spur der Darstellungsmatrix

$$\mathrm{tr}(f) = \mathrm{tr}(\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)).$$

Wir zeigen in Korollar 13.3, daß die Spur eines Endomorphismus nicht von der Wahl der Basis für die Darstellungsmatrix abhängt.

Die Spur hat die folgenden algebraischen Eigenschaften.

Proposition 13.2. *(1) Die Spur* $\mathrm{tr} : \mathrm{M}_n(K) \to K$ *ist* K*-linear, und für alle* $A, B \in \mathrm{M}_n(K)$ *gilt*

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

(2) Die Spur $\mathrm{tr} : \mathrm{End}_K(V) \to K$ *ist linear, und für alle* $f, g \in \mathrm{End}_K(V)$ *gilt*

$$\mathrm{tr}(f \circ g) = \mathrm{tr}(g \circ f).$$

Beweis. (1) Sei $A = (a_{ij})$ und $B = (b_{ij})$. Aus der Definition der Spur und der K-linearen Struktur von $\mathrm{M}_n(K)$ folgt für $\mu \in K$, daß

$$\mathrm{tr}(\mu A + B) = \sum_{i=1}^{n} (\mu A + B)_{ii} = \sum_{i=1}^{n} \mu a_{ii} + b_{ii} = \mu \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \mu \mathrm{tr}(A) + \mathrm{tr}(B).$$

Weiter gilt

$$\mathrm{tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i,j=1}^{n} a_{ij} b_{ji} = \sum_{i,j=1}^{n} b_{ji} a_{ij} = \sum_{j=1}^{n} (BA)_{jj} = \mathrm{tr}(BA).$$

Aussage (2) folgt sofort aus (1), weil der Übergang zur Darstellungsmatrix linear ist und Komposition in Matrixmultiplikation übersetzt. □

Korollar 13.3. *Sei* K *ein Körper.*

(1) Ähnliche Matrizen haben gleiche Spur: sei $A \in \mathrm{M}_n(K)$ *und* $S \in \mathrm{GL}_n(K)$*. Dann gilt*

$$\mathrm{tr}(A) = \mathrm{tr}(SAS^{-1}).$$

(2) Sei V *ein endlichdimensionaler* K*-Vektorraum. Die Spur* $\mathrm{tr}(f)$ *eines Endomorphismus* $f : V \to V$ *ist wohldefiniert.*

Beweis. (1) Aus Proposition 13.2 folgt für die Faktoren S und AS^{-1}

$$\operatorname{tr}(SAS^{-1}) = \operatorname{tr}(AS^{-1}S) = \operatorname{tr}(A\mathbf{1}_n) = \operatorname{tr}(A).$$

(2) Basiswechsel ändert die Darstellungsmatrix A zu einem SAS^{-1}, siehe Proposition 12.24. Damit folgt (2) sofort aus (1). □

Spur und Determinante treten (bis auf ein Vorzeichen) als gewisse Koeffizienten des charakteristischen Polynoms auf.

Proposition 13.4. *Das charakteristische Polynom hat die folgende Form:*

(1) *Sei $A = (a_{ij}) \in \mathrm{M}_n(K)$. Das charakteristische Polynom hat Grad*

$$\deg(\chi_A(X)) = n$$

und die Form

$$\chi_A(X) = X^n - \operatorname{tr}(A) \cdot X^{n-1} + \ldots + (-1)^n \det(A).$$

(2) *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums. Dann gilt*

$$\deg(\chi_f) = \dim(V)$$

und, wenn $\dim(V) = n$, dann hat das charakteristische Polynom die Form

$$\chi_f(X) = X^n - \operatorname{tr}(f) \cdot X^{n-1} + \ldots + (-1)^n \det(f).$$

Beweis. (2) folgt sofort aus (1) angewandt auf eine Darstellungsmatrix von f. Für (1) analysieren wir den Grad der Summanden in der Leibniz-Formel zu $\det(X \cdot \mathbf{1} - A)$. Jeder Summand ist ein Produkt von n-Faktoren der Form $(X - a_{ii})$ oder $-a_{ij}$ mit $i \neq j$. Im Summanden zu $\sigma \in S_n$ ist damit der Exponent einer auftretenden X-Potenz nach Ausmultiplizieren des Pordukts höchstens

$$\text{Anzahl der Faktoren } (X - a_{ii}) = |\{i \ ; \ \sigma(i) = i\}| = \text{ Anzahl der Fixpunkte von } \sigma.$$

Das Maximum wird von der Identität $\sigma = \mathrm{id}$ (auch wirklich) erreicht. Alle anderen $\sigma \neq \mathrm{id}$ haben höchstens $n - 2$ Fixpunkte, somit beeinflussen diese Summanden die Koeffizienten von X^n und X^{n-1} nicht: die entsprechenden Summanden haben Grad $\leq n - 2$. Daher gilt

$$\begin{aligned}\chi_A(X) &= \prod_{i=1}^{n}(X - a_{ii}) + \text{ Terme vom Grad } \leq n - 2 \\ &= X^n - (a_{11} + a_{22} + \ldots + a_{nn})X^{n-1} + \text{ Terme kleineren Grades} \\ &= X^n - \operatorname{tr}(A) \cdot X^{n-1} + \text{ Terme kleineren Grades.}\end{aligned}$$

Es fehlt noch der konstante Term (mit X^0) in $\chi_A(X)$. Diesen bekommt man durch Auswertung in $X = 0$:

$$\chi_A(0) = \det(0 \cdot \mathbf{1} - A) = \det(-A) = (-1)^n \det(A).$$ □

Korollar 13.5. *Sei $A \in \mathrm{M}_n(K)$ eine diagonalisierbare quadratische Matrix mit Diagonalisierung $A = SDS^{-1}$ und Diagonalmatrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Dann gilt:*

(1) $\det(A) = \prod_{i=1}^{n} \lambda_i$,

(2) $\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i$,

(3) $\chi_A(X) = \prod_{i=1}^{n}(X - \lambda_i)$.

Beweis. Nach Proposition 12.17 und Beispiel 12.21 gilt

$$\chi_A(X) = \chi_{SDS^{-1}}(X) = \chi_D(X) = \prod_{i=1}^{n}(X - \lambda_i).$$

Die Formeln für $\mathrm{tr}(A)$ und $\det(A)$ folgen aus dem Koeffizientenvergleich mittels Propositiopn 13.4 und Ausmultiplizieren:

$$\chi_A(X) = X^n - \Big(\sum_{i=1}^{n} \lambda_i\Big) X^{n-1} + \ldots + (-1)^n \cdot \prod_{i=1}^{n} \lambda_i. \qquad \square$$

Für eine Blockmatrix mit quadratischen Diagonalmatrizen berechnet sich die Spur offensichtlich additiv aus den Spuren der diagonalen Blöcke:

$$\mathrm{tr}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathrm{tr}(A) + \mathrm{tr}(D).$$

Die entsprechende multiplikative Aussage für Determinanten gilt auch, sofern es sich um eine obere Blockdreiecksmatrrix handelt.

Proposition 13.6. *Für eine quadratische Matrix in oberer Blockdreiecksmatrixform*

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

mit quadratischen Diagonalen A und D gilt

$$\det(M) = \det(A) \cdot \det(D).$$

Beweis. Das kann man aus der Leibniz-Formel ableiten: nur die Permutationen σ tragen bei, welche die Spaltenindizes aus dem Bereich der Spalten von A nicht mit denen aus dem Bereich der Spalten von B und D mischen. Solche führen nämlich zu einem Faktor 0 aufgrund eines Faktors aus dem links-unteren Block. Die Leibniz-Formel wird dann zu einem Produkt der Leibnizformeln für A und für D. Etwas genauer:

Sei $A \in \mathrm{M}_r(K)$, $D \in \mathrm{M}_s(K)$ und $M = (a_{ij}) \in \mathrm{M}_n(K)$ mit $n = r + s$. Zu jedem Paar von $\sigma_1 \in S_r$ und $\sigma_2 \in S_s$ betrachten wir die Permutation $\sigma \in S_n$ definiert durch

$$\sigma(i) = \begin{cases} \sigma_1(i) & i \leq r \\ \sigma_2(i-r) + r & i > r, \end{cases}$$

welche σ_1 auf $\{1, \ldots, r\}$ mit σ_2 auf $\{r+1, \ldots, r+s\}$ (bis auf Translation mit r) zusammensetzt. Die so beschriebene Abbildung

$$S_r \times S_s \to S_n$$

ist injektiv, und das Bild sind genau die Permutationen, welche die ersten r Indizes nicht mit den hinteren s Indizes mischen. Es gilt mit dieser Notation

$$\mathrm{sign}(\sigma) = \mathrm{sign}(\sigma_1) \cdot \mathrm{sign}(\sigma_2),$$

wie man durch Abzählen der nötigen Transpositionen in einer Produktdarstellung sieht. Dann gilt

$$\begin{aligned} \det(A) \cdot \det(D) &= \Big(\sum_{\sigma_1 \in S_r} \mathrm{sign}(\sigma_1) \prod_{i=1}^{r} a_{i,\sigma_1(i)}\Big) \cdot \Big(\sum_{\sigma_2 \in S_s} \mathrm{sign}(\sigma_2) \prod_{i=r+1}^{n} a_{i,\sigma_2(i-r)+r}\Big) \\ &= \sum_{(\sigma_1,\sigma_2) \in S_r \times S_s} \mathrm{sign}(\sigma_1)\mathrm{sign}(\sigma_2) \prod_{i=1}^{r} a_{i,\sigma_1(i)} \cdot \prod_{i=r+1}^{n} a_{i,\sigma_2(i-r)+r} \\ &= \sum_{\sigma \in S_n, \text{ nicht mischend}} \mathrm{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = \det(M), \end{aligned}$$

weil die anderen Summanden in der Leibniz-Formel für $\det(M)$ jeweils einen Faktor im linken unteren Block haben, der 0 ist und das Produkt annihiliert.

Alternativ kann man auch mit elementaren Zeilenoperationen A zur oberen Driecksmatrix A' und D zur oberen Dreiecksmatrix D' machen. Dabei seien für A (bzw. D) a (bzw. d Vertauschungen vorgenommen worden. Dann gilt

$$\det(A) = (-1)^a \det(A') \quad \text{und} \quad \det(D) = (-1)^d \det(D').$$

Wenn man dieselben Zeilenoperationen auf M anwendet, dann wird mit $a+d$-vielen Vertauschungen aus M die Matrix

$$M' = \begin{pmatrix} A' & B' \\ 0 & D' \end{pmatrix},$$

wobei wir über B' nichts sagen müssen. Nach dem Beispiel der Determinante einer oberen Dreiecksmatrix folgt nun

$$\det(M) = (-1)^{a+d} \det M' = (-1)^{a+d} \det(A') \cdot \det(D') = \det(A) \cdot \det(D). \qquad \square$$

Korollar 13.7. *Für eine quadratische Matrix in oberer Blockdreiecksmatrixform*

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

mit quadratischen Diagonalen A und D gilt:

$$\chi_M(X) = \chi_A(X) \cdot \chi_D(X).$$

Beweis. Es gilt nach Proposition 13.6 angewandt auf $X \cdot \mathbf{1} - M$:

$$\chi_M(X) = \det(X \cdot \mathbf{1} - M) = \det(X \cdot \mathbf{1} - A) \cdot \det(X \cdot \mathbf{1} - D) = \chi_A(X) \cdot \chi_D(X). \qquad \square$$

Übungsaufgaben zu §13

Übungsaufgabe 13.1. Sei $A \in \mathrm{M}_{m\times n}(K)$ und $B \in \mathrm{M}_{n\times m}(K)$. Zeigen Sie

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

(Beachten Sie, daß AB und BA bei $n \neq m$ gar nicht dieselben Abmessungen haben!)

Übungsaufgabe 13.2. Sei $A \in \mathrm{M}_2(K)$. Zeigen Sie

$$\mathrm{tr}(A) = \det(\mathbf{1} + A) - 1 - \det(A).$$

Gilt dies auch allgemein für $A \in \mathrm{M}_n(K)$ und beliebiges n?

Teil 4. Symmetrische Bilinearformen

14. Bilinearformen und Orthogonalität

14.1. **Matrixbeschreibung von Bilinearformen.** Wir wiederholen den Begriff der multilinearen Abbildung aus Definition 11.23 im Spezialfall einer Bilinearform.

Definition 14.1. Sei K ein Körper und sein V ein K-Vektorraum. Eine **(K-)Bilinearform** auf V ist eine Abbildung von Mengen

$$f : V \times V \to K$$

mit den folgenden Eigenschaften: für alle $u, v, v_1, v_2 \in V$ und $\lambda \in K$ gilt

$$\begin{aligned} f(u, v_1 + v_2) &= f(u, v_1) + f(u, v_2), \\ f(v_1 + v_2, u) &= f(v_1, u) + f(v_2, u), \\ f(\lambda u, v) &= f(u, \lambda v) = \lambda f(u, v). \end{aligned}$$

Manchmal ziehen wir eine Notation mit Klammern vor. Wir schreiben dann

$$\langle u, v\rangle := f(u, v).$$

Bemerkung 14.2. Man sagt kurz: die Bilinearform ist **linear in beiden Argumenten**, und meint, wenn man ein Argument fixiert, so ist die verbleibende Abbildung im anderen Argument linear. Für eine Bilinearform $f : V \times V \to K$ gilt für alle $v \in V$

$$f(v, 0) = 0 = f(0, v),$$

denn lineare Abbildungen bilden 0 auf 0 ab.

Beispiel 14.3. (1) Das **Standardskalarprodukt** auf K^n aus Definition 8.13 ist das Standardbeispiel einer Bilinearform $K^n \times K^n \to K$. Die Bilinearität von

$$x \bullet y = x^t y = \sum_{i=1}^{n} x_i y_i$$

für $x, y \in K^n$ wurde in Proposition 8.16 notiert.

(2) Mittels der Spur quadratischer $n \times n$-Matrizen aus Definition 13.1 definieren wir die Spurform auf $\mathrm{M}_n(K)$ durch

$$\langle\ ,\ \rangle_{|tr} : \mathrm{M}_n(K) \times \mathrm{M}_n(K) \to K, \qquad \langle A, B\rangle_{\mathrm{tr}} := \mathrm{tr}(AB).$$

In Koordinaten $A = (a_{ij})$, $B = (b_{ij})$ gilt

$$\mathrm{tr}(AB) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji}.$$

In dieser Form ist die Spurform ersichtlich bilinear. Die Bilinearität folgt aber auch aus dem Distributivgesetz der Matrixmultiplikation und der Linearität der Spur.

Beispiel 14.4. Mit $A = (a_{ij}) \in \mathrm{M}_n(K)$ definieren wir eine Bilinearform $\langle\ ,\ \rangle_A$ auf K^n durch

$$\langle x, y\rangle_A := x \bullet Ay = x^t Ay.$$

für alle $x, y \in K^n$. In Koordinaten $x = (x_1, \ldots, x_n)^t$ und $y = (y_1, \ldots, y_m)^t$ ergibt sich

$$\langle x, y\rangle_A = x^t Ay = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j a_{ij}.$$

Wir rechnen exemplarisch für $x, y, y' \in K^n$ und $\lambda \in K$:

$$\langle x, \lambda y + y'\rangle_A = x^t A(\lambda y + y') = \lambda x^t Ay + x^t Ay' = \lambda \langle x, y\rangle_A + \langle x, y'\rangle_A.$$

Die Matrixeinträge bekommen wir als Werte der Paarung auf den Standardbasisvektoren zurück:

$$\langle e_i, e_j\rangle_A = e_i^t A e_j = a_{ij}. \tag{14.1}$$

Der Fall $A = \mathbf{1}_n$ der Einheitsmatrix führt erneut zum Standardskalarprodukt auf K^n.

Wir wollen nun einsehen, daß Beispiel 14.4 typisch ist: mittels Koordinaten läßt sich jede Bilinearform auf einem endlich-dimensionalen K-Vektorraum so beschreiben.

Definition 14.5. Die **Gram'sche Matrix** einer Bilinearform $f : V \times V \to K$ auf dem endlich-dimensionalen K-Vektorraum V bezüglich der Basis $\mathscr{B} = (b_1, \dots, b_n)$ ist die $n \times n$-Matrix

$$\mathrm{Gram}^{\mathscr{B},\mathscr{B}}(f) = (f(b_i, b_j)) \in \mathrm{M}_n(K).$$

Beispiel 14.6. Die Gleichung (14.1) besagt, daß die Gram'sche Matrix zu $\langle\ ,\ \rangle_A$ bezüglich der Standardbasis $\mathscr{E} = \{e_1, \dots, e_n\}$ gerade A ist:

$$\mathrm{Gram}^{\mathscr{E},\mathscr{E}}(\langle\ ,\ \rangle_A) = A.$$

Proposition 14.7. *Sei $f : V \times V \to K$ eine Bilinearform auf dem endlich-dimensionalen K-Vektorraum V mit Basis $\mathscr{B} = (b_1, \dots, b_n)$. Sei $A = \mathrm{Gram}^{\mathscr{B},\mathscr{B}}(f)$ die Gram'sche Matrix der Bilinearform. Dann gilt für alle $v, w \in V$:*

$$f(v, w) = \kappa_{\mathscr{B}}(v)^t A \kappa_{\mathscr{B}}(w) = \langle \kappa_{\mathscr{B}}(v), \kappa_{\mathscr{B}}(w)\rangle_A. \tag{14.2}$$

Die Gram'sche Matrix bezüglich $\mathscr{B}$ bestimmt die dargestellte Bilinearform auf V eindeutig, und jede Matrix aus $\mathrm{M}_n(K)$ kommt als Gram'sche Matrix bezüglich $\mathscr{B}$ einer Bilinearform auf V vor.

Bemerkung 14.8. Proposition 14.7 besagt, daß in Koordinaten für V die Bilinearform f durch die von der Gram'schen Matrix zu f definierte Bilinearform beschrieben wird: das folgende Diagramm kommutiert.

$$\begin{array}{ccc} V \times V & \xrightarrow{f} & K \\ {\scriptstyle \kappa_{\mathscr{B}} \times \kappa_{\mathscr{B}}}\downarrow & & \| \\ K^n \times K^n & \xrightarrow{\langle\ ,\rangle_A} & K. \end{array}$$

Beweis. Sei $A = (a_{ij}) \in \mathrm{M}_n(K)$ die Gram'sche Matrix von f. Wir müssen für $v = \sum_{i=1}^n x_i b_i$ und $w = \sum_{j=1}^n y_j b_j$ die Gleichung (14.2) nachweisen:

$$\begin{aligned} f(v, w) &= f\Big(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n y_j b_j\Big) = \sum_{i=1}^n x_i f\Big(b_i, \sum_{j=1}^n y_j b_j\Big) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j f(b_i, b_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} y_j = \sum_{i=1}^n x_i \big(A\kappa_{\mathscr{B}}(w)\big)_i = \kappa_{\mathscr{B}}(v)^t A \kappa_{\mathscr{B}}(w). \end{aligned}$$

Aus (14.2) folgt sofort, daß die Biliearform durch ihre Gram'sche Matrix eindeutig bestimmt wird. Für die letzte Behauptung nutzen wir aus, daß für jedes $A \in \mathrm{M}_n(K)$ durch (14.2) eine Bilinearform auf V definiert wird, deren Gram'sche Matrix A ist:

$$\kappa_{\mathscr{B}}(b_i)^t A \kappa_{\mathscr{B}}(b_j) = e_i \bullet A e_j = \ ij\text{-Eintrag von } A. \qquad \square$$

Die Gram'sche Matrix einer Bilinearform ändert sich, wenn man die verwendete Basis wechselt, durch eine Formel, welche die zugehörigen Basiswechselmatrizen benötigt.

Proposition 14.9. *Sei f eine Bilinearform auf dem endlich-dimensionalen K-Vektorraum V mit Basis $\mathscr{B} = (b_1, \ldots, b_n)$. Sei $A = \mathrm{Gram}^{\mathscr{B},\mathscr{B}}(f)$ die Gram'sche Matrix.*

Sei $\mathscr{B}' = (b'_1, \ldots, b'_n)$ eine weitere Basis von V, und sei $A' = \mathrm{Gram}^{\mathscr{B}',\mathscr{B}'}(f)$ die zugehörige Gram'sche Matrix. Dann gilt

$$A' = S^t A S$$

mit der Basiswechselmatrix

$$S = S^{\mathscr{B}'}_{\mathscr{B}} = \mathrm{M}^{\mathscr{B}'}_{\mathscr{B}}(\mathrm{id}_V) \in \mathrm{GL}_n(K).$$

Beweis. Die Basiswechselmatrix S transformiert Koordinaten als

$$\kappa_{\mathscr{B}}(b'_i) = S\kappa_{\mathscr{B}'}(b'_i) = Se_i.$$

Der ij-te Eintrag von A' sei a'_{ij}. Damit gilt nach Proposition 14.7

$$\begin{aligned} a'_{ij} &= f(b'_i, b'_j) = \kappa_{\mathscr{B}}(b'_i)^t A \kappa_{\mathscr{B}}(b'_i) = (Se_i)^t A (Se_j) \\ &= e_i^t S^t A S e_j = e_i \bullet (S^t A S) e_j = ij\text{-ter Eintrag von } S^t A S. \end{aligned}$$

□

14.2. **Vertauschungssymmetrie.**

Definition 14.10. (1) Eine **Bilinearform** $f : V \times V \to K$ auf einem K-Vektorraum V heißt **symmetrisch**, wenn für alle $v, w \in V$ gilt:

$$f(v, w) = f(w, v).$$

(2) Eine quadratische Matrix $A \in \mathrm{M}_n(K)$ heißt **symmetrisch**, wenn $A^t = A$.

Beispiel 14.11. Das Standardskalarprodukt $\langle x, y\rangle = x^t y$ auf K^n ist symmetrisch. Ebenso ist die Spurform auf $\mathrm{M}_n(K)$ symmetrisch nach Proposition 13.2.

Proposition 14.12. *Sei V ein K-Vektorraum mit Basis $\mathscr{B}$ und Bilinearform $f : V \times V \to K$. Sei $A = \mathrm{Gram}^{\mathscr{B},\mathscr{B}}(f)$ die Gram'sche Matrix. Dann gilt:*

$$f \text{ ist symmetrisch} \iff A \text{ ist symmetrisch.}$$

Beweis. Ist f symmetrisch, so gilt für alle $1 \le i, j \le n$

$$A_{ij} = f(b_i, b_j) = f(b_j, b_i) = (A^t)_{ij}, \tag{14.3}$$

Das bedeutet $A = A^t$ ist symmetrisch.

Für die Umkehrung müssen wir aus (14.3) folgern, daß f symmetrisch ist. Seien $v, w \in V$ beliebig. Wir setzen $x = \kappa_{\mathscr{B}}(v)$ und $y = \kappa_{\mathscr{B}}(w)$. Mit Proposition 14.7 rechnen wir wegen $A = A^t$ und $x = (x^t)^t$

$$f(v, w) = x^t A y = (x^t A y)^t = y^t A^t (x^t)^t = y^t A x = f(w, v),$$

wobei wir ausgenutzt haben, daß die 1×1-Matrix $x^t A y$ beim Transponieren ihren Wert nicht ändert. □

Bemerkung 14.13. Im Folgenden wird für einen Körper oft die Voraussetzung $2 \in K^\times$ gefordert werden. Diese Notation muß erklärt werden. Dazu erinnern wir daran, daß jeder Körper Elemente $0, 1 \in K$ hat, und damit für jedes $n \in \mathbb{N}_0$ ein Element (zur Unterscheidung kurzzeitig n_K benannt)

$$n_K = \underbrace{1 + \ldots + 1}_{n\text{-mal}} \in K.$$

Es ist also $0_K = 0$ und $1_K = 1$ im jeweiligen Körper. Man überzeugt sich leicht, daß für alle $n, m \in \mathbb{N}$ gilt

$$(n + m)_K = n_K + m_K.$$

Allgemeiner, mittels additiven inversen Elementen $(-n)_K = -(n_K)$, erhält man einen Ringhomomorphismus $\mathbb{Z} \to K$

$$n \mapsto n_K.$$

Wenn im Folgenden von $2 \in K$ die Rede ist, so ist stets $2_K = 1 + 1 \in K$ gemeint. Für die Körper $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ ist $2 \neq 0$, und damit gilt $2 \in K^\times$. Aber für $K = \mathbb{F}_2$ oder auch $\mathbb{F}_q$ mit einer 2er Potenz q ist $2 = 0$, und damit gilt $2 \notin K^\times$.

Je nach Anwendungsgebiet ist die Voraussetzung $2 \in K^\times$ oft, oder, wenn Sie zum Beispiel ein Computer sind, selten erfüllt.

Wir zeigen nun eine Variante der binomischen Formel. In der Tat handelt es bei dem folgenden Satz im Fall des $\mathbb{R}^1$ mit dem Stadardskalarprodukt um die erste binomische Formel. Und der Beweis ist auch derselbe.

Satz 14.14. *Sei f eine symmetrische Bilinearform auf dem K-Vektorraum V. Dann gilt für alle $v, w \in V$ die* ***Polarisationsformel****:*

$$2f(v, w) = f(v + w, v + w) - f(v, v) - f(w, w). \tag{14.4}$$

Beweis. Wegen Bilinearität gilt

$$f(v + w, v + w) = f(v, v + w) + f(w, v + w) = f(v, v) + f(v, w) + f(w, v) + f(w, w),$$

und die Symmetrie zeigt $f(v, w) + f(w, v) = 2f(v, w)$. Hieraus folgt die Polarisationsformel. □

Bemerkung 14.15. Die Polarisationsformel zeigt, daß eine symmetrische Bilinearform f durch die Werte $f(v, v)$ für alle $v \in V$ eindeutig bestimmt ist, sofern $2 \in K^\times$ gilt.

Proposition 14.16. *Sei $A \in \mathrm{M}_n(K)$. Dann gilt für alle $x, y \in K^n$*

$$Ax \bullet y = x \bullet A^t y.$$

Insbesondere gilt für symmetrische A bereits $Ax \bullet y = x \bullet Ay$.

Beweis. Das folgt aus den algebraischen Eigenschaften des Transponierens und der Rechnung

$$Ax \bullet y = (Ax)^t y = (x^t A^t) y = x^t (A^t y) = x \bullet A^t y.$$ □

14.3. **Orthogonalität und Orthogonalbasen.** Letztendlich wollen wir mit Bilinearformen Geometrie mit Längen und Winkeln betreiben. Der Begriff der Orthogonalität ist ein Vorgriff hierauf. Orthogonale Vektoren „stehen senkrecht aufeinander“, und, was das bedeutet, motivieren wir in der euklidischen Anschauungsebene $\mathbb{A}^2(\mathbb{R}) := \mathbb{R}^2$.

Wir beginnen mit dem **Satz des Pythagoras**. Sei z der Abstand der Punkte[4] $\binom{x}{y}$ und $\binom{0}{0}$ im $\mathbb{A}^2(\mathbb{R}) = \mathbb{R}^2$. Dann zeigt eine elementargeometrische Überlegung, daß

$$z^2 = x^2 + y^2.$$

Wir benutzen die Addition von Flächeninhalten bei disjunkter Überdeckung, sowie die Flächeninhaltsformel für Quadrate und Dreiecke.

[4]Ich habe hier absichtlich nicht die Buchstaben a, b und c gewählt, weil es darauf nicht ankommt.

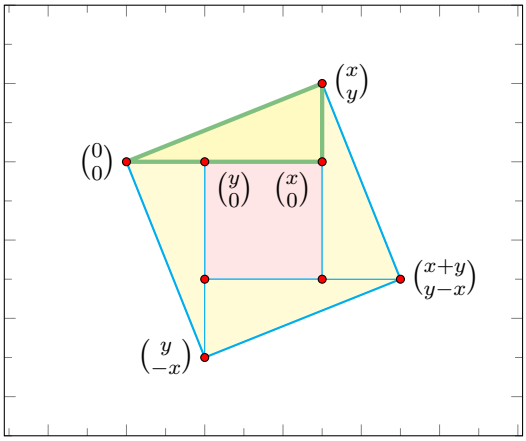


Abbildung 7. Geometrischer Beweis des Satz des Pythagoras.

Die Fläche des großen Quadrats z^2 entspricht der Fläche des kleinen Quadrats $(x-y)^2$ und viermal der Fläche des gelben Dreiecks $xy/2$:

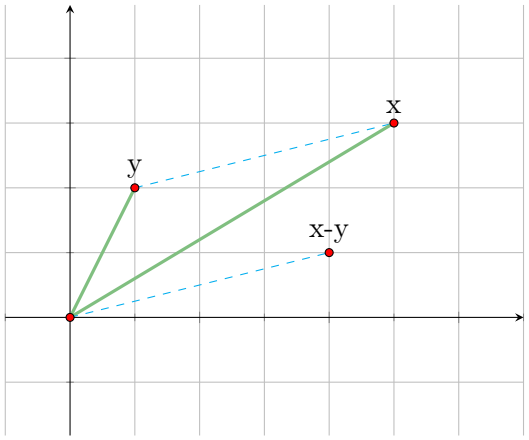
$$z^2 = (x-y)^2 + 4(xy/2) = x^2 + y^2.$$

Definition 14.17. Wir definieren die **Länge eines Vektors** $x = \binom{x_1}{x_2} \in \mathbb{R}^2$ als

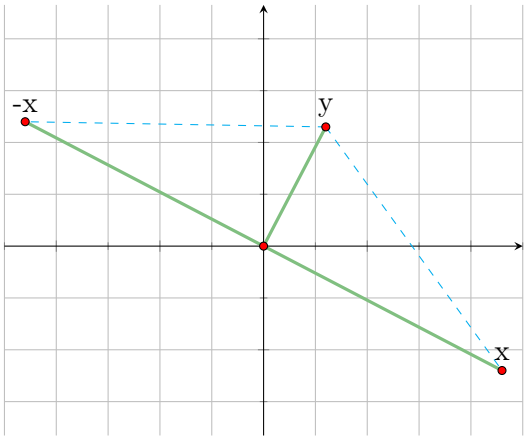
$$\|x\| = \sqrt{x \bullet x} = \Big(\sum_{i=1}^{2} x_i^2\Big)^{1/2} = \sqrt{(x_1)^2 + (x_2)^2},$$

und den **Abstand**[a] von $x, y \in \mathbb{A}^2(\mathbb{R})$ als $d(x,y) := \|x - y\|$.

[a] Das Symbol d steht für Distanz.



(a) Abstand von Punkten im $\mathbb{R}^n$, ein Parallelogramm.



(b) Orthogonale Vektoren.

Abbildung 8. Euklidische Motivation für Länge und Orthogonalität.

Elementargeometrisch motiviert sagen wir, daß Vektoren x und y aus $\mathbb{R}^2$ **orthogonal (senkrecht)** sind, wenn y denselben Abstand von x wie von $-x$ hat. Wir rechnen zunächst

$$\begin{aligned} d(y,-x)^2 - d(y,x)^2 &= \|y+x\|^2 - \|y-x\|^2 = \langle y+x, y+x\rangle - \langle y-x, y-x\rangle \\ &= \langle y, y+x\rangle + \langle x, y+x\rangle - \langle y, y-x\rangle + \langle x, y-x\rangle = \langle y, 2x\rangle + \langle x, 2y\rangle \\ &= 4\langle x, y\rangle, \end{aligned}$$

wobei wir die Symmetrie des Standardskalarproduktes benutzt haben. Dies führt zu

$$x \text{ orthogonal zu } y \iff d(y,-x)^2 = d(y,x)^2 \iff \langle x, y\rangle = 0.$$

Dies erheben wir in der allgemeinen Situation zur Definition.

Definition 14.18. Sei V ein K-Vektorraum mit einer symmetrischen Bilinearform $\langle\ ,\ \rangle$. Zwei Vektoren $v, w \in V$ heißen **orthogonal** (bezüglich $\langle\ ,\ \rangle$), wenn gilt:

$$\langle v, w\rangle = 0.$$

Als Notation vereinbaren wir $v \perp w$, wenn v und w orthogonal sind.

Bemerkung 14.19. (1) Orthogonalität ist kein Begriff des Vektorraums V alleine, sondern hängt ab von der gewählten Bilinearform $\langle\ ,\ \rangle$ auf V.

(2) Da wir $\langle\ ,\ \rangle$ als symmetrisch vorausgesetzt haben, gilt

$$v \perp w \iff w \perp v.$$

Die Relation ‚orthogonal' ist symmetrisch, aber weder reflexiv noch transitiv.

Beispiel 14.20. (1) Bezüglich des Standardskalarprodukts auf K^n sind verschiedene Standardbasisvektoren e_i und e_j für $i \neq j$ orthogonal, denn

$$e_i \bullet e_j = \delta_{ij}.$$

(2) Ein Vektor kann zu sich selbst orthogonal sein. Wir betrachten die symmetrische Bilinearform $\langle v, w\rangle_A = v^t A w$ auf K^2 bezüglich der Matrix

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

Dies ist eine symmetrische Bilinearform, weil $A = A^t$. Trotzdem ist für $i = 1, 2$

$$\langle e_i, e_i\rangle_A = 0.$$

Beide Standardbasisvektoren sind orthogonal zu sich selbst.

Definition 14.21. Eine **Orthogonalbasis** eines K-Vektorraums V mit Bilinearform $\langle\ ,\ \rangle$ ist eine Basis $\mathscr{B}$ von V, so dass für alle verschiedenen Basisvektoren $b \neq b'$ aus $\mathscr{B}$ gilt:

$$\langle b, b'\rangle = 0 = \langle b', b\rangle.$$

Wenn $\langle\ ,\ \rangle$ symmetrisch ist, bedeutet das $b \perp b'$, die Vektoren von $\mathscr{B}$ sind paarweise orthogonal.

Beispiel 14.22. Bezüglich des Standardskalarprodukts auf K^n ist die Standardbasis eine Orthogonalbasis: $e_i \bullet e_j = \delta_{ij}$.

Lemma 14.23. *Sei $\langle\ ,\ \rangle$ eine Bilinearform auf dem endlich-dimensionalen K-Vektorraum V mit Basis $\mathscr{B}$ und zugehöriger Gram'scher Matrix A. Dann sind äquivalent:*

(a) $\mathscr{B}$ ist Orthogonalbasis für V.

(b) A ist eine Diagonalmatrix.

Beweis. Dies folgt sofort aus der Definition von Orthogonalbasis und Gram'scher Matrix. □

Theorem 14.24 (Diagonalformensatz)**.** *Sei K ein Körper mit $2 \in K^\times$. Sei $f : V \times V \to K$ eine Bilinearform auf dem K-Vektorraum V mit $\dim_K(V) < \infty$. Dann sind äquivalent:*

(a) Die Bilinearform f ist symmetrisch.
(b) V hat eine Orthogonalbasis bezüglich f.
(c) Es gibt eine Basis $\mathscr{B}$ von V, so daß die zugehörige Gram'sche Matrix $\mathrm{Gram}^{\mathscr{B},\mathscr{B}}(f)$ eine Diagonalmatrix ist.

Beweis. Die Äquivalenz (b) $\iff$ (c) folgt aus Lemma 14.23. Da Diagonalmatrizen symmetrisch sind, folgt (c) $\implies$ (a) aus Proposition 14.12. Wir müssen nur noch (a) $\implies$ (b) zeigen.

Schritt 1: Wenn f die Nullform ist, also $f(v, w) = 0$ für alle $v, w \in V$, dann ist nichts zu tun: jede Basis ist Orthogonalbasis. Andernfalls muß es einen Vektor $b_1 \in V$ mit $f(b_1, b_1) \neq 0$ geben, denn ansonsten folgt aus der Polarisationsformel (Symmetrie!), Satz 14.14, für alle $v, w \in V$

$$f(v, w) = \frac{1}{2}\big(f(v + w, v + w) - f(v, v) - f(w, w)\big) = 0.$$

Schritt 2: Die Linearform $f(b_1, -) : V \to K$ hat wegen $f(b_1, b_1) \neq 0$ nichttriviales Bild. Der Unterraum $U = \ker(f(b_1, -))$ hat daher die Dimension

$$\dim(U) = \dim(V) - \dim\big(\mathrm{im}(f(b_1, -))\big) = \dim(V) - 1.$$

Weil $b_1 \notin U$ kann man eine Basis $\mathscr{C}$ von U durch b_1 zu einer Basis $\mathscr{B}$ die Basis V ergänzen ($\mathscr{B}$ ist Basis, weil die lineare Hülle ein Unterraum von V ist, der U echt enthält, und das ist aus Dimensionsgründen dann V; ferner hat $\mathscr{B}$ mit $\dim(V)$ die richtige Anzahl von Vektoren). Weil $f(b_1, u) = f(u, b_1) = 0$ für alle $u \in U$ gilt, ist $\mathscr{B}$ eine Orthogonalbasis von V genau dann, wenn $\mathscr{C}$ eine Orthogonalbasis von U ist.

Schritt 3: Jetzt argumentieren wir per Induktion nach $n = \dim(V)$. Der Induktionsanfang $\dim(V) = 0$ ist trivial.

Die Einschränkung von f auf U ist immer noch eine symmetrische Bilinearform. Per Induktionsannahme gibt es eine Orthogonalbasis von U und damit durch Ergänzen mit b_1 auch eine Orthogonalbasis von V. □

Bemerkung 14.25. Sei V wie im Theorem 14.24. Aus dem Beweis folgt, dass man jeden Vektor $b_1 \in V$ mit $\langle b_1, b_1\rangle \neq 0$ zu einer Orthogonalbasis ergänzen kann.

Korollar 14.26. *Sei K ein Körper mit $2 \in K^\times$. Sei $f : V \times V \to K$ eine symmetrische Bilinearform auf dem K-Vektorraum V mit $\dim_K(V) = n < \infty$. Dann gibt es eine Basis $\mathscr{B}$ und $\lambda_1, \ldots, \lambda_n \in K$ mit*

$$f(v, w) = \lambda_1 x_1 y_1 + \ldots + \lambda_n x_n y_n$$

für alle $v \in V$ und $w \in W$ dargestellt durch Koordinaten $x = \kappa_{\mathscr{B}}(v) = (x_1, \ldots, x_n)^t$ und $y = \kappa_{\mathscr{B}}(w) = (y_1, \ldots, y_n)^t$.

