Lineare Algebra

Studienjahr 2021/22

Christoph Schweigert
Universität Hamburg
Department Mathematik
Schwerpunkt Algebra und Zahlentheorie

(Stand: 05.01.2026)

Inhaltsverzeichnis

Literatur:

Literatur, die ich bei der Vorbereitung häufig herangezogen habe:

- Christian Bär: Lineare Algebra und analytische Geometrie Springer Fachmedien Wiesbaden, 2018. Volltextzugang Campus
  `https://doi.org/10.1007/978-3-658-22620-6`
- Catherine Meusburger: Skript zu den Vorlesungen "Lineare Algebra I und II"
  `https://www.math.fau.de/lie-gruppen/personen/catherine-meusburger/lehre/vorlesungsskripten/`
- Gerd Fischer, Boris Springborn: Lineare Algebra, Vieweg, 2025 Volltextzugang Campus
  `https://doi.org/10.1007/978-3-662-71261-0`
- Für Kapitel 7: Falko Lorenz: Lineare Algebra II, Spektrum Akad. Verl. 2008

Die aktuelle Version dieses Skriptes finden Sie unter
`https://christophschweigert.github.io/skripten/laskript.pdf` als pdf-Datei.

Bitte schicken Sie Korrekturen und Bemerkungen an `christoph.schweigert@uni-hamburg.de`!

Bei Frau N. Potylitsina-Kube möchte ich mich für Ihre große Hilfe bei der Erstellung dieses Skriptes und bei den Hamburger Studenten, insbesondere bei Max Demirdilek, Miká Kruschel und Marc Lange, für zahlreiche Hinweise bedanken.

1 Vorbereitung

In diesem Kapitel werden wir vorbereitende Betrachtungen durchführen: wir fangen ganz konkret an mit ein wenig elementarer Geometrie der Ebene. Konkrete Probleme führen dabei oft auf lineare Gleichungssysteme, deren systematische Lösung wir anschließend behandeln. Der letzte Teil des Kapitels ist etwas abstrakter: wir führen die Sprache von Mengen und Abbildungen ein, die auf Aussagen und der Verknüpfung von Aussagen beruht.

1.1 Geometrie von Geraden in der Ebene

Wir setzen in diesem einleitenden Kapitel voraus, dass Sie Folgendes wissen:

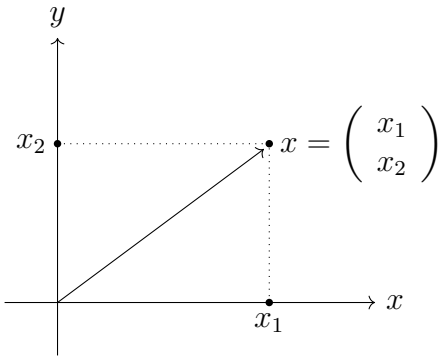
- Sie haben eine Vorstellung, was reelle Zahlen sind. In der Analysis wird dies noch einmal präzise eingeführt werden. Wir bezeichnen die Gesamtheit der reellen Zahlen mit $\mathbb{R}$. Wir sprechen auch von der *Menge* der reellen Zahlen. Wir können entscheiden, ob etwas ein Element der Menge der reellen Zahlen ist oder nicht.

- Reelle Zahlen können addiert und multipliziert werden. Für Addition und Multiplikation beliebiger reeller Zahlen gelten Assoziativ- und Kommutativgesetze. Es gibt eine reelle Zahl $0 \in \mathbb{R}$, so dass für alle reellen Zahlen, also alle $a \in \mathbb{R}$, gilt $0 + a = a + 0 = a$; für die Zahl $1 \in \mathbb{R}$ gilt bei Multiplikation $1 \cdot a = a \cdot 1 = a$ für alle $a \in \mathbb{R}$. Sie wissen sicher auch, welche Eigenschaften für eine gegebene reelle Zahl $a \in \mathbb{R}$ die Zahlen $-a$ und, falls $a \neq 0$ gilt, $1/a$ haben.

- Sie wissen auch, dass es auf $\mathbb{R}$ die Ordnungsrelationen $\geq$ oder $\leq$ gibt – diese werden in dieser Vorlesung zwar nur eine untergeordnete Rolle spielen, sind aber wichtig für die Analysis.

- Wir setzen voraus, dass sie die Veranschaulichung der reellen Zahlen als Punkte auf der Zahlengerade kennen.

Wir können nun die Menge $\mathbb{R}^2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1, x_2 \in \mathbb{R} \right\}$ von *geordneten* Paaren reeller Zahlen betrachten. Durch die Einführung kartesischer Koordinaten können wir diese als mathematisches Modell für die Ebene unserer Anschauung sehen:

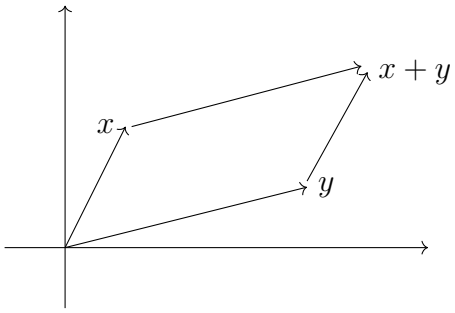


Wir wollen auf die Menge $\mathbb{R}^2$ noch andere Struktur aufprägen: Je zwei Elementen $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$,

$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ können wir durch komponentenweise Addition ihre Summe

$$x + y := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \in \mathbb{R}^2$$

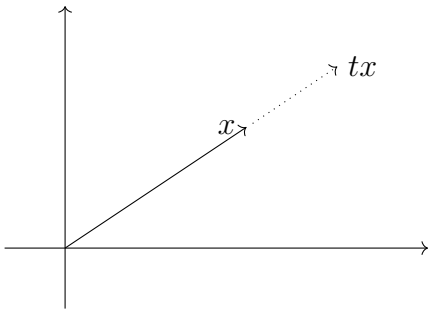
zuordnen. Doppelpunkte := werden immer anzeigen, dass ein Ausdruck auf der linken Seite definiert wird. Das Gleichheitszeichen = dagegen macht eine Aussage über schon definierte Größen. Wir stellen den Vektor $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ durch einen Pfeil vom Ursprung mit Spitze im Punkt mit Koordinaten (x_1, x_2) dar. Die Summe wird dann so veranschaulicht



Für jede reelle Zahl $t \in \mathbb{R}$ können wir durch Multiplikation aller Komponenten den Vektor um einen Faktor t strecken,

$$t \cdot x := \begin{pmatrix} tx_1 \\ tx_2 \end{pmatrix} \ .$$

Wir schreiben auch kurz tx. Bildlich für $t > 0$:



Vielleicht erwarten Sie hier Beispiele wie:

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} \qquad \text{und} \quad 3 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

Allerdings wird in Mathematikvorlesungen sehr schnell von Ihnen erwartet, dass Sie sich solche einfachen Beispiele selbst verschaffen. Man liest mathematische Literatur daher immer mit Papier und Stift zur Hand. Machen Sie sich auch ein Bild für $t < 0$. Manchmal werden in der Vorlesung solche Bilder auch spontan oder auf Nachfrage gezeichnet. Dieses Skript hat *nicht* den Anspruch, die Vorlesung ersetzen zu können. Daher dürfen solche Bilder auch im Skript fehlen.

Bemerkungen 1.1.1.

1. *Für alle* $x, y, z \in \mathbb{R}^2$ und *für alle* $t, t' \in \mathbb{R}$ gilt

   (a) $(x + y) + z = x + (y + z)$ [Assoziativität]

(b) Sei $0 := \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbb{R}^2$ der sogenannten Nullvektor. Dann gilt:
$0 + x = x = x + 0$ [Neutrales Element]
Man beachte, dass wir mit dem gleichen Symbol die reelle Zahl $0 \in \mathbb{R}$ und den Nullvektor $0 \in \mathbb{R}^2$ bezeichnen. Mathematische Formeln erschließen sich *nur* aus dem Kontext. Dieser Kontext ist auch wichtig, weil verschiedene Autoren verschiedene Konventionen benutzen.

(c) Zu jedem $x \in \mathbb{R}^2$ gibt es ein $-x \in \mathbb{R}^2$, so dass

$$x + (-x) = (-x) + x = 0 \,,$$

nämlich $-x = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$ [Additives Inverses]

(d) $x + y = y + x$ [Kommutativität]

(e) $(tt')x = t(t'x)$. (Man mache sich hier genau klar, welche Verknüpfung bei tt' und welche bei $t'x$ gemeint ist!)

(f) $1x = x$

(g) $t(x + y) = tx + ty$

(h) $(t + t')x = tx + t'x$. (Man mache sich hier genau klar, welche Verknüpfung mit $+$ bei $t + t'$ und welche bei $tx + t'x$ gemeint ist!)

2. Alle Gleichungen werden gezeigt, in dem man sie durch Betrachtung von Komponenten auf die entsprechenden Gesetze für die reellen Zahlen zurückführt. Ein Beispiel:

$$x + y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} = \begin{pmatrix} y_1 + x_1 \\ y_2 + x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = y + x \,.$$

Solche kleinen Argumente werden wir nicht immer im Skript ausführen, aber manchmal spontan oder auf Nachfrage beispielhaft in der Vorlesung.

3. Man beachte, dass es sich nicht um Aussagen über einzelne, spezielle $t, t' \in \mathbb{R}$ oder $x, y, z \in \mathbb{R}^2$ handelt. Wir haben vielmehr z.B. in 1. eine Aussage, die von den drei Elementen $x, y, z \in \mathbb{R}^2$ abhängt. Diese Aussagen sind dann für *alle* möglichen Wahlen solche Elemente wahr.

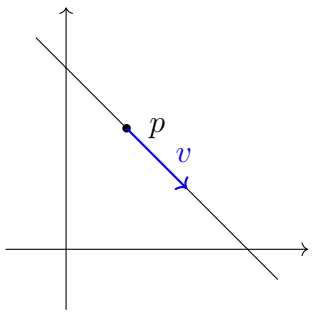
Allgemeiner hätten wir für beliebiges $n \in \mathbb{N} \setminus \{0\}$ auch $\mathbb{R}^n$, also die Menge aller geordneten n-Tupel reeller Zahlen, betrachten können, wobei insbesondere $\mathbb{R}^3$ ein Modell für den dreidimensionalen Raum liefert. Wir nennen auch ein Element $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ einen Vektor (in $\mathbb{R}^n$) und für $i = 1, 2, \ldots, n$ die reelle Zahl v_i die i-te Komponente oder Koordinate. Wir geben $\mathbb{R}^0$ in diesem einleitenden Kapitel noch keinen Sinn. Der $\mathbb{R}^n$ tritt in vielen Anwendungen auf; unsere Methoden sind so beschaffen, dass sie nicht vom Wert von n abhängen. (Wir werden uns auch vom Fall $\mathbb{R}$ frei machen.) Abstraktion in der Mathematik führt eigentlich immer zu größerer Anwendbarkeit.

Durch *Definitionen* schafft man sich neue mathematische Begriffe. $\mathbb{R}^2$ ist eine Menge und hat eine Klasse interessanter Teilmengen:

Definition 1.1.2
Seien $p, v \in \mathbb{R}^n$, $v \neq 0$. Dann heißt die Teilmenge von $\mathbb{R}^n$ der Form

$$G_{p,v} := p + \mathbb{R}v := \{p + \lambda v |\ \lambda \in \mathbb{R}\} \,,$$

die (affine) Gerade durch den Fußpunkt p mit Richtungsvektor v. Die Darstellung $p+\mathbb{R}v$ heißt Parameterdarstellung von $G_{p,v}$.



Machen Sie sich bitte klar, dass $v \neq 0$ bedeutet, dass die Komponenten v_1 und v_2 nicht beide Null sein dürfen.

In einer Definition unterstreichen wir die Begriffe, die definiert werden. Der Ausdruck $\{p + \lambda v | \lambda \in \mathbb{R}\}$ ist so zu lesen: wir betrachten die Teilmenge von $\mathbb{R}^2$, die aus den Elementen besteht, für die es ein $\lambda \in \mathbb{R}$ gibt, so dass sich das Element als $p + \lambda v$ schreiben lässt. Der Ausdruck $p+\mathbb{R}v$ sollte nicht so verstanden werden, dass man eine neue Addition einführt. Er ist, wie $G_{p,v}$, nur ein graphisches Hilfsmittel, eine Teilmenge zu bezeichnen, die durch p und v beschrieben wird.

Wir bringen noch einmal an einer solchen Stelle ein Beispiel: $G_{0,(1,1)}$ ist die Winkelhalbierende des ersten und dritten Quadranten. Bei einer solchen Definition sollte man sich sofort überlegen, warum $v = 0$ nicht zugelassen wurde! (Wie sähe denn $G_{p,0}$ aus?)

Wir fangen mit einer Hilfsaussage an. Die Tatsache, dass ein Autor eine Aussage als Hilfsaussage einschätzt, wird durch die Verwendung des Worts *Lemma* ausgedrückt.

Lemma 1.1.3.
Für $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ sowie für $p, q \in \mathbb{R}^n$ gilt:

Es gilt $G_{p,v} = G_{q,w}$ genau dann, wenn $q \in G_{p,v}$ gilt und es ein $\mu \in \mathbb{R} \setminus \{0\}$ gibt mit $w = \mu v$

Insbesondere ist die Parameterdarstellung einer gegebenen Geraden $G \subset \mathbb{R}^n$ nicht eindeutig. Es gibt mehrere Wahlen von (p, v), die die gleiche Gerade (also die gleiche Teilmenge von $\mathbb{R}^n$) beschreiben. (Oft ist es schwer, mathematische Objekte ohne Redundanz zu parametrisieren.)

Eine mathematische Aussage erfordert immer einen *Beweis*, in dem diese Aussage auf bekannte Aussagen zurückgeführt wird. Der Verweis auf ein Bild, auch wenn er für die Anschauung hilfreich ist, ist kein Beweis.

Beweis.

- Schauen wir uns genau an, was wir beweisen wollen. Es gibt zwei Aussagen, die von $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ abhängen:
  "Aussage 1" für $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$: es gilt $G_{p,v} = G_{q,w}$
  "Aussage 2" für $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$: es ist $q \in G_{p,v}$ und es gibt ein $\mu \in \mathbb{R} \setminus \{0\}$ mit $w = \mu v$.

  - Für eine gegebene Wahl von $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ sind die Aussagen entweder wahr oder falsch. Jede sinnvolle Aussage hat einen Wahrheitswert. Andere Wahrheitswerte als "wahr" oder "falsch" werden wir nicht betrachten.
  - In Aussage 2 werden zwei Teilaussage durch das Wort "und" zu einer neuen Aussage verknüpft. Diese neue Aussage ist genau dann wahr, wenn die beiden ursprünglichen Aussagen wahr sind. Dies ist die Definition der Verknüpfung "und".

– Wir behaupten, dass die Aussagen für *jede* erlaubte Wahl von v, w, p, q entweder beide wahr oder beide falsch sind. Wir sagen dann, dass die Aussagen 1 und 2 äquivalent sind und schreiben dann

$$\text{“Aussage 1”} \Leftrightarrow \text{“Aussage 2”}$$

und sagen, die beiden Aussagen sind äquivalent.

Dazu zeigen wir zweierlei: zum einen: ist Aussage 1 wahr, dann ist auch Aussage 2 wahr. Man schreibt dann ‘

$$\text{‘Aussage 1”} \Rightarrow \text{“Aussage 2”}.$$

(Wir machen hier keine Aussage darüber, was los ist, wenn Aussage 1 falsch ist!) Zum anderen zeigen wir: ist Aussage 2 wahr, so ist auch Aussage 1 wahr. Von den vier möglichen Kombinationen in der sogenannten Wahrheitstafel

Aussage 1	Aussage 2
wahr	wahr
wahr	falsch
falsch	wahr
falsch	falsch

eliminiert der erste Teil des Beweises, den wir mit ” $\Rightarrow$ ” bezeichnen, die zweite Zeile; der zweite Teil ” $\Leftarrow$ ” eliminiert die dritte Zeile. Damit ist die Äquivalenz gezeigt.

- ” $\Rightarrow$ ”. Es sollen also $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ so gewählt sein, dass Aussage 1 gilt. Wir nehmen also an, dass $G_{p,v} = G_{q,w}$ gilt. Aus $q \in G_{q,w}$ folgt mit Aussage 1 $G_{q,w} = G_{p,v}$ direkt die Aussage $q \in G_{p,v}$. Also gibt es ein $\mu_1 \in \mathbb{R}$ mit $q = p + \mu_1 v$. Ferner gilt auch $q + w \in G_{q,w} = G_{p,v}$, also gibt auch es ein $\mu_2 \in \mathbb{R}$ mit $q + w = p + \mu_2 v$. Es folgt

$$w = (q + w) - q = (p + \mu_1 v) - (p - \mu_2 v) = (\mu_2 - \mu_1)v \, .$$

Wir müssen noch zeigen, dass $\mu_2 - \mu_1 \neq 0$ ist. Dazu nehmen wir an, es würde $\mu_2 - \mu_1 = 0$ gelten. Wir nehmen also das *Gegenteil* dessen an, was wir beweisen wollen. Aus $\mu_2 - \mu_1 = 0$ folgt aber $w = 0v = 0$. Dies ist im Widerspruch zu unserer Voraussetzung, dass $w \neq 0$ ist. Damit kann unsere unsere *Widerspruchsannahme* $\mu_2 - \mu_1 = 0$ nicht gelten und die Behauptung ist gezeigt.

Dieses Argument ist ein einfaches Beispiel für einen *indirekten Beweis*, auch *Widerspruchsbeweis* genannt.

Warum haben wir die Aussage *für alle* $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ gezeigt? Wir haben im Beweis keine weiteren Annahmen über v, w, p, q verwendet, als die, dass dies Elemente in $\mathbb{R}^n$ sind und v und w nicht der Nullvektor sind. Eine solche Aussage über alle Elemente mit gewissen Eigenschaften kann man also keinesfalls beweisen, indem man nur ein oder einige Beispiel betrachtet - sei es mit Zahlenwerten oder in einer Zeichnung.

- ” $\Leftarrow$ ”. Es soll also für eine gewisse erlaubte Wahl von $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ Aussage 2 gelten. Zu zeigen ist die Gleichheit der zwei Teilmengen $G_{p,v}$ und $G_{q,w}$ von $\mathbb{R}^2$. Damit wir diese Aussage für alle Wahlen zeigen, dürfen wir auch wieder nichts anderes für v, w, p, q verwenden, als dass dies Elemente im $\mathbb{R}^n$ sind und $v \neq 0$ und $w \neq 0$ gilt.

- Wir überlegen uns erst, dass zwei Teilmengen T_1, T_2 einer Menge M, in Zeichen $T_1 \subset M$ und $T_2 \subset M$, genau dann gleich sind, wenn sie die gleichen Elemente von M enthalten. (Das ist eigentlich die Definition von Gleichheit von Teilmengen einer Menge, die wir aber in diesem einleitenden Abschnitt naiv verwenden.) Ist jedes Element von T_1 auch in T_2, so ist $T_1 \subset T_2$ Teilmenge. Es gilt also genau dann $T_1 = T_2$, wenn die Aussagen $T_1 \subset T_2$ und $T_2 \subset T_1$ *beide* gelten.

- Wir wollen zuerst die Inklusion $G_{q,w} \subset G_{p,v}$ zeigen.
  Wegen unserer Annahme, dass Aussage 2 gilt, haben wir $q \in G_{p,v}$. Also gibt es ein $\lambda_0 \in \mathbb{R}$, so dass $q = p + \lambda_0 v$ gilt. Sei $s \in G_{q,w}$ beliebig, also gilt $s = q + \lambda_1 w$ mit einem geeigneten [1] $\lambda_1 \in \mathbb{R}$. Einsetzen liefert

  $$s = p + \lambda_0 v + \lambda_1 \mu v = p + (\lambda_0 + \lambda_1 \mu) v \, ,$$

  woraus $s \in G_{p,v}$ folgt. Wir hatten über s keine weiteren Annahmen gemacht, als dass $s \in G_{q,w}$ gilt. Dann ist aber auch $s \in G_{p,v}$. Alle Elemente von $G_{q,w}$ sind somit auch Elemente von $G_{p,v}$ und wir haben die Inklusion $G_{q,w} \subset G_{p,v}$ gezeigt.

- Wir müssen noch die umgekehrte Inklusion $G_{p,v} \subset G_{q,w}$ zeigen.
  Sei $s \in G_{p,v}$ beliebig, also $s = p + \lambda_0 v$ mit einem geeigneten $\lambda_0 \in \mathbb{R}$. Aus $q \in G_{p,v}$ folgt, dass es $\lambda_1 \in \mathbb{R}$ gibt, so dass $q = p + \lambda_1 v$ gilt. Durch Umstellen folgt $p = q - \lambda_1 v$ mit $\lambda_1 \in \mathbb{R}$. Also

  $$s = p + \lambda_0 v = q - \lambda_1 v + \lambda_0 v = q + (\lambda_0 - \lambda_1) \mu^{-1} w \, .$$

  Man beachte, dass wir hier die Annahme $\mu \neq 0$ ausgenutzt haben. Hieraus folgt $s \in G_{q,w}$. Wie im vorhergehenden Spiegelpunkt sehen wir, dass wir die Inklusion $G_{p,v} \subset G_{q,w}$ gezeigt haben.

□

Beispiel 1.1.4.
Als weiteres Beispiel für einen indirekten Beweis beweisen wir die Aussage:
$\sqrt{2}$ ist irrational: es gibt keine rationale Zahl q mit $q^2 = 2$.

Wir setzen dabei als bekannt voraus, dass sich jede rationale Zahl als gekürzter Bruch beschreiben lässt und dass sich jede natürliche Zahl außer der 1 in Primfaktoren zerlegen lässt.

Angenommen es gibt eine rationale Zahl q mit $q^2 = 2$. Dann können wir q als gekürzten Bruch ausdrücken, $q = \frac{a}{b}$ mit teilerfremden ganzen Zahlen a, b. Aus $q^2 = 2$ folgt $a^2 = 2b^2$. Die Primzahl 2 teilt das Produkt a^2 und damit einen der Faktoren, also a. Also ist a gerade, schreibe $a = 2c$. Einsetzen in die Gleichung $a^2 = 2b^2$ liefert $b^2 = 2c^2$. Nach dem gleichen Argument ist dann aber auch b gerade, im Widerspruch dazu, dass der Bruch $q = \frac{a}{b}$ gekürzt ist. Also kann es keine rationale Zahl q mit $q^2 = 2$ geben.

Lemma 1.1.5.
Sei $G \subset \mathbb{R}^n$ eine Gerade und seien $a, b \in G$ und $a \neq b$. Dann ist $G = G_{a,b-a}$. Eine Gerade wird also durch zwei verschiedene Punkte, die auf ihr liegen, festgelegt.

Beweis.

[1] Das ist mathematischer Slang: "mit einem geeigneten…" ist eine Existenzaussage und gleichbedeutend mit "es gibt ein …".

- Aus Lemma 1.1.3 folgt, dass wir einen beliebigen Punkt auf G, also insbesondere a, als Fußpunkt wählen können. Es ist also $G = G_{a,v}$ mit einem geeignetem Richtungsvektor $v \in \mathbb{R}^n \setminus \{0\}$, den wir uns noch aus a, b beschaffen müssen. Aus $b \in G_{a,v}$ folgt, dass es ein $t_0 \in \mathbb{R}$ gibt mit $b = a + t_0 v$.
- Wir wollen zeigen, dass $t_0 \neq 0$ gilt. Dazu nehmen wir an, es würde $t_0 = 0$ gelten. Daraus würde aber $b = a + t_0 v = a + 0v = a + 0 = a$ folgen, im Widerspruch zur Voraussetzung $a \neq b$. Also kann unsere Annahme $t_0 = 0$ nicht gelten und ist zum Widerspruch geführt.
- Damit haben wir aber

  $$a \in G_{a,b-a} \quad \text{und} \quad b - a = t_0 v \quad \text{mit} \quad t_0 \neq 0\,,$$

  mit Lemma 1.1.3 also $G_{a,b-a} = G_{a,v}$.

□

Auf $\mathbb{R}^2$ – allgemeiner auf $\mathbb{R}^n$ – kann man noch eine weitere Struktur einführen. Sie wird im ersten Teil der Vorlesung keine Rolle spielen. Sie ermöglicht es uns aber in diesem einleitenden Kapitel den Anschluss an geometrische Anschauung und einfachere Schreibweisen.

Definition 1.1.6

1. *Je zwei Vektoren* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *und* $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ *des* $\mathbb{R}^2$ *ordnen wir ihr Skalarprodukt zu:*

   $$\langle x, y \rangle := x_1 y_1 + x_2 y_2 \in \mathbb{R}\,.$$

2. *Die Norm eines Elements* $x \in \mathbb{R}^2$ *ist definiert durch*

   $$\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{(x_1)^2 + (x_2)^2} \in \mathbb{R}_{\geq 0}\,.$$

Man mache sich an Hand der Zeichnung auf Seite 1 mit Hilfe des Satzes des Pythagoras klar, dass $\|x\|$ die Länge des Vektors $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, also der Abstand des Punktes mit Koordinaten $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ vom Ursprung ist.

Bemerkung 1.1.7.

1. Wir schreiben die Zuordnung, die Paaren von Vektoren ihr Skalarprodukt zuordnet, auch so:

   $$\begin{array}{rccc} \langle \cdot, \cdot \rangle : & \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ & (x, y) & \mapsto & \langle x, y \rangle \end{array}$$

   Man beachte die unterschiedlichen Pfeilsymbole. Sie hat die folgenden Eigenschaften, die man leicht nachrechnet: für alle $x, x', y \in \mathbb{R}^2$ und $t \in \mathbb{R}$ gilt

   (a) $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$ [Additivität im ersten Argument]

   (b) $\langle tx, y \rangle = t\langle x, y \rangle$

   (c) $\langle x, y \rangle = \langle y, x \rangle$ [Symmetrie des Skalarprodukts]

   (d) $\langle x, x \rangle \geq 0$ und $\langle x, x \rangle = 0 \Leftrightarrow x = 0$. [Das Skalarprodukt ist positiv definit.]

Man wird in einer Vorlesung nie alle solchen Aussagen nachrechnen können. Hier ist Ihre eigene Nacharbeit gefordert! Sie werden auch nicht hier im Skript alle Kommentare, die zu solchen Gleichungen in der Vorlesung gemacht werden, nachlesen können.

2. Es gilt für die Norm aller $x \in \mathbb{R}^2$ die Ungleichung $\|x\| \geq 0$ und Gleichheit $\|x\| = 0$ genau für $x = 0$. Ferner gilt für alle $x \in \mathbb{R}^2$ und $t \in \mathbb{R}$

$$\|tx\| = \sqrt{\langle tx, tx\rangle} = \sqrt{t^2\langle x, x\rangle} = |t|\sqrt{\langle x, x\rangle} = |t|\|x\| \,,$$

wobei wir die Identität $\sqrt{r^2} = |r|$ für $r \in \mathbb{R}$ ausgenutzt haben.

Schätzt ein Autor eine mathematische Aussage als etwas wichtiger ein, so nennt er sie üblicherweise nicht Lemma, sondern *Satz*. *Theoreme* sind Sätze, die der Autor für besonders wichtig hält.

Satz 1.1.8.

1. Für alle $x, y \in \mathbb{R}^2$ gilt die sogenannte Cauchy–Schwarz'sche Ungleichung

$$\Big|\langle x, y\rangle\Big| \leq \|x\| \cdot \|y\|$$

2. Für alle $x, y \in \mathbb{R}^2$ gilt

$$\|x + y\| \leq \|x\| + \|y\| \qquad \text{[Subadditivität der Norm]}$$

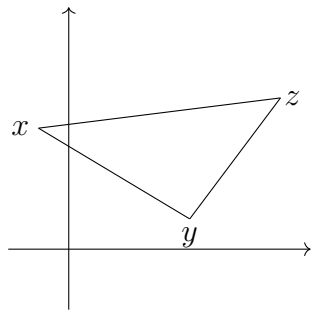
3. Für den euklidischen Abstand

$$d(x, y) := \|x - y\|$$

gilt die Dreiecksungleichung

$$d(x, z) \leq d(x, y) + d(y, z)$$

für alle $x, y, z \in \mathbb{R}^2$.



Beweis.

1. Wir rechnen:

$$\Big(\|x\|\|y\|\Big)^2 - \langle x, y\rangle^2 = \langle x, x\rangle\langle y, y\rangle - \langle x, y\rangle^2 = (x_1^2 + x_2^2)(y_1^2 + y_2^2) - (x_1y_1 + x_2y_2)^2$$
$$= x_1^2y_2^2 + x_2^2y_1^2 - 2x_1y_1x_2y_2 = (x_1y_2 - x_2y_1)^2 \geq 0 \,.$$

Hieraus folgt die Ungleichung $\Big(\|x\|\|y\|\Big)^2 \geq \langle x, y\rangle^2$; Wurzelziehen aus nicht-negativen reellen Zahlen erhält Ungleichungen; unter Beachtung der Regel $\sqrt{r^2} = |r|$ für $r \in \mathbb{R}$ finden wir die Cauchy-Schwarzsche Ungleichung:

$$\|x\| \cdot \|y\| \geq |\langle x, y\rangle| \,.$$

Man beachte auf der rechten Seite, dass die Norm nicht-negativ ist und daher die Betragsstriche nicht nötig sind.

Nebenbemerkung: es ist bei Ungleichungen oft sehr instruktiv, sich zu fragen, wann genau Gleichheit gilt. Dies ist hier der Fall, wenn $x_1y_2 = x_2y_2$ gilt. Das gilt für $x = 0$ der Fall. Ist $x \neq 0$ und $x_1 \neq 0$, so folgt $y_2 = \frac{y_1}{x_1}x_2$ und damit $y = \frac{y_1}{x_1}(x_1, x_2)$. Analog schließt man für $x_2 \neq 0$, dass $y = \frac{y_2}{x_2}(x_1, x_2)$ gelten muss. Die Gleichheit gilt also genau dann, wenn die Vektoren x, y kollinear sind.

2. Dies folgt durch Wurzelziehen aus

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \qquad \text{[Symmetrie des Skalarprodukts]} \\ &\leq \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 \qquad \text{[nach Cauchy-Schwarz, also 1.]} \\ &= (\|x\| + \|y\|)^2 \,.\end{aligned}$$

3. Wir rechnen:

$$\begin{aligned}d(x,z) &= \|x-z\| = \|x-y+y-z\| \\ &\leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z) \qquad \text{[nach 2.]}\end{aligned}$$

□

Wir können jetzt mehr Geometrie machen, weil wir einen Abstandsbegriff haben.

Lemma 1.1.9.
Sei $G \subset \mathbb{R}^2$ eine Gerade und seien $a, b \in G$ mit $a \neq b$. Dann existiert genau ein Punkt $c \in G$, so dass gilt $d(a,c) = d(b,c)$. Für diesen Punkt c gilt ferner

$$c = \frac{1}{2}(a+b) \qquad \text{und} \quad d(a,c) = d(b,c) = \frac{1}{2}d(a,b)\,.$$

Beweis.
Man beachte, dass wir hier eine *Existenzaussage* und eine *Eindeutigkeitsaussage* machen.
Um die Existenzaussage zu beweisen, müssen wir nur nachrechnen, dass $c := \frac{1}{2}(a+b)$ die gewünschten Eigenschaften hat:

$$d(a,c) = \|\frac{1}{2}(a+b) - a\| = \|\frac{1}{2}(b-a)\| = \|\frac{1}{2}(a-b)\| = \|\frac{1}{2}(a+b) - b\| = d(b,c)\,.$$

Da G durch a, b nach Lemma 1.1.5 festgelegt ist, $G = G_{a,b-a}$, folgt aus $c = a + \frac{1}{2}(b-a)$ auch, dass $c \in G$ gilt. Es ist für den Beweis sogar unerheblich zuwissen, wie man auf den Punkt c kommt. Das ist für einen Existenzbeweis logisch nicht notwendig. Diese Darstellungsweise ist ökonomisch, erfordert aber beim Lesen mathematischer Texte eine große Anstrengung.

Die Eindeutigkeitsaussage ist etwas aufwändiger: für $t \in \mathbb{R}$ betrachten wir den Punkt $c(t) := a + t(b-a)$. So erhalten wir alle Punkte auf der Gerade durch a und b. Für den Punkt $c(t)$ gilt

$$d(a,c(t)) = \|c(t)-a\| = \ |t|\,\|b-a\| \quad \text{und} \quad d(b,c(t)) = |t-1|\,\|b-a\|$$

(So etwas müssen Sie dann selbst nachrechnen!). Für Gleichheit brauchen wir $|t| = |t-1|$, was die eindeutige Lösung $t = \frac{1}{2}$ hat. (Nachrechnen!) □

Bemerkung 1.1.10.

1. Der vorangehende Beweis ist ein Beispiel für einen konstruktiven Beweis. Man zeigt, dass ein mathematisches Objekt existiert, indem man es direkt angibt oder eine Konstruktionsvorschrift gibt, die es erlaubt, es zu konstruieren. (Eine solche Konstruktionsvorschrift muss nicht unbedingt rechnerisch effizient sein. Mathematische Objekte effizient zu konstruieren ist aber eine sinnvolle mathematische Aufgabe!)

2. Hier ist ein weiteres Beispiel für einen konstruktiven Beweis. Wir zeigen, dass das Quadrat einer geraden ganzen Zahl n gerade ist. Wir können $n = 2k$ schreiben, woraus $n^2 = (2k)^2 = 2(2k^2)$ folgt. Wir haben explizit angegeben, wie sich n^2 als das Doppelte einer ganzen Zahl schreiben lässt.

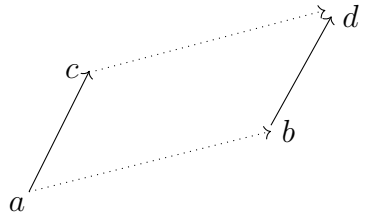
3. Ein nicht-konstruktiver Beweis zeigt, dass das Polynom $p(x) = x^5 - x^4 + 2x^3 + x - 1$ eine Nullstelle im offenen Intervall $(-1, 1)$ hat. Denn es ist $p(1) = 2 > 0$ und $p(-1) = -6 < 0$. Die Existenz, aber keine Konstruktionsvorschrift, folgt aus dem Zwischenwertsatz der Analysis.

Lemma 1.1.9 motiviert die folgende Definition, für die man eigentlich nicht den Begriff des Skalarprodukts braucht:

Definition 1.1.11

1. *Gegeben zwei Punkte $a, b \in \mathbb{R}^2$, so heißt der Punkt $\frac{1}{2}(a + b)$ Mittelpunkt von a und b.*

2. *Ein Parallelogramm ist ein 4–Tupel (a, b, c, d) von Punkten in $\mathbb{R}^2$, so dass $c - a = d - b$ gilt:*



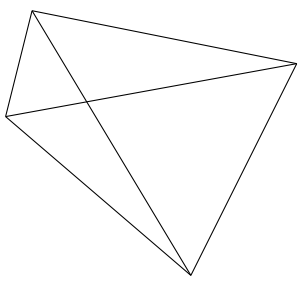
   Es folgt dann auch $b - a = d - c$. (Nachrechnen! Geometrisch veranschaulichen!)

3. *Ein Parallelogramm heißt nicht-ausgeartet, falls keine drei Punkte auf einer Geraden liegen.*

Satz 1.1.12 (Diagonalensatz)**.**
In einem nicht-ausgearteten Parallelogramm halbieren sich die Diagonalen gegenseitig.

Man beachte, dass mit dieser Formulierung eine Aussage für *alle* nicht-ausgearteten Parallelogramme gemacht wird.

Für allgemeine Vierecke ist die Aussage nicht unbedingt richtig, etwa:



(Bei der Nachbereitung einer Mathematik-Vorlesung muss man sich auch mit Gegenbeispielen beschäftigen, sich diese auch einmal selbst verschaffen! Die Betrachtung von Gegenbeispielen schafft oft auch Ansatzpunkte und Überprüfungsmöglichkeiten für Beweise.)

Beweis.
Der Mittelpunkt der Diagonale von a nach d ist $\frac{1}{2}(a+d)$, der Mittelpunkt der Diagonale von b nach c ist $\frac{1}{2}(b+c)$.
Wir rechnen:

$$\frac{1}{2}(a+d) - \frac{1}{2}(b+c) = \frac{1}{2}(a+d-b-c) = 0\,.$$

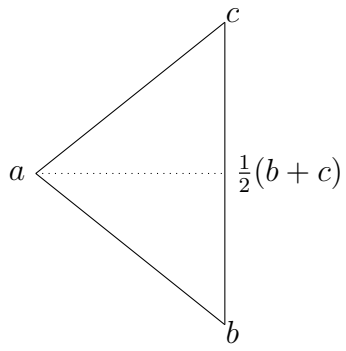
Also sind die Mittelpunkte gleich und der Schnittpunkt der beiden Diagonalen. □

Auch für den folgenden Satz wird weder für Aussage noch für Beweis das Skalarprodukt auf $\mathbb{R}^2$ benötigt.

Definition 1.1.13

1. *Ein Dreieck ist ein Tripel (a,b,c) von Punkten in $\mathbb{R}^2$. Es heißt nicht-ausgeartet, falls die Eckpunkte a,b,c nicht auf einer Geraden liegen.*

2. *Sei (a,b,c) ein nicht-ausgeartetes Dreieck. Eine Seitenhalbierende ist eine Gerade durch eine der Ecken und den Mittelpunkt der gegenüberliegenden Seite:*



Satz 1.1.14 (Schwerpunktsatz)**.**
In einem nicht-ausgearteten Dreieck (a,b,c) schneiden sich alle drei Seitenhalbierenden in dem Punkt $\frac{1}{3}(a+b+c)$, dem Schwerpunkt des Dreiecks. Die Seitenhalbierenden zerlegen sich im Verhältnis $1:2$.

Dies ist wieder ein Satz über *alle* nicht-ausgearteten Dreiecke.

Beweis.

- Die Seitenhalbierende durch a enthält a und den Seitenmittelpunkt $\frac{b+c}{2}$. Sie ist in Parameterform wegen Lemma 1.1.5 gegeben durch

$$a + \mathbb{R}\left(\frac{1}{2}(b+c) - a\right)$$

Wähle den Parameter $t = \frac{2}{3}$ und finde auf dieser Geraden den Punkt

$$q = a + \frac{2}{3}\left(\frac{1}{2}(b+c) - a\right) = \frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c\,.$$

Der Ausdruck ist symmetrisch in a, b, c, man kann also die Rollen von a, b und c vertauschen. Also liegt q auch auf den anderen beiden Seitenhalbierenden. Man beachte, dass für die erste Aussage des Satzes das Skalarprodukt nicht benötigt wird. Der Ausdruck $\frac{a+b+c}{3}$ erklärt auch die Benennung des Schnittpunkts als Schwerpunkt.

- Bemerkung: Man kann das Problem des Schnitts der drei Seitenhalbierenden auch so umformulieren: suche t_1, t_2 und $t_3 \in \mathbb{R}$, so dass gilt:

$$a + t_1(\frac{b+c}{2} - a) = b + t_2(\frac{c+a}{2} - b)$$
$$a + t_1(\frac{b+c}{2} - a) = c + t_3(\frac{a+b}{2} - c)$$

  Dies ist äquivalent zu einem System von 4 (warum vier?) inhomogenen linearen Gleichungen für die 3 reellen Unbestimmten t_1, t_2 und t_3:

$$\frac{b+c-2a}{2}t_1 + \frac{2b-c-a}{2}t_2 = b - a$$
$$\frac{b+c-2a}{2}t_1 + \frac{2c-a-b}{2}t_3 = c - a$$

- Für den letzten Teil der Aussage brauchen wir das Skalarprodukt, denn wir sprechen von Längen. Wir berechnen wir den Abstand vom Schwerpunkt $\frac{a+b+c}{3}$ zum Seitenmittelpunkt $\frac{b+c}{2}$ zu

$$d\left(\frac{a+b+c}{3}, \frac{b+c}{2}\right) = \left\|\frac{a+b+c}{3} - \frac{b+c}{2}\right\| = \left\|\frac{a}{3} - \frac{b}{6} - \frac{c}{6}\right\| = \frac{1}{6}\|2a - b - c\|$$

  und den Abstand vom Schwerpunkt zur Ecke a zu

$$d\left(\frac{a+b+c}{3}, a\right) = \left\|\frac{a+b+c}{3} - a\right\| = \left\|-\frac{2}{3}a + \frac{b}{3} + \frac{c}{3}\right\| = \frac{1}{3}\|2a - b - c\|$$

□

Wir bringen noch einige ergänzende Bemerkungen zur geometrischen Interpretation des Skalarprodukts: wegen der Cauchy-Schwarz'schen Ungleichung 1.1.8.1 ist

$$\frac{|\langle x, y\rangle|}{\|x\|\|y\|} \leq 1 \qquad \text{für alle } x, y \in \mathbb{R}^2 \text{ mit } \quad x \neq 0 \text{ und } \; y \neq 0$$

Definition 1.1.15

1. *Seien $x, y \in \mathbb{R}^2$, $x \neq 0$ und $y \neq 0$. Dann heißt*

$$\alpha(x, y) := \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|} \in [0, \pi]$$

   der Innenwinkel der Vektoren x und y.

2. *Gilt $\alpha(x, y) = \frac{\pi}{2}$, so heißen x und y orthogonal. Dies ist genau für $\langle x, y\rangle = 0$ der Fall.*

Wir wollen den Schnitt von Geraden in der Ebene untersuchen. Dazu setzen wir erst einmal hilfsweise für $a = (a_1, a_2) \in \mathbb{R}^2$

$$a^\perp := (-a_2, a_1) \in \mathbb{R}^2 \, .$$

(Wir wollen im Interesse eines übersichtlicheren Schriftsatzes hier nicht Spalten- und Zeilenvektoren unterscheiden.) Beschreiben Sie $a^\perp$ geometrisch! Dann gilt für $a, b \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$, wie man am besten selbst nachrechnen sollte:

1. $(a+b)^\perp = a^\perp + b^\perp \qquad (\lambda a)^\perp = \lambda a^\perp$
2. $\langle a, b^\perp \rangle = -\langle a^\perp, b \rangle$, also insbesondere $\langle a, a^\perp \rangle = 0$, was für $a \neq 0$ heißt, dass $a^\perp$ auf a senkrecht steht. Ferner gilt

$$\|a^\perp\| = \|a\| \qquad \text{und} \quad (a^\perp)^\perp = -a \, .$$

Betrachtung 1.1.16.
Wir betrachten den Schnitt $G_{p,a} \cap G_{q,b}$ zweier Geraden $G_{p,a}$ und $G_{q,b}$ im $\mathbb{R}^2$. Dieser besteht aus allen Elementen $s \in \mathbb{R}^2$, für die $s \in G_{p,a}$ und $s \in G_{q,b}$ gilt. Wir schreiben dann $s \in G_{p,a} \cap G_{q,b}$; gilt dies, so gibt es $\lambda, \mu \in \mathbb{R}$ mit

$$s = p + \lambda a = q + \mu b \, .$$

- Dies schreiben wir in der Form $\lambda a + \mu(-b) = q - p$. Die beiden Komponenten liefern zwei lineare Gleichungen für die zwei Unbestimmten λ, μ. Wir werden solche linearen Gleichungssysteme in Kapitel 1.2 systematisch untersuchen.
- Hier ist ein ad hoc Lösungsweg: wir bilden das Skalarprodukt mit $a^\perp$ und erhalten

$$\langle p, a^\perp \rangle = \langle q, a^\perp \rangle + \mu \langle b, a^\perp \rangle \qquad .(*)$$

  Dies ist eine lineare Gleichung in einer Variable μ.

  – Ist $\langle b, a^\perp \rangle \neq 0$, so gibt es für μ die eindeutige Lösung

$$\mu = \frac{\langle p - q, a^\perp \rangle}{\langle b, a^\perp \rangle} \, ,$$

    die zum Schnittpunkt

$$q + \frac{\langle p - q, a^\perp \rangle}{\langle b, a^\perp \rangle} b$$

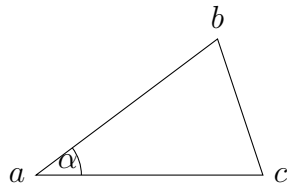
    führt.

  – Ist $\langle b, a^\perp \rangle = 0$, so sieht man
    * Für $\langle p - q, a^\perp \rangle \neq 0$ gibt es keinen Schnittpunkt. Die Geraden sind parallel und verschieden.
    * Für $\langle p - q, a^\perp \rangle = 0$ ist jedes $\mu \in \mathbb{R}$ Lösung von $(*)$. Die Geraden fallen zusammen.

  Eine inhomogene lineare Gleichung wie $(*)$ muss also nicht unbedingt eine Lösung haben. Hat sie eine Lösung, so ist diese auch nicht unbedingt eindeutig.

Satz 1.1.17 (Kosinussatz)**.**

1. Sei (a, b, c) ein nicht-ausgeartetes Dreieck. Sei α der Winkel an der Ecke a, d.h. $\alpha = \alpha(b - a, c - a)$



Dann gilt

$$d(b, c)^2 = d(a, b)^2 + d(a, c)^2 - 2d(a, b)d(a, c) \cos \alpha \, .$$

2. Ist insbesondere (a, b, c) rechtwinklig mit $\alpha = \frac{\pi}{2}$, so folgt der Satz des Pythagoras:

$$d(b, c)^2 = d(a, b)^2 + d(a, c)^2 \, ,$$

das Quadrat der Länge der Hypotenuse ist gleich der Summe der Kathetenquadrate.

Beweis.

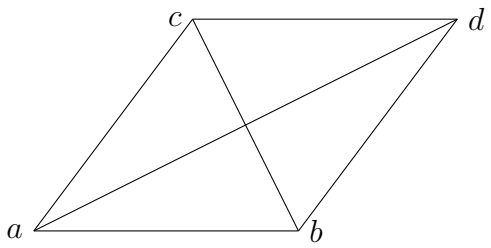
1. Durch direktes Nachrechnen:

$$\begin{aligned} & d(b, c)^2 - d(a, b)^2 - d(a, c)^2 \\ = & \langle b - c, b - c \rangle - \langle a - b, a - b \rangle - \langle a - c, a - c \rangle \\ = & \langle b, b \rangle - 2\langle b, c \rangle + \langle c, c \rangle - \langle a, a \rangle + 2\langle a, b \rangle - \langle b, b \rangle \\ & -\langle a, a \rangle + 2\langle a, c \rangle - \langle c, c \rangle \\ = & 2(-\langle b, c \rangle - \langle a, a \rangle + \langle a, b \rangle + \langle a, c \rangle) \\ = & -2\langle c - a, b - a \rangle = -2\|a - c\| \cdot \|b - a\| \cos \alpha = -2d(a, c)d(b, a) \cos \alpha \end{aligned}$$

2. Der Spezialfall für $\alpha = \frac{\pi}{2}$ folgt aus $\cos \frac{\pi}{2} = 0$.

□

Satz 1.1.18 (Rhombensatz)**.**
Die vier Seiten eines nicht-ausgearteten Parallelogramms sind genau dann gleich lang, wenn sich die beiden Diagonalen senkrecht schneiden:



Beweis.
Die Vektoren, die die beiden Seiten des Parallelogramms festlegen, sind $v := b - a$ und $w := c - a$. Wir berechnen die Diagonalen: die Diagonale, die b und c verbindet, ist durch den Vektor

$c - b = w - v$ gegeben. Im Parallelogramm gilt $b - a = d - c$ nach Definition 1.1.11.2. Daher ist die Diagonale, die a und d verbindet

$$v + w = b + c - 2a = d - a \ .$$

Für den Winkel zwischen den Diagonalen ist also das folgende Skalarprodukt wichtig:

$$\langle v + w, w - v \rangle = \|w\|^2 - \|v\|^2 \ .$$

Die Diagonalen schneiden sich senkrecht, genau dann, wenn das Skalarprodukt $\langle v + w, w - v \rangle$ verschwindet. Genau dann ist aber $\|w\| = \|v\|$. (Man beachte, dass wir hier wieder über eine Äquivalenz von Aussagen sprechen.) □

Wir lernen schließlich noch eine andere Beschreibung von Geraden im $\mathbb{R}^2$ kennen:

Betrachtung 1.1.19.
Sei $G_{p,a}$ eine Gerade mit Fußpunkt $p \in \mathbb{R}^2$ und Richtungsvektor $a \in \mathbb{R}^2 \setminus \{0\}$. Es liegt $x = (x_1, x_2)$ auf G genau dann, wenn es ein $\lambda \in \mathbb{R}$ gibt mit

$$\begin{aligned} x_1 &= p_1 + \lambda a_1 \\ x_2 &= p_2 + \lambda a_2 \ . \end{aligned}$$

Multiplizieren wir die erste Gleichung mit a_2, die zweite mit a_1 und subtrahieren die so erhaltenen Gleichungen, so finden wir die folgende Gleichung für die Koordinaten (x_1, x_2)

$$-a_2 x_1 + a_1 x_2 = -a_2 p_1 + a_1 p_2 \ .$$

Es gilt also mit der schon eingeführten Bezeichnung $(a_1, a_2)^\perp := (-a_2, a_1)$, dass

$$G_{p,a} \subset \{x \in \mathbb{R}^2 | \langle x, a^\perp \rangle = \langle p, a^\perp \rangle\} \ .$$

Definition 1.1.20
Sei $c \in \mathbb{R}^2 \setminus \{0\}$ und $\alpha \in \mathbb{R}$. Dann setzen wir

$$H_{c,\alpha} = \{x \in \mathbb{R}^2 | \langle x, c \rangle = \alpha\} = \{x \in \mathbb{R}^2 | c_1 x_1 + c_2 x_2 = \alpha\}$$

Satz 1.1.21.

1. Sei $c \in \mathbb{R}^2 \setminus \{0\}$ und $\alpha \in \mathbb{R}$. Dann ist $H_{c,\alpha}$ eine Gerade im $\mathbb{R}^2$. Genauer gilt
$$H_{c,\alpha} = G_{\frac{\alpha}{\|c\|^2} c, c^\perp} \ . \qquad (*)$$
2. Sei $p \in \mathbb{R}^2$ und $a \in \mathbb{R}^2 \setminus \{0\}$. Dann hat umgekehrt die Gerade $G_{p,a}$ die Gleichungsdarstellung
$$G_{p,a} = H_{a^\perp, \langle a^\perp, p \rangle} \ . \qquad (**)$$

$H_{c,\alpha}$ heißt die Gleichungsdarstellung der Geraden, denn $H_{c,\alpha}$ ist die Lösungsmenge einer inhomogenen linearen Gleichung $c_1 x_1 + c_2 x_2 = \alpha$ in den Unbestimmten x_1 und x_2.

Beweis.
Wir zeigen 1.; der Beweis von 2. ist analog und dem Leser zur Übung überlassen. Es ist die Gleichheit der zwei Teilmengen $G_{\frac{\alpha}{\|c\|^2} c, c^\perp}$ und $H_{c,\alpha}$ des $\mathbb{R}^2$ zu zeigen.

- Sei $x \in G_{\frac{\alpha}{\|c\|^2}c,c^\perp}$. Dann gibt es nach Definition von G ein $\lambda \in \mathbb{R}$, so dass gilt

$$x = \frac{\alpha}{\|c\|^2}c + \lambda c^\perp \ .$$

  Es folgt $\langle x, c\rangle = \frac{\alpha}{\|c\|^2}\langle c, c\rangle + \lambda\langle c^\perp, c\rangle = \alpha$, also $x \in H_{c,\alpha}$.

- Sei $x \in H_{c,\alpha}$. Dann gilt $\langle x, c\rangle = \alpha$ und somit

$$\langle x - \frac{\alpha}{\|c\|^2}c, c\rangle = \langle x, c\rangle - \frac{\alpha}{\|c\|^2}\langle c, c\rangle = \alpha - \alpha = 0 \, .$$

  Für $y := x - \frac{\alpha}{\|c\|^2}c$ gilt also $\langle y, c\rangle = 0$. Wir behaupten, dass es deswegen ein $\lambda \in \mathbb{R}$ gibt, mit dem $y = \lambda c^\perp$ gilt.
  Daraus folgt dann sofort

$$x = \frac{\alpha}{\|c\|^2}c + y = \frac{\alpha}{\|c\|^2}c + \lambda c^\perp \in G_{\frac{\alpha}{\|c\|^2}c,c^\perp} \ ,$$

  und somit (*).

- Wir müssen die Hilfsbehauptung zeigen: es seien $y, c \in \mathbb{R}^2$, $c \neq 0$, und es gelte

$$0 = \langle y, c\rangle = y_1c_1 + y_2c_2 \ .$$

  Ist $c_1 \neq 0$, so folgt $y_1 = -\frac{y_2}{c_1}c_2$. Es gilt sowieso $y_2 = \frac{y_2}{c_1} \cdot c_1$, also gilt $y = \frac{y_2}{c_1}c^\perp$. Ist $c_1 = 0$, so muss wegen $c \neq 0$ gelten $c_2 \neq 0$. Dann folgt aus $c_2y_2 = 0$, dass $y_2 = 0$. Es folgt $y = \left(-\frac{y_1}{c_2}\right)c^\perp$.

□

Wir bringen die folgende Anwendung:

Definition 1.1.22
Sei (a, b, c) ein nicht-ausgeartetes Dreieck im $\mathbb{R}^2$. Die Höhe H_a ist diejenige Gerade durch den Punkt a, die auf der gegenüberliegenden Seite des Dreiecks, also auf dem Vektor $b - c \in \mathbb{R}^2$, senkrecht steht. Daraus folgt die Gleichungsform

$$H_a = \{x \in \mathbb{R}^2 |\ \langle x, b - c\rangle = \langle a, b - c\rangle\} \ .$$

Analog werden die Höhen H_b durch b und H_c durch c definiert.

Satz 1.1.23 (Höhenschnittsatz)**.**
Sei (a, b, c) ein nicht-ausgeartetes Dreieck. Dann schneiden sich die drei Höhen in einem Punkt.

Beweis.

- H_a und H_b sind nicht parallel, da sonst die Seiten gegenüber den Eckpunkten a und b parallel sein müssten und das Dreieck ausgeartet wäre. Es gibt also einen Punkt $h \in H_a \cap H_b$.

- Wir zeigen: $h \in H_c$. Dazu bemerken wir:

$$\begin{aligned} &\text{Die Aussage } h \in H_a \text{ ist äquivalent zu } \langle h, b-c\rangle = \langle a, b-c\rangle && (*)\\ &\text{Die Aussage } h \in H_b \text{ ist äquivalent zu } \langle h, a-c\rangle = \langle b, a-c\rangle && (**) \end{aligned}$$

  Daraus folgt

$$\begin{array}{rcccl} \langle h, a-b\rangle & = & \langle h, a-c+c-b\rangle & = & \langle h, a-c\rangle - \langle h, b-c\rangle \\ & \overset{(*),(**)}{=} & \langle b, a-c\rangle - \langle a, b-c\rangle & = & \langle b, a\rangle - \langle b, c\rangle - \langle a, b\rangle + \langle a, c\rangle \\ & = & \langle -b+a, c\rangle & = & \langle c, a-b\rangle \ , \end{array}$$

  also $h \in H_c$.

□

Wir wollen noch einige Bemerkungen zu Ebenen und Geraden im dreidimensionalen Raum $\mathbb{R}^3$ bringen.

Bemerkungen 1.1.24.

1. Zwei Geraden im $\mathbb{R}^3$ sind im allgemeinen windschief, d.h. sie schneiden sich nicht und sind nicht parallel. Als Beispiel betrachte die Geraden

$$\mathbb{R}\begin{pmatrix}1\\0\\0\end{pmatrix} \quad \text{und} \quad \begin{pmatrix}0\\1\\0\end{pmatrix} + \mathbb{R}\begin{pmatrix}0\\0\\1\end{pmatrix}$$

   (Machen Sie sich hierzu selbst eine Skizze. Suchen Sie sich zwei windschiefe Geraden in Ihrer Umgebung!)

2. Eine lineare Gleichung im $\mathbb{R}^3$ beschreibt eine Ebene. Geraden können wir dann durch den Schnitt von zwei geeigneten Ebenen beschreiben, also ein lineares Gleichungssystem von 2 Gleichungen in 3 Unbestimmten.

3. Konkret sei $c \in \mathbb{R}^3 \setminus \{0\}$ und $\alpha \in \mathbb{R}$. Dann setzen wir

$$E_{c,\alpha} = \{x \in \mathbb{R}^3 | \langle x, c\rangle = \alpha\} = \{x \in \mathbb{R}^3 | c_1x_1 + c_2x_2 + c_3x_3 = \alpha\}$$

   Wir können Ebenen auch durch zwei reelle Parameter beschreiben: ist $c_3 \neq 0$, so gilt $x \in E_{c,\alpha}$, genau dann, wenn $x_3 = \frac{1}{c_3}(\alpha - c_1x_1 - c_2x_2)$ gilt. Es folgt

$$E = \begin{pmatrix}0\\0\\\frac{\alpha}{c_3}\end{pmatrix} + \mathbb{R}\begin{pmatrix}1\\0\\-\frac{c_1}{c_3}\end{pmatrix} + \mathbb{R}\begin{pmatrix}0\\1\\-\frac{c_2}{c_3}\end{pmatrix} \ .$$

   Dies ist eine Parameterform für eine Ebene in $\mathbb{R}^3$, $E = p + \mathbb{R}v_1 + Rv_2$. Bei einer Gerade hatten wir für eine Parameterform $v_1 \neq 0$ vorausgesetzt. Wir müssen eine Bedingung an das Paar (v_1, v_2) von Vektoren finden, die garantiert, dass wir wirklich eine Ebene parametrisieren. Dafür ist $v_1 \neq 0$ und $v_2 \neq 0$ nicht ausreichend; bei $v_2 = \lambda v_1$ würde man keine Ebene erhalten. Dies leistet der Begriff der linearen Unabhängigkeit, vielleicht der wichtigste Begriff der Vorlesung dieses Semesters und Thema von Kapitel 2.

1.2 Lineare Gleichungssysteme, Gauß'scher Algorithmus

Definition 1.2.1

1. *Ein (reelles) lineares Gleichungssystem ist ein System von Gleichungen der Form*

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

   mit gegebenen $a_{ij} \in \mathbb{R}$ und $b_i \in \mathbb{R}$. Gesucht sind alle reellen Lösungen $x_1, \ldots, x_n$.

2. *Gilt $b_1 = \ldots = b_m = 0$, so heißt das lineare Gleichungssystem homogen; sonst inhomogen.*

3. *Ersetzt man bei einem inhomogenen linearen Gleichungssystem alle b_i durch 0, so erhält man das zugehörige homogene lineare Gleichungssystem.*

4. *Wir nennen die rechteckige Anordnung reeller Zahlen*

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

   die Koeffizientenmatrix des linearen Gleichungssystems. Die reellen Zahlen auf der rechten Seite fassen wir zu $b = (b_1 \ldots, b_m) \in \mathbb{R}^m$ zusammen. Die Lösungsmenge des Gleichungssystems bezeichnen wir mit $\mathrm{Lsg}(A, b)$. Dies ist eine Teilmenge von $\mathbb{R}^n$. Die Matrix

$$(A, b) := \begin{pmatrix} a_{11} & \ldots & a_{1n} & \vert b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & \vert b_m \end{pmatrix}$$

   heißt erweiterte Koeffizientenmatrix des inhomogenen linearen Gleichungssystems.

Wir könnten natürlich auch lineare Gleichungssysteme über den rationalen Zahlen betrachten. Für Matrizen von gewisser Form ist der Lösungsraum einfach zu bestimmen:

Definition 1.2.2

1. *Eine Matrix A ist in Zeilenstufenform, falls für alle $i = 2, \ldots, m$ gilt: sind die ersten $(k-1)$ Einträge der $(i-1)$–ten Zeile gleich Null, so sind die ersten k Einträge der i–ten Zeile gleich Null, wobei $k = 1, \ldots, n$.*

2. *Eine Matrix ist in spezieller Zeilenstufenform, wenn sie in Zeilenstufenform ist und falls für alle $i = 1 \ldots m$ gilt: ist $a_{i1} = a_{i2} = \ldots = a_{i,k-1} = 0$ und $a_{ik} \neq 0$, so ist $a_{ik} = 1$.*

Beispiele 1.2.3.

- Matrizen, die nicht in Zeilenstufenform sind:

$$\begin{pmatrix} 0 & \mathbf{0} & 1 \\ 0 & \mathbf{1} & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & \mathbf{0} \\ 0 & 0 & \mathbf{1} \end{pmatrix}, \qquad \begin{pmatrix} 3 & 5 & 7 \\ \mathbf{1} & 2 & 4 \end{pmatrix}$$

- Matrizen in spezieller Zeilenstufenform: $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\quad$ $\begin{pmatrix} 1 & 5 & 9 \\ 0 & 0 & 0 \end{pmatrix}$
- Matrizen in Zeilenstufenform, aber nicht spezieller Zeilenstufenform: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \mathbf{5} \end{pmatrix}$

Wir zeigen nun an einem Beispiel, warum ein inhomogenes lineares Gleichungssystem, dessen Koeffizientenmatrix in Zeilenstufenform ist, sich einfach lösen lässt. Betrachte

$$(A, b) = \left(\begin{array}{ccc|c} 1 & 2 & 4 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right) .$$

Das steht für das Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 4 \\ x_3 &= 5 \\ 0 &= 0 . \end{aligned}$$

Die dritte Gleichung ist immer erfüllt, die zweite legt $x_3 = 5$ fest. Die erste Gleichung ergibt nach Substitution von x_3 die Gleichung $x_1 + 2x_2 = -16$. Wählt man x_2 als Parameter, so ist der Lösungsraum

$$\mathrm{Lsg}(A, b) = \left\{ x \in \mathbb{R}^3 |\ x_3 = 5,\ x_1 = -16 - 2x_2 \right\} = \begin{pmatrix} -16 \\ 0 \\ 5 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} .$$

Die Lösungsmenge hat die Geometrie einer Geraden in $\mathbb{R}^3$.

Lemma 1.2.4.
Sei A eine Matrix in Zeilenstufenform. Dann ist die Lösungsmenge eines linearen Gleichungssystems mit Koeffizientenmatrix A genau denn leer, wenn es einen Index $i \in \{1, \ldots, m\}$ gibt, so dass $a_{ij} = 0$ für alle j, aber $b_i \neq 0$ gilt.

Beweis.

" $\Leftarrow$ " Die i–te Gleichung ist dann

$$0 = a_{i1}x_1 + \ldots + a_{in}x_n = b_i \neq 0 ,$$

was offensichtlich keine Lösung hat.

" $\Rightarrow$ " Andernfalls finden wir, durch "Aufrollen von unten", wie oben beschrieben, mindestens eine Lösung. Dann kann aber die Gleichung $0 \neq 0$ nicht auftreten.

□

Nun ist die Strategie, Umformungen vorzunehmen, die die Lösungsmenge nicht verändern, also neue, einfachere Gleichungssysteme in Zeilenstufenform zu produzieren, die die gleiche Lösungsmenge haben.

Satz 1.2.5.
Es gibt die folgenden elementaren *Zeilen*umformungen:

1. Multiplikation einer Zeile mit $\lambda \in \mathbb{R} \setminus \{0\}$
2. Vertauschung zweier Zeilen
3. Addition des Vielfachen einer Zeile zu einer anderen Zeile.

Entsteht das lineare Gleichungssystem $(\widetilde{A}, \widetilde{b})$ aus (A, b) durch eine Folge elementarer Zeilenumformungen, so ändert sich die Lösungsmenge nicht, $\mathrm{Lsg}(\widetilde{A}, \widetilde{b}) = \mathrm{Lsg}(A, b)$.

Beweis.
Es kommt offensichtlich nicht auf die die Reihenfolge der Gleichungen an; damit ist 2. klar. Auch 1. sieht man sofort. Beim dritten Typ kommt es auf nur zwei Zeilen i, k an. Es reicht also aus zu zeigen, dass die Gleichungssysteme

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n &= b_i \\ a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n &= b_k \end{aligned}$$

und

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 \qquad\quad + \ldots + a_{in}x_n \qquad\quad &= b_i \\ (a_{k1} + \lambda a_{i1})x_1 + (a_{k2} + \lambda a_{i2})x_2 + \ldots + (a_{kn} + \lambda a_{in})x_n &= b_k + \lambda b_i \end{aligned}$$

gleiche Lösungsräume haben. Erfüllt $x = (x_1, \ldots, x_n)$ das erste Gleichungssystem, so auch die erste Gleichung des zweiten Systems und durch Addition des λ-fachen der ersten Gleichung zur zweiten Gleichung des ersten Systems auch die zweite Gleichung des zweiten Systems. Umgekehrt subtrahiert man das λ-fache der der ersten Gleichung des zweiten Systems von der zweiten Gleichung und sieht, dass jede Lösung des zweiten Systems auch Lösung des ersten Systems ist. □

Betrachtung 1.2.6.
Rezept zur Überführung einer beliebigen Matrix in spezielle Zeilenstufenform durch elementare Zeilenumformungen:

1. Vertausche Zeilen so, dass in der ersten Zeile das erste von Null verschiedene Element nicht weiter rechts steht als bei allen anderen Zeilen.
2. Multipliziere alle Zeilen, bei denen der erste nicht verschwindende Eintrag in der gleichen Spalte wie bei der ersten Zeile steht, mit $\lambda \in \mathbb{R} \setminus \{0\}$, so dass dieser Eintrag gleich 1 wird.
3. Subtrahiere die erste Zeile von genau diesen Zeilen.
4. Ist spezielle Zeilenstufenform noch nicht erreicht, so wende die Schritte 1.3. auf die Untermatrix an, die durch Streichung der ersten Zeile entsteht.

Beispiel 1.2.7.
Wir betrachten das inhomogene lineare Gleichungssystem von drei Gleichungen in drei Variablen mit erweiterter Koeffizientenmatrix $\begin{pmatrix} 0 & 1 & 1 & \vert 1 \\ 5 & 10 & -20 & \vert 5 \\ 2 & 8 & 4 & \vert 14 \end{pmatrix}$. Vertauschen der ersten beiden Zeilen liefert das äquivalente System $\begin{pmatrix} 5 & 10 & -20 & \vert 5 \\ 0 & 1 & 1 & \vert 1 \\ 2 & 8 & 4 & \vert 14 \end{pmatrix}$. Wir teilen die erste Zeile durch 5

und die dritte durch 2 und erhalten $\begin{pmatrix} 1 & 2 & -4 & |1 \\ 0 & 1 & 1 & |1 \\ 1 & 4 & 2 & |7 \end{pmatrix}$. Nun ziehen wir die erste Zeile von der dritten ab: $\begin{pmatrix} 1 & 2 & -4 & |1 \\ 0 & 1 & 1 & |1 \\ 0 & 2 & 6 & |6 \end{pmatrix}$ und dividieren die dritte Zeile durch 2: $\begin{pmatrix} 1 & 2 & -4 & |1 \\ 0 & 1 & 1 & |1 \\ 0 & 1 & 3 & |3 \end{pmatrix}$. Wir ziehen nun die zweite Zeile von der dritten Zeile ab: $\begin{pmatrix} 1 & 2 & -4 & |1 \\ 0 & 1 & 1 & |1 \\ 0 & 0 & 2 & |2 \end{pmatrix}$ und dividieren schließlich die dritte Zeile durch 2, um spezielle Zeilenstufenform zu erhalten: $\begin{pmatrix} 1 & 2 & -4 & |1 \\ 0 & 1 & 1 & |1 \\ 0 & 0 & 1 & |1 \end{pmatrix}$.

Insgesamt finden wir wegen Satz 1.2.5 , dass die beiden inhomogenen linearen Gleichungssysteme

$$\begin{array}{rcl} x_2 + x_3 &=& 1 \\ 5x_1 + 10x_2 - 20x_3 &=& 5 \\ 2x_1 + 8x_2 + 4x_3 &=& 14 \end{array} \qquad \text{und} \qquad \begin{array}{rcl} x_1 + 2x_2 - 4x_3 &=& 1 \\ x_2 + x_3 &=& 1 \\ x_3 &=& 1 \end{array}$$

die gleichen Lösungsmengen haben. Das rechte System lösen wir, wie in Lemma 1.2.4 beschrieben, direkt von unten nach oben: aus $x_3 = 1$ folgt durch Einsetzen $x_2 = 0$ und durch weiteres Einsetzen $x_1 - 4 = 1$, also $x_1 = 5$. Wir finden in diesem Beispiel eine eindeutige Lösung.

Wir fassen den Gauß'schen Algorithmus zur Lösung inhomogener linearer Gleichungssysteme zusammen:

1. Stelle die erweiterte Koeffizientenmatrix (A, b) auf.
2. Überführe diese Matrix (A, b) durch die elementaren Zeilenumformungen aus Satz 1.2.5 in Zeilenstufenform $(\widetilde{A}, \widetilde{b})$.
3. Löse das lineare Gleichungssystem $\widetilde{A}x = \widetilde{b}$ in Zeilenstufenform wie in Lemma 1.2.4 beschrieben sukzessive von unten nach oben.