Beweis. Dazu eignet sich jede Orthogonalbasis $\mathscr{B}$, deren Existenz durch Theorem 14.24 gesichtert ist. Die Gram'sche Matrix $D = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ ist dann eine Diagonalmatrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ und die Formel ergibt sich sofort aus (14.2):

$$f(v, w) = x^t D y = \lambda_1 x_1 y_1 + \ldots + \lambda_n x_n y_n.$$ □

Sei $D \in \mathrm{M}_n(K)$ eine Diagonalmatrix und $S \in \mathrm{M}_n(K)$ beliebig. Dann ist $S^t DS$ einne symmetrische Matrix, denn

$$(S^t DS) = S^t D^t (S^t)^t = S^t DS.$$

Derr Diagonalformensatz hat als Korollar die Strukturaussage, daß jede symmetrische Matrix diese Gestalt hat, undn zwar mit invertierbarem S.

Korollar 14.27 (Normalformensatz)**.** *Sei $2 \in K^\times$. Zu einer symmetrischen Matrix $A \in \mathrm{M}_n(K)$ gibt es $S \in \mathrm{GL}_n(K)$ und eine Diagonalmatrix $D \in \mathrm{M}_n(K)$, so daß*

$$A = S^t D S.$$

Beweis. Wir betrachten die Bilinearform $\langle x, y\rangle_A = x^t A y$ auf K^n. Nach Proposition 14.12 ist $\langle\ ,\ \rangle_A$ symmetrisch, weil A symmetrisch ist. Theorem 14.24 zeigt die Existenz einer Orthogonalbasis $\mathscr{B}$ für $\langle\ ,\ \rangle_A$. Dann ist nach Lemma 14.23

$$D = \mathrm{Mat}^{\mathscr{B},\mathscr{B}}(\langle -, -\rangle_A)$$

eine Diagonalmatrix. Sei $S = \mathrm{M}^{\mathscr{E}}_{\mathscr{B}}(\mathrm{id})$ die Basiswechselmatrix aus der Standardbasis $\mathscr{E}$. Die Transformationsformel für Gram'sche Matrizen aus Proposition 14.9 liefert dann

$$A = \mathrm{Gram}^{\mathscr{E},\mathscr{E}}(\langle -, -\rangle_A) = S^t D S. \qquad \square$$

Algorithmus 14.28. Eine symmetrische Gram'sche Matrix kann leicht diagonalisiert werden, indem man das Gauß'sche Eliminationsverfahren, das die Zeilenstufenform einer Matrix liefert, auf evidente Art und Weise symmetrisch gleichzeitig durch Zeilen- und Spaltenoperationen anwendet. Es sei daran erinnert, daß die Zeilenoperation „Addiere das λ-fache der i-ten Zeile zur j-ten Zeile" durch Multiplikation von links mit der Elementarmatrix $E_{ji}(\lambda)$ erreicht wird. Analog liefert die Multiplikation von rechts mit der Elementarmatrix $E_{ij}(\lambda) = E_{ji}(\lambda)^t$ die Spaltenoperation „Addiere das λ-fache der i-ten Spalte zur j-ten Spalte".

Sei K ein Körper mit $2 \in K^\times$. Sei $A \in \mathrm{M}_n(K)$ eine symmetrische Matrix. Wie im Beweis des Diagonalformensatzes, Theorem 14.24, unterscheiden wir den Fall $A = 0$, wo nichts zu tun ist, und $A \neq 0$.

Schritt 1: Aus der Polarisationsformel (14.4) schließen wir wie im Beweis von Theorem 14.24, daß es einen Vektor b_1 gibt, für den $\langle b_1, b_1\rangle_A = \alpha \neq 0$. Wir ergänzen b_1 zu einer Basis und transformieren A entsprechend. Damit hat A die Blockform

$$A = \begin{pmatrix} \alpha & x^t \\ x & B \end{pmatrix}$$

mit $x \in K^{n-1}$ und $B \in \mathrm{M}_{n-1}(K)$. Der Eintrag x^t wird durch die Symmetrie von A bedingt und B ist ebenso deshalb symmetrisch.

Bemerkung: typischerweise ist der Schritt 1 gar nicht nötig, weil der Eintrag $e_1^t A e_1$ in Zeile 1 und Spalte 1 bereits von Null verschieden ist. Sollte dies nicht der Fall sein, versuchen wir als nächstes eine Permutation der Basisvektoren: ist ein Diagonaleintrag $e_i^t A e_i \neq 0$ von Null verschieden, dann tauschen wir e_1 und e_i, das entspricht der Vertauschung der Zeilen 1 und i und gleichzeitig der Spalten 1 und i. Sind auch alle Diagonaleinträge von 0 verschieden, dann suchen wir einen Matrixeintrag ungleich 0, etwa $e_i^t A e_j$. Wir tauschen nun 1 und i, sowie 2 und j, also oBdA $i = 1$ und $j = 2$, und ersetzen e_1, e_2 durch $e' = e_1 + e_2, e_2$ mittels $E_{12}(1)$. Dann gilt

$$e^t A e = e_1^t A e_1 + 2 e_1^t A e_2 + e_2^t A e_2 = 2 e_1^t A e_2 \neq 0,$$

denn die Diagonalterme haben wir als 0 angenommen, und $2 \in K^\times$ setzen wir voraus.

Schritt 2: Sei in Koordinaten

$$x = \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix},$$

und sei

$$S^t = E_{n1}(-x_n/\alpha) \cdot \ldots \cdot E_{21}(-x_2/\alpha).$$

Dann ist $S^t A$ das Ergebnis der Zeilenoperationen im Gauß'schen Eliminationsverfahren, und man rechnet

$$S^t A = \begin{pmatrix} \alpha & x^t \\ 0 & B - \alpha^{-1} x x^t \end{pmatrix}.$$

Man beachte, daß xx^t und damit auch $B - \alpha^{-1}xx^t$ eine symmetrische Matrix in $\mathrm{M}_{n-1}(K)$ ist.

Schritt 3: Es bleibt S^tAS zu berechnen. Per Transposition erhalten wir

$$S = E_{12}(-x_2/\alpha) \cdot \ldots \cdot E_{1n}(-x_n/\alpha),$$

somit ist S^tAS das Ergebnis von Spaltenoperationen, die analog zum Gauß'schen Eliminationsverfahren für Einträge 0 in der ersten Zeile in den Spalten $2, 3 \ldots, n$ sorgen. Es folgt

$$S^tAS = \begin{pmatrix} \alpha & 0 \\ 0 & B - \alpha^{-1}xx^t \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & A' \end{pmatrix},$$

mit der symmetrischen Matrix $A' = B - \alpha^{-1}xx^t$.

Schritt 4: Wir verfahren weiter mit A', und zwar genauer per Induktion nach der Abmessung n der Matrix. Dabei ist zu beachten, daß die Zeilen- und Spaltenoperationen, die man für die Untermatrizen ausführt (hier A' in A), durch Zeilen- und Spaltenoperationen der umgebenden Matrix (hier A) realisiert werden.

14.4. **Orthonormalbasen und orthogonale Matrizen.** Vektorräume haben lineare Struktur und man betrachtet Basen, um Koordinaten zu bekommen. Alle Basen sind gleich gut.

In Vektorräumen mit symmetrischer Bilinearform möchten wir Basen betrachten, die an die Bilinearform angepaßt sind. Wenn sie existieren, sind dies die Orthonormalbasen. Die später zu behandelnden euklidischen Vektorräume sind genau die reellen Vektorräume mit Orthonormalbasen. Alle Orthonormalbasen sind gleich gut.

Definition 14.29. Ein **Orthonormalsystem** in einem K-Vektorraum V mit symmetrischer Bilinearform $\langle\ ,\ \rangle$ ist ein Tupel von Vektoren $(v_1, \ldots, v_r)$, für die

$$\langle v_i, v_j \rangle = \delta_{ij} \text{ für alle } 1 \leq i, j \leq r.$$

Lemma 14.30. *Die Vektoren eines Orthonormalsystems $(v_1, \ldots, v_r)$ sind linear unabhängig.*

Beweis. Wenn $\sum_{i=1}^r a_i v_i = 0$ mit $a_i \in K$, dann gilt

$$a_j = \langle v_j, a_j v_j \rangle = \langle v_j, \sum_{i=1}^r a_i v_i \rangle = 0.$$

Die Linearkombination ist also trivial. Damit ist das Orthonormalsystem linear unabhängig. □

Definition 14.31. Eine **Orthonormalbasis** eines K-Vektorraums V mit symmetrischer Bilinearform $\langle\ ,\ \rangle$ ist eine Basis $\mathscr{B} = (b_1, \ldots, b_n)$, die ein Orthonormalsystem ist. Wir sagen kurz $\mathscr{B}$ ist eine **ONB**.

Bemerkung 14.32. Offensichtlich ist jede ONB eine Orthogonalbasis.

Proposition 14.33. *Sei $(V, \langle\ ,\ \rangle)$ ein K-Vektorraum mit symmetrischer Bilinearform und $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V. Wir statten K^n mit dem Standardskalarprodukt aus. Dann sind die folgenden Aussagen äquivalent.*

(a) $\mathscr{B}$ ist ONB.

(b) Die Gram'sche Matrix von $\langle\ ,\ \rangle$ bezüglich $\mathscr{B}$ ist die Einheitsmatrix.

*(c) Der Koordinatenisomorphismus $\kappa_{\mathscr{B}} : V \to K^n$ ist eine **isometrische Abbildung**: d.h. für alle $v, w \in V$ gilt*

$$\langle v, w \rangle = \kappa_{\mathscr{B}}(v) \bullet \kappa_{\mathscr{B}}(w).$$

Beweis. (a) $\iff$ (b) ist klar per Definition. Und (b) $\implies$ (c) folgt sofort aus der Formel in Proposition 14.7.

(c) $\implies$ (a): Wenn $\kappa_{\mathscr{B}}$ eine isometrische Abbildung ist, dann ist $\mathscr{B}$ eine ONB, weil

$$\langle b_i, b_j\rangle = \kappa_{\mathscr{B}}(b_i) \bullet \kappa_{\mathscr{B}}(b_j) = e_i \bullet e_j = \delta_{ij}.$$ □

Der Basiswechsel zwischen ONB bedient sich spezieller Matrizen, den orthogonalen Matrizen.

Definition 14.34. Sei K ein Körper. Eine **orthogonale Matrix** ist ein $A \in \mathrm{M}_n(K)$, so daß gilt

$$A^tA = \mathbf{1}_n.$$

Satz 14.35. *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann sind äquivalent:*

(a) A ist eine orthogonale Matrix.

(b) Die Linksmultiplikation mit A ist eine isometrischer Endomorphismus von K^n ausgestattet mit dem Standardskalarprodukt, d.h. für alle $v, w \in K^n$ gilt:

$$Av \bullet Aw = v \bullet w.$$

Beweis. Wenn A orthogonal ist, dann gilt mit Proposition 14.16

$$Av \bullet Aw = v \bullet A^tAw = v \bullet \mathbf{1}_n w = v \bullet w.$$

Umgekehrt rechnen wir den ij-Eintrag von A^tA mit Proposition 14.16 aus als

$$e_i \bullet A^tAe_j = Ae_i \bullet Ae_j = e_i \bullet e_j = \delta_{ij}.$$

Folglich ist $A^tA = \mathbf{1}_n$ und A ist orthogonal. □

Lemma 14.36. *Eine orthogonale Matrix A ist invertierbar mit $A^{-1} = A^t$.*

Beweis. Das folgt sofort aus $A^tA = \mathbf{1}_n$, und der Tatsache, daß eine quadratische Matrix invertierbar ist, sobald sie ein Linksinverses besitzt. □

Beispiel 14.37. Wir betrachten die Drehung $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ des $\mathbb{R}^2$ um den Winkel φ. Die Drehung R_φ ist bezüglich Standardkoordinaten gegeben durch die Matrixmultiplikation mit

$$D_\varphi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Die Transponierte ist

$$D_\varphi^t = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} = \begin{pmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{pmatrix} = D_{-\varphi}.$$

Die aus der Analysis bekannte Gleichung $\cos^2(\varphi) + \sin^2(\varphi) = 1$ führt zu

$$D_{-\varphi} = D_\varphi^{-1}.$$

Somit ist die Drehmatrix D_φ ist eine orthogonale Matrix.

Lemma 14.38. *Sei $A \in \mathrm{M}_n(K)$. Dann*

$$A \text{ orthogonal} \iff A^t \text{ orthogonal.}$$

Beweis. Weil $(A^t)^t = A$ gilt, reicht eine Richtung. Sei A orthogonal. Dann ist $A^t = A^{-1}$ und wegen $(A^t)^tA^t = AA^{-1} = \mathbf{1}_n$ auch A^t orthogonal. □

Proposition 14.39. *Für alle orthogonalen $A \in \mathrm{M}_n(K)$ gilt $\det(A) = 1$ oder $\det(A) = -1$.*

Beweis. Das folgt sofort aus

$$\det(A)^2 = \det(A^t) \cdot \det(A) = \det(A^t A) = \det(\mathbf{1}_n) = 1,$$

denn $\det(A)$ ist Nullstelle des Polynoms $X^2 - 1 = (X-1)(X+1)$. □

Proposition 14.40. *Sei $A \in \mathrm{M}_n(K)$. Dann sind äquivalent:*

(a) A ist orthogonal.
(b) Die Spalten von A sind eine ONB von K^n mit dem Standardskalarprodukt.
(c) Die transponierten Zeilen von A sind eine ONB von K^n mit dem Standardskalarprodukt.

Beweis. Sei $A = [s_1, \ldots, s_n]$ die Spaltenschreibweise mit den Spalten s_j von A und $A^t = [z_1^t, \ldots, z_n^t]$ entsprechend mit den Zeilen z_i von A. Dann gilt für alle i, j

$$(A^t A)_{ij} = s_i \bullet s_j \quad \text{und} \quad (AA^t)_{ij} = z_i^t \bullet z_j^t.$$

Dann ist $(s_1, \ldots, s_n)$ eine ONB genau dann, wenn $A^t A = \mathbf{1}_n$ gilt. Das wiederum ist äquivalent zu $AA^t = \mathbf{1}_n$, was genau dann gilt, wenn $(z_1^t, \ldots, z_n^t)$ eine ONB ist. □

Bemerkung 14.41. Proposition 14.40 hat die folgende Parallele für Basen und invertierbare Matrizen. Eine Matrix $A = [v_1, \ldots, v_n]$ in $\mathrm{M}_n(K)$ mit Spaltenvektoren v_j ist genau dann in $\mathrm{GL}_n(K)$, d.h. sie ist invertierbar, wenn die Spalten eine Basis $(v_1, \ldots, v_n)$ von K^n bilden.

Satz 14.42. *Sei K ein Körper. Seien $\mathscr{B}$ und $\mathscr{C}$ Basen eines K-Vektorraums V mit symmetrischer Bilinearform $\langle\ ,\ \rangle$, und sei $S = \mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\mathrm{id})$ die Basiswechselmatrix. Dann gelten alle drei der folgenden Ausssagen, sobald zwei davon eintreffen (eine „2 aus 3“ Eigenschaft).*

(a) $\mathscr{B}$ ist ONB.
(b) $\mathscr{C}$ ist ONB.
(c) S ist orthogonal.

Beweis. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ und $\mathscr{C} = (c_1, \ldots, c_n)$ und $S = (s_{ij})$. Dann ist $b_j = \sum_{k=1}^n s_{kj} c_k$. Wenn $\mathscr{C}$ ONB ist, dann folgt

$$\langle b_i, b_j \rangle = \langle \sum_{k=1}^n s_{ki} c_k, \sum_{k=1}^n s_{kj} c_k \rangle = \sum_{k=1}^n s_{ki} s_{kj} = (S^t S)_{ij},$$

somit ist S orthogonal genau dann, wenn $\mathscr{B}$ eine ONB ist.

Jetzt müssen wir noch zeigen, daß mit S orthogonal und $\mathscr{B}$ eine ONB dann auch $\mathscr{C}$ eine ONB ist. Aber das folgt aus Symmetrie, wenn wir $\mathscr{B}$ mit $\mathscr{C}$ und dadurch S mit $\mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(\mathrm{id}) = S^{-1} = S^t$ ersetzen. Das bereits bewiesene zeigt, daß dann S^t eine orthogonale Matrix ist, und damit auch S nach Lemma 14.38. □

Definition 14.43. Die **orthogonale Gruppe** (zum Standardskalarprodukt auf K^n) ist die Untergruppe von $\mathrm{GL}_n(K)$ der orthogonalen Matrizen

$$\mathrm{O}_n(K) = \{A \in \mathrm{M}_n(K)\ ;\ A^t A = \mathbf{1}_n\}.$$

Im speziellen Fall $K = \mathbb{R}$ schreibt man auch $\mathrm{O}(n) = \mathrm{O}_n(\mathbb{R})$.

Beweis. Wir müssen zeigen, daß $\mathrm{O}_n(K)$ eine Untergruppe von $\mathrm{GL}_n(K)$ ist. Ein $A \in \mathrm{O}_n(K)$ ist wegen $A^t A = \mathbf{1}_n$ invertierbar, also in $\mathrm{GL}_n(K)$. Insbesondere folgt automatisch auch $AA^t = \mathbf{1}_n$. Wenn $A, B \in \mathrm{O}_n(K)$, dann ist $AB \in \mathrm{O}_n(K)$ wegen

$$(AB)^t(AB) = B^t A^t AB = B^t (A^t A) B = B^t B = \mathbf{1}_n,$$

Für das Inverse A^{-1} gilt

$$(A^{-1})^t A^{-1} = (A^t)^{-1} A^{-1} = (AA^t)^{-1} = \mathbf{1}_n,$$

und somit $A^{-1} \in \mathrm{O}_n(K)$. Weil $\mathbf{1}_n \in \mathrm{O}_n(K)$ gilt, ist $\mathrm{O}_n(K) \neq \emptyset$. Damit ist $\mathrm{O}_n(K)$ eine Untergruppe von $\mathrm{O}_n(K) \subseteq \mathrm{GL}_n(K)$. □

15. Skalarprodukte

15.1. **Positiv definite symmetrische Bilinearformen.** Die zentrale Definition diese Kapitels ist die folgende.

Definition 15.1. Eine symmetrische Bilinearform $\langle\ ,\ \rangle$ auf einem $\mathbb{R}$-Vektorraum V heißt

(1) **positiv definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v\rangle > 0$,
(2) **positiv semi-definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v\rangle \geq 0$,
(3) **negativ definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v\rangle < 0$,
(4) **negativ semi-definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v\rangle \leq 0$,
(5) **indefinit**, wenn keiner der Fälle (1)-(4) eintritt.

Eine symmetrische reelle Matrix $A \in \mathrm{M}_n(\mathbb{R})$ heißt **positiv/negativ (semi-)definit** oder **indefinit**, wenn die zugehörige Bilinearform $\langle\ ,\ \rangle_A$ auf $\mathbb{R}^n$ entsprechend positiv/negativ (semi-)definit oder indefinit ist.

Wir nennen $\langle\ ,\ \rangle$ ein **Skalarprodukt** (oder **inneres Produkt**), wenn $\langle\ ,\ \rangle$ eine symmetrische und positiv definite Bilinearform ist.

Beispiel 15.2. Hier sind die typischen Beispiele.

(1) Das Standardskalarprodukt auf $\mathbb{R}^n$ ist positiv definit: für alle $x = (x_1, \ldots, x_n)^t \neq 0$ gilt

$$x \bullet x = \sum_{i=1}^{n} x_i^2 > 0$$

(2) Die Matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist indefinit: $\langle e_1, e_1\rangle_A = 1$ während $\langle e_2, e_2\rangle_A = -1$.

(3) Für $K \neq \mathbb{R}$ ist das Standardskalarprodukt auf K^n kein Skalarprodukt. Um von *positiv definit* sprechen zu können, muß der Körper eine Anordnung „$>$“ haben, also für uns der Körper $\mathbb{R}$ sein.

Bemerkung 15.3. Am einfachsten erkennt man, ob eine symmetrische Bilinearform $\langle\ ,\ \rangle$ (oder eine symmetrische Matrix A) positiv (oder negativ) (semi-)definit ist, indem man zu einer Diagonalform übergeht, also zu einer Gram'schen Matrix $D = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ einer Orthogonalbasis $\mathscr{B}$. Dann ist $\langle\ ,\ \rangle$ (bzw. A)

(1) positiv definit, wenn $\alpha_i > 0$ für alle $i = 1, \ldots, n$,
(2) positiv semi-definit, wenn $\alpha_i \geq 0$ für alle $i = 1, \ldots, n$,
(3) negativ definit, wenn $\alpha_i < 0$ für alle $i = 1, \ldots, n$,
(4) negativ semi-definit, wenn $\alpha_i \leq 0$ für alle $i = 1, \ldots, n$,
(5) indefinit, wenn unter den α_i für $i = 1, \ldots, n$ beide Vorzeichen vorkommen.

Dieses Kriterium folgt sofort, wenn man sich die Formel

$$\langle v, v\rangle = \alpha_1(x_1)^2 + \ldots + \alpha_n(x_n)^2 \quad \text{für } \kappa_{\mathscr{B}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

anschaut. Die Behauptung folgt, weil $x^2 \geq 0$ für alle $x \in \mathbb{R}$ gilt, und weil $x^2 = 0$ genau dann, wenn $x = 0$.

Satz 15.4. *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform $\langle\ ,\ \rangle$. Dann gilt*

$$V \text{ hat ONB} \iff \langle\ ,\ \rangle \text{ ist positiv definit.}$$

Beweis. Sei $\mathscr{B}$ eine ONB. Weil das Standardskalarprodukt auf $\mathbb{R}^n$ positiv definit ist, gilt für alle $v \in V$ nach Proposition 14.33 (c)

$$\langle v, v\rangle = \kappa_{\mathscr{B}}(v) \bullet \kappa_{\mathscr{B}}(v) \geq 0$$

mit Gleichheit genau für $\kappa_{\mathscr{B}}(v) = 0$, also $v = 0$. Damit ist $\langle\ ,\ \rangle$ auf V positiv definit.

Sei umgekehrt $\langle\ ,\ \rangle$ positiv definit. Weil $2 \in \mathbb{R}^\times$ gibt es nach dem Diagonalformensatz, Theorem 14.24 eine Orthogonalbasis $\mathscr{B}' = (b'_1, \ldots, b'_n)$. Weil $\langle\ ,\ \rangle$ positiv definit ist und man in $\mathbb{R}$ Wurzeln aus positiven reellen Zahlen ziehen kann, gibt für es $i = 1, \ldots, n$ Zahlen $\lambda_i \in \mathbb{R}$, $\lambda_i \neq 0$ mit

$$\langle b'_i, b'_i\rangle = \lambda_i^2.$$

Dann sind $b_i := \lambda_i^{-1} b'_i$ die Vektoren einer ONB. □

Definition 15.5. Ein **euklidischer Vektorraum** ist ein endlichdimensionaler $\mathbb{R}$-Vektorraum zusammen mit einem Skalarprodukt.

Korollar 15.6. *Ein euklidischer Vektorraum V der Dimension n ist isometrisch zu $\mathbb{R}^n$ mit dem Standardskalarprodukt.*

Beweis. Satz 15.4 liefert eine ONB $\mathscr{B}$ auf V. Die zugehörige Koordinatenabbildung $\kappa_{\mathscr{B}} : V \to \mathbb{R}^n$ ist eine isometrische Abbildung nach Proposition 14.33 und ein Isomorphismus. □

15.2. **Die euklidische Norm.** In Definition 14.17 haben wir die Länge eines Vektors $x = \binom{x_1}{x_2} \in \mathbb{R}^2$ definiert als

$$\|x\| = \sqrt{\langle x, x\rangle} = \Big(\sum_{i=1}^{2} x_i^2\Big)^{1/2} = \sqrt{(x_1)^2 + (x_2)^2}.$$

Dies übertragen wir nun auf den $\mathbb{A}^3(\mathbb{R}) = \mathbb{R}^3$ und die Punkte

$$O = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\ P = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix},\ Q = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix},\ R = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Das Quadrat des Abstands von O und Q berechnen wir in der Ebene gegeben durch die Gleichung $z = 0$ zu

$$a^2 + b^2.$$

Da die Punkte O, Q und R ebenfalls ein rechtwinkliges Dreieck bilden, jetzt in der Ebene gegeben durch die Gleichung $bx - ay = 0$, folgt für das Quadrat des Abstands von R zu O

$$(a^2 + b^2) + c^2.$$

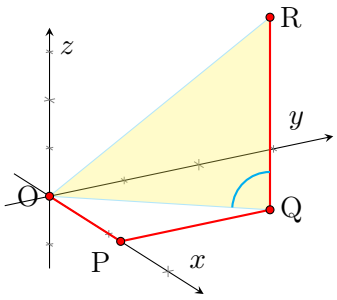


ABBILDUNG 9. Abstand R zu O im $\mathbb{A}^3(\mathbb{R})$.

Diese elementargeometrischen Überlegungen, die an unsere Anschauung appellieren und daher nicht der axiomatischen Methode entsprechen, motivieren die folgenden Definitionen.

Definition 15.7. Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum.

(1) Die **Norm** (oder **Länge**) eines Vektors $v \in V$ ist die reelle Zahl
$$\|v\| = \sqrt{\langle v, v\rangle}.$$
Hier bezeichne $\sqrt{x}$ die nicht-negative Wurzel in $\mathbb{R}$ von $x \geq 0$.

(2) Ein **normierter** Vektor (oder auch **Einheitsvektor**) in V ist ein Vektor $v \in V$, so daß
$$\|v\| = 1.$$

Bemerkung 15.8. (1) Für den $\mathbb{R}^n$ mit dem Standardskalarprodukt ist mit Koordinaten $x^t = (x_1, \ldots, x_n)$ die Norm durch die folgende Formel gegeben:
$$\|x\| = \sqrt{x \bullet x} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

(2) Es ist wichtig, daß wir mit einem Skalarprodukt $\langle\ ,\ \rangle$ arbeiten, weil nur dann $\langle v, v\rangle \geq 0$ ist und eine Wurzel in $\mathbb{R}$ existiert.

Proposition 15.9. *Sei V ein euklidischer Vektorraum.*

(1) Für alle $v \in V$ gilt $\|v\| \geq 0$ und
$$\|v\| = 0 \iff v = 0.$$

(2) Für alle $v \in V$ und $\lambda \in \mathbb{R}$ gilt
$$\|\lambda v\| = |\lambda| \cdot \|v\|.$$

Beweis. (1) Es ist $\langle v, v\rangle \geq 0$, weil $\langle\ ,\ \rangle$ ein Skalarprodukt ist. Daher gibt es eine eindeutige nicht-negative Wurzel. Positive Definitheit zeigt
$$v = 0 \iff \langle v, v\rangle = 0 \iff \|v\| = 0.$$

(2) Wir ziehen die (nicht-negative) Wurzel aus
$$\|\lambda v\|^2 = \langle \lambda v, \lambda v\rangle = \lambda^2 \langle v, v\rangle = \lambda^2 \|v\|^2. \qquad \square$$

Beispiel 15.10. Jedes $v \in V$, $v \neq 0$ hat genau zwei normierte Vielfache. Die Gleichung $\|\lambda v\| = 1$ ist äquivalent zu $|\lambda| \cdot \|v\| = 1$. Gesucht sind die normierten Vektoren
$$\pm \|v\|^{-1} \cdot v.$$
Dieser Trick kam schon im Beweis von Satz 15.4 vor.

Satz 15.11 (Cauchy–Schwarz Ungleichung)**.** *Für alle Vektoren v, w in einem euklidischen Vektorraum V gilt*

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|.$$

Gleichheit gilt genau dann, wenn v und w linear abhängig sind.

Beweis. Wenn v und w linear abhängig sind, also OBdA $w = \lambda v$ mit $\lambda \in K$, dann gilt

$$|\langle v, w\rangle| = |\langle v, \lambda v\rangle| = |\lambda| \cdot \|v\|^2 = \|v\| \cdot \|\lambda v\| = \|v\| \cdot \|w\|.$$

Seien nun v und w linear unabhängig. Dann ist für alle $t \in \mathbb{R}$ der Vektor $v + tw \neq 0$ und daher

$$0 < \|v + tw\|^2 = \langle v + tw, v + tw\rangle = \|v\|^2 + 2t\langle v, w\rangle + t^2\|w\|^2.$$

Als Funktion von t ist die rechte Seite quadratisch und nimmt sein Minimum (Ableiten, Nullsetzen, ...) bei

$$t_0 = -\langle v, w\rangle / \|w\|^2$$

an. Der Wert bei t_0 ist

$$0 < \|v\|^2 - (\langle v, w\rangle)^2 / \|w\|^2,$$

und daraus folgt die Cauchy-Schwarz Ungleichung durch Umstellen und Wurzelziehen. □

Korollar 15.12. *Seien $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$ reelle Zahlen. Dann gilt*

$$\Big(\sum_{i=1}^{n} x_i y_i\Big)^2 \leq \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2.$$

Beweis. Das ist das Quadrat der Cauchy–Schwarz Ungleichung für $x, y \in \mathbb{R}^n$ bezüglich des Standardskalarprodukts und der üblichen Notation für die Koordinaten. Man darf quadrieren, weil beide Seiten der Cauchy–Schwarz'schen Ungleichung nicht-negativ sind. □

Satz 15.13 (Dreiecksungleichung)**.** *Sei V ein euklidischer Vektorraum. Für alle $x, y \in V$ gilt*

$$\|x + y\| \leq \|x\| + \|y\|.$$

Beweis. Da beide Seiten nicht-negativ sind, dürfen wir die Ungleichung quadrieren. Wir haben zu zeigen:

$$\langle x + y, x + y\rangle \leq \langle x, x\rangle + 2\|x\|\|y\| + \langle y, y\rangle.$$

Dies ist nach der Polarisationsformel aus Satz 14.14 äquivalent zu

$$2\langle x, y\rangle \leq 2\|x\|\|y\|,$$

was aus der Cauchy–Schwarz Ungleichung folgt. □

Definition 15.14. (1) Der Abstand zweier Vektoren $v, w \in V$ ist

$$d(v, w) = \|v - w\|.$$

(2) Sei A ein affiner Raum mit Translationen durch V. Dann ist der Abstand der Punkte $P, Q \in A$ definiert als

$$d(P, Q) = \|v\|$$

wenn $v \in V$ der eindeutige Vektor ist mit $v + Q = P$.

Bemerkung 15.15. (1) Daß wir die Norm $\|v\|$ als Länge ansprechen, liegt an der Formel im affinen Raum $\mathbb{A}(V)$:

$$d(v, 0) = \|v\|.$$

(2) Für den affinen Raum $A = \mathbb{A}(V) = V$ stimmen beide Definitionen des Abstands überein, denn
$$(v - w) + w = v.$$
(3) Im $\mathbb{A}^2(\mathbb{R})$ und im $\mathbb{A}^3(\mathbb{R})$ paßt unsere Definition der Länge mit der aus der Anschauung über den Satz des Pythagoras hergeleiteten Länge zusammen.

Satz 15.16 (Eigenschaften des Abstands)**.** *Sei A ein affiner Raum mit Vektorraum V der Translationen. Der Abstand hat die folgenden Eigenschaften:*

(1) ***Positiv definit****: für alle $P, Q \in A$ gilt $d(P, Q) \geq 0$, und*
$$d(P, Q) = 0 \iff P = Q.$$
(2) ***Symmetrie****: für alle $P, Q \in A$ ist*
$$d(P, Q) = d(Q, P).$$
(3) ***Dreiecksungleichung****: für alle P, Q und $R \in A$ gilt*
$$d(P, Q) \leq d(P, R) + d(R, Q).$$

Beweis. (1) Sei $v \in V$ der eindeutige Vektor mit $v + Q = P$. Dann ist $d(P, Q) = \|v\| \geq 0$. Und genauer
$$d(P, Q) = 0 \iff \|v\| = 0 \iff v = 0 \iff P = Q.$$
(2) Weiter ist dann $-v + P = -v + (v + Q) = (-v + v) + Q = Q$ und daher
$$d(P, Q) = \|v\| = |-1|\|-v\| = \|-v\| = d(Q, P).$$
(3) Seien nun $x, y \in V$ die eindeutigen Vektoren mit $x + R = P$ und $y + Q = R$. Dann ist $(x + y) + Q = x + (y + Q) = x + R = P$. Damit folgt die Dreiecksungleichung für den Abstand aus der Dreiecksungleichung für die Norm aus Satz 15.13:
$$d(P, Q) = \|x + y\| \leq \|x\| + \|y\| = d(P, R) + d(R, Q). \qquad \square$$

Bemerkung 15.17. Eine Funktion $d : X \times X \to \mathbb{R}$ auf einer Menge X mit den Eigenschaften von Satz 15.16 nennt man eine Metrik auf X. Das ist ein wichtiger Begriff für die Analysis.

15.3. **Gram-Schmidt.** Ein Skalarprodukt auf V erlaubt für einen Unterraum $U \subseteq V$ eine ausgezeichnete Projektion auf den Unterraum U: die orthogonale Projektion. In der folgenden Formel (15.1) benötigen wir dazu zunächst eine ONB für den Unterraum U. Proposition 15.18 funktioniert als Lemma für den eigentlichen Satz 15.20, das Gram-Schmidt'sche Orthonormalisierungsverfahren. Im Laufe des Verfahrens wird die benötigte Basis zur Verfügung gestellt.

Proposition 15.18. *Sei K en Körper. Sei V ein K-Vektorraum und $\langle\ ,\ \rangle$ eine symmetrische Bilinearform auf V. Sei $U \subseteq V$ ein Unterraum, und sei $(b_1, \ldots, b_r)$ eine Orthonormalbasis von U bezüglich der auf U eingeschränkten symmetrischen Bilinearform $\langle\ ,\ \rangle$. Die* ***orthogonale Projektion*** *$p_U : V \to V$ auf U ist gegeben durch*
$$p_U(v) = \sum_{i=1}^{r} \langle v, b_i \rangle b_i \qquad \text{für alle } v \in V, \tag{15.1}$$
und hat die folgenden Eigenschaften.

(1) $V = U \oplus \ker(p_U)$, und für $v \in V$ mit $v = u + x$ mit $u \in U$ und $x \in \ker(p_U)$ ist $p_U(v) = u$.
(2) Für alle $v \in V$ und $u \in U$ gilt
$$v - p_U(v) \perp u.$$

Beweis. (1) Es gilt für $j = 1, \dots, r$

$$p_U(b_j) = \sum_{i=1}^{r} \langle b_j, b_i \rangle b_i = \sum_{i=1}^{r} \delta_{ij} b_i = b_j.$$

Daraus folgt für alle $v \in V$

$$p_U(p_U(v)) = p_U\big((\sum_{i=1}^{r} \langle v, b_i \rangle b_i\big) = \sum_{i=1}^{r} \langle v, b_i \rangle p_U(b_i) = p_U(v).$$

Also ist $p_U \circ p_U = p_U$ und p_U ein Projektor, siehe Kapitel 7.3. Es bleibt das Bild vonn p_U zu bestimmen. Offenbar ist $p_U(v) \in U$ für alle $v \in V$. Weil die Basis $(b_1, \dots, b_r)$ im Bild liegt, folgt $\mathrm{im}(p_U) = U$. Damit ist p_U ein Projektor auf U.

(2) Für $v \in V$ ist mit der Notation aus (1) $x = v - p_U(v) \in \ker(p_U)$. Das bedeutet

$$0 = p_U(x) = \sum_{i=1}^{r} \langle x, b_i \rangle b_i.$$

Weil $(b_1, \dots, b_r)$ eine Basis von U ist, folgt $\langle x, b_i \rangle = 0$ für alle $i = 1, \dots, r$. Jedes $u \in U$ können wir schreiben als Linearkombination $\sum_{i=1}^{r} y_i b_i$, und damit gilt

$$\langle v - p_U(v), u \rangle = \langle x, \sum_{i=1}^{r} y_i b_i \rangle = \sum_{i=1}^{r} y_i \langle x, b_i \rangle = 0. \qquad \square$$

Beispiel 15.19. Sei V ein $\mathbb{R}$-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Zu Vektoren $b, c \in V$ ist ein $\lambda \in \mathbb{R}$ zu bestimmen, so daß $c + \lambda b$ orthogonal zu b wird.

Die Anforderung ist nichts anderes als eine lineare Gleichung für die Unbekannte λ:

$$0 = \langle c + \lambda b, b \rangle = \langle c, b \rangle + \lambda \langle b, b \rangle.$$

Wenn $b \neq 0$, dann ist $\langle b, b \rangle > 0$ und es gibt eine eindeutige Lösung

$$\lambda = -\frac{\langle c, b \rangle}{\langle b, b \rangle}.$$

Im Gram-Schmidt Orthonormalisierungsverfahren passiert genau das.

Der Diagonalformensatz, Theorem 14.24, hat für Skalarprodukte einen konstruktiven Beweis. Der Beweis beinhaltet einen Algorithmus, der aus einer beliebigen Basis eine Orthogonalbasis macht, und das ist das Gram–Schmidt'sche Orthonormalisierungsverfahren.

Satz 15.20 (Gram–Schmidt'sches Orthonormalisierungsverfahren)**.** *Sei V ein $\mathbb{R}$-Vektorraum mit Basis $\mathscr{C} = (c_1, \dots, c_n)$, und sei $\langle\ ,\ \rangle$ ein Skalarprodukt auf V. Die Vektoren $b_1, \dots, b_n$ seien rekursiv für $i = 1, \dots, n$ definiert durch*

$$b_i' = c_i - \sum_{j=1}^{i-1} \langle c_i, b_j \rangle b_j,$$

$$b_i = \frac{1}{\|b_i'\|} b_i'.$$

Dann ist $\mathscr{B} = (b_1, \dots, b_n)$ eine Orthonormalbasis von V bezüglich $\langle\ ,\ \rangle$.