Man beachte, dass bei der Reihenfolge des Ausräumens im Gauß'schen Algorithmus die Nummerierung der Variablen ausschlaggebend ist. Man macht nur *Zeilen*umformungen.

1.3 Aussagen

Wir werden jetzt einige der Vorgehensweisen in der Meta-Sprache der Mathematik zusammenfassen. Konkrete Beispiele für diese Konzepte haben wir schon in den vorhergehenden Abschnitten gesehen.

Definition 1.3.1
Unter einer Aussage A verstehen wir ein sprachliches Gebilde, das entweder wahr (w) oder falsch (f) ist.

Durch den undefinierten Ausdruck "sprachliches Gebilde" ist das so natürlich keine mathematische Definition. Wenn Sie sehen wollen, wie man das mit Hilfe formaler Systeme besser machen kann, schauen Sie zum Beispiel in Dirk W. Hoffmann, Grenzen der Mathematik, Springer Spektrum 2018.

Beispiele 1.3.2.
Die Aussage: "Die Geraden G und G' im $\mathbb{R}^n$ schneiden sich" ist entweder wahr oder falsch.
Die Aussage "Es gibt einen Studierenden im Hörsaal H1" hat Wahrheitswert w.
Die Ausssage "$3 \cdot 4 = 4$" hat Wahrheitswert f.
Die Ausdrücke "$5 + 7$", "Guten Tag" und "Wie heißen Sie?" sind keine Aussagen.
Der Satz "Dieser Satz ist falsch." ist keine Aussage. Denn wäre er wahr, so wäre er falsch und umgekehrt. Man kann hier (wegen der Selbstbezüglichkeit) keinen Wahrheitswert wahr oder falsch zuordnen.

Wir bauen nun aus Aussagen neue Aussagen:

Definition 1.3.3
Für $n \in \{1, 2, 3, \ldots\}$ ist eine n-stellige Verknüpfung von gegebenen Aussagen $A_1, A_2, \ldots, A_n$ eine Aussage $V(A_1, \ldots, A_n)$, deren Wahrheitswert durch die Wahrheitswerte der gegebenen Aussagen $A_1, \ldots A_n$ eindeutig bestimmt ist. Sie wird durch eine Wahrheitstafel beschrieben, die die Wahrheitswerte in Abhängigkeit der Wahrheitswerte der gegebenen Aussagen angibt.

Insbesondere definieren wir für zwei Aussagen A und B:

1. *Konjunktion: A und B, in Zeichen $A \wedge B$.*

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

2. *Disjunktion: A oder B, in Zeichen $A \vee B$*

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

Es handelt sich also um ein nicht ausschließendes "oder". Das Zeichen $\vee$ kommt vom lateinischen Wort für oder, vel.

3. *Implikation: aus A folgt B, auch "Wenn A, dann B", in Zeichen $A \Rightarrow B$*

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

Um die vierte Zeile dieser Tafel zu illustrieren, beachte man, dass die Aussage "Wenn Ptolemäus Recht hat, ist die Erde eine Scheibe." eine wahre Aussage ist, obwohl die Erde keine Scheibe ist. Die Aussage "Wenn es regnet, ist die Straße nass." ist nur falsch, wenn es regnet, aber die Straße trocken ist. Dazu steht nicht im Widerspruch, dass eine Straße auch durch den Einsatz eines Reinigungsfahrzeugs naß sein kann. Auch die Aussage "Wenn Kühe fliegen können, dann können Delphine klettern." ist wahr.

4. *Äquivalenz: A äquivalent zu B, auch "A genau dann, wenn B", in Zeichen $A \Leftrightarrow B$*

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

Man vergleiche hierzu auch noch einmal mit dem Beweis von Lemma 1.1.3.

5. *Negation: nicht A, in Zeichen $\neg A$*

A	$\neg A$
w	f
f	w

Zwei theoretisch wichtige Verknüpfungen sind xor und nand:

A	B	A xor B	A nand B
w	w	f	f
w	f	w	w
f	w	w	w
f	f	f	w

Man kann zeigen: unter Verwendung der (immer) wahren und der (immer) falschen Aussage lassen sich alle elementaren Verknüpfungen durch nand darstellen.

Wir vereinbaren die Reihenfolge, ähnlich wie "Punkt vor Strich" in der normalen Arithmetik, $\neg$ vor $\wedge$ vor $\vee$ vor $\Rightarrow$ vor $\Leftrightarrow$.

Beispiel 1.3.4.
Seien A und B zwei Aussagen. Die Aussage

$$(\neg A) \vee B =: \neg A \vee B$$

hat die folgende Wahrheitstafel:

A	B	$\neg A$	$\neg A \vee B$
w	w	f	w
w	f	f	f
f	w	w	w
f	f	w	w

Dies ist dieselbe Wahrheitstafel wie die der Verknüpfung $A \Rightarrow B$. Die beiden verknüpften Aussagen $\neg A \vee B$ und $A \Rightarrow B$ sind also durch die Wahrheitstafel nicht zu unterscheiden; sie unterscheiden sich nur darin, wie sie aus elementaren Verknüpfungen aufgebaut sind.

Definition 1.3.5
Gegeben seien mehrere Aussagen $A, B, C, \ldots$ und zwei Aussagen X und Y, die durch die Verknüpfung dieser Aussagen entstanden sind. Wenn die Aussage

$$X \Leftrightarrow Y$$

für alle möglichen Wahrheitswerte der Aussagen $A, B, C, \ldots$ den Wahrheitswert w annimmt, so sagt man, X und Y sind logisch gleichwertig. Die Aussage $X \Leftrightarrow Y$ heißt dann eine Tautologie.

Satz 1.3.6.
Wenn A, B, C Aussagen sind, dann sind die folgenden Aussagen Tautologien:

1. (Doppelnegationsgesetz) $\neg(\neg A) \Leftrightarrow A$
2. (Kommutativgesetze) $A \wedge B \Leftrightarrow B \wedge A$ und $A \vee B \Leftrightarrow B \vee A$
3. (Assoziativgesetze) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$ und $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$
4. (Distributivgesetze) $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ und $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$
5. (de Morgansche Gesetze) $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$ und $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$
6. (Kontrapositionsgesetz) $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$

Beweis.
Der Beweis dieser Aussagen geschieht durch die Betrachtung der relevanten Wahrheitstafeln. Wir führen dies am Beispiel der ersten Aussage in 5., der de Morganschen Regel, vor:

A	B	$A \wedge B$	$\neg(A \wedge B)$	$\neg A$	$\neg B$	$(\neg A) \vee (\neg B)$	(5.1)
w	w	w	f	f	f	f	w
w	f	f	w	f	w	w	w
f	w	f	w	w	f	w	w
f	f	f	w	w	w	w	w

□

Bemerkungen 1.3.7.

1. Die Tautologie 1.3.6.6 liegt der Beweistechnik des "indirekten Beweises" zugrunde. Man beachte beachte, dass sich die Richtung der Implikation umkehrt. Wir haben dies im Beweis von Lemma 1.1.5 und Beispiel 1.1.4 gesehen. Um indirekte Beweise führen zu können, sollte man also das Verneinen von Aussagen sehr gut beherrschen.
2. Man kann alle n-stelligen Verknüpfungen als verschachtelte Verknüpfung dieser elementaren Verknüpfungen darstellen. Dafür reichen sogar die Negation $\neg$, "und" $\wedge$ oder "oder" $\vee$.

Wir wollen abschließend noch Aussagen betrachten, deren Wahrheitswert von einem Element einer gewissen Menge M abhängt. Solche Aussagen waren schon im Beweis von Lemma 1.1.3 aufgetreten. Sei zum Beispiel M die Menge aller Hörer der Vorlesung und für $x \in M$ die Aussage $C(x)$ "x hatte vertieft Mathematik". Den Begriff Menge und Abbildung behandeln wir hier weiterhin noch naiv.

Definition 1.3.8
Eine Aussageform oder Prädikat ist eine Abbildung $Q: M \to \{w, f\}$ von einer Menge M in die Menge $\{w, f\}$ der Wahrheitswerte.

Bemerkungen 1.3.9.

1. Zum Beispiel sei $M = \mathbb{N}$ die Menge der natürlichen Zahlen. Betrachte das Prädikat, das $n \in \mathbb{N}$ die Aussage "Die Zahl n ist eine Primzahl." zuordnet.
2. Eine n-stellige Verknüpfung ist ein Prädikat $\{w, f\}^n \to \{w, f\}$.

3. Man kann Prädikate mit Negationen, Konjunktionen, Disjunktionen, Implikationen und Äquivalenzen kombinieren und neue Prädikate erhalten. Zum Beispiel ordnet für ein gegebenes Prädikat $P : M \to \{w, f\}$ das Prädikat $\neg P : M \to \{w, f\}$ dem Element $m \in M$ per Definitioin den Wahrheitswert $\neg P(m)$ zu.

Bemerkungen 1.3.10.
Seien M, N Mengen und seien $A(x), B(x)$ und $C(x, y)$ Aussagen, deren Wahrheitswert davon abhängt, welche Elemente $x \in M$ bzw. $y \in N$ man einsetzt. Dann bedeutet:

$$\begin{array}{rl} \forall x \in M : A(x) & \text{Die Aussage } A(x) \text{ gilt für alle } x \in M\,. \\ \exists x \in M : A(x) & \text{Es gibt ein } x \in M, \text{ für das } A(x) \text{ gilt}\,. \\ \exists! x \in M : A(x) & \text{Es gibt } \textit{genau} \text{ ein } x \in M, \text{ für das } A(x) \text{ gilt}\,. \end{array}$$

Man nennt auch $\forall$ den Allquantor und $\exists$ den Existenzquantor. Es gelten die folgenden Regeln:

1. $\neg\big(\forall x \in M :\ A(x)\big) \Leftrightarrow \big(\exists x \in M : \neg A(x)\big)$.
2. $\neg\big(\exists x \in M :\ A(x)\big) \Leftrightarrow \big(\forall x \in M : \neg A(x)\big)$.
3. $\Big(\forall x \in M :\ A(x)\Big) \wedge \Big(\forall x \in M :\ B(x)\Big) \Leftrightarrow \forall x \in M :\ A(x) \wedge B(x)$.
4. $\Big(\forall x \in M :\ A(x)\Big) \vee \Big(\forall x \in M :\ B(x)\Big) \Rightarrow \forall x \in M :\ A(x) \vee B(x)$.
5. $\exists x \in M :\ A(x) \vee B(x) \Leftrightarrow \Big(\exists x \in M :\ A(x)\Big) \vee \Big(\exists x \in M :\ B(x)\Big)$.
6. $\exists x \in M :\ A(x) \wedge B(x) \ \Rightarrow \Big(\exists x \in M :\ A(x)\Big) \wedge \Big(\exists x \in M :\ B(x)\Big)$
7. $\exists x \in M :\ \exists y \in N :\ C(x, y) \Leftrightarrow \exists y \in N : \exists x \in M :\ C(x, y)$
8. $\exists x \in M :\ \forall y \in N :\ C(x, y) \Rightarrow \forall y \in N : \exists x \in M :\ C(x, y)$.

Bemerkungen 1.3.11.

1. zu 1.3.10.1: Die Verneinung der Aussage "Alle Schafe sind schwarz." ist eben nicht "Kein einziges Schaf ist schwarz." sondern "Es gibt (wenigstens) ein nicht-schwarzes Schaf." (Prinzip des Gegenbeispiels).

2. zu 1.3.10.4 und 1.3.10.6: wir machen klar, warum "$\Leftarrow$" jeweils nicht gilt: sei M die Menge aller lebenden Menschen. Sei $A(x)$ die Aussage: "x ist ein Vegetarier." und $B(x)$ die Aussage: "x ist ein Fleischesser.". Dann gilt die rechte Seite von 1.3.10.4, weil jeder Mensch in unserem Modell entweder Vegetarier oder Fleischesser ist; die linke Seite würde aber nur gelten, wenn es entweder nur Vegetarier oder nur Fleischesser gäbe.

   Ähnlich gilt die rechte Seite von 1.3.10.6, weil es wenigstens einen Vegetarier und wenigstens einen Fleischesser gibt, aber linke Seite würde nur gelten, wenn es einen Menschen gäbe, der gleichzeitig Vegetarier und Fleischesser ist, was wiederum in unserem Modell nicht vorgesehen ist.

3. zu 1.3.10.8, das klar macht, warum "$\Leftarrow$" nicht gilt: sei M die Menge aller lebenden Männer und N die Menge aller lebenden Frauen. Sei $C(x, y)$ die Aussage "Herr x ist mit Frau y verheiratet.". Die Aussage "$\exists x \in M\ \forall y \in N$ gilt $C(x, y)$" bedeutet: "Es gibt einen Riesenheiratsschwindler, der mit allen lebenden Frauen verheiratet ist."
   Die Aussage "$\forall y \in N\ \exists x \in M$ gilt $C(x, y)$" bedeutet dagegen: "Alle Frauen sind verheiratet, aber möglicherweise monogam."

Bemerkung 1.3.12.
Es ist nützlich, die gängigsten sprachlichen Formulierungen zu kennen, die die Erzeugung einer Aussage aus einem Prädikat mit Hilfe des Allquantors oder des Existenzquantors beschreiben. Sie werden häufig benutzt, um mathematische Texte lesbar zu machen.

Allquantor $\forall x \in M : \ P(x)$

- Für alle $x \in M$ gilt $P(x)$.
- Für jedes Element $x \in M$ gilt $P(x)$.
- Für ein beliebiges Element $x \in M$ gilt $P(x)$.
- Sei $x \in M$ (beliebig). Dann gilt $P(x)$.
- Ist $x \in M$, dann/so gilt $P(x)$.
- Wenn $x \in M$, dann folgt $P(x)$.
- Jedes Element von M erfüllt P.
- Alle Elemente von M erfüllen P.

Existenzquantor $\exists x \in M : \ P(x)$

- Es gibt (mindestens) ein $x \in M$ mit $P(x)$.
- Es existiert (mindestens) ein $x \in M$ mit $P(x)$.
- Die Menge M hat ein Element x, das P erfüllt.
- Ein Element von M erfüllt P.
- Für ein geeignetes Element $x \in M$ gilt $P(x)$.
- Man kann ein $x \in M$ wählen, so dass $P(x)$ gilt.

1.4 Mengen und Abbildungen

Ich folge hier dem Skript von Frau Meusburger.

"Definition" 1.4.1 (Cantor)**.**
"Eine Menge ist eine Zusammenfassung bestimmter, wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen."

Bemerkungen 1.4.2.
Die Definition ist inadäquat, aber dennoch wird aus ihr deutlich:

- Eine Menge besteht aus *Elementen* – den "Objekten unserer Anschauung oder unseres Denkens."
- Die Elemente sind bestimmt: es lässt sich entscheiden, ob etwas Element einer Menge ist oder nicht. Dies ist eine *Aussage*.
- Die Elemente sind wohlunterschieden, d.h. sie es kommt nicht mehrmals das gleiche Element in einer Menge vor. Wir vereinbaren daher $\{a_1, a_1, a_2\} = \{a_1, a_2\}$. Auch kommt es nicht nicht auf die Reihenfolge der Elemente an, also $\{a_2, a_1\} = \{a_1, a_2\}$.

- Die Menge ("das Ganze") ist durch ihre Elemente charakterisiert.
- Es war schon Cantor klar, dass es eine leere Menge geben sollte, die kein Element enthält.

Etwas moderner ausdrückt:

"Definition" 1.4.3.

1. Eine Menge ist etwas, das Elemente enthalten kann. Ist a ein Element einer Menge M, so schreibt man $a \in M$. Ist a kein Element der Menge M, so schreibt man $a \notin M$. Für eine endliche Menge M, die genau die endlich vielen Elemente $a_1, a_2, \ldots, a_n$ enthält, schreibt man $M = \{a_1, a_2, \ldots, a_n\}$.
2. Es gibt eine ausgezeichnete Menge, die leere Menge $\emptyset$, die keine Elemente enthält.
3. Eine Menge ist durch ihre Elemente eindeutig bestimmt. Diese können auch selbst Mengen sein. Zwei Mengen M, N sind gleich, genau dann, wenn sie die gleichen Elemente enthalten, also wenn aus $a \in M$ folgt, dass $a \in N$ und aus $a \in N$ folgt $a \in M$. Man schreibt dann $M = N$.

Es bleibt aber ein Selbstbezüglichkeitsproblem, das die Russelsche Antinomie thematisiert: sie betrachtet die (sogenannte) Menge R aller Mengen x, die sich nicht selbst als Element enthalten, $R = \{x | x \notin x\}$ und fragt, ob R sich selbst als Element enthält.

Der moderne Zugang startet von der Idee, dass man eine Menge als etwas definiert, was man aus anderen Mengen konstruieren kann. Dadurch lassen sich alle Mengen auf eine Menge, die leere Menge, zurückführen. (Auch eine Spielfigur eines Brettspiels ist das, was man mit ihr machen darf.) Die folgenden Axiome sind leicht redundant:

"Definition" 1.4.4 (Zermelo-Fraenkel Axiome)**.**

1. Bestimmtheitsaxiom: Zwei Mengen M, N sind gleich, wenn $a \in M \Rightarrow a \in N$ und $a \in N \Rightarrow a \in M$. Man schreibt dann $M = N$ und ansonsten $M \neq N$. Eine Menge N heisst Teilmenge einer Menge M, in Zeichen $N \subset M$, wenn $a \in N \Rightarrow a \in M$ gilt.

   Bild: Ein Behälter ist durch seinen Inhalt eindeutig bestimmt.

   Zwei Behälter werden als gleich betrachtet, wenn sie den gleichen Inhalt haben. Insbesondere ist der Inhalt der Behälter nicht in irgendeiner Weise geordnet und jedes Ding ist maximal einmal in einem Behälter enthalten.

2. Axiom der leeren Menge: Es gibt eine ausgezeichnete Menge, die leere Menge, die keine Elemente enthält und mit $\emptyset$ bezeichnet wird.

   Bild: Es gibt einen leeren Behälter, der keine Dinge enthält.

3. Paarungssaxiom: Zu zwei beliebigen Mengen M, N gibt es eine Menge X, die genau M und N als Elemente enthält. Man schreibt $X = \{M, N\}$, falls $M \neq N$ und $X = \{M\}$, wenn $M = N$.

   Bild: Man kann zwei Behälter mitsamt ihrem Inhalt in einen weiteren Behälter packen.

   Beispiel: durch wiederholte Anwendung des Paarungsaxioms konstruiert man aus der leeren Menge die Mengen $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ ...

4. Vereinigungssaxiom: Zu jeder Menge M von Mengen gibt es eine Menge X, die genau die Elemente der Elemente von M als Elemente enthält. Man schreibt $X = \cup M$ und statt $\cup \{A_1, A_2, \ldots, A_N\}$ auch $A_1 \cup A_2 \cup \ldots\ldots \cup A_n$.

   Bild: Man kann alle in einem Behälter enthaltenen Behälter ausleeren und die dabei zum Vorschein kommenden Dinge in einen neuen Behälter packen.

5. Aussonderungsaxiom: Für jede Menge M und jedes Prädikat $P : M \to \{w, f\}$ gibt es eine Menge X, die genau die Elemente von M mit $P(m) = w$ enthält. Man schreibt $X = \{m \in M | P(m)\}$ oder $X = \{m \in M : P(m)\}$.

   Bild: Man kann einen Behälter öffnen, die Dinge herausnehmen, die eine bestimmte Eigenschaft haben, und sie in einen neuen Behälter packen.

6. Unendlichkeitsaxiom: Es gibt eine Menge X, so dass $\emptyset \in X$ und für jedes Element $x \in X$ auch $\{x\} \in X$ gilt.

   Bild: Es gibt einen Universalbehälter, der den leeren Behälter enthält und für jedes Ding, das er enthält, auch den Behälter enthält, der dieses Ding enthält. Insbesondere enthält X alle Mengen der Form $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$ $\{\{\{\emptyset\}\}\}$,...

7. Potenzmengenaxiom: Für jede Menge M gibt es eine Menge $\mathcal{P}(M)$, die Potenzmenge von M, deren Elemente genau die Teilmengen von M sind.

   Bild: Man kann einen gegebenen Behälter ausleeren, einen Teil seines Inhalts auswählen und in einen neuen Behälter füllen - die Teilmenge. Dann kann man einen Behälter bauen, der alle auf diese Weise gefüllten Behälter enthält.

8. Ersetzungsaxiom: Ist M eine Menge und f eine fest gewählte "Konstruktionsvorschrift", die aus Mengen und deren Elementen durch endliches Verschachteln der Ausdrücke $=, \in, \wedge, \vee, \neg, \Leftrightarrow, \Rightarrow, \forall, \exists$ neue Elemente 'konstruiert" und so jedem Element '$m \in M$ genau ein Element $f(m)$ irgendeiner Menge zuordnet, dann gibt es eine Menge X, die genau die Elemente $f(m)$ für $m \in M$ enthält.

   Bild: Kann man jedem Ding in einem gegebenen Behälter durch eine Vorschrift, die aus "ist gleich", "ist enthalten in", "und", "oder", "nicht", "genau dann, wenn", "wenn, dann", "für alle" und "es gibt ein" besteht, eindeutig ein Ding in irgendeinem anderen oder dem gleichen Behälter zuordnen, dann kann man diese zugeordneten Dinge aus ihren Behältern herausnehmen und zusammen in einen neuen Behälter packen.

9. Fundierungsaxiom: In jeder nichtleeren Menge M gibt es ein Element $m \in M$, so dass m und M keine Elemente gemeinsam haben.

   Bild: In jedem nicht-leeren Behälter gibt es ein Ding, das keinen gemeinsamen Inhalt mit dem Behälter hat.

10. Auswahlaxiom: Ist M eine Menge, so dass alle Elemente von M nicht-leere Mengen sind und je zwei Elemente von M keine gemeinsamen Elemente haben, dann gibt es eine Menge X, die genau ein Element aus jedem Element $m \in M$ enthält.

    Bild: Hat man einen Behälter, der nur nicht-leere Behälter enthält, die untereinander keinerlei Inhalt gemeinsam haben, dann kann man aus jedem dieser nicht-leeren Behälter genau ein Ding auswählen und diese Dinge in einen neuen Behälter packen.

Bemerkungen 1.4.5.

1. Die Zermelo-Fraenkel-Axiome schließen die Russelsche Antinomie aus. Konkret geschieht das durch das Fundierungsaxiom und das Paarungsaxiom. Für jede Menge M kann man mit dem Paarungsaxiom die Menge $\{M\}$ bilden, die als einziges Element die Menge M enthält. Nach dem Fundierungsaxiom, muss dann gelten, dass die Menge M mit der Menge $\{M\}$ kein Element gemeinsam hat. Da $M \in \{M\}$ gilt, muss somit $M \notin M$ gelten. Keine Menge kann sich also selbst enthalten.

Somit existieren die "Menge aller Mengen" oder die "Menge aller Mengen, die sich selbst als Element enthalten" nicht, da sie jeweils sich selbst als Element enthalten würden. Die "Menge aller Mengen, die sich nicht selbst als Element enthalten" kann es ebenfalls nicht geben, da keine Menge sich selbst als Element enthalten kann, und somit diese Menge gleich der "Menge aller Mengen" wäre.

2. Wir zeigen, wie sich beispielsweise die Menge $\mathbb{N} = \{0, 1, \ldots\}$ [2] der natürlichen Zahlen aus den Zermelo-Fraenkel Axiomen konstruieren lässt, die zunächst nur verschachtelte Ausdrücke mit leeren Mengen liefern. Dazu genügt es, eine eindeutige Zuordnung einer mittels der Zermelo-Fraenkel Axiome konstruierbaren Menge M_n zu jeder Zahl $n \in \mathbb{N}_0$ anzugeben, so dass für alle $n; m \in \mathbb{N}_0$ gilt: $n < m \Leftrightarrow M_n \in M_m$. Eine solche Zuordnung liefert die von Neumannsche Zahlenreihe, die durch $M_0 = \emptyset$; und $M_{n+1} := M_n \cup \{M_n\}$ für alle $n \in N_0$ definiert ist. Dies ergibt $M_0 = \emptyset$, $M_1 = \{\emptyset\}$, $M_2 = \{\emptyset, \{\emptyset\}\}$, $M_3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, ...

   Jetzt führen weitere Konstruktionen auf die Menge $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, \ldots\}$ der ganzen Zahlen, die Menge $\mathbb{Q} = \{\frac{p}{q} | \, p, \, q \in \mathbb{Z}, q \neq 0\}$ der rationalen Zahlen. Die Menge $\mathbb{R}$ der reellen Zahlen wird in der Vorlesung Analysis eingeführt.

Mit Hilfe dieser Betrachtung können wir nun die wichtigsten Konstruktionen mit Mengen in einer handlicheren Definition zusammenfassen, die wir im Folgenden als Ausgangspunkt nehmen werden. Wir werden die Zermelo-Fraenkel Axiome im Verlauf der Vorlesung nicht mehr benutzen. In der Tat kommt man in der Mathematik erstaunlich weit, ohne je von diesen Axiomen gehört zu haben. Im Prinzip reicht es aus, zu wissen, dass eine Menge nie Element von sich selbst sein kann, und die grundlegenden Konstruktionen mit Mengen zu beherrschen.

Definition 1.4.6

1. *Seien A, B Mengen; dann heißt A Teilmenge von B bzw. B Obermenge von A, falls jedes Element von A auch Element von B ist. Wir schreiben $A \subset B$ oder $B \supset A$ genau dann, wenn für alle $a \in A$ auch $a \in B$ gilt, in Formeln $a \in A \Rightarrow a \in B$. Es gilt zum Beispiel:*

$$\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \, .$$

   Aus $x \in A$ folgt $\{x\} \subset A$. Man sollte aber die Menge $\{x\}$ nicht mit dem Element x verwechseln: eine Schachtel mit einem Hut ist eben etwas anderes als ein Hut.

2. *Zwei Mengen heißen gleich:*

$$A = B \underset{Def}{\Leftrightarrow} A \subset B \wedge B \subset A \, .$$

3. *Sei A eine Menge. Dann ist die Potenzmenge $\mathcal{P}(A)$ die Menge aller Teilmengen von A, d.h.*

$$\mathcal{P}(A) := \{B | B \subset A\} \, .$$

   Beispiel: Die Potenzmenge $\mathcal{P}(\emptyset) = \{\emptyset\}$ der leeren Menge hat genau ein Element, nämlich die leere Menge.

   Eine Menge kann also ein Element einer anderen Menge sein. Das ist oft so: zum Beispiel ist eine Gerade in $\mathbb{R}^2$ eine Menge von Punkten des $\mathbb{R}^2$, aber auch ein Element der Menge der Geraden in $\mathbb{R}^2$.

[2] Wenn wir deutlich machen wollen, dass 0 in der Menge der natürlichen Zahlen liegt, schreiben wir auch $\mathbb{N}_0$.

4. *Elementare Operationen aus der Verknüpfung von Aussagen: seien A, B Mengen*

   (a) *Schnittmenge, Durchschnitt*
   $$A \cap B = \{a | a \in A \wedge a \in B\} .$$

   (b) *Vereinigung*
   $$A \cup B = \{a \mid a \in A \vee a \in B\} .$$
   Man zeige: $A \cap B \subset A \cup B$.

   (c) *Mengendifferenz*
   $$A \setminus B = \{a | a \in A \wedge a \notin B\} .$$
   Falls $B \subset A$ *gilt, heißt*
   $$\complement B = A \setminus B$$
   auch Komplement von B *(bezüglich* A*).*

Bemerkung 1.4.7.

1. Man kann die natürlichen Zahlen durch die Peano-Axiome charakterisieren.

   (a) 0 ist eine natürliche Zahl.

   (b) Jede natürliche Zahl n hat eine natürliche Zahl n' als Nachfolger.

   (c) 0 ist kein Nachfolger einer natürlichen Zahl.

   (d) Natürliche Zahlen mit gleichem Nachfolger sind gleich.

   (e) Prinzip der vollständigen Induktion:
   Sei $M \subset \mathbb{N}_0$ eine Teilmenge mit den beiden Eigenschaften, dass M die Null enthält, $0 \in M$, und dass mit n auch der Nachfolger in M liegt, $n \in M \Rightarrow n' \in M$. Dann ist $M = \mathbb{N}$.

   Es ist wichtig zu verstehen, dass das Prinzip der vollständigen Induktion eine *Eigenschaft* der natürlichen Zahlen ist.

2. Das Prinzip der vollständigen Induktion liefert ein wichtiges *Beweisverfahren*, um eine Aussage der Form $(x \in \mathbb{N} : P(x))$ (oder für eine unendliche Teilmenge von $\mathbb{N}$) zu beweisen. Beispiele sind Formeln, deren Gültigkeit für alle natürlichen Zahlen (oder eine unendliche Teilmenge von $\mathbb{N}$) bewiesen werden soll. Dabei geht man wie folgt vor. Man zeigt zunächst, dass die Aussage für die kleinste natürliche Zahl in der Menge M wahr ist. Dies bezeichnet man als den Induktionsanfang. Danach zeigt man, dass aus $P(n) = w$ für eine Zahl $n \in M$ folgt, dass auch $P(m) = w$ für die nächstgrößere Zahl $m \in M$, also dass die Induktionsbehauptung wahr ist. Dies nennt man den Induktionsschritt. Er beweist die Wahrheit der Aussage für alle $m \in M$.

Beispiel 1.4.8.
Für alle natürlichen Zahlen n gilt die Aussage $A(n)$, nämlich $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$.
1. Induktionsanfang: Die Aussage gilt für $n = 1$, denn $1 = \frac{1 \cdot 2}{2}$.
2. Induktionsschritt: Angenommen die Aussage gilt für die natürliche Zahl n. Dann ergibt sich für die Zahl $n + 1$:

$$\begin{aligned} 1 + 2 + \ldots + n + n + 1 &= (1 + 2 + \ldots + n) + (n + 1) = \frac{n(n+1)}{2} + n + 1 \\ &= \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2} \end{aligned}$$

Also gilt dann die Aussage auch für die natürliche Zahl $n+1$. Sei M die Menge aller natürlichen Zahlen $n \in \mathbb{N}$, für die $A(n)$ wahr ist. Wegen der Induktionsannahme ist $1 \in M$. Wegen des Induktionsschritts folgt aus $n \in M$, dass $n+1 \in M$. Aus dem Prinzip der vollständigen Induktion folgt, dass M alle natürlichen Zahlen ≥ 1 enthält, also dass die Aussage für alle $n \in \mathbb{N}_{\geq 1}$ wahr ist.

Bemerkung 1.4.9.
Das Prinzip der vollständigen Induktion liefert aber auch ein *Definitionsverfahren.*

1. Mit vollständiger Induktion lassen sich Addition und Multiplikation auf $\mathbb{N}$ rekursiv definieren:

   - Addition: $0 + m := m$ und $m + n' := (m + n)'$.
   - Multiplikation: $0 \cdot m := 0$ und $n' \cdot m := (n \cdot m) + m$

   Die Eins definiert man als Nachfolger der Null, $1 := 0'$. Aus dem Additionsaxiom folgt dann $n' = n + 1$.

2. Zum Beispiel sei für jedes $i \in \mathbb{N}$ ein Ausdruck $f(i)$ gegeben, etwa $f(i) = i^2$. Um

   $$\sum_{i=0}^{k} f(i)$$

   für natürliche Zahlen zu definieren, setzen wir als Induktionsanfang für $k = 0$, dass $\sum_{i=0}^{0} f(i) = f(0)$. Dann setzen wir

   $$\sum_{i=0}^{k+1} f(i) = \left(\sum_{i=0}^{k} f(i) \right) + f(k+1) \ .$$

   Nach dem Prinzip der vollständigen Induktion ist dann der Ausdruck $\sum_{i=0}^{k} f(i)$ für alle natürlichen Zahlen definiert.

Satz 1.4.10.
Seien A, B, C Mengen. Dann gilt

1. Kommutativgesetze: $A \cap B = B \cap A$ und $A \cup B = B \cup A$
2. Assoziativitätsgesetze: $(A \cap B) \cap C = A \cap (B \cap C)$ und $A \cup (B \cup C) = (A \cup B) \cup C$
3. Distributivgesetze: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
4. Komplementgesetz: Ist A Teilmenge einer Menge C, so gilt für das Komplement bezüglich C: $\complement(\complement A) = A$
5. De Morgansche Regeln: $\complement(A \cup B) = \complement A \cap \complement B \qquad \complement(A \cap B) = \complement A \cup \complement B$.

Beweis.

Alle Aussagen folgen aus den entsprechenden Aussagen in Satz 1.3.6. Wir führen dies am Beispiel des ersten Distributivgesetzes 1.4.10.3 vor:

$$\begin{aligned} x \in A \cap (B \cup C) \quad &\Leftrightarrow\ x \in A \wedge x \in (B \cup C) \\ &\Leftrightarrow x \in A \wedge (x \in B \vee x \in C) \\ &\underset{1.3.6.4}{\Leftrightarrow} (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C) \\ &\Leftrightarrow (x \in A \cap B) \vee (x \in A \cap C) \\ &\Leftrightarrow x \in (A \cap B) \cup (A \cap C)\,. \end{aligned}$$

Machen Sie sich diesen Sachverhalt auch am Beispiel eines Bildes klar. □

Vorsicht: man sollte die Operationen $\subset$, $\cap$, $\cup$ für Mengen und die Verknüpfungen $\Rightarrow$, $\wedge$, $\vee$ für Aussagen nicht verwechseln. Bitte unterscheiden Sie auch zwischen "$\in$" für "Element sein" und "$\subset$" für "Teilmenge sein". Also $\{0,1,2\} \in \{\emptyset, 1, \{0,1,2\}\}$ aber $\{0,1,2\} \subset \{0,1,2,3,4\}$.

Definition 1.4.11
Seien $A_1, \ldots, A_n$ Mengen. Dann ist das kartesische Produkt oder direkte Produkt $A_1 \times \ldots \times A_n$ die Menge der geordneten *n-Tupel mit Elementen in $A_1, \ldots, A_n$, d.h.*

$$A_1 \times \ldots \times A_n = \{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_1 \in A_1 \ldots a_n \in A_n \}\,.$$

Man schreibt im Fall $A_1 = A_2 = \ldots A_n = A$ auch

$$A^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \middle| a_i \in A,\ i = 1 \ldots n \right\}$$

für die geordneten n-Tupel von Elementen in A.

Beachte, dass alle Tupel geordnet sind, also $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ zwei verschiedene Elemente in $\mathbb{Z} \times \mathbb{Z}$ sind. Ein wichtiges Beispiel eines kartesischen Produkts ist natürlich $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Eine wichtige Anwendung des kartesischen Produktes ist

Definition 1.4.12

1. *Eine Relation ist ein Tripel (M, N, R), bestehend aus zwei Mengen M, N und einer Teilmenge $R \subset M \times N$. Gilt $(m,n) \in R$, so schreiben wir $m \sim_R n$ und sagen, dass m in Relation R mit n steht. Gilt $M = N$, so sprechen wir von einer Relation auf der Menge M.*

2. *Eine Relation auf einer Menge X heißt Äquivalenzrelation, wenn für alle $x, y, z \in X$ gilt:*

$$\begin{aligned} x \sim x \qquad &\textit{(reflexiv)} \\ x \sim y \Rightarrow y \sim x \qquad &\textit{(symmetrisch)} \\ x \sim y \wedge y \sim z \Rightarrow x \sim z \qquad &\textit{(transitiv)} \end{aligned}$$

3. *Gegeben eine Menge X mit Äquivalenzrelation $\sim$, so heißt eine Teilmenge $A \subset X$ Äquivalenzklasse, falls gilt*

$$A \neq \emptyset$$
$$x, y \in A \Rightarrow x \sim y$$
$$x \in A \text{ und } y \in X \text{ und } x \sim y \Rightarrow y \in A\,.$$

Beispiele 1.4.13.

1. Sei X irgendeine Menge und $\Delta_X = \{(x,x) \,|\, x \in X\} \subset X \times X$ die sogenannte Diagonale. Sie definiert als Relation die Gleichheit von Elementen in X. Dies ist eine Äquivalenzrelation.

2. X = Menge aller Menschen und $x \sim y :\Leftrightarrow x$ kennt y. Dies ist keine Äquivalenzrelation, denn die Relation ist nicht transitiv: sie müssten ja sonst nie ihre Freunde einander vorstellen und Franz Beckenbauer würde alle Deutschen über 40 Jahren kennen.

3. M = Menge aller Männer, N = Menge aller Frauen, R = Menge aller heterosexuellen Ehepaare.

4. $X = \mathbb{R}$ und $x \sim y :\Leftrightarrow x < y$. Diese Relation ist nicht reflexiv und nicht symmetrisch, also keine Äquivalenzrelation.

5. $X = \mathbb{R}^n$ und $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \sim \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} :\Leftrightarrow x_1^2 + \ldots + x_n^2 = y_1^2 + \ldots + y_n^2$. Dies ist eine Äquivalenzrelation. Die Äquivalenzklassen sind Kugeln um den Ursprung.

6. Sei $X = \mathbb{Z}$; wir geben uns $m \in \mathbb{N} \setminus \{0\}$ vor und setzen: $x \sim y :\Leftrightarrow y - x$ ist durch m teilbar. Dies ist eine Äquivalenzrelation auf der Menge der ganzen Zahlen.

Lemma 1.4.14.
Ist R eine Äquivalenzrelation auf einer Menge X, so gehört jedes $a \in X$ zu genau einer Äquivalenzklasse. Insbesondere gilt für zwei beliebige Äquivalenzklassen A, A' entweder $A = A'$ oder $A \cap A' = \emptyset$.

Beweis.
Für $a \in X$ definieren wir

$$A_a := \{x \in X |\, x \sim a\}\,.$$

- Wir zeigen: A_a ist eine Äquivalenzklasse, die a enthält.
  - $a \sim a \Rightarrow a \in A_a$. Insbesondere gilt $A_a \neq \emptyset$
  - Sei $x, y \in A_a$, so gilt $x \sim a$ und $y \sim a$, also wegen der Symmetrie auch $a \sim y$. Somit $x \sim y$ wegen Transitivität.
  - Sei $x \in A_a$, $y \in X$ und $x \sim y$. Dann gilt $x \sim a$ und $x \sim y$, also auch $y \sim x$ wegen Symmetrie und somit wegen Transitivität $y \sim a$. Nach der Definition von A_a folgt $y \in A_a$.

- Es bleibt zu zeigen, dass zwei Äquivalenzklassen entweder gleich oder disjunkt sind. Angenommen $A \cap A' \neq \emptyset$, dann gibt es $a \in A \cap A'$. Ist $x \in A$, folgt aus der zweiten definierenden Eigenschaft der Äquivalenzklasse A, dass $x \sim a$. Zusammen mit $a \in A'$ folgt aus der dritten definierenden Eigenschaft der Äquivalenzklasse A, dass $x \in A'$. Also $A \subset A'$. Die umgekehrte Inklusion $A' \subset A$ folgt analog und somit $A = A'$.

□

Bemerkungen 1.4.15.

1. Jede Äquivalenzrelation R auf einer Menge X liefert eine Zerlegung von X in paarweise disjunkte, also paarweise elementfremde, Äquivalenzklassen. Man sagt auch, dass die Äquivalenzklassen eine Partition von X bilden. Eine Zerlegung oder Partition ist eine Menge von disjunkten Teilmengen, die nicht die leere Menge enthält und deren Vereinigung X ist.

2. Die Äquivalenzklassen fasst man als Elemente einer neuen Menge X/R auf. Deren Elemente sind also Teilmengen von X. Die Menge X/R heißt Quotientenmenge von X nach R.

3. Es gibt eine kanonische (d.h. in dieser Situation ausgezeichnete) Abbildung

$$\begin{aligned} X &\longrightarrow X/R \\ a &\mapsto A_a \end{aligned}$$

4. Jedes $a \in A$ heißt ein Repräsentant der Äquivalenzklasse A. Im allgemeinen gibt es aber keine ausgezeichneten Repräsentanten.

Beispiele 1.4.16.

1. Wir betrachten auf der Menge X aller Schüler einer Schule die Relation $R \ni (a, b)$ genau dann, wenn a und b in die gleiche Klasse gehen. Dies ist eine Äquivalenzrelation. Die Quotientenmenge X/R ist dann genau die Menge aller Klassen der Schule. Stundenplanmacher sind gewohnt, mit dieser Quotientenmenge zu rechnen. Die kanonische Abbildung ordnet einem Schüler seine Klasse zu. Sie wird bei der Beschriftung von Schulheften oft benutzt. Der Klassensprecher ist ein Repräsentant der Klasse.

2. Für jedes $n \in \mathbb{N}$ ist $x \sim_n y$ genau dann, wenn n teilt $x - y$ eine Äquivalenzrelation. Die Äquivalenzklassen sind die Restklassen $[z] := \{z + kn | k \in \mathbb{Z}\}$. Es gibt n Restklassen mit Repräsentanten $0, 1, \ldots n-1$. Die kanonische Abbildung ordnet einer Zahl die Restklasse ihres Rests nach Division durch n zu, etwa für $n = 12$ haben wir $23 \mapsto [11]$. Einfache Uhren mit einem Stundenzeiger, der stündlich springt, realisieren diese Restklassenabbildung.

3. Rationale Zahlen werden als Äquivalenzklassen auf der Menge $M := \{(m, n)\ m, n \in \mathbb{Z}, n \neq 0\}$ mit der Äquivalenzrelation $(a, b) \sim (c, d) \Leftrightarrow ad = bc$ eingeführt. Man bezeichnet die Äquivalenzklasse mit $\frac{a}{b}$. Ein Bruch ist also eine Äquivalenzklasse, $\frac{a}{b} = [(1, 2), (2, 4), (-3, -6), \ldots]$. Gekürzte Brüche mit positivem Nenner bilden ein System von Repräsentanten für die Äquivalenzklasse. Die rationalen Zahlen sind die Quotientenmenge. (Im übrigen sind die reellen Zahlen auch eine Menge von Äquivalenzklassen von Dezimalbrüchen, z.B. muss man $0, \overline{9}$ und 1 identifizieren.)

Definition 1.4.17

1. *Seien A, B Mengen. Eine Abbildung f von einer Menge A in eine Menge B ist eine Relation $R \subset A \times B$, so dass es zu jedem $a \in A$ genau ein $b \in B$ mit $(a, b) \in R$ existiert. Wir schreiben auch $f : A \to B$ oder $A \xrightarrow{f} B$.*

   Anschaulich ist dies eine Vorschrift, die jedem Element $a \in A$ genau ein Element $f(a) \in B$ zuordnet, das Bild von a unter f. Wir schreiben auch $a \mapsto f(a)$.

   Die Menge A heißt Definitionsbereich oder Urbildbereich, B Bild- oder Wertebereich und die Teilmenge $R \subset A \times B$ der Graph der Abbildung. Diese Mengen gehören zur Definition einer Abbildung.

2. *Sei $f : A \longrightarrow B$ eine Abbildung und $A' \subset A$ eine Teilmenge. Dann heißt die Menge*

$$f(A') := \{f(a) \in B \mid a \in A'\}$$

   Bildmenge von A' unter f oder kürzer das Bild der Teilmenge A' unter f. Das Bild ist eine Teilmenge von B, also $f(A') \subset B$. Die Teilmenge $f(A) \subset B$ heißt das Bild der Abbildung f.
   Für eine Teilmenge $B' \subset B$ heißt die Menge

$$f^{-1}(B') := \{a \in A \mid f(a) \in B'\}$$

   Urbildmenge oder kurz Urbild von B' unter f. Das Urbild ist eine Teilmenge von A, also $f^{-1}(B') \subset A$.

3. *Seien $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Die Verkettung oder Komposition $g \circ f$ von g mit f ist die durch*

$$\begin{aligned} g \circ f : \quad & A \longrightarrow C \\ & (g \circ f)(a) := g\big(f(a)\big) \end{aligned}$$

   definierte Abbildung.

Einer der wichtigsten Sätze der Elementarmathematik besagt, dass die Verkettung von Abbildungen assoziativ ist: sei $h : C \to D$ eine weitere Abbildung, dann gilt

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

Beweis: Sei $a \in A$. Wir rechnen:

$$\begin{array}{rcccl} (h \circ g) \circ f(a) & = & (h \circ g)\,(f(a)) & = & h\,(g(f(a))) \ , \\ h \circ (g \circ f)(a) & = & h\,((g \circ f)(a)) & = & h\,(g(f(a))) \ . \end{array}$$

Bemerkungen 1.4.18.

1. Die Abbildung $f_1 : \mathbb{R} \to \mathbb{R}$ mit $x \mapsto x^2$ und die Abbildung $f_2 : \mathbb{R} \to \mathbb{R}_{\geq 0}$ mit $x \mapsto x^2$ sind also verschiedene Abbildungen, da sie unterschiedlichen Bildbereich haben, obwohl sie die gleiche Rechenvorschrift benutzen.

2. Für gegebene Mengen M, N bilden die Abbildungen $f : M \to N$ eine Menge. Denn sie lassen sich durch ein logisches Prädikat aus der Potenzmenge $\mathcal{P}(M \times N)$ auswählen.

3. Eine besonders wichtige Abbildung ist $f = \mathrm{id}_A : A \to A$ mit $a \mapsto a$, die identische Abbildung oder die Identität von A. Ihr Graph ist die Diagonale in $\Delta \subset A \times A$, i.e. $\Delta = \{(a, a) | a \in A\}$.

4. Es gibt zu jeder Menge X eine leere Abbildung $f : \emptyset \to X$; es gibt aber für $X \neq \emptyset$ keine Abbildung $X \to \emptyset$.

5. Die Gerade $G_{p,v} \subset \mathbb{R}^2$ ist das Bild der Funktion $f : \mathbb{R} \to \mathbb{R}^2$ mit $f(\lambda) = p + \lambda v$.

6. Eine Abbildung kann durch eine Rechenvorschrift gegeben sein, etwa $f : \mathbb{Z} \to \mathbb{N}$ mit $x \mapsto x^2$, aber auch durch eine Fallunterscheidung, etwa
$f : \mathbb{R} \to \mathbb{N} \qquad f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \notin \mathbb{Q} \end{cases}$ oder auch durch eine Tafel. Zum Beispiel ist für die Abbildung, die jedem Tag des Jahres 2024 die Tageshöchsttemperatur in Hamburg zuordnet, eine Tafel, aber meines Wissens keine Rechenvorschrift bekannt. Man sollte aber keinesfalls eine Abbildung mit einer Rechenvorschrift verwechseln; die Angabe von Definitions- und Bildbereich ist sehr wichtig. Zum Beispiel gibt die Rechenvorschrift $x \mapsto 2x$ eine Abbildung $f : \mathbb{Q} \to \mathbb{Q}$, die eine Umkehrabbildung $g : \mathbb{Q} \to \mathbb{Q}$ besitzt, nämlich die Rechenvorschrift $x \mapsto \frac{1}{2}x$. Die Verkettungen $f \circ g$ und $g \circ f$ der beiden Abbildungen sind jeweils die Identität auf $\mathbb{Q}$. Die entsprechende Abbildung $\mathbb{Z} \to \mathbb{Z}$ hat aber keine Umkehrabbildung.

Definition 1.4.19
Sei $f : A \to B$ eine Abbildung und seien $A' \subset A$ und $B' \subset B$ Teilmengen.

1. *Die Einschränkung oder Restriktion von f auf A' ist die durch*

$$f_{|A'} : \quad A' \to B$$
$$f_{|A'}(a') = f(a') \qquad \forall a' \in A'$$

definierte Abbildung.

2. *Gilt $f(A) \subset B'$, so definieren wir die Korestriktion von f als die Abbildung $\tilde{f} : A \to B'$ mit $\tilde{f}(a) = f(a)$.*

Man beachte, dass die Korestriktion auf das Bild $f(A)$ von f immer eine surjektive Abbildung ist.

Definition 1.4.20

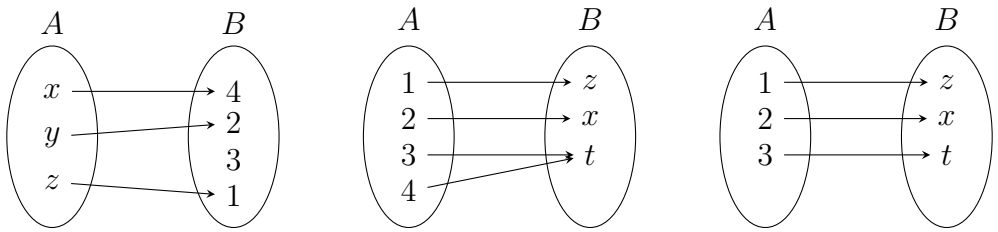
1. *Eine Abbildung $f : A \to B$ heißt surjektiv, falls es zu jedem $b \in B$ ein $a \in A$ gibt mit $f(a) = b$, d.h. falls für ihr Bild $f(A)$ gilt $f(A) = B$. Im Falle einer surjektiven Abbildung schreiben wir auch $f : A \twoheadrightarrow B$.*

2. *Eine Abbildung $f : A \to B$ heißt injektiv, falls aus $f(a_1) = f(a_2)$ folgt $a_1 = a_2$, d.h. aus $a_1 \neq a_2$ folgt stets $f(a_1) \neq f(a_2)$. Im Falle einer injektiven Abbildung schreiben wir auch $f : A \hookrightarrow B$.*

3. *Eine Abbildung heißt bijektiv, wenn sie surjektiv und injektiv ist. Im Falle einer bijektiven Abbildung schreiben wir auch $f : A \xrightarrow{\sim} B$.*

Bemerkungen 1.4.21.

1. Schema für eine injektive Abbildung (links), eine surjektive Abbildung (Mitte) und für eine bijektive Abbildung (rechts):



2. Für jede Menge M ist die Identitätsabbildung $\mathrm{id}_M : M \to M$ bijektiv.

3. Für jede Teilmenge $A \subset M$ ist die Inklusionsabbildung $\iota_A : A \to M$ injektiv.

Satz 1.4.22.
Sei $A \neq \emptyset$ und $f : A \to B$ eine Abbildung. Dann gilt:

1. f ist genau dann surjektiv, wenn es eine Abbildung $g : B \to A$ gibt, so dass $f \circ g = \mathrm{id}_B$ gilt. Man sagt dann auch, f habe ein Rechtsinverses.

2. f ist genau dann injektiv, wenn es eine Abbildung $g : B \to A$ gibt, so dass $g \circ f = \mathrm{id}_A$ gilt. Man sagt dann auch, f habe ein Linksinverses.

3. f ist genau dann bijektiv, wenn es eine Abbildung $g : B \to A$ gibt, so dass $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$ gilt. Man sagt dann auch, f habe eine inverse Abbildung oder Umkehrabbildung.

Beweis.

1. "$\Leftarrow$" Es gebe eine Abbildung $g : B \to A$, so dass $f \circ g = \mathrm{id}_B$ gilt. Sei $b \in B$ beliebig, setze $a := g(b) \in A$ und finde

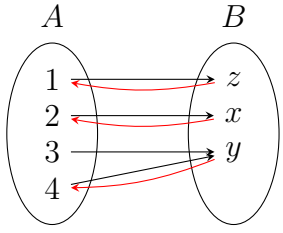
$$f(a) = f \circ g(b) = \mathrm{id}_B(b) = b\,.$$

Wir haben also mit a ein Urbild von b gefunden. Da b beliebig war, haben alle Elemente $b \in B$ Urbilder, die Abbildung f ist surjektiv.

"$\Rightarrow$" Sei f surjektiv, konstruiere g. Zu jedem $b \in B$ betrachte die Menge $f^{-1}(b) := \{a \in A \,|\, f(a) = b\}$. Sie ist nicht leer, weil f surjektiv sein soll. Wegen des Auswahlaxioms können wir ein Element $a_b \in f^{-1}(b)$ auswählen. Es gilt $f(a_b) = b$. Setze $g(b) := a_b$ und rechne

$$f \circ g(b) = f(a_b) = b\,.$$

Schematisch sieht das so aus:



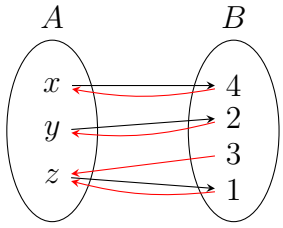
Schematisch sieht das so aus:
die schwarzen Pfeile zeigen die surjektive Abbildung f, die roten zeigen eine mögliche Wahl für das Rechtsinverse g. Wir haben als Urbild unter f für das Element $y \in B$ das Element $4 \in A$ gewählt; wir hätten aber auch $3 \in A$ wählen können.

2. "$\Leftarrow$" Seien $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$. Dann folgt $g(f(a_1)) = g(f(a_2)) \Leftrightarrow \mathrm{id}_a(a_1) = \mathrm{id}_A(a_2) \Leftrightarrow a_1 = a_2$. Also ist f injektiv.
"$\Rightarrow$" Definiere g wie folgt: wähle ein beliebiges $a_0 \in A$ und setze

$$g(b) := \begin{cases} a_0 & \text{falls } b \notin f(A) \\ \text{das eindeutige } a \in A \text{ mit } f(a) = b & \text{falls } b \in f(A). \end{cases}$$

Für jedes $a \in A$ ist $f(a) \in f(A)$, also ist $g \circ f(a) = g(f(a)) = a$ für alle $a \in A$. Es sollte klar sein, dass die Abbildung g nicht notwendigerweise eindeutig bestimmt ist.

Schematisch sieht das so aus:



Schematisch sieht das so aus:
die schwarzen Pfeile zeigen die injektive Abbildung f, die roten zeigen eine mögliche Wahl für das Linksinverse g. Wir haben als Bild unter g für die Elemente in der Bildmente B, die nicht im Bild $f(A)$ liegen, das Element $a_0 = z$ gewählt; wir hätten aber auch jedes andere Element von A wählen können.

3. "$\Leftarrow$" folgt sofort aus 1. und 2.
"$\Rightarrow$" Aus 1. folgt sofort, da f als bijektive Abbildung auch surjektiv ist, dass es eine Abbildung $g_1 : B \to A$ gibt, so dass $f \circ g_1 = \mathrm{id}_B$ gilt. Als bijektive Abbildung ist f auch injektiv; deshalb gibt es nach 2. eine Abbildung $g_2 : B \to A$, so dass $g_2 \circ f = \mathrm{id}_A$ gilt. Zu zeigen bleibt die Gleichheit $g_1 = g_2$. Diese folgt aus

$$g_2 = g_2 \circ \mathrm{id}_B = g_2 \circ (f \circ g_1) = (g_2 \circ f) \circ g_1 = \mathrm{id}_A \circ g_1 = g_1 \,.$$

Man beachte, dass hier die Assoziativität der Verkettung von Abbildungen eingeht.

□

Die im Falle bijektiver Abbildungen eindeutig bestimmte Abbildung g heißt Umkehrabbildung von f. Man schreibt auch $g = f^{-1}$.

Definition 1.4.23

1. *Für $n \in \mathbb{N}_0$ setzen wir*

$$\begin{aligned} \underline{n} &:= \{m \in \mathbb{N} \mid m \leq n\} = \{1, 2, \ldots, n\} \\ \underline{0} &:= \emptyset \end{aligned}$$

2. *Eine Menge M heißt endlich, wenn es ein $n \in \mathbb{N}_0$ und eine Bijektion*

$$f : \underline{n} \longrightarrow M$$

gibt. Induktiv zeigt man, dass $n \in \mathbb{N}$ eindeutig bestimmt ist. Die natürliche Zahl

$$\#(M) \cong |M| := n$$

heißt die Mächtigkeit oder Kardinalität der Menge M.

3. *Zwei Mengen A, B heißen gleichmächtig, wenn es eine bijektive Abbildung $f : A \to B$ gibt.*

Beispiele 1.4.24.

1. $\mathbb{N}$ und $\mathbb{Z}$ sind gleichmächtig, obgleich $\mathbb{N} \subsetneq \mathbb{Z}$. Eine Bijektion $f : \mathbb{N} \to \mathbb{Z}$ ist
$$f(a) = \begin{cases} \frac{a}{2} & \text{falls } a \text{ gerade} \\ -\frac{a+1}{2} & \text{falls } a \text{ ungerade} \end{cases}$$
2. Die Mengen $\mathbb{N}$ und $\mathbb{Q}$ sind gleichmächtig, nicht aber $\mathbb{N}$ und $\mathbb{R}$.
3. Wir betrachten Selbstabbildungen einer Menge Menge: Ist M eine endliche Menge, so gilt für alle Abbildungen $f : M \to M$: f bijektiv $\Leftrightarrow$ f injektiv $\Leftrightarrow$ f surjektiv. Ist aber M eine unendliche Menge, so gibt es Abbildungen $f : M \to M$, die injektiv, aber nicht surjektiv sind, und Abbildungen, die surjektiv, aber nicht injektiv sind.

   Wir zeigen als Beispiel, dass für eine endliche Menge $M \xrightarrow{f} M$ surjektiv ist, wenn f injektiv ist. Offenbar gibt die Korestriktion $\tilde{f} : M \to f(M)$ eine Abbildung, die injektiv und surjektiv ist, also bijektiv ist. Damit haben die M und $f(M)$ die gleiche Mächtigkeit. Haben eine endliche Menge und eine Teilmenge gleiche Mächtigkeit, so sind sie gleich. Also ist auch f surjektiv.

Mengen, die die gleiche Mächtigkeit wie $\mathbb{N}$ haben, heißen abzählbar unendlich.

Bemerkung 1.4.25.
Mit Hilfe von Abbildungen können wir eine weitere Bedeutung des kartesischen Produkts erklären. Seien M_1, M_2 und N Mengen. Dann gibt es für jede Menge N eine Bijektion von Mengen von Abbildungen
$$\mathrm{Abb}(M_1 \times M_2, N) \cong \mathrm{Abb}(M_1, \mathrm{Abb}(M_2, N)) \ .$$
Hierbei wird die Abbildung $\phi : M_1 \times M_2 \to N$ auf die Abbildung $\tilde{\phi} : M_1 \to \mathrm{Abb}(M_2, N)$ geschickt, die $m_1 \in M_1$ die Abbildung $m_2 \mapsto \phi(m_1, m_2)$ zuordnet. Umgekehrt wird einer Abbildung $\psi : M_1 \to \mathrm{Abb}(M_2, N)$ die Abbildung $\hat{\psi} : M_1 \times M_2 \to N$ mit $\hat{\psi}(m_1, m_2) = \psi(m_1)(m_2)$ zugeordnet. Man rechne nach, dass man so eine Bijektion von Mengen von Abbildungen erhält. Hätte man also nach einer Menge X gesucht, so dass sich die Menge $\mathrm{Abb}(M_1, \mathrm{Abb}(M_2, N))$ für *jede* Menge N als Menge von Abbildungen $\mathrm{Abb}(X, N)$ schreiben lässt, so wären wir auf das kartesische Produkt $X = M_1 \times M_2$ geführt worden.

Wir beschreiben eine kleine Anwendung von Bemerkung 1.4.25: ein deterministischer endlicher Automat ist ein einfaches Modell für einen Computer. Er besteht

- Aus einer endlichen Zustandsmenge Z, deren Elemente als Zustände des Rechners interpretiert werden können.
- Einem Eingabealphabet Σ; wir nehmen $\Sigma \cap Z = \emptyset$ an.
- Einer Übergangsfunktion $\delta : \Sigma \times Z \to Z$.
- Einem Startzustand $z_0 \in Z$ und einer Menge von Endzuständen $F \subset Z$.

Die Übergangsfunktion liefert uns Bemerkung 1.4.25 eine Funktion $\hat{\delta} : \Sigma \to \mathrm{Abb}(Z, Z)$. Die Abbildung $\hat{\delta}(s) : Z \to Z$ gibt an, wie die Eingabe von $s \in \Sigma$ aus dem Eingabealphabet den Zustand des Rechners verändert.

2 Algebraische Grundbegriffe

Die zentralen Begriffe der linearen Algebra sind die Begriffe des Vektorraums und der linearen Abbildung. In diesem Kapitel führen wir den Begriff des Vektorraums ein. Dazu müssen wir zunächst einige elementare algebraische Grundbegriffe einführen.

2.1 Gruppen

Wir wollen zunächst einige Eigenschaften unseres Modells $\mathbb{R}^2$ für die Ebene der Anschauung abstrahieren. Hierbei beschränken wir uns zunächst auf Eigenschaften der *Addition* von Elementen aus $\mathbb{R}^2$.

Definition 2.1.1
Eine Gruppe ist ein Paar $(G, \cdot)$, bestehend aus einer Menge G und einer Abbildung

$$\begin{aligned} \cdot : \quad & G \times G \to G \,, \\ & (a, b) \mapsto a \cdot b \,, \end{aligned}$$

genannt Verknüpfung, so dass die folgenden Eigenschaften erfüllt sind:

(G1) Für alle $a, b, c \in G$ gilt: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (Assoziativgesetz)

(G2) Es gibt ein Element $e \in G$, so dass gilt

- *a) Für alle $a \in G$ gilt $e \cdot a = a$*
- *b) Für jedes $a \in G$ gibt es ein $a' \in G$, so dass $a' \cdot a = e$ gilt.*

Man nennt ein solches Element e auch ein linksneutrales Element und a' ein linksinverses Element zu a.

Bemerkungen 2.1.2.

1. Die Gültigkeit des Kommutativgesetzs $a \cdot b = b \cdot a$ für alle $a, b \in G$ wird nicht gefordert.
2. Aus dem Assoziativgesetz (G1) folgt, dass alle möglichen Klammerungen eines mehrfachen Produkts das gleiche Ergebnis liefern. Daher können wir Klammern in mehrfachen Produkten weglassen.
3. Wegen (G2) gibt es wenigstens ein Element $e \in G$; die Menge G kann also nicht leer sein.

Satz 2.1.3.
Sei $(G, \cdot)$ eine Gruppe. Dann gilt:

1. Für jedes $a \in G$ gilt auch $a \cdot e = a$: jedes linksneutrale Element ist auch rechtsneutral. Wir sprechen daher von einem neutralen Element.
2. Das neutrale Element ist eindeutig.
3. Sei e das neutrale Element. Für alle $a \in G$ gilt dann auch $aa^{-1} = e$. Jedes linksinverse Element ist auch ein rechtsinverses Element. Wir sprechen daher von einem inversen Element.
4. Ist $e \in G$ das neutrale Element, so ist für jedes gegebene $a \in G$ das zugehörige Element a' aus (G2b) eindeutig. Es heißt inverses Element zu a, wir schreiben a^{-1} statt a'.

Beweis.

- Sei e ein linksneutrales Element in G und $a \in G$ beliebig. Wegen des Gruppenaxioms (G2b) gibt es wenigstens ein $a' \in G$, so dass $a'a = e$ gilt. Wiederum wegen (G2b) gibt es auch ein Inverses a'' zu a', für das also $a''a' = e$ gilt. Es folgt

$$\begin{aligned} a \cdot a' &\underset{(G2a)}{=} e \cdot (a \cdot a') \underset{(G2b)}{=} (a'' \cdot a') \cdot (a \cdot a')) \underset{(G1)}{=} a'' \cdot (a' \cdot (a \cdot a')) \\ &\underset{(G1)}{=} a'' \cdot ((a' \cdot a) \cdot a') \underset{(G2b)}{=} a'' \cdot (e \cdot a') \underset{(G2a)}{=} a'' \cdot a' \underset{(G2b)}{=} e \,. \end{aligned}$$

  Damit ist 3. für jedes linksneutrale Element e gezeigt: jedes bezüglich e linkinverse Element ist auch ein rechtinverses Element bezüglich e. Wir sprechen nur noch von einem inversen Element (bezüglich eines linksneutralen Elements).

- Sei $a \in G$, e ein linksneutrales Element und a' ein Inverses bezüglich e zu a, d.h. es gilt $a' \cdot a = e$. Wir rechnen

$$a \cdot e \underset{(G2b)}{=} a \cdot (a' \cdot a) \underset{(G1)}{=} (a \cdot a') \cdot a \underset{1.}{=} e \cdot a \underset{(G2a)}{=} a$$

  Damit ist 1. gezeigt: jedes linksneutrale Element ist auch rechtsneutral. Wir sprechen nur noch von einem neutralen Element.

- Sei $\widetilde{e}$ ein weiteres neutrales Element, also gelte für alle $a \in G$ die Gleichung $\widetilde{e} \cdot a = a$. Setze $a = e$ und erhalte $\widetilde{e} \cdot e = e$. Nach 2. für e gilt aber auch $\widetilde{e} \cdot e = \widetilde{e}$, also $e = \widetilde{e}$, die Eindeutigkeit 2. des neutralen Elements. Wir sprechen nur noch von *dem* neutralen Element.

- Sei $a \in G$ und seien a' und $\widetilde{a}'$ zu a inverse Elemente, d.h. $a' \cdot a = \widetilde{a}' \cdot a = e$.
  Wir rechnen, um die Eindeutigkeit 4. zu zeigen:

$$\widetilde{a}' = \widetilde{a}' \cdot e \underset{1.}{=} \widetilde{a}' \cdot (a \cdot a') \underset{(G1)}{=} (\widetilde{a}' \cdot a) \cdot a' = e \cdot a' \underset{(G2a)}{=} a' \,.$$

  Damit sprechen wir nur noch von *dem* zu a inversen Element.

□

Aus unserem Beweis folgt auch, dass das Inverse von e wieder e ist.

Satz 2.1.4. [Lösbarkeit von Gleichungen in einer Gruppe]
Sei $(G, \cdot)$ eine Gruppe und seien $a, b \in G$. Dann gilt

1. Es gibt ein eindeutiges $x \in G$, für das $x \cdot a = b$ gilt.
2. Es gibt ein eindeutiges $y \in G$, für das $a \cdot y = b$ gilt.
3. Für alle $a \in G$ gilt $(a^{-1})^{-1} = a$.
4. $(ab)^{-1} = b^{-1}a^{-1}$

Beweis.

1. Wenn es eine Lösung x der Gleichung $x \cdot a = b$ gibt, so ist diese eindeutig:

$$x = x \cdot e = x \cdot (a \cdot a^{-1}) = (x \cdot a) \cdot a^{-1} = b \cdot a^{-1} \,.$$

   Hierbei haben wir das neutrale Element, das Inverse von a und die Assoziativität benutzt, bevor wir im letzten Schritt die Annahme verwendet haben, dass x eine Lösung ist.

   In der Tat ist $x := b \cdot a^{-1}$ auch wirklich eine Lösung, denn es gilt

$$x \cdot a = (b \cdot a^{-1}) \cdot a = b \cdot (a^{-1} \cdot a) = b \cdot e = b \,.$$

2. Analog folgt $y = a^{-1} \cdot b$. Man beachte die Reihenfolge von a und b in 1. und 2.

3. $(a^{-1})^{-1}$ ist das nach Satz 2.1.3 eindeutige Inverse von a^{-1}; also gilt $(a^{-1})^{-1}a^{-1} = e$ per Definition. Andererseits gilt wegen Satz 2.1.3.4 $a \cdot a^{-1} = e$, so dass auch a ein Inverses von a^{-1} ist. Da das Inverse eindeutig ist, folgt $a = (a^{-1})^{-1}$.
Beachten Sie, dass man aus Eindeutigkeitsaussagen algebraische Gleichungen folgern kann!

4. Wir rechnen mit Hilfe des Assoziativgesetzes:

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

Also ist $b^{-1} \cdot a^{-1}$ das eindeutige Inverse von $a \cdot b$, d.h.

$$(a \cdot b)^{-1} = b^{-1} \cdot a^{-1} \ .$$

Dies kennt man auch aus dem richtigen Leben: zieht man morgens zuerst die Socken und dann die Schuhe an, so zieht man abends zuerst die Schuhe aus und dann die Socken (und nicht umgekehrt).

□

Beispiele 2.1.5.

1. $(\mathbb{Z}, +)$ ist eine Gruppe mit $e = 0$ und $a^{-1} = -a$.

2. Ebenso sind $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ Gruppen bezüglich der Addition.

3. $(\mathbb{R}^2, +)$ ist eine Gruppe mit

$$e = 0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ und } x^{-1} = -x = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

4. $(\mathbb{N}, +)$ ist keine Gruppe, da es zu $a \neq 0$ kein Inverses in $\mathbb{N}$ gibt.

5. $(\mathbb{Q} \setminus \{0\}, \cdot)$ ist eine Gruppe mit $e = 1$ und $a^{-1} = \frac{1}{a}$. Dagegen ist $(\mathbb{Q}, \cdot)$ keine Gruppe, denn zu $0 \in \mathbb{Q}$ gibt es kein (multiplikatives) Inverses.

6. Sei $M \neq \emptyset$ eine nichtleere Menge. Setze

$$\mathrm{Sym}\,(M) = \{f : M \to M \mid f \ \text{ bijektiv}\}$$

Dann ist $(\mathrm{Sym}\,(M), \circ)$ eine Gruppe:

(G1) Die Komposition von Abbildungen ist assoziativ, siehe die Bemerkung nach Definition 1.4.17.

(G2a) $e = \mathrm{id}_M$

(G2b) Das zu $f \in \mathrm{Sym}\,(M)$ inverse Element ist die Umkehrabbildung aus Satz 1.4.22.3.
$(\mathrm{Sym}\,(M), \circ)$ heißt die symmetrische Gruppe von M oder die Permutationsgruppe von Elementen von M.

Speziell für die Menge $\underline{n} = \{1, 2, \ldots n\}$ schreibt man $\mathrm{Sym}\,(\underline{n}) =: S_n$.

Definition 2.1.6
Eine Gruppe $(G, \cdot)$ heißt abelsch oder kommutativ, falls für alle $a, b \in G$ gilt $a \cdot b = b \cdot a$.

Bemerkungen 2.1.7.

1. Abelsche Gruppen werden oft additiv geschrieben und das neutrale Element dann mit 0 bezeichnet.

2. Die Beispiele $1, 2, 3$ und 5 sind abelsche Gruppen.

3. Die symmetrische Gruppe ist im allgemeinen nicht abelsch. Gegenbeispiel: Sei $M = \underline{3} = \{1, 2, 3\}$ und seien f, g die beiden bijektiven Abbildungen

x	1	2	3
$f(x)$	2	1	3
$g(x)$	1	3	2

Dann ist

$$f \circ g(1) = f(g(1)) = f(1) = 2$$
$$g \circ f(1) = g(f(1)) = g(2) = 3\ ,$$

also ist $f \circ g \neq g \circ f$.

Definition 2.1.8
Sei $(G, \cdot)$ eine Gruppe. Eine Teilmenge $H \subset G$ heißt Untergruppe, falls sie den folgenden Axiomen genügt:

(UG1) $H \neq \emptyset$. *H ist nicht die leere Menge.*

(UG2) *Für alle $a, b \in H$ gilt $a \cdot b \in H$. Wir sagen auch, dass H unter der Verknüpfung $\cdot$ von G abgeschlossen ist.*

(UG3) *Für alle $a \in H$ gilt $a^{-1} \in H$. Wir sagen auch, dass H unter der Inversenbildung abgeschlossen ist.*

Satz 2.1.9.
Sei $(G, \cdot)$ eine Gruppe mit neutralem Element e und $H \subset G$ eine Untergruppe. Dann gilt

1. $e \in H$

2. Mit der Einschränkung der Multiplikation

$$\cdot : \ G \times G \to G$$

auf

$$\cdot : \ H \times H \to H$$

ist $(H, \cdot)$ selbst eine Gruppe.

Beweis.

1. Nach (UG1) ist $H \neq \emptyset$, somit gibt es wenigstens ein $a \in H$. Nach (UG3) ist dann auch $a^{-1} \in H$ und nach (UG2) ist $e = a^{-1} \cdot a \in H$.

2. Wegen (UG2) ist das Bild der Einschränkung der Multiplikation auf H in H.
   Die Gruppenaxiome (G1) und (G2) gelten für alle Elemente von G und somit insbesondere für alle Elemente von H; das Inverse liegt wegen des Axioms (UG3) in H.

□

Bemerkungen 2.1.10.

1. Untergruppen abelscher Gruppen sind abelsch.

2. Eine Untergruppe einer Untergruppe von G ist selbst Untergruppe von G.

Beispiele 2.1.11.

1. Sei $(G, \cdot) = (\mathbb{Q}, +)$ die additive Gruppe der rationalen Zahlen. Dann ist $H = \mathbb{Z}$ eine Untergruppe. Sei $(G, \cdot) = (\mathbb{R}, +)$. Dann ist $H = (\mathbb{Q}, +)$ eine Untergruppe. Wegen Bemerkung 2.1.10.2 ist $H = (\mathbb{Z}, +)$ Untergruppe von $(\mathbb{R}, +)$.

2. Sei $(G, \cdot) = (\mathbb{Z}, +)$. Dann ist $H = \mathbb{N}$ keine Untergruppe, denn das Axiom (UG3) ist nicht erfüllt.

3. Sei $(G, \cdot)$ eine beliebige Gruppe. Dann ist

   - $H = \{e\}$ eine Untergruppe, die sogenannte triviale Untergruppe.
   - $H = G$ eine Untergruppe. Jede Gruppe ist Untergruppe von sich selbst.

Wir wollen eine Klasse von Abbildungen zwischen Gruppen auszeichnen:

Definition 2.1.12
Seien $(G, \cdot)$ und $(H, \odot)$ Gruppen. Eine Abbildung

$$\Phi : G \to H$$

heißt Gruppenhomomorphismus, falls für alle $a, b \in G$ gilt

$$\Phi(a \cdot b) = \Phi(a) \odot \Phi(b) .$$

Beispiele 2.1.13.

1. $G = H = \mathbb{Z}$ mit Addition ganzer Zahlen. Wähle ein festes $m \in \mathbb{Z}$ und betrachte die Abbildung

   $$\Phi_m : \quad \mathbb{Z} \to \mathbb{Z}$$

   mit $\Phi_m(k) = mk$. In der Tat gilt für alle $k, l \in \mathbb{Z}$

   $$\Phi_m(k + l) = m(k + l) = mk + ml = \Phi_m(k) + \Phi_m(l)$$

2. Sei $G = (\mathbb{R}, +)$ und $H = (\mathbb{R}_{>0}, \cdot)$. Wähle ein festes ein $\alpha \in \mathbb{R}$ und betrachte die Exponentialabbildung

   $$\Phi_\alpha : \mathbb{R} \to \mathbb{R}_{>0}$$

   mit $\Phi_\alpha(x) = \mathrm{e}^{\alpha x}$. Dann gilt

   $$\Phi_\alpha(x + y) = \mathrm{e}^{\alpha(x+y)} = \mathrm{e}^{\alpha x}\mathrm{e}^{\alpha y} = \Phi_\alpha(x)\Phi_\alpha(y) .$$

Satz 2.1.14.
Seien G und H Gruppen mit neutralen Elementen e_G bzw. e_H. Sei $\Phi : G \to H$ ein Gruppenhomomorphismus. Dann gilt:

1. $\Phi(e_G) = e_H$
2. Für alle $g \in G$ gilt $\Phi(g^{-1}) = \Phi(g)^{-1}$.
3. Ist Φ bijektiv, so ist
$$\Phi^{-1} : \quad H \to G$$
ebenfalls ein Gruppenhomomorphismus.

Ein bijektiver Gruppenhomomorphismus heißt Gruppenisomorphismus. Zwei Gruppen G, H heißen isomorph, wenn es einen Gruppenisomorphismus $\Phi : G \to H$ gibt; in Zeichen: $G \cong H$. "Isomorph sein" ist eine Äquivalenzrelation auf der Menge der Gruppen: die Relation ist reflexiv, weil die Identität ein Isomorphismus ist. Wegen Satz 2.1.14.3 ist die Relation symmetrisch. Sind $\Phi : G \to H$ und $\Psi : H \to K$ Gruppenisomorphismen, so ist auch $\psi \circ \Phi : G \to K$ ein Gruppenisomorphismus; daher ist die Relation transitiv.

Es kann für zwei gegebene Gruppen durchaus mehrere Gruppenisomorphismen geben, von denen keiner ausgezeichnet ist. (Man sagt auch, die Gruppen sind nicht kanonisch isomorph.)

Beweis.

1. Wir rechnen zunächst: $\Phi(e_G) = \Phi(e_G \cdot e_G) = \Phi(e_G) \cdot \Phi(e_G)$. Daraus folgt
$$\begin{aligned} e_H &= \Phi(e_G)^{-1}\Phi(e_G) = \Phi(e_G)^{-1}[\Phi(e_G)\Phi(e_G)] \\ &= [\Phi(e_G)^{-1}\Phi(e_G)] \cdot \Phi(e_G) = e_H \cdot \Phi(e_G) = \Phi(e_G). \end{aligned}$$
2. Für alle $g \in G$ gilt
$$\Phi(g^{-1})\Phi(g) = \Phi(g^{-1}g) = \Phi(e_G) \underset{1.}{=} e_H \ .$$
Aus der Eindeutigkeit der Inversen $\Phi(g)^{-1}$ von $\Phi(g)$ folgt $\Phi(g^{-1}) = \Phi(g)^{-1}$.
3. Seien $h, h' \in H$. Wir rechnen, da Φ bijektiv ist
$$\begin{aligned} \Phi^{-1}(h) \cdot \Phi^{-1}(h') &= (\Phi^{-1} \circ \Phi)\Big(\Phi^{-1}(h) \cdot \Phi^{-1}(h')\Big) = \Phi^{-1}\left(\Phi\Big(\Phi^{-1}(h) \cdot \Phi^{-1}(h')\Big)\right) \\ &= \Phi^{-1}\left(\Phi\Big(\Phi^{-1}(h)\Big) \cdot \Phi\Big(\Phi^{-1}(h')\Big)\right) \text{ da } \Phi \text{ ein Gruppenhomomorphismus ist} \\ &= \Phi^{-1}(h \cdot h') \ . \end{aligned}$$

□

Beispiele 2.1.15.

1. Für den Gruppenhomomorphismus $\Phi_1 : \mathbb{R} \to R_{>0}$, $\Phi_1(x) = e^x$, die Exponentialfunktion aus Beispiel 2.1.13.2, bedeuten die Aussagen von Satz 2.1.14 folgendes:
   (a) $e^0 = 1$
   (b) $e^{-x} = (e^x)^{-1} = \frac{1}{e^x}$.

(c) Die Umkehrfunktion

$$\Phi_1^{-1} =: \log \,:\, \mathbb{R}_{>0} \to \mathbb{R}$$

ist ein Gruppenhomomorphismus, d.h. es gilt

$$\log(xy) = \log(x) + \log(y) \quad \text{für} \quad x, y \in \mathbb{R}_{>0} \,.$$

Insbesondere sind die Gruppen $(\mathbb{R}, +)$ und $(\mathbb{R}_{>0}, \cdot)$ isomorph. Allerdings liefert jede der Abbildungen Φ_α mit $\alpha \neq 0$ einen Isomorphismus; es gibt genau so wenig eine ausgezeichnete Isomorphie der additiven Gruppe $(\mathbb{R}, +)$ auf die multiplikative Gruppe $(\mathbb{R}_{>0}, \cdot)$ wie es für den Logarithmus eine ausgezeichnete Basis gibt.

2. Die Homomorphismen $\Phi_m : \mathbb{Z} \to \mathbb{Z} \quad k \mapsto mk$ aus Beispiel 2.1.13.1 sind für jedes $m \neq 0$ injektiv, aber nur für $m = \pm 1$ ein Gruppen*iso*morphismen.

Definition 2.1.16
Seien G und H Gruppen und sei $\Phi : G \to H$ ein Gruppenhomomorphismus. Dann heißt das Urbild des neutralen Elements $e_H \in H$

$$\ker(\Phi) := \Phi^{-1}(e_H) = \{g \in G |\, \Phi(g) = e_H\}$$

der Kern von Φ.

Satz 2.1.17.
Sei $\Phi : G \to H$ ein Gruppenhomomorphismus. Dann gilt:

1. Der Kern $\ker(\Phi)$ ist eine Untergruppe von G, enthält also insbesondere e_G.
2. Die Abbildung Φ ist injektiv $\Leftrightarrow \ker(\Phi) = \{e_G\}$
3. Das Bild $\mathrm{Im}\,(\Phi)$ ist eine Untergruppe von H.
4. Die Abbildung Φ ist surjektiv $\Leftrightarrow$ $\mathrm{Im}\,\Phi = H$.

Beweis.
1. Wir überprüfen die Untergruppenaxiome aus Definition 2.1.8:
   (UG1) Wegen $\Phi(e_G) = e_H$ ist $e_G \in \ker \Phi$, also $\ker \Phi \neq \emptyset$.
   (UG2) Seien $g, g' \in \ker \Phi$. Dann folgt

   $$\Phi(gg') = \Phi(g)\Phi(g') = e_H \cdot e_H = e_H \,,$$

   also ist $gg' \in \ker \Phi$.
   (UG3) Sei $g \in \ker \Phi$. Dann folgt

   $$\Phi(g^{-1}) = \Phi(g)^{-1} = e_H^{-1} = e_H \,,$$

   also $g^{-1} \in \ker \Phi$.

2. "$\Rightarrow$" Sei Φ injektiv. Sei $g \in \ker \Phi$, d.h. $\Phi(g) = e_H = \Phi(e_G)$. Die Injektivität von Φ impliziert $g = e_G$, also $\ker \Phi = \{e_G\}$.
   "$\Leftarrow$" Sei $\ker \Phi = \{e_G\}$. Seien $g, g' \in G$ mit $\Phi(g) = \Phi(g')$. Zu zeigen ist $g = g'$. Wir rechnen

   $$\Phi(g(g')^{-1}) = \Phi(g)\Phi(g')^{-1} = \Phi(g)\Phi(g)^{-1} = e_H \,.$$

   Also $g(g')^{-1} \in \ker \Phi = \{e_G\}$. Das heißt $g(g')^{-1} = e_G$, also $g = g'$. Also ist Φ injektiv.

3. Wir überprüfen die Untergruppenaxiome aus Definition 2.1.8:

   (UG1) $e_H = \Phi(e_G) \in \operatorname{Im}\Phi$, also ist $\operatorname{Im}\Phi$ nicht leer.
   (UG2) Seien $h, h' \in \operatorname{Im}\Phi$. Dann gibt es $g, g' \in G$ mit $h = \Phi(g)$ und $h' = \Phi(g')$. Es folgt

   $$hh' = \Phi(g)\Phi(g') = \Phi(g \cdot g') \in \operatorname{Im}\Phi \ .$$

   (UG3) Sei $h \in \operatorname{Im}\Phi$, finde $g \in G$ mit $\Phi(g) = h$. Dann gilt

   $$h^{-1} = \Phi(g)^{-1} = \Phi(g^{-1}) \in \operatorname{Im}\Phi \ .$$

4. Klar nach der Definition von Surjektivität.

□

Betrachtung 2.1.18.

- Wir wählen ein festes $m \in \mathbb{N}$. Betrachte auf der der Gruppe $(\mathbb{Z}, +)$ zu Grunde liegenden Menge wie in Beispiel 1.4.13.3 die Äquivalenzrelation

  $$R_m = \{(a, b) \mid \ m \ \text{ teilt } \ a - b\} \subset \mathbb{Z} \times \mathbb{Z} \ .$$

  Die Quotientenmenge $\mathbb{Z}/R_m$ wird auch mit $\mathbb{Z}/m\mathbb{Z} = \mathbb{Z}_m$ bezeichnet.
  Eine Äquivalenzklasse besteht für $m \neq 0$ aus genau denjenigen ganzen Zahlen, die bei Division durch m den gleichen Rest aus $\{0, 1, \ldots, m-1\}$ lassen. Eine Äquivalenzklasse wird auch Restklasse modulo m genannt. Für jedes $0 \leq r \leq m-1$ ist

  $$r + m\mathbb{Z} = \{x \in \mathbb{Z} | \ m \text{ teilt } x - r\}$$

  eine Restklasse.

  Wir schreiben $a = a' \mod m$ und sagen "a ist kongruent a' modulo m", wenn a und a' in der gleichen Restklasse liegen. Für $m = 0$ ist die Äquivalenzrelation die Gleichheit und $\mathbb{Z}/0\mathbb{Z} \cong \mathbb{Z}$.

- Betrachte die kanonische Abbildung

  $$\text{can} : \qquad \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \quad ,$$

  die jeder ganzen Zahl a ihre Äquivalenzklasse

  $$a + m\mathbb{Z} =: \overline{a} = a \bmod m/Z = a \bmod m$$

  zuordnet. a ist also ein Repräsentant der Klasse $\overline{a}$.

- Nun ist $(\mathbb{Z}, +)$ auch eine abelsche Gruppe. Wir wollen auch auf der Quotientenmenge eine Gruppenstruktur definieren, und zwar so, dass die kanonische Abbildung ein Gruppenhomomorphismus wird. Dann soll gelten:

  $$\overline{a + b} = \text{can}(a + b) = \text{can}(a) + \text{can}(b) = \overline{a} + \overline{b} \ .$$

  Versuchsweise definieren wir also eine Addition auf der Restklassenmenge folgendermaßen: für zwei Restklassen wählen wir Repräsentanten $a \in \overline{a}$ und $b \in \overline{b}$. Dann setzen wir als Wert der Verknüpfung versuchsweise die Restklasse von $a + b$. Das ist nur sinnvoll, wenn

die Addition *nicht* von der Auswahl der Repräsentanten abhängt. Man sagt dann, dass die Verknüpfung *wohldefiniert* ist.

Dass dies hier so ist, sieht man folgendermaßen: ist $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, so ist $a - a' = mk$ und $b - b' = ml$ mit $k, l \in \mathbb{Z}$.
Dann ist

$$a + b = a' + b' + mk + ml = a' + b' + m(k + l) \ ,$$

also ist $\overline{a + b} = \overline{a' + b'}$.

Satz 2.1.19.

1. Ist $m \in \mathbb{N}$, so ist die Restklassenmenge $\mathbb{Z}/m\mathbb{Z}$ mit der oben erklärten Addition eine abelsche Gruppe, die zyklische Gruppe der Ordnung m.

2. Die kanonische Abbildung

$$\begin{array}{rcl} \mathbb{Z} & \to & \mathbb{Z}/m\mathbb{Z} \\ a & \mapsto & \overline{a} = a + m\mathbb{Z} \end{array}$$

ist ein surjektiver Gruppenhomomorphismus mit Kern $m\mathbb{Z}$.

Beweis.

1. Die Assoziativität vererbt sich von $\mathbb{Z}$:

$$\begin{aligned} (\overline{a} + \overline{b}) + \overline{c} = \overline{a + b} + \overline{c} = \overline{(a + b) + c} = \overline{a + (b + c)} \\ = \overline{a} + \overline{b + c} = \overline{a} + (\overline{b} + \overline{c}) \ . \end{aligned}$$

Man mache sich in diesen Gleichungen zur Übung klar, welche Pluszeichen die Addition in $\mathbb{Z}$ und welche die Addition in $\mathbb{Z}/m\mathbb{Z}$ bezeichnen. Das neutrale Element ist die Restklasse $\overline{0}$, denn

$$\overline{a} + \overline{0} = \overline{a + 0} = \overline{a} = \overline{0} + \overline{a} \ .$$

Das Inverse von $\overline{a}$ ist $\overline{-a}$. Auch die Kommutativität vererbt sich von $\mathbb{Z}$.

2. Ist nach Definition der Addition auf $\mathbb{Z}_m$ klar.

□

2.2 Ringe und Körper

Auf den Mengen $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$ sind zwei Operationen erklärt: eine Addition und eine Multiplikation. Wir formalisieren deren Eigenschaften.

Definition 2.2.1

1. Eine Menge R zusammen mit zwei Verknüpfungen

$$\begin{array}{rcccc} + : & R \times R \to R & (a, b) & \mapsto & a + b \\ \cdot : & R \times R \to R & (a, b) & \mapsto & a \cdot b \end{array}$$

heißt ein Ring, wenn gilt:

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

(R2) Die Multiplikation ist assoziativ.

(R3) Es gelten die beiden Distributivgesetze: für alle $a, b, c \in R$ gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c \qquad (a + b) \cdot c = a \cdot c + b \cdot c$$

2. *Ein Element $1 \in R$ heißt Einselement, wenn für alle $a \in R$ gilt $1 \cdot a = a \cdot 1 = a$. Ein Ring mit Einselement heißt auch unitärer Ring oder unitaler Ring.* [3]

3. *Ein Ring heißt kommutativ, wenn für alle $a, b \in R$ gilt $a \cdot b = b \cdot a$.*

Bemerkungen 2.2.2.

1. Man beachte, dass die Addition in einem Ring immer kommutativ ist.

2. Wir vereinbaren für alle Ringe die Regel "Punkt vor Strich".

3. Ist R ein Ring und $0 \in R$ das neutrale Element der abelschen Gruppe $(R, +)$, genannt das Nullelement, so gilt für alle $a \in R$

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

Dies folgt aus der Rechnung

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

Beispiele 2.2.3.

1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ sind unitäre kommutative Ringe. Der Ring $m\mathbb{Z}$ mit $m \in \mathbb{N}$ ein kommutativer Ring, aber ist für $m \neq \pm 1$ nicht unitär.

2. Ist $I \subset \mathbb{R}$ ein Intervall und

$$R := \{f : \ I \to \mathbb{R}\}$$

die Menge der reellwertigen Funktionen, so definieren wir Verknüpfungen durch Operationen auf den Funktionswerten:

$$\begin{aligned}(f + g)(x) :=& f(x) + g(x) \\ (f \cdot g)(x) :=& \ f(x) \cdot g(x)\end{aligned}$$

Sie versehen R mit der Struktur eines kommutativen unitären Rings. Allgemeiner sei M eine Menge und R ein Ring. Dann kann man die Menge der Abbildungen $\{f : M \to R\}$ mit Hilfe der Ringstruktur auf dem Bild der Abbildungen mit der Struktur eines Ringes versehen.

3. Auf der abelschen Gruppe $\mathbb{Z}/m\mathbb{Z}$ mit $m \in \mathbb{N}$ aus Betrachtung 2.1.18 kann man durch

$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$

eine Multiplikation definieren. Sie ist wohldefiniert, denn gilt

$$a - a' = mk \qquad \text{und} \qquad b - b' = ml$$

[3] Die Bezeichnung "unitär" ist gebräuchlicher, sollte aber nicht mit unitären Matrizen verwechselt werden.

mit $k, l \in \mathbb{Z}$, so folgt

$$a \cdot b = (a' + mk)(b' + ml) = a'b' + m(kb' + a'l + mkl) \,,$$

so dass die Multiplikation nicht von der Wahl der Repräsentanten abhängt. Die Assoziativität der Multiplikation und die Distributivgesetze vererben sich von $\mathbb{Z}$. Es liegt ein kommutativer Ring mit Eins $\overline{1}$ vor. Wir rechnen zum Beispiel in $\mathbb{Z}/4\mathbb{Z}$:

$$\overline{2} \cdot \overline{2} = \overline{4} = \overline{0} \,.$$
$$\overline{2} \cdot \overline{1} = \overline{2} \qquad \overline{2} \cdot \overline{3} = \overline{6} = \overline{2} \,.$$

Die Menge $(\mathbb{Z}/m\mathbb{Z} \setminus \{\overline{0}\}, \cdot)$ ist also nicht immer eine Gruppe, denn für $m = 4$ hat die Restklasse $\overline{2}$ offenbar kein (multiplikatives) Inverses.

Definition 2.2.4
Ein kommutativer Ring R heißt nullteilerfrei oder integer, wenn für $a, b \in R$ aus $a \cdot b = 0$ stets $a = 0$ oder $b = 0$ folgt.

Lemma 2.2.5.
Der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ ist genau dann integer, wenn m eine Primzahl ist.

Beweis.

- Ist m keine Primzahl, so gibt es $1 < k,\ l < m$ mit $m = k \cdot l$. Also ist $\overline{k}, \overline{l} \neq \overline{0}$, aber $\overline{0} = \overline{m} = \overline{k \cdot l} = \overline{k} \cdot \overline{l}$. (Zum Beispiel gilt modulo 6, dass $\overline{2} \cdot \overline{3} = \overline{0}$.)
- Sei umgekehrt m prim und gelte $\overline{k} \cdot \overline{l} = \overline{0}$, so ist

  $$kl = r \cdot m$$

  für ein $r \in \mathbb{Z}$. Wegen der Eindeutigkeit der Primzahlzerlegung in $\mathbb{Z}$ teilt die Primzahl m entweder k oder l, also $\overline{k} = \overline{0}$ oder $\overline{l} = \overline{0}$.

□

Definition 2.2.6

1. *Ist R ein Ring und $R' \subset R$ eine Teilmenge, so heißt R' Unterring, wenn $(R', +)$ Untergruppe von $(R, +)$ ist und R' bezüglich der Multiplikation abgeschlossen ist: mit $a, b \in R'$ gilt stets $a \cdot b \in R'$. (Bei unitären Unterringen unitärer Ringe fordert man zusätzlich, dass das Einselement von R' gleich dem Einselement von R ist.)*
2. *Seien $(R, +, \cdot)$ und $(S, \oplus, \odot)$ Ringe, so heißt eine Abbildung*

   $$\varphi : \qquad R \to S$$

   Ringhomomorphismus, wenn für alle $a, b \in R$ gilt

   $$\varphi(a + b) = \varphi(a) \oplus \varphi(b)$$
   $$\varphi(a) \odot \varphi(b) = \varphi(a \cdot b)$$

   Für unitäre Ringe fordern wir noch $\varphi(1_R) = 1_S$.

Zum Beispiel ist $m\mathbb{Z}$ ein Unterring von $\mathbb{Z}$, aber für $m \neq \pm 1$ kein Unterring mit Eins. Die kanonische Abbildung von $\mathbb{Z}$ auf $\mathbb{Z}/m\mathbb{Z}$ aus Betrachtung 2.1.18, die durch $a \mapsto a + m\mathbb{Z}$ gegeben ist, ist ein (unitärer) Ringhomomorphismus.
In einem nullteilerfreien Ring R ist die Multiplikation auf $R\setminus\{0\}$ abgeschlossen. Aber $(R\setminus\{0\}, \cdot)$ ist deshalb noch nicht unbedingt eine Gruppe: zum Beispiel gibt es in $\mathbb{Z}\setminus\{0\}$ es keine multiplikativen Inversen. Die Existenz solcher Inversen fordert man in der folgenden mathematischen Struktur:

Definition 2.2.7
Ein Körper ist eine Menge K mit zwei Verknüpfungen $+, \cdot$

$$+, \cdot : \qquad K \times K \to K \ ,$$

für die die folgenden Axiome gelten:

(K1) $(K, +)$ *ist eine abelsche Gruppe. Das neutrale Element der Addition wird mit* 0 *bezeichnet.*

(K2) $(K \setminus \{0\}, \cdot)$ *ist eine abelsche Gruppe, deren neutrales Element wir mit* 1 *bezeichnen.*

(K3) Es gilt das Distributivgesetz: für alle $a, b, c \in K$ gilt

$$(a + b) \cdot c = a \cdot c + b \cdot c \ .$$

Bemerkungen 2.2.8.

1. Wie in Bemerkung 2.2.2.2 folgt aus der Rechnung

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

   dass $a \cdot 0 = 0$ für alle $a \in K$. Weil $(K\setminus\{0\}, \cdot)$ eine Gruppe ist, gilt das Assoziativitätsgesetz $(ab)c = a(bc)$ für alle von Null verschiedenen $a, b, c \in K$. Ist wenigstens eines der Elemente a, b, c gleich Null, so reduziert sich das Assoziativitätsgesetz der Multiplikation auf die Gleichheit $0 = 0$.

   Ein Körper ist also insbesondere ein kommutativer Ring. Er ist integer, weil $(K \setminus \{0\}, \cdot)$ eine Gruppe ist und somit das Produkt von zwei von Null verschiedenen Elementen wieder von Null verschieden ist.

2. Seien K und K' Körper. Ein Körperhomomorphismus ist ein Ringhomomorphismus $\varphi : K \to K'$. Da er ein Gruppenhomomorphismus der Gruppen $(K, +)$ und $(K \setminus \{0\}), \cdot)$ ist, erhält er nach Satz 2.1.14 die neutralen Elemente, $\varphi(0_K) = 0_{K'}$ und $\varphi(1_K) = 1_{K'}$ und additive und multiplikative Inverse, $\varphi(-a) = -\varphi(a)$ für alle $a \in K$ und $\varphi(a^{-1}) = \varphi(a)^{-1}$ für $a \neq 0$.

   Ist K ein Körper und E ein unitärer Ring, so ist jeder Ringhomomorphismus $\varphi : K \to E$ injektiv. Da φ insbesondere ein Gruppenhomomorphismus der additiven Gruppen ist, reicht es wegen Satz 2.1.17.2, den Kern von φ zu berechnen. Angenommen, es wäre $\varphi(a) = 0$ für $a \neq 0$. Dann gilt

   $$\varphi(1) = \varphi(aa^{-1}) = \varphi(a)\varphi(a)^{-1} = 0\varphi(a)^{-1} = 0 \ ,$$

   was für einen unitären Ringhomomorphismus nicht gelten kann.

3. Ein Unterkörper $K' \subset K$ eines Körper K ist eine nicht-leere Teilmenge, die unter Addition und Multiplikation abgeschlossen sowie unter Bildung additiver und multiplikativer Inverser abgeschlossen ist.

4. Ist $(K \setminus \{0\}, \cdot)$ eine nicht-abelsche Gruppe und $(K, +, \cdot)$ ein Ring, so heißt K ein Schiefkörper.

5. Wir nennen die Verknüpfung $+$ Addition und die Verknüpfung $\cdot$ Multiplikation. Wir lassen auch oft den Punkt "$\cdot$" weg, wenn wir die Multiplikation schreiben:

$$a \cdot b =: ab$$

Das zu $a \in K$ bezüglich der Addition $+$ inverse Element schreiben wir als $-a$. Das zu $a \in K \setminus \{0\}$ bezüglich der Multiplikation $\cdot$ inverse Element schreiben wir als $\frac{1}{a} = a^{-1}$. Wir setzen wie von den rationalen Zahlen her vertraut

$$a + (-b) =: a - b \quad \text{ und } a \cdot (\frac{1}{b}) =: \frac{a}{b}\,.$$

Das neutrale Element bezüglich $+$ schreiben wir als 0, das neutrale Element bezüglich $\cdot$ als 1. Es ist $1 \in K \setminus \{0\}$, also $1 \neq 0$. Ein Körper hat also wenigstens zwei Elemente.

Beispiele 2.2.9.

1. $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper. Die rationalen Zahlen $\mathbb{Q}$ sind ein Unterkörper des Körpers $\mathbb{R}$ der reellen Zahlen.

2. $(\mathbb{Z}, +, \cdot)$ ist kein Körper, sondern nur ein integrer kommutativer Ring mit Eins, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist.

3. Sei $p \in \mathbb{Z}$ prim. Wir wissen schon, dass $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ ein integrer Ring mit Eins $\bar{1}$ ist. Das Distributivgesetz (K3) vererbt sich unmittelbar von $\mathbb{Z}$. Es bleibt zu zeigen, dass jedes $\bar{r} \neq \bar{0}$, $\bar{r} \in \mathbb{Z}/p\mathbb{Z}$ ein multiplikatives Inverses hat. Betrachte hierzu für gegebenes $\bar{r}$ die Selbstabbildung der endlichen Menge

$$\begin{array}{rccc} \varphi_{\bar{r}}: & \mathbb{Z}/p\mathbb{Z} \setminus \{\bar{0}\} & \to & \mathbb{Z}/p\mathbb{Z} \setminus \{\bar{0}\} \\ & \bar{s} & \mapsto & \bar{r} \cdot \bar{s} \end{array}$$

Sie ist injektiv, denn $\bar{r} \cdot \bar{s} = \bar{r} \cdot \bar{s'}$ ist äquivalent zu $\bar{r} \cdot (\bar{s} - \bar{s'}) = 0$. Wir wissen aus Lemma 2.2.5, dass der Ring $\mathbb{Z}/p\mathbb{Z}$ integer ist, was impliziert $\bar{s} - \bar{s'} = 0$, also $\bar{s} = \bar{s'}$ ist. Als injektive Abbildung einer *endlichen* Menge ist $\varphi_{\bar{r}}$ nach Beispiel 1.4.24.3 auch surjektiv, also liegt $\bar{1}$ in ihrem Bild.

$\mathbb{Z}/p\mathbb{Z}$ ist also ein Körper; er hat p Elemente. In diesem Körper ist die p-fache Summe der Eins mit sich selbst gleich Null, $1 + \ldots + 1 = 0$. Wir schreiben auch $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ wenn $\mathbb{Z}/p\mathbb{Z}$ als Körper aufgefasst wird. In der Algebra zeigt man, dass endliche Körper nur p^s Elemente haben können, wobei p prim ist und $s \in \mathbb{N} \setminus \{0\}$ liegen muss. Für jede natürliche Zahl der Form p^s, eine sogenannte Primzahlpotenz, gibt es bis auf Isomorphie genau einen Körper mit p^s Elementen.

Satz 2.2.10.
Sei K ein Körper. Dann gilt für alle $a, b, c \in K$:

1. $a(-b) = -a \cdot b$ und $(-a)(-b) = a \cdot b$ und $(-a)b = -ab$.

2. Aus $b \cdot a = c \cdot a$ und $a \neq 0$ folgt $b = c$. Man kann also in Körpern durch von Null verschiedene Elemente kürzen.

Beweis.

1. $ab + (-a)b = \Big(a + (-a)\Big) \cdot b = 0 \cdot b = 0$, also $(-a)b = -a \cdot b$ wegen der Eindeutigkeit des inversen Elements $-ab$ der Addition. Daraus folgt

$$(-a)(-b) = -a(-b) = -\Big(-(ab)\Big) = ab \; .$$

2. Da $a \neq 0$ gilt, gibt es ein multiplikatives Inverses a^{-1}. Wir rechnen wie im Beweis von Satz 2.1.4

$$b = b(aa^{-1}) = (ba)a^{-1} = (ca)a^{-1} = c(aa^{-1}) = c \; .$$

□

Beispiele 2.2.11.
Wir führen den Körper der komplexen Zahlen ein. Er hat zahllose theoretische und praktische Anwendungen. (Jeder Elektroingenieur kennt ihn!)

- Wir wissen schon, dass $(\mathbb{R}^2, +)$ eine abelsche Gruppe ist.
- Für die Multiplikation kann man *nicht* die komponentenweise Multiplikation benutzen, um einen Körper zu erhalten. Denn es gilt für die komponentenweise Multiplikation

$$\begin{pmatrix}1\\0\end{pmatrix} \times \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \; ,$$

  d.h. es gäbe Nullteiler. Wir definieren vielmehr

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} \cdot \begin{pmatrix}y_1\\y_2\end{pmatrix} := \begin{pmatrix}x_1y_1 - x_2y_2\\x_1y_2 + x_2y_1\end{pmatrix} \; ,$$

  Man rechnet leicht nach, dass $(\mathbb{R}^2 \setminus \{0\}, \cdot)$ mit dieser Verknüpfung eine abelsche Gruppe ist. Das Assoziativitätsgesetz überlassen wir als Übung. Das neutrale Element ist $\begin{pmatrix}1\\0\end{pmatrix}$:

$$\begin{pmatrix}1\\0\end{pmatrix} \cdot \begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1 \cdot x_1 - 0\\0 + 1 \cdot x_2\end{pmatrix} = \begin{pmatrix}x_1\\x_2\end{pmatrix} \quad \forall x \in \mathbb{R}^2$$

  und das zu $x = \begin{pmatrix}x_1\\x_2\end{pmatrix} \neq 0$ inverse Element ist

$$\frac{1}{x} = x^{-1} = \begin{pmatrix}\frac{x_1}{x_1^2+x_2^2}\\\frac{-x_2}{x_1^2+x_2^2}\end{pmatrix} \; .$$

  In der Tat gilt:

$$x^{-1} \cdot x = \begin{pmatrix}x_1\\x_2\end{pmatrix} \cdot \begin{pmatrix}\frac{x_1}{x_1^2+x_2^2}\\\frac{-x_2}{x_1^2+x_2^2}\end{pmatrix} = \begin{pmatrix}\frac{x_1^2}{x_1^2+x_2^2} + \frac{x_2^2}{x_1^2+x_2^2}\\\frac{-x_1x_2}{x_1^2+x_2^2} + \frac{x_2x_1}{x_1^2+x_2^2}\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} \; .$$

Die Überprüfung der Distributivgesetze überlassen wir als Übung. Wir haben also einen Körper $\mathbb{C}$, dessen Elemente komplexe Zahlen heißen. Es ist

$$0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad 1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .$$

Fassen wir die Ebene $\mathbb{R}^2$ dermaßen als Körper der komplexen Zahlen auf, so sprechen wir auch von der komplexen Zahlenebene. Wir setzen $\mathrm{i} := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{C}$. Diese komplexe Zahl heißt imaginäre Einheit. Es gilt

$$\mathrm{i}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -1 .$$

Jede komplexe Zahl $x \in \mathbb{C}$ lässt sich in der folgenden Form schreiben:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1 \cdot 1 + x_2 \mathrm{i} .$$

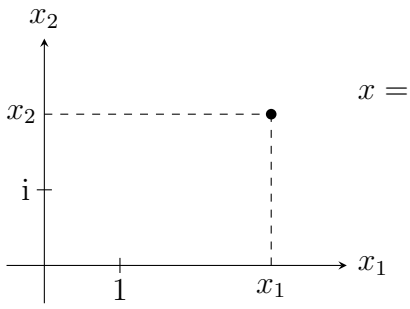
Diese Zerlegung ist sogar eindeutig. Die injektive Abbildung

$$\begin{aligned} \mathbb{R} &\to \mathbb{C} \\ \lambda &\mapsto \lambda \cdot 1 = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \end{aligned}$$

ist ein Körperhomomorphismus und erlaubt es, die reellen Zahlen mit einem Unterkörper der komplexen Zahlen zu identifizieren. Schreiben wir

$$x = x_1 + \mathrm{i}x_2 ,$$

so nennen wir $x_1 = \mathrm{Re}\,(x)$ den Realteil und $x_2 =: \mathrm{Im}\,(x)$ den Imaginärteil von x:



$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \mathrm{Re}(x) \\ \mathrm{Im}(x) \end{pmatrix}$$

Die reellen Zahlen liegen dann auf der horizontalen Achse, der reellen Achse. Mit der Schreibweise $x = x_1 + \mathrm{i}x_2$ ist die Regel für die Multiplikation komplexer Zahlen leicht zu merken:

$$(x_1 + x_2\mathrm{i}) \cdot (y_1 + \mathrm{i}y_2) = x_1y_1 + x_2y_2\mathrm{i}^2 + (x_1y_2 + x_2y_1)\mathrm{i} = (x_1y_1 - x_2y_2) + (x_1y_2 + x_2y_1)\mathrm{i}$$

Definition 2.2.12

1. *Die Abbildung*

$$\begin{aligned} \mathbb{C} &\to \mathbb{C} \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\mapsto \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \end{aligned}$$

heißt komplexe Konjugation. Wir schreiben einfacher

$$\overline{x} = \overline{x_1 + \mathrm{i}x_2} = x_1 - \mathrm{i}x_2 .$$

2. *Für eine komplexe Zahl* $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}$ *heißt*

$$|z| := \left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\| = \sqrt{z_1^2 + z_2^2}$$

der Absolutbetrag von z.

Bemerkungen 2.2.13.

1. Geometrisch ist komplexe Konjugation die Spiegelung an der reellen Achse, der x-Achse der komplexen Zahlenebene: $\mathbb{R} \cong \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \middle| \, x \in \mathbb{R} \right\}$

   Für alle $z, w \in \mathbb{C}$ gilt

2. $\overline{1} = 1, \overline{0} = 0, \overline{\mathrm{i}} = -\mathrm{i}$
3. $\overline{\overline{z}} = z$
4. $\overline{z + w} = \bar{z} + \bar{w}$
5. $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$
6. $|z + w| \leq |z| + |w|$
7. $|z \cdot w| = |z||w|$
8. $|z|^2 = z \cdot \bar{z}$.

Beweis.

1.-3. sind klar.

4. Wir rechnen:

$$\overline{z + w} = \overline{\begin{pmatrix} z_1 + w_1 \\ z_2 + w_2 \end{pmatrix}} = \begin{pmatrix} z_1 + w_1 \\ -(z_2 + w_2) \end{pmatrix} = \begin{pmatrix} z_1 \\ -z_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ -w_2 \end{pmatrix} = \bar{z} + \bar{w}$$

5. Übung
6. Folgt aus Satz 1.1.8.2 für die Norm.
7. Übung
8. Wir rechnen:

$$z \cdot \bar{z} = (z_1 + z_2\mathrm{i})(z_1 - z_2\mathrm{i}) = z_1^2 - (z_2\mathrm{i})^2 = z_1^2 + z_2^2 = |z|^2 \, .$$

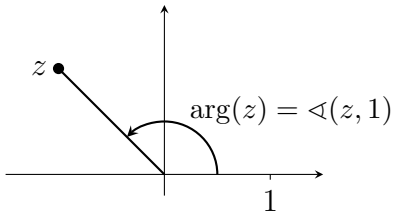
□

Definition 2.2.14
Sei $z \in \mathbb{C} \setminus \{0\}$; dann heißt die Zahl aus $[0, 2\pi)$, die durch

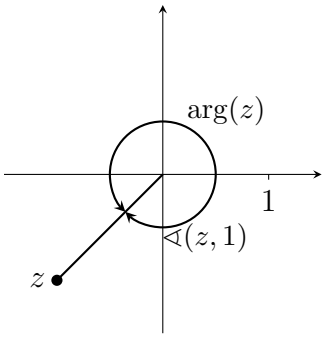
$$\arg(z) := \begin{cases} \sphericalangle(z,1) & \textit{falls } \operatorname{Im}(z) \geq 0 \\ 2\pi - \sphericalangle(z,1) & \textit{falls } \operatorname{Im}(z) < 0 \end{cases}$$

definiert wird, das Argument von z.

Zeichnung:



$\operatorname{Im}(z) \geq 0$



$\operatorname{Im}(z) < 0$

Satz 2.2.15.
Sei $z \in \mathbb{C}, z \neq 0$. Dann gilt

1. $$\operatorname{Re}(z) = |z| \cos(\arg(z))$$
$$\operatorname{Im}(z) = |z| \sin(\arg(z))$$

2. Das Argument von z ist die eindeutig bestimmte reelle Zahl $\varphi \in [0, 2\pi)$, für die gilt
$$z = |z|\Big(\cos\varphi + \mathrm{i}\sin\varphi\Big)$$

Beweis.

1. Wir betrachten die beiden Fälle $\operatorname{Im} z \geq 0$ und $\operatorname{Im} z < 0$ getrennt. Sei zunächst $\operatorname{Im} z \geq 0$:

$$|z| \cos\arg(z) = |z| \cos\Big(\sphericalangle(z,1)\Big)$$
$$= |z| \cos\arccos\frac{\langle z, 1\rangle}{|z||1|} = |z| \frac{\left\langle \begin{pmatrix} \operatorname{Re} z \\ \operatorname{Im} z \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle}{|z| \cdot 1} = \operatorname{Re}(z)\,.$$

$$|z| \sin\arg(z) = |z| \sin\Big(\sphericalangle(z,1)\Big) = |z| \sin\arccos\frac{\langle z, 1\rangle}{|z||1|}\,.$$

Es ist $\sin t = \pm\sqrt{1-\cos^2 t}$ für $t \in \mathbb{R}$. Wegen $\mathrm{Im}\,(z) \geq 0$ ist $\arg(z) \in [0,\pi]$, also $\sin\arg(z) \geq 0$. Daher gilt:

$$\begin{aligned}|z|\sin\arg(z) &= |z|\sqrt{1-\cos^2\arccos\frac{\langle z,1\rangle}{|z|\cdot 1}}\\ &= |z|\sqrt{1-\frac{\langle z,1\rangle^2}{|z|^2}} = \sqrt{|z|^2-\langle z,1\rangle^2}\\ &= \sqrt{(\mathrm{Re}\,z)^2+(\mathrm{Im}\,z)^2-(\mathrm{Re}\,z)^2} = |\mathrm{Im}\,(z)| = \mathrm{Im}\,(z)\ ,\end{aligned}$$

denn nach Vorraussetzung ist $\mathrm{Im}\,(z) \geq 0$.
Den Beweis für $\mathrm{Im}\,(z) < 0$ überlassen wir als Übung.

2. Nach 1. erfüllt $\varphi = \arg(z)$ diese Bedingungen. Zu zeigen ist noch die Eindeutigkeit. Gelte

$$z = |z|(\cos\varphi + \mathrm{i}\sin\varphi) = |z|(\cos\varphi' + \mathrm{i}\sin\varphi'),$$

so folgt

$$\cos\varphi = \cos\varphi' \qquad \sin\varphi = \sin\varphi'\ .$$

Wir rechnen

$$\sin(\varphi-\varphi') = \sin\varphi\cos\varphi' - \cos\varphi\sin\varphi' = 0\ ,$$

wegen der Lage der Nullstellen der Sinusfunktion muss also gelten

$$\varphi = \varphi' + k\pi \quad \text{mit} \quad k \in \{0,\pm 1\}\ .$$

Also

$$\cos(\varphi-\varphi') = \cos k\pi = (-1)^k\ .$$

Andererseits folgt auch

$$\cos(\varphi-\varphi') = \cos\varphi\cos\varphi' + \sin\varphi\sin\varphi' = \cos^2\varphi + \sin^2\varphi = 1\ ,$$

also $k = 0$ und es ist $\varphi = \varphi'$.

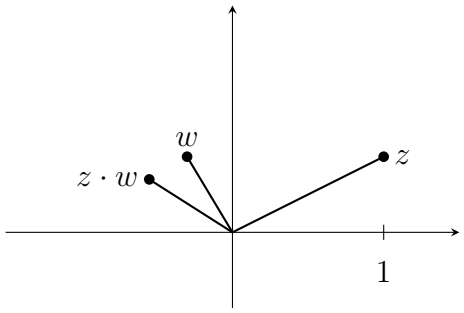
□

Satz 2.2.16.
Seien $z, w \in \mathbb{C}\setminus\{0\}$. Dann gilt

$$\arg(z\cdot w) = \begin{cases}\arg(z)+\arg(w) & \text{falls } \arg(z)+\arg(w) < 2\pi\\ -2\pi+\arg(z)+\arg(w) & \text{sonst}\end{cases}$$

In der komplexen Zahlenebene hat die Multiplikation in $\mathbb{C}$ also die folgende geometrische Bedeutung:

- Die Beträge werden multipliziert: $|zw| = |z||w|$.
- Die Argumente von z und w werden modulo 2π addiert.



Beweis.
Wir rechnen mit $\varphi := \arg(z)$ und $\psi := \arg(w)$

$$\begin{aligned} z \cdot w &= |z||w|\Big(\cos\varphi + \mathrm{i}\sin\varphi\Big)\Big(\cos\psi + \mathrm{i}\sin\psi\Big) \\ &= |z||w|\Big(\cos\varphi\cos\psi - \sin\varphi\sin\psi + \mathrm{i}(\cos\varphi\sin\psi + \sin\varphi\cos\psi)\Big) \\ &= |z||w|\Big(\cos(\varphi+\psi) + \mathrm{i}\sin(\varphi+\psi)\Big) \end{aligned}$$

Wegen $\varphi, \psi \in [0, 2\pi)$ liegt $\varphi + \psi \in [0, 4\pi)$. Es folgt die Behauptung, da für $\varphi + \psi \in [2\pi, 4\pi)$ $\varphi + \psi - 2\pi \in [0, 2\pi)$ und

$$\cos(\varphi + \psi - 2\pi) + \mathrm{i}\sin(\varphi + \psi - 2\pi) = \cos(\varphi + \psi) + \mathrm{i}\sin(\varphi + \psi)$$

gilt. □

Wir brauchen noch einen speziellen Ring, den Polynomring. Sei K ein Körper oder, allgemeiner, ein kommutativer Ring.

Betrachtung 2.2.17.

1. Wir betrachten die Menge aller Folgen von Elementen eines Körpers K, in der nur endlich viele Elemente ungleich 0 sind, oder, allgemeiner, eines kommutativen Rings. Die Folge $(a_0, a_1, \ldots, a_n, 0, \ldots)$ schreiben wir auch suggestiv $a_0 + a_1 t + a_2 t^2 + \ldots a_n t^n$, ohne uns zu fragen, was t ist. Wir nennen die Elemente Polynome.

   Der Grad eines Polynoms $f = a_0 + a_1 t + a_2 t^2 + \ldots a_n t^n$ ist -1, falls $f = 0$, sonst gleich dem $\max_i\{i \mid a_i \neq 0\}$. Zum Beispiel gilt $\mathrm{grad}(t^3 + 2t + 1) = 3$.

   Ein Polynom, dessen Koeffizienten bis auf einen verschwinden, heißt Monom. Zum Beispiel ist $2t^3$ ein Monom, nicht aber $2t^3 + 3t^2$.

2. Seien

   $$f := a_0 + a_1 t + a_2 t^2 + \ldots a_n t^n \quad \text{und} \quad g := b_0 + b_1 t + b_2 t^2 + \ldots b_m t^m$$

   Polynome. Wir addieren Polynome, indem wir die Koeffizienten addieren,

   $$f + g = (a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \ldots (a_n + b_n)t^n + \ldots \ .$$

   Zum Beispiel ist $(x^3 + 2x^2) + (x^2 + 4x) = x^3 + 3x^2 + 4x$. Wir multiplizieren Polynome, indem wir formal das Distributivgesetz anwenden und die Exponenten von t addieren,

   $$f \cdot g = a_0 b_0 + (a_1 b_0 + a_0 b_1)t + (a_2 b_0 + a_1 b_1 + a_0 b_2)t^2 + \ldots$$

Der Koeffizient von t^n in $f \cdot g$ ist also $\sum_{k=0}^{n} a_k b_{n-k}$. Zum Beispiel ist $(t-1)(t+1) = t^2 - 1$

Damit bildet die Menge $K[t]$ der Polynome einen kommutativen Ring, den Polynomring über K. Wir werden ihn konzeptioneller in Kapitel 5 einführen. Er ist integer, wenn K integer ist, also insbesondere für Körper.

3. In einem Polynom $f \in K[t]$ können wir t durch ein beliebiges $\lambda \in K$ ersetzen und erhalten einen Wert $f(\lambda) \in K$. Zum Beispiel erhalten wir für $K = \mathbb{R}$ und $f(t) = t^4 - 3$ für $\lambda = \sqrt{2}$ die Zuordnung $\sqrt{2} \mapsto \sqrt{2}^4 - 3 = 4 - 3 = 1$. So liefert jedes Polynom $f \in K[t]$ eine Funktion $K \to K$, eine polynomiale Funktion.

4. Wir können nun auch formal Brüche $\frac{f}{g}$ betrachten, wobei $f, g \in K[t]$ und $g \neq 0$ gilt. Diese bilden einen Körper $K(t)$; durch Einsetzen erhält man gebrochen-rationale Funktionen, die Polstellen haben können. Zum Beispiel hat $\frac{1}{t^2-5t+6} \in \mathbb{Q}(t)$ die Polstellen 2 und 3.

2.3 Vektorräume

Wir kommen nun zum Begriff des Vektorraums, der für die lineare Algebra zentral ist. Genauer gesagt werden wir für jeden gegebenen Körper K den Begriff eines K-Vektorraums einführen. Der Körper K wird bei unseren Betrachtungen (fast) immer festgehalten werden.

Definition 2.3.1

1. *Sei K ein Körper. Ein K–Vektorraum oder auch Vektorraum über K ist ein Tripel $(V, +, \cdot)$ bestehend aus einer Menge V und Abbildungen*

$$+ : \quad V \times V \to V \qquad \cdot : \quad K \times V \to V \ ,$$

die den folgenden Axiomen genügen:

(V1) $(V, +)$ *ist eine abelsche Gruppe*

Für alle $v, w \in V$ und $\alpha, \beta \in K$ gilt:

(V2a) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$
(V2b) $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$
(V2c) $(\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$
(V2d) $1 \cdot v = v$

2. *Die Elemente eines Vektorraums heißen Vektoren. Das neutrale Element 0 von $(V, +)$ heißt Nullvektor. Im Zusammenhang mit K-Vektorräumen nennt man Elemente von K auch Skalare. Die Verknüpfung $\cdot$ heißt Skalarmultiplikation. (Sie darf nicht mit dem Skalarprodukt aus dem einleitenden Kapitel verwechselt werden!)*

Das neutrale Element der Addition im Körper K und die Inversen in K treten in der Definition nicht auf, spielen aber in der Theorie der Vektorräume eine große Rolle. Über einem beliebigen Ring gibt es den Begriff eines Moduls.

Beispiele 2.3.2.

1. Sei K ein beliebiger Körper. Definiere auf dem kartesischen Produkt

$$V := K^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in K \right\} \ ,$$

also den n-tupeln von Körperelementen, die Addition komponentenweise durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

und die skalare Multiplikation $\cdot : K \times V \to V$ durch komponentenweise Multiplikation

$$\alpha \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

Dann ist $(K^n, +, \cdot)$ ein K–Vektorraum mit $0 := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ und $-x = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}$.

2. Setzt man speziell $n = 1$, so folgt, dass jeder Körper K auch ein K–Vektorraum ist.

3. $(\mathbb{C}, +, \cdot_{|\mathbb{R}\times\mathbb{C}}) = (\mathbb{R}^2, +, \cdot)$ ist ein $\mathbb{R}$–Vektorraum.

4. $(\mathbb{R}, +, \cdot_{|\mathbb{Q}\times\mathbb{R}})$ ist ein $\mathbb{Q}$–Vektorraum.

5. Sei K ein beliebiger Körper und $M \neq \emptyset$ eine Menge. Setze $V := \mathrm{Abb}(M, K) = \{f : M \to K\}$ und definiere Verknüpfungen punktweise: für $f, g \in V$ und $\alpha \in K$

$$(f + g)(m) := f(m) + g(m)$$
$$(\alpha \cdot f)(m) := \alpha \cdot f(m)$$

Das neutrale Element ist die konstante Abbildung mit Wert 0,

$$0_V(m) := 0 \quad \text{für alle } m \in M \ ,$$

das additive Inverse ist

$$(-f)(m) := -f(m) \ .$$

Dann ist $(V, +, \cdot) = (\mathrm{Abb}(M, K), +, \cdot)$ ein K–Vektorraum:

- In Anbetracht von Beispiel 2.2.3.3 ist klar, dass $(V, +)$ eine abelsche Gruppe ist.
- Wir zeigen exemplarisch (V2a): für alle $f \in V$ und $\alpha, \beta \in K$ gilt

$$\begin{aligned} ((\alpha + \beta) \cdot f)(m) &\underset{\text{def}}{=} (\alpha + \beta) \cdot f(m) \\ \underset{\text{(K3b)}}{=} \alpha \cdot f(m) + \beta \cdot f(m) &\underset{\text{def}}{=} (\alpha \cdot f)(m) + (\beta \cdot f)(m) \\ &\underset{\text{def}}{=} (\alpha \cdot f + \beta \cdot f)(m) \ . \end{aligned}$$

Der Vektorraum $\mathrm{Abb}(\{1, 2, \ldots, n\}, K) = \mathrm{Abb}(\underline{n}, K)$ kann in natürlicher Weise mit K^n, also dem ersten Beispiel identifiziert werden.

6. Haben die Menge M oder der Körper K mehr Struktur, so gibt es oft interessante Teilmengen von $\mathrm{Abb}(M, K)$, die auch Vektorräume sind. Beispiele sind

- $M = K$, V=polynomiale Funktionen auf K.

- M topologischer Raum, $K = \mathbb{R}$ oder $K = \mathbb{C}$: stetige Funktionen.
- $M = \mathbb{R}^n$, $K = \mathbb{R}$ oder $K = \mathbb{C}$: differenzierbare Funktionen.

7. Ist $V = (\{0_V\}, +, \cdot)$, so heißt V der Nullvektorraum. Da jeder Vektorraum zumindest den Nullvektor enthält, hat jeder Vektorraum mindestens ein Element.

Satz 2.3.3.
Sei K ein Körper und V ein K–Vektorraum. Dann gilt für alle $v, w \in V$ und $\alpha \in K$

1. $0_K \cdot v = 0_V$
2. $\alpha \cdot 0_V = 0_V$
3. Aus $\alpha \cdot v = 0_V$ folgt $\alpha = 0_K$ oder $v = 0_V$.
4. $(-1) \cdot v = -v$

Beweis.

1. $0_K \cdot v = (0_K + 0_K) \cdot v = 0_K \cdot v + 0_K \cdot v$. Hieraus folgt $0_K \cdot v = 0_V$.
2. $\alpha \cdot 0_V = \alpha \cdot (0_V + 0_V) = \alpha \cdot 0_V + \alpha \cdot 0_V$. Hieraus folgt $\alpha \cdot 0_V = 0_V$.
3. Sei $\alpha \cdot v = 0$ und $\alpha \neq 0$. Wir rechnen:

$$v \underset{\text{(V2d)}}{=} 1 \cdot v = (\alpha^{-1} \cdot \alpha)v \underset{\text{(V2c)}}{=} \alpha^{-1}(\alpha \cdot v) = \alpha^{-1} \cdot 0_V \underset{1.}{=} 0_V \ .$$

   Man beachte, dass hier die Existenz multiplikativer Inverser im Körper K eingeht.

4. Wir berechnen

$$(-1)v + v \underset{\text{(V2d)}}{=} (-1)v + 1 \cdot v \underset{\text{(V2a)}}{=} (-1 + 1)v = 0_K \cdot v \underset{1.}{=} 0_V \ .$$

□

Ab sofort unterscheiden wir in der Notation nicht mehr $0_K \in K$ und $0_V \in V$.

Definition 2.3.4
Sei $(V, +, \cdot)$ ein K-Vektorraum. Eine Teilmenge $W \subset V$ heißt Untervektorraum, falls sie den folgenden Axiomen genügt:

(UV1) $W \neq \emptyset$

(UV2) *Für alle $v, w \in W$ gilt $v + w \in W$. Wir sagen auch, W sei unter der Addition von V abgeschlossen.*

(UV3) *Für alle $\alpha \in K$ und $v \in W$ gilt $\alpha \cdot v \in W$. Wir sagen auch W sei unter der skalaren Multiplikation abgeschlossen.*

Satz 2.3.5.
Sei $(V, +, \cdot)$ ein K–Vektorraum. Sei $W \subset V$ ein Untervektorraum. Sei $+_W$ die Einschränkung von

$$\begin{aligned} + : \quad & V \times V \to V \\ \text{auf} \quad +_W : \quad & W \times W \to W \end{aligned}$$

Sei $\cdot_W$ die Einschränkungen von

$$\begin{aligned} \cdot : \quad & K \times V \to V \\ \text{auf} \quad \cdot_W : \quad & K \times W \to W \end{aligned}$$

Dann ist $(W, +_W, \cdot_W)$ ein K–Vektorraum. Die 0 in W stimmt mit der 0 in V überein.

Beweis.

1. Wir beweisen zunächst, dass $(W, +_W)$ eine abelsche Untergruppe von $(V, +)$ ist: Offenbar gilt impliziert (UV1) das Untergruppenaxiom (UG1), und (UV2) impliziert (UG2). Mit $v \in W$ ist mit $\alpha = -1$ nach (UV3) auch

   $$(-1)v = -v \in W \ ,$$

   so dass (UG3) erfüllt ist. Insbesondere ist $(W, +_W)$ eine abelsche Gruppe und $0_W = 0_V$.

2. Die Axiome (V2a-d) gelten sogar für alle Elemente in V, also erst recht für alle Elemente in der Teilmenge W.

□

Beispiele 2.3.6.

1. $K = \mathbb{R}$ und $(\mathbb{C}, +, \cdot)$. Dann ist

   $$W_r = \{x_1 + 0 \cdot \mathrm{i} \mid \ x_1 \in \mathbb{R}\}$$

   ein Untervektorraum, ebenso

   $$W_i = \{0 + x_2 \cdot \mathrm{i} \mid \ x_2 \in \mathbb{R}\}$$

2. Aber für $K = \mathbb{C}$ sind dies keine Untervektorräume: etwa für $\alpha = \mathrm{i} \in K$ und $v = 1 + 0 \cdot \mathrm{i} \in W_r$ ist

   $$\alpha \cdot v = (0 + \mathrm{i}) \cdot (1 + 0 \cdot \mathrm{i}) = 1 \cdot \mathrm{i} \notin W_r \ .$$

3. Sei K ein beliebiger Körper und $V = K^n$ wie im Beispiel 2.3.2.1. Dann ist

   $$W = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| x_i \in K \right\} \qquad \text{ein Untervektorraum}$$

   Die Axiome (UV1)-(UV3) sind offensichtlich.

4. Sei K ein beliebiger Körper, $M \neq \emptyset$ eine Menge und $V = \mathrm{Abb}(M, K)$. Sei $M' \subset M$ eine Teilmenge, dann ist

$$W := \{f : M \to K \mid f(m') = 0 \text{ für alle } m' \in M'\}$$

ein Untervektorraum, der Annihilator von M':

(UV1) folgt, da die Abbildung 0_V mit $0_V(m) = 0$ für alle $m \in M$ in W liegt.

(UV2) Seien $f, g \in W$, rechne für $m' \in M'$ beliebig

$$(f + g)(m') = f(m') + g(m') = 0 + 0 = 0 \text{ , also } f + g \in W$$

(UV3) Sei $f \in W$, $\alpha \in K$, rechne für $m' \in M'$

$$(\alpha f)(m) = \alpha \cdot f(m) = \alpha \cdot 0 = 0 \text{ , also } \alpha f \in W$$

5. Sei speziell $M = K$ und $V = \mathrm{Abb}(K, K)$. Dann ist die Menge der polynomialen Funktionen

$$\{f : K \to K |\ f(x) = p(x) \text{ mit } p \in K[X] \text{ Polynom}\}$$

ein Untervektorraum von V. Ebenso ist für $k \in \mathbb{N}$

$$K_k[x] := \{f :\ K \to K |\ f \text{ ist Polynom vom Grad } \leq k\}$$

ein Untervektorraum, aber für $k \geq 1$ ist

$$\widetilde{K}_k[x] := \{f :\ K \to K |\ f \text{ ist Polynom vom Grad } = k\}$$

kein Untervektorraum. Denn zum Beispiel sind $f(X) = X^3 + X^2 + 1 \in \mathbb{C}[X]$ und $g(X) = -X^3 + X^2 + 1 \in \mathbb{C}[X]$ beides Polynome vom Grad genau gleich 3, aber ihre Summe $f + g = 2X^2 + 2$ hat Grad 2. Also ist (UV3) nicht erfüllt.

6. In jedem Vektorraum V sind V selbst und $W := \{0\}$ Untervektorräume. $W = \{0\}$ heißt der triviale Untervektorraum.

Bemerkungen 2.3.7.

1. Jeder Untervektorraum eines Untervektorraums von V ist selbst ein Untervektorraum von V.

2. Sei V ein Vektorraum und seien W_1, W_2 Untervektorräume von V. Dann ist $W_1 \cap W_2$ ein Untervektorraum.

   Beweis.
   (UV1) Aus $0_V \in W_1$ und $0_V \in W_2$ folgt $0_V \in W_1 \cap W_2$.
   (UV2) folgt, da mit $v, w \in W_i$ auch $v + w \in W_i$ liegt, also aus $v, w \in W_1 \cap W_2$ folgt $v + w \in W_1 \cap W_2$. (UV3) folgt analog. □

3. $W_1 \cup W_2$ ist im Allgemeinen kein Untervektorraum.
   Als Gegenbeispiel betrachte $\mathbb{C}$ als reellen Vektorraum, $K = \mathbb{R}$ und $V = \mathbb{C}$, und

$$W_r \cup W_i = \{x_1 + x_2\,\mathrm{i} \mid\ x_1 = 0 \text{ oder } x_2 = 0\}\,.$$

   Dies ist das Achsenkreuz bestehend aus der reellen und imaginären Achse. Dann ist $x = 1 + 0\,\mathrm{i} \in W_r$ auf der reellen Achse und $y = 0 + 1 \cdot \mathrm{i} \in W_i$ auf der imaginären Achse. Die Summe $x + y = 1 + 1 \cdot \mathrm{i} \notin W_r \cup W_i$ ist aber nicht auf dem Achsenkreuz.

Definition 2.3.8

1. *Sei V ein K–Vektorraum und seien endlich viele Elemente $v_1, \ldots, v_m \in V$ gegeben, die nicht unbedingt verschieden sind. Ein Element $w \in V$ der Form*

$$w = \lambda_1 v_1 + \ldots + \lambda_m v_m$$

mit $\lambda_1, \ldots, \lambda_m \in K$ heißt Linearkombination der Vektoren $v_1, \ldots, v_m$.

2. *Die Menge aller Linearkombinationen der Vektoren $v_1, \ldots, v_m$*

$$\mathrm{span}_K(v_1, \ldots, v_m) := \{\lambda_1 v_1 + \ldots + \lambda_m v_m | \, \lambda_i \in K\}$$

heißt der von den Vektoren $v_1, \ldots, v_m$ aufgespannte Raum oder das Erzeugnis dieser Vektoren.

3. *Sei V ein K–Vektorraum und $M \subset V$, $M \neq \emptyset$ eine nichtleere Teilmenge von V. Dann heißt*

$$\mathrm{span}_K(M) := \{w = \lambda_1 v_1 + \ldots + \lambda_m v_m | \, \lambda_i \in K, \, v_i \in M, \, m \in \mathbb{N}\}$$

der von der Teilmenge M aufgespannte Raum oder ihr Erzeugnis.

Man beachte, dass alle Summen *endliche* Summen sind - unendliche Summen sind in unserem Kontext gar nicht definiert.

Satz 2.3.9.
Sei V ein K–Vektorraum, $v_1, \ldots, v_m \in V$ und $M \subset V$ eine nicht-leere Teilmenge, $M \neq \emptyset$. Dann gilt:

1. $\mathrm{span}_K(v_1, \ldots, v_m)$ ist ein Untervektorraum von V.
2. $\mathrm{span}_K(M)$ ist ein Untervektorraum von V und es gilt $M \subset \mathrm{span}_K(M)$.
3. Ist $W \subset V$ ein Untervektorraum und $M \subset W$, so ist auch $\mathrm{span}_K(M) \subset W$.

Beweis.

1. (UV1): Folgt, da

$$0 = 0v_1 + \ldots + 0v_m \in \mathrm{span}_K(v_1, \ldots, v_m)$$

(UV2) Seien $x, y \in \mathrm{span}_K(v_1, \ldots, v_m)$. Dann gibt es $\lambda_1, \ldots, \lambda_m \in K$ und $\mu_1, \ldots, \mu_m \in K$, so dass gilt

$$x = \lambda_1 v_1 + \ldots + \lambda_m v_m \qquad y = \mu_1 v_1 + \ldots + \mu_m v_m \, .$$

Somit ist

$$x + y = (\lambda_1 + \mu_1)v_1 + \ldots + (\lambda_m + \mu_m)v_m \in \mathrm{span}_K(v_1, \ldots, v_m)$$

(UV3) folgt analog aus

$$\alpha x = (\alpha\lambda_1)v_1 + \ldots + \alpha\lambda_m v_m \in \mathrm{span}_K(v_1, \ldots, v_m) \quad \text{für } \alpha \in K \, .$$

2. Der Beweis für eine beliebige Teilmenge $M \subset V$ geht analog.

3. Ist $M \subset W$ und W ein Untervektorraum, so liegen nach (UV2) und (UV3) auch alle Linearkombinationen von Elementen in M in W.

□

Es folgt nun, dass

$$\mathrm{span}_K(M) = \bigcap_{M \subset W, W \subset V \text{ Untervektorraum}} W$$

Denn da $\mathrm{span}_K(M)$ nach Satz 2.3.9.2 selbst ein Untervektorraum von V ist, der die Teilmenge M enthält, ist er einer der Unterräume, über die der Schnitt genommen wird, enthalten, also gilt $\cap W \subset \mathrm{span}_K(M)$. Nach Satz 2.3.9.3 ist aber $\mathrm{span}_K(M)$ in jedem der Untervektorräume, über die der Schnitt genommen wird, enthalten, also gilt auch die umgekehrte Inklusion $\mathrm{span}_K(M) \subset \cap W$. Anders gesagt: $\mathrm{span}_K(M)$ ist also bezüglich der Inklusion der kleinste Untervektorraum von V, der M enthält.

Beispiele 2.3.10.

1. Speziell für $M = \emptyset$ erhalten wir

$$\mathrm{span}_K(\emptyset) = \bigcap_{\emptyset \subset W, W \subset V \text{ Untervektorraum}} W = 0$$

den Nullvektorraum.

2. Sei K ein beliebiger Körper, den wir als Vektorraum über sich selbst betrachten, und sei $v \in K$. Für $v = 0$ ist $\mathrm{span}_K(0) = \{0\}$ der triviale Untervektorraum von K; für $v \neq 0$ ist $\mathrm{span}_K(v) = K$.

3. Sei K ein beliebiger Körper und $V = K^n$. Setze

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \ldots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

dann ist $\mathrm{span}_K(e_1, \ldots, e_n) = V$, denn für jedes $x \in K^n$ gilt

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \ldots + x_n e_n \in \mathrm{span}_K(e_1, \ldots, e_n)$$

Definition 2.3.11

1. *Seien Λ, M Mengen. Man nennt eine Abbildung*

$$\phi : \Lambda \to M$$

auch eine durch die Indexmenge Λ indizierte Familie von Elementen von M. Man nennt dann Λ die Indexmenge, schreibt $a_\lambda = \phi(\lambda)$ für $\lambda \in \Lambda$ und $(a_\lambda)_{\lambda \in \Lambda}$ für die Familie. Ist $\Lambda = \underline{n}$ für $n \in \mathbb{N}$, so heißt die Familie endlich.