Beweis. Wir setzen $U_i = \langle c_1, \dots, c_i \rangle_{\mathbb{R}}$ und erhalten eine aufsteigende Folge von Unterräumen

$$0 = U_0 \subseteq U_1 \subseteq U_2 \subseteq \dots \subseteq U_n = V$$

mit $\dim(U_i) = i$ für alle $i = 1, \dots, n$. Wir zeigen per Induktion nach r, daß $(b_1, \dots, b_r)$ eine ONB von U_r ist. Der Induktionsanfang mit $r = 0$ ist trivial.

Sei die Aussage bis r bewiesen. Dann ist $(b_1, \dots, b_r)$ eine ONB für U_r. Nach Proposition 15.18 ist die orthogonale Projektion p_{U_r} auf U_r derart, daß

$$b'_{r+1} = c_{r+1} - p_{U_r}(c_{r+1}) \in \ker(p_{U_r}).$$

Insbesondere ist $\langle b_i, b'_{r+1}\rangle = 0$ für alle $i = 1, \dots, r$, denn diese b_i sind in U_r.

Weil b'_{r+1} und c_{r+1} sich nur um das Element $p_{U_r}(c_{r+1}) \in U_r$ unterscheiden, ist

$$U_{r+1} = U_r + \mathbb{R} \cdot c_r = U_r + \mathbb{R} \cdot b'_{r+1} = \langle b_1, \dots, b_r, b'_{r+1}\rangle_{\mathbb{R}}.$$

Damit ist $(b_1, \dots, b_r, b'_{r+1})$ eine Basis von U_{r+1} nach Satz 5.31. Damit ist $b'_{r+1} \neq 0$ und b_{r+1} ist wohldefiniert und wie in Beispiel 15.10 normiert. Außerdem folgt aus $b'_{r+1} \perp b_i$ für $i = 1, \dots, r$ auch

$$\langle b_i, b_{r+1}\rangle = \langle b_i, \frac{1}{\|b'_{r+1}\|} b'_{r+1}\rangle = \frac{1}{\|b'_{r+1}\|}\langle b_i, b'_{r+1}\rangle = 0.$$

Das zeigt, daß $(b_1, \dots, b_{r+1})$ eine ONB von U_{r+1} ist, denn $(b_1, \dots, b_r)$ ist ja bereits eine ONB von U_r. Dies beendet den Beweis des Induktionsschritts und damit den Beweis des Satzes. □

Das Gram–Schmidt'sche Orthonormalisierungsverfahren zeigt in Wirklichkeit mehr.

Satz 15.21. *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und $\langle\ ,\ \rangle$ eine positiv definite symmetrische Bilinearform auf V.*

(1) Es gibt Orthonormalbasen für V.

(2) (orthonormaler Basisergänzungssatz) Jedes Orthonormalsystem in V läßt sich zu einer ONB ergänzen.

Beweis. (1) Der Vektorraum V hat eine Basis. Der Algorithmus aus Satz 15.20 transformiert diese in eine ONB, die es damit auch gibt.

(2) Sei $(c_1, \dots, c_r)$ eine Orthonormalsystem. Dann sind diese Vektoren nach Lemma 14.30 linear unabhängig und können deshalb zu einer Basis $\mathscr{C} = (c_1, \dots, c_r, c_{r+1}, \dots, c_n)$ ergänzt werden. Nun wenden wir auf $\mathscr{C}$ den Algorithmus aus Satz 15.20 an. Da $(c_1, \dots, c_r)$ bereits orthonormal sind, wird dabei (per Induktion zu zeigen) $b_i = c_i$ für $1 \leq i \leq r$, d.h. es passiert zunächst nichts. Der Algorithmus spuckt also eine ONB aus, die das gegebene Orthonormalsystem fortsetzt. □

Proposition 15.22. *Sei $P \in \mathrm{M}_n(\mathbb{R})$ eine quadratische Matrix. Dann sind äquivalent:*

(a) P ist symmetrisch und positiv definit (bzw. semi-definit).

(b) Es gibt eine Matrix $A \in \mathrm{GL}_n(\mathbb{R})$ (bzw. $A \in \mathrm{M}_n(\mathbb{R})$) mit

$$P = A^t A.$$

Beweis. (b) $\Longrightarrow$ (a): Die Matrix $P = A^t A$ ist wegen

$$P^t = (A^t A)^t = A^t (A^t)^t = A^t A = P$$

offensichtlich symmetrisch. Weiter ist für alle $v \in \mathbb{R}^n$

$$\langle v, v\rangle_P = v^t A^t A v = (Av)^t (Av) = Av \bullet Av \geq 0.$$

Wenn A sogar invertierbar ist, dann folgt für $v \neq 0$ sogar $\langle v, v\rangle_P > 0$.

(a) $\Longrightarrow$ (b): Sei umgekehrt P positiv definit. Die symmetrische Bilinearform $\langle\ ,\ \rangle_P$ besitzt nach Satz 15.21 eine ONB $\mathscr{B}$. Sei $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{E}}(\mathrm{id}) \in \mathrm{GL}_n(\mathbb{R})$ die Basiswechselmatrix von der Standardbasis $\mathscr{E}$ zu $\mathscr{B}$. Dann ist

$$P = \mathrm{Gram}^{\mathscr{E},\mathscr{E}}(\langle\ ,\ \rangle_P) = A^t\, \mathrm{Gram}^{\mathscr{B},\mathscr{B}}(\langle\ ,\ \rangle_P) A = A^t \mathbf{1}_n A = A^t A.$$

Ist P nur positiv semi-definit, dann wählen wir eine Orthogonalbasis $\mathscr{B} = (b_1, \ldots, b_n)$ mit $\langle b_i, b_i\rangle = 1$ für $1 \le i \le r$ und $\langle b_i, b_i\rangle = 0$ für $r+1 \le i \le n$. Es folgt, mit der Matrix in Blockform der Größe $(r, n-r)$

$$E_{r,n-r} := \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} = E_{r,n-r}^t E_{r,n-r}$$

wie im positiv definiten Fall nur

$$P = A^t E_{r,n-r} A = A^t E_{r,n-r}^t E_{r,n-r} A = B^t B$$

mit $B = E_{r,n-r}A \in \mathrm{M}_n(\mathbb{R})$. □

Bemerkung 15.23. Der Basiswechsel $S = \mathrm{M}_{\mathscr{B}}^{\mathscr{C}}(\mathrm{id})$ von der Basis $\mathscr{C}$ nach $\mathscr{B}$ aus Satz 15.20 ist eine obere Dreiecksmatrix. In der Tat gilt

$$c_j = \|b_j'\| \cdot b_j + \sum_{i=1}^{j-1} \langle c_j, b_i\rangle \cdot b_i,$$

so daß $S = (s_{ij})_{1 \le i,j \le n}$ mit

$$s_{ij} = \begin{cases} \langle c_j, b_i\rangle & i < j, \\ \|b_j'\| & i = j, \\ 0 & i > j. \end{cases}$$

Mit S ist auch S^{-1}, die Basiswechselmatrix $\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(\mathrm{id})$, eine obere Dreiecksmatrix, denn die Menge der oberen Dreiecksmatrizen ist eine Untergruppe von $\mathrm{GL}_n(K)$.

Bemerkung 15.24. Die Matrix A in Proposition 15.22 kann als obere Dreiecksmatrix gewählt werden. Dazu wählt man wie im Beweis die ONB $\mathscr{B}$ nach dem Gram-Schmidt'schen Orthonormalisierungsverfahren angewandt auf die Standardbasis. Der Basiswechsel hat dann als Basiswechselmatrix $R = A$ eine obere Dreiecksmatrix, vgl. Bemerkung 15.23. Die Zerlegung

$$P = R^t R$$

heißt **Cholesky-Zerlegung** und ist eindeutig, wenn man noch fordert, daß die Diagonaleinträge positiv sind. Die Zerlegung wird auch mit einer unteren Dreiecksmatrix L als $P = LL^t$ geschrieben. Das ist mit $L = R^t$ offensichtlich äquivalent.

16. Spektraltheorie reller symmetrischer Matrizen

16.1. **Das charakteristische Polynom reeller symmetrischer Matrizen.** Ein komplexes Polynom $P(X) \in \mathbb{C}[X]$ positiven Grades, $\deg(P) \ge 1$, hat nach dem Fundamentalsatz der Algebra, Theorem F.1, stets eine Nullstelle: es gibt $z_0 \in \mathbb{C}$ mit $P(z_0) = 0$. Daraus folgern wir in Satz F.2, daß $P(X)$ sogar vollständig in Linearfaktoren zerfällt. Dies gilt dann auch für reelle Polynome, allerdings typischerweise nicht in $\mathbb{R}[X]$, also mit reellen Linearfaktoren $X - a$ und $a \in \mathbb{R}$, sondern mit komplexen Linearfaktoren. Vor diesem Hintergrund betrachtet ist der folgende Satz bemerkenswert.

Satz 16.1. *Das charakteristische Polynom einer reellen symmetrischen Matrix zerfällt in $\mathbb{R}[X]$ vollständig in ein Produkt von Linearfaktoren.*

Korollar 16.2. *Sei $A \in \mathrm{M}_n(\mathbb{R})$ ein reelle symmetrische Matrix. Aufgefaßt als komplexe Matrix $A \in \mathrm{M}_n(\mathbb{C})$ hat A nur reelle Eigenwerte.*

Beweis. Das folgt sofort aus Satz 16.1 und dem Zusammenhang zwischen Eigenwerten und Nullstellen des charakteristischen Polynoms, beziehungsweise dessen Linearfaktoren. □

Wir führen zwei Beweise für Satz 16.1. Der erste elegante Beweis zeigt zuerst das Korollar 16.2 und benutzt dann den Fundamentalsatz der Algebra, Theorem F.1. Der zweite Beweis benutzt ein wenig Analysis.

Erster Beweis von Satz 16.1. Sei $A \in \mathrm{M}_n(\mathbb{R})$ eine symmetrische Matrix. Diese können wir als komplexe Matrix auffassen. Das charakteristische Polynom ändert sich dadurch nicht, was sofort aus der Leibniz-Formel für die Determinante folgt. Als komplexe Matrix zerfällt das charakteristische Polynom $\chi_A(X) \in \mathbb{C}[X]$ im Wesentlichen nach dem Fundamentalsatz der Algebra, genauer Satz F.2, in ein Produkt aus Linearfaktoren

$$\chi_A(X) = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n).$$

Wir müssen daher nur zeigen, daß jede komplexe Nullstelle $\lambda \in \mathbb{C}$ von $\chi_A(X)$ in Wahrheit bereits reell ist. Sei $\lambda \in \mathbb{C}$ mit $\chi_A(\lambda) = 0$. Dann gibt es einen entsprechenden komplexen Eigenvektor $z \in \mathbb{C}^n$, $z \neq 0$, mit

$$Az = \lambda z.$$

Unter $\bar{z}$ verstehen wir den koordinatenweise komplexkonjugierten Vektor. Dann gilt $\overline{Az} = A\bar{z}$, weil A reelle Einträge hat. Weiter gilt $\overline{\lambda z} = \bar{\lambda}\bar{z}$. Daraus folgt

$$\bar{\lambda}(\bar{z} \bullet z) = (\bar{\lambda}\bar{z}^t)z = (\bar{\lambda}\bar{z})^t z = \overline{\lambda z}^t z = \overline{Az}^t z = (A\bar{z})^t z = \bar{z}^t A^t z = \bar{z}^t A z = \bar{z}^t(\lambda z) = \lambda(\bar{z} \bullet z).$$

Seien $z_1, \ldots, z_n$ die Koordinaten von z. Dann ist wegen $z \neq 0$

$$\bar{z} \bullet z = \sum_{i=1}^n \bar{z}_i z_i = \sum_{i=1}^n |z_i|^2 > 0,$$

also insbesondere ist $\bar{z} \bullet z \neq 0$ als komplexe Zahl. Nach Division durch $\bar{z} \bullet z$ folgt $\bar{\lambda} = \lambda$ und damit

$$\lambda \in \mathbb{R}.$$

Dies zeigt die Behauptung. □

Bemerkung 16.3. Die komplexen Zahlen haben geholfen, den rein reellen Satz 16.1 zu beweisen. Es ist nicht unüblich, daß man „outside the box“ denken muß.

Zweiter Beweis von Satz 16.1. Sei $A \in \mathrm{M}_n(\mathbb{R})$ symmetrisch. Die $(n-1)$-Sphäre

$$\mathrm{S}^{n-1} = \{v \in \mathbb{R}^n \ ; \ \|v\| = 1\}$$

ist abgeschlossen und beschränkt, also kompakt (Analysis), und daher nimmt die stetige Funktion

$$f : \mathrm{S}^{n-1} \to \mathbb{R}, \qquad f(v) = v \bullet Av \text{ für alle } v \in \mathrm{S}^{n-1}$$

auf S^{n-1} ihr Supremum als Maximum an. Das ist ein Fakt aus der Analysis.

Sei also $v_0 \in \mathrm{S}^{n-1}$ normiert und für alle $w \in \mathrm{S}^{n-1}$

$$v_0 \bullet Av_0 \geq w \bullet Aw.$$

Wir behaupten, daß v_0 ein Eigenvektor ist zum Eigenwert

$$\lambda_0 = v_0 \bullet Av_0.$$

Dazu nutzen wir aus, daß für alle $w \in \mathbb{R}^n$ die Funktion

$$F_w(t) = \frac{(v_0 + tw) \bullet A(v_0 + tw)}{\|v_0 + tw\|^2} = f\Big(\frac{1}{\|v_0 + tw\|}(v_0 + tw)\Big)$$

in einer Umgebung von 0 stetig differenzierbar ist und in $t = 0$ ein Maximum hat. Die Ableitung bei $t = 0$ ist demnach mit Proposition 14.16

$$\begin{aligned} 0 = F_w'(0) &= \frac{d}{dt}\Big|_{t=0} \left(\frac{v_0 \bullet Av_0 + t\big(w \bullet Av_0 + v_0 \bullet Aw\big) + t^2 w \bullet Aw}{1 + t\big(w \bullet v_0 + v_0 \bullet w\big) + t^2 w \bullet w} \right) \\ &= w \bullet Av_0 + v_0 \bullet Aw - v_0 \bullet Av_0 \cdot \big(w \bullet v_0 + v_0 \bullet w\big) \\ &= 2Av_0 \bullet w - \lambda_0 \cdot \big(2v_0 \bullet w\big) \qquad (A = A^t \text{ und } - \bullet - \text{ symmetrisch}) \\ &= 2 \cdot \big(Av_0 \bullet w - \lambda_0 \cdot v_0 \bullet w\big) = 2(Av_0 - \lambda_0 v_0) \bullet w. \end{aligned}$$

Es gilt also für alle $w \in \mathbb{R}^n$

$$(Av_0 - \lambda_0 v_0) \bullet w = 0.$$

Mit der Wahl von $w = Av_0 - \lambda_0 v_0$ folgt $\|Av_0 - \lambda_0 v_0\|^2 = 0$. Daraus folgt

$$Av_0 - \lambda_0 v_0 = 0$$

und somit die Behauptung, daß v_0 Eigenvektor zum Eigenwert λ_0 ist. Damit ist gezeigt, daß jede reelle symmetrische Matrix einen reellen Eigenwert hat.

Wir zeigen nun Satz 16.1 per Induktion nach der Dimension n. Ohne Einschränkung nehmenn wir v_0 als normiert an: $\|v_0\| = 1$. Wir ergänzen $b_1 = v_0$ mittels Satz 15.21 zu einer ONB $\mathscr{B} = (b_1, \ldots, b_n)$ und betrachten die Basiswechselmatrix $S = S^{\mathscr{B}}_{\mathscr{E}} = [b_1, \ldots, b_n] \in \mathrm{O}_n(\mathbb{R})$, deren Spalten genau aus dieser ONB $\mathscr{B}$ bestehen, die erste aus v_0. Als Basiswechsel zwischen ONBen ist S orthogonal. Somit gilt $S^{-1} = S^t$. Dann erhalten wir wegen $Av_0 = \lambda_0 v_0$ eine Blockform

$$A' = S^{-1}AS = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(L_A) = \begin{pmatrix} \lambda_0 & \beta_2 & \cdots & \beta_n \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

mit $B \in \mathrm{M}_{n-1}(\mathbb{R})$ und $\beta_2, \ldots, \beta_n \in \mathbb{R}$. Es ist A' symmetrisch:

$$A'^t = (S^{-1}AS)^t = S^t A^t (S^{-1})^t = S^{-1}AS = A'.$$

Daher gilt $B^t = B$ und $\beta_2 = \ldots = \beta_n = 0$. Damit ist A' eine Blockdiagonalmatrix und nach Proposition 12.17 und Korollar 13.7 folgt

$$\chi_A(X) = \chi_{A'}(X) = (X - \lambda_0) \cdot \chi_B(X).$$

Nach Induktionsvoraussetzung hat das symmetrische B ein vollständig in reelle Linearfaktoren zerfallendes charakteristisches Polynom. Dasselbe gilt daher auch für A. □

16.2. **Die Hauptachsentransformation.** Wir zeigen nun, daß reelle symmetrische Matrizen stets diagonalisierbar sind, und zwar auf orthogonale Art und Weise. Konzeptionell gehört diese Aussage zum Spektralsatz im Kontext eines euklidischen Vektorraums.

Satz 16.4 (Spektralsatz im euklidischen $\mathbb{R}^n$)**.** *Sei $A \in \mathrm{M}_n(\mathbb{R})$ eine symmetrische Matrix. Dann ist A durch eine **orthogonale** Matrix diagonalisierbar: es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ und eine orthogonale Matrix $S \in \mathrm{O}(n)$, so daß*

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Addendum: Die $\lambda_1, \ldots, \lambda_n$ sind die Eigenwerte der Matrix A. Die Spalten von S bilden eine ONB aus Eigenvektoren von A.

Beweis. Nach Satz 16.1 zerfällt $\chi_A(X)$ vollständig in Linearfaktoren. Sei λ_1 eine Nullstelle und b_1 ein normierter Eigenvektor von A zum Eigennwert λ_1. Wie im zweiten Beweis von Satz 16.1 finden ergänzen wir b_1 zu einer ONB $\mathscr{C}$ und finden einen orthogonalen Basiswechsel $T = S^{\mathscr{C}}_{\mathscr{E}}$ (dort S), so daß

$$A' = T^{-1}AT = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

mit symmetrischem $B = B^t \in \mathrm{M}_{n-1}(\mathbb{R})$. Per Induktion nach n finden wir für B eine orthogonale Matrix $U \in \mathrm{O}_{n-1}(\mathbb{R})$ so daß

$$U^{-1}BU = \begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Die Matrix

$$S = T \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix}$$

erfüllt dann die Anforderungen aus dem Theorem, denn S ist als Produkt orthogonaler Matrizen selbst orthogonal und

$$\begin{aligned} S^{-1}AS &= \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix}^{-1} T^{-1}AT \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & U^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & U^{-1}BU \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \end{aligned}$$

Sei $\mathscr{B} = (b_1, \ldots, b_n)$ die Basis aus den Spalten von S. Weil S orthogonal ist, stellt S nach Satz 14.42 die Basiswechselmatrix von einer ONB $\mathscr{B}$ in die Standardbasis $\mathscr{E}$ dar. Das Addendum folgt aus Proposition 12.28. □

Der folgende geometrische Satz und zu verstehen, was der geometrische Gehalt genau sein soll, ist das eigentliche Ziel des Kapitels.

Theorem 16.5 (Hauptachsentransformation)**.** *Sei $A \in \mathrm{M}_n(\mathbb{R})$ eine symmetrische Matrix. Dann gibt es $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ und eine orthogonale Matrix $S \in \mathrm{O}(n)$, so daß*

$$S^tAS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Addendum: Die $\lambda_1, \ldots, \lambda_n$ sind die Eigenwerte der Matrix A. Die Spalten von S bilden eine ONB aus Eigenvektoren von A.

Beweis. Das ist wegen $S^{-1} = S^t$ für alle $S \in \mathrm{O}(n)$ genau Satz 16.4. □

Bemerkung 16.6. Sei $\langle\ ,\ \rangle_A$ die symmetrische Bilinearform auf $\mathbb{R}^n$ mit Gram'scher Matrix A. Dann besagt Theorem 16.5, daß es einen orthogonalen Basiswechsel mittels $S \in \mathrm{O}(n)$ gibt in eine ONB $\mathscr{B}$ von $\mathbb{R}^n$, die auch für $\langle\ ,\ \rangle_A$ orthogonal (aber in der Regel nicht orthonormal) ist. Es gibt somit orthonormale Koordinaten $x = \kappa_{\mathscr{B}}(v)$ auf $\mathbb{R}^n$, so daß

$$\langle v, v\rangle_A = \lambda_1 \cdot (x_1)^2 + \ldots + \lambda_n \cdot (x_n)^2$$

in den Koordinaten bezüglich der neuen ONB Diagonalgestalt annimmt.

Das Addendum beschreibt, wie man die Diagonaleinträge $\lambda_1, \ldots, \lambda_n$ und die Spalten von S bestimmen kann.

Bemerkung 16.7. (1) Der Spektralsatz Satz 16.4 diagonalisiert eine symmetrische Matrix A nach orthogonalem Basiswechsel. Der Satz über die Hauptachsentransformation, Theorem 16.5, diagonalisiert eine zweite symmetrische Bilinearform $\langle\ ,\ \rangle_A$ ebenfalls nach orthogonalem Basiswechsel. Beide Situationen werden durch eine quadratische Matrix dargestellt. Aber der Effekt der Basiswechseltransformation auf die Darstellungsmatrix des Endomorphismus bzw. der symmetrischen Bilinearform ist bei beliebigem (nicht notwendig orthogonalem) Basiswechsel verschieden.

Betrachten wir eine Matrix $A \in \mathrm{M}_n(\mathbb{R})$ als Endomorphsimus von $\mathbb{R}^n$, oder besser als Darstellungsmatrix bezüglich einer Basis $\mathscr{B}$ eines Endomorphismus $f \in \mathrm{End}_{\mathbb{R}}(V)$ für einen $\mathbb{R}$-Vektorraum V, dann hat der Basiswechsel zu einer Basis $\mathscr{C}$ den Effekt

$$A = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f) \rightsquigarrow S^{-1}AS = \mathrm{M}^{\mathscr{C}}_{\mathscr{C}}(f)$$

mit $S = \mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id}_V)$. Faßt man hingegen A als Gram'sche Matrix einer Bilinearform $\langle\ ,\ \rangle$ auf V zur Basis $\mathscr{B}$ auf, so hat der Basiswechsel zu $\mathscr{C}$ den Effekt

$$A = \mathrm{Gram}^{\mathscr{B},\mathscr{B}}(\langle\ ,\ \rangle) \rightsquigarrow S^tAS = \mathrm{Gram}^{\mathscr{C},\mathscr{C}}(\langle\ ,\ \rangle).$$

Unter einem orthogonalen Basiswechsel, und nur für S orthogonal, stimmen wegen

$$S^{-1} = S^t \in \mathrm{O}(n)$$

die beiden Transformationen überein! Ein Normalformensatz für einen Endomorphismus übersetzt sich in einen Normalformensatz für eine Bilinearform und umgekehrt unter der Voraussetzung, daß von einem orthogonalen Basiswechsel die Rede ist.

(2) Für symmetrische Matrizen haben wir bereits eine Diagonalisierbarkeitsaussage. Wir erklären nun die besondere Bedeutung der spezielleren Aussage der Hauptachsentransformation.

Der Diagonalformensatz Theorem 14.24 diagonalisiert symmetrische Bilinearformen über einem beliebigen Körper mit $2 \in K^\times$. Der auftretende Basiswechsel ist aus $\mathrm{GL}_n(K)$.

Im Vergleich dazu studiert die Hauptachsentransformation, Theorem 16.5, vor dem Hintergrund eines euklidischen Vektorraums ($\mathbb{R}^n$ mit dem Standardskalarprodukt) eine weitere symmetrische Bilinearform $\langle\ ,\ \rangle_A$, die mittels bezüglich der euklidischen Struktur orthogonalem Basiswechsel (im Fall von $\mathbb{R}^n$ aus $\mathrm{O}(n)$) in eine Diagonalform gebracht wird.

Diagonalformensatz		**Hauptachsentransformation**
Körper mit $2 \in K^\times$		über $K = \mathbb{R}$
eine symmetrische Bilinearform	$\longleftrightarrow$	zwei symmetrische Bilinearformen eine davon Skalarprodukt
Basiswechsel aus $\mathrm{GL}_n(K)$		Basiswechsel aus $\mathrm{O}_n(\mathbb{R})$

Teil 5. Polynome und Normalformen für Endomorphismen

17. Polynome

17.1. **Der Polynomring.** Für ein Element $x \in K$ eines Körpers und $n \in \mathbb{N}_0$ kürzen wir ab:

$$x^n = \underbrace{x \cdot \ldots \cdot x}_{n\text{-mal}},$$

wobei $x^0 = 1$ sein soll. Es gelten dann die üblichen Potenzgesetze:

$$x^n \cdot x^m = x^{n+m}, \quad (x^n)^m = x^{nm}, \quad (xy)^n = x^n \cdot y^n.$$

Negative Exponenten sind bei $x \neq 0$ zugelassen mit der Bedeutung x^{-n} ist das multiplikative Inverse zu x^n. Insbesondere stimmen die beiden Bedeutungen von x^{-1} überein.

Definition 17.1. Eine **Polynomfunktion** auf einem Körper K ist eine Abbildung $f : K \to K$, so daß es $n \in \mathbb{N}_0$ und $a_0, a_1, \ldots, a_n \in K$ gibt mit

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n, \quad \text{für alle } x \in K.$$

Die Reihenfolge der Quantoren ist wichtig! Die Koeffizienten a_i hängen nicht von x ab.

Ein Polynom (in einer Variablen) ist eine formale Blaupause für eine Polynomfunktion.

Definition 17.2. Ein **Polynom (in einer Variablen X)** mit Koeffizienten im Körper K besteht aus Koeffizienten $a_i \in K$, $i \in \mathbb{N}_0$, so daß es ein $n \in \mathbb{N}_0$ gibt mit

$$a_i = 0 \text{ für alle } i > n.$$

Das zugehörige Polynom $P = P(X)$ wird dann notiert als

$$P = P(X) = a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n = \sum_{i=0}^{n} a_i X^i,$$

wobei n so groß gewählt wird, daß $a_i = 0$ für alle $i > n$. Welche n man dabei wählen darf, hängt vom Polynom $P(X)$ ab. Wenn $a_i = 0$ ist, dann darf der Term $a_i X^i$ in der Notation weggelassen werden. Insbesondere sind, mit n und $P(X)$ wie oben, für alle $m \geq n$ durch

$$\sum_{i=0}^{n} a_i X^i = a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n + 0 \cdot X^{n+1} + \ldots + 0 \cdot X^m = \sum_{i=0}^{m} a_i X^i$$

verschiedene Darstellungen desselben Polynoms $P(X)$ gegeben.

Die Menge aller Polynome mit Koeffizienten aus K bezeichnen wir mit $K[X]$. Formal präzise (aber leider sehr unintuitiv) besteht $K[X]$ einfach aus der Menge der Folgen $(a_i)_{i \in \mathbb{N}_0}$ von Elementen $a_i \in K$, die für hinreichend große i zur Nullfolge $a_i = 0$ werden.

Definition 17.3. Seien $P(X) = \sum_{i=0}^{n} a_i X^i$ und $Q(X) = \sum_{i=0}^{m} b_i X^i$ Polynome mit Koeffizienten aus dem Körper K. Per Definition müssen wir uns a_i für $i > n$ und b_i für $i > m$ durch 0 fortgesetzt denken.

(1) **Addition**: Wir definieren die Summe von P und Q als

$$(P + Q)(X) = \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) X^i.$$

(2) **Multiplikation**: Wir definieren das Produkt von P und Q als

$$(P \cdot Q)(X) = \sum_{k=0}^{n+m} \Big(\sum_{i=0}^{k} a_i b_{k-i} \Big) X^k.$$

(3) Zu $a \in K$ setzen wir

$$a(X) = a = a + 0 \cdot X^1 + 0 \cdot X^2 + \ldots = \sum_{i=0}^{n} a\delta_{i,0} X^i$$

($n \geq 0$ beliebig!) und nennen dies das **konstante Polynom mit Wert** a.

Bemerkung 17.4. Zum Nachweis der Eigenschaften der Multiplikation ist es günstig, wenn man den Koeffizienten vor X^k symmetrischer schreibt als

$$\sum_{i=0}^{k} a_i b_{k-i} = \sum_{i+j=k} a_i b_j.$$

Dabei benutzen wir die Konvention, daß unter der Summe eine Bedingung geschrieben wird und über alle Werte der im Summanden auftretenden Variablen zu summieren ist, welche den Bedingungen genügen. Hier sind das i und j, die von Natur aus ≥ 0 sind, und die Bedingung ist, daß beide zu k summieren. Das sind offensichtlich dieselben Paare (i, j) wie $(i, k-i)$, wenn $i = 0, \ldots, k$ läuft. Die Notation ist etwas suggestiv, aber hilfreich und soll deshalb verwendet werden.

Proposition 17.5. *Sei K ein Körper. Die Menge der Polynome $K[X]$ bildet mit der angegebenen Addition und Multiplikation einen Ring, den* ***Polynomring*** *mit Koeffizienten aus K. Die Eins ist das konstante Polynom* 1.

Beweis. Summe und Produkt von Polynomen sind wieder Polynome und Addition und Multiplikation sind wohldefiniert. Hier müssen wir uns klarmachen, daß die Definition von Summe und Produkt nicht davon abhängt, bei welcher (hinreichend großen) Potenz X^n man die Darstellung der Polynome abgebrochen hat! Da die fraglichen Koeffizienten nur 0 sind, folgt diese Unabhängigkeit sofort.

Wir weisen nun die Ringaxiome nach. Seien dazu beliebig gewählte Elemente aus $K[X]$ gegeben durch

$$P = \sum_{i=0}^{n} a_i X^i, \quad Q = \sum_{i=0}^{m} b_i X^i, \quad R = \sum_{i=0}^{r} c_i X^i.$$

Additive Gruppe: Die Menge der Polynome $K[X]$ kann als eine Teilmenge des K-Vektorraums K^∞ der Folgen $(a_i)_{i \in \mathbb{N}_0}$ verstanden werden und die Addition in $K[X]$ genau die Addition in K^∞ ist. Ferner können wir für $\lambda \in K$ eine Skalarmultiplikation wie folgt definieren

$$\lambda \cdot P = \lambda(X) \cdot P(X) = \sum_{i=0}^{n} (\lambda a_i) X^i.$$

Dies ist nichts anderes als die Einschränkung der Skalarmultiplikation von K^∞ auf $K[X]$. Damit ist $K[X] \subseteq K^\infty$ ein K-Untervektorraum, insbesondere $(K[X], +)$ eine abelsche Gruppe.

Die Multiplikation ist assoziativ: Es gilt

$$\begin{aligned}
P \cdot (Q \cdot R) &= P \cdot \sum_{\ell=0}^{m+r} \Big(\sum_{j+k=\ell} b_j c_k \Big) X^\ell = \sum_{s=0}^{n+m+r} \Big(\sum_{i+\ell=s} a_i \cdot \Big(\sum_{j+k=\ell} b_j c_k \Big) \Big) X^s \\
&= \sum_{s=0}^{n+m+r} \Big(\sum_{i+j+k=s} a_i \cdot \big(b_j c_k \big) \Big) X^s = \sum_{s=0}^{n+m+r} \Big(\sum_{i+j+k=s} \big(a_i b_j \big) \cdot c_k \Big) X^s \\
&= \sum_{s=0}^{n+m+r} \Big(\sum_{\ell+k=s} \Big(\sum_{i+j=\ell} a_i b_j \Big) \cdot c_k \Big) X^s = \sum_{\ell=0}^{n+m} \Big(\sum_{i+j=\ell} a_i b_j \Big) X^\ell \cdot R = (P \cdot Q) \cdot R.
\end{aligned}$$

Die Multiplikation ist kommutativ: Es gilt

$$P(X)\cdot Q(X) = \sum_{k=0}^{n+m} \Big(\sum_{i+j=k} a_i b_j \Big) X^k = \sum_{k=0}^{n+m} \Big(\sum_{i+j=k} b_i a_j \Big) X^k = Q(X)\cdot P(X).$$

Distributivgesetz: Mit P, Q, R wie oben gilt

$$\begin{aligned} P\cdot(Q+R) &= P\cdot \sum_{j=0}^{\max\{m,r\}} (b_j + c_j) X^j \\ &= \sum_{s=0}^{n+\max\{m,r\}} \Big(\sum_{i+j=s} a_i(b_j+c_j) \Big) X^s = \sum_{s=0}^{\max\{n+m,n+r\}} \Big(\sum_{i+j=s} (a_i b_j + a_i c_j) \Big) X^s \\ &= \sum_{s=0}^{n+m} \Big(\sum_{i+j=s} a_i b_j \Big) X^s + \sum_{s=0}^{n+r} \Big(\sum_{i+j=s} a_i c_j \Big) X^s = P\cdot Q + P\cdot R. \end{aligned}$$

Eins: Wir weisen $1 = \sum_{j=0}^{m} \delta_{j,0} X^j$ als die Eins der Multiplikation nach:

$$1\cdot P(X) = \sum_{k=0}^{n} \sum_{i+j=k} a_i \delta_{j,0} X^k = \sum_{k=0}^{n} a_k X^k = P(X). \qquad \square$$

Korollar 17.6. *Der Polynomring $K[X]$ ist ein K-Vektorraum und ein Untervektorraum des K-Vektorraums der Folgen.*

Beweis. Das haben wir im Beweis von Proposition 17.5 nachgewiesen. $\square$

Bemerkung 17.7. Der K-Vektorraum $K[X]$ hat eine Basis gegeben durch die Potenzen von X:

$$1 = X^0, X = X^1, X^2, X^3, \ldots.$$

Damit ist $K[X]$ ein Vektorraum der Dimension ∞, von dem wir eine konkrete Basis kennen.

Bemerkung 17.8. Wenn man in der Definition eines Polynoms die Bedingung unterdrückt, daß in der Folge der Koeffizienten ultimativ nur noch die 0 auftritt, dann bekommt man durch die gleichen Rechnungen wie oben den **formalen Potenzreihenring**

$$K[[X]] = \{ \sum_{i=0}^{\infty} a_i X^i \ ; \ a_i \in K \}.$$

Man darf gar nicht der Versuchung erliegen, hier eine Reihe mit irgendeinem Konvergenzbegriff (außer dem X-adischen, aber das gehört in eine Vorlesung zur kommutativen Algebra) aufsummieren und einen Grenzübergang machen zu wollen. Diese unendlich langen Polynome

$$P(X) = \sum_{i=0}^{\infty} a_i X^i \quad \text{und} \quad Q(X) = \sum_{i=0}^{\infty} b_i X^i$$

sind rein formale Ausdrücke, die mit folgender Addition und Multiplikation

$$\begin{aligned} (P+Q)(X) &= \sum_{i=0}^{\infty} (a_i + b_i) X^i \\ (P\cdot Q)(X) &= \sum_{k=0}^{\infty} \Big(\sum_{i+j=k} a_i b_j \Big) X^k. \end{aligned}$$

einen Ring bilden. Hier muß man sich überlegen, daß der Koeffizient im Produkt nur eine endliche Summe und damit wieder in K definiert ist. Der formale Potenzreihenring enthält den Polynomring als Unterring.

Um ein Bespiel für eine Eigenschaft des formalen Potenzreihenring zu geben: in $K[X]$ hat das Polynom $1 - X$ kein Inverses. In $K[[X]]$ schon:

$$(1 - X)^{-1} = \sum_{i=0}^{\infty} X^i.$$

Definition 17.9. Sei K ein Körper und sei $x \in K$. Die **Auswertung(sabbildung)** in x ist die Abbildung $K[X] \to K$ gegeben durch

$$P = \sum_{i=0}^{n} a_i X^i \quad \mapsto \quad P(x) := a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n.$$

Die Auswertung ist wohldefiniert, weil unterschiedliche Darstellungen $\sum_{i=0}^{n} a_i X^i$ sich nur um Terme mit $a_i = 0$ unterscheiden, für die der Beitrag $a_i x^i = 0 \cdot x^i = 0$ verschwindet.

Proposition 17.10. *Seien K ein Körper und $x \in K$. Die Auswertung in x ist ein **Ringhomomorphismus**: es gilt $1(x) = 1$ und für alle $P, Q \in K[X]$ und $x \in K$ gilt*

$$(P + Q)(x) = P(x) + Q(x)$$
$$(P \cdot Q)(x) = P(x) \cdot Q(x).$$

Für alle $a \in K$ hat das konstante Polynom $a(X)$ den Wert $a(x) = a$.