2. *Ist eine Familie durch $\mathbb{N}$ oder $\underline{n} = \{1, 2, \ldots, n\}$ indiziert, so liefert die Ordnung auf der Indexmenge eine Ordnung der Familie. Wir sprechen von einer geordneten Familie, wenn wir diese Ordnung als weitere Struktur betrachten.*

Wir vereinbaren, dass wir bei einer endlichen Familie die Indexmenge nicht explizit schreiben und lassen die Schreibweise $(a_1, a_2, \ldots, a_n)$ zu. Man beachte, dass hier Elemente gleich sein können, etwa ist $a_1 = a_2$ möglich.

Definition 2.3.12
Sei V ein K–Vektorraum.

1. *Eine endliche Familie $(v_1, \ldots, v_r)$ von Vektoren aus V heißt linear unabhängig, falls gilt: sind $\lambda_1, \ldots, \lambda_r \in K$ und gilt*
$$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0 \ ,$$
so folgt daraus $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$. Die Familie $(v_1, \ldots, v_r)$ heißt also genau dann linear unabhängig, wenn der Nullvektor sich nur trivial als Linearkombination von $v_1, \ldots, v_k$ darstellen lässt.

2. *Eine beliebige Familie von Vektoren aus V heißt linear unabhängig, wenn jede* endliche *Teilfamilie linear unabhängig ist.*

3. *Andernfalls heißt die Familie linear abhängig; dann gibt es eine Darstellung des Nullvektors als nicht triviale Linearkombination, d.h. es gibt $\lambda_i \in K$ mit*
$$0 = \lambda_1 v_1 + \ldots + \lambda_r v_r \ ,$$
wobei nicht alle $\lambda_i \in K$ verschwinden.

4. *Eine Teilmenge $M \subset V$ heißt linear unabhängig, falls für jede* endliche *Teilmenge $\{v_1, \ldots, v_m\} \subset M$ die Familie der Vektoren $(v_1, \ldots, v_m)$ linear unabhängig ist.*
Andernfalls heißt sie linear abhängig. Dann enthält M eine endliche Teilmenge $\{v_1, \ldots, v_m\}$ für die die Familie $(v_1, \ldots, v_m)$ linear abhängig ist.

5. *Wir setzen $\mathrm{span}_K(\emptyset) = \{0\}$ und nennen die leere Familie linear unabhängig.*

Lineare Unabhängigkeit ist eine Eigenschaft einer Familie, nicht einzelner Vektoren. Insbesondere ist nicht ein Vektor linear (un)abhängig von einem anderen Vektor.

Lemma 2.3.13.
Für eine Familie $(v_1, \ldots, v_r)$ von Vektoren eines K–Vektorraums sind die folgenden Bedingungen äquivalent:

1. Die Familie (v_i) ist linear unabhängig.

2. Jeder Vektor $v \in \mathrm{span}_K(v_i)$ lässt sich in *eindeutiger* Weise als Linearkombination von Vektoren der Familie (v_i) schreiben.

Beweis.
2. $\Rightarrow$ 1. klar, denn insbesondere lässt sich der Nullvektor eindeutig als Linearkombination

schreiben.
1. $\Rightarrow$ 2. Aus

$$v = \sum_i \lambda_i v_i = \sum_i \mu_i v_i \qquad \text{mit } \lambda_i, \mu_i \in K$$

folgt

$$0 = \sum_i (\lambda_i - \mu_i) v_i$$

Wegen der vorausgesetzten linearen Unabhängigkeit folgt $\lambda_i - \mu_i = 0$, also $\lambda_i = \mu_i$ für alle i. $\square$

Bemerkungen 2.3.14.

1. Die leere Teilmenge $\emptyset \subset V$ eines Vektorraums V ist linear unabhängig, da hier die Prämisse nicht erfüllt ist.
2. In K^n ist jede Teilmenge der Menge $\{e_1, \ldots, e_n\}$ aus Beispiel 2.3.10.2 linear unabhängig.
3. Die Familie bestehnd aus einem einzigen Vektor $v \in V$ ist genau dann linear unabhängig, wenn $v \neq 0$ gilt.
   Denn wegen $1 \cdot 0 = 0$ ist 0 linear abhängig; ist v linear abhängig, so gibt es $\lambda \in K \setminus \{0\} = K^n$ mit $\lambda v = 0$, aus Satz 2.3.3.3 folgt nun $v = 0$.
4. Jede Untermenge einer linear unabhängigen Menge von Vektoren ist linear unabhängig. (Übung.)
5. Enthält eine Familie von Vektoren eines Vektorraums eine linear abhängige Unterfamilie, so ist sie linear abhängig. (Übung.)
6. Insbesondere ist jede Familie, die den Nullvektor oder zwei gleiche Vektoren enthält, linear abhängig. Denn gilt $v_1 = v_2$, so finden wir die nicht-triviale Linearkombination $1 \cdot v_1 + (-1)v_2 = 0$.
7. Enthält die Familie mehr als zwei Vektoren, so ist sie genau dann linear abhängig, wenn ein Vektor der Familie Linearkombination der anderen Vektoren ist.

Beweis.
Ist die Familie $(v_1, \ldots, v_r)$ linear abhängig, so finden wir eine nicht-triviale Linearkombination

$$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$$

für die etwa $\lambda_i \neq 0$ ist. Dann ist

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \ldots - \frac{\lambda_r}{\lambda_i} v_r \ .$$

(Beachten Sie, dass wir hier die Division im Körper K benutzen.) Ist umgekehrt v_i eine Linearkombination der anderen Vektoren der Familie

$$v_i = \mu_1 v_1 + \ldots + \mu_r v_r \quad \text{mit } \mu_i \in K \ ,$$

so ist

$$\mu_1 v_1 + \ldots + \mu_{i-1} v_{i-1} - v_i + \mu_{i+1} v_{i+1} + \ldots + \mu_r v_r = 0$$

eine nicht–triviale Linearkombination, die den Nullvektor darstellt. $\square$

2.4 Basis und Dimension

Definition 2.4.1
Sei K ein Körper und V ein K–Vektorraum.

1. *Eine Teilmenge $M \subset V$ heißt Erzeugendensystem von V, falls $\operatorname{span}_K(M) = V$ gilt.*
2. *Eine Teilmenge $M \subset V$ heißt Basis von V, falls M ein linear unabhängiges Erzeugendensystem ist.*
3. *Eine geordnete Basis von V ist eine endliche oder abzählbare linear unabhängige Familie $(v_\lambda)_{\lambda\in\Lambda}$ von Vektoren in V, die ein Erzeugendensystem von V ist und als Indexmenge $\Lambda = \mathbb{N}$ oder $\Lambda = \underline{n}$ für $n \in \mathbb{N}$ hat.*

Beispiele 2.4.2.

1. Jeder Vektorraum V besitzt ein Erzeugendensystem, zum Beispiel sich selbst, $M = V$. Es ist an dieser Stelle noch nicht klar, dass jeder Vektorraum eine Basis besitzt.
2. $V = K^n$ besitzt die Basis $\{e_i\}_{i=1\ldots n}$ aus Beispiel 2.3.10.2. Wir nennen diese Basis die Standardbasis des K^n. Wir fassen den Vektorraum K^n oft als Vektorraum mit einer ausgezeichneten Basis auf.
3. Allgemeiner hat der Vektorraum $\operatorname{Abb}(M, K)$ aus Beispiel 2.3.6.5 für M eine endliche Menge eine Basis, die aus den Funktionen f_m besteht, die durch $f_m(m') = 0$ für $m \neq m'$ und $f_m(m) = 1$ definiert sind. (Warum bilden diese Funktionen für unendliche Mengen keine Basis?)
4. Endliche geordnete Basen schreiben wir auch in der Form $(v_1, v_2, \ldots, v_n)$. Die Standardbasis des K^n ist durch $(e_1, e_2, \ldots, e_n)$ eine geordnete Basis. Die Ordnung ist also eine *zusätzliche* Struktur, nämlich die Wahl einer Reihenfolge der Basisvektoren. (Eine Basis ist eine Teilmenge, da kommt es nicht auf eine Reihenfolge der Elemente an.) Wenn wir die Standardbasis als geordnete Basis sehen, dann immer mit der Ordnung $(e_1, e_2, \ldots, e_n)$.
5. $K = \mathbb{R}$ und $V = \mathbb{C}$, dann ist $M = \{1, \mathrm{i}\}$ eine $\mathbb{R}$–Basis von $\mathbb{C}$. Auch $\{1, -\mathrm{i}\}$ ist eine $\mathbb{R}$–Basis von $\mathbb{C}$.
6. $K = \mathbb{R}$ und $V = \mathbb{R}_k[X]$ der Raum der Polynome mit reellen Koeffizienten vom Grad höchstens k, dann ist die Folge der Monome $(1, X, \ldots, X^k)$ eine (geordnete) Basis von V.

Satz 2.4.3.
Sei K ein Körper und $V \neq \{0\}$ ein K–Vektorraum und $M \subset V$ eine Teilmenge. Dann sind äquivalent:

1. M ist eine Basis von V.
2. M ist ein minimales Erzeugendensystem, d.h. M ist ein Erzeugendensystem und für jedes $v \in M$ ist $M \setminus \{v\}$ kein Erzeugendensystem.
3. M ist eine maximale linear unabhängige Teilmenge, d.h. M ist linear unabhängig und für jedes $v \in V \setminus M$ ist $M \cup \{v\}$ linear abhängig.

Beweis.

1.⇒ 2. Sei M eine Basis; nach Definition ist M insbesondere ein Erzeugendensystem. Zu zeigen ist noch, dass M ein *minimales* Ereugendensystem ist. Angenommen, schon die kleinere Familie $M \setminus \{v\}$ für ein $v \in M$ wäre ein Erzeugendensystem. Dann ist

$$v = \sum_i \lambda_i v_i$$

mit geeigneten $\lambda_i \in K$, $v_i \in M \setminus \{v\}$. Daraus folgt die Relation

$$\sum_i \lambda_i v_i - v = 0 \ ,$$

die den Nullvektor als nicht-triviale Linearkombination von Elementen von M darstellt, im Widerspruch dazu, dass M als Basis linear unabhängig ist.

2.⇒ 1. Sei M minimales Erzeugendensystem. Angenommen, $M = \{v_i\}_{i \in I}$ ist linear abhängig, so finde wegen Bemerkung 2.3.14.6

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

für ein $v \in M$ und $v_i \in M \setminus \{v\}$. Dann ist aber auch schon $M \setminus \{v\}$ auch ein Erzeugendensystem, im Widerspruch zur Annahme, dass M minimal sei.

1.⇒3. Zu zeigen ist, dass eine Basis M unter den linear unabhängigen Teilmengen maximal ist. Sei $v \in V \setminus M$ beliebig, so ist, da M ein Erzeugendensystem ist

$$v = \sum_{i \in I} \lambda_i v_i \ ,$$

also ist die Menge $M \cup \{v\}$ und damit jede echte Obermenge von M linear abhängig.

3.⇒1. Sei M eine maximale linear unabhängige Teilmenge. Zu zeigen ist, dass M Erzeugendensystem ist. Sei $v \in V$. Wenn $v \in M$, ist klar, dass $v \in \mathrm{span}_K(M)$. Sei also $v \in V \setminus M$. Dann ist $\{v\} \cup M$ linear abhängig. Finde also eine Linearkombination

$$\alpha v + \sum \lambda_i v_i = 0$$

mit geeigneten $\alpha, \lambda_i \in K$, $v_i \in M$. Hierbei dürfen nicht α und alle λ_i gleich Null sein. Wäre $\alpha = 0$, so folgt $\sum \lambda_i v_i = 0$, nicht alle $\lambda_i = 0$, im Widerspruch zur linearen Unabhängigkeit von M. Es ist also $\alpha \neq 0$, woraus mit Hilfe der Division im Körper K folgt

$$v = \sum_{j=1}^{m} \Big(-\frac{\lambda_j}{\alpha}\Big) v_j \ ,$$

d.h. M ist ein Erzeugendensystem.

□

Definition 2.4.4

Ein K–Vektorraum V heißt endlich erzeugt, falls V ein endliches Erzeugendsystem besitzt, d.h. falls es eine endliche Teilmenge $M = \{v_1, \ldots, v_m\}$ von V gibt, so dass $V = \mathrm{span}_K(M)$ gilt.

Lemma 2.4.5.
Sei V ein K–Vektorraum, der nicht endlich erzeugt ist. Dann gibt es zu jeder natürlichen Zahl $n \in \mathbb{N} \setminus \{0\}$ Vektoren $v_1, \ldots, v_n \in V$, so dass die Familie $(v_1, v_2, \ldots, v_n)$ linear unabhängig ist.

Beweis.
Durch vollständige Induktion nach n.

- Induktionsanfang $n = 1$. Wähle $v_1 \in V$, $v_1 \neq 0$. Dies existiert, da sonst $V = \{0\}$ wäre und V somit endlich erzeugt wäre, nämlich von der leeren Menge $\emptyset$, vgl. Beispiel 2.3.10.1.

- Induktionsschritt:
  Sei $n \in \mathbb{N} \setminus \{0\}$ und seien $v_1, \ldots, v_n \in V$ linear unabhängig. Wähle $v_{n+1} \in V \setminus \operatorname{span}_K(v_1, \ldots, v_n)$. Solch ein v_{n+1} existiert, da andernfalls V von den $(v_1, \ldots, v_n)$, also endlich erzeugt wäre. Es bleibt zu zeigen, dass auch die Familie $(v_1, \ldots, v_n, v_{n+1})$ linear unabhängig ist. Sei
  $$0 = \sum_{j=1}^{n+1} \alpha_j v_j \qquad \alpha_j \in K$$
  eine Linearkombination des Nullvektors. Wäre $\alpha_{n+1} \neq 0$, so würden wir die Relation
  $$v_{n+1} = \sum_{j=1}^{n} \left(-\frac{\alpha_j}{\alpha_{n+1}} \right) v_j$$
  erhalten, die im im Widerspruch zu unserer Wahl von $v_{n+1} \notin \operatorname{span}_K(v_1, \ldots, v_n)$ steht. Also muss $\alpha_{n+1} = 0$ gelten; daraus folgt
  $$\sum_{j=1}^{n} \alpha_j v_j = 0$$
  und hieraus nach Induktionsannahme $\alpha_j = 0$ für alle $j = 1, \ldots, n$.

□

Nicht jeder Vektorraum hat ein endliches Erzeugendensystem, zum Beispiel nicht der K–Vektorraum der Polynome $K[X]$ über einem beliebigen Körper K.

Satz 2.4.6 (Basisauswahlsatz)**.**
Sei V ein K–Vektorraum und $M \subset V$ ein endliches Erzeugendensystem. Dann gibt es eine Teilmenge $\mathcal{B} \subset M$, die eine Basis von V ist. Insbesondere besitzt also jeder endlich erzeugte Vektorraum eine Basis.

Beweis.

- Sei $M = \{v_1, \ldots, v_n\}$ ein Erzeugendensystem von V. Ist M auch linear unabhängig, so können wir $\mathcal{B} = M$ wählen und sind fertig.

- Ist M nicht linear unabhängig, so ist nach nach Satz 2.4.3.2 das Erzeugendensystem M nicht minimal. Es gibt also ein $v \in M$ mit $v \in \operatorname{span}_K(M \setminus \{v\})$ wieder ein Erzeugendensystem ist, das aber ein Element weniger enthält.
  So fährt man fort, bis man nach endlich vielen Schritten ein minimales Erzeugendensystem erhält. Dies ist nach Satz 2.4.3 dann eine Basis.

□

Im Fall von Vektorräumen, die nicht endlich-erzeugt sind, benutzt man das Zornsche Lemma, das zu dem Auswahlaxiom der Mengenlehre äquivalent ist. Wir brauchen erst eine Definition:

Definition 2.4.7

1. *Eine partielle Ordnung $\leq$ auf einer Menge X ist eine Relation $\leq$ auf X mit den folgenden Eigenschaften:*

(PO1) *reflexiv: $x \leq x$ für alle $x \in X$,*

(PO2) *transitiv: aus $x \leq y$ und $y \leq z$ folgt $x \leq z$,*

(PO3) *antisymmetrisch: aus $x \leq y$ und $y \leq x$ folgt $x = y$.*

2. *Eine Kette in einer partiell geordneten Menge X ist eine Teilmenge $\mathfrak{K} \subset X$, so dass $(\mathfrak{K}, \leq)$ total geordnet ist, d.h. für alle $h, k \in \mathfrak{K}$ gilt $h \leq k$ oder $k \leq h$.*

3. *Die Menge X heißt induktiv geordnet, wenn jede Kette $\mathfrak{K}$ in X eine obere Schranke besitzt, d.h. es existiert ein $m \in X$ mit $k \leq m$ für alle $k \in \mathfrak{K}$.*

Beispiele 2.4.8.

1. Sei M eine Menge. Dann ist die Potenzmenge $\mathcal{P}(M)$, vgl. Definition 1.4.6.3, partiell geordnet durch Inklusion. M selbst ist eine obere Schranke und ein maximales Element, d.h. es gibt kein $y \in \mathcal{P}(M) \setminus \{M\}$ mit $M \leq y$.

2. Sei $m \in \mathbb{N}$ und T_m die Menge der positiven Teiler von m ungleich 1. Dann definiert $p \leq q$ wenn q p teilt, eine partielle Ordnung, aber keine totale Ordnung. Eine Kette in T_{12} zum Beispiel $\{3, 6, 9, 12\}$, aber nicht ganz T_{12}, wie das Teilerpaar $3, 4$ zeigt. Eine obere Schranke der Kette $\{3, 6, 9, 12\}$ ist der Teiler 3; die Elemente $2, 3$ sind maximal.

Das folgende Lemma ist äquivalent zum Auswahlaxiom, worauf wir aber hier nicht eingehen können.

Lemma 2.4.9 (Zornsches Lemma)**.**
Jede nichtleere induktiv geordnete Menge X besitzt ein maximales Element, d.h. es gibt ein $x \in X$, so dass kein $y \in X \setminus \{x\}$ mit $x \leq y$ existiert.

Wir zeigen nun:

Satz 2.4.10.
Sei V ein K-Vektorraum, $E \subset V$ ein Erzeugendensystem von V und $M \subset E$ eine linear unabhängige Teilmenge von V. Dann gibt es eine Basis B von V mit $M \subset B \subset E$.

Beweis.

- Wir betrachten die Menge $X(M, E) := \{A \,|\, \text{linear unabhängig}, M \subset A \subset E\}$. Sie enthält M selbst und ist daher nicht leer. Sie ist durch Inklusion partiell geordnet.

- Wir zeigen, dass $X(M,E)$ induktiv geordnet ist, indem wir nachweisen, dass für jede Kette $\mathfrak{K} \subset X(M,E)$ die Vereinigung $\cup\mathfrak{K} := \cup_{k\in\mathfrak{K}} k$ in $X(M,E)$ liegt. Es ist dann klar, dass für jedes $A \in \mathfrak{K}$ gilt $A \subset \cup\mathfrak{K}$, also $\cup\mathfrak{K}$ eine obere Schranke ist.

  Aus $M \subset A \subset E$ für alle $A \in \mathfrak{K}$ folgt $M \subset \cup\mathfrak{K} \subset E$. Zu zeigen ist noch, dass die Vereinigung $\cup\mathfrak{K}$ linear unabhängig ist. Seien dazu $\lambda_1, \ldots, \lambda_n \in K$ und $v_1, \ldots v_n \in \cup\mathfrak{K}$ mit $\sum_{j=1}^n \lambda_j v_j = 0$. Dann existieren $A_1 \ldots, A_n \in \mathfrak{K}$ mit $v_j \in A_j$. Durch Umnummerieren können wir erreichen, dass $A_1 \subset A_2 \subset \ldots \subset A_n$ gilt. Daraus folgt $v_1, \ldots, v_n \in A_n$. Aus der linearen Unabhängigkeit von A_n ergibt sich $\lambda_1 = \ldots = \lambda_n = 0$. Also ist $\cup\mathfrak{K}$ linear unabhängig und eine obere Schranke der Kette $\mathfrak{K}$ in $X(M,E)$. Somit ist $X(M,E)$ induktiv geordnet.

- Da $X(M,E)$ nicht leer und induktiv geordnet ist, existiert nach dem Zornschen Lemma ein maximales Element $B \in X(M,E)$. Dieses ist per Definition eine linear unabhängige Teilmenge von V mit $M \subset B \subset E$. Wegen der Maximalität von B muss für jeden Vektor $e \in E \setminus B$ die Menge $B \cup \{e\}$ linear abhängig sein.

  Es existieren also $\lambda, \lambda_1, \ldots, \lambda_n \in K$, nicht alle Null und $b_1, \ldots, b_n \in B$, so dass $\lambda e + \sum_{j=1}^n \lambda_j b_j = 0$ gilt. Aus der linearen Unabhängigkeit von B folgt $\lambda \neq 0$. Damit gilt $e \in \mathrm{span}_K(B)$. Da dies für jedes $e \in E \setminus B$ gilt, folgt $E = B \cup (E \setminus B) \subset \mathrm{span}_K(B)$ und somit $V = \mathrm{span}(E) \subset \mathrm{span}(B)$. Also ist B eine Basis mit allen geforderten Eigenschaften.

□

Wir haben insbesondere auch gezeigt:

Korollar 2.4.11.

1. Basisauswahlsatz, der in Satz 2.4.6 für endlich erzeugte Vektorräume gezeigt wurde: Aus jedem Erzeugendensystem E eines Vektorraums kann man eine Basis auswählen. Hier wählt man einfach die nach Beispiel 2.3.14.1 linear unabhängige Teilmenge $M = \emptyset$.

2. Jeder K-Vektorraum hat eine Basis. Denn wähle einfach $E = V$ als Erzeugendensystem und wieder $M = \emptyset$ als linear unabhängige Teilmenge.

3. Basisergänzungssatz: Jede linear unabhängige Teilmenge $M \subset V$ lässt sich zu einer Basis von V ergänzen. Hier wählt man einfach $V = E$ als Erzeugendensystem.

Satz 2.4.12 (Austauschlemma)**.**
Sei V ein K–Vektorraum, $\mathcal{B} = \{v_1, \ldots, v_r\} \subset V$ eine Basis. Sei $w = \sum_{j=1}^r \alpha_j v_j \in V$ mit $\alpha_i \in K$. Dann gilt für jedes $k \in \{1, \ldots, r\}$ mit $\alpha_k \neq 0$: auch die Menge

$$\mathcal{B}'_k := \{v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_r\}$$

ist eine Basis, d.h. jedes Basiselement v_k mit $\alpha_k \neq 0$ kann gegen w ausgetauscht werden.

Beweis.

- Nach Umnummerierung können wir $k = 1$ annehmen.

- Wir zeigen: $\mathcal{B}'$ ist auch ein Erzeugendensystem. Für jedes gegebene $v \in V$ existieren $\beta_i \in K$, so dass

$$v = \sum_{j=1}^r \beta_j v_j \qquad (*)$$

gilt. Aus $\alpha_1 \neq 0$ folgt

$$v_1 = \frac{1}{\alpha_1} w + \sum_{j=2}^{r} \Big(-\frac{\alpha_j}{\alpha_1} \Big) v_j \, .$$

Dies setzen wir in (*) ein und erhalten

$$v = \frac{\beta_1}{\alpha_1} w + \sum_{j=2}^{r} \Big(\beta_j - \alpha_j \frac{\beta_1}{\alpha_1} \Big) v_j \, .$$

Somit ist $V \subset \mathrm{span}_K(\mathcal{B}')$, also ist auch $\mathcal{B}'$ Erzeugendensystem.

- Wir zeigen: $\mathcal{B}'$ ist linear unabhängig. Seien $\beta, \beta_i \in K$ mit

$$\beta \cdot w + \sum_{j=2}^{r} \beta_j v_j = 0$$

Wir setzen hier den Ausdruck $w = \sum_{j=1}^{r} \alpha_j v_j$ ein:

$$\beta\alpha_1 v_1 + \sum_{j=2}^{r} \Big(\beta\alpha_j + \beta_j \Big) v_j = 0$$

Da die Familie $\{v_1, \ldots, v_r\}$ linear unabhängig ist, folgen die Gleichungen

$$\beta\alpha_1 = 0 \qquad \text{und} \qquad \beta\alpha_j + \beta_j = 0$$

Aus $\alpha_1 \neq 0$ folgt im Körper K dass $\beta = 0$ und daraus $\beta_j = 0$.

□

Wir wollen mehr als einen Vektor austauschen:

Satz 2.4.13 (Austauschsatz)**.**
Sei V ein K–Vektorraum, und $\mathcal{B} = \{v_1, \ldots, v_r\}$ eine Basis von V. Sei $\{w_1, \ldots, w_n\}$ eine linear unabhängige Teilmenge von V. Dann gilt $n \leq r$, und es gibt Indizes $i_1, \ldots, i_n \in \{1, \ldots, r\}$, so dass der Austausch

$$\begin{aligned} &\text{von } v_{i_1} \text{ gegen } w_1, \ldots \\ &\text{von } v_{i_n} \text{ gegen } w_n \end{aligned}$$

eine neue Basis $\mathcal{B}^*$ von V liefert, die die vorgegebene linear unabhängige Menge $\{w_1, \ldots, w_n\}$ als Teilmenge enthält. Nach Umnummerierung zu $i_1 = 1, \ldots, i_n = n$ haben wir für die neue Basis

$$\mathcal{B}^* = \{w_1, \ldots, w_n, v_{n+1}, \ldots, v_r\} \, .$$

Beweis.
Vollständige Induktion nach n.

- Induktionsanfang: für $n = 0$ ist nichts zu zeigen. Sei die Aussage für $n - 1 \in \mathbb{N}$ gültig. Zu zeigen ist, dass die Aussage für n gültig ist.
  Sei also $\{w_1, \ldots, w_n\}$ linear unabhängig. Dann ist auch die Teilmenge $\{w_1, \ldots, w_{n-1}\}$ linear unabhängig. Nach Induktionsvoraussetzung gilt $n - 1 \leq r$ und (gegebenenfalls nach Umnummerierung) ist die Menge

$$\bar{\mathcal{B}} := \{w_1, \ldots, w_{n-1}, v_n, \ldots, v_r\}$$

eine Basis von V.

- Wir zeigen $n \leq r$. Nach Induktionsvoraussetzung ist $n-1 \leq r$; wir müssen nur $n-1 = r$ ausschließen. Dann wäre aber nach Induktionsvoraussetzung die Menge

$$\bar{\mathcal{B}} := \{w_1, \ldots, w_{n-1}\}$$

eine Basis von V, also eine maximale lineare unabhängige Teilmenge im Widerspruch zur Voraussetzung, dass auch noch die Teilmenge $\{w_1, \ldots, w_{n-1}, w_n\}$ linear unabhängig ist. Also ist $n-1 < r$, also $n \leq r$.

- Zu zeigen ist, dass es ein $i_n \in \{n, \ldots, r\}$ gibt, so dass man v_{i_n} gegen w_n austauschen kann. Da $\bar{\mathcal{B}}$ eine Basis von V ist, finde mit $\alpha_k \in K$

$$w_n = \sum_{j=1}^{n-1} \alpha_j w_j + \sum_{j=n}^{r} \alpha_j v_j \ .$$

Wären alle $\alpha_n, \ldots, \alpha_r$ gleich Null, so wäre w_n Linearkombination der $\{w_1, \ldots, w_{n-1}\}$, im Widerspruch zur vorausgesetzten linearen Unabhängigkeit von $\{w_1, \ldots, w_n\}$. Also gibt es einen Wert $i_n \in \{n, \ldots, r\}$ mit $\alpha_{i_n} \neq 0$. Wende nun das Austauschlemma 2.4.12 an und erhalte eine Basis $\mathcal{B}^* = \{w_1, \ldots, w_n, v_{n+1}, \ldots, v_r\}$.

□

Definition 2.4.14
Sei V ein endlich erzeugter K–Vektorraum und $\mathcal{B} = \{v_1, \ldots, v_r\}$ eine Basis. Die Zahl r heißt Länge der Basis $\mathcal{B}$.

Korollar 2.4.15.

1. Hat ein K–Vektorraum V eine endliche Basis, so ist jede Basis von V endlich.
2. Je zwei Basen eines endlich erzeugten K–Vektorraums V sind gleich lang.
3. Je zwei Basen eines beliebigen K–Vektorraums V sind gleich lang.

Beweis.

1. Sei $\{v_1, \ldots, v_r\}$ eine endliche Basis. Wäre eine weitere Basis $\mathcal{B}$ nicht endlich, gäbe es eine linear unabhängige Teilmenge $\{w_1, \ldots, w_{r+1}\} \subset \mathcal{B}$, im Widerspruch zum Austauschsatz 2.4.13.
2. Sind $\mathcal{B} = \{v_1, \ldots, v_r\}$ und $\mathcal{B}' = \{w_1, \ldots, w_k\}$ Basen von V, dann folgt aus dem Austauschsatz, da $\mathcal{B}'$ linear unabhängig und $\mathcal{B}$ Basis ist, $k \leq r$ und, indem man die Rollen von $\mathcal{B}'$ und $\mathcal{B}$ vertauscht, auch $r \leq k$, also $k = r$.
3. Wird mit einem Argument mit dem Zornschen Lemma gezeigt.

□

Definition 2.4.16
Für einen K–Vektorraum V setzen wir

$$\dim_K(V) = \begin{cases} r, & \text{falls } V \text{ eine Basis der Länge } r \text{ besitzt.} \\ \infty, & \text{falls } V \text{ keine endliche Basis besitzt.} \end{cases}$$

Für den Nullvektorraum setzen wir $\dim_K(\{0\}) = 0$ *und betrachten die leere Menge als Basis. Die Zahl*

$$\dim_K(V) \in \{0, 1, \ldots, \infty\}$$

heißt Dimension *des Vektorraums* V*.*

Der Nullvektorraum wird zwar von der leeren Menge $\emptyset$ erzeugt, die ihm zu Grunde liegende Menge hat aber genau ein Element, den Nullvektor.

Beispiele 2.4.17.

1. Sei K ein beliebiger Körper und $V = K^n$. Dann hat die Standardbasis $e_1, \ldots, e_n$ die Länge n; und daher ist $\dim_K K^n = n$.

2. $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, denn $\{1, \mathrm{i}\}$ ist eine $\mathbb{R}$–Basis. $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, denn $\{1\}$ ist eine $\mathbb{C}$–Basis; allgemein ist $\dim_K(K) = 1$.

3. Für den Vektorraum der Polynome gilt $\dim_{\mathbb{R}}(\mathbb{R}[X]) = \infty$, denn die abzählbar unendliche Menge $\{1, X, X^2, \ldots\}$ der Monome ist eine Basis.

Satz 2.4.18.
Sei V ein endlich erzeugter K–Vektorraum und $W \subset V$ ein Untervektorraum.

1. Dann ist W endlich erzeugt und es gilt

$$\dim_K(W) \leq \dim_K(V)$$

2. Falls Gleichheit gilt, $\dim_K(W) = \dim_K(V)$, so ist $W = V$.

Beweis.

1. Setze $n := \dim_K(V) < \infty$. Wäre W nicht endlich erzeugt, so gäbe es nach Lemma 2.4.5 sicher $n + 1$ linear unabhängige Vektoren $v_1, \ldots, v_{n+1} \in W \subset V$, im Widerspruch zum Austauschsatz 2.4.13.
   Also besitzt W eine endliche Basis $\mathcal{B} = \{w_1, \ldots, w_r\}$; da diese Familie linear unabhängig ist, folgt wiederum nach dem Austauschsatz $r = \dim_K(W) \leq \dim_K(V)$.

2. Sei nun $\dim_K(V) = \dim_K(W) = n$ und $\mathcal{B} = \{w_1, \ldots, w_n\}$ eine Basis von W. Gäbe es $v \in V \setminus \mathrm{span}_K(\mathcal{B})$, so ware $\mathcal{B} \cup \{v\}$ linear unabhängig, im Widerspruch zum Austauschsatz 2.4.13.

□

Korollar 2.4.19.
Sei V ein endlich erzeugter K–Vektorraum mit $n = \dim_K(V)$. Dann bilden je n linear unabhängige Vektoren $(v_1, \ldots, v_n)$ eine Basis.

Beweis.
Sei die Teilmenge $M = \{v_1, \ldots, v_n\}$ von V linear unabhängig. Dann ist $W := \mathrm{span}_K(M)$ ein Untervektorraum. Da W von M erzeugt ist, ist M ein Erzeugendensystem von W; da W linear unabhängig ist, ist M ein linear unabhängiges Erzeugendensystem von W, also eine Basis von W. Es gilt daher $\dim_K(W) = n$. Also haben V und der Untervektorraum $W \subset V$ die gleiche Dimension. Nach Satz 2.4.18.2 folgt $W = V$. Also ist M auch ein Erzeugendensystem, also eine Basis von V. □

2.5 Summen von Untervektorräumen

Definition 2.5.1
Sei V ein K–Vektorraum und seien $W_1, \ldots, W_r \subset V$ Untervektorräume. Dann heißt

$$W_1 + \ldots + W_r := \left\{ v \in V \mid \exists w_i \in W_i \ \text{ mit } v = \sum_{i=1}^{r} w_i \right\}$$

die (innere) Summe der Untervektorräume $W_1, \ldots, W_r$.

Satz 2.5.2.

1. Es gilt $W_1 + \ldots + W_r = \operatorname{span}_K(W_1 \cup \ldots \cup W_r)$. Insbesondere ist die innere Summe $W_1 + \ldots + W_r$ ein Untervektorraum von V.

2. $\dim_K(W_1 + \ldots W_r) \leq \dim_K W_1 + \ldots \dim_K W_r$. Hierbei verwenden wir für alle $n \in \mathbb{N}$ die Konvention:
$$\infty + \infty = \infty, \quad \infty + n = \infty, \quad \infty > n \ .$$

Beweis.

1. Jeder Vektor $v \in W_1 + W_2 + \ldots + W_r$ lässt sich als Summe $v = \sum_{i=1}^{r} w_i$ mit $w_i \in W_i$ schreiben und liegt daher im Erzeugnis $\operatorname{span}_K(W_1 \cup \ldots \cup W_r)$. Damit ist die Inklusion "$\subset$" klar.

   Sei $v \in \operatorname{span}_K(W_1 \cup \ldots \cup W_k)$; d.h. es gibt $u_1^{(i)}, \ldots, u_{n_i}^{(i)} \in W_i$ und $\alpha_1^{(i)}, \ldots, \alpha_{n_i}^{(i)} \in K$, so dass
   $$v = \sum_{i=1}^{r} \underbrace{\sum_{j=1}^{n_i} \alpha_j^{(i)} u_j^{(i)}}_{\in W_i} \text{ gilt.}$$

   Die Linearkombination $w_i := \sum_{j=1}^{n_i} \alpha_j^{(i)} u_j^{(i)}$ liegt in W_i, also liegt $v = w_1 + \ldots + w_r \in W_1 + \ldots + W_r$.

2. Ist eine der Dimensionen $\dim_K W_i$ gleich ∞, so ist nichts zu zeigen. Seien also alle Unterräume W_i endlich–dimensional und $\mathcal{B}_i \subset W_i$ Basen. Wir setzen
   $$\mathcal{B} := \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r \qquad \text{und} \qquad W := W_1 + \ldots + W_r \ .$$

   Wegen $\mathcal{B}_i \subset W_i \subset W$ gilt auch für die Vereinigung der Basen $\mathcal{B} \subset W$. Wir zeigen zunächst, dass $\mathcal{B}$ ein Erzeugendensystem von W ist.

   Da W ein Vektorraum ist, folgt aus $\mathcal{B} \subset W$, dass $\operatorname{span}_K(\mathcal{B}) \subset W$. Umgekehrt kann jedes $w \in W$ als Summe $w = w_1 + \ldots + w_r$ mit $w_i \in W_i$ geschrieben werden. Jedes w_i ist Linearkombination von Vektoren in $\mathcal{B}_i$ und damit wegen $\mathcal{B}_i \subset \mathcal{B}$ erst recht Linearkombination von Vektoren in $\mathcal{B}$, also in $\operatorname{span}_K(\mathcal{B})$. Damit ist auch die Summe $w = w_1 + \ldots + w_r$ mit $w_i \in W_i$ im Unterraum $\operatorname{span}_K(\mathcal{B})$.

   Daraus folgen die Ungleichungen:
   $$\dim_K(W_1 + \ldots + W_r) \leq |\mathcal{B}| \leq |\mathcal{B}_1| + \ldots + |\mathcal{B}_r| = \dim_K(W_1) + \ldots + \dim_K(W_r) \ .$$

   Die erste Ungleichung gilt, weil $\mathcal{B}$ ein Erzeugendensystem von $W_1 + \ldots + W_r$ ist. Die zweite Ungleichung folgt aus $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ und die dritte Gleichheit, weil die Menge $\mathcal{B}_i$ nach Voraussetzung eine Basis von W_i ist.

□

Wir verschärfen das letzte Resultat von einer Ungleichung zu einer Gleichung:

Satz 2.5.3.
Sei V ein K–Vektorraum und $W_1, W_2 \subset V$ zwei endlich–dimensionale Untervektorräume. Dann gilt:

$$\dim_K(W_1 + W_2) = \dim_K(W_1) + \dim_K(W_2) - \dim_K(W_1 \cap W_2) \; .$$

Als Beispiel betrachten wir mit $K = \mathbb{R}$ im Vektorraum $V = \mathbb{R}^3$ die Untervektorräume

$$\begin{aligned} W_1 = \mathrm{span}(e_1, e_2): &\quad x\text{-}y\text{-Ebene} \\ W_2 = \mathrm{span}(e_1, e_3): &\quad x\text{-}z\text{-Ebene} \\ W_1 \cap W_2 = \mathrm{span}(e_1): &\quad x\text{-Achse} \end{aligned}$$

Es gilt $W_1 + W_2 = \mathbb{R}^3$; in der Tat liefert die Dimensionsformel aus Satz 2.5.3 die Dimension $3 = 2 + 2 - 1$. Man kann auch die Dimension der inneren Summe von mehr als zwei Untervektorräumen beschreiben; dies ist eine gute Übung.

Beweis.

- Ergänze eine Basis $\mathcal{B} = \{v_1, \ldots, v_m\}$ von $W_1 \cap W_2$ zu einer Basis $\mathcal{B}_1 := \{v_1, \ldots, v_m, w_1, \ldots, w_k\}$ von W_1 und $\mathcal{B}_2 := \{v_1, \ldots, v_m, u_1, \ldots, u_l\}$ zu einer Basis von W_2. Wir behaupten, dass dann

  $$\widetilde{\mathcal{B}} := \mathcal{B}_1 \cup \mathcal{B}_2 = \{v_1, \ldots, v_m, w_1, \ldots, w_k, u_1 \ldots u_l\}$$

  eine Basis von $W_1 + W_2$ ist. Daraus folgt dann sofort

  $$\dim_K(W_1+W_2) = m+k+l = (m+k)+(m+l)-m = \dim_K W_1 + \dim_K W_2 - \dim_K W_1 \cap W_2 \; .$$

- Wir zeigen: $\widetilde{\mathcal{B}}$ ist ein Erzeugendensystem von $W_1 + W_2$. Es gilt: $\mathcal{B}_i \subset \tilde{\mathcal{B}}$, woraus folgt $W_i = \mathrm{span}_K(\mathcal{B}_i) \subset \mathrm{span}_K(\widetilde{\mathcal{B}})$. Daher ist $W_1 \cup W_2 \subset \mathrm{span}_K(\widetilde{\mathcal{B}})$. Es folgt

  $$W_1 + W_2 = \mathrm{span}_K(W_1 \cup W_2) \subset \mathrm{span}_K(\widetilde{\mathcal{B}}) \subset W_1 + W_2 \; ,$$

  wobei die letzte Inklusion ohnehin klar ist.

- Wir zeigen: $\widetilde{\mathcal{B}}$ ist linear unabhängig. Betrachte eine Darstellung des Nullvektors als Linearkombination von Vektoren in $\tilde{\mathcal{B}}$:

  $$\sum_{p=1}^{m} \alpha_p v_p + \sum_{q=1}^{k} \beta_q w_q + \sum_{r=1}^{l} \gamma_r u_r = 0 \; . \qquad (*)$$

  Setze

  $$v := \sum_{p=1}^{m} \alpha_p v_p + \sum_{q=1}^{k} \beta_q w_q \in W_1 \qquad (**)$$

  Aus $(*)$ folgt sofort

  $$v = -\sum_{r=1}^{l} \gamma_r u_r \in W_2 \qquad (***) \; .$$

  Also liegt $v \in W_1 \cap W_2$. In $(**)$ ist v in der Basis $\mathcal{B}_1$ von W_1 dargestellt. Wegen der Eindeutigkeit der Darstellung bezüglich dieser Basis folgt $\beta_q = 0$ für alle q. In $(***)$ ist v in der Basis $\mathcal{B}_2$ dargestellt. Mit dem gleichen Argument folgt $\gamma_r = 0$ für alle r und somit $v = 0$. Da $\mathcal{B}$ eine Basis von $W_1 \cap W_2$ ist, müssen auch alle α_i verschwinden.

□

Lemma 2.5.4.
Sei V ein endlich–dimensionaler K–Vektorraum, seien $W_1, W_2 \subset V$ Untervektorräume mit $W_1 + W_2 = V$. Dann sind äquivalent:

1. $W_1 \cap W_2 = \{0\}$
2. Jedes Element $v \in V$ lässt sich in eindeutige Weise als Summe $v = w_1 + w_2$ mit $w_i \in W_i$ schreiben.
3. $\dim_K(W_1 + W_2) = \dim_K W_1 + \dim_K W_2$.

Beweis.
Wir zeigen 3.⇒ 1. ⇒ 2. ⇒3.

3.⇒1. Aus der Dimensionsformel aus Satz 2.5.3 folgt $\dim_K(W_1 \cap W_2) = 0$, also $W_1 \cap W_2 = \{0\}$.

1.⇒2. Wegen $V = W_1 + W_2$ lässt sich jedes $v \in V$ schreiben als $v = w_1 + w_2$ mit $w_i \in W_i$. Angenommen, es gäbe eine weitere Darstellung:

$$v = w_1 + w_2 = w_1' + w_2' \qquad w_i' \in W_i \ .$$

Daraus folgt

$$w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{0\} \ ,$$

also $w_1 = w_1'$ und $w_2 = w_2'$.

2.⇒3. Wäre $\dim_K(W_1 \cap W_2) > 0$, so gäbe es $w \in W_1 \cap W_2$ mit $w \neq 0$. Dann finde für $v = 0 \in V$ die beiden verschiedenen Darstellungen

$$v = 0 + 0 = w - w$$

im Widerspruch zur Eindeutigkeitsaussage in 2. Aus der Dimensionsformel aus Satz 2.5.3 folgt nun 3.

□

Definition 2.5.5
Ist $V = W_1 + W_2$ und gilt $W_1 \cap W_2 = \{0\}$ (oder jede andere der äquivalenten Bedingungen aus Lemma 2.5.4), so sagt man, V sei die (innere) direkte Summe der Untervektorräume W_1 und W_2, in Zeichen

$$V = W_1 \oplus W_2 \ .$$

Beispiel 2.5.6.
Betrachte $K = \mathbb{R}$ und den Vektorraum $V = \mathbb{R}^3$ mit den beiden Untervektorräumen $W_1 = \mathrm{span}_{\mathbb{R}}(e_1, e_2)$ und $W_2 = \mathrm{span}_{\mathbb{R}}(e_3)$. Dann ist $V = W_1 \oplus W_2$. Im Beispiel nach Satz 2.5.3 ist die Summe nicht direkt.

Satz 2.5.7.
Sei V ein endlich–dimensionaler Vektorraum. Seien $W_1, W_2 \subset V$ Untervektorräume. Dann sind äquivalent:

1. $V = W_1 \oplus W_2$
2. Für jede Basis $\mathcal{B}_1$ von W_1 und $\mathcal{B}_2$ von W_2 ist die Vereinigung $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ disjunkt, und $\mathcal{B}$ ist eine Basis von V.
3. $V = W_1 + W_2$ und $\dim_K V = \dim_K W_1 + \dim_K W_2$.

Beweis.

1.⇒3. Klar wegen der Dimensionsformel aus Satz 2.5.3.

3.⇒2. $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ ist Erzeugendensystem von $\operatorname{span}_K(W_1 \cup W_2) = W_1 + W_2 = V$. Es ist also

$$\dim_K V \leq |\mathcal{B}| \leq |\mathcal{B}_1| + |\mathcal{B}_2| = \dim_K W_1 + \dim_K W_2 = \dim_K V \ .$$

Hieraus folgt

$$\dim_K V = |\mathcal{B}| = |\mathcal{B}_1| + |\mathcal{B}_2|$$

Also ist $\mathcal{B}$ als minimales Erzeugendensystem eine Basis und $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ wegen $|\mathcal{B}| = |\mathcal{B}_1| + |\mathcal{B}_2|$.

2.⇒1. Sei $\mathcal{B}_i$ Basis von W_i, dann ist $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ Basis von V. Es folgt

$$V = \operatorname{span}_K(\mathcal{B}) = \operatorname{span}_K(\mathcal{B}_1 \cup \mathcal{B}_2) = \operatorname{span}_K(W_1 \cup W_2) = W_1 + W_2 \ .$$

und $\dim_K V = |\mathcal{B}| = |\mathcal{B}_1| + |\mathcal{B}_2| = \dim_K W_1 + \dim_K W_2$. Die Dimensionsformel aus Satz 2.5.3 liefert $\dim_K W_1 \cap W_2 = 0$, also ist die Summe direkt, $V = W_1 \oplus W_2$.

□

3 Lineare Abbildungen

Vektorräume sind die zentralen mathematischen Objekte der linearen Algebra. So wie wir schon für Gruppen in Definition 2.1.12 und Ringe in Definition 2.2.6.2 eine passende Klasse von Abbildungen ausgezeichnet haben, nämlich die Gruppenhomomorphismen, so müssen wir auch für Vektorräume eine Klasse von Abbildungen finden, die mit der Vektorraumstruktur verträglich sind.

3.1 Definition, Dimensionsformel

Definition 3.1.1
Sei K ein Körper. Seien V,W zwei K–Vektorräume. Eine Abbildung

$$\Phi: \quad V \to W$$

heißt K–linear oder K–Vektorraumhomomorphismus, falls gilt

(L1) *Φ ist Gruppenhomomorphismus bezüglich der Addition, d.h. für alle $v, v' \in V$ gilt $\Phi(v + v') = \Phi(v) + \Phi(v')$*

(L2) *Φ ist mit der skalaren Multiplikation verträglich, d.h. für alle $v \in V$ und für alle $\lambda \in K$ gilt $\Phi(\lambda v) = \lambda\Phi(v)$.*

Bemerkungen 3.1.2.

1. Eine Abbildung $\Phi : V \to W$ ist genau dann linear, wenn gilt
   (L) für alle $v, v' \in V$ und für alle $\lambda, \lambda' \in K$ ist $\Phi(\lambda v + \lambda' v') = \lambda\Phi(v) + \lambda'\Phi(v')$.

2. Induktiv zeigt man, dass dann für alle endlichen Familie $v_1, \ldots, v_n \in V$ und $\lambda_1, \ldots, \lambda_n \in K$ gilt
$$\Phi(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1\Phi(v_1) + \ldots + \lambda_n\Phi(v_n) \; .$$

Beweis.
Um zu sehen, dass aus (L) die Bedingung $(L1)$ folgt, wähle in (L) die Familie v_1, v_2 und für die Skalare $\lambda = \lambda' = 1$; um $(L2)$ zu sehen, setze $\lambda' = 0$. Dass aus $(L1)$ und $(L2)$ die Gleichung (L) folgt, zeigt die folgende Rechnung

$$\Phi(\lambda v + \lambda' v') \underset{(L1)}{=} \Phi(\lambda v) + \Phi(\lambda' v') \underset{(L2)}{=} \lambda\Phi(v) + \lambda'\Phi(v') \; .$$

□

Satz 3.1.3.
Sei $\Phi: \quad V \to W$ eine lineare Abbildung. Dann gilt

1. $\Phi(0) = 0$ und $\Phi(v - v') = \Phi(v) - \Phi(v')$ für alle $v, v' \in V$.

2. Ist $V' \subset V$ ein Untervektorraum, so ist das Bild $\Phi(V')$ ein Untervektorraum von W. Insbesondere ist $\Phi(V) = \mathrm{Im}\,(\Phi)$ ein Untervektorraum von W.

3. Sei $W' \subset W$ ein Untervektorraum; dann ist das Urbild $\Phi^{-1}(W')$ ein Untervektorraum von V. Insbesondere ist $\Phi^{-1}(0) = \ker \Phi$ ein Untervektorraum von V.

4. Ist Φ linear und bijektiv, so ist die Umkehrabbildung $\Phi^{-1} : W \to V$ ebenfalls linear.

Beweis.

1. $\Phi(0) = 0$ folgt aus Satz 2.1.14.1 für Gruppenhomomorphismen, angewandt auf die Gruppe $(V, +)$. Wir rechnen

$$\Phi(v - v') = \Phi(1 \cdot v + (-1) \cdot v') \underset{(L)}{=} 1 \cdot \Phi(v) + (-1)\Phi(v') = \Phi(v) - \Phi(v') \; .$$

2. Da $V' \subset V$ ein Untervektorraum ist, folgt $0 \in V'$, also $\Phi(0) = 0 \in \Phi(V')$. Also $\Phi(V') \neq \emptyset$. Seien $w_1, w_2 \in \Phi(V')$ und $\lambda_1, \lambda_2 \in K$. Zu zeigen ist:

$$\lambda_1 w_1 + \lambda_2 w_2 \in \Phi(V') \; .$$

   Wähle Urbilder $v_i \in V'$ mit $\Phi(v_i) = w_i$. Dann ist

$$v := \lambda_1 v_1 + \lambda_2 v_2 \in V' \; ,$$

   da V' ein Untervektorraum von V ist. Rechne

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 \Phi(v_1) + \lambda_2 \Phi(v_2) \underset{(L)}{=} \Phi(\lambda_1 v_1 + \lambda_2 v_2) = \Phi(v) \in \Phi(V') \; .$$

3. Sei $W' \subset W$ ein Untervektorraum. Also $0 \in W'$, daher $0 \in \Phi^{-1}(\{0\}) \subset \Phi^{-1}(W')$. Seien $v_1, v_2 \in \Phi^{-1}(W')$ und $\lambda_1, \lambda_2 \in K$. Es gilt

$$\Phi(\lambda_1 v_1 + \lambda_2 v_2) \underset{(L)}{=} \lambda_1 \Phi(v_1) + \lambda_2 \Phi(v_2) \in W' \; ,$$

   also $\lambda_1 v_1 + \lambda_2 v_2 \in \Phi^{-1}(W')$.

4. Nach Satz 2.1.14.3 ist Φ^{-1} Gruppenhomomorphismus bezüglich der Addition. Für $w \in W$ und $\lambda \in K$ gilt

$$\lambda \Phi^{-1}(w) = \Phi^{-1}\Phi\left(\lambda \Phi^{-1}(w)\right) = \Phi^{-1}\lambda\left(\Phi\Phi^{-1}(w)\right) = \Phi^{-1}(\lambda w) \; .$$

□

Beispiele 3.1.4.

1. Sei $V = W = K$; definiere eine lineare Abbildung $\Phi : V \to W$ durch $\Phi(\beta) := \beta\gamma$ für ein festes $\gamma \in K$.

   (L1) folgt aus $\Phi(\beta + \beta') = (\beta + \beta')\gamma = \beta\gamma + \beta'\gamma = \Phi(\beta) + \Phi(\beta')$, wegen des Distributivgesetzes.

   (L2) folgt aus $\Phi(\lambda\beta) = (\lambda\beta)\gamma = \lambda(\beta\gamma) = \lambda\Phi(\beta)$ mit dem Assoziativgesetz der Multiplikation.

Wir berechnen

$$\ker \Phi = \Big\{\beta \in K |\, \beta\gamma = 0\Big\} = \begin{cases} \{0\}, & \text{falls } \gamma \neq 0 \\ K, & \text{falls } \gamma = 0 \end{cases}$$

d.h. Φ ist injektiv genau für $\gamma \neq 0$.

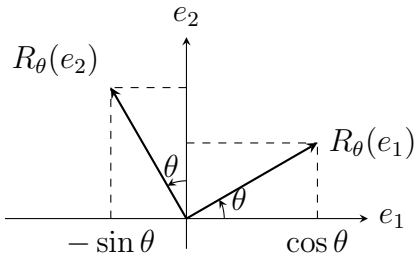
$$\operatorname{Im} \Phi = \Big\{\beta\gamma \in K |\, \beta \in K\Big\} = \begin{cases} \{0\}, & \text{falls } \gamma = 0 \\ K, & \text{falls } \gamma \neq 0 \end{cases}$$

d.h. Φ ist surjektiv genau für $\gamma \neq 0$.

2. $K = \mathbb{R}$ und $V = W = \mathbb{R}^2$. Wir wählen ein festes $\theta \in \mathbb{R}$ und betrachten

$$\begin{aligned} R_\theta :&\quad \mathbb{R}^2 \to \mathbb{R}^2 \\ R_\theta \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos\theta x - \sin\theta y \\ \sin\theta x + \cos\theta y \end{pmatrix} \end{aligned}$$

R_θ ist eine Drehung um den Winkel θ gegen den Uhrzeigersinn:



zu (L1)

$$\begin{aligned} R_\theta\left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix}\right) = R_\theta \begin{pmatrix} x + x' \\ y + y' \end{pmatrix} &= \begin{pmatrix} \cos\theta(x + x') - \sin\theta(y + y') \\ \sin\theta(x + x') + \cos\theta(y + y') \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta x - \sin\theta y \\ \sin\theta x + \cos\theta y \end{pmatrix} + \begin{pmatrix} \cos\theta x' - \sin\theta y' \\ \sin\theta x' + \cos\theta y' \end{pmatrix} \\ &= R_\theta \begin{pmatrix} x \\ y \end{pmatrix} + R_\theta \begin{pmatrix} x' \\ y' \end{pmatrix} . \end{aligned}$$

(L2) rechnet man analog nach:

$$\begin{aligned} R_\theta\left(\lambda \begin{pmatrix} x \\ y \end{pmatrix}\right) = R_\theta \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} &= \begin{pmatrix} \cos\theta\lambda x - \sin\theta\lambda y \\ \sin\theta\lambda x + \cos\theta\lambda y \end{pmatrix} \\ &= \lambda \begin{pmatrix} \cos\theta x - \sin\theta y \\ \sin\theta x + \cos\theta y \end{pmatrix} = \lambda R_\theta \begin{pmatrix} x \\ y \end{pmatrix} . \end{aligned}$$

3. Seien der Körper K und ein K-Vektorraum $V = W$ beliebig. Dann ist $\Phi = \mathrm{id}_V$ eine lineare Abbildung.

4. Sei $V = K$ und W beliebig. Eine lineare Abbildung $\Phi : K \to W$ liefert den Vektor $\Phi(1_K) \in W$. Dieser Vektor in W reicht aus, um die Abbildung zu kennen, denn es gilt $\Phi(\lambda) = \Phi(\lambda 1_K) = \lambda\Phi(1_K)$ für alle $\lambda \in K$. Dies liefert eine Bijektion von linearen Abbildungen von K nach W auf Vektoren in W.

5. Sei K ein beliebiger Körper und $V = W = K[X]$ gleich dem Polynomring über K. Definiere
$$\text{diff}: \quad K[X] \to K[X]$$
durch
$$\text{diff}\Big(\sum_{j=0}^{n} \alpha_j X^j\Big) = \sum_{j=0}^{n} j\alpha_j X^{j-1} .$$
Man nennt diese Abbildung auch die formale Anleitung von Polynomen. Zur Übung zeige man, dass die formale Ableitung linear ist.

6. Sei K ein beliebiger Körper, und seien V, W, Z drei K–Vektorräume. Seien $\Phi: V \to W$ und $\Psi: W \to Z$ zwei lineare Abbildungen, dann ist auch ihre Verknüpfung $\Phi \circ \Psi$ eine lineare Abbildung. Denn seien $\lambda, \lambda' \in K$ und $v, v' \in V$:
$$\begin{aligned} \Psi \circ \Phi(\lambda v + \lambda' v') &= \Psi(\Phi(\lambda v + \lambda' v')) \\ &= \Psi(\lambda\Phi(v) + \lambda'\Phi(v')) \qquad \text{da } \Phi \text{ linear} \\ &= \lambda\Psi(\Phi(v) + \lambda'\Psi(\Phi(v'))) \qquad \text{da } \Psi \text{ linear} \\ &= \lambda(\Psi \circ \Phi)(v) + \lambda'(\Psi \circ \Phi)(v') . \end{aligned}$$

Definition 3.1.5

1. *Eine lineare Abbildung $\Phi: \quad V \to W$ heißt*

Monomorphismus falls	*Φ injektiv*
Epimorphismus	*Φ surjektiv*
Isomorphismus	*Φ bijektiv*
Endomorphismus	*$V = W$*
Automorphismus	*$V = W$ und Φ bijektiv*

ist.

2. *Zwei K–Vektorräume V, W heißen isomorph, in Zeichen $V \cong W$, falls ein Vektorraumisomorphismus $\Phi: V \to W$ existiert.*

3. *Sei $\Phi: V \to W$ eine lineare Abbildung. Dann heißt*
$$\text{rg}\,(\Phi) := \dim_K \text{Im}\,\Phi$$
der Rang der linearen Abbildung Φ.

Isomorphie ist eine Äquivalenzrelation: sie ist reflexiv, denn id ist wegen Beispiel 3.1.4.3 ein Isomorphismus. Transitivität folgt aus Beispiel 3.1.4.5. und Symmetrie, da nach Satz 3.1.3.4 das Inverse eines Isomorphismus ein Isomorphismus ist.

Bemerkung 3.1.6.
Sei $\Phi: V \to W$ eine lineare Abbildung und $E \subset V$ ein Erzeugendensystem von V. Dann ist $\Phi(E)$ ein Erzeugendensystem des Bildes $\Phi(V)$. Denn für $w \in \Phi(V)$ existiert $v \in V$ mit

$\Phi(v) = w$. Weil E ein Erzeugendensystem von V ist, können wir für dieses $v \in V$ Skalare $\lambda_1, \ldots, \lambda_n \in K$ sowie $v_1, \ldots, v_n \in E$ finden, so dass

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

gilt, woraus

$$w = \Phi(v) = \sum \lambda_i \Phi(v_i)$$

folgt.

Satz 3.1.7 (Dimensionsformel)**.**
Sei $\Phi : V \to W$ eine lineare Abbildung und $\dim_K V < \infty$. Dann gilt die Dimensionsformel

$$\dim_K V = \dim_K \ker \Phi + \operatorname{rg}(\Phi) .$$

Beweis.

- $\ker \Phi$ ist als Untervektorraum von V endlich–dimensional; wähle eine Basis $\{v_1, \ldots, v_k\}$ von $\ker \Phi$ mit $k := \dim_K \ker \Phi$ und ergänze diese nach dem Basisergänzungssatz 2.4.11.2 zu einer Basis $\{v_1, \ldots, v_n\}$ von V mit $n = \dim_K V$. Nach Bemerkung 3.1.6 ist $\{\Phi(v_1), \ldots, \Phi(v_n)\}$ ein Erzeugendensystem des Bildes $\Phi(V)$. Aber es gilt, da $v_1, \ldots, v_r$ in $\ker \Phi$ liegt, $0 = \Phi(v_1) = \ldots = \Phi(v_k)$, also ist schon die Teilmenge

  $$\{\Phi(v_{k+1}), \ldots, \Phi(v_n))\}$$

  ein Erzeugendensystem von $\Phi(V)$.

- Wir zeigen, dass dieses Erzeugendensystem des Bildes $\Phi(V)$ linear unabhängig und somit eine Basis von $\Phi(V)$ ist. Seien $\lambda_{k+1}, \ldots, \lambda_n \in K$ und gelte

  $$\sum_{j=k+1}^{n} \lambda_i \Phi(v_i) = 0 .$$

  Hieraus folgt $\Phi\Big(\sum_{i=k+1} \lambda_i v_i\Big) = 0$, und somit $\sum_{i=k+1}^{n} \lambda_i v_i \in \ker \Phi$. Daher gibt es $\lambda_1, \ldots, \lambda_k \in K$ mit

  $$\sum_{i=k+1}^{n} \lambda_i v_i = \sum_{i=1}^{k} (-\lambda_i) v_i$$
  $$\Leftrightarrow \sum_{i=1}^{n} \lambda_i v_i = 0$$

  Da $\{v_i\}$ als Basis von V linear unabhängig ist, folgt $\lambda_i = 0$ für $i = 1, \ldots, n$. Insbesondere ist die Familie $\Big\{\Phi(v_i)\Big\}_{i=k+1,\ldots,n}$ von Vektoren in W linear unabhängig. Damit gilt $\operatorname{rg} \Phi = n - k = \dim_K V - \dim_K \ker \Phi$.

□

Korollar 3.1.8.
Sei $\Phi : V \to W$ eine lineare Abbildung und $\dim_K V = \dim_K W < \infty$. Dann sind äquivalent:

1. Φ ist injektiv
2. Φ ist surjektiv
3. Φ ist bijektiv

Achtung: Diese Aussage gilt *nicht* für unendlich–dimensionale Vektorräume! Ein Gegenbeispiel ist für $K = \mathbb{R}$ und $V = \mathbb{R}[x]$ die Abbildung diff : $\mathbb{R}[x] \to \mathbb{R}[x]$ aus Beispiel 3.1.4.4, die als Kern die Polynome vom Grad Null hat. Sie ist also nicht injektiv; aber sie ist surjektiv.

Beweis.
Es genügt, die Äquivalenz von 1. und 2. zu zeigen. Φ ist genau dann injektiv, wenn der Kern trivial ist, $\ker \Phi = \{0\}$. Dies ist genau dann der Fall, wenn $\dim_K \ker \Phi = 0$ gilt. Wegen der Dimensionsformel 3.1.7 ist dies äquivalent zu $\dim_K W = \dim_K V = \dim_K \operatorname{Im} \Phi$. Aus Satz 2.4.18.2 folgt, dass dies äquivalent zu $W = \operatorname{Im} \Phi$ ist, also dazu, dass die lineare Abbildung Φ surjektiv ist. □

Satz 3.1.9.
1. Sei X eine Menge und W ein K–Vektorraum. Dann wird die Menge der Abbildungen $\mathrm{Abb}(X, W)$ durch die Operationen auf den Funktionswerten zu einem K–Vektorraum, vergleiche Beispiel 2.3.2.5.
2. Ist auch $X = V$ ein K–Vektorraum, so setzen wir
$$\mathrm{Hom}_K(V, W) := \{\Phi : V \to W : \Phi \quad \text{ist linear}\} .$$
Es ist
$$\mathrm{Hom}_K(V, W) \subset \mathrm{Abb}(V, W)$$
ein K-Untervektorraum.
3. Im Spezialfall $V = W$ führen wir die Bezeichnung
$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V) .$$
ein. Die Hintereinanderausführung von Abbildungen versieht $\mathrm{End}_K(V)$ mit einem Produkt
$$\circ : \mathrm{End}_K(V) \times \mathrm{End}_K(V) \to \mathrm{End}_K(V) ,$$
durch die $(\mathrm{End}_K(V), +, \circ)$ zu einem Ring mit Eins wird, dem Endomorphismenring.

Beweis.
1. Gegeben $f, g \in \mathrm{Abb}(X, W)$, setze $(f + g)(x) := f(x) + g(x)$. Man erhält eine abelsche Gruppe, vgl. Beispiel 2.2.3.2, mit neutralem Element $0(x) = 0$ und Inversen $(-f)(x) = -f(x)$. Die skalare Multiplikation ist durch $(\lambda f)(x) := \lambda f(x)$ definiert.
2. Sicher ist die Nullabbildung $\Phi(v) = 0$ für alle $v \in V$ linear und daher in $\mathrm{Hom}_K(V, W)$, daher ist $\mathrm{Hom}_K(V, W) \neq \emptyset$. Die Summe $\Phi + \Psi$ linearer Abbildungen ist wieder linear:
$$\begin{aligned}(\Phi + \Psi)(\lambda v + \lambda' v') &= \Phi(\lambda v + \lambda' v') + \Psi(\lambda v + \lambda' v') = \lambda\Phi(v) + \lambda'\Phi(v') + \lambda\Psi(v) + \lambda'\Psi(v') \\ &= \lambda(\Phi + \Psi)(v) + \lambda'(\Phi + \Psi)(v')\end{aligned}$$
Ähnlich rechnet man nach, dass auch die Abbildung $\lambda\Phi$ wieder linear ist.
3. Die Eins im Ring $\mathrm{End}_K(V)$ ist die Identitätsabbildung id_V.

□

Wir werden sehen, dass der Endomorphismenring für $\dim_K V > 1$ nicht kommutativ ist.

3.2 Matrizen

Wir wollen nun lineare Abbildungen zwischen K-Vektorräumen mit Hilfe von Basen explizit durch Elemente in K beschreiben. Zentral hierfür ist

Satz 3.2.1.
Gegeben seien endlich–dimensionale Vektorräume V und W sowie Vektoren $v_1, \ldots, v_r \in V$ und $w_1, \ldots, w_r \in W$. Dann gilt:

1. Ist die Familie $(v_1, \ldots, v_r)$ in V linear unabhängig, so gibt es mindestens eine lineare Abbildung $\Phi: \ V \to W$ mit $\Phi(v_i) = w_i$ für $i = 1, \ldots, r$. Man kann also die Werte einer linearen Abbildung auf einer linear unabhängigen Menge beliebig vorschreiben.

2. Ist $(v_1, \ldots, v_r)$ sogar eine Basis von V, so gibt es genau eine lineare Abbildung $\Phi: \ V \to W$ mit $\Phi(v_i) = w_i$. Diese hat die beiden Eigenschaften:

   (a) $\Phi(V) = \operatorname{Im} \Phi = \operatorname{span}_K(w_1, \ldots, w_r)$

   (b) Die Abbildung Φ ist genau dann injektiv, wenn die Familie $(w_1, \ldots, w_r)$ in W linear unabhängig ist.

Beweis.

- Wir zeigen erst 2: da die Familie $(v_1, \ldots, v_r)$ eine Basis ist, hat jedes $v \in V$ eine eindeutige Darstellung $v = \sum_{i=1}^r \lambda_i v_i$. Aus der Linearität von Φ und der Bedingung $\Phi(v_i) = w_i$ folgt, dass
$$\Phi(v) = \sum_{i=1}^r \lambda_i \Phi(v_i) = \sum_{i=1}^r \lambda_i w_i \qquad (*)$$
gelten muss. Es gibt also höchstens eine lineare Abbildung mit den genannten Eigenschaften. Man rechnet leicht nach, dass die durch $(*)$ definierte Abbildung auch linear ist.

- Aus $(*)$ folgt auch die Inklusion $\operatorname{Im} \Phi \subset \operatorname{span}_K(w_1, \ldots, w_r)$. Umgekehrt gilt für ein beliebiges $w = \sum_{i=1}^r \mu_i w_i \in \operatorname{span}_K(w_1, \ldots, w_r)$, dass
$$w = \Phi\Big(\sum_{i=1}^r \mu_i v_i\Big) \in \operatorname{Im}(\Phi)$$
Somit folgt $\operatorname{Im}(\Phi) = \operatorname{span}(w_1, \ldots, w_r)$.

- Sei die Familie $(w_1, \ldots, w_r)$ linear abhängig. Dann gibt es Skalare $\mu_i \in K$ mit $\sum_{i=1}^r \mu_i w_i = 0$, wobei nicht alle μ_i gleich Null sind. Weil aber die Familie $(v_1, \ldots, v_r)$ linear unabhängig ist, ist $v := \sum_{i=1}^r \mu_i v_i \neq 0$. Es gilt
$$\Phi\Big(\sum_{i=1}^r \mu_i v_i\Big) = \sum_{i=1}^r \mu_i \Phi(v_i) = \sum_{i=1}^r \mu_i w_i = 0 \ ,$$
also ist $v \in \ker \Phi$. Damit ist $\ker \Phi$ nicht trivial und Φ nicht injektiv.

- Sei nun $v \in \ker \Phi$, also $\Phi(v) = 0$. Wir schreiben v als eindeutig bestimmte Linearkombination
$$v = \sum_{i=1}^r \lambda_i v_i$$

und erhalten die Gleichung

$$0 = \Phi(v) = \sum_{i=1}^{r} \lambda_i w_i \ .$$

Ist die Familie (w_i) linear unabhängig, so folgt $\lambda_i = 0$ für alle $i = 1, \ldots, r$, also $v = 0$. Also ist Φ injektiv, wenn die Familie (w_i) linear unabhängig ist.

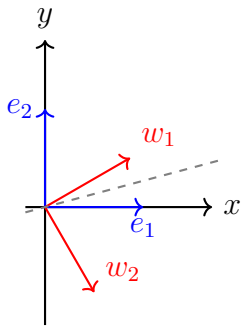
- Ist die Familie $(v_1, \ldots, v_r)$ nur linear unabhängig, aber keine Basis, so können wir mit Hilfe des Basisergänzungssatzes 2.4.11.3 die Familie zu einer Basis von V ergänzen:

  $$(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$$

  und ein Φ durch Vorgabe beliebiger Werte $w_{r+1}, \ldots, w_n \in W$ für $v_{r+1}, \ldots, v_n$ wie in 2. festlegen.

□

Beispiel 3.2.2.
Wir haben somit eine Möglichkeit, uns konkrete lineare Abbildungen zu verschaffen. Sei zum Beispiel $K = \mathbb{R}$ und $V = W = \mathbb{R}^2$. Wir wählen für $(v_1, v_2) = (e_1, e_2)$ die Standardbasis und als Bilder für eine lineare Abbildung $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ die Vektoren $w_1 = (\cos\varphi, \sin\varphi)$ und $w_2 = (\sin\varphi, -\cos\varphi)$. Die beiden Vektoren stehen bezüglich des Standardskalarprodukts auf einander senkrecht. Wie wir später sehen werden, haben wir eine Spiegelung beschrieben:



Korollar 3.2.3.
Isomorphe Vektorräume haben die gleiche Dimension.

Beweis.
Wir zeigen den Satz für endlich-dimensionale Vektorräume. Sei $\Phi : V \to W$ ein Isomorphismus von K-Vektorräumen und $(v_1, \ldots, v_n)$ eine Basis von V. Dann ist die Familie $(\Phi(v_1), \ldots, \Phi(v_n))$ linear unabhängig, weil Φ injektiv ist, und ein Erzeugendensystem, weil Φ surjektiv ist, also eine Basis von W. Die Vektorräume V und W haben also Basen gleicher Länge, also die gleiche Dimension. □

Korollar 3.2.4.
Ist V ein K–Vektorraum mit geordneter Basis $\mathcal{B} = (v_1, \ldots, v_n)$, so gibt es genau einen Isomorphismus

$$\Phi_{\mathcal{B}} : \quad K^n \to V$$

mit $\Phi_{\mathcal{B}}(e_i) = v_i$ für $i = 1, \ldots, n$ wobei e_i die kanonische Basis von K^n bezeichnet. Es ist $n = \dim_K V$; insbesondere ist jeder endlich–erzeugte K–Vektorraum V zu einen Vektorraum der Form K^n für genau ein n isomorph.

Beweis.
Wende Satz 3.2.1.2 auf die Basen $(e_1, \ldots, e_n)$ von K^n und $(v_1, \ldots, v_n)$ von V an. $\Phi_{\mathcal{B}}$ ist surjektiv nach Teilaussage (a), weil die Familie (v_i) ein Erzeugendensystem von V ist, und injektiv, weil (v_i) linear unabhängig ist. Da isomorphe Vektorräume nach Korollar 3.2.3 gleiche Dimension haben, aber für $n \neq m$ gilt $n = \dim_K K^n \neq \dim_K K^m = m$, folgt auch die Eindeutigkeitsaussage. □

Nebenbei haben wir auch alle endlich erzeugten Vektorräume bis auf Isomorphie klassifiziert: zwei solche Vektorräume sind genau dann isomorph, wenn sie die gleiche Dimension haben. Dies führt uns darauf, zunächst lineare Abbildungen

$$\Phi : \ K^n \to K^m$$

zu beschreiben.

Definition 3.2.5
Sei K ein Körper.

1. *Ein rechteckiges Schema der Form*

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in K$ heißt eine $m \times n$–Matrix mit Einträgen in K. Die Menge der $m \times n$ Matrizen mit Einträgen in K bezeichnen wir mit $M(m \times n, K)$.

2. *Sei $\Phi : \ K^n \to K^m$ eine lineare Abbildung. Es seien*

$$\Phi(e_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, \ \Phi(e_n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

mit $a_{ij} \in K$ die Bilder der Vektoren $(e_1, \ldots, e_n)$ der Standardbasis von K^n. Dann heißt

$$M(\Phi) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

die darstellende Matrix von Φ.

Man beachte, dass wir hier die Standardbasis als geordnete Basis auffassen, damit wir wissen, was die erste Spalte der darstellenden Matrix ist, was die zweite Spalte etc.

Aus Satz 3.2.1 folgt, das $M(\Phi)$ und Φ sich umkehrbar eindeutig entsprechen. Sei nun

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in K^n$$

beliebig. Dann rechnen wir mit $v = \sum_{i=1}^{n} v_i e_i$, also

$$\Phi(v) = \Phi\Big(\sum_{i=1}^{n} v_i e_i\Big) = \sum_{i=1}^{n} v_i \Phi(e_i) = v_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + v_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} v_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} v_j \end{pmatrix} .$$

Definition 3.2.6
Wir definieren daher für eine Matrix $A \in M(m \times n, K)$ und einen Vektor $v \in K^n$ die Multiplikation von Matrizen mit Vektoren durch

$$A \cdot v = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^{n} a_{1j} v_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} v_j \end{pmatrix}$$

Beispiel 3.2.7.
Wir setzen $K = \mathbb{R}$ und $n = m = 2$ und betrachten Drehungen um den Ursprung, vgl. Beispiel 3.1.4.2. Mit Hilfe des Produkts einer Matrix mit einem Vektor erhält man mit $\theta \in \mathbb{R}$

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta x - \sin\theta y \\ \sin\theta x + \cos\theta y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} ,$$

und somit die darstellende Matrix

$$M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Betrachtung 3.2.8.
Seien $\Phi : K^n \to K^m$ und $\Psi : K^m \to K^l$ lineare Abbildungen. Nach Bemerkung 3.1.4.5 ist dann $\Psi \circ \Phi$ wieder eine lineare Abbildung. Wir wollen deren darstellende Matrix $M(\Psi \circ \Phi)$ bestimmen. Dazu benutzen wir die Standardbasen $(e_1, \dots, e_n)$ von K^n und $(e'_1, \dots, e'_m)$ von K^m. Seien

$$M(\Psi) = (a_{ij}) , \; M(\Phi) = (b_{ij}) , \; M(\Psi \circ \Phi) = (c_{ij})$$

die darstellenden Matrizen. Dann ist

$$\begin{pmatrix} c_{1j} \\ \vdots \\ c_{lj} \end{pmatrix} = \Psi \circ \Phi(e_j) = \Psi\Big(\Phi(e_j)\Big) = \Psi \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix} = \Psi\Big(\sum_{k=1}^{m} b_{kj} e'_k\Big)$$
$$= \sum_{k=1}^{m} b_{kj} \Psi(e'_k) = \sum_{k=1}^{m} b_{kj} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{lk} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} a_{1k} b_{kj} \\ \vdots \\ \sum_{k=1}^{m} a_{lk} b_{kj} \end{pmatrix} ,$$

also

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

Definition 3.2.9

Wir definieren daher für $A \in M(l\times m, K)$ und $B \in M(m\times n, K)$ das Produkt $A\cdot B \in M(l\times n, K)$ durch die Formel

$$(A \cdot B)_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \ .$$

Die Definition stellt sicher, dass $M(\Psi \circ \Phi) = M(\Psi) \cdot M(\Phi)$ gilt. Die Komposition linearer Abbildung wird also in die in Definition 3.2.9 eingeführte Multiplikation der darstellenden Matrizen überführt.

Beispiel 3.2.10.
Wir berechnen $M(R_{\theta_1} \circ R_{\theta_2})$:

$$\begin{aligned} M(R_{\theta_1} \circ R_{\theta_2}) = M(R_{\theta_1}) \cdot M(R_{\theta_2}) &= \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 & -\cos\theta_1\sin\theta_2 - \sin\theta_1\cos\theta_2 \\ \sin\theta_1\cos\theta_1 + \cos\theta_1\sin\theta_2 & -\sin\theta_1\sin\theta_2 + \cos\theta_1\cos\theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{pmatrix} = M(R_{\theta_1+\theta_2}) \ . \end{aligned}$$

Die Drehwinkel zweier Drehungen um den Ursprung addieren sich also.

Bemerkungen 3.2.11.

1. Im allgemeinen ist die Verkettung linearer Abbildungen und somit die Matrizenmultiplikation nicht kommutativ. Zum Beispiel finden wir für

$$\begin{aligned} A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

2. Die Verkettung von Abbildungen und somit auch die Matrizenmultiplikation sind aber assoziativ.

3. Sei $\Phi: K^n \to K^m$ linear. Dann ist (vgl. Definition 3.1.5.3):

$$\begin{aligned} \operatorname{rg}(\Phi) &= \dim_K \operatorname{Im}\Phi = \dim_K \Phi(K^n) \\ &= \dim_K \operatorname{span}_K(\Phi(e_1), \ldots, \Phi(e_n)) \qquad \text{[Satz 3.2.1.2 (a)]} \ . \end{aligned}$$

Dies ist wegen des Basisauswahlsatzes 2.4.6 gleich der maximalen Anzahl linear unabhängiger Spaltenvektoren der Matrix $M(\Phi)$. Als Beispiel betrachten wir mit $K = \mathbb{R}$ die linearen Abbildungen $\Phi, \Psi: \mathbb{R}^3 \to \mathbb{R}^2$ mit

$$\begin{aligned} M(\Phi) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \qquad \operatorname{rg}(\Phi) = 1 \\ M(\Psi) = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 5 \end{pmatrix} \qquad \operatorname{rg}(\Phi) = 2 \end{aligned}$$

Wir wollen noch etwas mehr mit Matrizen rechnen. Nach Satz 3.1.9 ist die Menge $\operatorname{Hom}_K(K^n, K^m)$ der linearen Abbildungen ein K–Vektorraum. Durch Übergang zu den darstellenden Matrizen wird auch die Menge $M(m \times n, K)$ der $m \times n$ Matrizen zu einem K–Vektorraum.

Definition 3.2.12

1. *Die Summe zweier Matrizen $A, B \in M(m \times n, K)$ ist komponentenweise erklärt:*

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

Sei $\lambda \in K$; für die Skalarmultiplikation setzen wir

$$\lambda A = \lambda(a_{ij}) = (\lambda a_{ij})$$

2. *Die Transponierte einer Matrix $A \in M(m \times n, K)$ ist die durch*

$$A^t = (a_{ij}^t) = (a_{ji}) \in M(n \times m, K)$$

definierte $n \times m$ Matrix. Zum Beispiel ist:

$$\begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \end{pmatrix}^t = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 0 & 1 \end{pmatrix}$$

3. *Wir setzen*

$$E_n = M(\mathrm{id}_{K^n}) = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

und nennen E_n die Einheitsmatrix für K^n. Wir schreiben $E_n = (\delta_{ij})$ mit

$$\delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \ . \end{cases}$$

δ_{ij} heißt des Kroneckersche δ–Symbol.

Lemma 3.2.13.
Es gelten die folgenden Rechenregeln: sind $A, A' \in M(m \times n, K)$, $B, B' \in M(n \times r, K)$, $C \in M(r \times s, K)$ und $\lambda \in K$, so gilt

1. $A \cdot (B + B') = AB + A \cdot B'$ und $(A + A') \cdot B = AB + A'B$ (Distributivgesetze)
2. $A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$
3. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ (Assoziativgesetz)
4. $(A \cdot B)^t = B^t \cdot A^t$
5. $E_m \cdot A = A \cdot E_n = A$

Beweis.
1.,2. und 5. zeigt man durch einfaches Hinschreiben.
3. folgt aus dem Assoziativitätsgesetz für die Verkettung von Abbildungen
4. rechnen wir vor: ist $A = (a_{ij})$ und $B = (b_{jk})$, so ist $A \cdot B = (c_{ik})$ mit

$$c_{ik} = \sum_j a_{ij} b_{jk} \ .$$

Also ist $(AB)^t = (c'_{ki})$ mit $c'_{ki} = c_{ik} = \sum_j a_{ij} b_{jk}$. Weiter ist

$$B^t = (b'_{kj}) \quad \text{mit} \quad b'_{kj} = b_{jk}$$
$$A^t = (a'_{ji}) \quad \text{mit} \quad a'_{ji} = b_{ij}$$

Hieraus folgt

$$B^t \cdot A^t = (d_{ki}) \quad \text{mit}$$
$$d_{ki} = \sum_j b'_{kj} \cdot a'_{ji} = \sum_j b_{jk} \cdot a_{ij} = c_{ik} = c'_{ki} \; .$$

□

Korollar 3.2.14.
Die Menge $M(n \times n, K)$ der quadratischen Matrizen über einen Körper K bildet mit den Operationen $(+, \cdot)$ einen Ring. Dieser ist für $n \geq 2$ nicht–kommutativ.

Unter der Entsprechung

$$\mathrm{Hom}(K^n, K^n) \to M(n \times n, K)$$

entsprechen den Isomorphismen die folgenden Matrizen:

Definition 3.2.15
Eine Matrix $A \in M(n \times n, K)$ heißt invertierbar, wenn es eine Matrix $A' \in M(n \times n, K)$ gibt mit

$$A \cdot A' = A' \cdot A = E_n \; .$$

Korollar 3.2.16.
Die Menge

$$GL(n, K) := \{A \in M(n \times n, K) : \; A \text{ invertierbar}\}$$

mit der Multiplikation als Verknüpfung bildet eine Gruppe mit neutralem Element E_n. Sie heißt allgemeine lineare Gruppe, englisch general linear group.

Beweis.

- Mit $A, B \in GL(n, K)$ ist auch $A \cdot B \in GL(n, K)$. Denn gelte für $A', B' \in M(n \times n, K)$

  $$AA' = A'A = E_n \qquad \text{und} \quad BB' = B'B = E_n \; ,$$

  so ist wegen der Assoziativität der Matrizenmultiplikation

  $$(B'A')AB = B'(A'A)B = E_n \qquad \text{und} \quad AB(B'A') = A(BB')A' = E_n \; .$$

- Die Gruppenaxiome sind klar, das neutrale Element ist die Einheitsmatrix E_n.

□

Wir behandeln nun lineare Abbildungen zwischen beliebigen Vektorräumen:

Satz 3.2.17.
Gegeben seien K–Vektorräume

$$\begin{aligned} V &\quad \text{mit geordneter Basis } \mathcal{A} = (v_1, \ldots, v_n) \\ W &\quad \text{mit geordneter Basis } \mathcal{B} = (w_1, \ldots, w_m) \ . \end{aligned}$$

Dann gibt es zu jeder linearen Abbildung

$$\Phi : \ V \to W$$

genau eine Matrix $A = (a_{ij}) \in M(m \times n, K)$, so dass

$$\Phi(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad\qquad (*)$$

gilt. Die so erhaltene Abbildung

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{A}} : \quad & \mathrm{Hom}(V, W) \to M(m \times n, K) \\ & \Phi \qquad\qquad \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \end{aligned}$$

ist ein Isomorphismus von K–Vektorräumen. Insbesondere gilt

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{A}}(\Phi + \Psi) &= M_{\mathcal{B}}^{\mathcal{A}}(\Phi) + M_{\mathcal{B}}^{\mathcal{A}}(\Psi) \\ M_{\mathcal{B}}^{\mathcal{A}}(\lambda\Phi) &= \lambda M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \ . \end{aligned}$$

Nach Wahl von geordneten Basen von V und von W kann man also lineare Abbildungen durch Matrizen beschreiben. Man sagt, die Matrix $M_{\mathcal{B}}^{\mathcal{A}}(\Phi)$ stelle die lineare Abbildung Φ bezüglich der geordneten Basen $\mathcal{A}$, $\mathcal{B}$ von V und von W dar.

Beweis.

- Da $\mathcal{B} = (w_1, \ldots, w_m)$ eine geordnete Basis von W ist, sind die Linearkombinationen aus $(*)$ und somit die Spalten der Matrix eindeutig bestimmt.
- Gehört zur Abbildung Ψ die Matrix $B = (b_{ij})$, so rechnen wir:

$$\begin{aligned} (\Phi + \Psi)(v_j) &= \Phi(v_j) + \Psi(v_j) \\ &= \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i \\ &= \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i \end{aligned}$$

  und für $\lambda \in K$

$$(\lambda\Phi)(v_j) = \lambda \cdot \Phi(v_j) = \lambda \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} (\lambda_i a_{ij}) w_i \ .$$

  Also ist die Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ eine K–lineare Abbildung.

- Nach Satz 3.2.1.2 ist die Abbildung sogar bijektiv.

□

Korollar 3.2.18.
Mit den gleichen Bezeichnungen betrachte für jedes $i = 1 \dots n$ und $j = 1 \dots m$ die lineare Abbildung

$$\begin{aligned} & F_i^j : && V \to W \\ \text{mit} \quad & F_i^j(v_k) := \begin{cases} w_j & \text{für } k = i \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Diese $m \cdot n$ Abbildungen bilden eine Basis von $\mathrm{Hom}(V, W)$. Insbesondere gilt

$$\dim_K \mathrm{Hom}(V, W) = \dim_K V \cdot \dim_K W \ .$$

Beweis.

- Es ist $M_{\mathcal{B}}^{\mathcal{A}}(F_i^j) = E_i^j$, also gleich der Matrix, deren Einträge alle gleich Null sind, außer einem Eintrag in der i–ten Spalte und j–ten Zeile, der gleich Eins ist. Man nennt so eine Matrix auch eine Basismatrix.
- Die Familie (E_i^j) bildet eine Basis des K–Vektorraums $M(m \times n, K)$. Da $M_{\mathcal{B}}^{\mathcal{A}}$ ein Isomorphismus ist, bildet auch (F_i^j) eine Basis von $\mathrm{Hom}(V, W)$.

□

Die Frage, wie die Matrix $M_{\mathcal{B}}^{\mathcal{A}}(F)$ sich ändert, wenn man die geordneten Basen $\mathcal{A}, \mathcal{B}$ ändert, werden wir erst am Ende dieses Kapitels angehen.

Bemerkung 3.2.19.
Ist $V = W$, d.h. liegt ein Endomorphismus vor, so ist es zweckmäßig, mit nur einer geordneten Basis zu arbeiten, also $\mathcal{A} = \mathcal{B} = (v_1, v_2, \dots, v_n)$ zu wählen. Man schreibt dann $M_{\mathcal{B}} := M_{\mathcal{B}}^{\mathcal{B}}$. Der Vektorraumisomorphismus

$$M_{\mathcal{B}} : \mathrm{End}(V) \to M(n \times n, K)$$

ist dann definiert durch die Gleichungen

$$\Phi(v_j) = \sum_{i=1}^{n} a_{ij} v_i \ .$$

Die Einheitsmatrix $E_n = (\delta_{ij})$ beschreibt in jeder Basis $\mathcal{B}$ von V die identische Abbildung, $M_{\mathcal{B}}(\mathrm{id}_V) = E_n$. Die Frage, wie durch Wahl einer geeigneten Basis die darstellende Matrix eines Endomorphismus auf eine Standardform gebracht werden kann, werden wir erst später, in den Kapiteln 5 und 7, beantworten können.

3.3 Affine Unterräume und affine Abbildungen

Definition 3.3.1
Sei V ein K–Vektorraum. Sei $M \subset V$ eine Teilmenge und $v \in V$ ein Vektor. Dann bezeichnen wir mit $v + M$ die folgende Teilmenge von V:

$$v + M := \{v + m |\ m \in M\} \ .$$

Eine Teilmenge $A \subset V$ der Form $A = v + W$ mit $v \in V$ und $W \subset V$ Untervektorraum heißt affiner Unterraum von V.

Lemma 3.3.2.
Ist $A = v + W$ ein affiner Unterraum eines K–Vektorraums V, so gilt

1. $W = \{a - a' |\ a, a' \in A\}$
2. $v + W = v' + W$ genau dann, wenn $v - v' \in W$ gilt.

Beweis.

1. Für die Inklusion " $\subset$ " betrachte $w \in W$ und setze $a = v + w$ und $a' = v + 0$. Dann sind $a, a' \in A$ und $a - a' = w$. Für die umgekehrte Inklusion " $\supseteq$ " betrachte $a, a' \in A$. Schreibe $a = v + w$ und $a' = v + w'$. Dann ist
$$a - a' = w - w' \in W \ .$$
2. Wir zeigen zuerst die Implikation " $\Rightarrow$ ". Es ist $v = v + 0 \in v + W = v' + W$. Also gibt es $w \in W$ mit
$$v = v' + w \ .$$
Daher gilt $v - v' = w \in W$.
Um die umgekehrte Implikation " $\Leftarrow$ " zu sehen, sei nun $v - v' \in W$. Aus $v + w \in v + W$, folgt $v + w = v' + (v - v') + w \in v' + W$. Daher gilt $v + W \subset v' + W$. Die umgekehrte Inklusion folgt analog.

□

Insbesondere ist der Untervektorraum W durch den affinen Unterraum A eindeutig bestimmt. Er heißt der zu A gehörende Untervektorraum. Wir nennen
$$\dim A := \dim_K W$$
die Dimension des affinen Unterraums A.

Als "Fußpunkt" v eines affinen Unterraums A eignet sich aber jeder Punkt $v \in A$. Denn ist $A = v + W$, so ist $v = v + 0 \in v + W = A$. Ist nun $v' \in A$ beliebig, so ist $v - v' \in W$ und nach Lemma 3.3.2.2 gilt $v' + W = v + W = A$. Es gibt in einem affinen Unterraum keinen ausgezeichneten Punkt; in einem Untervektorraum ist dagegen immer der Nullvektor ein ausgezeichneter Vektor.

Beispiele 3.3.3.
Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$. Wie sehen die affinen Unterräume A von V aus?

- $\dim A = 0$, also $\dim_{\mathbb{R}} W = 0$, $W = \{0\}$.
$$A = v + W = \{v + 0\} = \{v\}$$
Nulldimensionale affine Unterräume nennt man auch Punkte. Nicht nur der Ursprung ist ein nulldimensionaler affiner Unterraum.

- $\dim A = 1$, also $\dim_{\mathbb{R}} W = 1$. Der Untervektorraum $W = \{\lambda w |\ \lambda \in \mathbb{R},\ w \neq 0\}$ ist eine Ursprungsgerade.
$$A = v + W = \{v + \lambda w |\ \lambda \in \mathbb{R}\} = G_{v,w}$$
ist eine Gerade, die nicht notwendigerweise den Ursprung enthält, eine sogenannte affine Gerade.

- $\dim A = 2$, also $W = 2$, also $\dim_{\mathbb{R}} W = V$. Daher

$$A = v + V = V, \quad \text{da} \quad v - 0 \in V \ .$$

Hier sind die zwei–dimensionalen affinen Unterräume gleich V. Zwei-dimensionale affine Unterräume des $\mathbb{R}^n$ nennt man affine Ebenen , $(n-1)$-dimensionale affine Unterräume des $\mathbb{R}^n$ nennt man affine Hyperebenen.

Bemerkungen 3.3.4.

1. Ein affiner Unterraum A ist genau dann ein Untervektorraum, wenn er den Nullvektor enthält, $0 \in A$.
2. Sind $A_i \subset V$, $i \in I$ affine Unterräume, so ist ihr Schnitt $\bigcap_{i\in I} A_i$ entweder leer oder ein affiner Unterraum.

Beweis.

1. Ein Untervektorraum enthält immer den Nullvektor. Enthält ein affiner Unterraum A den Nullvektor, so können wir diesen als Fußpunkt wählen und finden $A = 0 + U = U$ mit einem Untervektorraum $U \subset V$.
2. Sei $\cap_{i\in I} A_i \neq \emptyset$; dann wähle $v \in \cap_{i\in I} A_i$. Also $v \in A_i$ für alle $i \in I$. Alle A_i lassen sich schreiben als

$$A_i = v + W_i \ ,$$

wobei W_i ein Untervektorraum von V ist. Daher gilt

$$\begin{aligned} \bigcap_{i\in I} A_i &= \bigcap_{i\in I} (v + W_i) = \{v + w |\ w \in W_i \ \text{ für alle } i \in I\} \\ &= \{v + w |\ w \in \bigcap_{i\in I} W_i\} = v + \bigcap_{i\in I} W_i \ . \end{aligned}$$

Aber der Durchschnitt $\bigcap_{i\in I} W_i$ einer Familie von Untervektorräumen ist ein Untervektorraum von V, vgl. Bemerkung 2.3.7.2.

□

So wie wir für Vektorräume lineare Abbildungen als die passenden Abbildungen betrachtet haben, führen wir jetzt auch noch eine Klasse von Abbildungen von Vektorräumen ein, die sich mit affinen Unterräumen verträgt.

Definition 3.3.5
Seien V, W zwei K–Vektorräume. Eine Abbildung

$$F : \ V \to W$$

heißt affine Abbildung, falls es eine lineare Abbildung

$$\Phi : \ V \to W$$

und ein $w \in W$ gibt, so dass für alle $v \in V$ gilt

$$F(v) = \Phi(v) + w \ .$$

Die lineare Abbildung Φ und der Vektor w sind durch die affine Abbildung F eindeutig bestimmt, denn es gilt $w = F(0)$ und $\Phi(v) = F(v) - F(0)$. Lineare Abbldungen sind spezielle affine Abbildungen: man setze einfach $w = 0$.

Satz 3.3.6.
Seien V, W, Z drei K–Vektorräume und

$$F : V \to W \quad \text{und} \quad G : W \to Z$$

affine Abbildungen. Dann gilt:

1. $G \circ F$ ist affin.
2. Ist $A \subset V$ ein affiner Unterraum, so ist auch $F(A) \subset W$ ein affiner Unterraum.
3. Ist $B \subset W$ ein affiner Unterraum, so ist $F^{-1}(B)$ entweder ein affiner Unterraum oder leer.
4. Ist $F(v) = \Phi(v) + w$, wobei Φ eine lineare Abbildung ist, so gilt

$$\begin{aligned} F \text{ injektiv} &\Leftrightarrow \Phi \text{ injektiv} \\ F \text{ surjektiv} &\Leftrightarrow \Phi \text{ surjektiv} . \end{aligned}$$

Beweis.

1. Schreibe

$$\begin{aligned} F(v) = \Phi(v) + w \quad &\text{mit} \quad w \in W \text{ und } \Phi : V \to W \quad \text{linear} \\ G(w') = \Psi(w') + z \quad &\text{mit} \quad z \in Z \text{ und } \Psi : W \to Z \quad \text{linear} \end{aligned}$$

   Rechne

$$\begin{aligned} G \circ F(v) = G\Big(F(v)\Big) &= \Psi\Big(F(v)\Big) + z \\ &= \Psi\Big(\Phi(v) + w\Big) + z = \Psi \circ \Phi(v) + \Psi(w) + z . \end{aligned}$$

   Da $\Psi \circ \Phi$ nach Beispiel 3.1.4.5 linear ist und $\Psi(w) + z \in Z$ liegt, ist die Abbildung $G \circ F : V \to Z$ eine affine Abbildung.

2. Schreibe $A = v_0 + V'$ mit $v_0 \in V$ und V' Untervektorraum von V. Schreibe $F(v) = \Phi(v) + w$ wie oben.

$$\begin{aligned} F(A) &= \Phi\Big(v_0 + V'\Big) + w \\ &= \Phi(v_0) + w + \Phi(V') ; \end{aligned}$$

   dies ist ein affiner Unterraum mit Fußpunkt $\Phi(v_0) + w \in W$ und Untervektorraum $\Phi(V') \subset W$.

3. Sei $F^{-1}(B) \neq \emptyset$; wähle $v_0 \in F^{-1}(B)$. Setze $w_0 := F(v_0) = \Phi(v_0) + w \in B$. Schreibe den affinen Unterraum B als $B = w_0 + W'$ mit W' Untervektorraum zu B. Nun gilt

$$\begin{aligned} F^{-1}(B) &= \{v \in V |\ F(v) \in B\} \\ &= \{v \in V |\ \Phi(v) + w \in w_0 + W'\} \\ &= \{v \in V |\ \Phi(v) + w \in \Phi(v_0) + w + W'\} \\ &= \{v \in V |\ \Phi(v - v_0) \in W'\} \\ &= \{v \in V |\ v - v_0 \in \Phi^{-1}(W')\} \\ &= v_0 + \Phi^{-1}(W') \ , \end{aligned}$$

wobei $\Phi^{-1}(W')$ nach Satz 3.1.3.3 ein Untervektorraum von V ist.

4. Es ist $F = T \circ \Phi$ mit $T(x) = x + w$ der Translation um $w \in W$. Eine Translation T ist bijektiv mit Umkehrabbildung $T^{-1}(x) = x - w$. Daraus folgt sofort die Aussage.

□

Bemerkung 3.3.7.
Als wichtigen Spezialfall von Satz 3.3.6.3 betrachten wir im Bildbereich einen Punkt $B = \{b\}$ als affinen Unterraum und nehmen an, dass $F = \Phi$ sogar linear ist. Dann ist das Urbild $\Phi^{-1}(b)$ entweder leer oder ein affiner Unterraum. Das Urbild $\Phi^{-1}(b)$ heißt die Faser der linearen Abbildung Φ über b.

Ist $\Phi^{-1}(b) \neq \emptyset$ und $a \in \Phi^{-1}(b)$, so gilt

$$\Phi^{-1}(b) = a + \ker \Phi \ .$$

Denn sei $v \in \ker \Phi$, so ist $\Phi(a + v) = \Phi(a) + \Phi(v) = b + 0 = b$, also gilt $a + \ker \Phi \subset \Phi^{-1}(b)$. Sei nun umgekehrt $a' \in \Phi^{-1}(b)$ beliebig. Es ist $a' = a + (a' - a)$ und

$$\Phi(a' - a) = \Phi(a') - \Phi(a) = b - b = 0 \ ,$$

also gilt auch die umgekehrte Inklusion $\Phi^{-1}(b) \subset a + \ker \Phi$.

Bemerkung 3.3.8.

1. Wir haben affine *Unter*räume behandelt; man kann auch den Begriff eines affinen Raums über einem Körper K einführen. Dies ist ein Tripel (A, V, τ) bestehend aus einer nichtleeren Menge A, deren Elemente Punkte heißen, einem K-Vektorraum V, dessen Elemente Translationen heißen, und einem Gruppenhomomorphismus $\tau : V \to S(A)$ in die Gruppe der Bijektionen von A, die einfach transitiv ist in dem Sinne, dass es zu je zwei Punkten $x, y \in A$ genau ein $g \in V$ gibt mit $\tau(g)(x) = y$.

   Die durch zwei Punkte $x, y \in A$ eindeutig bestimmte Translation bezeichnen wir mit $\overrightarrow{xy} \in T$.

2. Im Gegensatz zu einem Vektorraum, der mit dem Nullvektor immer ein ausgezeichnetes Element enthält, gibt es in einem affinen Raum kein ausgezeichnetes Element, also keinen "Ursprung".

3. Seien $(A, T(A), \tau)$ und $(A', T(A'), \tau')$ affine Räume. Eine Abbildung $f : A \to A'$ heißt affine Abbildung, wenn es eine lineare Abbildung $T(f) : T(A) \to T(A')$ gibt, so dass

$$\overrightarrow{f(a_1)f(a_2)} = T(f)(\overrightarrow{a_1 a_2})$$

   gilt.

Mehr Information findet man zum Beispiel in Kapitel 1 von Gerd Fischer, Analytische Geometrie, Vieweg 2001.

3.4 Lineare Gleichungssysteme, Gauß'scher Algorithmus

Die folgenden Überlegungen hatten im im speziellen Fall des Körpers $\mathbb{R}$ der reellen Zahlen schon einmal in Abschnitt 1.2 gesehen. Wir bringen sie nun mit der Geometrie affiner Unterräume und der Struktur der Gruppe $GL(n, K)$ zusammen.

Definition 3.4.1

1. *Sei K ein Körper. Ein lineares Gleichungssystem ist ein System von Gleichungen der Form*

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

   mit $a_{ij} \in K$ und $b_i \in K$. Gesucht sind $x_1, \ldots, x_n \in K$.

2. *Gilt $b_1 = \ldots = b_m = 0$, so heißt das lineare Gleichungssystem homogen; sonst inhomogen.*

3. *Ersetzt man bei einem inhomogenen linearen Gleichungssystem alle b_i durch 0, so erhält man das zugehörige homogene lineare Gleichungssystem.*

4. *Wir nennen*

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \in M(m \times n, K)$$

   die Koeffizientenmatrix des linearen Gleichungssystems. Mit $b := (b_1 \ldots, b_m) \in K^m$ nennen wir die Matrix

$$(A, b) := \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{pmatrix} \in M(m \times (n+1), K)$$

   die erweiterte Koeffizientenmatrix des inhomogenen linearen Gleichungssystems.

5. *Die Lösungsmenge des linearen Gleichungssystems ist nun in Matrixschreibweise*

$$\mathrm{Lsg}(A, b) := \{x \in K^n | \; Ax = b\}$$

Betrachtung 3.4.2.
Wir können jetzt die erarbeitete Theorie anwenden:

- Für ein gegebenes lineares Gleichungssystem $Ax = b$ führen wir wie in Definition 3.2.5.2 die lineare Abbildung

$$\Phi : \; K^n \to K^m \qquad \text{mit darstellender Matrix} \;\; M(\Phi) = A$$

  ein. Die Lösungsmenge des linearen Gleichungssystems $Ax = b$ ist gleich dem Urbild von b unter der linearen Abbildung Φ, also der Faser von Φ über b:

$$\mathrm{Lsg}(A, b) = \Phi^{-1}(b) \;\; ,$$

  und daher nach Bemerkung 3.3.7 entweder leer oder ein affiner Unterraum des K^n.

- Ist $x_0 \in \mathrm{Lsg}(A,b)$, so folgt aus der allgemeinen Form affiner Unterräume (vergleiche Lemma 3.3.2)
$$\mathrm{Lsg}(A,b) = x_0 + \mathrm{Lsg}(A,0)\ .$$
Man erhält also die allgemeine Lösung x des inhomogenen linearen Gleichungssystems $Ax = b$ durch Addition der Lösungen h des homogenen linearen Gleichungssystems $Ah = 0$ zu einer speziellen Lösung x_0 des inhomogenen linearen Gleichungssystems $Ax_0 = b$. In der Tat ist dann $A(x_0+h) = Ax_0+Ah = b+0 = b$, also ist $x_0+h \in \mathrm{Lsg}(A,b)$. Umgekehrt folgt für eine Lösung x des inhomogenen linearen Gleichungssystems aus $Ax = b$, dass $A(x-x_0) = Ax - Ax_0 = b - b = 0$, also ist jede Lösung von der beschriebenen Form.

Wir wollen nun die Lösbarkeit eines inhomogenen Gleichungssystems untersuchen. Dazu brauchen wir den folgenden Begriff:

Definition 3.4.3
Sei $X \in M(m \times n, K)$. Die maximale Anzahl linear unabhängiger

Spalten von X heißt	*Spaltenrang* $\mathrm{rg}\,(X)$
Zeilen von X heißt	*Zeilenrang* $\widetilde{\mathrm{rg}}\,(X)$

Wir erinnern an die Definition 3.1.5.3 des Rangs einer linearen Abbildung: $\Phi : V \to W$ als der Dimension des Bildes:
$$\mathrm{rg}\ \Phi := \dim_K \mathrm{Im}\ \Phi$$
Für eine linearen Abbildung $\phi : K^n \to K^m$ zwischen Standardvektorräumen hatten wir in Bemerkung 3.2.11.3 gesehen:
$$\mathrm{Im}\ \phi = \mathrm{span}_K(\phi(e_1), \phi(e_2), \ldots \phi(e_n))$$
Rechts steht das Erzeugnis der Spalten der darstellenden Matrix $M(\phi)$. Daher ist die Dimension des Bildes von ϕ gleich der maximalen Anzahl der linear unabhängigen Spalten, also gleich dem Spaltenrang der darstellenden Matrix $M(\phi)$. Also gilt
$$\mathrm{rg}\ M(\phi) = \mathrm{rg}\ \phi\ , \qquad (*)$$
d.h. der Spaltenrang der darstellenden Matrix $M(\phi)$ ist gleich dem Rang der linearen Abbildung ϕ.

Offenbar ist
$$\mathrm{rg}\ (A,b) = \begin{cases} \mathrm{rg}\ (A), & \text{falls } b \text{ Linearkombination der Spaltenvektoren ist.} \\ \mathrm{rg}\ (A)+1, & \text{sonst.} \end{cases}$$
Im ersten Fall existieren $x_1, \ldots, x_n \in K$, so dass
$$b = \sum_{i=1}^{n} x_i \Phi(e_i) = \Phi\left(\sum_{i=1}^{n} x_i e_i\right)\ ,$$
d.h. $\begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} \in \mathrm{Lsg}(A,b)$; das inhomogene lineare Gleichungssystem $Ax = b$ hat also eine Lösung.

Ist umgekehrt $\begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} \in \mathrm{Lsg}(A,b)$, so folgt
$$b = \Phi \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i \Phi(e_i)\ ,$$

d.h. aber der Vektor b ist eine Linearkombination der Spaltenvektoren von A. Also gilt $\operatorname{rg}(A,b) = \operatorname{rg}(A)$.

Satz 3.4.4.
Sei $\Phi: K^n \to K^m$ linear und $A = M(\Phi) \in M(m \times n, K)$, $b \in K^m$. Dann gilt:

1. $\operatorname{Lsg}(A,b) \neq \emptyset \Leftrightarrow b \in \operatorname{Im}\Phi \Leftrightarrow \operatorname{rg}(A,b) = \operatorname{rg}(A)$.
2. $\operatorname{Lsg}(A,b) = \emptyset \Leftrightarrow b \notin \operatorname{Im}\Phi \Leftrightarrow \operatorname{rg}(A,b) = \operatorname{rg}(A) + 1$.
3. Ist $\operatorname{Lsg}(A,b) \neq \emptyset$ und $x_0 \in \operatorname{Lsg}(A,b)$, so ist $\operatorname{Lsg}(A,b)$ ein affiner Unterraum von K^n der Dimension $n - \operatorname{rg}(A)$, der x_0 enthält.

Beweis.
Nur die Dimension des affinen Unterraums $\operatorname{Lsg}(A,b)$ ist noch zu berechnen:

$$\dim \operatorname{Lsg}(A,b) = \dim(x_0 + \ker\Phi) = \dim_K \ker\Phi = n - \operatorname{rg}\Phi = n - \operatorname{rg} A\ ,$$

wobei im vorletzten Schritt die Dimensionsformel 3.1.7 einging. □

Beispiel 3.4.5.
Betrachte die beiden inhomogenen linearen Gleichungssysteme mit erweiterter Koeffizientenmatrix

$$\left(\begin{array}{cc|c} 2 & 4 & 0 \\ 1 & 2 & 1 \end{array}\right) \quad \text{und} \quad \left(\begin{array}{cc|c} 2 & 4 & 6 \\ 1 & 2 & 3 \end{array}\right)$$

In beiden Fällen ist $\operatorname{rg} A = 1$. Im ersten Fall ist $\operatorname{rg}(A,b) = 2$, das System hat keine Lösung. Im zweiten Fall ist $\operatorname{rg}(A,b) = 1$, das System ist lösbar und die Lösungsmenge ein affiner Unterraum der Dimension $\dim \operatorname{Lsg}(A,b) = n - \operatorname{rg} A = 2 - 1 = 1$. Das inhomogene lineare Gleichungssystem mit erweiterter Koeffizientenmatrix

$$\left(\begin{array}{cc|c} 1 & 3 & a \\ 1 & 3 & b \end{array}\right)$$

hat für alle Werte von a, b die Werte $\operatorname{rg}(A) = 2$ und somit $\operatorname{rg}(A,b) = 2 = \operatorname{rg}(A)$ und ist lösbar. Wegen $\dim \operatorname{Lsg}(A,b) = n - \operatorname{rg} A = 2 - 2 = 0$ ist die Lösung eindeutig.

Die Lösungstheorie im Falle von Koeffizientenmatrizen in Zeilenstufenform überträgt sich auf den Fall allgemeiner Körper:

Definition 3.4.6

1. *Eine Matrix $A \in M(m \times n, K)$ ist in Zeilenstufenform, falls für alle $i = 2, \ldots, m$ gilt: sind die ersten $(k-1)$ Einträge der $(i-1)$–ten Zeile gleich Null, so sind die ersten k Einträge der i–ten Zeile gleich Null, wobei $k = 1, \ldots, n$.*
2. *Eine Matrix ist in spezieller Zeilenstufenform, wenn sie in Zeilenstufenform ist und falls für alle $i = 1 \ldots m$ gilt: ist $a_{i1} = a_{i2} = \ldots = a_{i,k-1} = 0$ und $a_{ik} \neq 0$, so ist $a_{ik} = 1$.*

Wie Lemma 1.2.4 zeigt man auch im Fall allgemeiner Körper:

Lemma 3.4.7.
Sei A eine Matrix in Zeilenstufenform. Dann ist die Lösungsmenge eines linearen Gleichungssystems mit Koeffizientenmatrix A genau denn leer, wenn es einen Index $i \in \{1, \ldots, m\}$ gibt, so dass $a_{ij} = 0$ für alle j, aber $b_i \neq 0$ gilt.

Wir bringen noch einen leicht anderen Blick auf den Gaußalgorithmus mit Hilfe der allgemeinen linearen Gruppe $GL(m, K)$ aus Korollar 3.2.16:

Lemma 3.4.8.
Ist $T \in GL(m, K)$, so ist

$$\mathrm{Lsg}(A, b) = \mathrm{Lsg}(TA, Tb) \equiv \mathrm{Lsg}(T \cdot (A, b))$$

Beweis.
Wir haben eine Lösung $x \in \mathrm{Lsg}(A, b)$ genau dann, wenn $Ax = b$ gilt. Da T invertibel sein soll, gilt dies, genau dann, wenn $TAx = Tb$ gilt, was aber heißt, dass $x \in \mathrm{Lsg}(Ta, Tb)$ liegt. □

Wir müssen uns also nun einen Satz nützlicher Elemente in der allgemeinen linearen Gruppe $GL(m, K)$ verschaffen.

Lemma 3.4.9.
Die folgenden Matrizen liegen stets in $GL(n, K)$; sie heißen Elementarmatrizen.

- Die Diagonalmatrix $\Delta(1, 1, \ldots, \lambda, 1, \ldots, 1)$ mit $\lambda \in K \setminus \{0\}$

$$\begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

  denn sie hat offenbar die Diagonalmatrix $\Delta(1, 1, \ldots, \lambda^{-1}, 1, \ldots, 1)$ als Inverses. Dann ist $TAx = Tb$ das lineare Gleichungssystem, bei dem die i–te Zeile mit $\lambda \neq 0$ multipliziert wurde.

- Die Matrix $\tau(i, j)$ für $i \neq j$ mit Einträgen

$$\tau(i,j)_{kl} = \begin{cases} \delta_{kl} & \text{für } k, l \notin \{i, j\} \\ 0 & \text{für } k = l = i \text{ und } k = l = j \\ 1 & \text{sonst} \end{cases}$$

$$\begin{pmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & 0 & \ldots & 1 & & \\ & & & & \vdots & & & & \\ & & & & 1 & \ldots & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix}$$

  Wegen $\tau(i, j)\tau(i, j) = E_n$ ist $\tau(i, j)$ in $GL(n, K)$. Die Matrix $\tau(i, j)A$ unterscheidet sich von A durch die Vertauschung der i–ten und j–ten Zeile.

- Schließlich

$$\delta(i,j,\lambda) = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & \lambda & & 1 & 0 & \\ 0 & 0 & \dots & 0 & 1 & \end{pmatrix} \leftarrow j\text{–te Zeile}$$

$$\uparrow\ i\text{–te Spalte}$$

  Wegen $\delta(i,j,\lambda)\delta(i,j,-\lambda) = E_n$ ist dies in $GL(n,K)$. Die Matrix $\delta(i,j,\lambda)A$ entsteht aus A, indem das λ–fache der i–ten Zeile zur j–ten Zeile addiert wird.

Da alle Elementarmatrizen in $GL(n,K)$ liegen, haben wir wiederum folgenden Satz gezeigt:

Satz 3.4.10.
Es gibt die folgenden elementaren *Zeilen*umformungen:

1. Multiplikation einer Zeile mit $\lambda \in K \setminus \{0\}$
2. Vertauschung zweier Zeilen
3. Addition des Vielfachen einer Zeile zu einer anderen Zeile.

Entsteht das lineare Gleichungssystem $(\widetilde{A}, \widetilde{b})$ aus (A,b) durch sogenannte elementare Zeilenumformungen, so ändert sich die Lösungsmenge nicht, $\mathrm{Lsg}(\widetilde{A}, \widetilde{b}) = \mathrm{Lsg}(A,b)$.

Damit haben wir nun für beliebige Körper das aus Betrachtung 1.2.6 bekannte Rezept hergeleitet, eine Matrix in Zeilenstufenform zu bringen.

Bemerkungen 3.4.11.

1. Jede Matrix $A \in GL(n,K)$ ist ein endliches Produkt dieser Elementarmatrizen ist: durch elementare Zeilenumformungen bringen wir A zunächst auf spezielle Zeilenstufenform, was aber heißt, dass die Matrix $T_s \cdot \ldots T_1 \cdot A$ mit T_i Elementarmatrizen spezielle Zeilenstufenform hat.

   Diese Matrix ist immer noch invertibel und kann daher auf der Diagonale nur Einsen haben. Durch weitere Zeilenumformungen erreichen wir die Einheitsmatrix. Also gibt es Elementarmatrizen T_i, so dass $T_n \cdot \ldots T_1 \cdot A = E_n$ gilt. Lösen wir nach A auf, so haben wir A als Produkt von Elementarmatrizen geschrieben, $A = T_1^{-1} \ldots T_n^{-1}$. Wir sehen also einen engen Zusammenhang zwischen dem Gauß-Algorithmus und der Struktur der Gruppe $GL(n,K)$.

2. Das Inverse einer Matrix $A \in \mathrm{GL}(n,K)$ kann mit Hilfe des Gauß-Algorithmus bestimmt werden. Dazu lösen wir n inhomogene lineare Gleichungssysteme $Ax^{(i)} = e_i$ und finden $x^{(i)} = A^{-1}e_i$, so dass $x^{(i)}$ die Spalten der darstellenden Matrix von A^{-1} sind.

   Rechnung am Beispiel der Matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -3 & -2 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \end{array}\right)$$

3.5 Quotientenvektorräume, äußere direkte Summe und Produkte

Lemma 3.5.1.
Sei V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann ist die Relation auf V, die definiert ist durch $v \sim w$ für $v, w \in V$ genau dann, wenn $v-w \in U$ liegt, eine Äquivalenzrelation.

Beweis.
Die Relation ist reflexiv, $v \sim v$, denn $v - v = 0 \in U$ für alle $v \in V$. Sie ist symmetrisch, denn $v \sim w$ gilt genau dann, wenn $v - w \in U$. Dies ist aber genau dann der Fall, wenn $w - v = -(v - w) \in U$ liegt, was aber gleichbedeutend zu $w \sim v$ ist. Die Transitivität der Relation folgt, da $v \sim w$ und $w \sim z$ bedeuten, dass $v - w \in U$ und $w - z \in U$ liegen; wegen $v - z = v - w + w - z \in U$ folgt aber auch $v \sim z$. □

Satz 3.5.2.
Sei V ein K-Vektorraum und $U \subset V$ ein Untervektorraum.

1. Die Menge der Äquivalenzklassen

   $$V/U := \{[v] \mid \ v \in V\}$$

   unter der Äquivalenzrelation aus Lemma 3.5.1 wird durch die Verknüpfungen

   $$[v] + [w] := [v + w] \quad \text{und} \quad \alpha[v] := [\alpha v]$$

   zu einem K–Vektorraum.

2. Die kanonische Surjektion (vgl. Bemerkung 1.4.15.3)

   $$\text{can} : \ V \to V/U \quad \text{mit} \ \ \pi(v) = [v]$$

   ist linear. Es ist $\ker \text{can} = U$.

3. Ist V endlich-dimensional, so ist V/U endlich-dimensional mit Dimension $\dim_K V/U = \dim_K V - \dim_K U$.

Beweis.

- Wie in Betrachtung 2.1.18 ist zunächst die Wohldefiniertheit der Verknüpfungen zu zeigen: dazu wählen wir äquivalente Vektoren, $v_1 \sim v_2$ und $w_1 \sim w_2$. Es gilt dann $v_1 - v_2 \in U$ und $w_1 - w_2 \in U$. Daraus folgt

  $$(v_1 + w_1) - (v_2 + w_2) = (v_1 - v_2) + (w_1 - w_2) \in U \ ,$$

  mithin $v_1 + w_1 \sim v_2 + w_2$. Ähnlich sehen wir für die Multiplikation mit Skalaren

  $$v \sim w \Rightarrow v - w \in U \Rightarrow \alpha(v - w) \in U \Rightarrow \alpha v \sim \alpha w \ .$$

- Das Nachrechnen, dass sich die Vektorraumaxiome für V/U von V vererben, sowie der Linearität von can ist dann Routine und geht wie in Betrachtung 2.1.18. Die Surjektivität von can ist nach Bemerkung 1.4.15.3 ohnehin klar. Es ist $v \in \ker \text{can}$ genau dann, wenn $[v] = 0$, was aber äquivalent zu $v \in U$ ist.

- Aus der Dimensionsformel 3.1.7 folgt wegen $U = \ker \mathrm{can}$ und der Surjektivität von can sofort $\dim_K V - \dim_K U = \dim_K V/U$.

□

Definition 3.5.3
Sei V ein K-Vektorraum und $U \subset V$ ein Untervektorraum. Der in Satz 3.5.2 eingeführte Vektorraum V/U mit der kanonische Surjektion $\mathrm{can} : V \to V/U$ *heißt der Quotientenvektorraum von V nach U.*

Theorem 3.5.4 (Homomorphiesatz/kanonische Faktorisierung)**.**
Seien V und W K–Vektorräume und sei $\Phi : V \to W$ linear. Dann existiert ein eindeutiger Isomorphismus

$$\bar{\Phi} : V/\ker\Phi \to \mathrm{Im}\,\Phi$$

so dass gilt

$$\Phi = \imath \circ \bar{\Phi} \circ \mathrm{can} \ ;$$

wobei

$$\mathrm{can} : V \to V/\ker\Phi$$

die kanonische Surjektion und

$$\imath : \mathrm{Im}\,\Phi \to W$$

die kanonische Einbettung von $\mathrm{Im}\,\Phi$ in W ist.

Man kann also jede lineare Abbildung zerlegen in eine kanonische Projektion, einen Isomorphismus und eine Inklusion. Man schreibt dies auch als kommutierendes Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ {\scriptstyle \mathrm{can}}\downarrow & & \uparrow{\scriptstyle \imath} \\ V/\ker\phi & \underset{\exists!\bar{\Phi}}{\xrightarrow{\sim}} & \mathrm{Im}\,\Phi \end{array}$$

Beweis.

- Wenn solch eine Abbildung $\bar{\Phi}$ existiert, ist sie eindeutig bestimmt. Denn es muss dann für alle $v \in V$
$$\Phi(v) = \bar{\Phi}(\mathrm{can}(v)) = \bar{\Phi}([v])$$
gelten, was $\bar{\Phi}$ auf der Äquivalenzklasse $[v]$ eindeutig festlegt.

- Dies ist wohldefiniert: gilt $[v_1] = [v_2]$, so ist $v_1 - v_2 \in \ker(\Phi)$, woraus folgt
$$\Phi(v_1) - \Phi(v_2) = \Phi(v_1 - v_2) = 0 \ .$$

- Wegen $\mathrm{Im}\,\bar{\Phi} = \mathrm{Im}\,\Phi$ ist $\bar{\Phi}$ trivialerweise surjektiv.

- Die Linearität von $\bar{\Phi}$ folgt sofort aus der Linearität von Φ:
$$\bar{\Phi}([v] + [w]) = \bar{\Phi}([v + w]) = \Phi(v + w) = \Phi(v) + \Phi(w) = \bar{\Phi}([v]) + \bar{\Phi}([w]) \ .$$

- $\bar{\Phi}$ ist auch injektiv: denn gilt für $v \in V$

$$0 = \bar{\Phi}([v]) = \Phi(v) \ ,$$

so ist $v \in \ker \Phi$, also $[v] = 0$.

□

Betrachtung 3.5.5.
Wir untersuchen die Situation explizit durch Wahl angepasster Basen:

- Seien V, W endlich–dimensionale K-Vektorräume und

$$\Phi : \ V \to W$$

eine lineare Abbildung. Wie im Beweis von Satz 3.1.7 ergänze eine geordnete Basis $(v_1, \ldots, v_k)$ des Untervektorraums $\ker \Phi \subset V$ zu einer geordneten Basis $(v_1, \ldots, v_n)$ von V.

- Nach Satz 3.2.1 ist dann

$$(\Phi(v_{k+1}), \ldots, \Phi(v_n)))$$

ein geordnetes Erzeugendensystem von $\operatorname{Im} \Phi$. Die Familie ist auch linear unabhängig, denn Φ, eingeschränkt auf den Unterraum $\operatorname{span}_K(v_{k+1}, \ldots v_n) \subset V$ ist injektiv, da der Schnitt dieses Unterraums mit dem Kern von Φ trivial ist.

- Wir wenden die gleiche Betrachtung auch auf die kanonische Projektion

$$\text{can} : \ V \to V/\ker \Phi$$

an: es gilt ker
$can = \ker \Phi$; daher folgt, dass die Familie

$$([v_{k+1}], \ldots, [v_n])$$

eine Basis von $\operatorname{Im} \text{can} = V/\ker \Phi$ ist.

- Die Abbildung $\bar{\Phi} : V/\ker \Phi \to \operatorname{Im} \Phi$ ist dann auf der Basis $\{[v_{k+1}], \ldots, [v_n]\}$ von $V/\ker \Phi$ durch

$$\bar{\Phi}([v_i]) = \Phi(v_i)$$

definiert und liefert nach Satz 3.2.1.2 als Bijektion von Basen einen Isomorphismus von $V/\ker \Phi$ auf $\operatorname{Im} \Phi$.

- Ergänzt man die linear unabhängige Familie

$$(\Phi(v_{k+1}), \ldots, \Phi(v_n))$$

in W in irgendeiner Weise zu einer geordneten Basis

$$\mathcal{B} = (\Phi(v_{k+1}), \ldots, \Phi(v_n), w_1, \ldots, w_{m-r})$$

von W mit $m := \dim_K W$ und $r := \operatorname{rg} \Phi$ und wählt bequemerweise als geordnete Basis von V die umgeordnete Familie

$$\mathcal{A} = (v_{k+1}, \ldots, v_n, v_1, \ldots, v_k) \ ,$$

so hat man für Φ die folgende Blockmatrix als darstellende Matrix:

$$M_{\mathcal{B}}^{\mathcal{A}}(\Phi) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) \begin{array}{l} \Big\} m-r \\ \underbrace{\qquad}_{n-r=k} \end{array}$$

Durch *unabhängige* Wahl von Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W kann also die darstellende Matrix einer linearen Abbildung immer auf eine sehr einfache Form gebracht werden.

Beispiel 3.5.6.
Sei ein K-Vektorraum V die direkte Summe der Untervektorräume U, W, also $V = U \oplus W$. Betrachte die lineare Abbildung

$$\Phi : V \to W$$

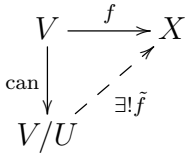
mit $\Phi(u+w) := w$ für alle $u \in U$ und $w \in W$. Dann ist $\ker \Phi = U$ und $\operatorname{Im} \Phi = W$. Nach Theorem 3.5.4 gibt es einen eindeutigen Isomorphismus

$$\bar{\Phi} : V/U \xrightarrow{\sim} W$$

mit $\bar{\Phi}([w+u]) = \Phi(u+w) = w$, also gilt $V/U \cong W$. Genauer gesagt gilt $(V/U, \text{can}) \cong (W, \Phi)$, denn wir fassen den Quotientenvektorraum als einen Vektorraum mit einer kanonischen Surjektion auf.

Satz 3.5.7 (Universelle Eigenschaft des Quotientenvektorraums)**.**
Sei K ein Körper, V ein K-Vektorraum und $U \subset V$ ein Untervektorraum und $\text{can} : V \to V/U$ die kanonische Surjektion.

1. Dann existiert für jeden K-Vektorraum X und jede lineare Abbildung $f : V \to X$ mit $f_{|U} = 0$ eine eindeutig bestimmte lineare Abbildung $\tilde{f} : V/U \to X$ mit $f = \tilde{f} \circ \text{can}$. Man sagt, dass f über V/U eindeutig faktorisiert. Als Diagram:



2. Sei Q ein K-Vektorraum und $\hat{\pi} : V \to Q$ eine lineare Abbildung mit $\hat{\pi}_{|U} = 0$, so dass für jede lineare Abbildung $f : V \to X$ mit $f_{|U} = 0$ eine eindeutig bestimmte lineare Abbildung $\hat{f} : Q \to X$ mit $f = \hat{f} \circ \hat{\pi}$ existiert. Dann gibt es einen eindeutig bestimmten Isomorphismus $\hat{f}_Q : Q \to V/U$, so dass $\text{can} = \hat{f}_Q \circ \hat{\pi}$ gilt.

Wir können die universelle Eigenschaft des Quotientenvektorraums auch so umformulieren: sei $\text{Hom}_U(V, X) := \overline{\{f : V \to X \,|\, \text{linear und } f|_U = 0\}}$. Dann ist

$$\begin{array}{rcl} \text{Hom}_K(V/U, X) & \to & \text{Hom}_U(V, X) \\ \tilde{f} & \mapsto & \tilde{f} \circ \text{can} \end{array}$$

ein Isomorphismus von Vektorräumen. Die Surjektivität ist die Aussage, dass man zu jedem $f \in \text{Hom}_U(V, X)$ ein $\tilde{f} \in \text{Hom}_K(V/U, X)$ finden kann, so dass $f = \tilde{f} \circ \text{can}$ gilt, also die Existenzaussage für $\tilde{f}$. Die Injektivität ist die Eindeutigkeitsaussage für $\tilde{f}$.

Beweis.

1. Sei $v \in V$. Für jede solche Abbildung $\tilde{f}$ muss gelten:

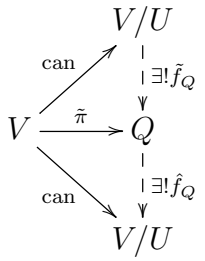
$$\tilde{f}([v]) = \tilde{f} \circ \pi(v) = f(v) \ ,$$

so dass $\tilde{f}$ eindeutig festgelegt ist. Dies ist wegen $f(v+u) = f(v) + f(u) = f(v)$ für alle $u \in U$ wohldefiniert, d.h. unabhängig vom Repräsentanten v von $[v]$. Aus der Linearität von f folgt wieder leicht die Linearität von $\tilde{f}$, vgl. hierzu auch den Beweis von Satz 3.5.2.1.

2. Wenden wir 1. auf die lineare Abbildung $\hat{\pi} : V \to Q$ an, so finden wir eine eindeutige lineare Abbildung $\tilde{f}_Q : V/U \to Q$ mit $\hat{\pi} = \tilde{f}_Q \circ \text{can}$. Zum zweiten wenden wir die für Q geforderte Eigenschaft auf die lineare Abbildung $\text{can} : V \to V/U$ an und finden $\hat{f}_Q : Q \to V/U$ mit $\text{can} = \hat{f}_Q \circ \hat{\text{can}}$. Es folgt

$$\hat{f}_Q \circ \tilde{f}_Q \circ \text{can} = \hat{f}_Q \circ \hat{\pi} = \text{can} \, .$$

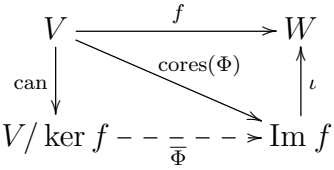
Man sieht dies besser mit Diagrammen:



Natürlich gilt aber auch $\text{id}_{V/U} \circ \text{can} = \text{can}$. Aber die universelle Eigenschaft von V/U, angewandt auf $\text{can} : V \to V/U$ selbst, sagt, dass eine solche Abbildung eindeutig ist, also gilt $\hat{f}_Q \circ f_Q = \text{id}_{V/U}$. Analog zeigt man auch $\tilde{f}_Q \circ \hat{f}_Q = \text{id}_Q$.

□

Bei einer universellen Eigenschaft charakterisieren wir ein Objekt, hier den Quotientvektorraum, durch das, was es leisten soll. Wir leiten noch einmal den Homomorphiesatz Theorem 3.5.4 im Diagramm her:



Im Rest dieses Kapitels werden wir eine weitere Begriffsbildung kennenlernen und auch durch eine universelle Eigenschaft verstehen. Sie verallgemeinert, wie $\mathbb{R}^n$ durch Betrachtung von n-Tupeln aus n-Kopien des Vektorraums $\mathbb{R}$ gebildet wurde.

Definition 3.5.8
Gegeben sei eine nicht notwendigerweise endliche Familie $(V_\lambda)_{\lambda\in\Lambda}$ von K–Vektorräumen. Wir bilden zwei neue K–Vektorräume:
das <u>Produkt</u>

$$\prod_{\lambda\in\Lambda} V_\lambda = \Big\{ (v_\lambda)_{\lambda\in\Lambda} |\ v_\lambda \in V_\lambda \Big\}$$

und die (äußere) <u>direkte Summe</u>

$$\bigoplus_{\lambda\in\Lambda} V_\lambda = \Big\{ (v_\lambda)_{\lambda\in\Lambda}, \quad \textit{nur endlich viele } v_\lambda \in V_\lambda \textit{ ungleich Null} \Big\}$$

Die K–Vektorraum–Struktur ist dabei komponentenweise erklärt.

Für endliche Familien, $|\Lambda| < \infty$, fallen die beiden Vektorräume offenbar zusammen, $\bigoplus_{\lambda\in\Lambda} V_\lambda = \prod_{\lambda\in\Lambda} V_\lambda$. Ist überdies auch die Dimension aller Vektorräume endlich, so haben direkte Summe und direktes Produkt die Dimension $\sum_{\lambda\in\Lambda} \dim V_\lambda$.

Lemma 3.5.9.
Definiere für jedes $\mu \in \Lambda$ die kanonische Injektion mit

$$\begin{array}{rccl} \imath_\mu : & V_\mu & \hookrightarrow & \bigoplus_{\lambda\in\Lambda} V_\lambda \\ & v_\mu & \mapsto & (0, 0, \ldots, v_\mu, 0, \ldots) \end{array}$$

Ist W nun ein beliebiger K–Vektorraum, so ist die folgende Abbildung ein Isomorphismus von K–Vektorräumen:

$$\begin{array}{rcl} \mathrm{Hom}_K\Big(\bigoplus_{\lambda\in\Lambda} V_\lambda, W\Big) & \tilde{\to} & \prod_{\lambda\in\Lambda} \mathrm{Hom}_K(V_\lambda, W) \qquad (*) \\ f & \mapsto & (f \circ \imath_\lambda)_{\lambda\in\Lambda} \end{array}$$

Bemerkungen 3.5.10.
Wir denken daher über die direkte Summe nicht als einen Vektorraum, sondern als einen Vektorraum mit einer Familie von Abbildungen $(\imath_\mu)_{\mu\in\Lambda}$. Wir reformulieren diesen Sachverhalt, den man die universelle Eigenschaft der (äußeren) direkten Summe nennt:

- Ist für jedes $\mu \in \Lambda$ eine lineare Abbildung

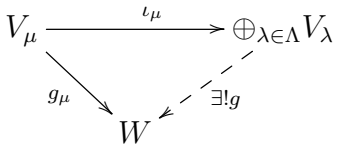
  $$g_\mu : V_\mu \to W$$

  in einen beliebigen, aber festen Vektorraum W gegeben, folgt aus der Surjektivität in $(*)$ die Existenz einer einzigen linearen Abbildung

  $$g : \oplus V_\lambda \to W \ ,$$

  so dass $g_\mu = g \circ \imath_\mu$ für alle $\mu \in \Lambda$ gilt. Diese ist wegen der Injektivität in $(*)$ eindeutig. Man kann also eine ganze Familie $(g_\mu)_{\mu\in\Lambda}$ von Abbildungen in ein und denselben K-Vektorraum W eindeutig durch eine einzige lineare Abbildung g aus der direkten Summe heraus beschreiben.

- Dies sei noch einmal in dem folgenden kommutierenden Diagramm dargestellt:



für alle $\mu \in \Lambda$

Beweis.

1. Vorüberlegung: jedes Element der direkten Summe lässt sich eindeutig schreiben in der Form

   $$v = \sum_{\lambda\in\Lambda} \imath_\lambda(v_\lambda)$$

   mit $v_\lambda \in V_\lambda$, nur endlich viele $v_\lambda \neq 0$. Deswegen ist insbesondere die Summe endlich und definiert.

2. Injektivität von $(*)$: gelte $f \circ \imath_\lambda = 0$ für alle $\lambda \in \Lambda$. Dann gilt für einen beliebigen Vektor $v \in \oplus_{\lambda\in\Lambda} V_\lambda$
$$f(v) = f\left(\sum_{\lambda\in\Lambda} \imath_\lambda(v_\lambda)\right) = \sum_{\lambda\in\Lambda} f \circ \imath_\lambda(v_\lambda) = 0\ ,$$
also ist f die Nullabbildung, der Nullvektor in $\mathrm{Hom}(\oplus_\Lambda V_\lambda, W)$.

3. Um die Surjektivität von $(*)$: für eine gegebene Familie von Abbildungen
$$g_\lambda :\ V_\lambda \to W \qquad \text{für alle } \lambda \in \Lambda$$
zu zeigen, setzen wir
$$g\left(\sum_{\lambda\in\Lambda} \imath_\lambda(v_\lambda)\right) := \sum_{\lambda\in\Lambda} g_\lambda(v_\lambda)\ .$$
Dies definiert eine K–lineare Abbildung g mit den gewünschten Eigenschaften.

□

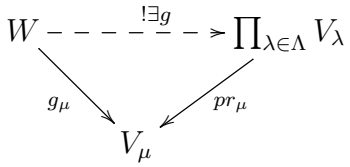
Eine analoge Aussage gilt auch für das direkte Produkt:

Lemma 3.5.11.
Im Falle des Produktes betrachten wir die kanonischen Surjektionen
$$pr_\mu :\ \prod_{\lambda\in\Lambda} V_\lambda \twoheadrightarrow V_\mu$$
auf die μ–te Komponente der Familie und erhalten einen Isomorphismus von K-Vektorräumen:
$$\begin{array}{rl} \mathrm{Hom}_K\Big(W, \prod_{\lambda\in\Lambda} V_\lambda\Big) & \xrightarrow{\sim} \prod_{\lambda\in\Lambda} \mathrm{Hom}_K(W, V_\lambda) \qquad (**) \\ f & \mapsto \Big(pr_\lambda \circ f\Big)_{\lambda\in\Lambda}\ . \end{array}$$
Man beachte, dass im Verhältnis zu Lemma 3.5.9 die Rolle der beiden Argumente von Hom_K vertauscht ist.

Bemerkungen 3.5.12.
Wir reformulieren auch diesen Sachverhalt, den man die universelle Eigenschaft des Produkts nennt:

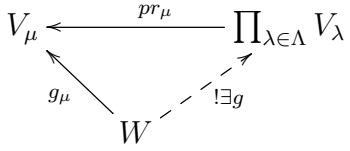
1. Ist für jedes $\mu \in \Lambda$ eine lineare Abbildung
$$g_\mu :\ W \to V_\mu$$
aus einem beliebigen, aber festen Vektorraum W heraus gegeben, folgt aus der der Surjektivität in $(**)$ die Existenz einer linearen Abbildung
$$g :\ W \to \prod_{\lambda\in\Lambda} V_\lambda$$
so dass $g_\mu = pr_\mu \circ g$ für alle $\mu \in \Lambda$ gilt. Diese ist wegen der Injektivität in $(**)$ eindeutig. Man kann also eine ganze Familie von Abbildungen aus einen K-Vektorraum W eindeutig durch eine einzige eindeutig bestimmte lineare Abbildung beschreiben.

2. Wiederum als Diagram:



für alle $\mu \in \Lambda$

3. Wenn man dies so schreibt



für alle $\mu \in \Lambda$

ist beim Vergleich zu $(**)$ klar, dass beim Produkt lediglich alle Pfeile umgedreht werden. Deshalb kann auch der Beweis leicht übertragen werden.

Wir denken daher auch über das Produkt als einen Vektorraum, zusammen mit einer Familie von Abbildungen.

Satz 3.5.13.
Sei wieder $(V_\lambda)_{\lambda\in\Lambda}$ eine Familie von K–Vektorräumen und W ein K–Vektorraum, für den eine Familie von linearen Abbildungen

$$h_\mu : V_\mu \to W ,$$

gegeben ist, die die gleiche universelle Eigenschaft wie die direkte Summe in Bemerkung 3.5.9 haben: für jeden K-Vektorraum V ist die lineare Abbildung

$$\begin{aligned} \mathrm{Hom}_K(W, V) &\to \prod_{\lambda\in\Lambda} \mathrm{Hom}_K(V_\lambda, V) \\ f &\mapsto (f \circ h_\lambda)_{\lambda\in\Lambda} \end{aligned}$$

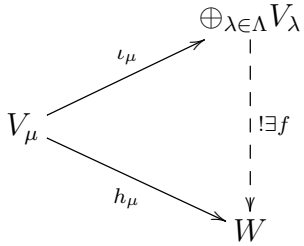
ein Isomorphismus. Dann gibt es genau eine lineare Abbildung

$$f : \bigoplus_{\lambda\in\Lambda} V_\lambda \to W$$

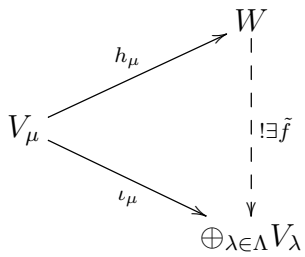
mit $f \circ \imath_\mu = h_\mu$. Diese ist ein Isomorphismus von Vektorräumen.