Beweis. Die Auswertung der Konstanten ist offensichtlich und damit die Verträglichkeit der Auswertung mit 1. Zu zeigen ist noch Verträglichkeit mit Addition und Multiplikation. Seien $P(X) = \sum_{i=0}^{n} a_i X^i$ und $Q(X) = \sum_{i=0}^{m} b_i X^i$. Dann gilt

$$\begin{aligned}(P + Q)(x) &= \Big(\sum_{i=0}^{\max\{n,m\}} (a_i + b_i) X^i \Big)(x) = \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) x^i \\ &= \sum_{i=0}^{\max\{n,m\}} a_i x^i + \sum_{i=0}^{\max\{n,m\}} b_i x^i = P(x) + Q(x).\end{aligned}$$

$$\begin{aligned}(P \cdot Q)(x) &= \Big(\sum_{k=0}^{n+m} \big(\sum_{i+j=k} a_i b_j \big) X^k \Big)(x) = \sum_{k=0}^{n+m} \big(\sum_{i+j=k} a_i b_j \big) x^k \\ &= \Big(\sum_{i=0}^{n} a_i x^i \Big) \cdot \Big(\sum_{j=0}^{m} b_j x^j \Big) = P(x) \cdot Q(x).\end{aligned}$$

□

Bemerkung 17.11. Jedes Polynom $P \in K[X]$ definiert eine Polynomfunktion $K \to K$ durch $x \mapsto P(x)$ mittels Auswertung von Polynomen. Die resultierende Abbildung

$$K[X] \to \mathrm{Abb}(K, K)$$

ist eine K-lineare Abbildung und das Bild ist genau der Unterraum der Polynomfunktionen. Abbildungen mit Werten in K kann man punktweise additieren und multiplizieren: zu $f, g \in \mathrm{Abb}(K, K)$ ist

$$(f + g)(x) = f(x) + g(x) \quad \text{und} \quad (f \cdot g)(x) = f(x) \cdot g(x).$$

Dies macht $\mathrm{Abb}(K, K)$ zu einem Ring und $K[X] \to \mathrm{Abb}(K, K)$ zu einem Ringhomomorphismus.

17.2. **Arithmetik im Polynomring.** Polynome kann man durch ihren Grad sortieren.

Definition 17.12. Sei K ein Körper. Der **Grad** eines Polynoms $P = \sum_{i=0}^{n} a_i X^i \in K[X]$ ist der Index des höchsten auftretenden nichttrivialen Koeffizienten:

$$\deg(P) = \begin{cases} \max\{i \ ; \ a_i \neq 0\} & P \neq 0, \\ -\infty & P = 0. \end{cases}$$

Für alle $P \neq 0$ gibt es einen Koeffizienten ungleich 0. Ist $\deg(P) = d \neq -\infty$, so wird der Koeffizient a_d auch **Leitkoeffizient** oder **führender Koeffizient** genannt.

Beispiel 17.13. (1) Ein $P \in K[X]$ ist genau dann ein konstantes Polynom, wenn $\deg(P) < 1$.
(2) Für alle $\alpha \in K$ gilt

$$\deg(X - \alpha) = 1.$$

Proposition 17.14. *Für alle $P, Q \in K[X]$ gelten die folgenden Aussagen.*
(1) Mit der Konvention $-\infty + d = -\infty$ für alle $d \in \{-\infty\} \cup \mathbb{N}_0$ gilt

$$\deg(P \cdot Q) = \deg(P) + \deg(Q).$$

(2) Mit der Konvention, daß $-\infty < n$ für alle $n \in \mathbb{N}_0$:

$$\deg(P + Q) \leq \max\{\deg(P), \deg(Q)\}.$$

Wenn $\deg(P) \neq \deg(Q)$, dann gilt sogar

$$\deg(P + Q) = \max\{\deg(P), \deg(Q)\}.$$

Beweis. (1) Wenn $P = 0$ oder $Q = 0$, dann ist $PQ = 0$, und beide Seiten der Grad-Gleichung sind $-\infty$. Wir nehmen daher jetzt an, daß $P \neq 0 \neq Q$. Seien $n = \deg(P)$, $m = \deg(Q)$ und

$$P(X) = \sum_{i=0}^{n} a_i X^i \quad \text{und} \quad Q(X) = \sum_{j=0}^{m} b_j X^j$$

mit $a_n \neq 0 \neq b_m$. Wir berechnen die Koeffizienten c_k von PQ mit Index $k \geq n + m$. Diese sind Summen von $a_i b_j$ mit $i + j = k$. Summanden $\neq 0$ erfordern $i \leq n$ und $j \leq m$. Diese treten nur für $k = i + j \leq n + m$ auf. Bei Gleichheit $k = n + m$ müssen auch die Abschätzungen Gleichheiten sein: also nur für $i = n$ und $j = m$. Daher gilt:

$$c_k = \sum_{i+j=k} a_i b_j = \begin{cases} a_n b_m & k = n + m \\ 0 & k > n + m, \end{cases}$$

Weil K ein Körper ist, folgt aus $a_n \neq 0 \neq b_m$ auch $a_n b_m \neq 0$. Daraus folgt $\deg(PQ) = n + m$.

(2) Das ist offensichtlich. □

Korollar 17.15 (Keine Nullteiler)**.** *Seien $P, Q \in K[X]$. Wenn $P \cdot Q = 0$, dann gilt $P = 0$ oder $Q = 0$.*

Beweis. Falls $P \neq 0$ und $Q \neq 0$, dann gilt nach Proposition 17.14 für den Grad des Produkts $\deg(PQ) = \deg(P) + \deg(Q) \geq 0$. Also kann PQ nicht das Nullpolynom sein. □

Korollar 17.16 (Kürzungsregel)**.** *Seien $P_1, P_2, Q \in K[X]$, $Q \neq 0$. Dann gilt*

$$QP_1 = QP_2 \quad \Longrightarrow \quad P_1 = P_2.$$

Beweis. Aus $QP_1 = QP_2$ folgt $Q(P_1 - P_2) = 0$. Weil $Q \neq 0$ muß laut Korollar 17.15 dann $P_1 - P_2 = 0$ gelten. Aber das ist $P_1 = P_2$. □

Satz 17.17 (Polynomdivision, Division mit Rest)**.** *Sei K ein Körper. Der Polynomring $K[X]$ hat* ***Division mit Rest bezüglich der Grad-Funktion****, d.h. zu allen $P, D \in K[X]$, $D \neq 0$ gibt es $Q, R \in K[X]$ mit*
(i) $P = Q \cdot D + R$
(ii) $\deg(R) < \deg(D)$.
Zusatz: die Polynome Q und R sind eindeutig durch P, D bestimmt.

Beweis. Wir zeigen zunächst per Induktion nach $\deg(P)$ die Existenz von Q, R. Als Induktionsanfang[5] beweisen wir die Existenz für alle P mit $\deg(P) < \deg(D)$. Dieser Fall ist nicht leer: für $P = 0$ gilt $\deg(P) = -\infty < \deg(D)$ (per Konvention). Falls $\deg(P) < \deg(D)$, so wählen wir $Q = 0$ und $R = P$, fertig.

Jetzt behandeln wir den Induktionsschritt. Sei dazu $n = \deg(P) \geq d = \deg(D)$, und die Existenz von Q und R wie im Satz gilt für Polynome vom Grad kleiner als n. Wir schreiben

$$P(X) = a_n X^n + \dots \text{ Terme kleineren Grades}$$
$$D(X) = b_d X^d + \dots \text{ Terme kleineren Grades}$$

mit $a_n \neq 0 \neq b_d$. Dann ist

$$\begin{aligned}\tilde{P} &= P - a_n b_d^{-1} X^{n-d} D \\ &= a_n X^n - a_n b_d^{-1} X^{n-d} \cdot b_d X^d + \dots \text{ Terme vom Grad } < n \;=\; \text{Terme vom Grad } < n,\end{aligned}$$

also $\deg(\tilde{P}) < \deg(P)$. Per Induktionsannahme gibt es Polynome $\tilde{Q}, \tilde{R}$ mit $\tilde{P} = \tilde{Q}D + \tilde{R}$ und $\deg(\tilde{R}) < \deg(D)$. Dann gilt

$$P = \tilde{P} + a_n b_d^{-1} X^{n-d} D = \tilde{Q}D + \tilde{R} + a_n b_d^{-1} X^{n-d} D = \big(\tilde{Q} + a_n b_d^{-1} X^{n-d}\big) D + \tilde{R},$$

und $Q = \tilde{Q} + a_n b_d^{-1} X^{n-d}$ und Rest $R = \tilde{R}$ sind wie im Satz gefordert.

Jetzt kümmern wir uns um die Eindeutigkeit. Angenommen es gäbe auch Q', R' mit denselben Eigenschaften. Dann ist

$$(Q - Q')D = QD - Q'D = (P - R) - (P - R') = (R' - R). \tag{17.1}$$

Wenn $Q \neq Q'$, dann folgt aus Proposition 17.14 (mit $\deg(R') = \deg(-R')$)

$$\begin{aligned}\deg(D) &> \max\{\deg(R), \deg(-R')\} \geq \deg(R' - R) \\ &= \deg((Q - Q')D) = \deg(Q - Q') + \deg(D) \geq \deg(D),\end{aligned}$$

ein Widerspruch. Also ist Q eindeutig: $Q = Q'$. Aber dann ist $R = R'$ wegen (17.1). Dies zeigt die Eindeutigkeit der Polynome Q und R. □

Bemerkung 17.18. Der Beweis von Satz 17.17 liefert in Wahrheit den Algorithmus der Polynomdivision zur Bestimmung von Q und R. Die Induktion im Beweis von Satz 17.17 liefert den Quotienten Q als

$$Q = a_n b_d^{-1} X^{n-d} + \tilde{Q}.$$

Aus Proposition 17.14 folgt

$$\deg(Q) = \deg(P) - \deg(D) = n - d,$$

[5] Der Induktionsanfang umfaßt mehr als nur die Aussage für den kleinsten Wert des Induktionsparameters. Das ist ungewöhnlich, aber paßt zur zu beweisenden Aussage. Man mache sich als Übungsaufgabe klar, daß unser Induktionsparameter $\deg(P)$ durch die Werte

$$-\infty, 0, 1, 2, 3, \dots$$

variiert, und daß es für die vollständige Induktion ausreicht, diese in einem nichtleeren(!) Anfangsstück dieser Paramterwerte zu verankern.

und ebenso $\deg(\tilde{Q}) = \deg(\tilde{P}) - d < n - d$. Demnach berechnet der Induktionsschritt den Koeffizienten in $Q(X)$ von X^{n-d}. Polynomdivision berechnet man also mit dem folgenden Algorithmus (Notation wie oben):

Solange $\deg(P) \geq \deg(D)$*:*

> Man bestimme den Quotienten $c_{n-d} = a_n b_d^{-1}$ der Leitkoeffizienten von P und D, notiere
> $$+c_{n-d}X^{n-d}.$$
> Man bestimme $\tilde{P} = P - c_{n-d}X^{n-d}D$, und starte erneut mit $\tilde{P}$ anstelle von P.

Wenn $\deg(P) < \deg(D)$*:*

> dann ist $R = P$ der Rest, und Q ist die Summe der notierten Terme. Der Algorithmus endet dann.

Beispiel 17.19. Wir berechnen in $\mathbb{Q}[X]$ die Division mit Rest von $P(X) = 2X^4 + 1$ durch $D(X) = X^2 + 1$.

- Der Quotient der Leitkoeffizienten ist $2/1 = 2$, also notieren wir $2X^2$ und berechnen
$$(2X^4 + 1) - 2X^2(X^2 + 1) = -2X^2 + 1.$$
- Der Quotient der Leitkoeffizienten ist $-2/1 = -2$, also notieren wir $-2X^0 = -2$ und berechnen
$$(-2X^2 + 1) - (-2)(X^2 + 1) = 3.$$
- $\deg(3) = 0 < 2 = \deg(X^2 + 1)$, also ist $R = 3$ und $Q = 2X^2 - 2$ und der Algorithmus stoppt. Es gilt
$$P(X) = (2X^2 - 2)D(X) + 3.$$

17.3. **Nullstellen, Linearfaktoren und Multiplizitäten.** Für die Anwendung von Polynomen in der linearen Algebra müssen wir etwas über die Nullstellen von Polynomen wissen.

Definition 17.20. Eine **Nullstelle** eines Polynoms $P \in K[X]$ ist ein $x_0 \in K$ mit $P(x_0) = 0$.

Beispiel 17.21. Das Polynom $X^2 + 1$ hat als Polynom in $\mathbb{R}[X]$ keine Nullstellen. Für jedes $x \in \mathbb{R}$ ist
$$x^2 + 1 > 0.$$
Aber als Polynom in $\mathbb{C}[X]$ kann man schreiben
$$X^2 + 1 = (X + i)(X - i).$$
Hieraus folgt sofort (auch sonst klar), daß $x = i$ und $x = -i$ Nullstellen von $X^2 + 1$ in $\mathbb{C}$ sind. Es folgt aber auch, daß dies die einzigen Nullstellen sind, denn für jedes $z \neq \pm i$ sind die Faktoren $z \pm i$ in
$$(X^2 + 1)(z) = z^2 + 1 = (z + i)(z - i)$$
von 0 verschieden. Da $\mathbb{C}$ ein Körper ist, kann dann das Produkt nicht 0 werden.

Proposition 17.22 (Nullstellen versus Linearfaktoren)**.** *Sei K ein Körper und P ein Polynom aus $K[X]$. Sei $\alpha \in K$. Dann sind äquivalent:*

(a) *α ist Nullstelle: $P(\alpha) = 0$.*

(b) *$X - \alpha$ ist ein Faktor von P: es gibt ein $Q \in K[X]$ mit*
$$P(X) = (X - \alpha)Q(X).$$

Beweis. (b) $\Longrightarrow$ (a): Hier müssen wir nur einsetzen und benutzen, daß die Auswertung ein Ringhomomorphismus ist:

$$P(\alpha) = \big((X-\alpha)Q(X)\big)(\alpha) = (\alpha-\alpha)Q(\alpha) = 0.$$

(a) $\Longrightarrow$ (b): Sei nun α eine Nullstelle von P. Wir berechnen Polynomdivison von $P(X)$ durch $D(X) = (X-\alpha)$ nach Satz 17.17. Das liefert Polynome Q, R mit

$$P(X) = (X-\alpha)Q(X) + R(X)$$

und $\deg(R) < \deg(X-\alpha) = 1$. Folglich ist $R(X)$ ein konstantes Polynom. Seinen Wert bestimmen wir durch Auswerten als

$$R(\alpha) = \big(P - (X-\alpha)Q\big)(\alpha) = P(\alpha) - (\alpha-\alpha)Q(\alpha) = P(\alpha).$$

Aber weil α eine Nullstelle von P ist, handelt es sich bei $R(X)$ um das Nullpolynom. Aus $R = 0$ folgt aber

$$P(X) = (X-\alpha)Q(X). \qquad \square$$

Der folgende Satz liefert eine auf Linearfaktorenm beschränkte Aussage über die eindeutige Faktorisierung eines Polynoms in Primpolynome. Die allgemeine Aussage über Hauptidealringe, die sich gleichzeitig um den Polynomring $K[X]$ und die ganzen Zahlen $\mathbb{Z}$ kümmert, beweisen wir in der Vorlesung *Lineare Algebra 2*.

Satz 17.23. *Sei $P \in K[X]$, $P \neq 0$ ein Polynom.*

(1) Das Polynom P hat eine Darstellung

$$P(X) = (X-\alpha_1)^{e_1} \cdot \ldots \cdot (X-\alpha_r)^{e_r} \cdot Q(X) \tag{17.2}$$

mit $r \in \mathbb{N}_0$, paarweise verschiedenen $\alpha_1, \ldots, \alpha_r \in K$, Exponenten $e_1, \ldots, e_r \in \mathbb{N}$ und einem Polynom $Q \in K[X]$ ohne Nullstellen.

(2) Die Nullstellen von P sind genau $\alpha_1, \ldots, \alpha_r$.

(3) Die Darstellung (17.2) ist eindeutig bis auf Permutation der Faktoren $(X-\alpha_i)^{e_i}$.

Beweis. Wir zeigen zuerst die Existenz einer solchen Darstellung per Induktion über $\deg(P)$. Wenn P keine Nullstelle in K hat, dann ist $r = 0$ und $P = Q$. Dieser Fall schließt den Induktionsanfang ein, wenn $\deg(P) = 0$.

Nehmen wir nun an, daß alle Polynome vom Grad kleiner als $\deg(P)$ bereits eine Darstellung (17.2) besitzen. Außerdem dürfen wir annehmen, daß P eine Nullstelle $\alpha \in K$ hat, denn der Fall ohne Nullstellen ist ja bereits abgehandelt. Nach Proposition 17.22 gibt es dann ein Polynom $\tilde{P}(X)$ mit

$$P(X) = (X-\alpha)\tilde{P}(X).$$

Es gilt nach Proposition 17.14

$$\deg(\tilde{P}) = \deg(P) - \deg(X-\alpha) = \deg(P) - 1 < \deg(P).$$

Daher besitzt nach Induktionsannahme $\tilde{P}(X)$ eine Darstellung wie behauptet. Die gesuchte Darstellung (17.2) von P ergibt sich nun durch Multiplikation mit dem Faktor $X-\alpha$. Dies zeigt die Existenz.

(2) Als nächstes zeigen wir, daß die in (17.2) vorkommenden $\alpha_1, \ldots, \alpha_r$ genau die Menge der Nullstellen von P sind. Nun, $P(x) = 0$ ist für $x \in K$, weil Auswerten ein Ringhomomorphismus ist, äquivalent zu

$$0 = P(x) = (x-\alpha_1)^{e_1} \ldots (x-\alpha_r)^{e_r} Q(x).$$

In einem Körper ist ein Produkt genau dann 0, wenn es einen Faktor gibt, der 0 ist. Der Faktor $Q(x)$ ist nach Voraussetzung stets $\neq 0$, also kommt nur einer der Faktoren $(x-\alpha_i)$ in Frage. Jede Nullstelle von P findet sich also unter den $\alpha_1, \ldots, \alpha_r$. Umgekehrt ist auch jedes α_i, $i = 1, \ldots, r$ eine Nullstelle wegen $e_i \geq 1$ und Proposition 17.22.

(3) Nun zur Eindeutigkeit. Dies zeigen wir wieder per Induktion über $\deg(P)$. Der Induktionsanfang wird wieder durch den Fall der Polynome ohne Nullstellen gegeben. Weil wir gerade gesehen haben, daß die $\alpha_1, \ldots, \alpha_r$ genau die Nullstellen von P sind, kommen diese hier nicht vor und notwendigerweise $P = Q$. Das ist eindeutig.

Nehmen wir nun an, daß die Darstellung (17.2) für alle Polynome mit Grad kleiner $\deg(P)$ bereits als eindeutig bewiesen ist. Und nehmen wir an, daß

$$P(X) = (X - \alpha_1)^{e_1} \cdot \ldots \cdot (X - \alpha_r)^{e_r} \cdot Q(X)$$

eine weitere Darstellung

$$P(X) = (X - \beta_1)^{f_1} \cdot \ldots \cdot (X - \beta_s)^{f_s} \cdot R(X)$$

besitzt. Da P eine Nullstelle hat, ist $r \geq 1$. Die Nullstelle α_1 kommt dann unter den β_i vor. Nach Anwendung einer Permutation der β_i dürfen wir annehmen, daß $\alpha_1 = \beta_1$. Wir kürzen einen Faktor, schließen mit Korollar 17.16 auf

$$(X - \alpha_1)^{e_1 - 1} \cdot \ldots \cdot (X - \alpha_r)^{e_r} \cdot Q(X) = (X - \beta_1)^{f_1 - 1} \cdot \ldots \cdot (X - \beta_s)^{f_s} \cdot R(X)$$

und dies ist ein Polynom vom Grad $\deg(P) - 1$. Aus der Induktionsannahme folgt nun die Eindeutigkeit für dieses Polynom (bis auf Permuation der Faktoren) und damit die Eindeutigkeit für P. □

Korollar 17.24. *Ein Polynom $P \in K[X]$, $P \neq 0$ hat höchstens $\deg(P)$-viele verschiedene Nullstellen.*

Beweis. Nach Satz 17.23 und mit dessen Notation hat P genau r Nullstellen und

$$\deg(P) = \deg(Q) + \sum_{i=1}^{r} e_i \geq 0 + \sum_{i=1}^{r} 1 = r.$$ □

Definition 17.25. Mit der Notation von Satz 17.23 bezeichnen wir den Exponenten e_i als die **Multiplizität (Vielfachheit)** der Nullstelle α_i.

Definition 17.26. Ein Polynom $P(X) \in K[X]$, $P(X) \neq 0$, **zerfällt vollständig (in Linearfaktoren)**, wenn es $c, \alpha_1, \ldots, \alpha_r \in K$ und $e_1, \ldots, e_r \in \mathbb{N}$ gibt mit

$$P(X) = c \cdot (X - \alpha_1)^{e_1} \cdot \ldots \cdot (X - \alpha_r)^{e_r}.$$

Korollar 17.27. *Ein Polynom $P \in K[X]$, $P \neq 0$ hat höchstens $\deg(P)$-viele verschiedene Nullstellen, wenn man jede mit ihrer Multiplizität gewichtet zählt. Es gilt Gleichheit von $\deg(P)$ und der Summe der Multiplizitäten der Nullstellen genau dann, wenn P vollständig zerfällt.*

Beweis. Nach Satz 17.23 und mit dessen Notation gilt

$$\deg(P) = \deg(Q) + \sum_{i=1}^{r} e_i \geq \sum_{i=1}^{r} e_i.$$

Es gilt Gleichheit genau dann, wenn Q konstantes Polynom ist, also genau dann wenn P vollständig zerfällt. □

Beispiel 17.28. Für jeden Körper K gibt es in $K[X]$ das Polynom $P(X) = X^2 + 1$. Wir bestimmen Nullstellen und Multiplizitäten für verschiedene Wahlen von K.

(1) Sei $K = \mathbb{C}$. Dann kennen wir die Nullstelle $x_0 = i$. Daher hat $X^2 + 1$ den Linearfaktor $X - i$. In der Faktorisierung $X^2 + 1 = (X - i) \cdot Q(X)$ hat das Polynom Q den Grad

$$\deg(Q) = \deg(X^2 + 1) - \deg(X - i) = 2 - 1 = 1,$$

ist also auch ein Linearfaktor. In der Tat, auch $-i$ ist Nullstelle und (sofort aus der dritten Binomischen Formel)

$$X^2 + 1 = (X - i)(X + i).$$

Die Multiplizitäten beider Nullstellen sind 1.

(2) Sei $K = \mathbb{R}$. Dann ist für alle $x \in \mathbb{R}$ das Quadrat $x^2 \geq 0$ nicht negativ und daher $P(x) = x^2 + 1 \geq 1$. Hier ist $P(X)$ selbst ein Polynom ohne Nullstellen.

(3) Sei $K = \mathbb{F}_5$. Wir schreiben zu $n \in \mathbb{Z}$ die Restklasse $[n] \in \mathbb{F}_5$ zur Notationsvereinfachung einfach als n. Dann ist $P(\pm 2) = (\pm 2)^2 + 1 = 5 = 0$ und

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

Weil $2 \neq -2$ in $\mathbb{F}_5$ gilt, haben wir zwei Nullstellen jeweils mit Multiplizität 1.

(4) Sei nun $K = \mathbb{F}_2$. Dann ist $P(0) = 1$ und $P(1) = 0$. Also hat P nur eine Nullstelle $x_0 = 1$. Nach Abspalten eines Faktors $X - 1$ bleibt ein Polynom vom Grad 1. Dies führt zu einer weiteren Nullstelle, die aber auch nur 1 sein kann. Wir rechnen leicht nach, daß in $\mathbb{F}_2[X]$ gilt

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

somit haben wir eine Nullstelle mit Multiplizität 2.

Für eine Teilmenge $S \subseteq K$ führt die Kollektion aller Auswertungen in den $x \in S$ zu einem Ringhomomorphismus

$$K[X] \to \mathrm{Abb}(S, K), \qquad P \mapsto \big(x \in S \mapsto P(x)\big).$$

Diese Abbildung wollen wir nun untersuchen, und zwar insbesondere im Fall $S = K$.

Korollar 17.29. *Sei K ein Körper mit unendlich vielen Elementen. Dann ist die Auswertungsabbildung*

$$K[X] \to \mathrm{Abb}(K, K)$$

injektiv aber nicht surjektiv.

Beweis. Wenn die Polynomfunktion zu $P \in K[X]$ die Nullfunktion ist, dann hat P unendlich viele Nullstellen. Also folgt aus Korollar 17.24, daß $P = 0$. Die Abbildung hat daher einen trivialen Kern und ist injektiv.

Betrachten wir weiter die charakteristische Abbildung $\mathbf{1}_a$ von $a \in K$, also $\mathbf{1}_a(x) = 0$ für $a \neq x$ und $\mathbf{1}_a(a) = 1$. Dann hat $\mathbf{1}_a$ unendlich viele Nullstellen und dasselbe Argument zeigt, daß nur das Nullpolynom $\mathbf{1}_a$ darstellen könnte. Aber $\mathbf{1}_a \neq 0$ ist nicht die Nullfunktion, also ist die Abbildung $K[X] \to \mathrm{Abb}(K, K)$ nicht surjektiv. □

Satz 17.30 (Lagrangeinterpolation)**.** *Sei K ein Körper und $S = \{x_0, \ldots, x_d\} \subseteq K$ eine endliche Teilmenge der Mächtigkeit $d + 1$.*

(1) Seien $b_0, \ldots, b_d \in K$ beliebeig gegeben. Dann gibt es genau ein Polynom $P \in K[X]$ vom Grad $\deg(P) \leq d$ mit

$$P(x_i) = b_i, \quad \text{für alle } x_i \in S.$$

(2) Mit $\mathrm{Pol}_{K,\leq d}$ bezeichnen wir den Untervektorraum von $K[X]$ der Polynome vom Grad höchstens d. Die Auswertungsabbildung in S ist ein K-Vektorraumisomorphismus

$$\mathrm{Pol}_{K,\leq d} \xrightarrow{\sim} \mathrm{Abb}(S, K).$$

Beweis. Aussage (1) und (2) sind unmittelbar äquivalent. Der Kern der Auswertungsabbildung besteht aus Polynomen P vom Grad $\leq d$, die in allen $x \in S$ den Wert Null haben, also $d+1$ Nullstellen haben. Aus Korollar 17.24 folgt $P = 0$. Damit ist die Auswertungsabbildung injektiv.

Zur Surjektivität benutzen wir die Formel der Lagrangeinterpolation. Das Polynom

$$P(X) = \sum_{i=0}^{d} b_i \cdot \prod_{j \neq i} \frac{X - x_j}{x_i - x_j}$$

löst die gestellte Interpolationsaufgabe, wie man durch Einsetzen sofort nachrechnet. □

Korollar 17.31. *Sei K ein endlicher Körper. Dann ist die Auswertungsabbildung*

$$K[X] \to \mathrm{Abb}(K, K)$$

surjektiv aber nicht injektiv.

Beweis. Das Polynom

$$P(X) = \prod_{\alpha \in K} (X - \alpha)$$

hat Grad

$$\deg(P) = \deg\Big(\prod_{\alpha \in K} (X - \alpha)\Big) = \sum_{\alpha \in K} \deg(X - \alpha) = \sum_{\alpha \in K} 1 = |K|,$$

und daher ist $P \neq 0$. Andererseits ist die zugehörige Polynomfunktion die Nullfunktion. Die Abbildung des Satzes ist somit nicht injektiv.

Lagrange-Interpolation, Satz 17.30, zeigt im Fall $S = K$ die Surjektivität. □

17.4. Der Polynomring ist ein Hauptidealring.

Definition 17.32. Ein **Ideal** in einem Ring R ist eine Teilmenge $I \subseteq R$, $I \neq \emptyset$ mit den folgenden Eigenschaften:

(i) für alle $x, y \in I$ ist $x + y \in I$,
(ii) für alle $x \in I$ und $a \in R$ gilt $ax \in I$.

Beispiel 17.33. (1) Seien R ein Ring und $x \in R$ ein beliebiges Element. Dann ist

$$(x) := Rx = \{ax \ ; \ a \in R\}$$

ein Ideal. In der Tat gibt es für alle $y_1, y_2 \in (x)$ Elemente $a_1, a_2 \in R$ mit $y_1 = a_1 x$ und $y_2 = a_2 x$. Dann gilt

$$y_1 + y_2 = a_1 x + a_2 x = (a_1 + a_2)x \in (x),$$

und für alle $b \in R$

$$by_1 = b(a_1 x) = (ba_1)x \in (x).$$

Ideale dieser Art heißen **Hauptideale** von R.

(2) In jedem Ring R gibt es die folgenden Ideale:

$$R = (1) \quad \text{und} \quad (0) = \{0\}.$$

Definition 17.34. Ein **Hauptidealring** ist ein Ring R, in dem

(i) jedes Ideal von R ein Hauptideal ist,
(ii) $0 \neq 1$,
(iii) und die Kürzungsregel gilt: für alle $a, x, y \in R$, $a \neq 0$, folgt aus $ax = ay$ bereits $x = y$.

Satz 17.35. *Sei K ein Körper. Der Polynomring $K[X]$ ist ein Hauptidealring.*

Beweis. Wir müssen nur zeigen, daß jedes Ideal I von $K[X]$ ein Hauptideal ist. Das Ideal $\{0\} = (0)$ ist ein Hauptideal. Wir dürfen daher davon ausgehen, daß es $P \in I$ gibt mit $\deg(P) \geq 0$. Damit ist

$$d = \min\{\deg(P) \ ; \ P \in I, \ P \neq 0\} \geq 0$$

wohldefiniert.

Sei $P \in I$ mit $\deg(P) = d$ und $F \in I$ ein beliebiges Element. Division mit Rest für Polynome liefert $Q, R \in K[X]$ mit $F = QP + R$ und

$$\deg(R) < \deg(P).$$

Weil

$$R = F - QP \in I,$$

folgt aus der Minimalität des Grades von P unter den Elementen von $I \setminus \{0\}$, daß

$$R = 0.$$

Dies zeigt $F = QP$, also $F \in (P)$. Weil F beliebig war, ist $I = (P)$ ein Hauptideal. □

Korollar 17.36. *Für jedes Ideal $0 \neq I \subseteq K[X]$ gibt es ein eindeutiges normiertes $0 \neq P \in K[X]$ mit $I = (P)$. Dies ist das eindeutige normierte Polynom $P \in I$ mit*

$$\deg(P) = \min\{\deg(F) \ ; \ F \in I, \ F \neq 0\}.$$

Beweis. Indem wir mit dem Inversen des Leitkoeffizienten multiplizieren (damit bleibt man im Ideal!), können wir annehmen, daß das P im Beweis von Satz 17.35 ein normiertes Polynom mit dem angegebenen Grad ist. Dies zeigt die Existenz. Angenommen $Q \neq P$ wäre ein weiteres solches Polynom. Dann ist auch $P - Q \in I$, und $\deg(P - Q) < \deg(P)$, ein Widerspruch. □

Ideale erlauben eine alternative Sichtweise auf das Minimalpolynom wie folgt.

Proposition 17.37. *Sei $A \in \mathrm{M}_n(K)$. Dann ist die Menge der Polynome, die A als „Nullstelle“ haben*

$$I_A = \{P \in K[X] \ ; \ P(A) = 0\},$$

ein Ideal in $K[X]$, das nicht das Nullideal ist.

Beweis. Aus $P(A) = 0$, $Q(A) = 0$ und $R \in K[X]$ beliebig, folgt

$$(P + Q)(A) = P(A) + Q(A) = 0$$
$$(RP)(A) = R(A)P(A) = 0.$$

Wir müssen noch einsehen, daß $I_A \neq 0$. Dazu betrachten wir die $(n^2 + 1)$-vielen Matrizen

$$\mathbf{1} = A^0, A, A^2, A^3, \ldots, A^{n^2}$$

als Vektoren des Vektorraums $\mathrm{M}_n(K)$. Weil es mehr Vektoren als $n^2 = \dim \mathrm{M}_n(K)$ sind, folgt aus Satz 5.31, daß diese Matrixpotenzen linear abhängig sind. Zu einer nichttrivialen Linearkombination $\sum_{i=0}^{n^2} a_i A^i = 0$ gehört das Polynom

$$P(X) = \sum_{i=0}^{n^2} a_i X^i \neq 0$$

mit $P(A) = 0$. Also ist $P \in I_A$ und damit I_A nicht das Nullideal. □

Definition 17.38. Das **Minimalpolynom** einer Matrix $A \in \mathrm{M}_n(K)$ ist das eindeutige normierte Polynom $m_A(X)$ mit
$$(m_A(X)) = \{P \in K[X] \;;\; P(A) = 0\}.$$
Das **Minimalpolynom** eines Endomorphismus $f : V \to V$ ist das eindeutige normierte Polynom $m_f(X)$ mit
$$(m_f(X)) = \{P \in K[X] \;;\; P(f) = 0\}.$$

17.5. **Das charakteristische Polynom.** Jetzt kommen wir der systematischen Bestimmung von Eigenwerten einen Schritt näher.

Definition 17.39. Sei K ein Körper. Unter der Determinante einer Matrix
$$A(X) = (a_{ij}(X)) \in \mathrm{M}_n(K[X])$$
mit Einträgen $a_{ij}(X) \in K[X]$ im Polynomring verstehen wir das Polynom
$$\det(A(X)) \in K[X],$$
das sich aus der Leibniz-Formel ergibt:
$$\det(A(X)) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}(X).$$

Seien $A(X) = (a_{ij}(X)) \in \mathrm{M}_n(K[X])$ und $\lambda \in K$. Mit $A(\lambda) = (a_{ij}(\lambda)) \in \mathrm{M}_n(K)$ bezeichnen wir die Auswertung von $A(X)$, genauer die Matrix der Auswertungen der Matrixeinträge von $A(X)$ in λ.

Lemma 17.40. *Auswerten vertauscht mit der Determinante, d.h., für alle $\lambda \in K$ gilt:*
$$\det(A(\lambda)) = \det(A(X))(\lambda).$$

Beweis. Das ist eine einfache Rechnung:
$$\det(A(\lambda)) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}(\lambda) = \Big(\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}(X) \Big)(\lambda) = \det(A(X))(\lambda). \qquad \square$$

Bemerkung 17.41. Lemma 17.40 sagt, daß das Diagramm
$$\begin{array}{ccc} \mathrm{M}_n(K[X]) & \xrightarrow{\det} & K[X] \\ {\scriptstyle A(X)\mapsto A(\lambda)}\downarrow & & \downarrow{\scriptstyle P(X)\mapsto P(\lambda)} \\ \mathrm{M}_n(K) & \xrightarrow{\det} & K \end{array}$$
kommutiert, d.h., es ist egal ob man die Abbildungen von links oben nach rechts unten entlang der Abbildung des einen oder des anderen Weges komponiert.

Definition 17.42. Sei K ein Körper. Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Das **charakteristische Polynom** von A ist
$$\chi_A(X) = \det(X \cdot \mathbf{1} - A) \in K[X].$$

Proposition 17.43. *Ähnliche Matrizen haben das gleiche charakteristische Polynom: für alle $A \in \mathrm{M}_n(K)$ und $S \in \mathrm{GL}_n(K)$ ist*
$$\chi_{SAS^{-1}}(X) = \chi_A(X).$$

Beweis. Wir rechnen $S(X \cdot \mathbf{1} - A)S^{-1} = X \cdot \mathbf{1} - SAS^{-1}$ und daher:

$$\begin{aligned}\chi_{SAS^{-1}}(X) &= \det(X \cdot \mathbf{1} - SAS^{-1}) = \det(S(X \cdot \mathbf{1} - A)S^{-1}) \\ &= \det(S)\det(X \cdot \mathbf{1} - A)\det(S)^{-1} = \chi_A(X).\end{aligned}$$

Dabei haben wir den Determinantenmultiplikationssatz für Matrizen mit Einträgen im Polynomring benutzt. Dies rechtfertigen wir in den folgenden Bemerkungen 17.44 und 17.45. □

Bemerkung 17.44. Den Ring der ganzen Zahlen $\mathbb{Z}$ kann man in den Körper $\mathbb{Q}$ einbetten. Genauer konstruiert man $\mathbb{Q}$ als Körper der Brüche mit Zähler und Nenner aus $\mathbb{Z}$, siehe Abschnitt 3.1.3.

Dieselbe Konstruktion kann man für einen Körper K auf den Polynomring $K[X]$ anwenden. Es entsteht der Körper $K(X)$, dessen Elemente **rationale** Funktionen $f = p/q$ mit $p, q \in K[X]$, und Nenner $q \neq 0$, sind. Formal korrekt sind rationale Funktionen Äquivalenzklassen nach der Äquivalenzrelation die von Kürzen und Erweitern erzeugt wird: für $f = p/q$ und $g = r/s$ gilt

$$f = g \text{ in } K(X) \quad \Longleftrightarrow \quad ps = rq \text{ in } K[X].$$

So wie man $\mathbb{Z}$ als Unterring von $\mathbb{Q}$ auffaßt, kann man Polynome als rationale Funktionen (mit Nenner 1) auffassen und erhält so den Polynomring als Unterring

$$K[X] \subseteq K(X)$$

des Körpers der rationalen Funktionen. Gleichungen in $K[X]$ kann man daher in dem Körper $K(X)$ prüfen.

Bemerkung 17.45. Der Determinantenmultiplikationssatz[6] gilt auch für Matrizen mit Einträgen in Polynomen. Sind $A, B \in \mathrm{M}_n(K[X])$ polynomwertige Matrizen, dann folgt

$$\det(AB) = \det(A)\det(B),$$

weil diese Gleichung im Körper $L = K(X)$ geprüft werden kann, denn man kann A und B als Matrizen in $\mathrm{M}_n(L)$ auffassen, wobei der Wert der Determinante sich nicht ändert. Schließlich gilt für Matrizen mit Einträgen aus einem Körper (hier L) der Determinantenmultiplikationssatz bereits.

Definition 17.46. Sei K ein Körper. Sei V ein K-Vektorraum und $f : V \to V$ ein Endomorphismus. Das **charakteristische Polynom** von f ist

$$\chi_f(X) = \chi_A(X) \in K[X]$$

für eine (jede) Darstellungsmatrix $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ von f.