In Worten: die universelle Eigenschaft charakterisiert die direkte Summe bis auf eine ausgezeichnete Isomorphie.

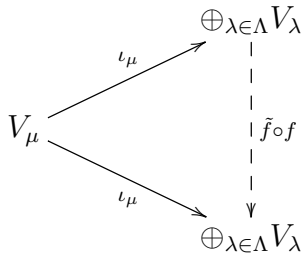
Beweis.
Wegen der universellen Eigenschaft der direkten Summe gibt eine eindeutige lineare Abbildung f, so dass für alle $\mu \in \Lambda$ das Diagramm



kommutiert. Die geforderte universelle Eigenschaft von W liefert ebenso eine eindeutige lineare Abbildung $\widetilde{f}$, so dass das Diagramm



kommutiert. Für die Verkettungen findet man für jedes $\mu \in \Lambda$ ein kommutierendes Diagramm:



Eine solche Abbildung ist aber – wiederum wegen der universellen Eigenschaft der direkten Summe – eindeutig. Da die Identität auch diese Eigenschaft hat, folgt $\widetilde{f} \circ f = \mathrm{id}_{\oplus V_\lambda}$. Ähnlich folgt $f \circ \widetilde{f} = \mathrm{id}_W$. Also ist f ein Isomorphismus von Vektorräumen. □

Korollar 3.5.14.
Sei ein K–Vektorraum Vdie innere direkte Summe zweier Untervektorräume U, W im Sinne von Definition 2.5.5, also $V = U \oplus W$. Dann ist V isomorph zur äußeren direkten Summe von U und W.

Beweis.
Wir müssen wegen Satz 3.5.13 nur zeigen, dass V zusammen mit den Inklusionen von Untervektorräumen

$$i_U : U \to V \quad \text{und} \quad i_W : W \to V$$

die universelle Eigenschaft der äußeren direkten Summe hat: sei Z ein beliebiger weiterer K-Vektorraum; seien ferner zwei beliebige lineare Abbildungen

$$g_U : U \to Z \qquad \text{und} \qquad g_W : W \to Z$$

gegeben. Wir suchen $g : V \to Z$, so dass $g \circ i_U = g_U$ und $g \circ i_W = g_W$ gilt.

Nach Lemma 2.5.4 lässt sich jedes $v \in V$ *eindeutig* als $v = u + w$ mit $u \in U$ und $w \in W$ schreiben. Daher gilt für jedes $g : V \to Z$, das diesen Forderungen genügt:

$$g(v) = g(u + w) = g \circ i_U(u) + g \circ i_W(w) = g_U(u) + g_W(w) ;$$

also ist g eindeutig. Die so definierte lineare Abbildung $g : V \to Z$ ist linear und erfüllt $g \circ i_U = g_U$ und $g \circ i_W = g_W$. Innere und äußere direkte Summe sind also nach Satz 3.5.13 kanonisch isomorphe Vektorräume. □

Beispiel 3.5.15.
Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$ und seien V_1 und V_2 zwei verschiedene eindimensionale Untervektorräume von V:

$$e_1 : V_1 \hookrightarrow V \quad \text{und} \quad e_2 : V_2 \hookrightarrow V \, .$$

Wir erhalten wegen der universellen Eigenschaft der direkten Summe eine Abbildung

$$e \, : V_1 \oplus V_2 \to V \, ,$$

deren Bild zweidimensional ist, die also surjektiv ist. Wegen $\dim_{\mathbb{R}}(V_1 \oplus V_2) = \dim_{\mathbb{R}} V_1 + \dim_{\mathbb{R}} V_2 = 2$ ist sie auch injektiv, also ein Isomorphismus. Betrachte nun die beiden Abbildungen

$$g_1 = e_1 : V_1 \hookrightarrow V \quad \text{und} \quad g_2 = -e_2 : V_2 \hookrightarrow V \, ,$$

die uns eine Abbildung

$$g \, : V_1 \oplus V_2 \to V$$

geben. Dann ist

$$g \circ e^{-1} \, : V \to V$$

ein Endomorphismus von V, der auf dem Unterraum V_1 die Identität und auf dem Unterraum V_2 gleich dem Negativen der Identität ist. Ist $\mathbb{R}^2$ zusätzlich mit dem Standardskalarprodukt versehen, und stehen die Unterräume V_1 und V_2 aufeinander senkrecht, so haben wir so eine Spiegelung an der Ursprungsgeraden V_1 definiert.

3.6 Koordinatentransformationen

Wir erinnern an Korollar 3.2.4: ist V ein n–dimensionaler K–Vektorraum, so liefert jede geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V einen Isomorphismus von K-Vektorräumen

$$\begin{array}{rccc} \Phi_{\mathcal{B}} \, : & K^n & \to & V \\ & e_i & \mapsto & v_i \, . \end{array}$$

vom Standardvektorraum K^n auf V.

Lemma 3.6.1.
Sei $\Phi : \quad V \to W$ eine lineare Abbildung und seien geordnete Basen $\mathcal{A} = (v_1, \ldots, v_n)$ von V und $\mathcal{B} = (w_1, \ldots, w_m)$ von W gegeben. Für die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}(\Phi)$ mit

$$\Phi(v_i) = \sum_{j=1}^{m} M_{\mathcal{B}}^{\mathcal{A}}(\Phi)_{ji} w_j \, .$$

(vgl. Satz 3.2.17) gilt:

$$M_{\mathcal{B}}^{\mathcal{A}}(\Phi) = M(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}}) \, ,$$

wobei M im Sinne von Definition 3.2.5 die Abbildung

$$\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}} : \; K^n \to K^m$$

bezüglich der Standardbasen der Vektorräume K^n und K^m darstellt.

Beweis.
Wir finden unter Verwendung von Definition 3.2.5 im dritten und der Linearität von $\Phi_{\mathcal{B}}$ im vierten Schritt

$$\Phi(v_i) = \Phi\Big(\Phi_{\mathcal{A}}(e_i)\Big) = \Phi_{\mathcal{B}}\Big(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}}(e_i)\Big)$$
$$\overset{3.2.5}{=} \Phi_{\mathcal{B}}\left(\sum_{j=1}^{m} M\Big(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}}\Big)_{ji} e_j\right) = \sum_{j=1}^{m} M\Big(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}}\Big)_{ji} w_j \ .$$

□

Satz 3.6.2.
Sei V ein n–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{A}$, W ein m–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{B}$ und Z ein k–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{C}$. Dann gilt für alle K–linearen Abbildungen

$$\Phi : \ V \to W \quad \text{und } \Psi : \ W \to Z$$

die Gleichung

$$M_{\mathcal{C}}^{\mathcal{A}}(\Psi \circ \Phi) = M_{\mathcal{C}}^{\mathcal{B}}(\Psi) \cdot M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \ ,$$

wobei "$\cdot$" für die Matrizenmultiplikation steht.

Beweis.

$$M_{\mathcal{C}}^{\mathcal{A}}(\Psi \circ \Phi) \overset{3.6.1}{=} M(\Phi_{\mathcal{C}}^{-1} \circ \Psi \circ \Phi \circ \Phi_{\mathcal{A}})$$
$$= M\left(\Big(\Phi_{\mathcal{C}}^{-1} \circ \Psi \circ \Phi_{\mathcal{B}}\Big) \circ \Big(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}}\Big)\right)$$
$$\overset{3.2.9}{=} M\Big(\Phi_{\mathcal{C}}^{-1} \circ \Psi \circ \Phi_{\mathcal{B}}\Big) \cdot M\Big(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}}\Big) \overset{3.6.1}{=} M_{\mathcal{C}}^{\mathcal{B}}(\Psi) \cdot M_{\mathcal{B}}^{\mathcal{A}}(\Phi)$$

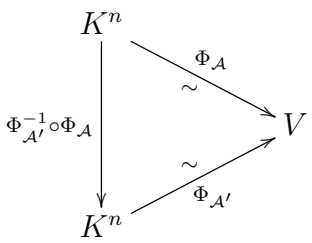
Hierbei haben wir im dritten Schritt die Einsicht aus Definition 3.2.9 verwendet, dass M die Komposition von Abbildungen in eine Matrizenmultiplikation überführt. □

Wir wollen uns jetzt der Frage zuwenden, wie die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}$ sich bei Änderungen der Basen $\mathcal{A}$, $\mathcal{B}$ ändert.

Definition 3.6.3
Seien $\mathcal{A}$ und $\mathcal{A}'$ zwei geordnete Basen von V. Dann heißt die quadratische Matrix

$$T_{\mathcal{A}'}^{\mathcal{A}} := M\Big(\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}}\Big) \in GL(n, K)$$

Transformationsmatrix des Basiswechsels von $\mathcal{A}$ nach $\mathcal{A}'$.



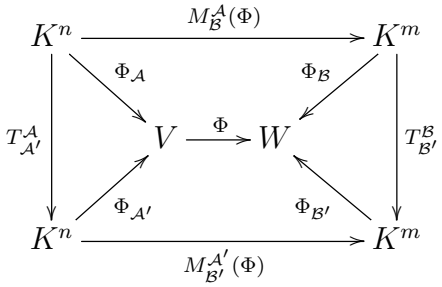
Bemerkungen 3.6.4.

1. Es ist
$$T^{\mathcal{A}}_{\mathcal{A}'} \overset{\text{def}}{=} M\Big(\Phi^{-1}_{\mathcal{A}'} \circ \Phi_{\mathcal{A}}\Big) = M\Big(\Phi^{-1}_{\mathcal{A}'} \circ \mathrm{id}_V \circ \Phi_{\mathcal{A}}\Big) \overset{3.6.1}{=} M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V) \ .$$
Schreibt man also mit geordneten Basen $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{A}' = (v'_1, \ldots, v'_n)$
$$v_j = \mathrm{id}(v_j) = \sum_{i=1}^{n} c_{ij} v'_i \ ,$$
so ist die Transformationsmatrix
$$T^{\mathcal{A}}_{\mathcal{A}'} = \begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \ldots & c_{nn} \end{pmatrix} \ .$$

2. Es gilt $T^{\mathcal{A}}_{\mathcal{A}'} \cdot T^{\mathcal{A}'}_{\mathcal{A}} = E_n$ und $T^{\mathcal{A}'}_{\mathcal{A}} \cdot T^{\mathcal{A}}_{\mathcal{A}'} = E_n$, denn zum Beispiel gilt
$$T^{\mathcal{A}}_{\mathcal{A}'} \cdot T^{\mathcal{A}'}_{\mathcal{A}} = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V) \cdot M^{\mathcal{A}'}_{\mathcal{A}}(\mathrm{id}_V) = M^{\mathcal{A}'}_{\mathcal{A}'}(\mathrm{id}_V \circ \mathrm{id}_V) = M^{\mathcal{A}'}_{\mathcal{A}'}(\mathrm{id}_V) = E_n \ .$$
Also sind Transformationsmatrizen invertierbar. Die Transformationsmatrizen $T^{\mathcal{A}}_{\mathcal{A}'}$ und $T^{\mathcal{A}'}_{\mathcal{A}}$ sind zueinander invers.

Satz 3.6.5 (Transformationsformel)**.**

1. Sei V ein endlich–dimensionaler K–Vektorraum mit geordneten Basen $\mathcal{A}$ und $\mathcal{A}'$ und W ein endlich–dimensionaler K–Vektorraum mit geordneten Basen $\mathcal{B}$ und $\mathcal{B}'$. Sei
$$\Phi: \quad V \to W$$
linear. Dann gilt
$$M^{\mathcal{A}'}_{\mathcal{B}'}(\Phi) = T^{\mathcal{B}}_{\mathcal{B}'} \cdot M^{\mathcal{A}}_{\mathcal{B}}(\Phi) \cdot \Big(T^{\mathcal{A}}_{\mathcal{A}'}\Big)^{-1} \ . \qquad (*)$$
Es kommutiert also das folgende Diagramm:



wobei hier Abbildungen zwischen Vektorräumen der Form K^n mit ihren darstellenden Matrizen identifiziert werden.

2. Speziell für Endomorphismen $\Phi: \ V \to V$ und zwei geordnete Basen $\mathcal{B}'$ und $\mathcal{B}$ von V ergibt sich
$$M_{\mathcal{B}'}(\Phi) = T^{\mathcal{B}}_{\mathcal{B}'} \cdot M_{\mathcal{B}}(\Phi) \cdot \Big(T^{\mathcal{B}}_{\mathcal{B}'}\Big)^{-1} \ . \qquad (**)$$

Beweis.

1. Wir rechnen:

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\Phi) = M_{\mathcal{B}'}^{\mathcal{A}'}(\mathrm{id}_W \circ \Phi \circ \mathrm{id}_V) \overset{3.6.2}{=} M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_W) \cdot M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \cdot M_{\mathcal{A}}^{\mathcal{A}'}(\mathrm{id}_V) \overset{3.6.4.1}{=} T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \cdot T_{\mathcal{A}}^{\mathcal{A}'} \ .$$

2. ist als Spezialfall klar.

□

Beispiel 3.6.6.

Sei $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Achse $\mathbb{R}\begin{pmatrix}1\\1\end{pmatrix}$, der Winkelhalbierenden des ersten und dritten Quadranten. Setze $b_1 = \begin{pmatrix}1\\1\end{pmatrix}$ und $b_2 = \begin{pmatrix}-1\\1\end{pmatrix}$. Dann ist $\mathcal{B} = (b_1, b_2)$ eine geordnete Basis von $\mathbb{R}^2$. Wegen

$$\Phi(b_1) = b_1 \qquad \Phi(b_2) = -b_2$$

wird der Endomorphismus Φ in der Basis $\mathcal{B}$ durch die Matrix

$$M_{\mathcal{B}}(\Phi) = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \ .$$

dargestellt. Wir wollen Φ in der geordneten Standardbasis $\mathcal{A} = (e_1, e_2)$ des $\mathbb{R}^2$ ausdrücken. Aus

$$b_1 = e_1 + e_2 \qquad b_2 = -e_1 + e_2$$

folgt

$$T_{\mathcal{A}}^{\mathcal{B}} = \begin{pmatrix}1 & -1\\1 & 1\end{pmatrix} \ .$$

und, zum Beispiel mit Hilfe des Gauß-Algorithmus, vgl. Bemerkung 3.4.11.2

$$\left(T_{\mathcal{A}}^{\mathcal{B}}\right)^{-1} = \frac{1}{2}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix} \ .$$

Es folgt

$$M_{\mathcal{A}}(\Phi) = T_{\mathcal{A}}^{\mathcal{B}} M_{\mathcal{B}}(\Phi)\left(T_{\mathcal{A}}^{\mathcal{B}}\right)^{-1} = \begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\cdot\frac{1}{2}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \ .$$

Dies entspricht der Tatsache, dass die Spiegelung an der Winkelhalbierenden des ersten und dritten Quadranten die beiden Vektoren der Standardbasis des $\mathbb{R}^2$ vertauscht.

Die Form von Gleichung $(*)$ beziehungsweise von Gleichung $(**)$ gibt Anlass zu der folgenden Definition:

Definition 3.6.7

1. *Zwei (nicht unbedingt quadratische) Matrizen $X, Y \in M(m \times n, K)$ heißen äquivalent, falls es quadratische Matrizen $S \in GL(m, K)$ und $T \in GL(n, K)$ gibt, so dass*

$$Y = SXT^{-1} \qquad \textit{gilt} \ .$$

2. *Zwei* quadratische *Matrizen $X, Y \in M(m \times m, K)$ heißen ähnlich, falls es* eine *Matrix $T \in GL(m, K)$ gibt, so dass*

$$Y = TXT^{-1} \qquad \textit{gilt} \ .$$

Bemerkungen 3.6.8.

1. Ähnliche Matrizen sind offenbar äquivalent: setze $S = T$. Die Umkehrung gilt nicht: betrachte $X = E_m$.

   - Sei $S \in GL(m, K)$ beliebig und setze $T := E_m$; dann gilt

     $$SXT^{-1} = SE_mE_m = S \ ,$$

     also sind die Matrizen S und E_m äquivalent: alle invertierbaren $m \times m$ Matrizen sind äquivalent zu E_m.
   - Die invertierbare Matrix S und E_m sind aber für $S \neq E_m$ nicht ähnlich: für alle $T \in GL(m, K)$ ist

     $$TXT^{-1} = TE_mT^{-1} = TT^{-1} = E_m \ ,$$

     also ist nur E_m zur Einheitsmatrix E_m ähnlich.

2. Äquivalenz und Ähnlichkeit sind Äquivalenzrelationen. Wir zeigen dies am Beispiel der Äquivalenz:

   - Reflexivität: setze $S = E_m$ und $T = E_n$.
   - Symmetrie: aus $X = SYT^{-1}$ folgt $Y = S^{-1}XT$.
   - Transitivität: $Y = S_1XT_1^{-1}$ und $Z = S_2YT_2^{-1}$ impliziert

     $$Z = S_2(S_1XT_1^{-1})T_2^{-1} = (S_2S_1)X(T_2T_1)^{-1} \ .$$

Lemma 3.6.9.

1. Zwei $m \times n$ Matrizen X, Y sind genau dann äquivalent, wenn sie dieselbe lineare Abbildung bezüglich verschiedener geordneter Basen beschreiben.

2. Zwei quadratische Matrizen sind genau dann ähnlich, wenn sie den gleichen Endomorphismus bezüglich verschiedener Basen beschreiben.

3. Jede Matrix X ist äquivalent zu einer Matrix der Form

   $$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \ ,$$

Das heißt ausführlicher: Es gibt einen n-dimensionalen K-Vektorraum V mit zwei geordneten Basen $\mathcal{A}, \mathcal{A}'$ und einen m-dimensionalen K-Vektorraum W mit zwei geordneten Basen $\mathcal{B}, \mathcal{B}'$ und eine lineare Abbildung $\Phi : V \to W$, so dass gilt

$$X = M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \quad \text{und} \quad Y = M_{\mathcal{B}'}^{\mathcal{A}'}(\Phi) \ .$$

Beweis.
Wir zeigen die Aussagen nur für die Relation "Äquivalenz".
"$\Leftarrow$" folgt sofort aus der Transformationsformel von Satz 3.6.5.
"$\Rightarrow$" Seien X, Y äquivalente $m \times n$ Matrizen. Dann gibt es invertiebare Matrizen S, T mit

$$Y = SXT^{-1} \ .$$

Sei $\mathcal{A} = (v_1, \ldots, v_n)$ eine beliebige geordnete Basis von V und $\mathcal{B} = (w_1, \ldots, w_m)$ eine beliebige geordnete Basis von W. Definiere mit Hilfe von Satz 3.2.1 eine lineare Abbildung $\Phi : V \to W$ durch ihre Werte auf der geordneten Basis $\mathcal{A}$:

$$\Phi(v_i) = \sum_{j=1}^{m} X_{ji} w_j \ .$$

Da T invertibel ist, gibt es eine inverse Matrix T^{-1}. Mit Hilfe von deren Matrixelementen definieren wir die Vektoren

$$v'_i := \sum_j T^{-1}_{ji} v_j \ .$$

Da ein Automorphismus eine Basis auf eine Basis abbildet, ist auch die Familie $\mathcal{A}' := (v'_1, \ldots, v'_n)$ eine geordnete Basis von V. Es gilt dann

$$v_j = \sum_{i=1}^{n} T_{ij} v'_i \ .$$

Analog finde eine geordnete Basis $\mathcal{B}' = (w'_1, \ldots, w'_m)$ von W mit

$$w_j = \sum_{i=1}^{m} S_{ij} w'_i$$

und rechne

$$\begin{aligned} \Phi(v'_i) &= \Phi(\textstyle\sum_j T^{-1}_{ji} v_j) = \sum_j T^{-1}_{ji} \Phi(v_j) \\ &= \textstyle\sum_{j,k} X_{kj} T^{-1}_{ji} w_k = \sum_{jkl} S_{lk} X_{kj} T^{-1}_{ji} w'_l = \sum_l Y_{li} w'_l \end{aligned}$$

Die dritte Aussage folgt aus Betrachtung 3.5.5, denn denn jede lineare Abbildung hat für eine geeignete Wahl von Basen eine darstellende Matrix dieser Form. □

Wir müssen diese Zahl r besser verstehen.

Definition 3.6.10
Sei $X \in M(m \times n, K)$. Die maximale Anzahl linear unabhängiger

$$\begin{aligned} \textit{Spalten von } X \textit{ heißt} \quad & \underline{\textit{Spaltenrang}}\ \mathrm{rg}\,(X) \\ \textit{Zeilen von } X \textit{ heißt} \quad & \underline{\textit{Zeilenrang}}\ \widetilde{\mathrm{rg}}\,(X) \end{aligned}$$

Wir erinnern an die Definition 3.1.5.3 des Rangs einer linearen Abbildung: $\Phi : V \to W$ als der Dimension des Bildes:

$$\mathrm{rg}\ \Phi := \dim_K \mathrm{Im}\ \Phi$$

Für eine linearen Abbildung $\phi : K^n \to K^m$ zwischen Standardvektorräumen hatten wir in Bemerkung 3.2.11.3 gesehen:

$$\mathrm{Im}\ \phi = \mathrm{span}_K(\phi(e_1), \phi(e_2), \ldots \phi(e_n))$$

Rechts steht das Erzeugnis der Spalten der darstellenden Matrix $M(\phi)$. Daher ist die Dimension des Bildes von ϕ gleich der maximalen Anzahl der linear unabhängigen Spalten, also gleich dem Spaltenrang der darstellenden Matrix $M(\phi)$. Also gilt

$$\mathrm{rg}\ M(\phi) = \mathrm{rg}\ \phi \ , \qquad (*)$$

d.h. der Spaltenrang der darstellenden Matrix $M(\phi)$ ist gleich dem Rang der linearen Abbildung ϕ.

Satz 3.6.11.
Seien V, W endlich–dimensionale K–Vektorräume und sei

$$\Phi : \ V \to W$$

linear. Sei $\mathcal{A}$ eine geordnete Basis von V und $\mathcal{B}$ eine geordnete Basis von W. Dann gilt

$$\text{rg } \ M_{\mathcal{B}}^{\mathcal{A}}(\Phi) = \text{rg } \ (\Phi) \ .$$

Hierbei steht links der Spaltenrang einer Matrix und rechts die Dimension des Bildes der linearen Abbildung Φ. Insbesondere ist der Spaltenrang der darstellenden Matrix $M_{\mathcal{B}}^{\mathcal{A}}(\Phi)$ von der Wahl der geordneten Basen $\mathcal{A}, \mathcal{B}$ unabhängig.

Beweis.
Mit den durch die geordneten Basen gelieferten Isomorphismen

$$\Phi_{\mathcal{A}} : \ K^n \xrightarrow{\sim} V \quad \text{ und } \quad \Phi_{\mathcal{B}} : \ K^m \xrightarrow{\sim} W$$

gilt

$$\begin{aligned} \text{rg } \left(M_{\mathcal{B}}^{\mathcal{A}}/\Phi \right) &\overset{3.6.2}{=} \text{rg } \left(M(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}} \right) \overset{(*)}{=} \text{rg } \left(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}} \right) \\ &\overset{\text{def3.1.5.3}}{=} \dim_K \text{Im } \left(\Phi_{\mathcal{B}}^{-1} \circ \Phi \circ \Phi_{\mathcal{A}} \right) = \dim_K \text{Im } \Phi \overset{\text{def3.1.5.3}}{=} \text{rg } (\Phi) \ , \end{aligned}$$

da $\Phi_{\mathcal{A}}$ und $\Phi_{\mathcal{B}}$ Isomorphismen sind.

□

Korollar 3.6.12.
Äquivalente Matrizen haben gleichen Spaltenrang. Insbesondere ist jede Matrix äquivalent zu einer blockdiagonalen Matrix der Form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \ ,$$

mit einem eindeutig bestimmten $r := \text{rg } A$.

Beweis.
Zwei äquivalente Matrizen $X, Y \in M(m \times n, K)$ können nach Lemma 3.6.9 als darstellende Matrizen ein und der derselben linearen Abbildung $\Phi : V \to W$ aufgefasst werden. Nach Satz 3.6.11 gilt $\text{rg } X = \text{rg } \Phi = \text{rg } Y$. (Hier ist es sehr nützlich, die Abbildung Φ als basisunabhängiges Objekt zur Verfügung zu haben - das hätte man nicht, wenn man versuchen würde, die lineare Algebra nur mit Matrizen zu entwickeln.)

Die angegebene blockdiagonale Matrix hat offensichtlich Spaltenrang r, daher ist r durch den Rang der Matrix eindeutig festgelegt. □

Die Äquivalenzklassen von $m \times n$ Matrizen sind also einfach zu beschreiben: die einzige Invariante einer Äquivalenzklasse linearer Abbildung ist ihr Rang. Ähnlichkeitsklassen werden wir erst vollständig in Kapitel 7 dieser Vorlesung beschreiben können.

Lemma 3.6.13.
Sei $X \in M(m \times n, K)$ und $S \in GL(m, K)$. Dann gilt

1. $(S^{-1})^t = (S^t)^{-1}$
2. $\operatorname{rg}(X) = \widetilde{\operatorname{rg}}(X^t)$ und $\widetilde{\operatorname{rg}}(X) = \operatorname{rg}(X^t)$.

Beweis.
2. ist offensichtlich, da die Transposition Zeilen und Spalten vertauscht.
1.: Aus $(S^{-1})^t \cdot S^t = (S \cdot S^{-1})^t = E_m^t = E_m$ folgt die Behauptung wegen der Eindeutigkeit des inversen Elements in einer Gruppe. □

Lemma 3.6.14.
Äquivalente Matrizen haben auch gleichen Zeilenrang. In Formeln: sind $X, Y \in M(m \times n, K)$ äquivalent, so ist

$$\widetilde{\operatorname{rg}}(X) = \widetilde{\operatorname{rg}}(Y) .$$

Beweis.
Wir wissen also, dass es $S \in GL(m, K)$ und $T \in GL(n, K)$ gibt, so dass

$$Y = SXT^{-1}$$

gilt. Die Transposition dieser Matrixgleichung liefert

$$Y^t = (SXT^{-1})^t = (T^{-1})^t X^t S^t = (T^t)^{-1} X^t S^t$$

nach Lemma 3.6.13.1. Somit sind auch die transponierten Matrizen X^t und Y^t äquivalent. Wegen Lemma 3.6.13.2 und Korollar 3.6.12 erhält man die Gleichungen

$$\widetilde{\operatorname{rg}}(X) \overset{3.6.13.2}{=} \operatorname{rg}(X^t) \overset{3.6.12}{=} \operatorname{rg}(Y^t) \overset{3.6.13.2}{=} \widetilde{\operatorname{rg}}(Y) .$$

□

Satz 3.6.15.
Zeilenrang und Spaltenrang einer Matrix sind gleich: für $X \in M(m \times n, K)$ gilt

$$\operatorname{rg}(X) = \widetilde{\operatorname{rg}}\ X .$$

Beweis.
Hat $X \in M(m \times n, K)$ Rang r, so ist X nach Bemerkung 3.6.8.3 äquivalent zu einer Matrix der Form $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Daher gilt

$$\operatorname{rg}(X) \underset{(3.6.12)}{=} \operatorname{rg} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r \quad \text{und} \quad \widetilde{\operatorname{rg}}(X) \underset{(3.6.14)}{=} \widetilde{\operatorname{rg}} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r$$

Hieraus folgt $\operatorname{rg}(X) = r = \widetilde{\operatorname{rg}}(X)$. □

Beispiel 3.6.16.

$$\operatorname{rg} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{pmatrix} = \widetilde{\operatorname{rg}} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{pmatrix} = 1 .$$

3.7 Kodierungstheorie

Betrachtung 3.7.1.

Ziele: Bei der Übertragung von Daten treten typischerweise Fehler auf. Dies führt zu den beiden Zielen der Fehlererkennung und Fehlerkorrektur.

Ansatz: Wir gehen davon aus, dass die Daten in einem Binärkode vorliegen, d.h. als eine Folge der Symbole 0 oder 1. Ein Datensatz fester Länge n ist also ein Vektor im $\mathbb{F}_2$-Vektorraum $V = (\mathbb{F}_2)^n$. Dies ist ein Vektorraum mit einer ausgezeichneten Basis, der Standardbasis.

Definition 3.7.2

Sei $K = \mathbb{F}_2$ und $V = K^n$. Die Abbildung

$$d_H : V \times V \to \mathbb{N}$$
$$d_H(v, w) := |\{j \in \{1, \ldots, n\} |\, v_j \neq w_j\}|$$

heißt Hamming-Abstand. Sie gibt die Zahl der Komponenten an, in der sich die beiden Argumente v und w unterscheiden.

Lemma 3.7.3.

Der Hamming-Abstand hat für alle $u, v, w \in V$ die folgenden Eigenschaften:

1. $d_H(v, w) \geq 0$ und $d_H(v, w) = 0$ genau für $v = w$
2. $d_H(v, w) = d_H(w, v)$ (Symmetrie)
3. $d_H(u, w) \leq d_H(u, v) + d_H(v, w)$ (Dreiecksungleichung)
4. $d_H(v, w) = d_H(v + u, w + u)$ (Translationsinvarianz)

Beweis.

1,2 und 4 sind trivial. Für 3. beachten wir: nur für $u_j \neq w_j$ trägt die j–te Komponente den Wert 1 zum Hamming-Abstand $d(u, v)$ bei. Dann ist aber entweder $v_j \neq u_j$ oder $v_j \neq w_j$. □

Definition 3.7.4

Sei $\lambda \in \mathbb{N}$. Eine Teilmenge $C \subset (\mathbb{F}_2)^n$ heißt λ–fehlerkorrigierender Kode, falls für alle $u, v \in C$, $u \neq v$ gilt

$$d_H(u, v) \geq 2\lambda + 1$$

Zum Beispiel ist für $n = 3$ die zweielementige Teilmenge von K^3

$$C = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

wegen $d_H\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = 3$ ein 1–fehlerkorrigierender Kode.

Die Benennung erklärt sich aus dem folgenden Lemma:

Lemma 3.7.5.
Sei $C \subset V$ ein λ–fehlerkorrigierender Kode. Dann gibt es zu jedem $v \in V$ *höchstens* ein $w \in C$ mit $d_H(v, w) \leq \lambda$.

Beweis.
Sei $v \in V$ gegeben und seien $w_1, w_2 \in C$ mit $d_H(v, w_i) \leq \lambda$. Dann gilt wegen der Dreiecksungleichung 3.7.3.3

$$d_H(w_1, w_2) \leq d_H(w_1, v) + d_H(v, w_2) \leq 2\lambda \; .$$

Da der Kode C als λ–fehlerkorrigierend vorausgesetzt wurde, folgt $w_1 = w_2$. □

Betrachtung 3.7.6.

- Idee: Verwende Elemente eines λ–fehlerkorrigierenden Kodes $C \subset (\mathbb{F}_2)^n$ als Sendedaten. Treten bei der Übermittlung des Elements λ oder weniger Fehler auf, so kann die Nachricht (nämlich das Element aus C) aus dem empfangenen Datum (nämlich ein Element in V) eindeutig rekonstruiert werden.
- Es treten aber die folgenden Probleme auf:
  - Große Kodes C brauchen viel Speicherplatz!
  - Die Dekodierung, die darin besteht, ein $w \in C$ wie in Lemma 3.7.5 zu finden, erfordert viele Vergleiche der empfangenen Nachricht mit Elementen in C.

Definition 3.7.7
Ein λ–korrigierender Kode $C \subset V$ heißt linear, falls C ein Untervektorraum von V ist.

Lineare Kodes bieten einen Vorteil beim Speicherplatz: ist $\dim_{\mathbb{F}_2} C = k$, so hat eine Basis k Elemente, der Untervektorraum C selbst aber 2^k Elemente!
Nun zur Fehlerkorrektur: Sei nun $C \subset (\mathbb{F}_2)^n$ ein linearer λ–korrigierender Kode der Dimension k. Wähle einen Epimorphismus:

$$\Phi : \; (\mathbb{F}_2)^n \to (\mathbb{F}_2)^{n-k}$$

mit $\ker \Phi = C$. Solche Epimorphismen existieren: der Basisergänzungssatz 2.4.19 erlaubt es uns, eine Basis $b_1, \ldots, b_k$ von C zu einer Basis $b_1, \ldots, b_n$ von $(\mathbb{F}_2)^n$ zu ergänzen. Setze

$$\Phi\left(\sum_{j=1}^{n} \alpha_j b_j\right) := \begin{pmatrix} \alpha_{k+1} \\ \vdots \\ \alpha_n \end{pmatrix} \; .$$

Die darstellende Matrix $M(\Phi) \in M((n-k) \times n, \mathbb{F}_2)$ heißt eine Kontrollmatrix des Kodes. Ein Element $y \in \mathbb{F}_2^{n-k}$ heißt zulässig, wenn es ein $x \in \Phi^{-1}(y)$ gibt mit $d_H(x, 0) \leq \lambda$.

Wir überlegen uns, dass dieses $x \in (\mathbb{F}_2)^n$ für ein gegebenes zulässiges $y \in \mathbb{F}_2^{n-k}$ eindeutig ist:
Seien $x, x' \in \Phi^{-1}(y)$ mit $d_H(x, 0) \leq \lambda$ und $d_H(x', 0) \leq \lambda$. Dann ist $x - x' \in \ker \Phi = C$ und

$$d_H(x - x', 0) = d_H(x, x') \leq d_H(x, 0) + d_H(x', 0) \leq 2\lambda \; .$$

Da der Kode C aber λ-fehlerkorrigierend sein soll, folgt $x - x' = 0$.

Die Dekodierung geschieht nun folgendermaßen: der Empfänger speichert eine Liste der zulässigen Elemente $y \in (\mathbb{F}_2)^{n-k}$ mit den zugehörigen eindeutig bestimmten $x_y \in \Phi^{-1}(y) \subset (\mathbb{F}_2)^n$ mit $d_H(x_y, 0) \leq \lambda$. Für jede empfangene Nachricht $v \in (\mathbb{F}_2)^n$ berechnet der Empfänger mit Hilfe der Kontrollmatrix das Element

$$y = \Phi(v) \in (\mathbb{F}_2)^{n-k} .$$

Ist y nicht zulässig, so sind so viele Fehler bei der Übertragung aufgetreten, dass eine Korrektur nicht möglich ist. Ist y dagegen zulässig, so ist $w := v - x_y$ die ursprüngliche Nachricht. Um dies zu sehen, berechnen wir

$$\Phi(w) = \Phi(v - x_y) = \Phi(v) - \Phi(x_y) = y - y = 0 ;$$

wegen $\ker \Phi = C$ liegt also $w \in C$ und stellt wirklich eine mögliche Nachricht dar. Da gilt

$$d_H(w, v) = d_H(w - v, 0) = d_H(x_y, 0) \leq \lambda ,$$

ist nach Lemma 3.7.5 w eindeutig und somit die gesendete Nachricht.

3.8 Kurzzusammenfassung der Kapitel 2 und 3

Wir wollen kurz den bisher entwickelten Stoff zusammenfassen.

1. Ausgehend vom Begriff der Gruppe haben wir Ringe und als spezielle Ringe Körper eingeführt. Wichtige Beispiele für Ringe, die nicht Körper sind, sind die ganzen Zahlen $\mathbb{Z}$, die Restklassenringe $\mathbb{Z}/n\mathbb{Z}$ für n nicht prim, und die Polynome über einem Körper. Wichtige Beispiele für Körper sind die Körper der rationalen Zahlen $\mathbb{Q}$, der reellen Zahlen $\mathbb{R}$, der komplexen Zahlen $\mathbb{C}$ und der Körper $\mathbb{F}_p$, der genau p Elemente hat mit p prim. Endliche Körper haben p^n Elemente, mit p prim und $n \in \mathbb{N}$.

2. Für einen gegebenen Körper K haben wir den Begriff des K-Vektorraums kennengelernt. Weitere wichtige Begriffe sind:

   - Untervektorraum, affiner Unterraum.
   - Innere Summe von Untervektorräumen: es gilt $W_1 + W_2 = \mathrm{span}_K(W_1 \cup W_2)$, vergleiche Satz 2.5.2; die Dimension ist $\dim_K(W_1 + W_2) = \dim_K(W_1) + \dim_K(W_2) - \dim_K(W_1 \cap W_2)$. Die innere direkte Summe ist der Spezialfall einer inneren Summe mit $W_1 \cap W_2 = \{0\}$. Einen Vektor in einer inneren direkten Summe kann man eindeutig als Summe von Vektoren $w_i \in W_i$ schreiben.
   - Äußere direkte Summe und Produkt einer Familie von Vektorräumen.
   - Zu einem Untervektorraum $U \subset V$ können wir den Quotienvektorraum V/U konstruieren. Im Falle endlich erzeugter Vektorräume ist dessen Dimension $\dim_K V/U = \dim_K V - \dim_K U$.

3. Aus den Vektoren eines Vektorraums kann man Linearkombinationen bilden. Dies führt auf zwei wichtige Begriffe

   - Erzeugnis /Erzeugendensystem $\mathcal{B} \subset V$: es gilt $\mathrm{span}_K(\mathcal{B}) = V$.
   - Lineare Unabhängigkeit

Linear unabhängige Erzeugendensysteme heißen Basen; sie existieren für jeden Vektorraum; die Anzahl ihrer Elemente ist eindeutig und heißt die Dimension des Vektorraums. Basen sind maximale linear unabhängige Familien und minimale Erzeugendensysteme.

- Der Basisauswahlsatz 2.4.6 erlaubte es uns, aus jedem Erzeugendensystem eine Basis auszuwählen.
- Der Basisergänzungssatz 2.4.11.2 erlaubt es uns, jede linear unabhängige Familie zu einer Basis zu ergänzen.

4. Für K-lineare Abbildungen $\Phi : V \to W$ sahen wir:

   - Urbilder von Untervektorräumen sind Untervektorräume von V und somit nie die leere Menge. Insbesondere ist der Kern von Φ als Urbild der $0 \in W$ ein Untervektorraum von V. Die lineare Abbildung Φ ist genau dann injektiv, wenn $\ker \Phi = \{0\}$ gilt.
   - Urbilder affiner Unterräume sind affine Unterräume oder die leere Menge.
   - Der Raum $\mathrm{Hom}_K(V, W)$ der K-linearen Abbildungen ist ein K-Vektorraum der Dimension $\dim_K V \cdot \dim_K W$.
   - Den Homomorphiesatz 3.5.4: für eine lineare Abbildung $\Phi : V \to W$ erhalten wir einen einduetigen Isomorphismus $V/\ker \Phi \xrightarrow{\sim} \operatorname{Im} \Phi$ so dass

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ {\scriptstyle\pi}\downarrow & & \uparrow{\scriptstyle\iota} \\ V/\ker\Phi & \dashrightarrow & \Phi(W) \end{array}$$

   Hieraus folgt insbesondere die Dimensionsformel 3.1.7:

$$\dim_K V - \dim_K \ker \Phi = \dim_K V/\ker \Phi = \dim_K \operatorname{Im} \Phi \stackrel{\text{def}}{=} \operatorname{rg} \Phi \,.$$

5. Explizite Beschreibung von linearen Abbildungen $\Phi : V \to W$ für endlich-dimensionale Vektorräume:

   - Situation:
     $\dim_K V < \infty \quad \mathcal{A} = (v_1, \ldots v_n)$ geordnete Basis von V
     $\dim_K W < \infty \quad \mathcal{B} = (w_1, \ldots w_m)$ geordnete Basis von W
   - Jede geordnete Basis $\mathcal{A}$ von V gibt einen Isomorphismus vom Standardvektorraum K^n mit Standardbasis auf V:

$$\Phi_{\mathcal{A}} : K^n \to V \text{ mit } \Phi_{\mathcal{A}}(e_i) = v_i \,.$$

   - Die darstellenden Matrizen geben Isomorphismen

$$M_{\mathcal{B}}^{\mathcal{A}} : \ \mathrm{Hom}_K(V, W) \to M(m \times n, K)$$

     von K Vektorräumen (mit Multiplikationen, wo diese definiert sind).
   - Ein Basiswechsel wird durch invertierbare Transformationsmatrizen

$$T_{\mathcal{A}'}^{\mathcal{A}} = M_{\mathcal{A}'}^{\mathcal{A}}(\mathrm{id}_V)$$

     beschrieben. Für lineare Abbildungen gilt die Transformationsformel 3.6.5

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\Phi) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \cdot \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1} .$$

- Zwei (nicht notwendigerweise quadratische) Matrizen X, Y heißen äquivalent, wenn es invertible (und damit quadratische) Matrizen S, T gibt mit $Y = SXT^{-1}$. Jede Matrix A ist äquivalent zu einer Matrix der Form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

  mit $r = \operatorname{rg}(A)$. Zeilenrang und Spaltenrang einer Matrix sind gleich.
- Zwei *quadratische* Matrizen X, Y heißen ähnlich, wenn es *eine* invertible Matrix S gibt mit $Y = SXS^{-1}$.

6. Aus den entwickelten Begriffen folgt eine Theorie für die Lösungsmenge

$$\operatorname{Lsg}(A, b) := \{x \in K^n \mid Ax = b\}$$

eines inhomogenen linearen Gleichungssystems:

- $\operatorname{Lsg}(A, b) = \emptyset$ genau dann, wenn $b \notin \operatorname{Im} A$, was genau dann gilt, wenn $\operatorname{rg}(A, b) = \operatorname{rg}(A) + 1$.
- $\operatorname{Lsg}(A, b)$ ist entweder leer oder affiner Unterraum der Dimension $n - \operatorname{rg} A$. Man erhält alle Lösungen des inhomogenen linearen Gleichungssystems, indem man zu einer speziellen Lösung alle Lösungen des zugehörigen homogenen Gleichungssystems addiert.
- Der Gauß'sche Algorithmus erlaubt es, lineare Gleichungssysteme systematisch zu lösen.

4 Determinanten

4.1 Das Vektorprodukt auf $\mathbb{R}^3$

Definition 4.1.1
Die Abbildung

$$\begin{array}{rcl} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (v, w) & \mapsto & v \times w := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} \end{array}$$

heißt Vektorprodukt.

Das Vektorprodukt kann nur für drei-dimensionale reelle Vektorräume mit Wahl einer geordneten Basis (eigentlich: Wahl einer euklidischen Struktur mit Orientierung) definiert werden. Das Vektorprodukt wird von vielen Anwendern (Ingenieure, Physiker, Schule,...) gerne benutzt.

Lemma 4.1.2.
Zusammen mit dem schon behandelten euklidischen Skalarprodukt

$$\langle v, w \rangle := v_1 w_1 + v_2 w_2 + v_3 w_3$$

erfüllt das Vektorprodukt für alle $u, v, u', v' \in \mathbb{R}^3$ und $\alpha, \beta \in \mathbb{R}$ die folgenden Identitäten:

1. Bilinearität

$$\begin{aligned} (\alpha v + \beta v') \times w &= \alpha(v \times w) + \beta(v' \times w) \\ v \times (\alpha w + \beta w') &= \alpha(v \times w) + \beta(v \times w') \end{aligned}$$

2. Antisymmetrie

$$v \times w = -w \times v$$

3. Grassmann–Identität

$$u \times (v \times w) = \langle u, w \rangle v - \langle u, v \rangle w \ .$$

4. Jacobi-Identität

$$u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0$$

5. $\langle u \times v, w \rangle = \langle u, v \times w \rangle$

Beweis.
1.–3. durch einfaches Nachrechnen, ebenso 5. Wir zeigen, wie 4. aus 3. und der Symmetrie des Skalarprodukts folgt:

$$\begin{aligned} & u \times (v \times w) + v \times (w \times u) + w \times (u \times v) \\ \underset{3.}{=} & \langle u, w \rangle v - \langle u, v \rangle w + \langle v, u \rangle w - \langle v, w \rangle u + \langle w, v \rangle u - \langle w, u \rangle v = 0 \end{aligned}$$

□

Lemma 4.1.3 (Verschärfung der Cauchy–Schwarz'schen Ungleichung für den $\mathbb{R}^3$)**.**
Für alle $v, w \in \mathbb{R}^3$ gilt

$$\langle v, w\rangle^2 + \|v \times w\|^2 = \|v\|^2 \cdot \|w\|^2 \ .$$

Beweis.
Wir rechnen:

$$\begin{aligned}
\|v \times w\|^2 &= \langle v \times w, v \times w\rangle \\
&= \langle v, w \times (v \times w)\rangle \qquad \text{[wegen Lemma 4.1.2.5]} \\
&= \langle v, \langle w, w\rangle v - \langle w, v\rangle w\rangle \qquad \text{[wegen Lemma 4.1.2.3]} \\
&= \|w\|^2 \ \|v\|^2 - \langle v, w\rangle^2 \ .
\end{aligned}$$

□

Betrachtung 4.1.4.

- Aus der Antisymmetrie, Lemma 4.1.2.2, folgt $v \times v = -v \times v = 0$. Aus Lemma 4.1.2.5 folgt daher

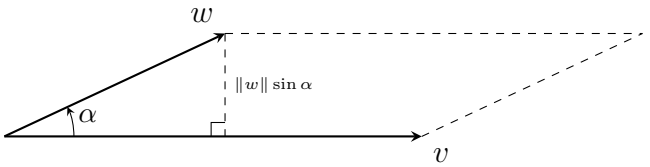
  $$\langle v, v \times w\rangle = \langle v \times v, w\rangle = 0 \ ,$$

  also $v \perp (v \times w)$. Ähnlich folgt auch $w \perp (v \times w)$. Der Vektor $v \times w$ steht also auf den beiden Vektoren v und w senkrecht.

- Wir wollen auch die Länge des Vektors $v \times w$ berechnen. Mit Lemma 4.1.3 folgt:

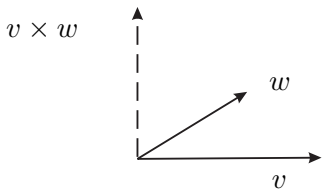
  $$\begin{aligned}
  \|v \times w\|^2 &= \|v\|^2\|w\|^2 - \langle v, w\rangle^2 \\
  &= \|v\|^2\|w\|^2(1 - \cos^2\alpha) \\
  &= \|v\|^2\|w\|^2 \sin^2\alpha \ .
  \end{aligned}$$

  $\|v \times w\|$ ist daher der Flächeninhalt des von v, w aufgespannten Parallelogramms:

  Man kann also mit Hilfe des Vektorprodukts Flächen von Parallelogrammen (und damit auch Dreiecken) ausrechnen.

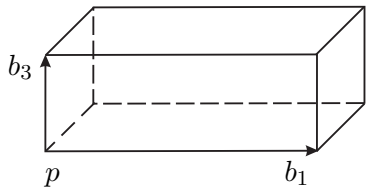
- Damit kennen wir Länge und Richtung des Vektors $v \times w$. Seine Orientierung wird durch die sogenannte "rechte Hand–Regel" angegeben: zeigt der Damen der rechten Hand in Richtung von v und der Zeigefinger in Richtung von w, so zeigt der Mittelfinger in Richtung von $v \times w$. Dies macht das Beispiel $v = e_1$ und $w = e_2$ mit Vektorprodukt $v \times w = e_3$ deutlich.

Definition 4.1.5
Sei (b_1, b_2, b_3) eine Basis von $\mathbb{R}^3$. Sei $p \in \mathbb{R}^3$. Dann heißt die Teilmenge

$$P = \{p + \alpha b_1 + \beta b_2 + \gamma b_3 |\, 0 \leq \alpha, \beta, \gamma \leq 1\} \ \subset \ \mathbb{R}^3$$

das von b_1, b_2 und b_3 aufgespannte Parallelotop oder Spat.



Satz 4.1.6.
Das Volumen des von den drei Vektoren $b_1, b_2, b_3 \in \mathbb{R}^3$ aufgespannten Spats ist

$$\mathrm{vol}(P) = |\langle b_1 \times b_2, b_3 \rangle| \ .$$

Deshalb heißt das Vektorprodukt auch Spatprodukt.

Beweis.
Das Volumen ist bekanntermaßen "Grundfläche mal Höhe". Wenn die Grundfläche ungleich Null ist, ist $b_1 \times b_2$ nicht der Nullvektor; wir bilden daher den Einheitsvektor

$$n := \frac{b_1 \times b_2}{\|b_1 \times b_2\|}$$

steht senkrecht auf den beiden Vektoren b_1 und b_2, die die Grundfläche aufspannen. Die Höhe des Spats ist daher

$$h = |\langle n, b_3 \rangle|$$

Wir erhalten

$$V = F \cdot h \overset{4.1.4}{=} \|b_1 \times b_2\| \cdot \ |\langle n, b_3 \rangle| = |\langle b_1 \times b_2, b_3 \rangle| \ .$$

□

4.2 Die Determinantenabbildung

Wir wollen einen Volumensbegriff für n–dimensionaler Parallelotope $P \subset \mathbb{R}^n$ einführen. Dazu betrachten wir eine Matrix, deren Zeilenvektoren die Transponierten der Basisvektoren sind, die den Spat aufspannen.
Wir erhalten so durch das Volumen eine Abbildung

$$M(n \times n, \mathbb{R}) \to \mathbb{R}$$

mit gewissen Eigenschaften. Die Eigenschaften dieser Abbildung führen auf den Begriff der Determinantenabbildung, die wir gleich für Matrizen über einem beliebigen Körper K einführen:

Definition 4.2.1

Sei K ein Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\begin{aligned} \det : \quad M(n \times n, K) &\to K \\ A &\mapsto \det(A) \end{aligned}$$

heißt Determinantenabbildung, falls gilt:

(D1) det *ist linear in jeder Zeile: es gilt*

$$\det \begin{pmatrix} (a_1)^t \\ \vdots \\ (\lambda a_i)^t \\ \vdots \\ (a_n)^t \end{pmatrix} = \lambda \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i)^t \\ \vdots \\ (a_n)^t \end{pmatrix}$$

und

$$\det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i + a_i')^t \\ \vdots \\ (a_n)^t \end{pmatrix} = \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i)^t \\ \vdots \\ (a_n)^t \end{pmatrix} + \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i')^t \\ \vdots \\ (a_n)^t \end{pmatrix}$$

Hierbei sind alle Vektoren $a_i \in K^n$ als Spaltenvektoren zu sehen, d.h. der transponierte Vektor $(a_i)^t$ steht für den i-ten Zeilenvektor.

(D2) det *ist alternierend, d.h. stimmen zwei Zeilen überein, so ist* $\det(A) = 0$.

(D3) det *ist normiert, d.h.* $\det(E_n) = 1$.

Ein Beispiel für eine Determinantenabbildung haben wir schon kennengelernt:

$$\begin{aligned} M(3 \times 3, \mathbb{R}) &\to \mathbb{R} \\ (a_1, a_2, a_3) &\mapsto \langle a_1, a_2 \times a_3 \rangle \end{aligned}$$

Wir untersuchen nun Determinantenabbildungen näher, um ihre Existenz und Eindeutigkeit zu verstehen und sie berechnen zu können:

Satz 4.2.2.
Sei K ein Körper und $n \in \mathbb{N}$. Sei $\det : M(n \times n, K) \to K$ eine Determinantenabbildung. Dann gilt für alle $A, B \in M(n \times n, K)$ und alle $\lambda \in K$:

1. $\det(\lambda A) = \lambda^n \det(A)$
2. Ist eine Zeile von A gleich 0, so ist $\det A = 0$.
3. Entsteht B aus A durch Vertauschung zweier Zeilen, so ist $\det B = -\det A$.
4. Entsteht B aus A durch Addition des Vielfachen einer Zeile zu einer anderen, so ist $\det B = \det A$.

5. Sei A eine obere Dreiecksmatrix der Form

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ 0 & & \lambda_n \end{pmatrix} .$$

Hierbei steht $*$ für beliebige Einträge oberhalb der Hauptdiagonalen. Dann ist

$$\det A = \lambda_1 \dots \lambda_n = \prod_{j=1}^{n} \lambda_j .$$

Die Determinante ist wegen 1. für $n \geq 2$ keine lineare Funktion; sie ist auch nicht additiv: zum Beispiel gilt für $A = B = E_2$, dass $\det A + \det B = 2 \det E_2 = 2$, aber $\det(A + B) = \det(2E_2) = 4 \det E_2 = 4$.

Beweis.

1. Wir rechnen:

$$\det(\lambda A) = \det \left(\lambda \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} \right) = \det \begin{pmatrix} \lambda(a_1)^t \\ \vdots \\ \lambda(a_n)^t \end{pmatrix} = \lambda \det \begin{pmatrix} (a_1)^t \\ \lambda(a_2)^t \\ \vdots \\ \lambda(a_n)^t \end{pmatrix}$$

$$= \lambda^2 \det \begin{pmatrix} (a_1)^t \\ (a_2)^t \\ \lambda(a_3)^t \\ \vdots \end{pmatrix} = \dots = \lambda^n \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} .$$

2. Aus der Zeilenlinearität (D1) folgt sofort:

$$\det A = \det \begin{pmatrix} (a_1)^t \\ \vdots \\ 0 \\ \vdots \\ (a_n)^t \end{pmatrix} = \det \begin{pmatrix} (a_1)^t \\ \vdots \\ 0 \cdot 0 \\ \vdots \\ (a_n)^t \end{pmatrix} = 0 \cdot \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} = 0 .$$

3. Sei $A = \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix}$. B gehe aus A durch Vertauschung der i–ten und j–ten Zeile hervor, mit $i > j$. Dann ist wegen (D2):

$$0 = \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i + a_j)^t \\ \vdots \\ (a_i + a_j)^t \\ \vdots \\ (a_n)^t \end{pmatrix} = \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i)^t \\ \vdots \\ (a_i)^t \\ \vdots \\ (a_n)^t \end{pmatrix} + \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_i)^t \\ \vdots \\ (a_j)^t \\ \vdots \\ (a_n) \end{pmatrix} + \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_j)^t \\ \vdots \\ (a_i)^t \\ \vdots \\ (a_n)^t \end{pmatrix} + \det \begin{pmatrix} (a_1) \\ \vdots \\ (a_j)^t \\ \vdots \\ (a_j)^t \\ \vdots \\ (a_n)^t \end{pmatrix}$$

$$= 0 + \det A + \det B + 0 .$$

4. B entstehe aus A durch Addition des λ–fachen der j–ten Zeile zur i–ten Zeile, mit $i \neq j$. Dann ist

$$\begin{aligned}\det B = \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_{i-1})^t \\ (a_i + \lambda a_j)^t \\ (a_{i+1})^t \\ \vdots \\ (a_n)^t \end{pmatrix} &= \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_{i-1})^t \\ (a_i)^t \\ (a_{i+1})^t \\ \vdots \\ (a_n)^t \end{pmatrix} + \lambda \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_{i-1})^t \\ (a_j)^t \\ (a_{i+1})^t \\ \vdots \\ (a_n)^t \end{pmatrix} \\ &= \det A + \lambda \cdot 0 = \det A \ ,\end{aligned}$$

   denn die zweite Matrix in der Summe hat die gleichen Einträge in der i-ten und j-ten Zeile. Damit wissen wir, wie sich die Determinantenabbildung unter den elementaren Zeilenumformungen aus Satz 3.4.10 verhält.

5. Sind alle $\lambda_i \neq 0$, so kann man A durch Zeilenumformungen in eine Diagonalmatrix überführen mit gleichen Diagonalelementen:

$$\det A = \det \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \cdot \det E_N \overset{\text{(D3)}}{=} \prod_{i=1}^{n} \lambda_i \ .$$

   Sind nicht alle $\lambda_j \neq 0$, so sei i der größte Index, für den $\lambda_i = 0$ ist. Wiederum durch elementare Zeilenumformungen finden wir

$$\det A = \det \begin{pmatrix} \lambda_1 & & & & & \\ 0 & \ddots & & & * & \\ & & \lambda_{i-1} & & & \\ 0 & 0 & 0 & 0 & \dots & 0 \\ & & & & \lambda_{i+1} & * \\ & 0 & & & & \ddots \end{pmatrix} = 0 = \prod_{j=1}^{n} \lambda_j \ .$$

□

Lemma 4.2.3.
Die Determinante einer Matrix verschwindet genau dann, wenn diese nicht maximalen Rang hat: $\det A = 0 \Leftrightarrow \operatorname{rg}\ A < n$.

Beweis.
Durch spezielle Zeilenumformungen, d.h. Zeilenvertauschungen und Additionen von Vielfachen von Zeilen zu anderen Zeilen, überführen wir A mit dem Gauß'schen Algorithmus in eine obere Dreiecksmatrix A'. Die Matrizen A und A' haben gleichen Rang, den sie gehen auseinander hervor durch Multiplikation mit einem Produkt von Elementarmatrizen, vgl. Lemma 3.4.9, das eine invertible Matrix ist. Aus Lemma 4.3.3.3 und 4 folgt

$$\det A' = \pm \det A \ .$$

Es ist

$$\det A' = \det \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} \lambda_i \ .$$

Daher verschwindet die Determinante, $\det A = 0$ genau dann, wenn wenigstens ein λ_j verschwindet, $\lambda_j = 0$. Das heißt aber, dass A' nicht invertierbar ist, was zu

$$\operatorname{rg} A = \operatorname{rg} A' < n$$

äquivalent ist. □

Korollar 4.2.4.
Für jeden Körper K und jedes $n \geq 1$ gibt es *höchstens* eine Determinantenabbildung

$$\det \ : \ M(n \times n, K) \to K \ .$$

Beweis.
Überführe eine Matrix A durch spezielle Zeilenumformung in eine obere Dreiecksmatrix A', wobei k Zeilenvertauschungen auftreten:

$$A' = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & * & \\ 0 & & & \lambda_n \end{pmatrix}$$

Für jede Determinantenabbildung muss also gelten

$$\det A = (-1)^k \det A' = (-1)^k \prod_{j=1}^{n} \lambda_j \ .$$

□

Beispiel 4.2.5.
Wir arbeiten über dem Körper $K = \mathbb{C}$ und betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 & \mathrm{i} \\ 1 & \mathrm{i} & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

Wir führen elementare Zeilenumformungen aus:

$$\begin{aligned} \det A &= -\det \begin{pmatrix} 1 & \mathrm{i} & 1 \\ 0 & 1 & \mathrm{i} \\ 2 & 3 & 4 \end{pmatrix} = -\det \begin{pmatrix} 1 & \mathrm{i} & 1 \\ 0 & 1 & \mathrm{i} \\ 0 & 3-2\mathrm{i} & 2 \end{pmatrix} \\ &= -\det \begin{pmatrix} 1 & \mathrm{i} & 0 \\ 0 & 1 & \mathrm{i} \\ 0 & 0 & 2-(3-2\mathrm{i})\mathrm{i} \end{pmatrix} = -(-3\mathrm{i}) = 3\mathrm{i} \ . \end{aligned}$$

Wir werden noch sehen, dass der Gauß-Algorithmus ein sehr effizienter Algorithmus für die Bestimmung von Determinanten ist.

Satz 4.2.6.
Für jeden Körper K und jedes $n \geq 1$ gibt es genau eine Determinantenabbildung

$$\det_n : \ M(n \times n, K) \to K$$

Beweis.
der Existenz durch vollständige Induktion nach n.

- Induktionsanfang: definiere

  $$\det_1(a) := a$$

  (D1)-(D3) sind in diesem Fall offensichtlich.

- Für den Induktionsschritt definiere für $A \in M(n \times n, K)$ die folgenden n^2 Streichungsmatrizen:

  $$A_{ij}^{Str} := \left(\begin{array}{ccc|c|ccc} a_{11} & \dots & a_{1j-1} & a_{1j} & a_{1j+1} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \hline a_{i1} & \dots & a_{ij-1} & a_{ij} & a_{ij+1} & \dots & a_{in} \\ \hline \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nj} & a_{nj+1} & \dots & a_{nn} \end{array}\right) \in M\left((n-1) \times (n-1), K\right)$$

  und definiere für beliebiges, aber festes $j \in \{1, \dots, n\}$

  $$\det_n(A) := \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}\left(A_{ij}^{Str}\right) \ .$$

  Wir müssen jetzt die Axiome (D1)-(D3) überprüfen.

(D1) Entstehe $\widetilde{A}$ aus A durch Multiplikation der k–ten Zeile mit $\lambda \in K$: $\tilde{a}_{kj} = \lambda a_{kj}$ und $\tilde{a}_{ij} = a_{ij}$ für $i \neq k$. Für die Streichungsmatrizen gilt:

- $\widetilde{A}_{kj}^{Str} = A_{kj}^{Str}$ da die einzige veränderte Zeile, nämlich die k-te Zeile, gestrichen wird.
- Für $i \neq k$ entsteht auch $\widetilde{A}_{ij}^{Str}$ aus A_{ij}^{Str} durch Multiplikation einer Zeile mit $\lambda \in K$, also gilt nach Induktionsannahme

  $$\det_{n-1}\left(\widetilde{A}_{ij}^{Str}\right) = \lambda \det_{n-1}\left(A_{ij}^{Str}\right) \ .$$

  Somit ist

  $$\det_n \widetilde{A} \stackrel{\text{def}}{=} \sum_{i \neq k} (-1)^{i+j} \widetilde{a}_{ij} \det_{n-1}\left(\widetilde{A}_{ij}^{Str}\right) + (-1)^{k+j} \widetilde{a}_{kj} \det_{n-1} \widetilde{A}_{kj}^{Str} = \lambda \det_n(A) \ .$$

  Die Additivität zeigt man analog.

(D2) Mögen die k–te und l–te Zeile übereinstimmen. Ohne Beschränkung der Allgemeinheit sei $k < l$. Ist $i \neq k$ und $i \neq l$, so hat die Streichungsmatrix A_{ij}^{Str} auch zwei identische Zeilen; nach Induktionsannahme folgt

$$\det_{n-1} A_{ij}^{Str} = 0 \ .$$

Also erhält man:

$$\det_n A = (-1)^{k+j} a_{kj} \det{}_{n-1} A_{kj}^{Str} + (-1)^{l+j} a_{lj} \det_{n-1} A_{lj}^{Str}$$

- Aus der Gleichheit der k-ten und l-ten Zeile folgt $a_{kj} = a_{lj}$.
- A_{lj}^{Str} geht aus A_{kj}^{Str} durch $(l-k-1)$ Zeilenvertauschungen hervor, also gilt

$$\det_{n-1} A_{lj}^{Str} = (-1)^{l-k-1} \det_{n-1} A_{kj}^{Str}$$

Insgesamt folgt

$$\det_n A = 0 \ .$$

(D3) Die Einheitsmatrix $A = E_n$ hat Einträge $a_{ij} = \delta_{ij}$. Daher gilt

$$\begin{aligned} \det_n(E_n) &= \sum_{i=1}^{n} (-1)^{i+j} \delta_{ij} \det_{n-1} A_{ij}^{Str} \\ &= (-1)^{j+j} \det_{n-1} A_{jj}^{Str} = \det_{n-1}(E_{n-1}) = 1 \ . \end{aligned}$$

Denn bei den Streichungsmatrizen A_{ij}^{Str} mit $i \neq j$ werden zwei verschiedene Einsen auf der Diagonale gestrichen, so dass die $(n-1) \times (n-1)$-Matrix A_{ij}^{Str} nur $n-2$ Einsen enthält, also eine Zeile enthalten muss, die Null ist. Somit verschwindet nach Satz 4.2.2.2 ihre Determinante.

□

Aus dem Satz folgt sofort der

Satz 4.2.7. Spaltenentwicklungssatz von Laplace
Für jede Matrix $A \in M(n \times n, K)$ gilt

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}^{Str}$$

für jedes feste $1 \leq j \leq n$.

Die Vorzeichen im Laplace'schen Entwicklungssatz folgen einem Schachbrettmuster.

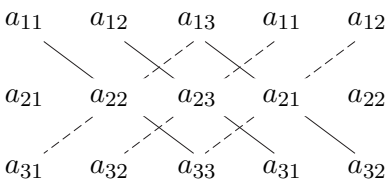
Beispiele 4.2.8.

1. K beliebig, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist $\det A = ad - cb$.

2. K beliebig, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$\det(A) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}$$

Für 3×3-Matrizen gibt es die Merkregel von Sarrus: : man berechnet für alle drei Parallelen zur Hauptdiagonalen – hier durchgehend eingezeichnet – die Produkte der Einträge und addiert die Ergebnisse auf. Davon zieht man die drei Produkte der Einträge auf den drei Parallelen der Nebendiagonalen – im Schema gestrichelt gezeichnet – ab.



Vorsicht: in der Entwicklung der Determinante einer $n \times n$–Matrix treten $n!$ Terme auf; es gibt also keine Verallgemeinerung der Regel von Sarrus für $n \geq 4$, die es erlaubt, nur mit Produkten auf Haupt- und Nebendiagonalen zu arbeiten!
Wir illustrieren die Berechnung durch Entwicklung nach der ersten Spalte:

$$\begin{aligned}\det \begin{pmatrix} 0 & 1 & \mathrm{i} \\ 1 & \mathrm{i} & 1 \\ 2 & 3 & 4 \end{pmatrix} &= 0 \cdot \det \begin{pmatrix} \mathrm{i} & 1 \\ 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & \mathrm{i} \\ 3 & 4 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix} \\ &= -(4 - 3\mathrm{i}) + 2(1 - \mathrm{i}^2) = 3\mathrm{i} \; .\end{aligned}$$

Rechnerisch ist die Berechnung von Determinanten mit Hilfe von Entwicklungssätzen weniger effizient als mit dem Gaußalgorithmus.

Satz 4.2.9.
Die Determinante ist auch linear in jeder Spalte; d.h. für alle Spaltenvektoren $a_1, \ldots, a_n, a'_j, a''_j \in K^n$ und $\lambda \in K$ gilt

$$\begin{aligned}\det(a_1 \ldots \lambda a_j \ldots a_n) &= \lambda \det(a_1 \ldots a_j \ldots a_n) \\ \det(a_1 \ldots \; a'_j + a''_j \ldots a_n) &= \det(a_1 \ldots a'_j \ldots a_n) + \det(a_1 \ldots a''_j \ldots a_n)\end{aligned}$$

Beweis.
Aus der Entwicklung nach der j–ten Spalte folgt:

$$\det(a_1 \ldots \lambda a_j \ldots a_n) = \sum_{i=1}^{n} (-1)^{i+j} (\lambda a_{ij}) \det A_{ij}^{Str} = \lambda \det A \; .$$

Analog zeigt man die Additivität. □

Satz 4.2.10.
Für alle $A \in M(n \times n, K)$ gilt $\det(A^t) = \det A$.

Beweis.
Wegen der Eindeutigkeit der Determinantenfunktion reicht es aus, zu zeigen, dass auch

$$\begin{aligned}\widetilde{\det} : \; M(n \times n, K) &\to K \\ A &\mapsto \det A^t\end{aligned}$$

eine Determinantenfunktion ist.

(D1) folgt aus der Spaltenlinearität in Satz 4.2.9.

(D2) Wenn A zwei gleiche Zeilen hat, so hat A^t zwei gleiche Spalten. Also ist $\mathrm{rg}\,(A^t) < n$, nach Lemma 4.2.3 muss $0 = \det A^t = \widetilde{\det} A$ gelten.

(D3) folgt aus

$$\widetilde{\det} E_n = \det E_n^t = \det E_n = 1 \; .$$

□

Korollar 4.2.11.

1. Zeilenentwicklungssatz von Laplace: für jedes $A \in M(n \times n, K)$ gilt

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}^{Str} \, , ,$$

   und zwar für jedes mögliche i, mit $1 \leq i \leq n$.

2. Entsteht $\widetilde{A}$ aus A durch Vertauschung zweier Spalten, so ist

$$\det \widetilde{A} = - \det A \, .$$

Beweis.

1. Berechne $\det A^t$ durch Entwicklung nach der i–ten Spalte, die ja genau die i–te Zeile von A ist. Beispiel: Entwicklung nach der ersten Zeile zur Berechnung der folgenden Determinante:

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & \mathrm{i} \\ 1 & \mathrm{i} & 1 \\ 2 & 3 & 4 \end{pmatrix} &= 0 \cdot \det \begin{pmatrix} \mathrm{i} & 1 \\ 3 & 4 \end{pmatrix} - \det \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} + \mathrm{i} \cdot \det \begin{pmatrix} 1 & \mathrm{i} \\ 2 & 3 \end{pmatrix} \\ &= -2 + \mathrm{i}(3 - 2\mathrm{i}) = 3\mathrm{i} \end{aligned}$$

2. Dann entsteht $\widetilde{A}^t$ aus A^t durch Vertauschen zweier Zeilen. Aus Satz 4.2.2.3 folgt

$$\det \widetilde{A} = \det \widetilde{A}^t = - \det A^t = - \det A \, .$$

□

Satz 4.2.12. Determinantenmultiplikationssatz
Für alle $A, B \in M(n \times n, K)$ gilt $\det(A \cdot B) = \det A \cdot \det B$.