Lemma 17.47. *Das charakteristische Polynom eines Endomorphismus ist wohldefiniert.*

Beweis. Wir müssen begründen, warum $\chi_f(X)$ von der Wahl der Basis unabhängig ist. Die Darstellungsmatrizen zu verschiedenen Basen sind ähnlich, siehe Proposition 12.24, und ähnliche Matrizen haben nach Proposition 17.43 das gleiche charakteristische Polynom. □

[6] Der Determinantenmultiplikationssatz gilt sogar für Matrizen mit Einträgen in einem beliebigen Ring! Das ist etwas aufwendiger. Entweder beweist man den Determinantenmultiplikationssatz $\det(A)\det(B) = \det(AB)$ nochmals durch eine wenig erhellende Indexschlacht mit der Leibniz-Formel für die jeweiligen Determinanten. Oder man überlegt sich, daß es ausreicht, $\det(A)\det(B) = \det(AB)$ zu zeigen für universelle Matrizen, deren Einträge a_{ij} und b_{ij} unabhängige Variablen sind. Man arbeitet dann im Polynomring $R = \mathbb{Z}[a_{ij}, b_{ij} \; ; \; i, j = 1, \ldots, n]$ mit n^2-vielen Variablen. Dieser Polynomring R ist in einem Körper K enthalten. Damit wird der Determinantenmultiplikationssatz $\det(A)\det(B) = \det(AB)$ für die universellen A und B zu einem Spezialfall des Determinantenmultiplikationssatzes für Matrizen mit Einträgen aus K. Dieser Fall mit Koeffizienten aus einem Körper K ist aber schon bewiesen!

Satz 17.48. *Seien K ein Körper, $A \in \mathrm{M}_n(K)$ eine quadratische Matrix und $f : V \to V$ ein Endomorphismus des endlichdimensionalen K-Vektorraums V.*
(1) Die Eigenwerte von A sind genau die Nullstellen von $\chi_A(X)$ aus K.
(2) Die Eigenwerte von f sind genau die Nullstellen von $\chi_f(X)$ aus K.

Beweis. Weil nach Lemma 17.40 das Auswerten mit Determinante vertauscht, folgt die Behauptung sofort aus Satz 12.8 und Korollar 12.9. □

Die Vielfachheit eines Eigenwerts kann man auf zwei Arten angeben.

Definition 17.49. (1) Die **algebraische Multiplizität (Vielfachheit)** eines Eigenwerts λ der Matrix A (bzw. des Endomorphismus f) ist die Multiplizität von λ als Nullstelle des charakteristischen Polynoms $\chi_A(X)$ (bzw. $\chi_f(X)$). Wir bezeichnen die algebraische Multiplizität von λ mit

$$\mathrm{am}_A(\lambda) \quad (\text{bzw. } \mathrm{am}_f(\lambda)).$$

(2) Die **geometrische Multiplizität (Vielfachheit)** eines Eigenwerts λ der Matrix A (bzw. des Endomorphismus f) ist die Dimension

$$\dim(V_\lambda)$$

des zugehörigen Eigenraums. Wir bezeichnen die geometrische Multiplizität von λ mit

$$\mathrm{gm}_A(\lambda) \quad (\text{bzw. } \mathrm{gm}_f(\lambda)).$$

Beispiel 17.50. Sei A eine obere Dreiecksmatrix

$$A = \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \alpha_n \end{pmatrix}.$$

Dann ist

$$\chi_A(X) = \det \begin{pmatrix} X-\alpha_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & X-\alpha_n \end{pmatrix} = \prod_{i=1}^{n} (X - \alpha_i)$$

und die Menge der Eigenwerte ist $\{\alpha_1, \ldots, \alpha_n\}$. Einen Eigenvektor zum Eigenwert α_1 kann man sehen: das ist e_1. Wenn die $*$-Einträge von A alle verschieden von 0 sind, dann sind für $i > 1$ die Vektoren e_i keine Eigenvektoren zum Eigenwert α_i. Trotzdem wissen wir bereits, daß die entsprechenden Eigenräume $V_{\alpha_i} \neq 0$ sind.

Beispiel 17.51. Sei $\varphi \in \mathbb{R}$ und

$$D(\varphi) = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}).$$

Die durch Matrixmultiplikation mit $D(\varphi)$ vermittelte lineare Abbildung nennen wir D_φ. Dies ist die Drehung um den Ursprung $0 \in \mathbb{R}^2$ im mathematisch positiven Drehsinn um den Winkel φ. Dazu beschreiben wir einen Vektor $x \in \mathbb{R}^2$ durch **Polarkoordinaten** als

$$x = r \begin{pmatrix} \cos(\psi) \\ \sin(\psi) \end{pmatrix}$$

und rechnen dann

$$D_\varphi(x) = D(\varphi) \cdot r \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} = r \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} \cos(\psi) \\ \sin(\psi) \end{pmatrix}$$
$$= r \begin{pmatrix} \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) \\ \sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi) \end{pmatrix} = r \begin{pmatrix} \cos(\varphi+\psi) \\ \sin(\varphi+\psi) \end{pmatrix}.$$

Hier haben wir die Additionstheoreme für Sinus und Cosinus verwendet. Dafür verweisen wir auf Analysis.

Wir berechnen nun die Eigenwerte von $D(\varphi)$. Das charakteristische Polynom ist

$$\chi_{D(\varphi)}(X) = (X - \cos(\varphi))(X - \cos(\varphi)) - \sin(\varphi)(-\sin(\varphi)) = X^2 - 2\cos(\varphi)X + 1.$$

(Dies braucht die Formel $\cos^2(\varphi) + \sin^2(\varphi) = 1$, wieder Analysis.) Die Nullstellen sind (in $\mathbb{C}$)

$$\lambda_{1/2} = \cos(\varphi) \pm i\sin(\varphi).$$

Also hat $D(\varphi)$ nur dann Eigenwerte, wenn $\varphi = k\pi$ mit $k \in \mathbb{Z}$ (Winkel im Bogenmaß!), und zwar $\lambda = 1$, wenn k gerade ist, oder $\lambda = -1$, wenn k ungerade ist. Die entsprechenden Abbildungen D_φ sind die Identität als Drehung um $\varphi = 0$ und die Punktspiegelung als Drehung $D_\varphi = -\operatorname{id}$.

Eine Drehung mit $\varphi \notin \pi \cdot \mathbb{Z}$ hat keine Eigenvektoren! Das erwarten wir auch geometrisch.

Nun betrachten wir $D(\varphi)$ durch $\mathbb{R} \subseteq \mathbb{C}$ als Matrix in $\mathrm{M}_2(\mathbb{C})$ mit komplexen Einträgen. In diesem Fall hat das charakteristische Polynom stets Lösungen, und wenn $\varphi \notin \pi \cdot \mathbb{Z}$ sogar zwei verschiedene. Als lineare Abbildung

$$D_\varphi : \mathbb{C}^2 \to \mathbb{C}^2$$

gibt es also stets Eigenwerte und entsprechend Eigenvektoren.

18. Invariante Unterräume

18.1. **Invariante Unterräume.** Eigenräume sind Spezialfälle von invarianten Unterräumen.

Definition 18.1. Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V. Ein **f-invarianter Unterraum** von V ist ein Unterraum $U \subseteq V$, so daß

$$f(U) \subseteq U.$$

Statt f-invariant verwendet man auch synonym die Bezeichnung **f-stabil**.

Bemerkung 18.2. Wenn $U \subseteq V$ ein f-invarianter Unterraum ist, dann bezeichnen wir die durch Einschränkung auf U in Definitionsbereich und(!) Wertebereich entstehende Abbildung ebenfalls mit

$$f|_U : U \to U.$$

Es ergibt sich dann ein kommutatives Diagramm

$$\begin{array}{ccc} U & \longrightarrow & V \\ {\scriptstyle f|_U}\downarrow & & \downarrow{\scriptstyle f} \\ U & \longrightarrow & V \end{array}$$

Beispiel 18.3. (1) Ein Eigenraum zu $f : V \to V$ ist ein f-invarianter Unterraum.
(2) Ein f-invarianter Unterraum $U \subseteq V$ mit $\dim(U) = 1$ ist nichts anderes als der von einem Eigenvektor $v \in V$ aufgespannte Unterraum $\langle v \rangle_K = K \cdot v$.

Ein invarianter Unterraum bedeutet, daß wir in einer der Situation angepaßten Basis eine obere Blockdreiecksmatrix haben:

Satz 18.4. *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums mit $\dim(V) = n$. Sei $U \subseteq V$ ein Unterraum der Dimension r. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V, so daß $U = \langle b_1, \ldots, b_r \rangle_K$. Dann sind äquivalent:*

(a) U ist ein f-invarianter Unterraum von V.

(b) Die Darstellungsmatrix von f bezüglich $\mathscr{B}$ ist eine obere Blockdreiecksmatrix:

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

mit $A \in \mathrm{M}_r(K)$, $B \in \mathrm{M}_{r\times(n-r)}(K)$ und $D \in \mathrm{M}_{n-r}(K)$.

In diesem Fall ist A die Darstellungsmatrix zu $f|_U$ bezüglich der Basis $(b_1, \ldots, b_r)$.

Beweis. (a) $\Longrightarrow$ (b): Für $j = 1, \ldots, r$ ist $f(b_j) \in f(U) \subseteq U$. Damit ist $f(b_j)$ eine Linearkombination der $b_1, \ldots, b_r$. Also hat die j-te Spalte der Darstellungsmatrix, das ist $f(b_j)$ in Koordinaten bezüglich $\mathscr{B}$, Einträge $\neq 0$ höchstens für die Zeilen $i \leq r$. Dies zeigt bereits die 0 im linken unteren Block der Darstellungsmatrix.

(b) $\Longrightarrow$ (a): Für alle $j = 1, \ldots, r$ gilt $f(b_j)$ ist Linearkombination der $b_1, \ldots, b_r$, folglich $f(b_j) \in U$. Da U von den $b_1, \ldots, b_r$ aufgespannt wird, folgt $f(U) \subseteq U$. Dieser Teil des Beweises zeigt bereits, daß A die Darstellungsmatrix zu $f|_U$ bezüglich der Basis $(b_1, \ldots, b_r)$ ist. □

Bemerkung 18.5. Ein invarianter Unterraum hat nicht notwendigerweise ein invariantes Komplement. Nehmen wir $V = K^2$ und $f : K^2 \to K^2$ die Multiplikation mit

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

dann ist $U = \langle e_1 \rangle_K = K \cdot e_1$ ein invarianter Unterraum, aber es gibt kein invariantes Komplement. Ein solches wäre eindimensional und für ein $x \in K$ von einem Vektor

$$v = e_2 + xe_1$$

aufgespannt. Aber

$$f(v) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} x+1 \\ 1 \end{pmatrix} \notin \langle v \rangle_K.$$

Daher ist keines der möglichen Komplemente $\langle v \rangle_K$ invariant unter f.

Alternativ kann man argumentieren, daß ein invariantes Komplement als invarianter eindimensionaler Unterraum von einem Eigenvektor aufgespannt wäre. Das charakteristische Polynom ist

$$\chi_A(X) = (X-1)^2,$$

und somit 1 der einzige Eigenwert. Der zugehörige Eigenraum ist

$$\ker(1 \cdot \mathbf{1} - A) = \ker(\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}) = K \cdot e_1.$$

Alle Eigenvektoren liegen bereits in U: es gibt kein invariantes Komplement.

18.2. **Diagonalisierbarkeit II.**

Satz 18.6. *Sei $f : V \to V$ ein diagonalisierbarer Endomorphismus eines endlichdimensionalen K-Vektorraums. Sei $U \subseteq V$ ein f-invarianter Unterraum. Dann ist $f|_U : U \to U$ ebenfalls diagonalisierbar.*

Beweis. Sei V_λ der Eigenraum zum Eigenwert λ von f auf V. Dann ist $U_\lambda = U \cap V_\lambda$ der Eigenraum zum Eigenwert λ von $f|_U$. Wir müssen nach Korollar 12.47 nur zeigen, daß U von allen seinen Eigenräumen U_λ zusammen erzeugt wird.

Wir numerieren dazu zunächst die Einwerte von f als $\lambda_1, \dots, \lambda_s$. Sei $u \in U$ beliebig. Weil f diagonalisierbar ist, gibt es in $V = \bigoplus_{i=1}^{s} V_{\lambda_i}$ eine Zerlegung

$$u = \sum_{i=1}^{s} v_i$$

mit $v_i \in V_{\lambda_i}$. Wenn wir zeigen, daß $v_i \in U$ gilt, so sind wir fertig. Wir ignorieren Eigenwerte λ_i mit $v_i = 0$, und sortieren so um, daß $v_1, \dots, v_t$ die Summanden ungleich 0 sind. Diese Vektoren sind nun linear unabhängig aufgrund der Eigenraumzerlegung von f.

Wir setzen $u_0 = u$ und für alle $m \geq 1$

$$u_m := f^m(u) = \sum_{i=1}^{t} f^m(v_i) = \sum_{i=1}^{t} \lambda_i^m v_i.$$

Weil U ein f-invarianter Unterraum ist, finden wir $u_m = f^m(u) \in U$ für alle $m \geq 0$. Außerdem sind die $u_m \in W := \langle v_1, \dots, v_t \rangle_K$. Um die lineare Hülle der $u_0, \dots, u_{t-1}$ zu bestimmen, drücken wir diese in W durch die Basis $v_1, \dots, v_t$ aus und suchen demnach den Spaltenraum der Vandermonde–Matrix

$$A = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{(t-1)} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{(t-1)} \\ \vdots & \vdots & & & \vdots \\ 1 & \lambda_t & \lambda_t^2 & \dots & \lambda_t^{(t-1)} \end{pmatrix}$$

Da die λ_i paarweise verschieden sind, hat die Vandermonde–Matrix eine Determinante

$$\det(A) = \prod_{1 \leq i < j \leq t} (\lambda_j - \lambda_i) \neq 0.$$

Daher ist der Spaltenraum gleich W. Es folgt für alle $1 \leq i \leq t$

$$v_i \in W = \langle u_0, \dots, u_{t-1} \rangle_K \subseteq U. \qquad \square$$

Bemerkung 18.7. Die im Beweis von Satz 18.6 auftretende Vandermonde-Matrix ist die Abbildungsmatrix der Auswertungsabbildung für $S = \{\lambda_1, \dots, \lambda_t\}$

$$\mathrm{Pol}_{K,\leq t-1} \to \mathrm{Abb}(S, K)$$

bezüglich der Monombasis $(1, X, \dots, X^{t-1})$ der Polynome und bezüglich der charakteristischen Funktionen der Elemente von S für den Vektorraum der Abbildungen. Die Vandermonde–Matrix ist invertierbar, weil die angegebene lineare Abbildung nach Satz 17.30 ein Vektorraumisomorphismus ist.

Satz 18.8. *Sei K ein Körper und seien $A, B \in \mathrm{M}_n(K)$. Dann sind äquivalent:*

(a) Die Matrizen A und B sind simultan diagonalisierbar: Es gibt $S \in \mathrm{GL}_n(K)$ und Diagonalmatrizen D und E mit

$$A = SDS^{-1} \qquad \text{und} \qquad B = SES^{-1}.$$

(b) Es gibt eine Basis von K^n, deren Vektoren gleichzeitig Eigenvektoren von A und B sind.

(c) A und B sind diagonalisierbar und

$$AB = BA.$$

Beweis. (b) $\Longrightarrow$ (a): Sei S eine Matrix, deren Spalten eine Basis von Eigenvektoren zu gleichzeitig A und B sind. Dann sind $D = S^{-1}AS$ und $E = S^{-1}BS$ Diagonalmatrizen, und demnach sind $A = SDS^{-1}$ und $B = SES^{-1}$ simultan diagonalisiert, siehe Proposition 12.28(b).

(a) $\Longrightarrow$ (c): Zu zeigen ist nur, daß A und B miteinander kommutieren. Diagonalmatrizen kommutieren stets, also $DE = ED$, und damit folgt

$$\begin{aligned} AB &= (SDS^{-1})(SES^{-1}) = SD(S^{-1}S)ES^{-1} = S(DE)S^{-1} \\ &= S(ED)S^{-1} = SE(S^{-1}S)DS^{-1} = (SES^{-1})(SDS^{-1}) = BA. \end{aligned}$$

(c) $\Longrightarrow$ (b): Sei $V_{\lambda,A} \subseteq K^n$ der Eigenraum von A zum Eigenwert λ. Dies ist ein L_B-stabiler Unterraum, denn für $v \in V_{\lambda,A}$ und $w = Bv$ gilt

$$Aw = A(Bv) = (AB)v = (BA)v = B(Av) = B(\lambda v) = \lambda(Bv) = \lambda w.$$

Weil B diagonalisierbar ist, ist nach Satz 18.6 auch die Einschränkung von L_B auf $V_{\lambda,A}$ diagonalisierbar. Damit hat der Eigenraum $V_{\lambda,A}$ eine Basis von Eigenvektoren von der Einschränkung von L_A, also von Eigenvektoren von L_B und damit von B. Diese Vektoren sind gleichzeitig Eigenvektoren zu A und B, und eine Basis von $V_{\lambda,A}$.

Weil A diagonalisierbar ist, folgt aus der Eigenraumzerlegung, daß Basen von $V_{\lambda,A}$ für alle λ zusammen eine Basis von K^n liefern. □

Satz 18.9. *Die geometrische Multiplizität eines Eigenwerts ist stets kleiner oder gleich der algebraischen Multiplizität:*

$$\mathrm{gm}(\lambda) \leq \mathrm{am}(\lambda).$$

Beweis. Es reicht, dies für einen Endomorphismus $f : V \to V$ mit $\dim(V) = n$ zu beweisen. Sei λ ein Eigenwert von f und V_λ der zugehörige Eigenraum. Sei $m = \dim(V_\lambda)$ die geometrische Multiplizität.

Als Eigenraum ist V_λ ein f-invarianter Unterraum. Zur Berechnung des charakteristischen Polynoms benutzen wir eine Basis wie in Satz 18.4 und erhalten

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

wobei A die Darstellungsmatrix von $f|_{V_\lambda}$ ist. Weil $f|_{V_\lambda} = \lambda \cdot \mathrm{id}_{V_\lambda}$, ergibt sich $A = \lambda \cdot \mathbf{1}$ und aus Korollar 13.7 mit Beispiel 17.50 dann

$$\chi_f(X) = \chi_A(X) \cdot \chi_D(X) = (X - \lambda)^{\dim(V_\lambda)} \cdot \chi_D(X).$$

Aus dieser Produktzerlegung lesen wir die Behauptung ab. □

Theorem 18.10 (Diagonalisierbarkeit)**.** *Sei $A \in \mathrm{M}_n(K)$, und seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A. Dann sind äquivalent:*

(a) *A ist diagonalisierbar.*
(b) *K^n hat eine Basis aus Eigenvektoren von A.*
(c) *$V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r} = K^n$.*
(d) *$\sum_{i=1}^r \dim(V_{\lambda_i}) = n$.*
(e) *Für jeden Eigenwert $\lambda \in K$ von A ist*

$$\mathrm{am}_A(\lambda) = \mathrm{gm}_A(\lambda)$$

und $\chi_A(X)$ zerfällt vollständig in Linearfaktoren.

Beweis. Die Äquivalenz von (a), (b), (c) und (d) ist die Aussage von Satz 12.45.

(e) $\Longrightarrow$ (d): Weil $\chi_A(X)$ vollständig in Linearfaktoren zerfällt und Nullstellen von $\chi_A(X)$ dasselbe wie Eigenwerte von A sind, folgt (d) aus der Rechnung

$$n = \deg(\chi_A(X)) = \sum_{\lambda \text{ Eigenwert}} \mathrm{am}_A(\lambda) = \sum_{\lambda \text{ Eigenwert}} \mathrm{gm}_A(\lambda) = \sum_{\lambda \text{ Eigenwert}} \dim(V_\lambda).$$

(b) $\Longrightarrow$ (e): Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis aus Eigenvektoren. Sei λ_i für $i = 1, \ldots, n$ der Eigenwert des Eigenvektors b_i, also $Ab_i = \lambda_i b_i$. Dazu gehört ein Basiswechsel $S \in \mathrm{GL}_n(K)$ mit

$$SAS^{-1} = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Damit gilt

$$\chi_A(X) = \chi_D(X) = \prod_{i=1}^{n}(X - \lambda_i),$$

und $\chi_A(X)$ zerfällt vollständig in Linearfaktoren. Außerdem ist für jeden Eigenwert λ

$$\mathrm{am}_A(\lambda) = |\{i \ ; \ \lambda_i = \lambda\}| = \dim\left(\langle b_i \ ; \ i \text{ mit } \lambda_i = \lambda\rangle_K\right) \leq \dim(V_\lambda) = \mathrm{gm}_A(\lambda) \leq \mathrm{am}_A(\lambda),$$

letzteres nach Satz 18.9. Damit gilt $\mathrm{am}_A(\lambda) = \mathrm{gm}_A(\lambda)$ wie behauptet. □

Korollar 18.11. *Sei $A = SDS^{-1}$ mit $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ und $S \in \mathrm{GL}_n(K)$. Dann ist für alle $\lambda \in K$ der Eigenraum V_λ von A zum Eigenwert λ genau*

$$V_\lambda = \langle Se_j \ ; \ j \text{ mit } \lambda_j = \lambda\rangle_K$$

und diese Spalten Se_j von S bilden eine Basis von V_λ.

Beweis. Das folgt aus dem Beweisteil (b) $\Longrightarrow$ (e) des Beweises von Theorem 18.10. □

Korollar 18.12 (Diagonalisierbarkeit von Endomorphismen)**.** *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums V, und seien $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A. Dann sind äquivalent:*

(a) *f ist diagonalisierbar.*
(b) *V hat eine Basis aus Eigenvektoren von f.*
(c) *$V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r} = V$.*
(d) *$\sum_{i=1}^{r} \dim(V_{\lambda_i}) = \dim(V)$.*
(e) *Für jeden Eigenwert $\lambda \in K$ von f ist*

$$\mathrm{am}_f(\lambda) = \mathrm{gm}_f(\lambda)$$

und $\chi_f(X)$ zerfällt vollständig in Linearfaktoren.

Beweis. Das ist die Übersetzung von Theorem 18.10 für Endomorphismen. □

18.3. **Das Minimalpolynom.** Die m-te Potenz einer Matrix A ist die Auswertung des Polynoms X^m in der Matrix A. Wir definieren nun allgemeiner die Auswertung eines Polynoms.

Definition 18.13. Die Auswertung eines Polynoms

$$P(X) = a_0 + a_1 X + \ldots + a_d X^d \in K[X]$$

in der Matrix $A \in \mathrm{M}_n(K)$ ist die Matrix

$$P(A) = a_0 \cdot \mathbf{1} + a_1 \cdot A + \ldots + a_d \cdot A^d \in \mathrm{M}_n(K).$$

Beispiel 18.14. Wenn $D = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ eine Diagonalmatrix ist, dann ist für $P(X) = \sum_{i=0}^{d} a_i X^i$

$$P(D) = \sum_{i=0}^{d} a_i \cdot D^i = \sum_{i=0}^{d} a_i \cdot \mathrm{diag}((\alpha_1)^i, \ldots, (\alpha_n)^i) = \mathrm{diag}(P(\alpha_1), \ldots, P(\alpha_n)).$$

Bemerkung 18.15. Matrixmultiplikation macht aus $R = \mathrm{M}_n(K)$ einen nichtkommutativen[7] Ring, den **Matrizenring** der $n \times n$-Matrizen mit Koeffizienten aus K. Ein **nichtkommutativer Ring**

[7] Nur falls $n \geq 2$, der Fall $n = 1$ ist isomorph zu K und daher kommutativ.

ist eine Menge mit Addition und Multiplikation, bei der alle Axiome eines Rings erfüllt sind mit Ausnahme der Kommutativität der Multiplikation. Genauer müßte es daher heißen: ein nicht notwendigerweise nichtkommutativer Ring, denn kommutative Ringe erfüllen ebenfalls diese Bedingungen. Hier rächt es sich, daß wir für Ringe die Kommutativität allgemein gefordert haben.

Genauer ist $\mathrm{M}_n(K)$ sogar eine K-Algebra. Eine **K-Algebra** ist ein K-Vektorraum R mit einer Multiplikation $R \times R \to R$, die in beiden Argumenten K-linear ist (bilinear), und für die R zusammen mit der Vektorraumaddition ein Ring mit Eins wird. Die Eins bezeichnen wir wie üblich mit $1 \in R$ wird. Mit Ausnahme des Nullrings, der K-Algebrastruktur auf dem Nullvektorraum, definiert die Abbildung

$$K \to R, \quad \lambda \mapsto \lambda \cdot 1$$

eine Einbettung des Körpers K in die K-Algebra R, mittels derer wir K als einen Teilring von R auffassen. Auch wenn R nichtkommutativ sein könnte, so gilt doch wegen der Bilinearität der Multiplikation für alle $x \in R$ und $a \in K$

$$ax = xa.$$

Die K-Vektorraumstruktur von R bekommt man allein aus der Multiplikation von R durch die Multiplikation mit Elementen von $K \subseteq R$. Die Struktur von K-Algebren kann man mit Mitteln der linearen Algebra studieren.

Als Beispiel für eine $\mathbb{R}$-Algebra haben wir den Körper der komplexen Zahlen $\mathbb{C}$ kennen gelernt.

Bemerkung 18.16. Genauso wie $\mathrm{M}_n(K)$ ist für einen K-Vektorraum V

$$\mathrm{End}_K(V) = \mathrm{Hom}_K(V, V)$$

eine K-Algebra. Die Multiplikation ist Komposition und diese ist bilinear (Distributivgesetz!) für die bekannte K-Vektorraumstruktur auf $\mathrm{End}_K(V) = \mathrm{Hom}_K(V, V)$. Für $V \neq 0$ ist $\mathrm{End}_K(V)$ nicht der Nullring und die Einbettung $K \to \mathrm{End}_K(V)$ kommt von der Identifikation von $a \in K$ mit der Streckung $a \cdot \mathrm{id}_V$.

Bemerkung 18.17. Allgemeiner kann man ein Polynom

$$P(X) = a_0 + a_1 X + \ldots + a_d X^d \in K[X]$$

in einem Element $x \in R$ einer K-Algebra R durch

$$P(x) = a_0 \cdot 1 + a_1 x + \ldots + a_d x^d$$

auswerten. Wir können demnach Polynome auch in Matrizen oder in Endomorphismen auswerten.

Proposition 18.18. *Die Auswertung von Polynomen in einer Matrix $A \in \mathrm{M}_n(K)$ ist ein Ringhomomorphismus*

$$K[X] \to \mathrm{M}_n(K),$$

das heißt, für alle $P, Q \in K[X]$ gilt

$$(P + Q)(A) = P(A) + Q(A),$$
$$(P \cdot Q)(A) = P(A) \cdot Q(A).$$

Beweis. Das folgt aus denselben Rechnungen wie beim Einsetzen von $x \in K$. □

Definition 18.19. Der **Leitkoeffizient** (oder **führender Koeffizient**) eines Polynoms $0 \neq P \in K[X]$ vom Grad $d = \deg(P)$ ist der Koeffizient vor X^d. Ein Polynom heißt **normiert** (oder **normiertes** Polynom), wenn der Leitkoeffizient 1 ist.

Satz 18.20. *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gibt es ein eindeutiges normiertes Polynom $m_A(X)$ mit*

(i) $m_A(A) = 0$, *und*

(ii) *für alle $P(X) \in K[X]$ mit $P(A) = 0$ gibt es ein $Q(X) \in K[X]$ mit $P(X) = Q(X) \cdot m_A(X)$.*

Beweis. Schritt 1: Wir müssen zuerst einsehen, daß es ein nichttrivales Polynom P gibt mit $P(A) = 0$. Dazu betrachten wir die $(n^2 + 1)$-vielen Matrizen

$$\mathbf{1} = A^0, A, A^2, A^3, \ldots, A^{n^2} \in \mathrm{M}_n(K)$$

als Vektoren des Vektorraums $\mathrm{M}_n(K)$. Weil es mehr Vektoren als $n^2 = \dim \mathrm{M}_n(K)$ sind, folgt aus Satz 5.31, daß diese Matrixpotenzen linear abhängig sind. Zu einer nichttrivialen Linearkombination $\sum_{i=0}^{n^2} a_i A^i = 0$ gehört das Polynom

$$P(X) = \sum_{i=0}^{n^2} a_i X^i \neq 0$$

mit $P(A) = 0$.

Schritt 2: Unter allen Polynomen $P \neq 0$ mit $P(A) = 0$ wählen wir nun eines mit minimalem Grad. Schritt 1 garantiert, daß es ein solches extremales Polynom gibt. Indem wir mit dem Inversen des Leitkoeffizienten multiplizieren (es bleibt die Bedingung des minimalen Grades und von $P(A) = 0$ erhalten), dürfen wir annehmen, daß wir ein normiertes Polynom gewählt haben. Dieses Polynom bezeichnen wir mit

$$m_A(X)$$

und zeigen nun, daß es Bedingung (ii) erfüllt. Bedingung (i) gilt per Kojnstruktion.

Schritt 3: Sei $P(X) \in K[X]$ mit $P(A) = 0$. Wir dividieren P durch m_A mit Rest, und erhalten

$$P(X) = Q(X)m_A(X) + R(X)$$

mit $\deg(R) < \deg(m_A)$. Da Auswerten in A ein Ringhomomorphismus ist, folgt

$$R(A) = (P - Qm_A)(A) = P(A) - Q(A) \cdot m_A(A) = 0 - Q(A) \cdot 0 = 0.$$

Aufgrund der Minimalität von $\deg(m_A)$ bleibt nur $R = 0$. Folglich ist $P = Qm_A$.

Schritt 4: Wir müssen nun noch die Eindeutigkeit von m_A zeigen. Angenommen $P(X)$ wäre ein weiteres Polynom, das (i) und (ii) erfüllt. Dann gilt $P(A) = m_A(A) = 0$, weshalb (ii) für m_A bezüglich P Anwendung findet: es gibt $Q(X) \in K[X]$ mit

$$m_A(X) = Q(X)P(X).$$

Damit ist

$$\deg(m_A) = \deg(Q) + \deg(P) \geq \deg(P) \geq \deg(m_A),$$

letzteres aufgrund der extremalen Wahl von m_A als Polynom minimalen Grades. Folglich haben P und m_A den gleichen Grad und sind beide normiert. Damit ist

$$D = m_A - P$$

ein Polynom mit $\deg(D) < \deg(m_A)$ und trotzdem $D(A) = m_A(A) - P(A) = 0$. Folglich ist $D = 0$ und damit $m_A = P$. Dies zeigt die Eindeutigkeit. □

Das in Satz 18.20 eindeutig bestimmte Polynom bekommt einen Namen.

Definition 18.21. Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Das eindeutiges normiertes Polynom $m_A(X)$ aus Satz 18.20 wird das **Minimalpolynom** von A genannt.

Eine analoge Definition gibt es für Endomorphismen:

Definition 18.22. Das **Minimalpolynom** eines Endomorphismus $f : V \to V$ ist das eindeutige normierte Polynom $m_f(X)$ mit
(i) $m_f(f) = 0$, und
(ii) für alle $P(X) \in K[X]$ mit $P(f) = 0$ gibt es ein $Q(X) \in K[X]$ mit $P(X) = Q(X) \cdot m_f(X)$.

Bemerkung 18.23. Es folgt sofort aus Satz 8.35 und Satz 10.4, daß für die Darstellungsmatrix A eines Endomorphismus f gilt:

$$m_A(X) = m_f(X).$$

Lemma 18.24. *Seien $P \in K[X]$, $A \in \mathrm{M}_n(K)$ und $S \in \mathrm{GL}_n(K)$. Dann gilt*

$$P(SAS^{-1}) = SP(A)S^{-1}.$$

Beweis. Beide Seiten sind linear in P. Es bleibt also nur der Fall $P(X) = X^m$. Das ist Proposition 12.33. □

Lemma 18.25. *Die Minimalpolynome ähnlicher Matrizen sind gleich.*

Beweis. Das folgt sofort aus der Definition und

$$P(A) = 0 \iff SP(A)S^{-1} = 0 \iff P(SAS^{-1}) = 0$$

nach Lemma 18.24. □

18.4. **Der Satz von Cayley–Hamilton.** Wir bringen nun das charakteristische Polynom mit dem Minimalpolynom zusammen. Als Vorarbeit beweisen wir das folgende für sich interessante Resultat.

Proposition 18.26. *Jedes normierte Polynom ist charakteristisches Polynom einer Matrix: sei $P(X) \in K[X]$ ein normiertes Polynom vom Grad n. Dann gibt es eine Matrix $A \in \mathrm{M}_n(K)$ mit*

$$\chi_A(X) = P(X).$$

Beweis. Im Fall $n = 0$ folgt die Proposition aus der Konvention, daß die Matrix zur Identität des Nullvektorraums das charakteristische Polynom konstant 1 hat.

Sei nun $n \geq 1$ und $P(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$ für Koeffizienten $a_i \in K$, $i = 0, \ldots, n-1$. Wir behaupten, daß die Matrix

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & \ddots & \ddots & \vdots & -a_1 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \in \mathrm{M}_n(K).$$

das charakteristische Polynom $\chi_A(X) = P(X)$ hat. Die nötige Determinante berechnen wir per Induktion nach $n \geq 1$ (der Induktionsanfang bei $n = 1$ ist trivial) und Laplace-Entwicklung nach

der ersten Zeile:

$$\chi_A(X) = \det(X \cdot \mathbf{1} - A) = \det \begin{pmatrix} X & 0 & \cdots & 0 & a_0 \\ -1 & \ddots & \ddots & \vdots & a_1 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & X & a_{n-2} \\ 0 & \cdots & 0 & -1 & X + a_{n-1} \end{pmatrix}$$

$$= X \cdot \det \begin{pmatrix} X & & & a_1 \\ -1 & \ddots & & \vdots \\ & \ddots & X & a_{n-2} \\ & & -1 & X + a_{n-1} \end{pmatrix} + (-1)^{n+1} a_0 \cdot \det \begin{pmatrix} -1 & X & & \\ & \ddots & \ddots & \\ & & \ddots & X \\ & & & -1 \end{pmatrix}$$

$$= X \cdot (X^{n-1} + a_{n-1}X^{n-2} + \ldots + a_2 X + a_1) + (-1)^{n+1} \cdot (-1)^{n-1} a_0$$

$$= X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0 = P(X). \qquad \square$$

Theorem 18.27 (Cayley–Hamilton). *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gilt $\chi_A(A) = 0$.*

Beweis. Schritt 1: Das Theorem ist äquivalent zu: für alle $x \in K^n$ gilt

$$\chi_A(A)x = 0,$$

weil die Matrix $\chi_A(A)$ genau dann 0 ist, wenn die zugehörige Matrixmultiplikationsabbildung

$$K^n \to K^n, \qquad x \mapsto \chi_A(A)x$$

die Nullabbildung ist. Der Fall $x = 0$ ist klar, daher nehmen wir nun $x \neq 0$ an.

Schritt 2: Wir fixieren nun ein $x \in K^n$, $x \neq 0$, und betrachten den Unterraum

$$U = \langle x, Ax, A^2x, \ldots, A^m x, \ldots \rangle_K \subseteq K^n.$$

Es ist U ein A-invarianter Unterraum, weil für beliebige $y_i \in K$ gilt:

$$A\Big(\sum_{i=0}^{m} y_i A^i x\Big) = \sum_{i=0}^{m} y_i A^{i+1} x \in U.$$

Schritt 3: Wir bestimmen nun eine geeignete Basis von U. Sei $d \geq 1$ minimal mit

$$A^d x \in \langle x, Ax, A^2x, \ldots, A^{d-1}x \rangle_K.$$

Dann gibt es $a_0, \ldots, a_{d-1} \in K$ mit

$$A^d x = a_0 x + a_1 Ax + a_2 A^2 x + \ldots + a_{d-1} A^{d-1} x \tag{18.1}$$

und

$$(x, Ax, \ldots, A^{d-1}x)$$

ist eine Basis von U:

- linear unabhängig: angenommen es gibt eine nichttriviale Linearkombination

  $$b_0 x + b_1 Ax + \ldots + b_{d-1} A^{d-1} x = 0.$$

  Sei $\delta = \max\{i \ ; \ b_i \neq 0\}$, was wohldefiniert ist, denn die Linearkombination ist als nichttrivial angenommen. Es gilt sogar $\delta \geq 1$, weil $x \neq 0$ ist. Man kann nun die Linearkombination nach $A^\delta x$ auflösen und erhält

  $$A^\delta x \in \langle x, Ax, A^2x, \ldots, A^{\delta-1}x \rangle_K$$

im Widerspruch zur Minimalität von d, weil $\delta \leq d-1$.

- Erzeugendensystem: Per Induktion nach m zeigen wir $A^m x \in \langle x, Ax, A^2x, \dots, A^{d-1}x\rangle_K$. Für $m < d$ ist nichts zu zeigen. Wenn $m \geq d$, dann multiplizieren wir (18.1) von links mit A^{m-d} und erhalten mit der Induktionsannahme
$$\begin{aligned} A^m x = A^{m-d} \cdot A^d x &= A^{m-d} \cdot \big(a_0 x + a_1 Ax + a_2 A^2 x + \dots + a_{d-1} A^{d-1} x\big) \\ &= a_0 A^{m-d} x + \dots + a_{d-1} A^{m-1} x, \end{aligned}$$
und das ist per Induktionsannahme in U enthalten.

Schritt 4: Mit dem Basisergänzungssatz erweitern wir $(x, Ax, \dots, A^{d-1}x)$ zu einer Basis
$$\mathscr{B} = (x, Ax, \dots, A^{d-1}x, b_{d+1}, \dots, b_n)$$
von K^n. Nach Satz 18.4 gibt es einen Basiswechsel, also ein $S \in \mathrm{GL}_n(K)$, und eine Blockmatrix
$$SAS^{-1} = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$
in der B die Darstellungsmatrix der Multiplikation mit A auf U bezüglich der Basis
$$(x, Ax, \dots, A^{d-1}x)$$
ist. Somit gilt:
$$B = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \ddots & \vdots & a_1 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & a_{d-1} \end{pmatrix} \in \mathrm{M}_d(K)$$
Das folgt sofort aus der definierenden Eigenschaft der Darstellungsmatrix.

Schritt 5: Nach Proposition 17.43 und Korollar 13.7 gilt (der Polynomring $K[X]$ ist kommutativ!)
$$\chi_A(X) = \chi_{SAS^{-1}}(X) = \chi_B(X) \cdot \chi_D(X) = \chi_D(X) \cdot \chi_B(X).$$
Im Beweis von Proposition 18.26 haben wir weiter
$$\chi_B(X) = X^d - a_{d-1}X^{d-1} - \dots - a_1 X - a_0$$
bestimmt. Daraus folgt nun mittels (18.1)
$$\begin{aligned} \chi_A(A)x &= \big(\chi_D(A) \cdot \chi_B(A)\big)x = \chi_D(A) \cdot (\chi_B(A)x) \\ &= \chi_D(A) \cdot \Big(\big(A^d - a_{d-1}A^{d-1} - \dots - a_1 A - a_0 \cdot \mathbf{1}\big)x\Big) \\ &= \chi_D(A) \cdot \big(A^d x - a_{d-1}A^{d-1}x - \dots - a_1 Ax - a_0 x\big) = \chi_D(A) \cdot 0 = 0. \end{aligned}$$
□

Bemerkung 18.28. (1) Der Satz von Cayley–Hamilton hat einen einzeiligen falschen(!) Beweis:
$$\chi_A(A) = \det(A \cdot \mathbf{1} - A) = \det(A - A) = \det(0) = 0.$$
Das Problem ist, daß das für die Variable substituierte A in $X \cdot \mathbf{1} - A$ nicht mit dem zweiten A identisch ist. Das substituierte A lebt nur in der Diagonale und die Substitution gelingt nur im Ring $\mathrm{M}_n(\mathrm{M}_n(K))$.

(2) Betrachten wir eine Variante mit der Spur anstelle der Determinante: für $A \in \mathrm{M}_n(K)$
$$\mathrm{tr}(X \cdot \mathbf{1} - A) = \mathrm{tr}(\mathbf{1}) \cdot X - \mathrm{tr}(A) = n \cdot X - \mathrm{tr}(A)$$
Ein analoger falscher(!) Beweis zeigt
$$\mathrm{tr}(X \cdot \mathbf{1} - A)(A) = \mathrm{tr}(A - A) = \mathrm{tr}(0) = 0.$$

Aber korrekt lautet die Auswertung bei A

$$nA - \operatorname{tr}(A) \cdot \mathbf{1}$$

und das ist genau für Vielfache der Einheitsmatrix gleich 0.

(3) Es gibt viele $A, B \in \mathrm{M}_n(K)$ mit

$$\det(B - A) = 0,$$

obwohl $\chi_A(B) \neq 0$. Wählen wir zum Beispiel

$$A = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \qquad \text{und} \qquad B = \begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix},$$

dann ist

$$B - A = \begin{pmatrix} 0 & x-a \\ 0 & y-b \end{pmatrix}$$

mit $\det(B - A) = 0$. Aber

$$\begin{aligned} \chi_A(B) &= \det(X \cdot \mathrm{id} - A)|_{X=B} = (X^2 - (1+b)X + b)|_{X=B} \\ &= \begin{pmatrix} 1 & x(1+y) \\ 0 & y^2 \end{pmatrix} - (1+b) \cdot \begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix} + b \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & x(y-b) \\ 0 & (y-b)(y-1) \end{pmatrix} \end{aligned}$$

und das ist nur 0, wenn $y = b$ oder ($x = 0$ und $y = 1$).

Bemerkung 18.29. Der Satz von Cayley–Hamilton ist eine polynomiale Identität für die Einträge einer $n \times n$-Matrix. Wenn K unendlich ist, kann man zeigen, daß eine solche Polynomidentität genau dann gilt, wenn sie für alle Matrizen A gilt, deren $\chi_A(X)$ in paarweise verschiedene Linearfaktoren zerfällt.[8] Für solche A gibt es n verschiedene Eigenwerte und somit ist A diagonalisierbar. Der Satz von Cayley–Hamilton ist leicht für diagonalisierbare Matrizen (Übungsaufgabe) und folgt daher für alle Matrizen aus dem allgemeinen Prinzip.

Definition 18.30. Ein **Teiler** eines Polynoms $P \in K[X]$ ist ein Polynom $Q \in K[X]$, so daß

$$P(X) = Q(X) \cdot R(X)$$

für ein Polynom $R \in K[X]$ gilt.

Satz 18.31. *Seien $A \in \mathrm{M}_n(K)$ (bzw. $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums).*

Dann ist das Minimalpolynom $m_A(X)$ (bzw. $m_f(X)$) ein Teiler des charakteristischen Polynoms $\chi_A(X)$ (bzw. $\chi_f(X)$).

Beweis. Das ist per Definition des Minimalpolynoms äquivalent zum Satz von Cayley–Hamilton $\chi_A(A) = 0$ (bzw. der analogen Version für Endomorphismen). □

Proposition 18.32. *Sei $x \in K^n$ ein Eigenvektor der Matrix $A \in \mathrm{M}_n(K)$ zum Eigenwert λ. Dann ist für jedes Polynom $P \in K[X]$*

$$P(A)x = P(\lambda) \cdot x.$$

Mit anderen Worten ist $P(\lambda)$ ein Eigenwert von $P(A)$, und x ist ein zugehöriger Eigenvektor.

[8] Man sagt, diese Menge von Matrizen ist Zariski-dicht.

Beweis. Die Aussage ist linear in P. Es reicht daher, sie für $P(X) = X^m$ für alle $m \geq 0$ zu beweisen. Wegen $Ax = \lambda x$ folgt per Induktion

$$A^m x = \lambda^m x.$$ □

Satz 18.33. *Seien $A \in \mathrm{M}_n(K)$ und $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums.*

Dann hat das Minimalpolynom $m_A(X)$ (bzw. $m_f(X)$) dieselben Nullstellen wie das charakteristische Polynom $\chi_A(X)$ (bzw. $\chi_f(X)$).

Beweis. Wir behandeln nur die Variante mit der quadratischen Matrix A. Weil das Minimalpolynom ein Teiler des charakteristischen Polynoms ist, sind Nullstellen von $m_A(X)$ auch Nullstellen von $\chi_A(X)$.

Für die umgekehrte Richtung betrachten wir eine Nullstelle $\chi_A(\lambda) = 0$. Dann ist λ ein Eigenwert von A. Sei $0 \neq x \in K^n$ ein Eigenvektor zum Eigenwert λ. Dann folgt

$$0 = m_A(A)x = m_A(\lambda)x$$

und somit $m_A(\lambda) = 0$. □

Beispiel 18.34. Nehmen wir an mit paarweise verschiedenen $\lambda_1, \ldots \lambda_s \in K$ und $e_1, \ldots, e_s \in \mathbb{N}$ sei für die Matrix $A \in \mathrm{M}_n(K)$

$$\chi_A(X) = \prod_{i=1}^{s} (X - \lambda_i)^{e_i}.$$

Dann gibt es $1 \leq f_i \leq e_i$ für $i = 1, \ldots, s$, so daß

$$m_A(X) = \prod_{i=1}^{s} (X - \lambda_i)^{f_i}.$$

Die f_i bestimmt man durch Zusatzinformation oder durch Ausprobieren der endlich vielen Möglichkeiten.

Beispiel 18.35. Eine Involution ist ein Automorphismus, der sein eigenes Inverses ist. Im Kontext von linearen Abbildungen bedeutet das für die Darstellungsmatrix $A \in \mathrm{M}_n(K)$ einer linearen Involution:

$$A^2 = \mathbf{1}.$$

Dies bedeutet, daß das Polynom $X^2 - 1$ in A ausgewertet 0 ergibt. Folglich ist $m_A(X)$ ein Teiler von

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

Dann gilt

$$m_A(X) = \begin{cases} 1 & \text{falls } A = \text{leere } 0 \times 0\text{-Matrix} \\ X - 1 & \text{falls } A = \mathbf{1} \\ X + 1 & \text{falls } A = -\mathbf{1} \\ X^2 - 1 & \text{sonst.} \end{cases}$$

Der letzte Fall tritt zum Beispiel bei einer Spiegelung auf.

Übungsaufgaben zu §18

Übungsaufgabe 18.1. Sei $A \in \mathrm{GL}_n(K)$. Drücken Sie das charakteristische Polynom von A^{-1} durch das charakteristische Polynom von A aus.

Übungsaufgabe 18.2. Beweisen Sie den Satz von Cayley–Hamilton mittels Beispiel 18.14 direkt für diagonalisierbare Matrizen.

Übungsaufgabe 18.3. Rechnen Sie den Satz von Cayley–Hamilton für eine beliebige 2×2-Matrix nach.

Übungsaufgabe 18.4. Finden Sie Beispiele von Matrizen $A, B \in \mathrm{M}_n(K)$ mit

$$\det(B - A) = 0,$$

obwohl

$$\chi_A(B) \neq 0.$$

19. Dévissage und Jordannormalform

19.1. **Faktorraum und Homomorphiesatz.** Seien K ein Körper und V ein K-Vektorraum. Zu einem Unterraum $U \subseteq V$ konstruieren wir einen Vektorraum V/U zusammen mit einer surjektiven linearen Abbildung

$$q : V \to V/U$$

mit Kern $\ker(q) = U$. Durch Anwenden von q werden so genau Vektoren in U „ignoriert“. Das ist oft eine hilfreiche Konstruktion.

Wir definieren ausgehend vom Unterraum $U \subseteq V$ eine Äquivalenzrelation auf V wie folgt. Zu $x, y \in V$ setzen wir

$$x \sim_U y :\iff x - y \in U.$$

Lemma 19.1. *Die Relation $\sim_U$ ist eine Äquivalenzrelation auf V.*

Beweis. Die Relation ist reflexiv, denn $x - x = 0 \in U$ zeigt $x \sim_U x$ für alle $x \in V$. Symmetrie folgt, denn aus $x \sim_U y$ folgt $x - y \in U$, also auch $y - x = -(x - y) \in U$ und damit $y \sim_U x$.

Zur Transitivität schließen wir aus $x \sim_U y$ und $y \sim_U z$ auf $x - y, y - z \in U$, also auch $x - z = (x - y) + (y - z) \in U$, und damit gilt $x \sim_U z$. □

Die Menge der Äquivalenzklassen für $\sim_U$ bezeichnen wir mit V/U, und die Quotientenabbildung mit

$$q : V \to V/U, \quad q(v) = [v].$$

Lemma 19.2. *Die Äquivalenzklasse bezüglich $\sim_U$ von $v \in V$ ist*

$$v + U = \{v + u \ ; \ u \in U\} \subseteq V$$

Beweis. Das ist klar aus der Definition. □

Definition 19.3 (Quotientenvektorraum)**.** Seien V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Auf V/U definieren wir wie folgt eine K-Vektorraumstruktur. Die **Addition** ist für alle $[v], [w] \in V/U$ gegeben durch

$$[v] + [w] := [v + w]$$

und die **Skalarmultiplikation** für alle $[v] in V/U$ und alle $\lambda \in K$ durch

$$\lambda \cdot [v] := [\lambda v].$$

Wir nennen V/U den **Quotientenvektorraum** (oder **Faktorraum**) von V nach U.

Proposition 19.4. *Der Quotient V/U ist mit den in Definition 19.3 angegebenen Addition und Skalarmultiplikation ein K-Vektorraum.*

Beweis. Wir müssen nachweisen, daß Addition und Multiplikation auf V/U wohldefiniert sind. Seien $a, a', b, b' \in V$ mit $[a] = [a']$ und $[b] = [b']$. Dann gibt es $x, y \in U$ mit $a - a' = x$ und $b - b' = y$.

- Addition: Dann gilt $[a + b] = [a' + b']$, weil

$$(a + b) - (a' + b') = (a - a') + (b - b') = x + y \in U.$$

- Skalarmultiplikation: Sei $\lambda \in K$. Dann gilt $[\lambda a] = [\lambda a']$, weil
$$\lambda a - \lambda a' = \lambda(a - a') = \lambda x \in U.$$

Da beide Verknüpfungen durch Repräsentanten im Vektorraum V berechnet werden, kann man die Vektorraumaxiome für V/U in V nachrechnen, wo sie gelten. □

Satz 19.5. *Die Quotientenabbildung $q : V \to V/U$ ist eine surjektive lineare Abbildung mit Kern $U = \ker(q)$.*

Beweis. Surjektivität von q ist klar per Definition. Seien $v, w \in V$ und $\lambda \in K$. Dann gilt
$$q(\lambda v + w) = [\lambda v + w] = [\lambda v] + [w] = \lambda[v] + [w] = \lambda q(v) + q(w),$$
also ist q linear. Der Kern von q bestimmt sich wie folgt:
$$x \in \ker(q) \iff q(x) = 0 \iff [x] = [0] \iff x - 0 \in U \iff x \in U.$$ □

Korollar 19.6 (Dimension des Faktorraums)**.** *Sei V ein endlichdimensionaler K-Vektorraum mit Unterraum $U \subseteq V$. Dann gilt*
$$\dim(V) = \dim(U) + \dim(V/U).$$

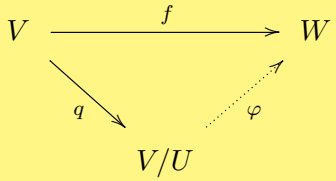
Beweis. Die Quotientenabbildung $q : V \to V/U$ ist surjektiv mit Kern U. Nach der Dimensionsformel für Kern und Bild, Satz 7.21, folgt
$$\dim(V/U) = \dim(\mathrm{im}(q)) = \dim(V) - \dim(\ker(q)) = \dim(V) - \dim(U).$$ □

Proposition 2.67 hat eine Entsprechung für die Quotientenabbildung von Vektorräumen.

Satz 19.7 (Quotienteneigenschaft)**.** *Sei V ein K-Vektorraum mit Unterraum $U \subseteq V$, und sei $f : V \to W$ eine lineare Abbildung. Dann sind äquivalent.*

(a) $U \subseteq \ker(f)$.

(b) *Es gibt eine lineare Abbildung $\varphi : V/U \to W$ mit $f = \varphi \circ q$, also ein kommutatives Diagramm*



Beweis. (b) $\Longrightarrow$ (a): Wenn f als $f = \varphi \circ q$ faktorisiert, dann gilt für alle $x \in U$, daß $f(x) = \varphi(q(x)) = \varphi(0) = 0$. Also ist $U \subseteq \ker(f)$.

(a) $\Longrightarrow$ (b): Wir definieren $\varphi : V/U \to W$ durch
$$\varphi([x]) = f(x).$$
Wenn $[x] = [y]$, dann ist $v = x - y \in U$ und
$$f(x) = f((x - y) + y) = f(x - y) + f(y) = 0 + f(y) = f(y),$$
weil $f(U) = 0$. Damit ist $\varphi : V/U \to W$ wohldefiniert. Diese Abbildung φ ist linear: zu $x, y \in V$ und $\lambda \in K$ gilt
$$\varphi(\lambda[x] + [y]) = \varphi([\lambda x + y]) = f(\lambda x + y) = \lambda f(x) + f(y) = \lambda\varphi([x]) + \varphi([y]).$$
Die Eigenschaft $f = \varphi \circ q$ ist offensichtlich. □

Satz 19.8 (Homomorphiesatz). *Sei $f : V \to W$ eine lineare Abbildung von K-Vektorräumen. Dann induziert f einen Isomorphismus*

$$\bar{f} : V/\ker(f) \xrightarrow{\sim} \operatorname{im}(f)$$
$$[x] \mapsto f(x).$$

Beweis. Nach Satz 19.7 angewandt auf $U = \ker(f)$ finden wir eine lineare Abbildung

$$\varphi : V/\ker(f) \to W$$

mit $\varphi([x]) = f(x)$ für alle $x \in V$. Dann gilt offensichtlich

$$\operatorname{im}(f) = \operatorname{im}(\varphi)$$

und wir können φ als eine lineare Abbildung $\bar{f} : V/U \to \operatorname{im}(f)$ auffassen. Diese Abbildung ist offensichtlich surjektiv, denn auch hier gilt $\bar{f}([x]) = f(x)$.

Um zu zeigen, daß $\bar{f}$ injektiv ist, müssen wir $\ker(\bar{f}) = 0$ zeigen. Sei

$$\bar{f}([x]) = 0.$$

Dann ist $f(x) = \bar{f}([x]) = 0$, also $x \in \ker(f)$. Damit ist aber $[x] = 0$. Damit ist $\bar{f}$ injektiv.

Als bijektive lineare Abbildung ist $\bar{f}$ ein Isomorphismus wie behauptet. □

Beispiel 19.9. Jedes normierte Polynom $0 \neq P \in K[X]$ ist Minimalpolynom. Dazu betrachten wir das Bild der Multiplikation mit P

$$(P) := \operatorname{im}(P\cdot : K[X] \to K[X]) = \{Q \cdot P \ ; \ Q \in K[X]\} \subseteq K[X]$$

als Unterraum und den zugehörigen Quotientenvektorraum

$$V = K[X]/(P).$$

Aus dem Satz über Division mit Rest, Satz 17.17, folgt, daß jede Restklasse einen eindeutigen Vertreter $R \in K[X]$ mit $\deg(R) < \deg(P) =: n$ hat. Daher ist

$$\bigoplus_{i=0}^{n-1} K \cdot X^i \xrightarrow{\sim} K[X]/(P)$$
$$R \mapsto R + (P)$$

ein Isomorphismus. Mit andern Worten können wir modulo (P) jedes Polynom eindeutig durch ein Polynom vom Grad $< \deg(P)$ darstellen.

Auf V betrachten wir nun den Effekt der Multiplikation mit X:

$$f : V \to V$$
$$Q + (P) \mapsto f(Q + (P)) = X \cdot Q + (P).$$

Dieser Endomorphismus hat Minimalpolynom $P(X)$. Zuerst ist für jedes Polynom $R(X)$

$$R(f) = \text{Multiplikation mit } R(X).$$

Damit ist $P(f) = 0$, denn für alle $Q + (P) \in V$ ist $P \cdot Q \in (P)$, also $P(f)(Q) = 0 \in V$. Das Minimalpolynom teilt also $P(X)$. Andererseits kann das Minimalpolynom nicht kleineren Grad haben, denn auf $1 + (P) \in V$ hat Multiplikation mit $R(X)$ bei $\deg(R) < \deg(P)$ das Bild

$$R(X) \cdot (1 + (P)) = R(X) + (P) \neq 0 \in V.$$

Proposition 19.10. *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums, und sei $U \subseteq V$ ein f-invarianter Unterraum, und sei* $\mathrm{pr} : V \to V/U$ *die Quotientenabbildung. Dann gibt es eine eindeutige lineare Abbildung*

$$\bar{f} : V/U \to V/U,$$

so daß für alle $v \in V$ gilt: $\bar{f}(\mathrm{pr}(v)) = \mathrm{pr}(f(v))$.

Beweis. Die Bedingung $\bar{f}(\mathrm{pr}(v)) = \mathrm{pr}(f(v))$ legt $\bar{f}$ eindeutig fest, weil jeder Vektor in V/U von der Form $\mathrm{pr}(v)$ für ein $v \in V$ ist. Zur Existenz betrachten wir

$$V \xrightarrow{f} V \xrightarrow{\mathrm{pr}} V/U.$$

Weil U ein f-invarianter Unterraum ist, gilt

$$\mathrm{pr}(f(U)) \subseteq \mathrm{pr}(U) = 0.$$

Nach der Quotienteneigenschaft von $\mathrm{pr} : V \to V/U$, Satz 19.7, folgt eine eindeutige Faktorisierung $\bar{f}$ wie im Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & V \\ {\scriptstyle \mathrm{pr}}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}} \\ V/U & \xrightarrow{\ \bar{f}\ } & V/U. \end{array}$$

□

Proposition 19.11. *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums der Dimension $\dim(V) = n$. Sei $U \subseteq V$ ein f-invarianter Unterraum der Dimension r. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V, so daß $U = \langle b_1, \ldots, b_r \rangle_K$. Der rechte untere Block $D \in \mathrm{M}_{n-r}(K)$ in der Darstellungsmatrix von f bezüglich $\mathscr{B}$*

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

siehe Satz 18.4, ist die Darstellungsmatrix zu

$$\bar{f} : V/U \to V/U$$

bezüglich der Basis $\bar{\mathscr{B}} = (\bar{b}_{r+1}, \ldots, \bar{b}_n)$ wobei $\bar{b}_i \in V/U$ das Bild von b_i unter der Quotientenabbildung $V \to V/U$ ist.

Beweis. Die Darstellungsmatrix zu $\bar{f}$ enthält $\bar{f}(\bar{b}_{r+j})$ für $j = 1, \ldots, n-r$ in Koordinaten bezüglich $\bar{\mathscr{B}}$. Diese berechnet man auf Urbildern b_{r+j} als $f(b_{r+j})$, in Koordinaten die $r+j$-te Spalte von $\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$. Wenn $B = (\beta_{ij})$ und $D = (d_{ij})$, dann gilt

$$f(b_{r+j}) = \sum_{i=1}^{r} \beta_{ij} b_i + \sum_{i=1}^{n-r} d_{ij} b_{r+i}.$$

Das anschließende Projizieren auf V/U vergißt die Koordinaten bezüglich $b_1, \ldots, b_r$ und behält die entsprechenden Spalten von D als Koordinatenvektor:

$$\bar{f}(b_{r+j}) = \sum_{i=1}^{n-r} d_{ij} \bar{b}_{r+i}.$$

□

Korollar 19.12. *Seien V ein endlichdimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Sei $U \subseteq V$ ein f-invarianter Unterraum und $\bar{f} : V/U \to V/U$ der von f induzierte Endomorphismus. Dann gilt*

$$\chi_f(X) = \chi_{f|_U}(X) \cdot \chi_{\bar{f}}(X).$$

Beweis. Das folgt sofort aus Satz 18.4, Proposition 19.11 und Korollar 13.7. □

19.2. **Fahnen und obere Dreiecksmatrizen.** Wir zeigen in diesem Abschnitt, welchen profunden Einfluß die arithmetischen Eigenschaften des charakteristischen Polynoms auf die mögliche Gestalt der Darstellungsmatrix haben.

Definition 19.13. Eine **Fahne** eines K-Vektorraums V ist eine aufsteigende Folge

$$W_0 \subseteq W_1 \subseteq W_2 \subseteq \ldots \subseteq W_n$$

von Unterräumen von V.

Eine **vollständige Fahne** ist eine Fahne

$$0 = W_0 \subset W_1 \subset W_2 \subset \ldots \subset W_n = V$$

mit $\dim(W_i) = i$ für alle $i = 0, \ldots, n$.

Beispiel 19.14. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis des K-Vektorraums V. Dann beschreiben die Unterräume

$$W_i = \langle b_1, \ldots, b_i \rangle_K$$

eine vollständige Fahne von V.

Aufgrund des Basisergänzungssatzes ist jede vollständige Fahne $0 = W_0 \subseteq W_1 \subseteq W_2 \subseteq \ldots \subseteq W_n = V$ von dieser Form: $b_1, \ldots, b_i$ sei konstruiert als Basis von W_i. Dann wählt man $b_{i+1} \in W_{i+1} \setminus W_i$ und erhält mit $b_1, \ldots, b_{i+1}$ eine Basis von W_{i+1}. Und weiter per Induktion ...

Proposition 19.15. *Sei $P(X) = P_1(X) \cdot P_2(X)$ für Polynome $P, P_1, P_2 \in K[X]$. Wenn P vollständig in Linearfaktoren zerfällt, dann auch P_1 und P_2.*

Beweis. Seien Q, Q_1 und Q_2 die Polynome ohne Nullstellen in K aus den Faktorisierungen von P, P_1 und P_2 wie in Satz 17.23. Dann hat $Q_1(X) \cdot Q_2(X)$ auch keine Nullstellen in K, denn für jede solche $\alpha \in K$ wäre bereits einer der Faktoren in $Q_1(\alpha) \cdot Q_2(\alpha)$ Null.

Daher erhält man durch Multiplizieren der Darstellungen für P_1 mit der von P_2 eine Darstellung für P. Die Eindeutigkeit in Satz 17.23 zeigt

$$Q(X) = Q_1(X) \cdot Q_2(X).$$

Nach Voraussetzung ist $\deg(Q) = 0$, also $0 = \deg(Q) = \deg(Q_1) + \deg(Q_2)$. Daher sind $Q_1(X)$ und $Q_2(X)$ konstant. □

Satz 19.16 (Fahnensatz)**.** *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums V. Dann sind äquivalent:*

(a) *V hat eine vollständige Fahne $0 = W_0 \subseteq W_1 \subseteq W_2 \subseteq \ldots \subseteq W_n = V$ aus f-invarianten Unterräumen W_i.*

(b) *Es gibt eine Basis $\mathscr{B} = (b_1, \ldots, b_n)$, bezüglich derer f eine obere Dreiecksmatrix als Darstellungsmatrix hat.*

(c) *$\chi_f(X)$ zerfällt vollständig in Linearfaktoren.*

Beweis. (a) $\Longrightarrow$ (b): Wir konstruieren wie in Beispiel 19.14 eine Basis $(b_1, \ldots, b_n)$ mit

$$W_i = \langle b_1, \ldots, b_i \rangle_K.$$

Weil $f(b_j) \in f(W_j) \subseteq W_j$ gilt, kann man $f(b_j)$ als Linearkombination der b_i mit $i \leq j$ schreiben. Die Darstellungsmatrix bezüglich $\mathscr{B}$ hat als j-te Spalte gerade $f(b_j)$ in Koordinaten bezüglich $\mathscr{B}$, also nur Einträge $\neq 0$ in Zeilen mit $i \leq j$. Diese Darstellungsmatrix ist eine obere Dreiecksmatrix.

(b) $\Longrightarrow$ (c): Das charakteristische Polynom kann man bezüglich jeder Darstellungsmatrix berechnen. Also wählen wir eine Basis $\mathscr{B}$, so daß $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ eine obere Dreiecksmatrix ist. In diesem Fall haben wir in Beispiel 17.50 bereits $\chi_f(X) = \chi_A(X)$ berechnet: ein Produkt von Linearfaktoren.

(c) $\Longrightarrow$ (a): Dies zeigen wir per Induktion über $\dim(V)$. Wenn $\dim(V) = 0$, dann ist nichts zu tun. Wenn $\deg(\chi_f(X)) = \dim(V) > 0$, dann hat $\chi_f(X)$ einen Linearfaktor $X - \lambda$ und daher eine Nullstelle $\lambda \in K$. Nullstellen von $\chi_f(X)$ sind Eigenwerte von f. Es gibt daher ein $0 \neq v \in V_\lambda$. Sei

$$W_1 = \langle v \rangle_K.$$

Da v ein Eigenvektor ist, ist W_1 ein f-invarianter Unterraum. Sei $f_1 = f|_{W_1}$. Es induziert f auf $\bar{V} = V/W_1$ einen Endomorphismus $\bar{f} : \bar{V} \to \bar{V}$. Es gilt nach Korollar 19.12

$$\chi_f(X) = \chi_{f_1}(X) \cdot \chi_{\bar{f}}(X).$$

Nach Proposition 19.15 zerfällt auch $\chi_{\bar{f}}(X)$ vollständig in Linearfaktoren. Per Induktionsvoraussetzung gibt es eine $\bar{f}$-invariante vollständige Fahne

$$0 = \bar{W}_0 \subseteq \bar{W}_1 \subseteq \ldots \subseteq \bar{W}_{n-1} = \bar{V}.$$

Sei $\mathrm{pr} : V \to \bar{V} = V/W_1$ die Quotientenabbildung. Dann ist mit $W_i = \mathrm{pr}^{-1}(\bar{W}_{i-1})$

$$0 \subseteq W_1 \subseteq W_2 \subseteq \ldots \subseteq W_{n-1} \subseteq W_n = V. \tag{19.1}$$

Wegen

$$W_1 = \ker(\mathrm{pr}\,|_{W_i} : W_i \to \bar{W}_{i-1})$$

folgt aus dem Kern/Bild-Dimensionssatz, Satz 7.21,

$$\dim(W_i) = \dim(W_1) + \dim(\bar{W}_{i-1}) = 1 + (i-1) = i.$$

Außerdem ist für alle $w \in W_i$ auch $f(w) \in W_i$:

$$\mathrm{pr}(f(w)) = \bar{f}(\mathrm{pr}(w)) \in \bar{f}(\bar{W}_{i-1}) \subseteq \bar{W}_{i-1}.$$

Dies zeigt, daß (19.1) eine f-invariante vollständige Fahne in V ist. □

Bemerkung 19.17. Nach dem Fundamentalsatz der Algebra zerfällt jedes Polynom $P \in \mathbb{C}[X]$ vollständig in Linearfaktoren. Daher erfüllt jeder Endomorphismus eines endlichdimensionalen $\mathbb{C}$-Vektorraums die äquivalenten Eigenschaften des Fahnensatzes.

Beispiel 19.18. Sei $K = \mathbb{R}$ und $f : V \to V$ habe die Darstellungsmatrix

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

Dann besitzt V keine vollständige f-invariante Fahne. Das charakteristische Polynom ist

$$\chi_f(X) = X^2 + 1$$

und dies zerfällt in $\mathbb{R}[X]$ nicht in Linearfaktoren.

Übungsaufgaben zu §19

Übungsaufgabe 19.1. Zeigen Sie: die Menge der invertierbaren oberen Dreiecksmatrizen ist eine Untergruppe von $\mathrm{GL}_n(K)$. So eine Gruppe in $\mathrm{GL}_n(K)$ heißt Borelsche Untergruppe.

Bemerkung: Der Begriff einer Untergruppe ist analog zum Begriff des Unterraums eines Vektorraums. Das ist eine Teilmenge $H \subseteq G$ einer Gruppe G, so daß die Verknüpfung von G eine Verknüpfung $H \times H \to H$ definiert, so daß damit H selbst eine Gruppe ist.

Übungsaufgabe 19.2. Sei V ein K-Vektorraum mit Unterraum U. Sei W ein Komplement von U. Zeigen Sie, daß die Einschränkung der Quotientenabbildung $q : V \to V/U$ auf W einen Isomorphismus

$$\begin{aligned} q|_W : W &\xrightarrow{\sim} V/U \\ x &\mapsto q(x) = [x] \end{aligned}$$

definiert.

Übungsaufgabe 19.3. Sei V ein K-Vektorraum mit Unterraum U. Zeigen Sie, daß

$$q^* : (V/U)^* \to V^*$$

injektiv ist und die durch Einschränkung auf U induzierte Abbildung

$$V^* \to U^*, \quad \pi \mapsto \pi|_U$$

einen Isomorphismus

$$\begin{aligned} V^*/q^*((V/U)^*) &\xrightarrow{\sim} U^* \\ [\pi] &\mapsto \pi|_U \end{aligned}$$

induziert.

19.3. **Jordanblöcke im zerfallenden Fall.** Eine Normalform für eine Matrix A beschreibt diese bis auf erlaubte Operationen

$$A \rightsquigarrow SAT \qquad \text{mit } S, T \text{ invertierbar}$$

durch eine einfache Gestalt, die dabei bis auf kontrollierte Mehrdeutigkeit — meist nur Permutationen — eindeutig ist.

Denkt man bei einer Matrix an die Darstellungsmatrix einer linearen Abbildung, so entspricht dies dem Übergang zur Darstellungsmatrix in einer geeigneten Basis. Diese Sichtweise hat den Vorteil, daß sich das zugrundeliegende Objekt, die lineare Abbildung, nicht verändert: nur die Koordinatenbeschreibung durch eine Matrix verändert sich. Eine der Abbildung angepaßte Basiswahl führt zu einer möglichst einfachen Matrix.

Bemerkung 19.19. Hier sind ein paar Beispiele von Normalformen, die wir bereits kennen.

(1) Beispiel 8.38: aus dem Beweis der Kern–Bild–Dimensionsformel in Satz 7.21 entstammt für eine lineare Abbildung $f : V \to W$ vom Rang r eine Darstellungsmatrix in besonders einfacher Form als Blockmatrix mit Zeilenaufteilung $n = r + t$ und Spaltenaufteilung $m = r + s$:

$$\mathrm{M}_{\mathscr{C}}^{\mathscr{B}}(f) = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}_{n\times m}(K).$$

Hier sind Definitionsbereich V und der Wertebereich W unabhängig, und damit auch die Basiswahl in diesen Vektorräumen unabhängig. Entsprechend findet man für $A \in \mathrm{M}_{n\times m}(K)$ vom Rang r invertierbare Matrizen $S \in \mathrm{GL}_n(K)$ und $T \in \mathrm{GL}_m(K)$ mit

$$SAT = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Die Darstellungsmatrix in einfacher Form hängt nur von $r = \mathrm{rg}(A)$ ab und ist als solche eindeutig.

(2) Das Gauß-Eliminationsverfahren liefert eine zweite Normalform für $A \in \mathrm{M}_{n\times m}(K)$. Es gibt ein Produkt S aus Elementarmatrizen und eine Matrix R in Zeilenstufenform mit

$$SA = R.$$

Wie in Satz 9.59 beschrieben findet man dann die LR-Zerlegung:

$$PA = LR$$

mit einer Permutationsmatrix P, einer unteren Dreiecksmatrix L und einer oberen Dreiecksmatrix R.

(3) Wenn das charakteristische Polynom von $f : V \to V$ vollständig in Linearfaktoren zerfällt, dann gibt es nach dem Fahnensatz, Satz 19.16, eine Basis $\mathscr{B}$, so daß

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) \text{ obere Dreiecksmatrix ist.}$$

Nun sind Definitionsbereich und Wertebereich identisch und sollen auch durch die gleiche Basis mit Koordinaten beschrieben werden. Dies schränkt die Möglichkeiten, die Darstellungsmatrix zu wählen, beträchtlich ein.

Außerdem gibt es im allgemeinen eine große Freiheit für diese Darstellung als obere Dreiecksmatrix. Insbesondere kann man über die Einträge oberhalb der Diagonalen nicht viel sagen. Das ändert der Satz über die Jordannormalform.

(4) Wenn das charakteristische Polynom nicht vollständig in Linearfaktoren zerfällt, gibt es ebenfalls eine Jordannormalform, siehe die Vorlesung Grundlagen der Algebra.

Definition 19.20. Seien K ein Körper, $\lambda \in K$ und $d \in \mathbb{N}$. Der **Jordanblock der Länge d zum Eigenwert** λ ist die Matrix

$$J_d(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in \mathrm{M}_d(K).$$

Alle fehlenden Einträge von $J_d(\lambda)$ sind wie üblich mit 0 aufzufüllen.

Beispiel 19.21. Oft haben wir die Matrix

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

als Beispiel herangezogen. Dies ist ein Jordanblock der Länge 2 zum Eigenwert 1. Geometrisch beschreibt die Multiplikation mit dieser Matrix eine Scherung mit dem Eigenraum zum Eigenwert 1 als Fixgerade.

Bemerkung 19.22. Ein Jordanblock $J_d(\lambda)$ gehört als Darstellungsmatrix zu einer linearen Abbildung $f : V \to V$ mit $\dim(V) = d$ und einer Basis $\mathscr{B} = (v_1, \ldots, v_d)$ mit

$$f(v_i) = \lambda v_i + v_{i-1} \qquad \forall i = 2, \ldots, d$$

und $f(v_1) = \lambda v_1$. Damit ist das charakteristische Polynom

$$\chi_{J_d(\lambda)} = (X - \lambda)^d$$

und λ der einzige Eigenwert. Zur Bestimmung des Eigenraums betrachten wir

$$J_d(\lambda) - \lambda \cdot \mathbf{1} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

mit der einfachen Interpretation, daß die Koordinaten $x = (x_1, \ldots, x_d)^t$ um eins verschoben werden:

$$\big(J_d(\lambda) - \lambda \cdot \mathbf{1}\big) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} = \begin{pmatrix} x_2 \\ \vdots \\ x_d \\ 0 \end{pmatrix}.$$

Der Eigenraum zu λ ist damit offensichtlich

$$V_\lambda = \langle v_1 \rangle_K$$

und so

$$\mathrm{am}_\lambda(f) = d, \qquad \mathrm{gm}_\lambda(f) = 1.$$

Außerdem kann man so auch das Minimalpolynom ablesen: als Teiler von $\chi_{J_d(\lambda)}(X)$ ist $m_{J_d(\lambda)} = (X-\lambda)^e$ für ein $1 \le e \le d$. Weil für $m < d$

$$(X-\lambda)^m|_{X=J_d(\lambda)} v_d = \big(J_d(\lambda) - \lambda \cdot \mathbf{1}\big)^m v_d = J_d(0)^m v_d = v_{d-m},$$

brauchen wir $e = d$, also

$$m_{J_d(\lambda)} = (X-\lambda)^d = \chi_{J_d(\lambda)}.$$

Lemma 19.23. *Seien $P \in K[X]$, und*

$$M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in \mathrm{M}_n(K).$$

Dann gilt

$$P(M) = \begin{pmatrix} P(A) & * \\ 0 & P(D) \end{pmatrix}.$$

Wenn $B = 0$, dann gilt sogar

$$P(M) = \begin{pmatrix} P(A) & 0 \\ 0 & P(D) \end{pmatrix}.$$

Beweis. Beide Seiten sind linear in P. Es bleibt also nur der Fall $P(X) = X^m$. Das folgt aus

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \cdot \begin{pmatrix} A' & B' \\ 0 & D' \end{pmatrix} = \begin{pmatrix} AA' & AB' + BD' \\ 0 & DD' \end{pmatrix},$$

was nach kurzer Rechnung aus der Definition der Matrixmultiplikation folgt. □

Theorem 19.24 (Jordannormalform, zerfallender Fall)**.** *Sei K ein Körper und sei $A \in \mathrm{M}_n(K)$ eine Matrix, so daß $\chi_A(X)$ vollständig in Linearfaktoren zerfällt.*
Dann gibt es $S \in \mathrm{GL}_n(K)$ und $\lambda_1, \ldots, \lambda_r \in K$ und $d_1, \ldots, d_r \in \mathbb{N}$ mit

$$SAS^{-1} = \begin{pmatrix} \boxed{J_{d_1}(\lambda_1)} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \boxed{J_{d_r}(\lambda_r)} \end{pmatrix}$$

Dabei gelten die folgenden Beziehungen für alle $\lambda \in K$.