Beweis.

- Wir behandeln erst den Fall $\det B = 0$. Dann ist nach Lemma 4.2.3 $\operatorname{rg} B < n$, nach der Dimensionsformel 2.5.3 ist dann $\ker B$ nicht trivial. Also gibt es $x \in K^n$, $x \neq 0$ mit $Bx = 0$. Es folgt erst recht $ABx = 0$, also $\operatorname{rg}(AB) < n$, wiederum nach der Dimensionsformel. Daraus folgt $\det(AB) = 0$ und somit die behauptete Gleichung.

- Wir halten B mit $\det B \neq 0$ fest und zeigen, dass auch die Funktion

$$\widetilde{\det} A := \frac{\det A \cdot B}{\det B}$$

  eine Determinantenfunktion ist, woraus die Behauptung wegen der Eindeutigkeitsaussage in Korollar 4.2.4 folgt.

(D1) Entsteht $\tilde{A}$ aus A durch Multiplikation der i–ten Zeile mit $\lambda \in K$, so ist

$$\widetilde{A} = \Delta_{\lambda,i} \cdot A$$

mit der Matrix

$$\Delta_{\lambda,i} := \begin{pmatrix} 1 & & & & & \\ & 1 & & 0 & & \\ & & \ddots & & & \\ & 0 & & \lambda & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

$\uparrow$
i–te Spalte.

Es folgt $\widetilde{A}B = \Delta_{\lambda,i}(A \cdot B)$, d.h. auch $\widetilde{A}B$ entsteht aus AB durch Multiplikation der i–ten Zeile mit λ. Nach dem Axiom (D1) für die Determinantenfunktion det folgt

$$\det(\widetilde{A}B) = \lambda \det AB$$

und somit

$$\widetilde{\det}\,\widetilde{A} = \frac{\det(\widetilde{A}B)}{\det B} = \lambda \frac{\det AB}{\det B} = \lambda \widetilde{\det}\, A.$$

Um die Additivität von $\widetilde{\det}$ zu zeigen, schreiben wir

$$A = \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} \qquad B = (b_1, \ldots, b_n) \ ,$$

drücken also die Matrix A durch Zeilenvektoren $a_j \in K^n$ und die Matrix B durch Spaltenvektoren $b_j \in K^n$ aus. Es gelte $a_i = a_i' + a_i''$. Dann ist

$$\begin{aligned} A \cdot B &= \begin{pmatrix} (a_1)^t b_1 & \ldots & (a_1)^t b_n \\ \vdots & & \vdots \\ (a_n)^t b_1 & \ldots & (a_n)^t b_n \end{pmatrix} \\ &= \big((a_i' + a_i'')^t b_1 \ldots (a_i' + a_i'')^t b_n\big) \end{aligned}$$

Somit folgt $\det AB = \det A'B + \det A''B$ und daraus die Additivität von $\widetilde{\det}$:

$$\widetilde{\det}\, A = \frac{\det AB}{\det B} = \widetilde{\det}\, A' + \widetilde{\det}\, A'' \ .$$

(D2) Hat A zwei gleiche Zeilen, so ist $\operatorname{rg} A < n$. Wegen $\operatorname{Im} AB \subset \operatorname{Im} A$ ist $\operatorname{rg} AB < n$, also wegen Lemma 4.2.3 auch $\det(AB) = 0$, also $\widetilde{\det} A = 0$.

(D3) Auch die Funktion $\widetilde{\det}$ ist normiert:

$$\widetilde{\det}\, E_n = \frac{\det(E_n \cdot B)}{\det B} = 1 \ .$$

□

Korollar 4.2.13.
Ist $A \in GL(n,K)$, so ist $\det A \neq 0$ und es gilt $\det A^{-1} = (\det(A))^{-1}$.

Beweis.

$$1 = \det(E_n) = \det\left(A \cdot A^{-1}\right) \overset{4.2.12}{=} \det A \cdot \det A^{-1} \ .$$

□

Satz 4.2.14.
Für $A \in M(n \times n, K)$ setze $B = (b_{ij}) \in M(n \times n, K)$ mit

$$b_{ij} := (-1)^{i+j} \det A_{ji}^{Str} \ .$$

Man beachte die Reihenfolgen der Indizes! Dann gilt

$$AB = BA = \det(A) E_n \ .$$

Ist insbesondere A invertierbar, so ist das Inverse von A gleich

$$A^{-1} = \frac{1}{\det A} B \ .$$

Beispiel: $n = 2$, K beliebig: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sei invertierbar. Dann ist

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^t = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \ .$$

Bitte rechnen Sie selbst die Produkt $A^{-1}A$ und AA^{-1} aus, um zu verifizieren, dass dies wirklich die inverse Matrix ist.
Beweis.

- Der (i,i)–te Eintrag von AB ist:

$$\sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det A_{ij}^{Str} = \det A$$

  nach dem Zeilenentwicklungssatz 4.2.11 für die i–te Zeile.

- Der (i,k)–te Eintrag von AB ist für $i \neq k$:

$$\sum_{j=1}^{n} a_{ij} b_{jk} = \sum_{j=1}^{n} a_{ij} (-1)^{j+k} \det\left(A_{kj}^{Str}\right) = \sum_{j=1}^{n} \widetilde{a}_{kj} (-1)^{j+k} \det\left(\widetilde{A}_{kj}^{Str}\right) \overset{4.2.11}{=} \det \widetilde{A} \ ,$$

  wobei $\widetilde{A}$ aus A entsteht, indem man die k-te Zeile durch die i–te Zeile ersetzt. Dies ändert nicht die Streichungsmatrix, da die k-te Zeile dort ohnehin gestrichen wird. Beim Matrixelement $\widetilde{a}_{kj}$ haben wir eine entsprechende Änderung des Index vorgenommen. Wegen $i \neq k$ hat $\widetilde{A}$ zwei Zeilen mit identischen Einträgen, also ist $\det \widetilde{A} = 0$.

- Aus dem Spaltenentwicklungssatz folgen die analogen Aussagen für das Produkt $B \cdot A$.

□

Satz 4.2.15. Cramersche Regel
Seien $a_1, \ldots, a_n, b \in K^n$ und sei die Matrix $A = (a_1, \ldots, a_n)$ invertierbar. Dann ist die eindeutige Lösung $x \in K^n$ des inhomogenen linearen Gleichungssystems von n Gleichungen für n Variablen

$$Ax = b$$

gegeben durch

$$x_i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A}$$

Beweis.

- Das lineare Gleichungssystem $Ax = b$ hat die eindeutige Lösung

$$x = A^{-1}b$$

- Mit dem Ergebnis von Satz 4.2.14

$$\left(A^{-1}\right)_{ij} = \frac{1}{\det A}(-1)^{i+j}\det A_{ji}^{Str}$$

folgt

$$x_i = \sum_{j=1}^{n} \left(A^{-1}\right)_{ij} b_j = \frac{1}{\det A}\sum_{j=1}^{n}(-1)^{i+j}\det\left(A_{ji}^{Str}\right) \cdot b_j \,.$$

Wir zeigen nun die Identität

$$(-1)^{i+j}\det A_{ji}^{Str} = \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) \,.$$

Sie folgt aus

$$(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) = \begin{pmatrix} a_{1,1} & \ldots & a_{i-1,1} & 0 & a_{i+1,1} & \ldots & a_{n,1} \\ \vdots & & & \vdots & & & \vdots \\ a_{1,j} & \ldots & a_{i-1,j} & 1 & a_{i+1,j} & \ldots & a_{n,j} \\ & & & 0 & & & \\ & & & \vdots & & & \\ a_{1,n} & \ldots & a_{i-1,n} & 0 & a_{i+1,n} & \ldots & a_{n,n} \end{pmatrix}$$

und Entwicklungssatz 4.2.7 nach der i-ten Spalte. Aus der Spalten-Linearität der Determinante und

$$b = \sum_{j=1}^{n} b_j e_j$$

folgt sofort

$$x_i = \frac{1}{\det A}\sum_{j=1}^{n}\det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n)b_j = \frac{1}{\det A}\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) \,.$$

□

Betrachtung 4.2.16.

1. Ähnliche Matrizen haben die gleiche Determinante. Denn gilt

$$\widetilde{A} = T \cdot A \cdot T^{-1}$$

mit $T \in GL(n, K)$, so ist nach dem Determinantenmultiplikationssatz 4.2.12

$$\det \widetilde{A} = \det T \cdot \det A \cdot \det T^{-1} = \det A \; .$$

2. Ist V ein n–dimensionaler K–Vektorraum und $\Phi : V \to V$ ein Endomorphismus, sind $\mathcal{B}$ und $\mathcal{B}'$ zwei geordnete Basen von V, so sind die beiden darstellenden Matrizen

$$M_{\mathcal{B}}^{\mathcal{B}}(\Phi) \quad \text{und} \quad M_{\mathcal{B}'}^{\mathcal{B}'}(\Phi)$$

ähnlich und haben somit die gleiche Determinante.

Definition 4.2.17
Für einen Endomorphismus $\Phi : V \to V$ eines endlich–dimensionalen Vektorraums V heißt

$$\det(\Phi) := \det\left(M_{\mathcal{B}}^{\mathcal{B}}(\Phi)\right)$$

Determinante von Φ, wobei $\mathcal{B}$ eine beliebige geordnete Basis von V ist.

Bemerkungen 4.2.18.

1. Für Endomorphismen $\Phi, \Psi : V \to V$ gilt:
   (a) $\det \Phi \neq 0 \Leftrightarrow \Phi$ ist Automorphismus.
   (b) $\det(\Phi^{-1}) = \frac{1}{\det \Phi}$ für alle Automorphismen Φ.
   (c) $\det(\Phi \circ \Psi) = \det(\Phi)\det(\Psi)$
2. Sei nun $V = \mathbb{R}^2$. Für eine Drehung $\Phi = R_\theta$ um den Ursprung ist

$$\det R_\theta = \det(M(R_\theta)) = \det \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \cos^2\theta + \sin^2\theta = 1 \; .$$

   Für eine Spiegelung $\Phi = S_\theta$ an einer Ursprungsgeraden ist

$$\det S_\theta = \det(M(S_\theta)) = \det \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} = -\cos^2 2\theta - \sin^2 2\theta = -1 \; .$$

4.3 Permutationen und Determinanten

Betrachtung 4.3.1.

- Wir erinnern an Beispiel 2.1.5.5: die Menge aller bijektiven Selbstabbildungen der Menge $\underline{n} := \{1, 2, \ldots, n\}$ bildet eine Gruppe, die symmetrische Gruppe S_n. Sie hat $|S_n| = n! = 1 \cdot 2 \cdot \ldots \cdot n$ Elemente und ist für $n \geq 3$ nicht abelsch.

- Für $\sigma \in S_n$ ist die Schreibweise

$$\sigma = \begin{bmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{bmatrix}$$

bequem. Das Produkt von $\sigma, \tau \in S_n$ schreiben wir als

$$\tau \cdot \sigma = \begin{bmatrix} 1 & \dots & n \\ \tau(1) & \dots & \tau(n) \end{bmatrix} \begin{bmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{bmatrix} = \begin{bmatrix} 1 & \dots & n \\ \tau(\sigma(1)) & \dots & \tau(\sigma(n)) \end{bmatrix}$$

Beispiel:

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} .$$

Definition 4.3.2

1. *Ein Element $\sigma \in S_n$ heißt auch Permutation.*
2. *Eine Permutation $\tau \in S_n$ heißt Transposition, falls τ zwei Elemente der Menge n vertauscht und alle übrigen Elemente fest lässt.*

Lemma 4.3.3.

1. Für jede Transposition $\tau \in S_n$ gilt $\tau^{-1} = \tau$.
2. Jede Permutation lässt sich als Produkt von Transpositionen schreiben, $\sigma = \tau_1 \dots \tau_k$. Diese Darstellung ist keinesfalls eindeutig!
3. Sei τ_0 die Transposition, die 1 und 2 vertauscht,

$$\tau_0 = \begin{bmatrix} 1 & 2 & 3 & \dots & n \\ 2 & 1 & 3 & \dots & n \end{bmatrix}$$

Dann ist jede andere Transposition $\tau \in S_n$ zu τ_0 konjugiert, d.h. es gibt ein $\sigma \in S_n$, so dass $\tau = \sigma\tau_0\sigma^{-1}$ gilt.

Beweis.
1. und 2. sind offensichtlich. Seien k, l die beiden verschiedenen von einer Transposition τ vertauschten Elemente. Wähle irgendein $\sigma \in S_n$, für das

$$\sigma(1) = k \qquad \text{und} \qquad \sigma(2) = l$$

gilt. Dann gilt

$$\sigma\tau_0\sigma^{-1}(k) = \sigma\tau_0(1) = \sigma(2) = l$$
$$\sigma\tau_0\sigma^{-1}(l) = \sigma\tau_0(2) = \sigma(1) = k$$

und für $i \neq k, l$ gilt $\sigma\tau_0\sigma^{-1}(i) = \sigma\sigma^{-1}(i) = i$. □

Betrachtung 4.3.4.

- Betrachte für eine Permutation $\sigma \in S_n$ die folgende quadratische Matrix

$$E_\sigma = \begin{pmatrix} (e_{\sigma^{-1}(1)})^t \\ \vdots \\ (e_{\sigma^{-1}(n)})^t \end{pmatrix} \in M(n \times n, K) \; .$$

  Da sie durch Vertauschungen von Zeilen aus der Einheitsmatrix $E_{\mathrm{id}} = E_n$ hervorgeht, ist

$$\det\,(E_\sigma) \in \{1, -1\} \; .$$

- Für die Matrixelemente gilt

$$(E_\sigma)_{ij} = \delta_{i,\sigma(j)}$$

  Es folgt für $\sigma, \tau \in S_n$

$$(E_\sigma E_\tau)_{ij} = \sum_{l=1}^{n} \delta_{i\sigma(l)}\delta_{l\tau(j)} = \delta_{i,\sigma(\tau(j))} = E_{\sigma\cdot\tau}$$

  Es folgt insbesondere

$$\det\,(E_{\sigma\tau}) = \det\,(E_\sigma E_\tau) = \det\,(E_\sigma)\det\,(E_\tau) \; .$$

Definition 4.3.5

1. *Die Abbildung*

$$\begin{aligned} \mathrm{sign}: \; & S_n \to \{\pm 1\} \\ & \sigma \mapsto \det E_\sigma =: \mathrm{sign}(\sigma) \end{aligned}$$

   heißt Signumsabbildung. Sie ist ein Gruppenhomomorphismus. Wegen Lemma 4.3.3.3 ist das Signum einer Transposition $\tau \in S_n$ gleich -1:

$$\mathrm{sign}(\tau) = \det E_\tau = \det\,(E_\sigma \cdot E_{\tau_0} \cdot E_{\sigma^{-1}}) = \det E_{\tau_0} = -1 \; .$$

2. *Der Kern der Signumsabbildung heißt alternierende Gruppe A_n:*

$$A_n := \{\sigma \in S_n | \, \mathrm{sign}(\sigma) = +1\}$$

   (Für $n \geq 1$ hat diese Gruppe $\frac{1}{2}n!$ Elemente und ist für $n \geq 4$ nicht abelsch.)

Satz 4.3.6. Leibniz'sche Regel
Sei K ein Körper und $A \in M(n \times n, K)$ mit $A = (a_{ij})$. Dann gilt

$$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdot \cdots \cdot a_{n\sigma(n)} \; .$$

Beweis.
Der i–te Zeilenvektor von A ist

$$(a_i)^t = \sum_{j=1}^{n} a_{ij} (e_j)^t \; .$$

Wir rechnen

$$\det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} \underset{(D1)}{=} \sum_{i_1=1}^{n} a_{1i_1} \det \begin{pmatrix} (e_{i_1})^t \\ (a_2)^t \\ \vdots \\ (a_n)^t \end{pmatrix}$$

$$\underset{(D1)}{=} \sum_{i_1,i_2=1}^{n} a_{1i_1}a_{2i_2} \det \begin{pmatrix} (e_{i_1})^t \\ (e_{i_2})^t \\ \vdots \\ (a_n)^t \end{pmatrix} \underset{(D1)}{=} \sum_{i_1 \ldots i_n=1}^{n} a_{1i_1} \cdot \ldots \cdot a_{ni_n} \det \begin{pmatrix} (e_{i_1})^t \\ (e_{i_2})^t \\ \vdots \\ (e_{i_n})^t \end{pmatrix}$$

Von den n^n Termen der Summe sind aber nur die $n!$ Terme nicht null, für die alle Indizes $i_1, i_2, \ldots, i_n$ paarweise verschieden sind. Dann gibt es $\sigma \in S_n$ mit $\sigma(j) = i_j$. Also

$$\begin{aligned} \det \begin{pmatrix} (a_1)^t \\ \vdots \\ (a_n)^t \end{pmatrix} &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \ldots a_{n\sigma(n)} \det \begin{pmatrix} (e_{\sigma(1)})^t \\ \vdots \\ (e_{\sigma(n)})^t \end{pmatrix} \\ &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \ldots a_{n\sigma(n)} \ \mathrm{sign}(\sigma) \ , \end{aligned}$$

wobei wir beachten, dass aus $\sigma\sigma^{-1} = 1$ die Gleichung $\mathrm{sign}(\sigma) = \mathrm{sign}(\sigma^{-1})$ folgt. □

Korollar 4.3.7.
Sei K ein Körper, $n_i \geq 1$ und $n := n_1 + n_2$. Sei $A \in M(n \times n, K)$ von der Gestalt

$$A = \begin{pmatrix} A^{(1)} & \vdots & * \\ \ldots & . & \ldots \\ 0 & \vdots & A^{(2)} \end{pmatrix}$$

mit $A^{(i)} \in M(n_i \times n_i, K)$. Dann ist

$$\det {}_{n_1+n_2} A = \det{}_{n_1} A^{(1)} \det{}_{n_2} A^{(2)} \ .$$

Beweis.

- In dem Ausdruck aus der Leibnizschen Regel 4.3.6

$$\det A = \sum_{\sigma \in S_n} a_{1\sigma(1)} \ldots a_{n\sigma(n)} \mathrm{sign}(\sigma)$$

  treten nur Terme auf, die von Permutationen σ kommen, für die gilt

$$\sigma(i) \in \{1, \ldots, n_1\} \Leftrightarrow i \in \{1, \ldots, n_1\}$$

- Betrachte die Einbettung von Gruppen:

$$\begin{aligned} \iota : \ & S_{n_1} \times S_{n_2} \hookrightarrow S_n \\ \iota(\sigma^{(1)}, \sigma^{(2)})(k) &= \begin{cases} \sigma^{(1)}(k) & \text{falls } k \leq n_1 \\ \sigma^{(2)}(k - n_1) + n_1 & \text{falls } k > n_1 \end{cases} . \end{aligned}$$

Aus

$$\iota\left(\sigma^{(1)},\sigma^{(2)}\right)=\iota\left(\sigma^{(1)},1\right)\cdot\iota\left(1,\sigma^{(2)}\right)$$

folgt

$$\operatorname{sign}\iota\left(\sigma^{(1)},\sigma^{(2)}\right)=\operatorname{sign}\iota\left(\sigma^{(1)},1\right)\cdot\operatorname{sign}\iota\left(1,\sigma^{(2)}\right)=\operatorname{sign}\sigma^{(1)}\cdot\operatorname{sign}\sigma^{(2)}\ .$$

- Damit folgt

$$\begin{aligned}\det A&=\sum_{\sigma_1\in S_{n_1}}a^{(1)}_{1\sigma^{(1)}(1)}\dots a^{(1)}_{n_1\sigma^{(1)}(n_1)}\operatorname{sign}\left(\sigma^{(1)}\right)\cdot\sum_{\sigma_2\in S_{n_2}}a^{(2)}_{1\sigma^{(2)}(1)}\dots a^{(2)}_{n_2\sigma^{(2)}(n_2)}\operatorname{sign}\left(\sigma^{(2)}\right)\\&=\det A^{(1)}\cdot\det A^{(2)}\ .\end{aligned}$$

□

Bemerkungen 4.3.8.

1. Wir wollen den Rechenaufwand für die Berechnung der Determinante einer $n\times n$-Matrix an Hand der Zahl der nötigen Multiplikationen einmal grob überschlagen. Bei der Leibnizschen Regel 4.3.6 oder den Entwicklungssätzen 4.2.7 bzw. 4.2.11 hat man sicher mehr als $n!$ Multiplikationen auszuführen.

   Beim Gauß-Algorithmus braucht man für die Elimination unterhalb der i-ten Zeile für jede Zeile eine Division zur Berechnung des Eliminationsfaktors und dann $n-i$ Multiplikationen in der Zeile, die verändert wird. Bei $n-i$ Zeilen ist der Aufwand

   $$(n-i)(n-i+1)=(n-i)^2+(n-i)$$

   Multiplikationen. Der Gesamtaufwand ist dann

   $$\sum_{i=1}^{n-1}\left((n-i)^2+(n-i)\right)=\sum_{i=1}^{n-1}(i^2-i)\sim\frac{n^3}{3}$$

   Multiplikationen.

2. Sie werden in einer Übungsaufgabe für das Signum einer beliebigen Permutation $\sigma\in S_n$ die Formel

   $$\operatorname{sign}(\sigma)=\prod_{1\le i<j\le n}\frac{\sigma(j)-\sigma(i)}{j-i}\qquad\qquad(*)$$

   zeigen. Ein Fehlstand einer Permutation σ ist ein Paar (i,j) mit $i<j$, für das $\sigma(i)>\sigma(j)$ gilt. Man kann jeder Permutation $\sigma\in S_n$ die Menge $\operatorname{inv}(\sigma)$ der Fehlstände zuordnen. Zum Beispiel hat die Permutation

   $$\sigma=\begin{pmatrix}1&2&3&4&5\\3&5&1&2&4\end{pmatrix}\in S_5$$

   die Menge der Fehlstände:

   $$\operatorname{inv}(\sigma)=\{(1,3),(2,3),(1,4),(2,4),(2,5)\}\ .$$

   Aus der Formel $(*)$ für das signum folgt sofort, dass das Signum genau dann -1 ist, wenn die Anzahl der Fehlstände ungerade ist. Dies ist in obigem Beispiel der Fall.

4.4 Orientierungen und Volumina

Wir wollen nun noch eine geometrische Anwendung der Determinante kennenlernen. Dazu arbeiten wir in diesem Unterkapitel über dem Körper $\mathbb{R}$ der reellen Zahlen, der bekanntlich ein angeordneter Körper ist.

Definition 4.4.1
Sei V ein endlich-dimensionaler reeller Vektorraum. (In der Folge sei stets $\dim_{\mathbb{R}} V \geq 1$ angenommen.) Zwei geordnete Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$ von V heißen gleich orientiert, falls für die Transformationsmatrix

$$\det T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} > 0$$

gilt. Andernfalls heißen die geordneten Basen entgegengesetzt orientiert.

Lemma 4.4.2.
Die Beziehung "gleich orientiert" liefert eine Äquivalenzrelation auf der Menge aller geordneten Basen eines gegebenen Vektorraums V.

Beweis.

- Reflexivität:
$$\det T_{\mathcal{B}}^{\mathcal{B}} = \det E_n = 1 > 0 \,.$$
- Symmetrie: für zwei geordnete Basen $\mathcal{B}, \widetilde{\mathcal{B}}$ von V gilt
$$T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} \cdot T_{\mathcal{B}}^{\widetilde{\mathcal{B}}} = E_n$$
Insbesondere folgt
$$\det T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} \cdot \det T_{\mathcal{B}}^{\widetilde{\mathcal{B}}} = 1 \,,$$
so dass beide Determinanten ungleich Null sind und gleiches Vorzeichen haben.
- Transitivität folgt aus Satz 3.6.2
$$\det T_{\mathcal{B}_3}^{\mathcal{B}_1} = \det \left(T_{\mathcal{B}_3}^{\mathcal{B}_2} T_{\mathcal{B}_2}^{\mathcal{B}_1}\right) = \det \left(T_{\mathcal{B}_3}^{\mathcal{B}_2}\right) \det \left(T_{\mathcal{B}_2}^{\mathcal{B}_1}\right) > 0 \,,$$
wenn die geordenten Basen $\mathcal{B}_1$, $\mathcal{B}_2$ und $\mathcal{B}_2$, $\mathcal{B}_3$ von V jeweils paarweise gleich orientiert sind.

□

Definition 4.4.3
Sei V ein endlich dimensionaler reeller Vektorraum. Eine Äquivalenzklasse von geordneten Basen bezüglich der Äquivalenzrelation 4.4.2 heißt eine Orientierung von V.

Lemma 4.4.4.
Ein endlich–dimensionaler reeller Vektorraum besitzt genau zwei Orientierungen.

Beweis.

- Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine geordnete Basis, setze $\widetilde{\mathcal{B}} = (-b_1, \ldots, b_n)$. Wegen

$$\det T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} = \det \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & & & 1 \end{pmatrix} = -1$$

  definieren die geordneten Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$ unterschiedliche Orientierungen. Es gibt also mindestens zwei unterschiedliche Orientierungen auf einem endlich-dimensionalen reellen Vektorraum.

- Es mögen die geordneten Basen $\mathcal{B}_1, \mathcal{B}_2$ und $\mathcal{B}_2, \mathcal{B}_3$ unterschiedliche Orientierungen besitzen:

$$\det T_{\mathcal{B}_2}^{\mathcal{B}_1} < 0 \quad \text{und} \quad \det T_{\mathcal{B}_3}^{\mathcal{B}_2} < 0 \, .$$

  Dann folgt, wiederum wegen Satz 3.6.2,

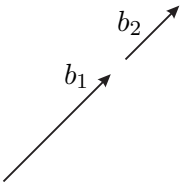
$$\det T_{\mathcal{B}_3}^{\mathcal{B}_1} = \det T_{\mathcal{B}_3}^{\mathcal{B}_2} \cdot \det T_{\mathcal{B}_2}^{\mathcal{B}_1} > 0 \, ,$$

  also haben $\mathcal{B}_1$ und $\mathcal{B}_3$ gleiche Orientierung.

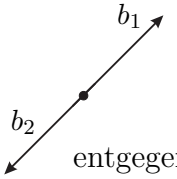
□

Beispiele 4.4.5.

1. $\dim V = 1$

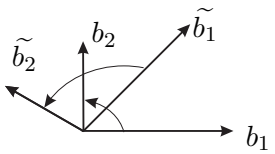


gleich orientiert



entgegengesetzt orientiert.

2. $\dim V = 2$: Hier kommt es auf den "Drehsinn der Basis" an: führt man den ersten Basisvektor durch die Drehung um einen betragsmäßig möglichst kleinen Winkel in den zweiten Basisvektor über, so sind die Basen genau dann gleich orientiert, wenn der Drehsinn dieser Drehungen gleich ist. Zum Beispiel haben



die gleiche Orientierung.

3. $\dim V = 3$: Analog ist hier die "Händigkeit der Basis" entscheidend.

4. Zwei Basen sind genau dann gleich orientiert, wenn sie sich stetig ineinander überführen lassen.

5. Einen Automorphismus $\Phi : V \to V$ eines endlich–dimensionalen reellen Vektorraums nennen wir orientierungserhaltend oder orientierungstreu, falls $\det(\Phi) > 0$ gilt. Die orientierungstreuen Automorphismen von V bilden wegen des Determinantenmultiplikationssatze 4.2.12 eine Untergruppe der Automorphismengruppe von V.

Definition 4.4.6

1. *Mit $\mathcal{K}_n$ bezeichnen wir die Menge aller kompakten Teilmengen des $\mathbb{R}^n$,*

$$\mathcal{K}_n = \{X \subset \mathbb{R}^n |\, X \text{ ist kompakt } \}\ .$$

Dies sind genau die beschränkten abgeschlossenen Teilmengen des $\mathbb{R}^n$.

2. *Das n–dimensionale Volumen ist eine Abbildung*

$$\mathrm{vol}_n :\ \mathcal{K}_n \to \mathbb{R}_{\geq 0}$$

mit den folgenden Eigenschaften:

(V1) Normierung: $\mathrm{vol}_n(W^n) = 1$, *wobei*

$$W^n := \{x \in \mathbb{R}^n |\, 0 \leq x_i \leq 1\}$$

der n–dimensionale Einheitswürfel ist.

(V2) Monotonie: $X \subset Y \Rightarrow \mathrm{vol}_n(X) \leq \mathrm{vol}_n(Y)$

(V3) Zerlegungseigenschaft: Es gilt $\mathrm{vol}_n(X \cup Y) = \mathrm{vol}_n(X) + \mathrm{vol}_n(Y) - \mathrm{vol}_n(X \cap Y)$.

(V4) Kovarianz unter affinen Abbildungen: Ist

$$\begin{aligned} F :\ \mathbb{R}^n &\to \mathbb{R}^n \\ F(x) &= Ax + b \end{aligned}$$

eine affine Abbildung, so gilt

$$\mathrm{vol}_n\left(F(X)\right) = |\det A|\ \ \mathrm{vol}_n(X)\ .$$

Dass die Determinante in Volumina auftritt, ist eine wichtige Rolle der Determinante. Wir zeigen hier nicht die Existenz einer Volumensfunktion, die Gegenstand der Analysis oder Maßtheorie ist.

Beispiel 4.4.7.
Sei $(b_1, \ldots, b_n)$ eine geordnete Basis des $\mathbb{R}^n$, $q \in \mathbb{R}^n$ und

$$X = \left\{ q + \sum_{i=1}^{n} t_i b_i |\, 0 \leq t_i \leq 1 \right\}$$

ein n–dimensionales Parallelotop, vgl. Definition 4.1.5. Betrachte die affine Abbildung

$$F(x) = Ax + q$$

mit der quadratischen Matrix $A = (b_1, \ldots, b_n)$, die es uns erlaubt, X als Bild des Standardwürfels W^n zu schreiben, $F(W^n) = X$. Somit ist wegen der Kovarianz (V4)

$$\mathrm{vol}_n(X) = \mathrm{vol}_n\left(F(W^n)\right) = |\det A| \mathrm{vol}_n(W^n) = |\det\left(b_1, \ldots, b_n\right)|\ .$$

Wir betrachten noch zwei Spezialfälle:

- In Dimension $n = 3$ finden wir $\mathrm{vol}(X) = |\langle b_1 \times b_2, b_3 \rangle|$, wie schon aus Satz 4.1.6 bekannt.
- Der n-dimensionale Quader

$$Q = \left\{ q + \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \mid \ 0 \le t_i \le a_i \right\} = \left\{ q + \sum t_i' a_i e_i \mid \ 0 \le t_i' \le 1 \right\}$$

mit $a_i > 0$ für $i = 1, \ldots, n$, hat das Volumen

$$\mathrm{vol}(Q) = \left| \det \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \right| = a_1 \cdot \cdots \cdot a_n \ .$$

Lemma 4.4.8.
Ist $X \in K_n$ in einem $(n-1)$– dimensionalen affinen Unterraum enthalten, so ist $\mathrm{vol}_n(X) = 0$.

Beweis.

- Sei $V \subset \mathbb{R}^n$ affiner Unterraum, $\dim V = n-1$ und $X \subset V$. Nach (V4) ändert eine Translation das Volumen nicht. Also können wir ohne Beschränkung der Allgemeinheit $0 \in V$ annehmen, d.h. V sei sogar Untervektorraum.
- Wähle eine lineare Abbildung $\Phi : \ \mathbb{R}^n \to \mathbb{R}^n$ mit $\mathrm{Im}\,\Phi = V$ und $\Phi|_V = \mathrm{id}_V$. Um zu sehen, dass solch eine lineare Abbildung existiert, wähle eine Basis $\{b_1, \ldots, b_{n-1}\}$ von V und ergänze sie zu einer Basis von $\mathbb{R}^n$. Betrachte dann

$$M(\Phi) = \begin{pmatrix} E_{n-1} & \vdots & 0 \\ \ldots & . & \ldots \\ 0 & \vdots & 0 \end{pmatrix} \in M(n \times n, \mathbb{R})$$

Es gilt

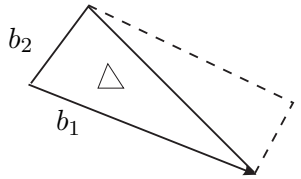
$$\mathrm{vol}_n(X) = \mathrm{vol}_n \Phi(X) \overset{(\mathrm{V3})}{=} |\det \Phi| \cdot \mathrm{vol}_n(X) = 0 \ ,$$

denn es ist $\mathrm{rg}\,(\Phi) = n-1$ und nach Lemma 4.2.3 daher $\det \Phi = 0$.

□

Beispiele 4.4.9.

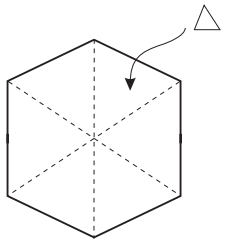
1. Dreieck: ergänze das Dreieck $\triangle$ durch ein Dreieck $\triangle'$ zu einem Parallelogramm:



$$|\det(b_1, b_2)| = \mathrm{vol}_2(\triangle \cup \triangle') \underset{(V3)}{=} \mathrm{vol}_2(\triangle) + \mathrm{vol}_2(\triangle') - \mathrm{vol}_2(\triangle \cap \triangle') \underset{(4.4.8)}{=} 2\mathrm{vol}_2(\triangle)$$

Also $\mathrm{vol}_2(\triangle) = \frac{1}{2} |\det(b_1, b_2)|$ Dies liefert uns mit Betrachtung 4.1.4 die bekannte Regel: Dreiecksfläche $= \frac{1}{2}$ Grundseite $\times$ zugehörige Höhe.

2. Regelmäßiges n–Eck $E_n(r)$ mit Radius r



$$\mathrm{vol}_2(\triangle) = \tfrac{r^2}{2}\sin\tfrac{2\pi}{n}\ .$$
$$\mathrm{vol}_2\left(E_n(r)\right) = \tfrac{nr^2}{2}\sin\tfrac{2\pi}{n}\ .$$

3. Kreisscheibe mit Radius r: $D(r)$. Wir leiten zunächst eine untere Schranke für die Fläche aus dem eingeschriebenen n-Eck $E_n(r)$ her. Aus $E_n(r) \subset D(r)$ für alle $n \in N$ folgt

$$\mathrm{vol}_2\left(D(r)\right) \geq \mathrm{vol}\left(E_n(r)\right) = \frac{nr^2}{2}\sin\frac{2\pi}{n}\ .$$

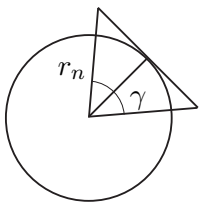
Also

$$\mathrm{vol}_2\left(D(r)\right) \geq \lim_{n\to\infty}\left(\frac{nr^2}{2}\sin\frac{2\pi}{n}\right)\ .$$

Um den Grenzwert auszurechnen, substituiere $x = \frac{2\pi}{n}$ und finde

$$\mathrm{vol}_2\left(D(r)\right) \geq \lim_{x\to 0}\left(\frac{2\pi}{x}\frac{r^2}{2}\sin x\right) = \pi r^2 \lim_{x\to 0}\frac{\sin x}{x} = \pi r^2\ .$$

Analog leiten wir eine obere Schranke durch umschriebene regelmäßige n-Ecke her. Wir betrachten $D(r) \subset E_n(r_n)$, wobei die Dreiecke aus (ii) nun Höhe r haben



$$r = r_n\cos\frac{\gamma}{2} = r_n\cos\frac{\pi}{n}\ .$$

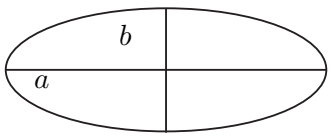
$$\mathrm{vol}_2\left(D(r)\right) \leq \mathrm{vol}_2\left(E_n(r_n)\right) = \frac{nr_n^2}{2}\sin\frac{2\pi}{n} = \frac{nr^2}{2\cos^2\frac{\pi}{n}}\sin\frac{2\pi}{n}\ .$$

Analog zur obigen Rechnung sieht man

$$\mathrm{vol}_2 D(r) \leq \lim_{x\to 0}\pi r^2\frac{1}{\cos^2\frac{x}{2}}\frac{\sin x}{x} = \pi r^2$$

Also ist das Volumen der Kreisscheibe gleich $\mathrm{vol}_2 D(r) = \pi r^2$.

4. Ellipse $E(a,b)$ mit halber Hauptachse a und halber Nebenachse b:



$E(a,b) = \Phi\left(D(1)\right)$ mit

$$M(\Phi) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

Also folgt aus dem Axiom $(V4)$:

$$\mathrm{vol}_2 E(a,b) = \left|\det\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\right| \cdot \mathrm{vol}_2 D(1) = a\cdot b\cdot\pi\ .$$

4.5 Minoren

Wir wissen schon aus Lemma 4.2.3, dass eine quadratische Matrix von maximalem Rang dadurch charakterisiert werden kann, dass ihre Determinante nicht verschwindet. Man kann auch einen nicht-maximalen Rang durch die Berechnung von (mehreren) Determinanten bestimmen. Hierfür betrachten wir nicht nur quadratische Matrizen.

Definition 4.5.1
Ist $A \in M(m \times n, K)$ und $k \leq \min(m, n)$, so heißt eine $k \times k$-Matrix A', die durch Streichen von $m - k$ Zeilen und $n - k$ Spalten aus A hervorgeht, eine k-reihige Untermatrix von A. Ihre Determinante $\det A' \in K$ *heißt ein k-reihiger Minor der Matrix A.*

Satz 4.5.2.
Sei $A \in M(m \times n, K)$ und $r \in \mathbb{N}$. Dann sind die folgenden Bedingungen äquivalent:

(i) $r = \operatorname{rg}(A)$.

(ii) Es gibt einen r-reihigen Minor ungleich Null, und für $k > r$ ist jeder k-reihige Minor gleich Null.

Beweis.
Wir zeigen zum Beweis, dass die folgenden beiden Bedingungen äquivalent sind:

(a) $\operatorname{rg}(A) \geq k$

(b) Es gibt eine k-reihige Untermatrix A' von A mit $\det A' \neq 0$.

Wir zeigen (b)⇒ (a): aus $\det A' \neq 0$ folgt nach Lemma 4.2.3 $\operatorname{rg}(A') = k$, und daraus $\operatorname{rg}(A) \geq k$, da der Rang einer Untermatrix durch den Rang der Matrix nach oben beschränkt ist. Denn die lineare Abhängigkeit von Zeilen (oder Spalten) der Matrix impliziert, dass auch die entsprechenden Zeilen (oder Spalten) der Untermatrix linear abhängig sind.

Um (a)⇒ (b) zu sehen, beachten wir, dass es wegen $\operatorname{rg}(A) \geq k$ sicher k linear unabhängige Zeilenvektoren von A gibt. Wir wählen k solche Zeilen aus; für die dadurch erhaltene rechteckige Matrix ist nach Satz 3.6.15 der Zeilenrang gleich dem Spaltenrang. Wir finden also k linear unabhängige Spalten dieser Untermatrix, die wir auswählen, so dass wir eine $k \times k$-Teilmatrix A' erhalten, die maximalen Rang k und somit nach Lemma 4.2.3 nicht-verschwindende Determinante hat. □

Wir beweisen schließlich noch eine Verallgemeinerung des Multiplikationssatzes 4.2.12 für Determinanten. Dafür drücken wir $A \in M(m \times n, K)$ mit $m \leq n$ durch die n Spaltenvektoren $a_1, a_2, \ldots, a_n \in K^m$ aus, $A = (a_1 \ldots, a_n)$. Wir schreiben $A^{k_1,\ldots,k_m} = (a_{k_1}, \ldots a_{k_m})$ für die m-reihige Untermatrix, deren erster Spaltenvektor der k_1-te Spaltenvektor von A ist, usw.

Satz 4.5.3 (Verallgemeinerter Determinantenmultiplikationssatz)**.**
Sei $m \leq n$; dann gilt für alle Matrizen $A, B \in M(m \times n, K)$

$$\det(AB^t) = \sum_{1 \leq k_1 < \ldots < k_m \leq n} \det(A^{k_1,\ldots,k_m}) \cdot \det(B^{k_1,\ldots,k_m}).$$

Beweis.

- Behauptung: Die Aussage gilt für den Fall, wenn die Zeilenvektoren von A Transponierte der Elemente der Standardbasis von K^n sind:

$$A = \begin{pmatrix} (e_{j_1})^t \\ (e_{j_2})^t \\ \vdots \\ (e_{j_m})^t \end{pmatrix} .$$

  Denn für $1 \leq k_1 < \ldots < k_m \leq n$ ist der Minor $\det A^{k_1,\ldots,k_m}$ genau dann von Null verschieden, wenn es eine Permutation $\sigma \in S_m$ gibt mit $j_i = k_{\sigma(i)}$. Dann ist der Minor nach Definition 4.3.5 gleich signσ. Andererseits gilt mit *Spalten*vektoren b^i von B

$$A \cdot B^t = \begin{pmatrix} (b_{j_1})^t \\ \vdots \\ (b_{j_m})^t \end{pmatrix} .$$

  Die Determinante dieses Produkts ist nur dann ungleich Null, wenn die Indizes paarweise verschieden sind. Dann gibt es eine Permutation $\sigma \in S_m$ mit $j_i = k_{\sigma(i)}$, und es gilt

$$\det AB^t = \text{sign}\sigma \cdot \det B^{k_1,\ldots,k_m} .$$

- Gilt die Aussage für eine Matrix A und entsteht die Matrix $\tilde{A}$ aus A durch Multiplikation der i-ten Zeile mit $\lambda \in K$, so gilt die Aussage auch für $\tilde{A}$, denn der Wert beider Seiten der Gleichung muss mit λ multipliziert werden.

- Aus der Linearität der Determinantenfunktion folgt ebenso: gilt die Aussage für eine Matrix A mit i-tem Zeilenvektor $(a_i)^t$ und eine Matrix $\tilde{A}$, die sich von A nur in der i-ten Zeile darin unterscheidet, dass der Spaltenvektor gleich $(\tilde{a}_i)^t$ ist, so gilt sie auch für die Matrix $A + \tilde{A}$.

- Man kann jede beliebige Matrix A nun mit Hilfe der vorangehenden Schritte aufbauen: fixiere beliebige $j_2, \ldots j_m \in \{1, \ldots, n\}$. Wähle $\lambda_{j_1} = a_{1j_1}$, so folgt die Aussage für Matrizen der Form

$$A_1 := \begin{pmatrix} (a_1)^t \\ (e_{j_2})^t \\ \vdots \\ (e_{j_m})^t \end{pmatrix} .$$

  Nun halte den Zeilenvektor $(a_1)^t$ fest und lasse j_2 von 1 bis n laufen. Es folgt die Behauptung für Matrizen der Form

$$A_2 := \begin{pmatrix} (a_1)^t \\ (a_2)^t \\ \vdots \\ (e_{j_m})^t \end{pmatrix} .$$

  Nach m solchen Schritten erhält man die Behauptung für die Matrix A.

□

Bemerkungen 4.5.4.

1. Der Fall $m > n$ braucht keine Formel: es ist $\operatorname{rg} A \leq n$ und $\operatorname{rg} B = \operatorname{rg} B^t \leq n$, somit

$$\operatorname{rg}(AB^t) \leq \min\{\operatorname{rg} A, \operatorname{rg} B\} \leq n < m$$

   Damit ist der Rang der $m \times m$-Matrix AB^t nicht maximal. Es gilt im Fall $m > n$ also immer $\det AB^t = 0$.

2. Man beachte, dass es im Fall $m \leq n$ genau $\begin{pmatrix} n \\ m \end{pmatrix}$ m-reihige Minoren gibt. Insbesondere hat für $m = n$, also für quadratische Matrizen, die Summe nur einen Term, und wir erhalten einen neuen Beweis des Multiplikationssatzes 4.2.12 für Determinanten.

3. Insbesondere gilt für jede Matrix $A \in M(m \times n, K)$ mit $m \leq n$

$$\det(AA^t) = \sum_{1 \leq k_1 < \ldots < k_m \leq n} \left(\det(A^{k_1,\ldots,k_m)}\right)^2 .$$

   Man nennt $\det(AA^t)$ eine Gramsche Determinante. Sie ist für Matrizen mit reellen Einträgen als Summe von Quadraten reeller Zahlen offensichtlich nicht-negativ. Sie ist genau dann gleich Null, wenn $\operatorname{rg} A < m$. Denn $\operatorname{rg} A < m \leq n$ ist nach Satz 4.5.2 äquivalent dazu, dass jeder m-reihige Minor gleich Null ist, was zu $\det(AA^t) = 0$ äquivalent ist. Gramsche Determinanten treten bei der Integration über Untermannigfaltigkeiten auf.

5 Eigenwerte

5.1 Definitionen

Die Klassen äquivalenter $m \times n$ Matrizen kennen wir aus Bemerkung 3.6.8.3: jede Matrix $A \in M(m \times n, K)$ ist äquivalent zu genau einer Matrix der Form

$$\begin{pmatrix} E_r & \vdots & 0 \\ \dots & . & \dots \\ 0 & \vdots & 0 \end{pmatrix} \qquad \text{mit } \; r = \operatorname{rg}(A) \; .$$

Dies erlaubt es uns, durch Wahl geeigneter Basen für V und für W eine besonders einfache Beschreibung einer gegebenen linearen Abbildung $\Phi : V \to W$ mit $\dim V = n$ und $\dim W = m$ zu finden.

Wir wollen in diesem Kapitel die Grundlagen für eine Beschreibung der Ähnlichkeitsklassen quadratischer Matrizen legen. Dies geht Hand in Hand mit dem Verständnis der Frage, auf welche Form die darstellende Matrix eines Endomorphismus eines endlich-dimensionalen Vektorraums durch geschickte Basiswahl gebracht werden kann. Entsprechend werden wir ab sofort frei zwischen der Sprache der (quadratischen) Matrizen und der linearen (Selbst-)Abbildungen endlich-dimensionaler Vektorräume wechseln.

Definition 5.1.1
Sei K ein Körper.

1. *Sei V ein K–Vektorraum und $\Phi : V \to V$ ein Endomorphismus. Ein Element $\lambda \in K$ heißt Eigenwert von Φ, falls es ein $v \in V \setminus \{0\}$ gibt, so dass $\Phi(v) = \lambda v$ gilt. Dann heißt v Eigenvektor von Φ zum Eigenwert λ.*
2. *Sei $A \in M(n \times n, K)$. Ein Element $\lambda \in K$ heißt Eigenwert von A, falls es ein $v \in K^n \setminus \{0\}$ gibt mit $A \cdot v = \lambda v$. Dann heißt v Eigenvektor von A zum Eigenwert λ.*
3. *Ein Endomorphismus $\Phi : V \to V$ heißt diagonalisierbar, falls es eine Basis $\mathcal{B}$ von V gibt, die nur aus Eigenvektoren von Φ besteht. Ist V endlich-dimensional, so ist danndie darstellende Matrix $M_{\mathcal{B}}^{\mathcal{B}}(\Phi)$ bezüglich jeder Ordnung dieser Basis $\mathcal{B}$ eine Diagonalmatrix.*

Satz 5.1.2.
Sei $\Phi : V \to V$ ein Endomorphismus. Seien $v_1, \dots, v_m$ Eigenvektoren von Φ zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m$. Dann ist die Familie $(v_1, \dots, v_m)$ linear unabhängig.

Beweis.
Wir benutzen vollständige Induktion nach m. Für $m = 1$ ist $v_1 \neq 0$ per Definition eines Eigenvektors klar. Gelte

$$0 = \sum_{i=1}^{m} \alpha_i v_i \qquad \text{mit} \qquad \alpha_i \in K \; . \qquad (*)$$

Es folgt

$$0 = \Phi(0) = \sum_{i=1}^{m} \alpha_i \Phi(v_i) = \sum_{i=1}^{m} \alpha_i \lambda_i v_i \; .$$

Die Multiplikation von $(*)$ mit λ_1 liefert:

$$0 = \sum_{i=1}^{m} \alpha_i \lambda_1 v_i \ .$$

Die Differenz der Gleichungen ist

$$\sum_{i=2}^{m} \alpha_i (\lambda_i - \lambda_1) v_i = 0 \ .$$

Nach Induktionsannahme ist die Familie $(v_2, \ldots, v_m)$ linear unabhängig, also $\alpha_i(\lambda_i - \lambda_1) = 0$ für alle $i = 2, \ldots, m$. Wegen $\lambda_i \neq \lambda_1$ folgt $\alpha_2 = \alpha_3 = \ldots = \alpha_m = 0$. Da $v_1 \neq 0$ ist, folgt auch $\alpha_1 = 0$. □

Korollar 5.1.3.

1. Ist $n := \dim_K V < \infty$, so hat jeder Endomorphismus $\Phi : V \to V$ höchstens n verschiedene Eigenwerte.

2. Ist $n := \dim_K V < \infty$ und hat Φ genau n verschiedene Eigenwerte, so ist Φ diagonalisierbar.

Beweis.

1. ist klar, weil jede Familie von mehr als n Vektoren linear abhängig ist.

2. Wähle zu jedem Eigenwert λ_i einen Eigenvektor v_i. Die Familie $(v_i)_{i=1\ldots n}$ ist nach Satz 5.1.2 linear unabhängig und wegen $\dim_K V = n$ eine Basis.

□

Definition 5.1.4
$\Phi : V \to V$ ein Endomorphismus eines K–Vektorraums V. Für $\lambda \in K$ heißt

$$\mathrm{Eig}(\Phi, \lambda) := \{v \in V | \ \Phi(v) = \lambda v\}$$

der Eigenraum von Φ zum Wert λ.

Bemerkungen 5.1.5.

1. Die Menge $\mathrm{Eig}(\Phi, \lambda) \setminus \{0\}$ besteht aus den Eigenvektoren zum Eigenwert λ. Ein Element $\lambda \in K$ ist also genau dann Eigenwert, wenn $\mathrm{Eig}(\Phi, \lambda) \neq \{0\}$.

2. $\mathrm{Eig}(\Phi, \lambda) = \ker(\lambda \mathrm{id}_V - \Phi)$, denn $\Phi(v) = \lambda v$ genau dann, wenn $(\lambda \mathrm{id}_V - \Phi)v = 0$. Insbesondere ist $\mathrm{Eig}(\Phi, \lambda)$ als Kern ein Untervektorraum von V.

3. Aus $\lambda_1 \neq \lambda_2$ folgt $\mathrm{Eig}(\Phi, \lambda_1) \cap \mathrm{Eig}(\Phi, \lambda_2) = \{0\}$. Denn für $v \in \mathrm{Eig}(\Phi, \lambda_1 \cap \mathrm{Eig}(\Phi, \lambda_2))$ folgt
$$\lambda_1 v = \Phi(v) = \lambda_2 v \ ,$$
also $(\lambda_1 - \lambda_2)v = 0$, woraus $v = 0$ wegen $\lambda_1 \neq \lambda_2$ folgt.

4. Aus 2. folgt sofort: λ ist genau dann Eigenwert, wenn die Abbildung $\lambda \cdot \mathrm{id}_V - \Phi$ nicht bijektiv ist. Dies ist genau dann der Fall, wenn

$$\det(\lambda \mathrm{id}_V - \Phi) = 0$$

gilt.

Definition 5.1.6
Ist

$$\mu_{geo}(\Phi, \lambda) := \dim_K \mathrm{Eig}(\Phi, \lambda)$$

nicht verschwindend, so heißt μ_{geo} geometrische Vielfachheit des Eigenwerts λ.

Definition 5.1.7
Die Funktion

$$\begin{aligned} P_\Phi : \quad & K \to K \\ & \lambda \mapsto \det(\lambda \mathrm{id}_V - \Phi) \end{aligned}$$

heißt das charakteristisches Polynom von Φ. Korrekter wäre hier die Bezeichnung charakteristische Funktion.

Bemerkungen 5.1.8.

1. Die Eigenwerte von Φ sind genau die Nullstellen von P_Φ.

2. Sei $\mathcal{B}$ eine geordnete Basis von V und

$$A = M_{\mathcal{B}}^{\mathcal{B}}(\Phi) \in M(n \times n, K) \ .$$

Es ist

$$M_{\mathcal{B}}^{\mathcal{B}}(\lambda \mathrm{id}_V - \Phi) = \lambda E_n - A \ ,$$

woraus folgt:

$$P_\Phi(\lambda) = \det(\lambda \mathrm{id}_V - \Phi) = \det(\lambda E_n - A) =: P_A(\lambda) \ .$$

Insbesondere haben ähnliche Matrizen das gleiche charakteristische Polynom, da sie den gleichen Endomorphismus bezüglich verschiedener Basen darstellen, vergleiche Lemma 3.6.9.

3. Beispiel: Drehungen des $\mathbb{R}^2$ um den Ursprung:

$$A = M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$P_A(\lambda) = \det \begin{pmatrix} \lambda - \cos\theta & \sin\theta \\ -\sin\theta & \lambda - \cos\theta \end{pmatrix} = \lambda^2 - 2\cos\theta\lambda + 1$$

Die komplexen Nullstellen sind

$$\lambda_{1,2} = \cos\theta \pm \mathrm{i}\sin\theta \ .$$

Diese sind nur für $\theta = 0, \pi$ reell. Nur dann gibt es reelle Eigenwerte und auch Eigenvektoren in $\mathbb{R}^2$. Bei $\theta = 0$ handelt es sich um die Identität, alle Vektoren sind Eigenvektoren zum Eigenwert 1; bei $\theta = \pi$ handelt es sich um die Punktspiegelung am Ursprung, alle Vektoren sind Eigenvektoren zum Eigenwert -1.

Satz 5.1.9.
Sei $A \in M(n \times n, K)$. Dann ist $P_A(\lambda)$ eine polynomiale Funktion vom Grad n,

$$P_A(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0 \qquad \text{mit } \; a_i \in K$$

und

$$a_n = 1, \quad a_{n-1} = -(a_{11} + a_{22} + \ldots + a_{nn}) \quad \text{und} \quad a_0 = (-1)^n \det A \; .$$

Beweis.

- Wir zeigen zunächst:

  $$P_A(\lambda) = (\lambda - a_{11}) \ldots (\lambda - a_{nn}) + Q_A(\lambda) \; ,$$

  wobei $Q_A(\lambda)$ ein Polynom vom Grad höchstens $n-2$ ist. Hierfür verwenden wir vollständige Induktion nach n:

  $$n = 1 : \; P_A(\lambda) = \det(\lambda - a_{11}) = \lambda - a_{11} \; .$$

  Induktionsschritt: Entwicklung nach der ersten Spalte:

  $$\begin{aligned} P_A(\lambda) &= (\lambda - a_{11}) \det \begin{pmatrix} \lambda - a_{22} & \ldots & -a_{2n} \\ \vdots & & \vdots \\ -a_{n2} & \ldots & \lambda - a_{nn} \end{pmatrix} \\ &\quad + \sum_{j=2}^{n} \underbrace{(-1)^{j+1}(-a_{j1}) \det\left((\lambda E_n - A)^{Str}_{1j}\right)}_{\text{Polynom der Ordnung } \leq n-2} \\ &= (\lambda - a_{11}) P_{A^{Str}_{11}}(\lambda) + \mathcal{O}(n-2) \, , \end{aligned}$$

  wobei $\mathcal{O}(n-2)$ ein Polynom der Ordnung höchstens $n-2$ ist. Hieraus folgt die Behauptung.

- Wir entwickeln daher

  $$\begin{aligned} P_A(\lambda) &= (\lambda - a_{11}) \ldots (\lambda - a_{nn}) + Q_A(\lambda) \\ &= \lambda^n - \sum_{i=1}^{n} a_{ii} \lambda^{n-1} + \ldots \end{aligned}$$

  woraus die Formeln für die Koeffizienten a_n und a_{n-1} des charakteristischen Polynoms folgen.

- Schließlich beachten wir

  $$a_0 = P_A(0) = \det(0 \cdot \mathrm{id}_V - A) = (-1)^n \det A \; .$$

□

Definition 5.1.10
Die Abbildung

$$\begin{aligned} \mathrm{Tr} \; : \; M(n \times n, K) &\to K \\ A &\mapsto \sum_{i=1}^{n} a_{ii} \end{aligned}$$

heißt Spur (englisch: trace).

Bemerkungen 5.1.11.

1. Ähnliche Matrizen haben nach Bemerkung 5.1.8.2 dasselbe charakteristische Polynom, also insbesondere dieselbe Spur. Für $\Phi \in \mathrm{End}_K(V)$ heißt

$$\mathrm{Tr}\ \Phi = \mathrm{Tr}\ M_{\mathcal{B}}^{\mathcal{B}}(\Phi)$$

die Spur des Endomorphismus Φ. Sie hängt nicht von der Wahl der geordneten Basis $\mathcal{B}$ von V ab.

2. Die Spur ist eine K–lineare Abbildung von $M(n \times n, K)$ mit Werten in K.

3. Es gilt für $A \in M(n \times m, K)$ und $B \in M(m \times n, K)$

$$\begin{aligned}\mathrm{Tr}\ AB &= \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}b_{ji} = \sum_{j=1}^{m}\sum_{i=1}^{n} b_{ji}a_{ij} \\ &= \sum_{j=1}^{m}(BA)_{jj} = \mathrm{Tr}\ BA\,.\end{aligned}$$

Oft wird dies als zyklische Invarianz der Spur formuliert: es gilt

$$\mathrm{Tr}\ A_1 \cdot A_2 \cdot \ldots \cdot A_n = \mathrm{Tr}\ A_n \cdot A_1 \cdot \ldots \cdot A_{n-1}\ .$$

4. Eine diagonalisierbare Matrix kann durch die folgenden Schritte diagonalisiert werden:

   (a) Berechne das charakteristische Polynom $P_A(\lambda)$ und bestimme seine Nullstellen.

   (b) Zu jedem Eigenwert λ bestimme den Eigenraum:

$$x \in \mathrm{Eig}(A, \lambda) = \ker(\lambda E_n - A) \Leftrightarrow (\lambda E_n - A)x = 0\,.$$

   Dieses lineare Gleichungssystem für x von n Gleichungen in n Unbestimmten kann zum Beispiel mit dem Gauß'schen Algorithmus gelöst werden.

   (c) Wähle Basen der Eigenräume, die zusammen eine Basis $(v_1, \ldots, v_n)$ von K^n bilden. Setze

$$S^{-1} := (v_1, \ldots, v_n)\ .$$

   Dann ist SAS^{-1} eine Diagonalmatrix, denn es gilt für den i-ten Vektor der Standardbasis

$$SAS^{-1}e_i = SAv_i = \lambda_i S v_i = \lambda_i e_i\ .$$

5. Beispiel: Spiegelungen des $\mathbb{R}^2$ an einer Ursprungsgeraden:

$$A = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

   (a)

$$P_A(\lambda) = (\lambda - \cos 2\theta)(\lambda + \cos 2\theta) - \sin^2 2\theta = \lambda^2 - 1$$

   Die beiden Eigenwerte sind $\lambda_{1,2} = \pm 1$.

(b) Eigenräume sind

$$\mathrm{Eig}(A,1) = \left\{ t \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \mid \ t \in \mathbb{R} \right\}$$
$$\mathrm{Eig}(A,-1) = \left\{ t \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix} \mid \ t \in \mathbb{R} \right\}$$

Man beachte, dass der Eigenraum zum Eigenwert $+1$ die Spiegelachse ist und der Eigenraum zum Eigenwert -1 senkrecht bezüglich des euklidischen Standardskalarprodukts auf $\mathbb{R}^2$ auf der Spiegelachse steht.

(c) Wir erhalten

$$S^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

und finden, wie erwartet,

$$SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

6. Aus

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

folgt

$$SA^kS^{-1} = SAS^{-1}SAS^{-1}\dots SAS^{-1} = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} .$$

7. Als Beispiel bestrachten wir die Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Ihr charakteristisches Polynom ist

$$P_A(\lambda) = \det \begin{pmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{pmatrix} = (\lambda - 1)^2 .$$

Also hat A nur den Eigenwert 1. Die Eigenvektoren bestimmen wir zu

$$x \in \mathrm{Eig}(A,1) \Leftrightarrow \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} x = 0 \Leftrightarrow x_2 = 0 .$$

Somit ist $\mathrm{Eig}(A,1) = \mathrm{span}_K e_1$ nur eindimensional. Die Matrix A ist also nicht diagonalisierbar!

5.2 Polynome

Bevor wir untersuchen können, welche Endomorphismen diagonalisierbar sind, müssen wir etwas mehr über Polynome lernen. Insbesondere wollen wir eine Definition des Begriffs Polynom geben, der ohne einen mysteriösen a priori Begriff einer Unbekannten auskommt.

In diesem Unterkapitel sei K ein kommutativer Ring mit Eins. (Bei einem ersten Durchlesen dürfen Sie sich unter K ruhig einen Körper vorstellen; allerdings brauchen wir später auch den allgemeineren Fall.)

Definition 5.2.1

Ein Quadrupel $(A, +, \circ, \cdot)$ mit einer Menge A und Verknüpfungen

$$\begin{aligned} + &\quad A \times A \to A \\ \circ &\quad A \times A \to A \\ \cdot &\quad K \times A \to A \end{aligned}$$

heißt K–Algebra, wenn gilt:

(A1) $(A, +, \circ)$ *ist ein Ring.*

(A2) Für alle $v, w \in A$ *und* $\alpha, \beta \in K$ *gilt*

$$\begin{aligned} (\alpha + \beta) \cdot v &= \alpha \cdot v + \beta \cdot v \\ \alpha \cdot (v + w) &= \alpha \cdot v + \alpha \cdot w \\ (\alpha\beta) \cdot v &= \alpha \cdot (\beta \cdot v) \\ 1 \cdot v &= v \end{aligned}$$

[Ist K ein Körper, so ist $(A, +, \cdot)$ ein K–Vektorraum.]

(A3) Für alle $v, w \in A$ *und* $\alpha, \beta \in K$ *gilt*

$$(\alpha \cdot v) \circ (\beta \cdot w) = (\alpha\beta) \cdot (v \circ w) \ .$$

Bemerkungen 5.2.2.

1. Oft schreiben wir "$a \cdot b$" oder "ab" statt "$a \circ b$".
2. Ist der Ring $(A, +, \circ)$ ein Ring mit Eins bzw. kommutativ, so heißt $(A, +, \circ, \cdot)$ eine Algebra mit Eins oder unitäre Algebra bzw. kommutative Algebra.
3. Beispiel
   - $A = M(n \times n, K)$ ist eine unitäre, aber für $n \geq 2$ nicht-kommutative K–Algebra.
   - Für jeden K–Vektorraum ist $A = \mathrm{End}(V)$ mit den Verknüpfungen

     $$\begin{aligned} (\varphi + \psi)(v) &= \varphi(v) + \psi(v) \\ (\alpha \cdot \varphi)(v) &= \alpha \cdot \varphi(v) \\ (\varphi \circ \psi)(v) &= \varphi\left(\psi(v)\right) \end{aligned}$$

     eine K–Algebra mit Einselement id_V. Sie ist für $\dim_K V \geq 2$ nicht kommutativ.
   - Sei X eine Menge und A eine K–Algebra. Dann ist $\mathrm{Abb}(X, A)$ ebenfalls eine K–Algebra.

Definition 5.2.3
Seien A, B zwei K–Algebren. Ein Ringhomomorphismus $\varphi : A \to B$ heißt K–Algebrenhomomorphismus, wenn zusätzlich gilt

$$\varphi\left(\lambda \cdot v\right) = \lambda \cdot \varphi(v) \ \text{ für alle } v \in A \text{ und } \lambda \in K \ .$$

Für unitäre Algebren A, B verlangt man zusätzlich $\varphi(1_A) = 1_B$.

Definition 5.2.4
Eine K–Algebra R mit Eins heißt ein Polynomring (korrekter wäre: Polynomalgebra) in der Unbestimmten X über K, wenn X ein Element von R ist und jedes Element $f \in R$ sich eindeutig *in der Gestalt*

$$f = a_0X^0 + a_1X^1 + \ldots + a_nX^n$$

mit $a_0, a_1, \ldots, a_n \in K$ darstellen lässt. Hierbei setzen wir $X^0 = 1 \in R$ und verstehen die Eindeutigkeit folgendermaßen: gilt in R

$$a_0 + a_1X^1 + \ldots + a_nX^n = b_0 + b_1X^1 + \ldots + b_mX^m$$

mit $m \geq n$, so ist $a_i = b_i$ für $0 \leq i \leq n$ und $b_i = 0$ für $i > n$. Die Elemente einer Polynomalgebra (mit Unbestimmter X) nennen wir Polynome (in X).

Man beachte, dass eine Polynomalgebra notwendigerweise kommutativ ist. Eine Polynomalgebra über K ist also ein Paar, bestehend aus einer K-Algebra und einem Element in dieser Algebra, genannt "Unbestimmte".

Satz 5.2.5. [universelle Eigenschaft der Polynomalgebra]
Sei R ein Polynomring in der Unbestimmten X über K. Für jede K–Algebra S und jede Wahl eines Elements $A \in S$ gibt es genau einen K–Algebrenhomomorphismus $\varphi_A : R \to S$ mit $\varphi_A(X) = A$.

Beweis.

- Eindeutigkeit: Da jedes Element $f \in R$ eine eindeutige Darstellung

  $$f = \sum_{i=0}^{n} a_iX^i$$

  besitzt, folgt aus $\varphi_A(X) = A$ sofort, dass

  $$\varphi_A(f) = \sum_{i=0}^{n} a_i\varphi_A(X)^i = \sum_{i=0}^{n} a_iA^i \qquad (*)$$

  gelten muss.

- Existenz: Die so definierte Abbildung φ_A ist wegen der Eindeutigkeit der Darstellung von f wohldefiniert. $\varphi_A(X) = A$ ist offensichtlich. Man rechnet leicht nach, dass die so definierte Abbildung φ_A ein K-Algebrenhomomorphismus ist.

□

Bemerkungen 5.2.6.

1. Wegen $(*)$ heißt φ_A auch der zum Element $A \in S$ gehörende Einsetzungshomomorphismus. Man schreibt auch für ein Polynom $f \in R$ auch

   $$\varphi_A(f) = f(A) \ .$$

2. Im Spezialfall des Polynomrings selbst, $S = R$, und $A = X$ folgt $\varphi_X = \mathrm{id}_R$ und somit die Schreibweise

   $$f = \varphi_X(f) = f(X) \ ,$$

   die Sie aus der Schule kennen.

3. Im Spezialfall der K-Algebra $S = K$ erhalten wir für jedes $\lambda \in K$ einen Wert

$$\varphi_\lambda(f) = f(\lambda) \in K \ .$$

Ein Polynom gibt also eine polynomiale Funktion $K \to K$. Man hat einen Ringhomomorphismus

$$\begin{array}{rcl} R & \to & \mathrm{Abb}(K,K) \ , \\ f & \mapsto & \tilde{f} \end{array}$$

den Auswertehomomorphismus. Dieser ist im allgemeinen *nicht* injektiv, d.h. ein Polynom kann nicht mit der induzierten polynomialen Funktion identifiziert werden:
zum Beispiel gibt es für $K = \mathbb{F}_2$ nur 4 verschiedene Funktionen, aber unendlich viele verschiedene Polynome.

4. Je zwei Polynomringe R, S in den Unbestimmten X bzw. Y sind isomorph:
Nach Satz 5.2.5 gibt es Algebrenhomomorphismen

$$\begin{array}{rcl} \Phi: \ R \to S & \text{und} & \Psi: \ S \to R \\ \text{mit} \ \ \Phi(X) = Y & \text{und} & \Psi(Y) = X \ . \end{array}$$

Dann sind

$$\Psi \circ \Phi: \ R \to R \ \text{ und } \ \Phi \circ \Psi: \ S \to S$$

Algebrenhomomorphismen mit

$$\Psi \circ \Phi(X) = X \ \text{ und } \ \Phi \circ \Psi(Y) = Y \ .$$

Aus der Eindeutigkeitsaussage in Satz 5.2.5 folgt

$$\Psi \circ \Phi = \mathrm{id}_R \ \text{ und } \ \Phi \circ \Psi = \mathrm{id}_S \ .$$

Wir müssen allerdings noch durch explizite Konstruktion zeigen, dass Polynomringe überhaupt existieren:

Satz 5.2.7.
Zu jedem kommutativen Ring K mit Eins existiert ein Polynomring in der Unbestimmten X über K.

Beweis.

- Die Menge R der Abbildungen in den kommutativen Ring K

$$\mathbb{N}_0 = \{0, 1, \ldots\} \ \to \ K \ ,$$

die nur für endlich viele natürliche Zahlen einen Wert ungleich Null annehmen, versehen wir durch Operationen auf den Abbildungswerten in K mit einer Addition und einer skalaren Multiplikation, so dass die Bedingungen (A2) aus Definition 5.2.1 erfüllt sind.