(1) Die algebraische Multiplizität von λ ist die Summe der Längen der Jordanblöcke zum Eigenwert λ:

$$\mathrm{am}_A(\lambda) = \sum_{i \text{ mit } \lambda_i = \lambda} d_i.$$

(2) Die geometrische Multiplizität von λ ist die Anzahl der Jordanblöcke zum Eigenwert λ:

$$\mathrm{gm}_A(\lambda) = |\{i \ ; \ \lambda_i = \lambda\}|.$$

(3) Die Multiplizität von λ als Nullstelle von $m_A(X)$ ist die maximale Länge eines Jordanblocks zum Eigenwert λ:

$$\max\{d_i \ ; \ \lambda_i = \lambda\}.$$

Beispiel 19.25. Sei $A \in \mathrm{M}_3(K)$ mit $\chi_A(X) = (X-a)^2(X-b)$ für $a \neq b \in K$. Dann ist

$$m_A(X) = (X-a)^e(X-b)$$

mit $e = 1$ oder $e = 2$.

- Falls $e = 1$, so ist A diagonalisierbar mit Jordannormalform
$$\begin{pmatrix} a & & \\ & a & \\ & & b \end{pmatrix}.$$
- Falls $e = 2$, so hat A die Jordannormalform
$$\begin{pmatrix} a & 1 & \\ & a & \\ & & b \end{pmatrix}.$$

Den Beweis der Existenz der Jordannormalform erbringen wir nach einigen Vorarbeiten.

Beweis der Formeln für die Multiplizitäten. Die betrachteten numerischen Werte stimmen für A und SAS^{-1} überein. Wir dürfen daher ohne Einschränkung $S = \mathbf{1}$ annehmen, d.h., A hat bereits die Gestalt der Jordannormalform.

(1) Die Jordannormalform ist eine Matrix in oberer Dreiecksgestalt, so daß wir direkt ablesen können
$$\chi_A(X) = \prod_{i=1}^{r} (X - \lambda_i)^{d_i}.$$
Die Formel für die algebraische Multiplizität $\mathrm{am}_A(\lambda)$ ergibt sich durch Zusammensortieren der Faktoren $X - \lambda$.

(2) Den Eigenraum V_λ bestimmen wir am besten in der Basis, in der A die Gestalt der Jordannormalform annimmt. Dann entkoppeln in
$$(\lambda \cdot \mathbf{1} - A)v = 0$$
die Gleichungen für die einzelnen Jordanblöcke. Das bedeutet, daß V_λ direkte Summe der entsprechenden Eigenräume der einzelnen Jordanblöcke ist. Diese haben wir in Bemerkung 19.22 diskutiert. Nur die Jordanblöcke $J_{d_i}(\lambda_i)$ mit $\lambda_i = \lambda$ tragen bei, und jeder davon mit einem eindimensionalen Eigenraum. Damit hat V_λ eine Basis, die aus den Vektoren, welche zu den jeweils ersten Spalten der Jordanblöcke mit $\lambda_i = \lambda$ gehören, gebildet wird. Folglich ist $\dim(V_\lambda)$ wie in (2) behauptet.

(3) Sei $W_i \subseteq K^n$ der Unterraum, auf dem A mit dem Jordanblock $J_{d_i}(\lambda_i)$ wirkt. Das Minimalpolynom $m_A(X)$ muß W_i nach Einsetzen von A annullieren. Aber auf W_i ist
$$0 = m_A(A)|_{W_i} = m_A(J_{d_i}(\lambda_i)).$$
(Hier vermischen wir die Notation für die Einschränkung einer linearen Abbildung mit der Matrixnotation!) Nach Bemerkung 19.22 und der definierenden Eigenschaft des Minimalpolynoms (zu $J_{d_i}(\lambda_i)$) ist damit $(X - \lambda_i)^{d_i}$ als Minimalpolynom von $J_{d_i}(\lambda_i)$ ein Teiler von $m_A(X)$.

Wenden wir dies für alle i an, so ergibt sich wegen Satz 17.23
$$e_\lambda := \max\{d_i \ ; \ \lambda_i = \lambda\}$$
als untere Schranke für die Vielfachheit von λ in $m_A(X)$.

Andererseits gilt für
$$P(X) = \prod_{\lambda} (X - \lambda)^{e_\lambda}$$
per Induktion nach Lemma 19.23
$$P(A) = \begin{pmatrix} P(J_{d_1}(\lambda_1)) & & \\ & \ddots & \\ & & P(J_{d_r}(\lambda_r)) \end{pmatrix}.$$

Für $\lambda_i = \lambda$ ist

$$P(J_{d_i}(\lambda_i)) = \Big(\prod_{\mu \neq \lambda_i} (J_{d_i}(\lambda_i) - \mu \cdot \mathbf{1})^{e_\mu} \Big) \cdot (J_{d_i}(\lambda_i) - \lambda_i \cdot \mathbf{1})^{e_\lambda} = 0,$$

denn es ist ja $d_i \leq e_\lambda$ und bereits

$$(J_{d_i}(\lambda_i) - \lambda_i \cdot \mathbf{1})^{d_i} = 0.$$

Folglich ist $P(X)$ ein Vielfaches von $m_A(X)$. Aber wenn zwei normierte Polynome gegenseitig Vielfaches voneinander sind, dann müssen sie gleich sein, und das beweist (3). □

Der Satz über die Jordannormalform hat eine analoge Version für Endomorphismen.

Korollar 19.26 (Jordannormalform für Endomorphismen, zerfallender Fall)**.** *Seien V ein endlichdimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus, so daß*

$$\chi_f(X) \text{ vollständig in Linearfaktoren zerfällt.}$$

Dann gibt es eine Basis $\mathscr{B}$ von V und $\lambda_1, \ldots, \lambda_r \in K$ und $d_1, \ldots, d_r \in \mathbb{N}$ mit

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} \boxed{J_{d_1}(\lambda_1)} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \boxed{J_{d_r}(\lambda_r)} \end{pmatrix}$$

Dabei gelten die folgenden Beziehungen für alle $\lambda \in K$.

(1) Die Multiplizität von λ als Nullstelle von $m_f(X)$ ist die maximale Länge eines Jordanblocks zum Eigenwert λ:

$$\max\{d_i \ ; \ \lambda_i = \lambda\}.$$

(2) Die geometrische Multiplizität von λ ist die Anzahl der Jordanblöcke zum Eigenwert λ:

$$\mathrm{gm}_f(\lambda) = |\{i \ ; \ \lambda_i = \lambda\}|.$$

(3) Die algebraische Multiplizität von λ ist die Summe der Längen der Jordanblöcke zum Eigenwert λ:

$$\mathrm{am}_f(\lambda) = \sum_{i \text{ mit } \lambda_i = \lambda} d_i.$$

19.4. **Verallgemeinerte Eigenräume und Kernzerlegung.** Sei $A \in \mathrm{M}_n(K)$ eine Matrix. Zu $r \in \mathbb{N}_0$ und $\lambda \in K$ betrachten wir

$$V_{\lambda,r} = \ker\big((A - \lambda \cdot \mathbf{1})^r\big) \subseteq K^n.$$

Für $r = 0$ ist per Konvention $(A - \lambda \cdot \mathbf{1})^0 = \mathbf{1}$, also

$$V_{\lambda,0} = (0).$$

Für $r = 1$ ist

$$V_{\lambda,1} = V_\lambda$$

der Eigenraum von f zum Eigenwert λ. Allgemein gilt

$$V_{\lambda,r} \subseteq V_{\lambda,r+1},$$

denn für $x \in V_{\lambda,r}$ folgt

$$(A - \lambda \cdot \mathbf{1})^{r+1}(x) = (A - \lambda \cdot \mathbf{1})\big((A - \lambda \cdot \mathbf{1})^r(x)\big) = (A - \lambda \cdot \mathbf{1})(0) = 0.$$

Zu festem λ bekommen wir so eine Fahne von Unterräumen

$$(0) = V_{\lambda,0} \subseteq V_{\lambda,1} \subseteq \ldots \subseteq V_{\lambda,r} \subseteq \ldots \subseteq V := K^n.$$

Weil dies eine aufsteigende Folge von Unterräumen ist, handelt es sich bei

$$V(\lambda) := \bigcup_{r \geq 0} V_{\lambda,r}$$

um einen Unterraum von V. Wegen

$$0 = \dim(V_{\lambda,0}) \leq \dim(V_{\lambda,1}) \leq \ldots \leq \dim(V_{\lambda,r}) \leq \dim(V_{\lambda,r+1}) \leq \ldots \leq \dim(V) = n.$$

kann die Dimension $\dim(V_{\lambda,r})$ nicht beliebig groß werden. Die Fahne wird damit für hinreichend große (spätestens mit $r \geq \dim(V)$; aber das muß man zeigen!) konstant, siehe Korollar 5.34, und

$$V(\lambda) = V_{\lambda,r} \qquad \text{für alle hinreichend großen } r.$$

Für einen Endomorphismus $f : V \to V$ betrachten wir entsprechend die analog konstruierten Unterräume $V_{\lambda,r} = \ker(f - \lambda \cdot \mathrm{id}_V)$.

Lemma 19.27. *Wenn λ kein Eigenwert von A ist, dann gilt $V_{\lambda,r} = 0$ für alle $r \geq 0$.*

Beweis. Wenn λ kein Eigenwert ist, dann ist $\det(\lambda \cdot \mathbf{1} - A) \neq 0$ und $\lambda \cdot \mathbf{1} - A$ ist invertierbar, siehe Satz 11.37. Dann ist aber auch $(A - \lambda \cdot \mathbf{1})^r$ für alle $r \geq 0$ invertierbar. Somit folgt

$$V_{\lambda,r} = \ker((A - \lambda \cdot \mathbf{1})^r) = 0. \qquad \square$$

Definition 19.28. Die Unterräume $V_{\lambda,r}$ wie oben heißen **verallgemeinerte Eigenräume der Stufe** r zum Eigenwert λ. Der Unterraum

$$V(\lambda)$$

heißt **verallgemeinerter Eigenraum** (oder **Hauptraum**) zum Eigenwert λ.

Beispiel 19.29. Betrachten wir einen Jordanblock $J_d(\lambda)$. Weil $J_d(0) = J_d(\lambda) - \lambda \cdot \mathbf{1}$ die Koordinaten um eins nach oben verschiebt, folgt sofort für alle $1 \leq r \leq d$

$$V_{\lambda,r} = \langle e_1, \ldots, e_r \rangle_K.$$

Die verallgemeinerten Eigenräume bilden daher eine vollständige Fahne von K^d.

Proposition 19.30. *Ein verallgemeinerter Eigenraum von A ist ein A-invarianter Unterraum.*

Beweis. Wenn $v \in V_{\lambda,r}$, dann ist

$$\begin{aligned}(A - \lambda \cdot \mathbf{1})^r (Av) &= \Big((X - \lambda)^r X\Big)\Big|_{X=A} v \\ &= \Big(X(X - \lambda)^r\Big)\Big|_{X=A} v = A\Big((A - \lambda \cdot \mathbf{1})^r v\Big) = A \cdot 0 = 0,\end{aligned}$$

also auch $Av \in V_{\lambda,r}$. $\square$

Die Zerlegung als innere direkte Summe im folgenden Satz heißt Kernzerlegung, weil der Hauptraum als Kern geschrieben werden kann.

Satz 19.31 (Kernzerlegung, zerfallender Fall)**.** *Sei $A \in \mathrm{M}_n(K)$ und seien $\lambda_1, \ldots, \lambda_s$ die paarweise verschiedenen Eigenwerte von A. Wenn $\chi_A(X)$ vollständig in Linearfaktoren zerfällt, dann ist (mit $V = K^n$)*

$$V = \bigoplus_{i=1}^{s} V(\lambda_i).$$

Beweis. Schritt 1: Nach Voraussetzung ist

$$\chi_A(X) = \prod_{i=1}^{s} (X - \lambda_i)^{\mathrm{am}_A(\lambda_i)}.$$

Nach dem Satz von Cayley–Hamilton bzw. genauer nach Satz 18.31 ist das Minimalpolynom $m_A(X)$ ein Teiler von $\chi_A(X)$. Somit gibt es aufgrund des Beweises von Proposition 19.15 Exponenten $m_i \in \mathbb{N}$ mit $m_i \leq \mathrm{am}_A(\lambda_i)$, so daß gilt:

$$m_A(X) = \prod_{i=1}^{s} (X - \lambda_i)^{m_i}.$$

Wir setzen nun

$$P_i(X) = \prod_{j=1,\ j\neq i}^{s} (X - \lambda_j)^{m_j},$$

so daß für alle $i = 1, \ldots, s$

$$m_A(X) = (X - \lambda_i)^{m_i} \cdot P_i(X).$$

Wir betrachten nun das Ideal aller $K[X]$-Linearkombinationen

$$\langle P_1(X), \ldots, P_s(X) \rangle_{K[X]} \subseteq K[X],$$

das ist die Menge

$$\langle P_1(X), \ldots, P_s(X) \rangle_{K[X]} = \{ \sum_{i=1}^{s} h_i(X) \cdot P_i(X) \ ; \ h_i \in K[X] \text{ für } i = 1, \ldots, s \}.$$

Die Idealeigenschaft ist offensichtlich.

Schritt 2: Weil $K[X]$ ein Hauptidealring ist, gibt es ein normiertes $F \in K[X]$ mit

$$\langle P_1(X), \ldots, P_s(X) \rangle_{K[X]} = (F),$$

d.h., alle $P_i(X)$ sind Vielfache von $F(X)$. Aufgrund der Eindeutigkeit der Linearfaktorzerlegung, Satz 17.23, ist $F(X)$ ein Produkt aus Linearfaktoren $(X - \lambda_i)$. Aber weil gerade dieser Faktor in $P_i(X)$ nicht auftritt, muß $F = 1$ sein.

Insbesondere ist $1 \in \langle P_1(X), \ldots, P_s(X) \rangle_{K[X]}$. Wir schließen so auf die Existenz von Polynomen $Q_1, \ldots, Q_s \in K[X]$ mit

$$1 = Q_1(X)P_1(X) + \ldots + Q_s(X)P_s(X). \tag{19.2}$$

Wir wollen nun zeigen daß

$$\pi_i := L_{Q_i(A)P_i(A)} : V \to V$$

ein Projektor auf den direkten Summanden $V(\lambda_i)$ ist. Dies zeigen wir nicht explizit, aber es folgt aus den Rechnungen, die wir für den Beweis anstellen, und diese Vorstellung soll einen roten Faden für das Folgende bieten.

Schritt 3: Für alle $i = 1, \ldots, s$ und alle $v \in V$ ist

$$v_i := \pi_i(v) = Q_i(A)P_i(A)v \subseteq V(\lambda_i),$$

denn

$$\begin{aligned}
(A - \lambda_i)^{m_i}(v_i) &= (A - \lambda_i)^{m_i}\big(Q_i(A)P_i(A)v\big) = \big((A - \lambda_i)^{m_i} Q_i(A)P_i(A)\big)v \\
&= \big((X - \lambda_i)^{m_i} \cdot Q_i(X)P_i(X)\big)\big|_{X=A}\, v \\
&= \big(Q_i(X)m_A(X)\big)\big|_{X=A}\, v = Q_i(A)\big(m_A(A)v\big) = Q_i(A)0 = 0.
\end{aligned}$$

Bedenkt man, daß die Einheitsmatrix $\mathbf{1}$ die Auswertung von $1 \in K[X]$ in der Matrix A ist, so folgt weiter aus (19.2)

$$v = \mathbf{1} \cdot v = 1|_{X=A} v = \big(Q_1(X)P_1(X) + \ldots + Q_s(X)P_s(X)\big)\big|_{X=A}\, v = \sum_{i=1}^{s} Q_i(A)P_i(A)v = \sum_{i=1}^{s} v_i.$$

Dies bedeutet, daß die Haupträume schon einmal V erzeugen, genauer:

$$V = V_{\lambda_1,m_1} + V_{\lambda_2,m_2} + \ldots + V_{\lambda_s,m_s}.$$

Schritt 4: Sei $1 \leq i \neq j \leq s$ und $v \in V_{\lambda_i,m_i}$. Setzen wir

$$R_{ij}(X) = Q_j(X) \cdot \prod_{k\neq i,j} (X-\lambda_k)^{m_k}$$

(R wie Rest). Dann folgt

$$\begin{aligned} Q_j(A)P_j(A)v &= \big(Q_j(X)\prod_{k\neq j}(X-\lambda_k)^{m_k}\big)\Big|_{X=A} v = \big(R_{ij}(X)(X-\lambda_i)^{m_i}\big)\big|_{X=A}\, v \\ &= R_{ij}(A)\big((A-\lambda_i)^{m_i}v\big) = R_{ij}(A)0 = 0. \end{aligned} \tag{19.3}$$

Mit (19.2) bewirkt (19.3) aber auch, daß

$$Q_i(A)P_i(A)v = \Big(\sum_{j=1}^{s} Q_j(A)P_j(A)\Big)v = \big(Q_1(X)P_1(X) + \ldots + Q_s(X)P_s(X)\big)\big|_{X=A}\, v = \mathbf{1}v = v.$$

Schritt 5: Bei jeder Darstellung

$$v = v_1 + \ldots + v_s$$

mit $v_i \in V_{\lambda_i,m_i}$ für alle $i = 1,\ldots,s$ gilt dann

$$Q_i(A)P_i(A)v = \sum_{j=1}^{s} Q_i(A)P_i(A)v_j = v_i,$$

woraus die Eindeutigkeit einer solchen Darstellung folgt. Somit haben wir eine innere direkte Summe

$$V = \bigoplus_{i=1}^{s} V_{\lambda_i,m_i}.$$

Schritt 6: Es fehlt nun nur noch die Aussage $V(\lambda_i) = V_{\lambda_i,m_i}$ für alle $i = 1,\ldots,s$. Das Minimalpolynom von A auf V_{λ_j,m_j} ist ein Teiler von (sogar gleich) $(X-\lambda_j)^{m_j}$. Folglich hat auch das charakteristische Polynom diese Gestalt (mit einem anderen Exponenten). Auf dem Komplement von V_{λ_i,m_i}, dem Unterraum

$$W_i = \bigoplus_{j=1,\ j\neq i}^{s} V_{\lambda_j,m_j}$$

haben wir als charakteristisches Polynom demnach ein Produkt von Potenzen von $(X-\lambda_j)$ mit $j \neq i$. Folglich ist λ_i keine Nullstelle von

$$\chi_{A|_{W_i}}(X).$$

Daher ist eingeschränkt auf W_i die Abbildung $A-\lambda_i$ invertierbar. Das gilt dann auch für $(A-\lambda_i)^r$ und alle $r \geq 0$. Damit ist W_i im Bild von $(A-\lambda_i)^r$ enthalten für alle r und somit

$$\dim(V_{\lambda_i,r}) = \dim(V) - \dim \operatorname{im}(A-\lambda_i)^r) \leq \dim(V) - \dim(W_i) = \dim(V_{\lambda_i,m_i}).$$

Daraus schließen wir, daß die Fahne der verallgemeinerten Eigenräume zum Eigenwert λ_i sich ab V_{λ_i,m_i} nicht mehr ändert, demnach

$$V(\lambda_i) = V_{\lambda_i,m_i}.$$ □

Satz 19.32. *Sei $A \in \mathrm{M}_n(K)$. Dann sind äquivalent:*

(a) A ist diagonalisierbar.

(b) $m_A(X)$ *zerfällt in Linearfaktoren und alle Nullstellen von* $m_A(X)$ *haben Multiplizität* 1.

Für Endomorphismen von endlichdimensionalen Vektorräumen gilt die analoge Aussage.

Beweis. (a) $\Longrightarrow$ (b): Sei A diagonalisierbar. Die Aussage von (b) hängt von A nur bis auf Ähnlichkeit ab, so daß wir ohne Einschränkung annehemen dürfen, daß $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ bereits eine Diagonalmatrix ist.

Sei $P(X) \in K[X]$ beliebig. Dann ist

$$P(A) = \operatorname{diag}(P(\lambda_1), \ldots, P(\lambda_n)),$$

so daß $P(A) = 0$ genau dann, wenn alle $\lambda_1, \ldots, \lambda_n$ Nullstellen von P sind. Das Minimalpolynom hat dann jeden Linearfaktor $X - \lambda$ für $\lambda \in \{\lambda_1, \ldots, \lambda_n\}$ genau einmal (sind zwei der λ_i gleich, dann kommt der entsprechende Faktor nur einmal vor). Das ist (b).

(b) $\Longrightarrow$ (a): Nach dem Beweis der Kernzerlegung, Satz 19.31, ist unter der Voraussetzung (b) für alle $\lambda \in K$

$$V(\lambda) = V_{\lambda,1} = V_\lambda$$

und

$$V = \bigoplus_{\lambda \text{ Eigenwert}} V_\lambda.$$

Damit ist A diagonalisierbar nach Theorem 18.10. □

Beweis der Existenz der Jordannormalform. Schritt 1: Wir setzen $V = K^n$. Weil $\chi_A(X)$ vollständig in Linearfaktoren zerfällt, haben wir nach Satz 19.31 die Kernzerlegung

$$V = \bigoplus_{\lambda \text{ Eigenwert}} V(\lambda)$$

als innere direkte Summe von A-invarianten Unterräumen, siehe Proposition 19.30. Damit nimmt A in einer angepaßten Basis Blockdiagonalgestalt an. Wir dürfen daher V und A durch $V(\lambda)$ und die Einschränkung von A, genauer der Linksmultiplikation mit A, auf $V(\lambda)$ ersetzen. Dabei ist wesentlich, daß die Voraussetzungen erhalten bleiben: nach Korollar 19.12 ist das charakteristische Polynom der Einschränkung ein Faktor von $\chi_A(X)$, und daher zerfällt es nach Proposition 19.15 auch vollständig in Linearfaktoren.

Schritt 2: Weil $V = V(\lambda)$, ist das Minimalpolynom von A nun von der Form $(X - \lambda)^m$. Wir wollen zeigen, daß wir in einer geeigneten Basis eine Blockdiagonalmatrix aus Jordanblöcken $J_{d_i}(\lambda)$ für $i = 1, \ldots, s$ haben. Dabei kommt nur noch ein Eigenwert vor.

Wir verschieben nun A um $\lambda \cdot \mathbf{1}$ zu

$$B = A - \lambda \cdot \mathbf{1}.$$

Weil für alle $P(X) \in K[X]$ gilt

$$P(X)|_{X=A} = 0 \iff P(X + \lambda)|_{X=B} = 0,$$

folgt für das Minimalpolynom von B

$$m_B(X) = m_A(X + \lambda) = \big((X + \lambda) - \lambda\big)^m = X^m.$$

Durch diese Verschiebung haben wir eine Matrix B bekommen, die nur noch den Eigenwert 0 hat. Es reicht, die Aussage für B zu beweisen, wie man sofort aus der Rechnung

$$SAS^{-1} = S(\lambda \cdot \mathbf{1} + B)S^{-1} = S(\lambda \cdot \mathbf{1})S^{-1} + SBS^{-1} = \lambda \cdot \mathbf{1} + SBS^{-1}$$

sieht: Die Jordannormalform zu B wird durch Addition von $\lambda \cdot \mathbf{1}$ zur Jordannormalform von A.

Anstatt mit der Notation B weiterzuarbeiten, nehmen wir an, daß von Anfang an $\lambda = 0$ war, und arbeiten mit A weiter.

Schritt 3: Nun arbeiten wir per Induktion nach $n = \dim(V)$. Für $n = 0$ oder $n = 1$ ist nichts zu tun. Betrachten wir nun

$$W = \mathrm{im}(A) \subseteq V.$$

Dies ist ein A-invarianter Unterraum: für $w \in W$ ist $Aw \in \mathrm{im}(A) = W$.

Wenn $m_A(X) = X^m$, dann ist

$$A^{m-1}(W) = A^{m-1}(A(V)) = A^m(V) = 0.$$

Das Minimalpolynom der Multiplikation mit A auf W ist also höchstens X^{m-1}. Dies zeigt $W \neq V$, und daher $\dim(W) < \dim(V)$. Es ist $W = \mathrm{im}(A)$ ein A-invarianter Unterraum, und $\chi_{A|_W}(X)$ ist als Faktor von $\chi_A(X)$ ebenfalls vollständig in Linearfaktoren zerfallend. Wir können daher nach Darstellungsmatrixwahl auf die Einschränkung von L_A auf W die Induktionsvoraussetzung anwenden. Übersetzen wir die Existenz der Jordannormalform in eine Basis, so erhalten wir $w_1, \ldots, w_r \in W$, so daß für gewisse $d_i \in \mathbb{N}$, $d_i > 1$

$$\mathscr{B}_W = (A^{d_1-1}w_1, \ldots, Aw_1, w_1, \ldots, A^{d_r-1}w_r, \ldots, Aw_r, w_r)$$

eine Basis von W ist. Dabei gehört der Teil

$$(A^{d_i-1}w_i, \ldots, Aw_i, w_i)$$

zu einem Jordanblock der Länge d_i.

Weil $W = \mathrm{im}(A)$ ist, gibt es $v_i \in V$, $i = 1, \ldots, r$ mit

$$w_i = Av_i \qquad i = 1, \ldots, r.$$

In dieser Notation wird

$$\mathscr{B}_W = (A^{d_1}v_1, \ldots, A^2v_1, Av_1, \ldots, A^{d_r}v_r, \ldots, A^2v_r, Av_r).$$

Außerdem kennen wir bereits eine Baisis von $\ker(A)$, den Eigenraum zum Eigenwert 0, eingeschränkt auf W, nämlich nach Bemerkung 19.22 und dem Beweis von Theorem 19.24(2)

$$(A^{d_1}v_1, \ldots, A^{d_r}v_r) \in \ker(A) \cap W.$$

Nach dem Basisergänzungssatz kann dies zu einer Basis von $\ker(A)$

$$(A^{d_1}v_1, \ldots, A^{d_r}v_r, v_{r+1}, \ldots, v_{r+s})$$

ergänzt werden.

Schritt 4: Wir behaupten nun, daß

$$\mathscr{B} = (A^{d_1}v_1, \ldots, A^2v_1, Av_1, v_1, \ldots, A^{d_r}v_r, \ldots, A^2v_r, Av_r, v_r, v_{r+1}, \ldots, v_{r+s})$$

eine Basis von V ist. In dieser Basis nimmt A die Gestalt einer Jordannormalform ein. Das ist klar, denn mit $d_i = 0$ für $i = r+1, \ldots, r+s$ sorgt der Teil

$$(A^{d_i}v_i, \ldots, Av_i, v_i)$$

für einen Jordanblock der Länge $d_i + 1$.

Zählen wir zunächst, so finden wir nach Satz 7.21

$$|\mathscr{B}| = |\{v_1, \ldots, v_r\}| + |\mathscr{B}_W| + |\{v_{r+1}, \ldots, v_{r+s}\}| = \dim(\mathrm{im}(A)) + \dim(\ker(A)) = \dim(V).$$

Es reicht also nun zu zeigen, daß $\mathscr{B}$ linear unabhängig ist. Dazu notieren wir eine Linearkombination mittels Polynome $P_i(X) \in K[X]$ vom Grad $\deg(P_i) \leq d_i$ (dabei ist $d_i = 0$ für $i = r+1, \ldots, r+s$) als

$$0 = \sum_{i=1}^{r+s} P_i(A)v_i. \tag{19.4}$$

Multiplizieren wir zunächst mit A, so gilt wegen $AP_i(A)v_i = P_i(A)Av_i = P_i(A)w_i$ auch

$$0 = \sum_{i=1}^{r} P_i(A)w_i. \tag{19.5}$$

Dabei verschwinden die Indizes $> r$ wegen $Av_{r+j} = 0$. Nach Voraussetzung an den Grad gibt es für $i = 1, \ldots, r$ Elemente $a_i \in K$ und $Q_i \in K[X]$ mit $\deg(Q_i) \leq d_i - 1$ und

$$P_i(X) = a_i X^{d_i} + Q_i(X).$$

Wegen $A^{d_i} w_i = 0$ folgt $P_i(A)w_i = a_i A^{d_i} w_i + Q_i(A)w_i = Q_i(A)w_i$, somit wird (19.5) zu

$$0 = \sum_{i=1}^{r} Q_i(A)w_i.$$

Dies ist eine Linearkombination in W der Vektoren aus $\mathscr{B}_W$. Weil $\mathscr{B}_W$ eine Basis ist, sind alle Koeffizienten von $Q_i(X)$ gleich 0, d.h. $P_i(X) = a_i X^{d_i}$. Es bleibt von der ursprünglichen Linearkombination (19.4) nur

$$0 = \sum_{i=1}^{r+s} a_i A^{d_i} v_i$$

übrig (bei den Indizes $i > 0$ ist $d_i = 0$, also $a_i A^{d_i} v_i = a_i v_i$). Dies ist eine Linearkombination der gewählten Basis von $\ker(A)$. Also sind auch die verbliebenen Koeffizienten $a_i = 0$ für alle $i = 1, \ldots, r + s$. Dies zeigt die lineare Unabhängigkeit. Nach Satz 5.31 ist $\mathscr{B}$ eine Basis und das Theorem über die Jordannormalform im zerfallenden Fall bewiesen. □

Übungsaufgaben zu §19

Übungsaufgabe 19.4. Zeigen Sie: die Summe von verallgemeinerten Eigenräumen zu paarweise verschiedenen Eigenwerten ist direkt.

Übungsaufgabe 19.5. Sei $f : V \to V$ ein diagonalisierbarer Endomorphismus, und sei $U \subseteq V$ ein f-invarianter Unterraum. Zeigen Sie, daß $f|_U$ ebenfalls diagonalisierbar ist.

Übungsaufgabe 19.6. Seien $\lambda, \mu \in K$. Zeigen Sie

$$J_d(\lambda) + \mu \cdot \mathbf{1} = J_d(\lambda + \mu).$$

Übungsaufgabe 19.7. Sei $A \in \mathrm{M}_n(K)$ eine Matrix mit vollständig zerfallendem charakteristischen Polynom. Zeigen Sie, daß es eine diagonalisierbare Matrix D und eine nilpotente Matrix N gibt mit den folgenden Eigenschaften:

(i) $A = D + N$.

(ii) Die Matrizen A, D und N kommutieren miteinander:

$$AD = DA, \quad AN = NA, \quad DN = ND.$$

Bemerkung: Eine Matrix N heißt **nilpotent**, wenn es ein $r \geq 0$ gibt mit $N^r = 0$.

Teil 6. Appendix

Anhang A. Griechische Buchstaben

Oft gehen dem Mathematiker die Buchstaben aus. Es wird dann auf Buchstaben anderer Schriftsysteme zurückgegriffen. Das griechische Alphabet wird dabei traditionell bevorzugt zu Hilfe genommen.

Ein anderer Grund für den Wechsel der Schrift besteht in dem Wunsch einer parallelen Notation. Zu Objekten A, B, C, ..., die bereits jeweils ein zugehöriges a, b, c, ... haben, und zu denen es Varianten A', B', C', ... gibt, möchte man zum Beispiel jeweils eine Abbildung $X \to X'$ beschreiben für alle $X \in \{A, B, C, \ldots\}$. Da ist $\alpha : A \to A'$, und β, γ, ... eine gute Idee für eine Notation.

α, A	Alpha	η, H	Eta	ν, N	Ny	τ, T	Tau
β, B	Beta	$\vartheta, \theta, \Theta$	Theta	ξ, Ξ	Xi	υ, Υ	Ypsilon
γ, Γ	Gamma	ι, I	Iota	o, O	Omicron	φ, ϕ, Φ	Phi
δ, Δ	Delta	κ, K	Kappa	π, Π	Pi	χ, X	Chi
ε, ϵ, E	Epsilon	λ, Λ	Lambda	ρ, ϱ, P	Rho	ψ, Ψ	Psi
ζ, Z	Zeta	μ, M	My	$\sigma, \varsigma, \Sigma$	Sigma	ω, Ω	Omega

Abbildung 10. Das griechische Alphabet

Anhang B. Quaternionen

Die Quaternionen $\mathbb{H}$ wurden 1843 von Hamilton beim Versuch gefunden, auf einem $\mathbb{R}$-Vektorraum von Dimension 3 eine nullteilerfreie Multiplikation zu etablieren. Letzteres ist ihm nicht gelungen, denn eine solche $\mathbb{R}$-bilineare Multiplikation auf $\mathbb{R}^3$ gibt es nicht. Die neu entdeckten Quaternionen haben reelle Dimension 4.

Definition B.1. Die Hamiltonschen Quaternionen $\mathbb{H}$ sind definiert als der $\mathbb{R}$-Vektorraum mit Basis $1, i, j, k$ und $\mathbb{R}$-linearer Multiplikation gegeben durch

$$i^2 = j^2 = -1, \quad \text{und} \quad ij = k = -ji.$$

Es folgen die Relationen

$$k^2 = -1, \quad jk = i = -kj, \quad ki = j = -ik.$$

Elemente von $\mathbb{H}$ sind gegeben durch eindeutige Koordinaten $x, a, b, c \in \mathbb{R}$ als

$$w = x + ai + bj + ck$$

und das Produkt mit $\omega = \xi + \alpha i + \beta j + \gamma k$ ist definiert durch

$$\begin{aligned} w\omega = {} & (x\xi - a\alpha - b\beta - c\gamma) + (x\alpha + a\xi + b\gamma - c\beta)i \\ & + (x\beta + b\xi - a\gamma + c\alpha)j + (x\gamma + c\xi + a\beta - b\alpha)k. \end{aligned}$$

Bemerkung B.2. Die Quaternionen $\mathbb{H}$ kann man auch als $\mathbb{C}$-Vektorraum beschreiben. Zunächst beobachten wir, daß durch

$$\mathbb{C} \hookrightarrow \mathbb{H}, \quad x + ai \mapsto x + ai$$

eine Einbettung der komplexen Zahlen gegeben ist, die mit Multiplikation und Addition verträglich ist. Wir finden für jedes Quaternion $w = x + ai + bj + ck$ eine eindeutige Darstellung

$$w = x + ai + bj + ck = z_1 + z_2 j$$

mit komplexen Zahlen $z_1 = x + ai$, $z_2 = b + ci$ hat. Damit ist $\mathbb{H}$ bezüglich der Basis $1, j$ ein $\mathbb{C}$-Vektorraum der Dimension 2. Man beachte für $z = x + ai \in \mathbb{C}$ gilt:

$$jz = j(x + ai) = xj + aji = xj - aij = (x - ai)j = \bar{z}j.$$

Die Multiplikation wird nun in komplexen Koordinaten ein wenig einfacher durch

$$(z_1 + z_2 j)(\zeta_1 + \zeta_2 j) = (z_1\zeta_1 - z_2\bar{\zeta}_2) + (z_1\zeta_2 + z_2\bar{\zeta}_1)j$$

beschrieben.

Bemerkung B.3. Noch besser wird es allerdings, wenn wir die folgende injektive lineare Abbildung betrachten:

$$\rho : \mathbb{H} \hookrightarrow \mathrm{M}_2(\mathbb{C}), \quad \rho(z_1 + z_2 j) = \begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix}.$$

Diese Abbildung ist $\mathbb{R}$-linear und übersetzt die Multiplikation von Quaternionen in Matrixmultiplikation: für $w = z_1 + z_2 j$ und $\omega = \zeta_1 + \zeta_2 j$ ist

$$\begin{aligned} \rho(w)\rho(\omega) &= \begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix} \begin{pmatrix} \zeta_1 & \zeta_2 \\ -\bar{\zeta}_2 & \bar{\zeta}_1 \end{pmatrix} = \begin{pmatrix} z_1\zeta_1 - z_2\bar{\zeta}_2 & z_1\zeta_2 + z_2\bar{\zeta}_1 \\ -\bar{z}_1\bar{\zeta}_2 - \bar{z}_2\zeta_1 & \bar{z}_1\bar{\zeta}_1 - \bar{z}_2\zeta_2 \end{pmatrix} \\ &= \begin{pmatrix} z_1\zeta_1 - z_2\bar{\zeta}_2 & z_1\zeta_2 + z_2\bar{\zeta}_1 \\ \overline{-z_1\zeta_2 + z_2\bar{\zeta}_1} & \overline{z_1\zeta_1 - z_2\bar{\zeta}_2} \end{pmatrix} = \rho(w\omega). \end{aligned}$$

Bemerkung B.4. Sei ρ die komplexe Darstellung $\rho : \mathbb{H} \hookrightarrow \mathrm{M}_2(\mathbb{C})$ von oben. Die Bilder der Basis $1, i, j, k$ sind die Einheitsmatrix $\rho(1) = \mathbf{1}_2$ und

$$\rho(i) = \begin{pmatrix} i & \\ & -i \end{pmatrix}, \quad \rho(j) = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}, \quad \rho(k) = \begin{pmatrix} & i \\ i & \end{pmatrix}.$$

Die letzten drei Matrizen sind bis auf einen Faktor $-i$ in der Physik als Pauli-Matrizen bekannt. Die Pauli-Matrizen spielen eine Rolle in der quantenmechanischen Behandlung von Spin und beim Dirac-Operator zur Beschreibung von Fermionen. Kehrt man diese Verbindung um, so findet man in der Physik Anwendungen der Quaternionen.

Definition B.5. Ein **Schiefkörper** ist eine Menge D mit einer Addition $+$ und einer Multiplikation $\cdot$, so daß $(D, +, \cdot)$ alle Axiome eines Körpers erfüllt bis eventuell auf das Axiom der Kommutativität der Multiplikation.

Satz B.6. *Die Hamiltonschen Quaternionen $\mathbb{H}$ sind ein Schiefkörper bezüglich der Addition als $\mathbb{R}$-Vektorraum und der Multiplikation von Quaternionen.*

Beweis. Es ist furchtbar mühsam, die Schiefkörperaxiome direkt nachzuweisen. Daher haben wir zunächst als strukturelles Argument die komplexe Darstellung $\rho : \mathbb{H} \hookrightarrow \mathrm{M}_2(\mathbb{C})$ eingeführt. Axiome eines Schiefkörpers, die mit Identitäten von Termen zu tun haben („Rechenregeln“) folgen nun aus den entsprechenden Regeln für Matrizenmultiplikation in $\mathrm{M}_2(\mathbb{C})$.

Es fehlt die Existenz eines Inversen bezüglich der Multiplikation für alle $w \in \mathbb{H}$, $w \neq 0$. Aber auch dieses kann man raten, wenn man die Formel für die inverse Matrix benutzt. Zunächst ist die Determinante von $w = z_1 + z_2 j = x + ai + bj + ck$ definiert als die Norm

$$N(w) := \det(\rho(w)) = z_1\bar{z}_1 + z_2\bar{z}_2 = N(z_1) + N(z_2) = x^2 + a^2 + b^2 + c^2.$$

Die Norm ist somit multiplikativ, weil die Determinante multiplikativ ist, und zweitens stets $N(w) \neq 0$, sofern $w \neq 0$. Das Inverse von $\rho(w)$ in $\mathrm{M}_2(\mathbb{C})$ ist nun für $w \neq 0$ gegeben durch

$$\rho(w)^{-1} = \frac{1}{N(w)} \begin{pmatrix} \bar{z}_1 & -z_2 \\ \bar{z}_2 & z_1 \end{pmatrix} = \frac{1}{N(w)} \cdot \rho(\bar{w})$$

mit dem konjugierten Element

$$\bar{w} = x - ai - bj - ck = \bar{z}_1 - z_2 j.$$

Somit gilt $N(w) = w \cdot \bar{w}$ und für alle $w \neq 0$ die Formel für das Inverse

$$w^{-1} = N(w)^{-1} \cdot \bar{w} = \frac{1}{x^2 + a^2 + b^2 + c^2}(x - ai - bj - ck). \qquad \square$$

Bemerkung B.7. Es gilt für $w = z_1 + z_2 j$

$$\rho(\bar{w}) = \rho(\bar{z}_1 - z_2 j) = \begin{pmatrix} \bar{z}_1 & -z_2 \\ \bar{z}_2 & z_1 \end{pmatrix} = \overline{\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix}}^t = \overline{\rho(w)}^t.$$

Daraus folgt

$$\rho(\overline{w\omega}) = \overline{\rho(w\omega)}^t = \overline{\rho(w)\rho(\omega)}^t = \overline{\rho(\omega)}^t \overline{\rho(w)}^t = \rho(\bar{\omega})\rho(\bar{w}) = \rho(\bar{\omega} \cdot \bar{w}),$$

also, weil ρ injektiv ist,

$$\overline{w\omega} = \bar{\omega} \cdot \bar{w}.$$

Definition B.8. Der Unterraum der rein imaginären Quaternionen ist das Komplement

$$V = \langle i, j, k \rangle_{\mathbb{R}} \subseteq \mathbb{H}.$$

von $\mathbb{R} \subseteq \mathbb{H}$. Die Eigenraumzerlegung der Involution $w \mapsto \bar{w}$ ist gerade $\mathbb{H} = \mathbb{R} \oplus V$.

Bemerkung B.9. Zu $z \in \mathbb{H}^\times = \mathbb{H} \setminus \{0\}$ betrachten wir die Konjugation

$$w \mapsto zwz^{-1}.$$

Für $w \in V$ gilt

$$\overline{zwz^{-1}} = \bar{z}^{-1}\bar{w}\bar{z} = zN(z)^{-1}(-w)N(z)z^{-1} = -(zwz^{-1}).$$

Daher liefert $z(-)z^{-1}|_V$ eine lineare Abbildung $V \to V$. Das Inverse ist die Konjugation mit z^{-1}, folglich ergibt sich ein Gruppenhomomorphismus

$$\sigma : \mathbb{H}^\times \to \mathrm{GL}(V), \quad z \mapsto \sigma_z = z(-)z^{-1}|_V.$$

Wegen

$$N(\sigma_z(w)) = N(zwz^{-1}) = N(z)N(w)N(z)^{-1} = N(w)$$

erhält σ_z die Norm von $w \in V$. Wir lernen in der Vorlesung Geometrie, daß eine solche normerhaltende lineare Abbildung eine Drehung oder eine Spiegelung von $V \simeq \mathbb{R}^3$ (bezüglich der Basis i, j, k) ist. Insbesondere ist $\det(\sigma_z) = \pm 1$. Dies hängt stetig von $z \in \mathbb{H}^\times$ ab. Weil wir jedes σ_z durch einen stetigen Weg in $\mathbb{H}^\times$ mit 1 verbinden können, gilt

$$\det(\sigma_z) = \det(\sigma_1) = 1.$$

Zusammengenommen haben wir einen Teil der folgenden Resultats bewiesen:

Satz B.10. *Die Abbildung $z \mapsto \sigma_z$ von oben definiert einen surjektiven Gruppenhomomorphismus*

$$\sigma : \mathbb{H}^\times \to \mathrm{SO}_3(\mathbb{R})$$

auf die dreidimensionale Drehgruppe $\mathrm{SO}_3(\mathbb{R})$, wobei $V \simeq \mathbb{R}^3$ mittels der Basis i, j, k identifiziert wurde. Der Kern von σ ist $\ker(\sigma) = \mathbb{R}^\times$.

Bemerkung B.11. Angeblich werden in Vektorgraphiken des Computers Rotationen mittels solcher Konjugationen durch Quaternionen beschrieben.

Bemerkung B.12. Die Untergruppe von $\mathbb{H}^\times$ der Quaternionen von Norm 1, also

$$\mathbb{H}^{N=1} = \ker(N : \mathbb{H}^\times \to \mathbb{R}^\times)$$

erzeugt zusammen mit $\ker(\sigma)$ die Gruppe $\mathbb{H}^\times$: jedes Quaternion ist Produkt eines zentralen Quaternions aus $\mathbb{R}^\times = \ker(\sigma)$ und eines Quaternions von Norm 1. Daher ist die Einschränkung von σ immer noch eine surjektive Abbildung

$$\sigma : \mathbb{H}^{N=1} \to \mathrm{SO}_3(\mathbb{R}).$$

Der Kern besteht nun nur noch aus ± 1. Die Gestalt von $\mathrm{SO}_3(\mathbb{R})$ kann nun verstanden werden als Paare von Vektoren der Norm 1 in $\mathbb{R}^4 \simeq \mathbb{H}$, eine 3-Sphäre, bis auf ± 1: das ist der reell-projektive Raum der Dimension 3:

$$\mathrm{SO}_3(\mathbb{R}) \approx \mathrm{S}^3 / \{\pm 1\} \approx \mathbb{RP}^3.$$

Bemerkung B.13. Das Zentrum eines Schiefkörpers D ist die Menge

$$Z(D) = \{x \in D \ ; \ xy = yx \text{ für alle } y \in D\},$$

die mit allen Elementen von D kommutiert. Man überzeugt sich schnell, daß das Zentrum ein Unterkörper ist. Das Zentrum von $\mathbb{H}$ ist

$$Z(\mathbb{H}) = \mathbb{R}.$$

Satz B.14. *Die einzigen Schiefkörper D, die den Körper $\mathbb{R}$ in ihrem Zentrum enthalten haben, und als solche ein endlichdimensionaler $\mathbb{R}$-Verktorraum sind, sind die folgenden:*

$$\mathbb{R}, \mathbb{C} \text{ und } \mathbb{H}.$$

Übungsaufgaben zu §B

Übungsaufgabe B.1. Zeigen Sie, daß es keinen $\mathbb{R}$-Vektorraum der Dimension 3 gibt, auf dem eine nullteilerfreie und $\mathbb{R}$-lineare Multiplikation existiert.

Anhang C. Das Zornsche Lemma

Um zu beweisen, daß jeder Vektorraum eine Basis hat, brauchen wir das Lemma von Zorn.

C.1. **Das Auswahlaxiom.** Zu einer Menge M bezeichnen wir mit

$$\mathscr{P}(M)^\times = \{S \subseteq M \ ; \ S \neq \emptyset\}$$

die Menge aller nichtleeren Teilmengen von M, also die Potenzmenge von M ohne die leere Menge.

Definition C.1. Eine **Auswahlfunktion** auf einer Menge M ist eine Abbildung

$$f : \mathscr{P}(M)^\times \to M$$

mit der Eigenschaft, daß

$$f(A) \in A$$

für alle $A \in \mathscr{P}(M)^\times$. Die Funktion f stellt also für jede nichtleere Teilmenge A von M ein Element aus A bereit.

Auswahlfunktionen sind nichts anderes als Elemente des kartesischen Produkts von Mengen

$$\prod_{\emptyset \neq A \subseteq M} A.$$

Axiom 1 (Auswahlaxiom)**.** Zu jeder Menge gibt es eine Auswahlfunktion.

Mittels der Interpretation als Elemente des kartesischen Produkts besagt das Auswahlaxiom also letztlich nur, daß das Produkt nichtleerer Mengen wieder nichtleer ist.

Dies klingt plausibel, muß aber im Rahmen der Mengenlehre axiomatisch gefordert werden. Wir verweisen auf mathematische Logik zur Klärung der logischen Zusammenhänge und beschränken uns hier darauf, aus dem Auswahlaxiom das nützliche Lemma von Zorn zu beweisen.

C.2. **Partielle, induktive Ordnungen.**

Definition C.2. (1) Eine **(partielle) Ordnung** auf einer Menge M ist eine Relation $\preceq$ auf der Menge M, so daß für alle $x, y, z \in M$ gilt:
   (i) transitiv: Wenn $x \preceq y$ und $y \preceq z$, dann gilt $x \preceq z$.
   (ii) antisymmetrisch: Wenn $x \preceq y$ und $y \preceq x$, dann gilt $x = y$.
   (iii) reflexiv: $x \preceq x$.

(2) Eine **totale Ordnung** auf einer Menge M ist eine partielle Ordnung auf M, so daß zusätzlich
   (iv) $x \preceq y$ oder $y \preceq x$ gilt.

Beispiel C.3. (1) Die Potenzmenge einer Menge M hat eine partielle Ordnung: die Inklusion $\subseteq$. Diese partielle Ordnung ist im Allgemeinen nicht total geordnet.
(2) Die Menge $\mathbb{Z}$ ist bezüglich der üblichen $\leq$-Relation total geordnet.

Definition C.4. (1) Sei M bezüglich $\preceq$ partiell geordnet. Ein Element $x \in M$ heißt **obere Schranke** für die Teilmenge $S \subseteq M$, wenn $y \preceq x$ für alle $y \in S$ gilt.
(2) Die Menge M heißt bezüglich der Ordnung $\preceq$ **induktiv geordnet**, wenn jede total geordnete Teilmenge $S \subseteq M$ eine obere Schranke in M besitzt.
(3) Ein Element $x \in M$ einer bezüglich $\preceq$ partiell geordneten Menge M heißt **maximales Element**, wenn für alle $y \in M$ mit $x \preceq y$ schon $x = y$ gilt.

Beispiel C.5. (1) Sei M eine Menge. Die Menge der Teilmengen von M ist bezüglich der Inklusion induktiv geordnet. Eine obere Schranke ergibt sich als Vereinigung.
(2) Sei M eine Menge. Die Menge der echten Teilmengen $U \subset M$, also $U \neq M$, ist bezüglich Inklusion partiell geordnet. Für jedes $a \in M$ ist $U_a = M \setminus \{a\}$ ein maximales Element.

Axiom 2 (Lemma von Zorn)**.** Sei M eine
(i) nicht-leere,
(ii) bezüglich $\preceq$ induktiv geordnete Menge.

Dann hat M bezüglich $\preceq$ ein maximales Element.

Notation C.6. Sei M eine bezüglich $\preceq$ partiell geordnete Menge, $x, y \in M$ und $S \subseteq M$ eine Teilmenge.
(1) Mit $x \prec y$ bezeichnen wir '$x \preceq y$ und $x \neq y$'.
(2) Weiter setzen wir

$$S_{\preceq x} = \{s \in S \ ; \ s \preceq x\}$$
$$S_{\prec x} = \{s \in S \ ; \ s \prec x\}.$$

Man kann nun zeigen, daß das Lemma von Zorn aus dem Auswahlaxiom folgt. In Wahrheit sind die beiden sogar logisch äquivalent (und auch noch äquivalent zum Wohlordnungssatz, auf den wir hier nicht weiter eingehen).

Theorem C.7. *Aus dem Auswahlaxiom folgt das Lemma von Zorn.*

C.3. **Existenz einer Basis für allgemeine Vektorräume.** Wir benutzten nun das Lemma von Zorn, um auch in unendlichdimensionalen Vektorräumen die Existenz einer Basis sicherzustellen.

Theorem C.8. *Jeder Vektorraum hat eine Basis.*

Beweis. Seien K ein Körper, V ein K-Vektorraum. Dann betrachten wir die Menge bestehend aus allen linear unabhängigen Teilmengen von Vektoren aus V:

$$\mathscr{M} = \{A \subseteq V \ ; \quad A \text{ ist linear unabhängig}\}.$$

Es ist $\mathscr{M}$ nicht leer, weil $\emptyset$ als Element in $\mathscr{M}$ enthalten ist, denn die leere Menge ist ja linear unabhängig. Außerdem ist $\mathscr{M}$ bezüglich Inklusion $\subseteq$ partiell geordnet. Diese Ordnung ist sogar induktiv, denn bei einer totalgeordneten Teilmenge von $A_i \in \mathscr{M}$ für $i \in I$ und eine Indexmange I ist auch

$$A := \bigcup_{i \in I} A_i$$

in $\mathscr{M}$ enthalten. Dies zeigen wir durch Widerspruch. Wäre A plötzlich linear abhängig, dann gibt es eine nichttriviale Linearkombination $a_1 v_1 + \ldots + a_n v_n = 0$, die das bezeugt. Die endlich vielen Vektoren $v_1, \ldots, v_n$ liegen in irgendwelchen der A_i, sagen wir $i_r \in I$ für $r = 1, \ldots, n$ und

$$v_r \in A_{i_r}.$$

Da aber die A_i's bezüglich $\subseteq$ total geordnet sind, gibt es unter den $A_{i_1}, \ldots, A_{i_n}$ eine bezüglich Inklusion größte Menge, in der dann alle v_r enthalten sind. Die Linearkombination kann also schon mit Vektoren von dort gebildet werden, und das ist ein Widerspruch dazu, daß alle A_i linear unabhängig sind.

Nach dem Lemma von Zorn gibt es also in $\mathscr{M}$ maximale Elemente. Jedes solche ist eine Basis, wie man Satz 5.18 entnimmt. Damit ist die Existenz einer Basis bewiesen. □

Anhang D. Affine Räume

Es gibt eine abstrakte Definition affiner Räume über einem Körper K. Wir beschreiben nur affine Unterräume eines K-Vektorraums.

Definition D.1. Sei K ein Körper und V ein K-Vektorraum.

(1) Zu einer Teilmenge $A \subseteq V$ setzen wir

$$U(A) = \langle v - w \ ; \ v, w \in A \rangle_K$$

für den von den Differenzen aus Vektoren aus A erzeugten Unterraum.

(2) Eine Teilmenge $A \subseteq V$ eines K-Vektorraums V ist ein **affiner** Unterraum, wenn $A = \emptyset$ oder

$$A = a_0 + U(A) = \{a_0 + v \ ; \ v \in U(A)\}$$

für ein (äquivalent dazu alle) $a_0 \in A$. In diesem Fall nennen wir $U(A)$ den **Translationsraum** von A.

(3) Als **Dimension** eines nichtleeren affinen Unterraums $A \subseteq V$ bezeichnen wir

$$\dim(A) := \dim U(A).$$

Beispiel D.2. Sei V ein K-Vektorraum mit Unterraum U. Die affinen Unterräume $A \subseteq V$ mit $U(A) = U$ sind gerade die Urbilder der Quotientenabbildung

$$q_U : V \to V/U.$$

In der Tat gilt für ein Urbild $a \in V$ von $\bar{a} := q_U(a) \in V/U$, daß

$$q_U^{-1}(\bar{a}) = \{x \in V \ ; \ a \sim_U x\} = a + U.$$

Dies zeigt auch die in der Definition implizit enthaltene Behauptung, daß $A = a_0 + U(A)$ für ein $a_0 \in A$ dieselbe Aussage für alle $a_0 \in A$ impliziert. Für die Äquivalenzklassen von $\sim_U$ ist das nämlich offensichtlich.

Beispiel D.3. (1) Im $\mathbb{R}$-Vektorraum $\mathbb{R}^3$ sind die affinen Unterräume:
- Punkte: $\{P\}$ für $P \in \mathbb{R}^3$. Die Dimension ist 0 und $U(P) = (0)$.
- Geraden: Zu $a \in \mathbb{R}^3$ und $0 \neq v \in \mathbb{R}^3$ die Geraden
$$L = a + Kv.$$
Die Dimension ist 1 und $U(L) = \langle v \rangle_{\mathbb{R}}$.
- Ebenen: Zu $a \in \mathbb{R}^3$ und linear unabhängigen Vektoren $v, w \in \mathbb{R}^3$ die Ebene
$$E = a + Kv + Kw.$$
Die Dimension ist 2 und $U(E) = \langle v, w \rangle_{\mathbb{R}}$.
- Der ganze $\mathbb{R}^3$: als einziger affiner Unterraum der Dimension 3. Es gilt $U(\mathbb{R}^3) = \mathbb{R}^3$.

(2) Sei $A \in \mathrm{M}_{n\times m}(K)$ und $b \in K^n$. Dann ist der Lösungsraum
$$L = \mathscr{L}(AX = b)$$
des inhomogenen linearen Gleichungssystems $AX = b$, falls er nicht leer ist, ein affiner Unterraum von K^m der Dimension $\dim \ker(A) = m - \mathrm{rg}(A)$ und $U(L) = \ker(A)$.

(3) Sei $f : V \to W$ eine lineare Abbildung. Dann sind die Urbildmengen $f^{-1}(y)$ zu $y \in W$ affine Unterräume von V. Der Translationsraum ist $\ker(f)$, sofern $f^{-1}(y)$ nicht leer ist.

Definition D.4. Eine **Translation** eines Vektorraums V ist eine Abbildung
$$T : V \to V,$$
so daß mit $a = T(0)$ für alle $v \in V$ gilt
$$T(v) = v + a.$$
Man sagt genauer, T ist die Translation um den Vektor a. Wir verwenden die Notation $T_a(v) = v + a$.

Bemerkung D.5. Die einzige Translation, die eine lineare Abbildung ist, ist die Translation um den Vektor 0.

Proposition D.6. *Sei V ein K-Vektorraum.*

(1) Die Komposition von Translationen von V ist wieder eine Translation von V. Genauer: zu $a, b \in V$ gilt
$$T_a \circ T_b = T_{a+b}$$

(2) Translationen sind bijektiv, und die inverse Abbildung ist die Translation mit dem negativen Vektor: für alle $a \in V$ gilt
$$T_a \circ T_{-a} = \mathrm{id}_V = T_{-a} \circ T_a.$$

(3) Die Menge der Translationen von V bildet mit Komposition eine Gruppe, die vermöge
$$a \mapsto T_a$$
isomorph zur additiven Gruppe $(V, +)$ des Vektorraums ist.

Beweis. Das folgt alles sofort aus den Definitionen. □

Bemerkung D.7.

(1) Jeder affine Unterraum $A \subseteq V$ ist das Bild des Translationsraums $U(A)$ unter der Translation um einen (jeden) Vektor $a \in A$.

(2) Die Translation $T_x : V \to V$ mit $x \in V$ bildet einen affinen Unterraum $A \subseteq V$ in sich (und dann automatisch bijektiv auf sich) ab genau dann, wenn $x \in U(A)$.

Übungsaufgaben zu §D

Übungsaufgabe D.1. Beweisen Sie die Behauptungen von Bemerkung D.7.

Übungsaufgabe D.2. Zeigen Sie, daß der Schnitt affiner Unterräume wieder ein affiner Unterraum ist.

Anhang E. PageRank

Der Erfolg der Suchmaschine Google basiert auf der Berechnung eines Eigenvektors. Der zugehörige Algorithmus heißt **PageRank**. Das Internet ist ein **gerichteter Graph**, das ist eine kombinatorische Struktur mit **Ecken**

$$W = \text{ Menge der Webseiten im Internet}$$

und mit **gerichteten Kanten** von $j \in W$ nach $i \in W$, wenn die Seite j auf die Seite i verlinkt.

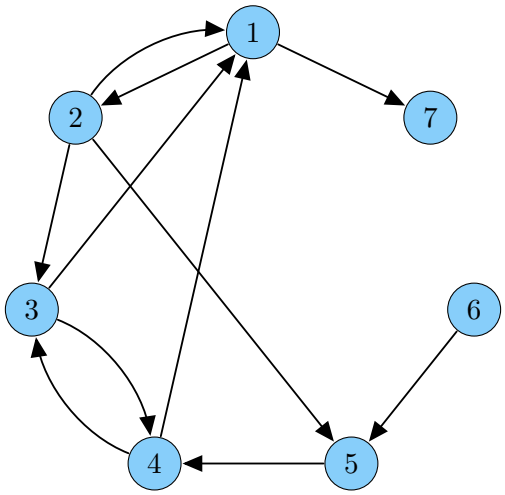


Abbildung 11. Bild eines gerichteten Graphen.

Sei $n = |W|$ die Anzahl der Webseiten. Wir definieren die **Nachbarschaftsmatrix (adjacency matrix)** des Graphen als $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ mit

$$a_{ij} = \begin{cases} 1 & j \text{ verlinkt auf } i \\ 0 & j \text{ verlinkt nicht auf } i. \end{cases}$$

Dabei haben wir stillschweigend die Webseiten mit einer Nummerierung der Webseiten gleichgesetzt. Im obigen Graphen ist A die Matrix

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

Die Spaltensummen von A sind

$$N_j = \sum_i a_{ij} = |\{\text{Links, die von } j \text{ ausgehen}\}|.$$

Wir nehmen der Einfachheit halber an, daß es keine toten Enden gibt: alle $N_j > 0$.

Wir wollen nun die Webseiten bewerten, und zwar soll die Wertung w_i der Seite $i \in W$ proportional sein zu den Wertungen der Webseiten j, die auf i verlinken. Dabei soll die Reputation w_j der Webseite j aber nur mit dem Gewicht $1/N_j$ eingehen. Die Reputation wird somit gleichmäßig auf die gelinkten Webseiten verteilt. Wir fordern also einen Proportionalitätsfaktor $\lambda \in \mathbb{R}$, so daß für alle $i \in W$

$$w_i = \lambda \cdot \sum_j a_{ij} \cdot \frac{w_j}{N_j}. \tag{E.1}$$

Die Wertungen sollen $w_i \geq 0$ und durch Skalieren normiert sein auf

$$\sum_i w_i = 1.$$

Sei $w \in \mathbb{R}^n$ der Spaltenvektor der Bewertungen. Dann können wir (E.1) mit

$$P = (\frac{a_{ij}}{N_j}) = A \cdot \operatorname{diag}(\frac{1}{N_1}, \ldots, \frac{1}{N_j}, \ldots) \in \mathrm{M}_n(\mathbb{R})$$

als

$$w = \lambda \cdot Pw.$$

Die Matrix P hat eine besondere Eigenschaft: sie ist eine stochastische Matrix.

Definition E.1. Eine **stochastische Matrix** ist eine quadratische Matrix mit Einträgen aus $\mathbb{R}$, die nicht negativ sind und die sich in jeder Spalte zu 1 addieren.

Lemma E.2. *Die Matrix P ist eine stochastische Matrix.*

Beweis. Dazu rechnet man für alle $j \in W$

$$\sum_i P_{ij} = \sum_i \frac{a_{ij}}{N_j} = \frac{1}{N_j} \sum_i a_{ij} = 1. \qquad \square$$

Lemma E.3. *Für jedes $x \in \mathbb{R}^n$ stimmen die Summen der Koordinaten von x und Px überein.*

Beweis. Seien x_i die Koordinaten von x und $(Px)_i$ die Koordinaten von Px. Dann ist

$$\sum_i (Px)_i = \sum_i \sum_j P_{ij} x_j = \sum_j x_j \sum_i P_{ij} = \sum_j x_j. \qquad \square$$

Korollar E.4. *Der Proportionalitätsfaktor in (E.1) muß $\lambda = 1$ sein.*

Beweis. Nach Annahme ist

$$1 = \sum_i w_i = \sum_i \lambda \cdot (Pw)_i = \lambda \cdot \sum_i (Pw)_i = \lambda \sum_i w_i = \lambda. \qquad \square$$

Der gesuchte Bewertungsvektor w ist also nichts anderes als ein (auf Summe der Koordinaten gleich 1 normierter) Eigenvektor von P zum Eigenwert 1:

$$w = Pw.$$

Wir überlegen uns zuerst, daß 1 tatsächlich ein Eigenwert von P ist. Da wir das Internet nicht als gerichteten Graphen und damit auch P nicht explizit kennen, muß das aus anderen allgemeinen Gründen folgen.

Proposition E.5. *Sei K ein Körper und sei $B \in \mathrm{M}_n(K)$ eine beliebige Matrix. Dann haben B und B^t die gleichen Eigenwerte.*

Beweis. Weil transponierte Matrizen die gleiche Determinante haben, ist

$$\chi_{B^t}(X) = \det(X \cdot \mathbf{1} - B^t) = \det((X \cdot \mathbf{1} - B)^t) = \det(X \cdot \mathbf{1} - B) = \chi_B(X).$$

Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms. □

Proposition E.6. *Jede stochastische Matrix Q hat den Eigenwert 1. Jeder andere komplexe Eigenwert λ von Q hat $|\lambda| \leq 1$.*

Beweis. Nach Proposition E.5 reicht es zu zeigen, daß Q^t den Eigenwert 1 hat. Sei e der Spaltenvektor, der in jeder Koordinate den Eintrag 1 hat. Dann ist

$$(Q^t e)_i = \sum_j Q^t_{ij} = \sum_j Q_{ji} = 1,$$

weil Q stochastisch ist. Damit ist $Q^t e = e$ und $e \neq 0$ ein Eigenvektor zum Eigenwert 1.

Sei nun λ ein beliebiger komplexer Eigenwert von Q^t und $x \in \mathbb{C}^n$ ein zugehöriger Eigenvektor. Wir wählen einen Index i, so daß die Koordinate x_i betragsmäßig maximal ist: $|x_j| \leq |x_i|$ für alle $j = 1, \ldots, n$. Dann ist

$$|\lambda| \cdot |x_i| = |\lambda x_i| = |(Q^t x)_i| = |\sum_j Q^t_{ij} x_j| \leq \sum_j Q^t_{ij} |x_j| \leq \Big(\sum_j Q^t_{ij}\Big)|x_i| = \Big(\sum_j Q_{ji}\Big)|x_i| = |x_i|.$$

Weil $x_i \neq 0$, folgt sofort $|\lambda| \leq 1$, also mit Proposition E.5 die Behauptung. □

Es bleibt das praktische Problem, wie man den Eigenvektor einer so großen Matrix P denn berechnet ($n \approx 2 \cdot 10^9$, Stand 2018). Das zugehörige Gleichungssystem zu lösen, kann man vergessen. Wir nehmen der Einfachheit halber an, daß P als Matrix in $\mathrm{M}_n(\mathbb{C})$ diagonalisierbar ist und $v_1, \ldots, v_n$ eine Basis von $\mathbb{C}^n$ aus Eigenvektoren von P zu den Eigenwerten $1 = \lambda_1, \ldots, \lambda_n$ ist. Wir nehmen weiter vereinfachend an, daß $|\lambda_i| < 1$ für alle $i \geq 2$.

Wir wählen nun einen Startvektor y_0, entwickeln in der Basis aus Eigenvektoren zu

$$y_0 = z_1 v_1 + \ldots + z_n v_n$$

und können fast sicher annehmen, daß die Koordinate z_1 nicht verschwindet. Dann gilt

$$y_k := P^k y_0 = z_1 (\lambda_1)^k v_1 + \ldots + z_n (\lambda_n)^k v_n. \tag{E.2}$$

Die Koordinate von v_1 bleibt $z_1 \lambda_1^k = z_1$, während für alle $i \geq 2$ wegen

$$z_i (\lambda_i)^k \to 0 \qquad \text{für } k \to \infty$$

die übrigen Koordinaten schnell gegen 0 konvergieren. Damit ist

$$y_k \approx z_1 v_1$$

für große k, etwa $k = 100$ liest man, eine gute Approximation an den gesuchten Eigenvektor v_1.

Wir interpretieren diese Rechnung als Beschreibung einer zufälligen Irrfahrt auf dem gerichteten Graphen des Internets. Ein Vektor $y \in \mathbb{R}^n$ beschreibe eine Liste der Wahrscheinlichkeiten y_i, mit der wir uns auf der Webseite $i \in W$ aufhalten. (Es gilt dann $0 \leq y_i \leq 1$ und $\sum_i y_i = 1$, denn irgendwo halten wir uns ja auf.) In jedem Schritt der Irrfahrt entscheiden wir uns ausgehend von der Seite j mit gleicher Wahrscheinlichkeit für einen der Links auf der Seite j. Damit ist die Wahrscheinlichkeit im nächsten Schritt auf der Seite i zu landen

$$\sum_j P_{ij} y_j = (Py)_i.$$

Die neue Verteilung der Wahrscheinlichkeiten ist somit nichts anderes als

$$Py.$$

Wenden wir das wiederholt an, so beschreibt der Vektor $P^k y$ die Verteilung der Aufenthaltswahrscheinlichkeiten nach k angeklickten Links. Die Folge

$$y, Py, P^2y, \dots, P^k y, \dots$$

beschreibt also die Wahrscheinlichkeitsverteilung bei **zufälliger Irrfahrt** (so heißt das wirklich) im Internet. Aus obiger Rechnung folgt unter unseren Annahmen, daß diese Wahrscheinlichkeitsverteilung gegen den auf Summe geich 1 (die Gesamtwahrscheinlichkeit ist 1) normierten Eigenvektor w zum Eigenwert 1 konvergiert. Dies ist der Fixpunkt der Iteration und wird **stationäre Verteilung** genannt.

Wir erhalten so eine neue Konzeption für die Bewertung von Webseiten: die Bewertung einer Webseite i soll die Aufenthaltswahrscheinlichkeit w_i auf der Seite i im Limes für eine lange zufällige Irrfahrt auf dem Internet sein.

Die Interpretation der Irrfahrt führt zu einer Variation. Das typische Surfverhalten startet mit einer gewissen Wahrscheinlichkeit (man liest 15%) bei jedem Schritt zufällig auf einer beliebigen Seite des Internets. Das führt zur folgenden modifizierten Transformationsmatrix

$$T = (1-d)\frac{1}{n}E + dP,$$

wobei $d \in [0,1]$ die Wahrscheinlichkeit ist, über einen Link weiterzuwandern, und E die Matrix ist, die an jeder Stelle eine 1 hat. Der Summand dP beschreibt das bisherige Modell des Surfens ohne Neuanfang, und der Summand $(1-d)\frac{1}{n}E$ beschreibt den zufälligen Neuanfang auf einer beliebigen Seite.

Die Matrix T ist weiter stochastisch und hat überdies ausschließlich positive Einträge, wenn $0 \leq d < 1$. Dann interessieren wir uns wieder für den Eigenvektor w zum Eigenwert 1 und bestimmen diesen mit Hilfe des Interationsverfahrens. Als Startwert benutzt man zum Beispiel den Eigenvektor des gestrigen Tages. Der Parameter d sorgt dafür, daß nun der Satz von Perron und Frobenius unsere Annahmen rechtfertigt.

Satz E.7 (Frobenius–Perron; einfache Version von Perron)**.** *Sei $T \in \mathrm{M}_n(\mathbb{R})$ mit lauter positiven Einträgen $T_{ij} > 0$. Dann hat T einen reellen Eigenwert $\lambda_{\max} > 0$ und einen zugehörigen Eigenvektor w, so daß gilt:*

(1) Die geometrische Vielfachheit von $\lambda_{\max}$ ist 1 und man kann w so wählen, daß alle Koordinaten positiv sind und die Summe der Koordinaten 1 ist.

(2) Jeder andere komplexe Eigenwert λ von T hat $|\lambda| < \lambda_{\max}$.

Die Annahme der Diagonalisierbarkeit brauchen wir nicht, wenn wir die Rechnung (E.2) zur Iteration

$$y, Ty, T^2y, \dots, T^k y, \dots$$

durch eine Rechnung in der Jordannormalform der Matrix T ersetzen. Wesentlich ist, daß der dominante Eigenwert, also der betragsmäßig größte Eigenwert, nur mit Multiplizität 1 auftritt. Der Beitrag der anderen Jordankästchen konvergiert gegen 0, und zwar umso schneller je größer die **spektrale Lücke**

$$\lambda_{\max} - \max\{|\lambda| \; ; \lambda \text{ Eigenwert }, |\lambda| < \lambda_{\max}\}$$

ist.

Anhang F. Polynome über den komplexe Zahlen

Die besondere Bedeutung der Körpererweiterung $\mathbb{R} \subseteq \mathbb{C}$ ist nicht zu unterschätzen:

Theorem F.1 (Fundamentalsatz der Algebra)**.** *Jedes nicht-konstante Polynom mit komplexen Koeffizienten hat mindestens eine komplexe Nullstelle: für jedes $P(X) \in \mathbb{C}[X]$ mit $\deg(P) \geq 1$ gibt es $z_0 \in \mathbb{C}$ mit $P(z_0) = 0$.*

Es gibt einen schönen Beweis mittels Funktionentheorie, den wir nicht behandeln, und einen Beweis mittels Galois-Theorie, der zum Stoff der Vorlesung Algebra gehört.

Per Induktion nach dem Grad beweisen wir aus dem Fundamentalsatz der Algebra die folgende präzisere Aussage.

Satz F.2. *Jedes Polynom mit komplexen Koeffizienten zerfällt in $\mathbb{C}[X]$ in ein Produkt aus Linearfaktoren.*

Beweis. Sei $P \in \mathbb{C}[X]$ beliebig. Für $\deg(P) \leq 0$ ist P konstant und die Aussage des Satzes ist klar. Wenn $\deg(P) \geq 1$ gilt, dann gibt es nach Theorem F.1 eine Nullstelle $z_0 \in \mathbb{C}$. Korollar 12.20 liefert $Q(X) \in \mathbb{C}[X]$ mit

$$P(X) = Q(X)(X - z_0).$$

und

$$\deg(Q) = \deg(P) - \deg(X - z_0) = \deg(P) - 1 < \deg(P).$$

Die Behauptung folgt nun per Induktion nach dem Grad des Polynoms. In der Tat, gilt der Satz für $Q(X)$, dann gilt er auch für $P(X)$. □

Aus Theorem F.1 leiten wir einen wichtigen algebraischen Fakt ab.

Theorem F.3. *Die irreduziblen Polynome in $\mathbb{R}[X]$ sind entweder linear oder quadratisch. Genauer sind die irreduziblen normierten Polynome in $\mathbb{R}[X]$ von der Form*

$$\begin{aligned} &X - \lambda, && \lambda \in \mathbb{R} \\ &X^2 - (\alpha + \bar{\alpha})X + \alpha\bar{\alpha}, && \alpha \in \mathbb{C} \setminus \mathbb{R}. \end{aligned}$$

Beweis. Sei $f(X) \in \mathbb{R}[X]$ irreduzibel und normiert. Nach dem Fundamentalsatz der Algebra, siehe Theorem F.1, hat jedes Polynom mit komplexen Koeffizienten (und damit auch mit reellen Koeffizienten) eine Nullstelle: es gibt ein $\alpha \in \mathbb{C}$ mit $f(\alpha) = 0$. Wenn $\alpha \in \mathbb{R}$ sogar reell ist, dann ist $X - \alpha$ ein Teiler von $f(X)$ in $\mathbb{R}[X]$ und damit

$$f(X) = X - \alpha.$$

Sein nun $\alpha \in \mathbb{C} \setminus \mathbb{R}$. Dann setzen wir mit dem komplex konjugierten $\bar{\alpha}$

$$h(X) = (X - \alpha)(X - \bar{\alpha}) = X^2 - (\alpha + \bar{\alpha})X + \alpha\bar{\alpha}.$$

Sei $d(X)$ der normierte ggT von $h(X)$ und $f(X)$. Dann gibt es nach dem Satz von Bézout Polynome $a(X), b(X) \in \mathbb{R}[X]$ mit

$$d(X) = a(X)h(X) + b(X)f(X).$$

Da $f(X)$ irreduzibel (also prim) ist, gibt es genau 2 Möglichkeiten: entweder $d(X) = 1$ oder $d(X) = f(X)$. Da

$$d(\alpha) = a(\alpha)h(\alpha) + b(\alpha)f(\alpha) = 0$$

scheidet die erste Möglichkeit aus. Damit ist $f(X) = d(X)$ ein Teiler von $h(X)$. Weil $h(X)$ nur die Lösungen $\alpha, \bar{\alpha} \in \mathbb{C} \setminus \mathbb{R}$ hat, kann $f(X)$ nicht linear sein, also

$$f(X) = h(X) = X^2 - (\alpha + \bar{\alpha})X + \alpha\bar{\alpha}.$$ □

Jakob Stix, Institut für Mathematik, Goethe–Universität Frankfurt, Robert-Mayer-Str. 6–8, 60325 Frankfurt am Main, Germany
Email address: stix@math.uni-frankfurt.de