- Mit den speziellen Abbildungen

$$e_i(j) = \begin{cases} 0 \in K & \text{für } j \neq i \\ 1 \in K & \text{für } j = i \end{cases}$$

hat jedes $f \in R$ die eindeutige Darstellung

$$f = \sum_{i=0}^{\infty} f(i) e_i \ .$$

(In dieser Summe sind nur endlich viele Summanden ungleich Null; die Summe ist daher definiert.) Wir definieren die Multiplikation auf R durch

$$e_i \circ e_j = e_{i+j} \ .$$

Das Distributivgesetz in R legt eindeutig fest, dass

$$\left(\sum_{i=0}^{\infty} a_i e_i \right) \circ \left(\sum_{j=0}^{\infty} b_j e_j \right) = \sum_{k=0}^{\infty} \left(\sum_{k=i+j} a_i b_j \right) e_k$$

gelten muss. Dies gibt eine kommutative K–Algebra mit 1. $X := e_1$ kann dann als Unbestimmte gewählt werden. (Weitere Details werden Sie in einer Übungsaufgabe ausarbeiten.)

□

Definition 5.2.8

1. *Da der Polynomring über dem kommutativen Ring K in der Unbestimmten X bis auf eindeutige Isomorphie eindeutig ist, bezeichnen wir ihn mit $K[X]$.*

2. *Besitzt $f \in K[X]$ die Gestalt*

$$f = a_0 + a_1 X + \ldots a_n X^n$$

mit $a_n \neq 0$, so heißt a_n der höchste Koeffizient von f. Die nicht-negative Zahl n heißt dann der Grad von f. Dem Nullpolynom ordnen wir zu

$$\operatorname{grad}(0) = -\infty \ .$$

Ist der höchste Koeffizient $a_n = 1$, so heißt das Polynom normiert.

Bemerkungen 5.2.9.

1. Es gilt

$$\begin{aligned} \operatorname{grad}(f+g) &\leq \max\left(\operatorname{grad}(f), \operatorname{grad}(g)\right) \\ \operatorname{grad}(fg) &\leq \operatorname{grad}(f) + \operatorname{grad}(g) \ . \end{aligned}$$

Ist der kommutative Ring K nullteilerfrei, d.h. folgt aus $ab = 0$ mit $a, b \in K$, dass $a = 0$ oder $b = 0$ gilt, so gilt

$$\operatorname{grad}(fg) = \operatorname{grad}(f) + \operatorname{grad}(g) \ .$$

Umgekehrt gilt diese Beziehung genau dann für alle Polynome f, g, wenn der kommutative Ring K nullteilerfrei ist.

2. Im folgenden werden wir nur den Fall betrachten, dass K ein nullteilerfreier kommutativer Ring mit Eins, also ein Integritätsring, ist. Körper sind insbesondere Integritätsringe.

Satz 5.2.10. Division mit Rest von Polynomen
Sei $f \neq 0$ ein Polynom in $K[X]$, dessen höchster Koeffizient in K ein multiplikatives Inverses hat. (Diese Bedingung ist automatisch erfüllt, wenn K ein Körper ist.) Zu jedem Polynom $g \in K[X]$ gibt es dann Polynome $q, r \in K[X]$ mit

$$g = qf + r \quad \text{und} \quad \operatorname{grad}(r) < \operatorname{grad}(f) \ .$$

Hierdurch sind q und r eindeutig bestimmt.

Ist $r = 0$, so sagen wir, dass f das Polynom g teilt, in Zeichen: $f|g$, bzw. dass g ein Vielfaches von f ist.

Beweis.
Den Beweis der Existenz von q und r für gegebenes f führen wir mit vollständiger Induktion nach $\operatorname{grad}(g)$.

- Für $\operatorname{grad}(g) < \operatorname{grad}(f)$ setze $q = 0$ und $r = g$.
- Gelte also $m := \operatorname{grad} g \geq n := \operatorname{grad} f$. Sei a der invertible höchste Koeffizient von f und c der höchste Koeffizient von g. Das Polynom

  $$h := g - ca^{-1} f \cdot X^{m-n}$$

  hat strikt kleineren Grad als g. Nach Induktionsannahme existieren $q_1, r \in K[X]$ mit

  $$h = q_1 f + r \qquad \text{und} \quad \operatorname{grad} r < \operatorname{grad} f \ .$$

  Daraus folgt eine Darstellung von g

  $$g = h + ca^{-1} f \cdot X^{m-n} = \left(q_1 + ca^{-1} \cdot X^{m-n}\right) f + r \quad \text{und} \quad \operatorname{grad} r < \operatorname{grad} f \ .$$

- Um die Eindeutigkeit dieser Darstellung zu zeigen, nehmen wir an, es gelte

  $$g = qf + r = \widetilde{q} f + \widetilde{r}$$

  mit $q, \widetilde{q}, r, \widetilde{r} \in K[X]$ und $\operatorname{grad} \widetilde{r} < \operatorname{grad} f$, $\operatorname{grad} r < \operatorname{grad} f$. Hieraus folgt

  $$r - \widetilde{r} = (\widetilde{q} - q) f$$

  Wäre $\widetilde{q} \neq q$, so wäre

  $$\operatorname{grad}(r - \widetilde{r}) = \operatorname{grad}(\widetilde{q} - q) + \operatorname{grad}(f) \geq \operatorname{grad}(f) \ ,$$

  im Widerspruch zur Bedingung an den Grad eines Rests. Also muss $\widetilde{q} = q$ gelten und somit $\widetilde{r} = r$.

□

Beispiel 5.2.11.
Es gilt: $(X^4 - 1) : (X - 1) = X^3 + X^2 + X + 1$.

Lemma 5.2.12.
Sei $f \in K[X]$. Ist $a \in K$ eine Nullstelle von f, d.h. gilt $\tilde{f}(a) = 0$, so gibt es genau ein Polynom $g \in K[X]$ mit

$$f = (X - a) g$$

und $\operatorname{grad}(g) = \operatorname{grad}(f) - 1$.

Man sagt auch, dass man einen Linearfaktor abspalten kann.

Beweis.
Die Polynomdivision mit Rest nach Satz 5.2.10 liefert uns Polynome g, r, die

$$f = (X - a) \cdot g + r$$

erfüllen mit $\text{grad} r < \text{grad}(X-a) = 1$; also ist r ein konstantes Polynom. Wegen $0 = \tilde{f}(a) = \tilde{r}(a)$ folgt für das konstante Polynom $r = 0$. Es ist also

$$\text{grad } f = \text{grad}(X - a) + \text{grad } g = 1 + \text{grad } g \; .$$

□

Korollar 5.2.13.
Sei $f \in K[X],\ f \neq 0$. Hat f genau k paarweise verschiedene Nullstellen, so ist

$$k \leq \text{grad}(f) \; .$$

Ein Polynom vom Grad n hat also höchstens n paarweise verschiedene Nullstellen.

Beweis.
Vollständige Induktion nach $n := \text{grad } f$.

- Für $n = 0$ ist f konstantes Polynom, hat also keine Nullstelle. Also ist $k = 0$.
- Induktionsschritt: hat f keine Nullstelle, so ist $k = 0$ und die Behauptung trivialerweise richtig. Hat f eine Nullstelle $\lambda \in K$, so gibt es nach Lemma 5.2.12 ein Polynom g vom Grad $\text{grad } g = \text{grad } f - 1$ mit
$$f = (X - \lambda)g$$
Jede von λ verschiedene Nullstelle ist dann auch Nullstelle von g. Nach Induktionsvoraussetzung hat aber g höchstens $n - 1$ verschiedene Nullstellen.

□

Korollar 5.2.14.
Ist K ein Integritätsring mit unendlich vielen Elementen, so ist der Auswertehomomorphismus

$$\begin{array}{lcl} K[X] & \to & \text{Abb}(K, K) \\ f & \mapsto & \tilde{f} \end{array}$$

injektiv.

Beweis.
Wäre f im Kern des Auswertehomomorphismus, aber nicht das Nullpolynom, so hätte f unendlich viele verschiedene Nullstellen. Dies ist im Widerspruch zu Korollar 5.2.13. □

Definition 5.2.15
Sei K ein Integritätsring, $f \in K[X]$ und $f \neq 0$. Für $\lambda \in K$ heißt

$$\mu(\lambda, f) := \max\{r \in \mathbb{N} \mid \exists g \in K[X] \ \textit{ mit } \ f = (X - \lambda)^r \cdot g\}$$

die Vielfachheit der Nullstelle λ von f.

Bemerkungen 5.2.16.

1. Es gilt $0 \leq \mu(\lambda, f) \leq \operatorname{grad}(f)$
2. Gilt $f = (X - \lambda)^{\mu(\lambda,f)} g$, so ist λ keine Nullstelle von g.
3. $\mu(\lambda, f) = 0$ genau dann, wenn λ keine Nullstelle von f ist.
4. Sind $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Nullstellen von f mit Vielfachheiten $r_1, \ldots, r_k$, so ist
$$f = (X - \lambda_1)^{r_1} (X - \lambda_2)^{r_2} \ldots (X - \lambda_k)^{r_k} g \, ,$$
wobei $g \in K[X]$ keine Nullstellen hat. Hierbei sind das Polynom g, die Nullstellen λ_i und ihre Vielfachheiten $, r_i$ bis auf Reihenfolge eindeutig. (Übung).
5. Beispiele:
   - Wir rechnen zunächst über dem Körper $\mathbb{R}$ der reellen Zahlen:
   $$\begin{aligned} f &= X^5 - X^4 + X^3 - X^2 \in \mathbb{R}[X] \\ f &= X^2 \left(X^3 - X^2 + X - 1\right) \\ &= X^2 (X - 1) \left(X^2 + 1\right) \, . \end{aligned}$$
   Hier ist $\lambda_1 = 0, r_1 = 2$, $\lambda_2 = 1, r_2 = 1$, $g = X^2 + 1$ Man beachte, dass die Summer der Vielfachheiten kleiner als der Grad von f ist.
   - Über dem Körper $\mathbb{C}$ der komplexen Zahlen dagegen erhalten wir
   $$\begin{aligned} f &= X^5 - X^4 + X^3 - X^2 \in \mathbb{C}[X] \\ f &= X^2 (X - 1) (X + \mathrm{i}) (X - \mathrm{i}) \, . \end{aligned}$$
   Hier ist $\lambda_1 = 0, r_1 = 2$, $\lambda_2 = 1, r_2 = 1, \lambda_3 = +\mathrm{i}, r_3 = 1$, $\lambda_4 = -\mathrm{i}, r_4 = 1$ und $g = 1$. Man beachte, dass hier die Summer der Vielfachheiten gleich dem Grad von f ist.

Definition 5.2.17
Sei K ein nullteilerfreier Ring. Man sagt, ein Polynom $f \in K[X]$ zerfalle in Linearfaktoren, falls es sich in der Form

$$f = a(X - \lambda_1)^{r_1} (X - \lambda_2)^{r_2} \ldots (X - \lambda_n)^{r_n}$$

mit $a, \lambda_1, \ldots, \lambda_n \in K$ und $r_i \in \mathbb{N}$ schreiben lässt. Ein solcher Ausdruck heißt Linearfaktorzerlegung des Polynoms f.

Bemerkungen 5.2.18.

1. Ein Polynom f zerfällt genau dann in ein Produkt von Linearfaktoren, wenn $\sum_{\lambda \in K} \mu(\lambda, f) = \operatorname{grad} f$ gilt. (Man beachte, dass es nach Korollar 5.2.13 nur endlich viele Nullstellen gibt und daher die Summe endlich ist.)
2. Existiert eine Zerlegung in Linearfaktoren, so ist diese eindeutig bis auf die Reihenfolge der Faktoren.

3. Der Fundamentalsatz der Algebra besagt:
   Ist $f \in \mathbb{C}[X]$ mit $\text{grad}(f) \geq 1$, so besitzt f wenigstens eine komplexe Nullstelle.
   Er wird mit den Hilfsmitteln der komplexen Analysis oder der Topologie beweisen, weshalb wir hier keinen Beweis bringen. (Literaturhinweis: W. Fischer, I. Lieb, *Funktionentheorie*, Vieweg 2005, Kapitel IV, §8; R. Stöcker, H. Zieschang, *Algebraische Topologie* Teubner, Stuttgart, 1994, Satz 2.2.9.)

4. Aus dem Fundamentalsatz der Algebra folgt sofort: jedes komplexe Polynom zerfällt in $\mathbb{C}[X]$ in Linearfaktoren. Denn schreibe $f = (X - \lambda_1)^{r_1} \dots (X - \lambda_n)^n g$ mit $\lambda_i \in \mathbb{C}$, wobei das komplexe Polynom g keine Nullstellen hat. Nach 3. ist g konstant.

5.3 Diagonalisierbarkeit

K bezeichne ab sofort wieder einen beliebigen Körper. Wir wollen untersuchen, welche Matrizen $A \in M(n \times n, K)$ diagonalisierbar sind, die also die Eigenschaft haben, dass K^n eine Basis aus Eigenvektoren von A besitzt. Wir haben schon gesehen:

- Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar (Korollar 5.1.3.2).

- Ist A diagonalisierbar, so ist

$$P_A(\lambda) = \det \begin{pmatrix} \lambda - \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda - \lambda_n \end{pmatrix} = (\lambda - \lambda_1) \cdot \ldots \cdot (\lambda - \lambda_n)$$

  d.h. das charakteristische Polynom zerfällt in Linearfaktoren.

Wir hatten das charakteristische Polynom eingeführt als polynomiale Abbildung

$$\begin{array}{rccl} \Phi_A : & K & \to & K \\ & \lambda & \mapsto & \det(\lambda E_n - A)\,. \end{array}$$

Das zugehörige Polynom ist definiert als Determinante der Matrix $XE_n - A$, deren Einträge im Polynomring $K[X]$ liegen. Determinanten lassen sich in der Tat wie in Kapitel 4 allgemeiner für Matrizen mit Einträgen in beliebigen kommutativen Ringen definieren. Es gelten die gleichen Aussagen wie in Kapitel 4; lediglich der Beweis der Multiplikativität der Determinantenfunktion muss wie in Satz 4.5.3 geführt werden. Wir definieren also das charakteristische Polynom als Determinante der Matrix $XE_n - A \in M(n \times n, K[X])$, deren Einträge Elemente im Polynomring $K[X]$ sind:

$$P_A(X) = \det(XE_N - A) \in K[X] \ .$$

Definition 5.3.1
Ist $\lambda \in K$ Eigenwert der Matrix $A \in M(n \times n, K)$, so heißt die Vielfachheit der Nullstelle λ des charakteristischen Polynom die algebraische Vielfachheit des Eigenwerts λ:

$$\mu_{alg}(A, \lambda) := \mu(\lambda, P_A(X)) \ .$$

Bemerkung 5.3.2.
Geometrische und algebraische Vielfachheit eines Eigenwerts einer Matrix können verschieden sein. Beispiel:

$$A = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix} \qquad \text{mit } \lambda_0 \in K$$

Dann ist das charakteristische Polynom $P_A(X) = (X - \lambda_0)^2$, also $\mu_{alg}(A, \lambda_0) = 2$, aber $\mu_{geo}(A, \lambda_0) = 1$, vgl. Beispiel 5.1.11.6

Lemma 5.3.3.
Sei $A \in M(n \times n, K)$ und λ Eigenwert von A. Dann gilt

$$1 \leq \mu_{geo}(A, \lambda) \leq \mu_{alg}(A, \lambda) \; .$$

Beweis.
Sei $k := \mu_{geo}(A, \lambda)$. Ergänze eine geordnete Basis $(v_1, \ldots, v_k)$ des Eigenraums $\mathrm{Eig}(A, \lambda)$ zu einer geordneten Basis $\mathcal{B} = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ von K^n. Für $S^{-1} = (v_1, \ldots, v_n)$ gilt $SAS^{-1}e_i = SAv_i = S(\lambda v_i) = \lambda e_i$ für $i = 1, \ldots, k$ und daher

$$SAS^{-1} = \begin{pmatrix} \lambda & & 0 & \vdots & \\ & \ddots & & \vdots & C \\ 0 & & \lambda & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & 0 & & \vdots & D \end{pmatrix}$$

Da nach Bemerkung 5.1.8.2 ähnliche Matrizen das gleiche charakteristische Polynom haben, folgt mit Korollar 4.3.7 für das charakteristische Polynom:

$$P_A(X) \overset{5.1.8.2}{=} P_{SAS^{-1}}(X) \overset{4.3.7}{=} (X - \lambda)^k \det\left(XE_{n-k} - D\right)$$

Daher ist $\mu_{alg}(A, \lambda) \geq k = \mu_{geo}(A, \lambda)$. □

Satz 5.3.4.
Sei V ein n–dimensionaler Vektorraum und sei $\Phi : V \to V$ ein Endomorphismus. Seien $\lambda_1, \ldots, \lambda_N$ die paarweise verschiedenen Eigenwerte von Φ. Dann sind die folgenden Aussagen äquivalent:

1. Φ ist diagonalisierbar.
2. $P_\Phi(X)$ zerfällt in Linearfaktoren und $\mu_{geo}(\Phi, \lambda) = \mu_{alg}(\Phi, \lambda)$ für alle Eigenwerte λ von Φ.
3. $$\sum_{i=1}^{N} \mu_{geo}(\Phi, \lambda_i) = n$$
4. Es gilt die folgenden Eigenraumzerlegung
$$V = \oplus_{i=1}^{N} \mathrm{Eig}(\Phi, \lambda_i) \; ,$$
d.h. jedes $v \in V$ kann man eindeutig in der Form
$$v = v_1 + v_2 + \ldots + v_N$$
mit $v_i \in \mathrm{Eig}(\Phi, \lambda_i)$ schreiben.

Beweis.

1. $\Rightarrow$ 2. Wähle eine geordnete Basis $\mathcal{B}$ von Eigenvektoren:

$$\begin{aligned} v_1, \ldots, v_{k_1} \quad & \text{zum Eigenwert } \lambda_1, \text{ also } k_1 = \mu_{geo}(\Phi, \lambda_1) \\ v_{k_1+1}, \ldots, v_{k_1+k_2} \quad & \text{zum Eigenwert } \lambda_2, \text{ also } k_2 = \mu_{geo}(\Phi, \lambda_2) \\ \ldots & \end{aligned}$$

Dann ist

$$M_{\mathcal{B}}^{\mathcal{B}}(\Phi) = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & 0 & \\ & & & \lambda_2 & & \\ & & & & \ddots & \\ & & 0 & & \lambda_2 & \\ & & & & & \ddots \end{pmatrix}$$

Für das charakteristische Polynom folgt

$$P_\Phi(X) = P_{M_{\mathcal{B}}^{\mathcal{B}}(\Phi)}(X) = (X - \lambda_1)^{k_1} \ldots (X - \lambda_N)^{k_N} \ ,$$

also

$$\mu_{geo}(\Phi, \lambda_i) = k_i = \mu_{alg}(\Phi, \lambda_i) \ .$$

2. $\Rightarrow$3. Zerfällt P_Φ in Linearfaktoren, so gilt

$$n = \operatorname{grad} P_\Phi = \sum_{i=1}^{N} \mu_{alg}(\Phi, \lambda_i) \ .$$

Aus der weiteren Annahme $\mu_{geo}(\Phi, \lambda_i) = \mu_{alg}(\Phi, \lambda_i)$ folgt sofort 3.

3. $\Rightarrow$ 4. Setze $W := \operatorname{Eig}(\Phi, \lambda_1) + \ldots + \operatorname{Eig}(\Phi, \lambda_N)$.
Die Summe ist direkt: denn gelte

$$w = w_1 + \ldots + w_N = w_1' + \ldots + w_N'$$

mit $w_i, w_i' \in \operatorname{Eig}(\Phi, \lambda_i)$, so liegt $w_i - w_i' \in \operatorname{Eig}(\Phi, \lambda_i)$ und es gilt $0 = \sum_{i=1}^{N}(w_i - w_i')$. Aus Satz 5.1.2 folgt $w_i - w_i' = 0$. Wegen

$$\dim W = \sum_{i=1}^{N} \dim \operatorname{Eig}(\Phi, \lambda_i) = \sum_{i=1}^{N} \mu_{geo}(\Phi, \lambda_i) = n = \dim V$$

folgt auch $V = W$, also 4.

4. $\Rightarrow$ 1. Wähle Basen von $\operatorname{Eig}(\Phi, \lambda_i)$, die zusammen eine Basis von V aus Eigenvektoren von Φ ergeben.

□

Korollar 5.3.5.
Sei $A \in M(n \times n, K)$. Dann sind äquivalent:

1. A ist diagonalisierbar, d.h. ähnlich zu einer Diagonalmatrix.

2. $P_A(X)$ zerfällt in Linearfaktoren und $\mu_{alg}(A,\lambda) = \mu_{geo}(A,\lambda)$ für alle Eigenwerte λ von A.

Beweis.
Wende Satz 5.3.4 auf die lineare Abbildung $\Phi: K^n \to K^n$ mit $\Phi(x) = Ax$ für $x \in K^n$ an. $\square$
Damit ist klar, dass die Matrix aus Beispiel 5.3.2 nicht diagonalisierbar ist. Die Klassifikation der Ähnlichkeitsklassen quadratischer Matrizen muss also komplizierter sein als die am Anfang des Kapitels rekapitulierte Klassifikation der Äquivalenzklassen belieber Matrizen aus Bemerkung 3.6.8.3.
Wir wollen noch die folgende Frage behandeln: gegeben seien *zwei* Endomorphismen $\Phi, \Psi: V \to V$. Wann kann man sie *gleichzeitig* diagonalisieren, d.h. wann existiert *eine* Basis $\mathcal{B}$ von V, so dass die *beiden* darstellenden Matrizen $M_{\mathcal{B}}(\Phi)$ und $M_{\mathcal{B}}(\Psi)$ Diagonalmatrizen sind?

Der folgende Begriff wird im Beweis von Satz 5.3.7 nützlich sein:

Definition 5.3.6
Es sei Φ ein Endomorphismus eines K-Vektorraums V. Ein Untervektorraumraum $W \subset V$ mit $\Phi(W) \subset W$ heißt ein Φ–invarianter Untervektorraum.

- Beispiele für Φ-invariante Unterräume sind die trivialen Untervektorräume $W = V$ und $W = \{0\}$ sowie $\mathrm{Eig}(\Phi, \lambda)$ für jedes $\lambda \in K$.
- Die Summe Φ-invarianter Untervektorräume ist Φ-invariant: Für $w_1 + w_2$ mit $w_i \in W_i$ gilt
$$\Phi(w_1 + w_2) = \Phi(w_1) + \Phi(w_2) \in W_1 + W_2 \, .$$
- Der Schnitt Φ-invarianter Untervektorräume ist ein Φ-invarianter Untervektorraum.

Satz 5.3.7.
Sei V ein endlich–dimensionaler K–Vektorraum und seien $\Phi, \Psi: V \to V$ diagonalisierbare Endomorphismen. Dann sind äquivalent:

1. Φ und Ψ sind gleichzeitig diagonalisierbar.

2. Φ und Ψ kommutieren, d.h. $\Phi \circ \Psi = \Psi \circ \Phi$.

Beweis.

1. $\Rightarrow$2. Sei $\mathcal{B}$ eine Basis von V, in der die beiden Matrizen $M_{\mathcal{B}}(\Phi)$ und $M_{\mathcal{B}}(\Psi)$ Diagonalmatrizen sind. Dann gilt:
$$M_{\mathcal{B}}(\Phi \circ \Psi) = M_{\mathcal{B}}(\Phi) \cdot M_{\mathcal{B}}(\Psi) = M_{\mathcal{B}}(\Psi) \cdot M_{\mathcal{B}}(\Phi) = M_{\mathcal{B}}(\Psi \circ \Phi) \, ,$$
wobei die zweite Gleichung daraus folgt, dass die darstellenden Matrizen Diagonalmatrizen sind und Diagonalmatrizen kommutieren. Hieraus folgt $\Phi \circ \Psi = \Psi \circ \Phi$, da $M_{\mathcal{B}}: \mathrm{End}(V) \to M(n \times n, K)$ ein *Iso*morphismus von K-Vektorräumen ist.

2.⇒1. Da Φ und Ψ diagonalisierbar sind, gibt es die beiden Eigenraumzerlegungen

$$V = \bigoplus_{\lambda \in K} \mathrm{Eig}(\Phi, \lambda) \ .$$
$$V = \bigoplus_{\mu \in K} \mathrm{Eig}(\Psi, \mu) \ .$$

Aus der Kommutativität von Φ und Ψ folgt, dass alle Eigenräume in diesen Zerlegungen unter Φ und Ψ gleichzeitig invariant sind. In der Tat: ist $v \in \mathrm{Eig}(\Phi, \lambda)$, so gilt $\Phi(v) = \lambda v$. Daraus folgt aber $\Phi(\Psi(v)) = \Psi(\Phi(v)) = \Psi(\lambda v) = \lambda \Psi(v)$. Also ist auch $\Psi(v) \in \mathrm{Eig}(\Phi, \lambda)$. Wir behaupten

$$V = \bigoplus_{\lambda, \mu} \mathrm{Eig}(\Phi, \lambda) \cap \mathrm{Eig}(\Psi, \mu) \ .$$

Da dies eine Zerlegung in gemeinsame Eigenräume von Φ und Ψ ist, folgt die Behauptung.

Da Eigenräume zu verschiedenen Eigenwerten sich nur im Nullvektorraum schneiden können, reicht es aus, für jeden festen Eigenwert $\lambda \in K$ von Φ die Summerzerlegung

$$\mathrm{Eig}(\Phi, \lambda) = \sum_{\mu_i} \mathrm{Eig}(\Phi, \lambda) \cap \mathrm{Eig}(\Psi, \mu_i)$$

zu zeigen. Die Summe ist dann automatisch direkt. Sei also $v \in \mathrm{Eig}(\Phi, \lambda)$; da Ψ diagonalisierbar ist, können wir v wie jeden Vektor aus V als Summe $v = v_1' + v_2' + \ldots + v_m'$ mit $v_i' \in \mathrm{Eig}(\Psi, \mu_i)$ schreiben. Dann gilt

$$\begin{aligned} \Phi(v) &= \Phi(v_1') + \Phi(v_2') + \ldots + \Phi(v_m') \\ &= \lambda v = \lambda v_1' + \lambda v_2' + \ldots + \lambda v_m' \ . \end{aligned}$$

Da $\Phi(v_i') \in \mathrm{Eig}(\Psi, \mu_i)$ wegen der Invarianz des Eigenraums $\mathrm{Eig}(\Psi, \mu_i)$ unter Φ, folgt aus der Eindeutigkeit der Zerlegung von $\Phi(v) = \lambda v$ in Komponenten bezüglich der direkten Summe $V = \oplus \mathrm{Eig}(\Psi, \mu_i)$, dass $\Phi(v_i') = \lambda v_i'$, also $v_i' \in \mathrm{Eig}(\Phi, \lambda) \cap \mathrm{Eig}(\Psi, \mu_i)$, was zu zeigen war.

□

5.4 Trigonalisierbarkeit

Wir wollen nun noch sehen, was sich über einen Endomorphismus aussagen lässt, wenn wir nur wissen, dass sein charakteristisches Polynom vollständig in Linearfaktoren zerfällt, aber die geometrischen Vielfachheiten nicht kennen.

Lemma 5.4.1.
Sei $\Phi : V \to V$ ein Endomorphismus eines endlich-dimensionalen K–Vektorraums V. Ist W ein Φ–invarianter Untervektorraum, so teilt das charakteristische Polynom $P_{\Phi|W}$ das charakteristische Polynom P_Φ.

Beweis.
Ergänze eine geordnete Basis $\mathcal{B}'$ von W zu einer geordneten Basis $\mathcal{B}$ von V. Die darstellende Matrix ist in dieser Basis

$$M_{\mathcal{B}}(\Phi) = \begin{pmatrix} M_{\mathcal{B}'}\left(\Phi_{|W}\right) & \vdots & * \\ \cdots\cdots\cdots\cdots & & \\ 0 & \vdots & A \end{pmatrix}$$

Es folgt mit $n := \dim_K V$ und $k := \dim_K W$ mit Korollar 4.3.7

$$P_\Phi(\lambda) = \det\ (\lambda E_n - M_\mathcal{B}(\Phi)) = \det\ \left(\lambda E_k - M_{\mathcal{B}'}(\Phi_{|W})\right) \cdot \det\ (\lambda E_{n-k} - A)$$

□

Definition 5.4.2
Sei $\Phi : V \to V$ ein Endomorphismus eines n–dimensionalen K–Vektorraums V.

1. *Eine Fahne von V ist eine Kette von Untervektorräumen*

$$\{0\} = V_0 \subset V_1 \subset \ldots \subset V_n = V$$

mit $\dim V_r = r$ für alle $r = 0, \ldots, n$.

2. *Eine Fahne von V heißt Φ–invariant, falls alle Untervektorräume Φ-invariant sind, also für alle $r = 0, \ldots, n$ die Inklusion $\Phi(V_i) \subset V_i$ gilt.*

Beispiel 5.4.3.
Jede geordnete Basis $(v_1, \ldots, v_n)$ eines Vektorraums V liefert mit $V_r := \text{span}_K\{v_1, \ldots, v_r\}$ eine Fahne von V.

Satz 5.4.4.
Sei $\Phi \in \text{End}(V)$, $\dim_K V = n$. Dann sind äquivalent:

1. Es existiert eine Φ–invariante Fahne von V.
2. Es gibt eine geordnete Basis $\mathcal{B}$ von V, in der die darstellende Matrix $M_\mathcal{B}(\Phi)$ eine obere Dreiecksmatrix ist.

Beweis.

2. $\Rightarrow$1. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ und $A = (a_{ij}) = M_\mathcal{B}(\Phi)$ eine obere Dreiecksmatrix. Betrachte die Fahne $V_i = \text{span}_K\{v_1, \ldots, v_i\}$ wie in Beispiel 5.4.3. Wegen

$$\Phi(v_i) = \sum_{j=1}^{n} a_{ji} v_j = \sum_{j=1}^{i} a_{ji} v_j$$

ist $\Phi(V_i) \subset V_i$, die Fahne ist also Φ–invariant.

1. $\Rightarrow$2. Wir konstruieren aus einer Φ–invarianten Fahne folgendermaßen eine geordnete Basis: sei $\{v_1\}$ eine Basis von V_1; wegen $V_1 \subset V_2$ ist es möglich, dies zu einer Basis $\{v_1, v_2\}$ von V_2 zu ergänzen. So fahren wir fort und ergänzen eine geordnete Basis $(v_1, v_2, \ldots, v_i)$ von V_i zu einer geordneten Basis $(v_1, v_2, \ldots, v_i, v_{i+1})$ von V_{i+1}.
Für alle $i \in \{1, \ldots, n\}$ gilt $\Phi(v_i) \in \Phi(V_i) \subset V_i$; also

$$\Phi(v_i) = \sum_{j=1}^{i} a_{ji} v_j$$

mit $a_{ji} \in K$. Die darstellende Matrix

$$M_\mathcal{B}(\Phi) =: A$$

erfüllt also $a_{ij} = 0$ für $i > j$, ist also eine obere Dreiecksmatrix.

□

Definition 5.4.5

1. *Sei Φ Endomorphismus eines endlich–dimensionalen K–Vektorraums V. Φ heißt trigonalisierbar, falls es eine geordnete Basis $\mathcal{B}$ von V gibt, so dass $M_{\mathcal{B}}(\Phi)$ eine obere Dreiecksmatrix ist.*

2. *Eine Matrix $A \in M(n \times n, K)$ heißt trigonalisierbar, falls sie zu einer oberen Dreiecksmatrix ähnlich ist.*

Satz 5.4.6.
Sei Φ Endomorphismus eines endlich–dimensionalen K–Vektorraums V. Dann sind äquivalent:

1. Φ ist trigonalisierbar.

2. Das charakteristische Polynom P_Φ zerfällt vollständig in Linearfaktoren.

Beweis.

1. $\Rightarrow$2. Wir rechnen das charakteristische Polynom in einer geordneten Basis $\mathcal{B}$ aus, in der $M_{\mathcal{B}}(\Phi)$ eine obere Dreiecksmatrix D ist:

$$P_\Phi(X) = P_D(X) = \det \begin{pmatrix} X-\lambda_1 & & & \\ & X-\lambda_2 & * & \\ & & \ddots & \\ & & & X-\lambda_n \end{pmatrix} \overset{4.2.2.5}{=} \prod_{i=1}^{n}(X-\lambda_i)$$

2. $\Rightarrow$1. Vollständige Induktion nach $n := \dim_K V$. Für $n = 1$ ist nichts zu zeigen. Sei $n > 1$; das charakteristische Polynom zerfalle,

$$P_\Phi(X) = \prod_{i=1}^{n}(X-\lambda_i) \ .$$

Wähle einen Eigenvektor $v_1 \in V$ zum Eigenwert λ_1 und ergänze zu einer Basis $\widetilde{\mathcal{B}} = (v_1, \ldots, v_n)$ von V. Dann ist

$$M_{\widetilde{\mathcal{B}}}(\Phi) = \left(\begin{array}{c|c} \lambda_1 & a_2 \ldots a_n \\ \hline 0 & \\ \vdots & \tilde{A} \\ 0 & \end{array}\right)$$

Es folgt wie in Lemma 5.4.1

$$P_\Phi(X) = (X-\lambda_1)P_{\tilde{A}}(X) \ .$$

Dann zerfällt auch das charakteristische Polynom $P_{\tilde{A}}(X)$. Nach Induktionsvoraussetzung ist die $(n-1) \times (n-1)$-Matrix $\widetilde{A}$ ähnlich zu einer oberen Dreiecksmatrix $\widetilde{D}$, d.h. es gibt eine invertible $(n-1) \times (n-1)$-Matrix $\tilde{S}$ mit:

$$\widetilde{S}\widetilde{A}\widetilde{S}^{-1} = \widetilde{D}$$

Setze

$$S := \left(\begin{array}{c|c} 1 & 0\dots0 \\ \hline 0 & \\ \vdots & \widetilde{S} \\ 0 & \end{array}\right) \in M(n\times n, K)$$

finde als inverse Matrix

$$S^{-1} = \left(\begin{array}{c|c} 1 & 0\dots0 \\ \hline 0 & \\ \vdots & \widetilde{S}^{-1} \\ 0 & \end{array}\right) \in M(n\times n, K)$$

und rechne:

$$\begin{aligned} SM_{\widetilde{\mathcal{B}}}(\Phi)S^{-1} &= \left(\begin{array}{c|c} 1 & 0\dots0 \\ \hline 0 & \\ \vdots & \widetilde{S} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} \lambda_1 & a_2\dots a_n \\ \hline 0 & \\ \vdots & \widetilde{A} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0\dots0 \\ \hline 0 & \\ \vdots & \widetilde{S}^{-1} \\ 0 & \end{array}\right) \\ &= \left(\begin{array}{c|c} \lambda_1 & a_2\dots a_n \\ \hline 0 & \\ \vdots & \widetilde{S}\widetilde{A} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0\dots0 \\ \hline 0 & \\ \vdots & \widetilde{S}^{-1} \\ 0 & \end{array}\right) \\ &= \left(\begin{array}{c|c} \lambda_1 & *\dots* \\ \hline 0 & \\ \vdots & \widetilde{S}\widetilde{A}\widetilde{S}^{-1} \\ 0 & \end{array}\right). \end{aligned}$$

Dies ist eine obere Dreiecksmatrix.

□

Korollar 5.4.7.
Jede Matrix $A \in M(n\times n, \mathbb{C})$ mit komplexen Einträgen ist trigonalisierbar.

Beweis.
Nach dem Fundamentalsatz der Algebra 5.2.18.3 zerfällt das charakteristische Polynom in Linearfaktoren. □

Betrachtung 5.4.8.
Rechenverfahren zur Triagonalisierung.
Sei $A \in M(n\times n, K)$ mit einem charakteristischem Polynom $P_A(X) = \prod_{i=1}^{n}(X-\lambda_i)$, das in Linearfaktoren zerfällt.

1. Schritt: Bestimme, zum Beispiel mit dem Gauß-Algorithmus, einen Eigenvektor

$$v_1 = \begin{pmatrix} v_{11} \\ \vdots \\ v_{1n} \end{pmatrix}$$

zum Eigenwert λ_1. Weil für den Eigenvektor $v_1 \neq 0$ gilt, gibt es ein i, so dass die i-te Komponente $v_{1i} \neq 0$ von v_1 ungleich Null ist; wähle ein solches und setze

$$S_1 = (v_1, e_1, \ldots, \widehat{e_i}, \ldots, e_n)^{-1} \in GL(n, K) \ .$$

Hierbei zeigt das Symbol "Hut" an, dass der Vektor e_i der Standardbasis ausgelassen wird. Dann gilt $(S_1)^{-1}e_1 = v_1$, denn das Bild von e_1 ist der erste Spaltenvektor von $(S_1)^{-1}$. Dann ist

$$S_1 A S_1^{-1} e_1 = S_1 A v_1 = S_1 \lambda_1 v_1 = \lambda_1 e_1 \ ,$$

also

$$S_1 A S_1^{-1} = \left(\begin{array}{c|c} \lambda_1 & *\ldots* \\ \hline 0 & A_2 \end{array}\right)$$

2. Schritt: Berechne einen Eigenvektor $\widetilde{v}_2 = \begin{pmatrix} v_{22} \\ \vdots \\ v_{2n} \end{pmatrix} \in K^{n-1}$ der $(n-1) \times (n-1)$-Matrix A_2 zum Eigenwert λ_2. Ergänze den Vektor $\widetilde{v}_2 \in K^{n-1}$ durch Zufügen einer beliebigen ersten Komponente zu einem Vektor $v_2 \in K^n$. Wähle $j \geq 2$ so dass $v_{2,j} \neq 0$ und setze

$$S_2 = (e_1, v_2, e_2, \ldots, \widehat{e_j}, \ldots, e_n)^{-1} \in GL(n, K)$$

Dann ist

$$S_2 S_1 A S_1^{-1} S_2^{-1} e_1 = S_2 S_1 A S_1^{-1} e_1 = \lambda_1 S_2 e_1 = \lambda_1 e_1$$

$$S_2 S_1 A S_1^{-1} S_2^{-1} e_2 = S_2 S_1 A S_1^{-1} v_2 = S_2 \begin{pmatrix} * \\ \lambda_2 v_2 \end{pmatrix} = \begin{pmatrix} * \\ \lambda_2 \\ 0 \\ \vdots \end{pmatrix}$$

Also

$$S_2 S_1 A S_1^{-1} S_2^{-1} = \left(\begin{array}{cc|c} \lambda_1 & * & *\ldots* \\ 0 & \lambda_2 & *\ldots* \\ \hline & 0 & A_3 \end{array}\right)$$

Nach $n-1$ Schritten liefert der Algorithmus eine obere Dreiecksmatrix.

Beispiel: Die Matrix

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

hat, wie eine kurze Rechnung zeigt, das charakteristische Polynom $P_A(X) = (X-2)^3$, also den einzigen Eigenwert 2.

1. Schritt: Wir bestimmen einen Eigenvektor zum Eigenwert 2:

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

und erhalten

$$S_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$S_1 A S_1^{-1} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 4 & 2 \\ 0 & -2 & 0 \end{pmatrix}$$

2. Schritt: Ein Eigenvektor zum Eigenwert 2 von $\begin{pmatrix} 4 & 2 \\ -2 & 0 \end{pmatrix}$ ist $\widetilde{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Setze

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{und} \quad S_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Man findet die zu A ähnliche obere Dreiecksmatrix

$$S_2 S_1 A S_1^{-1} S_2^{-1} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix} .$$

Diese ist natürlich nicht eindeutig bestimmt.

5.5 Das Minimalpolynom

Wir erinnern an den Einsetzungshomomorphismus für den Polynomring aus Bemerkung 5.2.6.1: sei K ein Körper und A eine K–Algebra mit Eins. Dann gibt es für jedes Element $a \in A$ einen eindeutigen Algebrenhomomorphismus

$$K[X] \to \mathrm{Abb}(A, A)$$
$$f \mapsto \widetilde{f}$$

mit $\tilde{f}(a) = \varphi_a(f) \in A$ für $a \in A$.

Definition 5.5.1

1. *Sei R ein Ring. Eine Teilmenge $I \subset R$ heißt <u>Linksideal</u> (bzw. <u>Rechtsideal</u>) von R, falls gilt:*

   (I1) I ist eine Untergruppe von $(R, +)$

   (I2) Für alle $r \in R$ und $a \in I$ gilt $r \cdot a \in I$ (bzw. $a \cdot r \in I$)

   I heißt ein <u>(beidseitiges) Ideal</u>, falls es ein Links- und Rechtsideal ist.
   Ideale von Algebren über Körpern sind entsprechend als Untervektorräume mit den gleichen multiplikativen Eigenschaften definiert.

2. *Sei K ein Körper und A eine unitäre K–Algebra. Sei $a \in A$; dann heißt*

$$\mathrm{Ann}(a) := \{f \in K[X] \mid \widetilde{f}(a) = 0\} \subset K[X]$$

   der <u>Annihilator</u> von a.

Beispiele für Ideale im Ring $\mathbb{Z}$ der ganzen Zahlen sind die geraden Zahlen $2\mathbb{Z}$ oder die durch 3 teilbaren Zahlen $3\mathbb{Z}$. In einem kommutativen Ring wie dem Polynomring oder dem Ring der ganzen Zahlen fallen die Begriffe Linsideal, Rechtsideal und beidseitiges Ideal zusammen.

Lemma 5.5.2.
Der Annihilator ist ein Ideal der Polynomalgebra $K[X]$.

Beweis.

- $\mathrm{Ann}(a)$ ist offenbar ein Untervektorraum von $K[X]$.
- Ist $f \in \mathrm{Ann}(a)$ und $g \in K[X]$ beliebig, so gilt

$$\widetilde{f \cdot g}(a) = \tilde{f}(a) \cdot \tilde{g}(a) = 0 \cdot \tilde{g}(a) = 0 \,,$$

  also $f \cdot g \in \mathrm{Ann}(a)$.

□

Beispiele 5.5.3.
Wir betrachten den Annhilator einer $n \times n$ Matrix in der K-Algebra $M(n \times n, K)$. Polynome $f \in K[X]$ schreiben wir in der Form $f = a_m X^m + \ldots + a_0$. Man beachte, dass wir hier keine Skalare, sondern Elemente der K-Algebra $M(n \times n, K)$ in ein Polynom einsetzen.

1. Sei $A = 0_n$ die Nullmatrix. Dann ist $\widetilde{f}(A) = a_0 E_n$, d.h. $f \in \mathrm{Ann}(A)$ genau dann, wenn $a_0 = 0$ gilt. Also finden wir

$$\mathrm{Ann}(0_n) = \{f \mid \ f = gX \ \text{ mit } \ g \in K[X]\} \,.$$

2. Sei $A = E_n$ die Einheitsmatrix. Dann ist $\widetilde{f}(A) = (a_m + \ldots + a_1 + a_0)E_n$, also

$$\begin{aligned} \mathrm{Ann}(E_n) &= \{f \in K[X] \mid \ \sum_{i=0}^{m} a_i = 0\} \\ &= \{f \mid \ f = (X-1)g \ \text{ mit } g \in K[X]\} \quad , \end{aligned}$$

   da $1 \in K$ Nullstelle aller Polynome im Annihilator ist.

3. Sei $n = 2$ und $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Es gilt $A^2 = 0$ und daher $\widetilde{f}(A) = a_1 A + a_0 E_2 = \begin{pmatrix} a_0 & a_1 \\ 0 & a_0 \end{pmatrix}$. Also finden wir für den Annihilator

$$\mathrm{Ann}(A) = \{f \in K[X] \mid \ f = gX^2 \ \text{ mit } g \in K[X]\} \,.$$

Wir wollen nun alle Ideale des Polynomrings $K[X]$ über einem Körper K beschreiben.

Satz 5.5.4.
Sei K ein Körper und $I \subset K[X]$ ein Ideal, $I \neq \{0\}$. Dann existiert genau ein normiertes Polynom $h \in K[X]$, so dass

$$I = \{f \in K[X] \mid \ f = gh \ \text{ mit } \ g \in K[X]\} \quad \text{gilt.}$$

Jedes Ideal in $K[X]$ besteht also aus den Vielfachen eines eindeutig bestimmten normierten Polynoms h.
Mit $g = 1$ finden wir insbesondere $h \in I$. Einen Ring, bei dem alle Ideale von der Form sind, dass sie Vielfache eines gewissen Rings sind, heißt auch Hauptidealring.

Beweis.
Zentral im Beweis ist die Division mit Rest aus Satz 5.2.10:

- Eindeutigkeit von h: gilt $I = hK[X]$, so gilt für alle $f \in I \setminus \{0\}$ wegen $f = gh$ mit $g \in K[X]$
$$\operatorname{grad}(f) = \operatorname{grad}(gh) = \operatorname{grad}(g) + \operatorname{grad}(h) \geq \operatorname{grad} h \ .$$
h hat also unter den normierten Polynomen in $I \setminus \{0\}$ minimalen Grad.
Gäbe es zwei normierte Polynome $h_1, h_2 \in I$ mit
$$I = h_1 K[X] = h_2 K[X] \ ,$$
so folgte $h_1 - h_2 \in I$, aber $\operatorname{grad}(h_1 - h_2) < \operatorname{grad}(h_1)$, also $h_1 - h_2 = 0$.

- Existenz von h. Sei $\widetilde{h} \in I \setminus \{0\}$ von minimalen Grad:
$$\widetilde{h} = a_n X^n + \ldots a_0 \quad \text{mit} \quad a_n \neq 0 \ .$$
Da I ein Ideal ist, folgt auch für das normierte Polynom $h := (a_n)^{-1} \widetilde{h} \in I$. Offenbar folgt aus der Idealeigenschaft von I sofort $K[X] \cdot h \subset I$. Wir zeigen die andere Inklusion: sei $f \in I \setminus \{0\}$. Polynomdivision durch h mit Rest liefert eine Darstellung
$$f = q \cdot h + r \quad \text{mit} \quad \operatorname{grad}(r) < \operatorname{grad}(h) \ .$$
Aber wegen $r = f - qh$ liegt auch r im Ideal I und hat kleineren Grad als h, was minimalen Grad haben soll; also gilt $r = 0$; das heißt aber $f \in K[X] \cdot h$.

□

Eine genauere Betrachtung des Beweises zeigt, dass im Argument lediglich die Tatsache eine Rolle spielt, dass der Polynomring ein Ring mit Division mit Rest ist. In jedem solchen Ring, also auch im Ring $\mathbb{Z}$ der ganzen Zahlen, gelten analoge Aussagen über die Ideale: sie bestehen aus den Vielfachen eines festen Ringelements.

Definition 5.5.5
Sei A eine K–Algebra mit Eins, $a \in A$. Dann heißt das eindeutig bestimmte normierte *Polynom $\mu_a \in K[X]$ mit*
$$\operatorname{Ann}(a) = \{f \in K[X] \mid \ \widetilde{f}(a) = 0\} = \mu_a \cdot K[X]$$
das Minimalpolynom von a.

Beispiele 5.5.6.
Wir fassen für die Beispiele aus 5.5.3 zusammen:

A	μ_A	P_A
0_n	X	X^n
E_n	$X-1$	$(X-1)^n$
$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	X^2	X^2

In all diesen Beispielen ist das Minimalpolynom μ_A ein Teiler des charakteristischen Polynoms P_A. Dies gilt für alle quadratischen Matrizen:

Theorem 5.5.7. Cayley-Hamilton
Sei $A \in M(n \times n, K)$. Dann gilt für das charakteristische Polynom

$$\widetilde{P_A}(A) = 0_n \ .$$

In anderen Worten: $P_A \in \mathrm{Ann}(A)$; das Minimalpolynom μ_A ist also ein Teiler des charakteristischen Polynoms P_A.

Beweis.

- Sei $v \in K^n \setminus \{0\}$ beliebig. Setze

$$v_i := A^i v \quad i = 0, 1, \ldots$$

Wähle m so, dass die Familie $(v_0, v_1, \ldots, v_{m-1})$ linear unabhängig, aber die Familie $(v_0, \ldots, v_m)$ linear abhängig ist. Es ist $m \leq n$, und es gilt

$$v_m = -\alpha_0 v_0 - \ldots - \alpha_{m-1} v_{m-1} \quad \text{mit gewissen } \ \alpha_i \in K \ . \qquad (*)$$

- Sei $W = \mathrm{span}_K(v_0, \ldots, v_{m-1})$. Dann ist offenbar W ein A–invarianter Untervektorraum, denn die darstellende Matrix von $A_{|W}$ bezüglich der Basis $(v_0, \ldots, v_{m-1})$ hat die Gestalt

$$\left(\begin{array}{ccccc|c} 0 & 0 & 0 & & & -\alpha_0 \\ 1 & 0 & 0 & & & -\alpha_1 \\ 0 & 1 & 0 & & & \vdots \\ 0 & 0 & 1 & & & \vdots \\ & & & \ddots & & \vdots \\ & & & & 1 & -\alpha_{m-1} \end{array}\right) .$$

- Wir berechnen das charakteristische Polynom der Matrix $A|_W$:

$$P_{A|_W}(X) = \det \left(\begin{array}{ccccc|c} X & & & & & \alpha_0 \\ -1 & X & & & & \\ & -1 & & & & \vdots \\ & & & \ddots & X & \alpha_{m-2} \\ & & & & -1 & X + \alpha_{m-1} \end{array}\right) .$$

Die Entwicklung nach der letzten Spalte liefert:

$$
\begin{aligned}
P_{A|W}(X) = (-1)^{m+1}\alpha_0 \cdot \det \begin{pmatrix} -1 & X & & & \\ & -1 & X & & \\ & & -1 & & \\ & & & \ddots & \\ & & & & X \\ & & & & -1 \end{pmatrix} \\
+(-1)^{m+2}\alpha_1 \cdot \det \begin{pmatrix} X & 0 & & & & \\ 0 & -1 & X & & & 0 \\ & 0 & -1 & X & & \\ & & & & \ddots & \\ & & & & & X \\ 0 & & & & & -1 \end{pmatrix} \\
+(-1)^{m+3}\alpha_2 \cdot \det \begin{pmatrix} X & 0 & & & & \\ -1 & X & 0 & & & 0 \\ & 0 & -1 & X & & \\ & & & & \ddots & \\ & & & & & X \\ 0 & & & & & -1 \end{pmatrix} \\
+ \ldots + (-1)^{m+(m-1)}\alpha_{m-2} \cdot \det \begin{pmatrix} X & & & & 0 & \\ -1 & X & & & & \\ & -1 & X & & & \\ & & & \cdots & & \\ & & & & X & \\ & & & & -1 & -1 \end{pmatrix} \\
+(-1)^{2m}\left(\alpha_{m-1} + X\right) \cdot \det \begin{pmatrix} X & & & 0 & \\ -1 & X & & & \\ & -1 & & & \\ & & & \ddots & \\ & & & -1 & X \end{pmatrix} \\
= \alpha_0 \cdot 1 + \alpha_1 X + \ldots + \alpha_{m-2}X^{m-2} + \alpha_{m-1}X^{m-1} + X^m
\end{aligned}
$$

- Nach Lemma 5.4.1 existiert ein Polynom $g \in K[X]$ mit

$$P_A(X) = g(X)P_{A|_W}(X) \; .$$

Es folgt

$$
\begin{aligned}
\widetilde{P_A}(A)v &= \widetilde{g}(A)\widetilde{P_{A|_W}}(A)v \\
&= \widetilde{g}(A)\left(A^m + \alpha_{m-1}A^{m-1} + \ldots + \alpha_1 A + \alpha_0\right)v \\
&= \widetilde{g}(A)\left(v_m + \alpha_{m-1}v_{m-1} + \ldots + \alpha_1 v_1 + \alpha_0 v\right) = 0 \; ,
\end{aligned}
$$

wobei wir im letzten Schritt $(*)$ verwendet haben. Da dies für alle $v \in V$ gilt, ist die Behauptung gezeigt.

□

Satz 5.5.8.
Sei $A \in M(n \times n, K)$. Dann haben das charakteristische Polynom $P_A(X)$ und das Minimalpolynom $\mu_A(X)$ dieselben Nullstellen. Die Vielfachheit einer Nullstelle $\lambda \in K$ als Nullstelle des Minimalpolynoms ist dabei kleiner als die Vielfachheit als Nullstelle des charakteristischen Polynoms.

Beweis.

- Nach Theorem 5.5.7 von Cayley-Hamilton gilt $P_A = g\mu_A$ mit $g \in K[X]$. Jede Nullstelle des Minimalpolynoms μ_A ist also auch Nullstelle des charakteristischen Polynoms P_A mit mindestens der gleichen Multiplizität.

- Sei $\lambda \in K$ eine Nullstelle von P_A, also ein Eigenwert. Wähle einen zugehörigen Eigenvektor v. Mit
$$\mu_A(X) = X^m + \alpha_{m-1}X^{m-1} + \ldots + \alpha_0$$
rechnen wir mit der definierenden Eigenschaft $\widetilde{\mu}_A(A) = 0_n$ des Minimalpolynoms:
$$\begin{aligned} 0 = 0_n \cdot v = \widetilde{\mu_A}(A)v &= \left(A^m + \alpha_{m-1}A^{m-1} + \ldots + \alpha_0\right) v \\ &= \left(\lambda^m + \alpha_{m-1}\lambda^{m-1} + \ldots + \alpha_0\right) v = \widetilde{\mu_A}(\lambda)v \ . \end{aligned}$$
Hieraus folgt aber wegen $v \neq 0$, dass $\widetilde{\mu_A}(\lambda) = 0$. Also ist der Eigenwert λ auch Nullstelle des Minimalpolynoms.

□

Satz 5.5.9.
Sei $A \in M(n \times n, K)$. Dann ist äquivalent:

1. Die Matrix A ist diagonalisierbar.
2. Das Minimalpolynom μ_A zerfällt in paarweise verschiedene Linearfaktoren.

Beweis.

2. $\Rightarrow$ 1. werden wir in Kapitel 7 aus der Jordanschen Normalform einer Matrix herleiten, vergleiche Bemerkung 7.3.8.5.

1. $\Rightarrow$2. Sei A diagonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_m$. Setze
$$f := \prod_{i=1}^{m}(X - \lambda_i) \in K[X] \ .$$
Alle Nullstellen von f sind Nullstellen des charakteristischen Polynoms P_A, also nach Satz 5.5.8 auch Nullstellen von μ_A. Somit teilt f das Minimalpolynom μ_A.

- Da A diagonalisierbar sein soll, existiert eine Eigenraumzerlegung
$$K^n \cong \oplus_{i=1}^{m}\mathrm{Eig}(A, \lambda_i) \ .$$
Schreibe einen beliebigen Vektor $v \in K^n$ als
$$v = \sum_{i=1}^{m} v_i \quad \text{mit} \ \ v_i \in \mathrm{Eig}(A, \lambda_i) \ .$$

Dann ist

$$\widetilde{f}(A)v_j = \prod_{\substack{i=1 \\ j\neq i}}^{m}(A - \lambda_i E_n) \cdot (A - \lambda_j E_n)v_j = 0 \,,$$

also $\widetilde{f}(A)v = 0 \quad$ für alle $v \in V$. Somit $\widetilde{f}(A) = 0$, und μ_A teilt f nach Definition 5.5.5 des Minimalpolynoms als Erzeuger des Annihilatorideals von A.

- Also gibt es $\alpha \in K \setminus \{0\}$ mit $\mu_A = \alpha f$. Da beide Polynome normiert sind, ist $\alpha = 1$.

□

Betrachtung 5.5.10.

- Schema zur Untersuchung der Diagonalisierbarkeit einer quadratischen Matrix $A \in M(n\times n, K)$:

1. Schritt: Berechne das charakteristische Polynom P_A und seine Zerlegung in Linearfaktoren. Zerfällt P_A nicht in Linearfaktoren, so ist A nach Satz 5.3.4 nicht diagonalisierbar. Zerfällt P_A in Linearfaktoren, gehe zu

2. Schritt: Setze $f(X) := \prod_{i=1}^{m}(X - \lambda_i) \in K[X]$, wobei $\lambda_1, \ldots, \lambda_m$ die verschiedenen Nullstellen des charakteristischen Polynoms P_A sind.
Gilt $\widetilde{f}(A) = 0_n$; so ist A nach Satz 5.5.9 diagonalisierbar.
Gilt $\widetilde{f}(A) \neq 0_n$; so ist A nach Satz 5.5.9 nicht diagonalisierbar.

- Als Beispiel betrachten wir die Matrix

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

mit charakteristischem Polynom

$$P_A(X) = \det \begin{pmatrix} X-2 & -1 \\ 0 & X-2 \end{pmatrix} = (X-2)^2 \,.$$

Dieses zerfällt über $\mathbb{Q}$ vollständig in Linearfaktoren. Für Schritt 2 setzen wir also $f(X) = X - 2$, und berechnen

$$\widetilde{f}(A) = A - 2E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq 0_n \,.$$

Die Matrix A ist daher nicht diagonalisierbar.

6 Bilinearformen; Euklidische und unitäre Vektorräume

6.1 Der Dualraum

Auch als Vorbereitung für das Studium von Bilinearformen wollen wir zunächst eine besonders wichtige Klasse linearer Abbildungen betrachten.

Definition 6.1.1
Sei K ein beliebiger Körper und V ein K–Vektorraum. Dann heißt der K-Vektorraum

$$V^* := \mathrm{Hom}_K(V, K) = \{\varphi : V \to K |\ \varphi \text{ linear}\}$$

der Dualraum von V. Die Elemente von V^* heißen Linearformen auf V.

Beispiele 6.1.2.
Sei $K = \mathbb{R}$ und $V = C^0([a, b], \mathbb{R})$ der Vektorraum der stetigen reellwertigen Funktionen auf dem abgeschlossenen Intervall $[a, b]$. Dann sind die Abbildungen

$$\varphi, \psi :\ V \to \mathbb{R}$$

mit

$$\psi(f) = f(a) \quad \text{und} \quad \varphi(f) = \int_a^b f(x)\mathrm{d}x$$

oder allgemeiner für $x \in [a, b]$ bzw. $g \in V$ die Abbildungen

$$\psi_x(f) = f(x) \quad \text{und} \quad \varphi_g(f) = \int_a^b f \cdot g\mathrm{d}x$$

Linearformen auf V.

Betrachtung 6.1.3.
Sei $\dim_K V < \infty$. Aus Korollar 3.2.18 folgt

$$\dim_K V^* = \dim_K \mathrm{Hom}_K(V, K) = \dim_K V \cdot \dim_K K = \dim_K V\ .$$

Sei nun $\mathcal{B}$ eine Basis von V, $\mathcal{B} = \{b_1, \ldots, b_n\}$. Definiere Linearformen $b_j^* \in V^*$ durch ihre Werte auf den Basisvektoren b_i, vgl. Satz 3.2.1,

$$b_j^*(b_k) = \delta_{j,k}\ ,$$

woraus sofort folgt

$$b_j^*\left(\sum_{k=1}^n \alpha_k b_k\right) = \alpha_j\ .$$

Wir behaupten, dass $\mathcal{B}^* = \{b_1^*, \ldots, b_n^*\}$ eine Basis von V^* ist. Wegen $\dim V^* = \dim V = n$ reicht es zu zeigen, dass $\mathcal{B}^*$ linear unabhängig ist. Gelte also

$$\sum_{k=1}^n \alpha_k b_k^* = 0 \quad \text{mit} \quad \alpha_k \in K\ ;$$

dann gilt aber für jedes $j = 1, \ldots, n$:

$$0 = \sum_{k=1}^n \alpha_k b_k^*(b_j) = \alpha_j\ .$$

Definition 6.1.4
Die Basis $\mathcal{B}^*$ von V^* heißt die zu $\mathcal{B}$ duale Basis. Ist $\mathcal{B}$ eine geordnete Basis, so ist auch die duale Basis $\mathcal{B}^*$ in natürlicher Weise eine geordnete Basis.

Lemma 6.1.5.
Sei V ein beliebiger K–Vektorraum und sei $v \in V \setminus \{0\}$. Dann gibt es eine Linearform $\varphi \in V^*$ mit $\varphi(v) \neq 0$.

Beweis.
Ergänze (v) zu einer Basis $(v, w_1, \ldots, w_{n-1})$ von V. Definiere die Linearform $\varphi \in V^*$ durch ihre Werte auf dieser Basis, vgl. Satz 3.2.1, und setze etwa

$$\varphi(v) = 1 \qquad \varphi(w_i) = 0 \qquad i = 1, \ldots, n-1 \; .$$

□

Bemerkung 6.1.6.
Wegen $\dim_K V = \dim_K V^*$ für $\dim_K V < \infty$ sind ein endlich-dimensionaler Vektorraum und sein Dualraum isomorph. Jede geordnete Basis $\mathcal{B}$ von V liefert einen Isomorphismus

$$\begin{array}{rccc} \Psi_{\mathcal{B}} : & V & \to & V^* \\ & b_i & \mapsto & b_i^* \end{array}$$

mit $\Psi_{\mathcal{B}}(b_i) := b_i^*$. Dieser Isomorphismus ist jedoch *nicht* kanonisch, d.h. für verschiedene Basen $\mathcal{B}, \mathcal{B}'$ sind die Isomorphismen $\Psi_{\mathcal{B}}$ und $\Psi_{\mathcal{B}'}$ im allgemeinen verschieden.

Sei etwa $\alpha \in K \setminus \{0\}$, und betrachte eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ und die reskalierte Basis $\mathcal{B}' = (\alpha b_1, \ldots, \alpha b_n)$. Für die Elemente der dualen Basis $(\mathcal{B}')^*$ gilt

$$\delta_{jk} = (b'_j)^*(b'_k) = (b'_j)^*(\alpha b_k) = \alpha (b'_j)^*(b_k) \; .$$

Andererseits ist $b_j^*(b_k) = \delta_{jk}$, so dass folgt $(b'_j)^* = \alpha^{-1} b_j^*$. Somit ist die duale Basis $(\mathcal{B}')^* = (\alpha^{-1} b_1^*, \ldots, \alpha^{-1} b_n^*)$.

Per Definition gilt für den von der Basis $\mathcal{B}'$ erzeugten Isomorphismus

$$\Psi_{\mathcal{B}'}(b'_i) = (b'_i)^*$$

und somit für alle $1 \leq i \leq n$

$$\Psi_{\mathcal{B}'}(b_i) = \Psi_{\mathcal{B}'}(\alpha^{-1} b'_i) = \alpha^{-1} \Psi_{\mathcal{B}'}(b'_i) = \alpha^{-1} (b'_i)^* = \alpha^{-2} b_i^* = \alpha^{-2} \Psi_{\mathcal{B}}(b_i) \; .$$

Es folgt $\Psi_{\mathcal{B}'} = \alpha^{-2} \Psi_{\mathcal{B}}$, so dass die Isomorphismen $\Psi_{\mathcal{B}'}$ und $\Psi_{\mathcal{B}}$ für $\alpha \neq \pm 1$ verschieden sind.

Da es also für eine Identifizierung eines Vektorraums mit seinem Dualraum keinen *ausgezeichneten* Isomorphismus gibt, dürfen die beiden Vektorräume V und V^* selbst für endlich-dimensionale Vektorräume nicht einfach miteinander identifiziert werden.

Definition 6.1.7
Sei V ein K–Vektorraum und $M \subset V$ eine Teilmenge. Dann heißt

$$M^0 := \{\varphi \in V^* | \; \varphi(m) = 0 \quad \textit{für alle } m \in M\} \subset V^*$$

der Annulator der Teilmenge M.

Lemma 6.1.8.
Es gilt:

1. Der Annulator M^0 ist ein Untervektorraum des Dualraums V^*.
2. $\operatorname{span}_K M = \{v \in V |\, \varphi(v) = 0 \text{ für alle } \varphi \in M^0\}$
   Ist insbesondere $M = U$ ein Untervektorraum von V, so gilt
   $$U = \{v \in V |\, \varphi(v) = 0 \text{ für alle } \varphi \in U^0\} .$$
3. Ist $\dim_K V < \infty$ und $U \subset V$ Untervektorraum, so gilt
   $$\dim_K U^\circ = \dim_K V - \dim_K U .$$

Beweis.

1. ist offensichtlich.
2. "$\subset$" Sei $v \in \operatorname{span}_K M$ beliebig. Finde $\alpha_i \in K$ und $m_i \in M$, so dass gilt $v = \sum_{i=1}^{j} \alpha_i m_i$. Für jede Linearform $\varphi \in M^0$ im Annulator folgt aus $\varphi(m_i) = 0$, dass
   $$\varphi(v) = \sum_i \alpha_i \varphi(m_i) = 0 .$$
   "$\supset$" Sei $v \in V$ mit $\varphi(v) = 0$ für alle $\varphi \in M^0$. Angenommen, es wäre $v \notin \operatorname{span}_K M$. Wähle eine geordnete Basis $(u_1, \ldots, u_m)$ von $\operatorname{span}_K M$; dann ist die Familie $(u_1, \ldots, u_m, v)$ linear unabhängig und kann zu einer geordneten Basis $(u_1, \ldots, u_m, v, w_1, \ldots, w_l)$ von V ergänzt werden. Definiere $\varphi \in V^*$ auf dieser Basis durch
   $$\varphi(u_i) = 0 \quad \varphi(v) = \varphi(w_j) = 1 .$$
   Dann ist $\varphi \in M^0$, aber $\varphi(v) = 1$. Dies ist im Widerspruch zur Annahme, dass $\varphi(v) = 0$ für alle $\varphi \in M^0$ gelten soll, also muss $v \in \operatorname{span}_K M$ gelten.
3. Ergänze eine geordnete Basis von U zu einer geordneten Basis $(u_1, \ldots, u_k, v_{k+1}, \ldots, v_n)$ von V. Sei $\mathcal{B}^* = (u_1^*, \ldots, u_k^*, v_{k+1}^*, \ldots, v_n^*)$ die duale Basis von V^*. Dann liegt
   $$\varphi = \sum_{i=1}^{k} \alpha_i u_i^* + \sum_{j=k+1}^{n} \alpha_i v_i^* \ \in \ V^*$$
   genau dann in U^0, wenn für alle $j = 1, \ldots, k$ gilt
   $$0 = \varphi(u_j) = \alpha_j$$
   Dies ist aber genau der Fall für $\varphi \in \operatorname{span}_K(v_{k+1}^*, \ldots, v_n^*)$. Somit ist $(v_{k+1}^*, \ldots, v_n^*)$ eine Basis von $U^0 \subset V^*$, und es gilt
   $$\dim_K U^0 = n - k = \dim_K V - \dim_K U .$$

□

Beispiel 6.1.9.

1. Sei $\dim_K V < \infty$ und $U \subset V$ eine Hyperebene, d.h. ein Untervektorraum der Dimension

$$\dim_K U = \dim_K V - 1 \ .$$

Nach Lemma 6.1.8.3 ist

$$\dim_K U^0 = n - (n-1) = 1 \ .$$

Also ist $U^0 = \text{span}_K(\varphi)$ für jedes $\varphi \in U^0 \setminus \{0\}$. Nach Lemma 6.1.8.2 gilt für jede solche Linearform φ

$$U = \{v \in V | \, \varphi(v) = 0\} \ .$$

2. Betrachten wir nun allgemeiner die affine Hyperebene $v_0 + U$, mit U wie oben und $v_0 \in V$, so gilt

$$\begin{aligned} v \in v_0 + U &\Leftrightarrow v - v_0 \in U \\ &\Leftrightarrow \varphi(v - v_0) = 0 \\ &\Leftrightarrow \varphi(v) = \varphi(v_0) =: c \end{aligned}$$

Wir finden also für jede affine Hyperebene eine Darstellung

$$U + v_0 = \{v \in V | \, \varphi(v) = c\}$$

Dies ist eine Hessesche Normalform der affinen Hyperebene $v_0 + U$. Sie ist eine koordinatenfreie Version der Gleichungsform einer Hyperebene, für die man kein Skalarprodukt auf V einführen muss und stattdessen mit dem Dualraum arbeitet.

Definition 6.1.10
Seien V, W zwei K–Vektorräume und sei $\Phi : V \to W$ eine lineare Abbildung. Dann heißt die Abbildung

$$\begin{aligned} \Phi^* : W^* &\to V^* \\ \varphi &\mapsto \varphi \circ \Phi \end{aligned}$$

die zu Φ duale Abbildung.

Man beachte, dass

$$\varphi \circ \Phi : \quad V \xrightarrow{\Phi} W \xrightarrow{\varphi} K$$

eine Linearform am V ist.

Lemma 6.1.11.

1. Φ^* ist eine lineare Abbildung.
2. Seien V, W, Z jeweils K-Vektorräume und $\Phi : V \to W$ und $\Psi : W \to Z$ lineare Abbildungen. Dann gilt $(\Psi \circ \Phi)^* = \Phi^* \circ \Psi^*$.

Beweis.

1. Seien $\varphi_1, \varphi_2 \in W^*$ und $\alpha_1, \alpha_2 \in K$. Wir rechnen für $v \in V$

$$\begin{aligned} \Phi^* \left(\alpha_1\varphi_1 + \alpha_2\varphi_2\right)(v) &= \left(\alpha_1\varphi_1 + \alpha_2\varphi_2\right)(\Phi v) \\ &= \alpha_1\varphi_1(\Phi v) + \alpha_2\varphi_2(\Phi v) \\ &= \alpha_1\Phi^*\varphi_1(v) + \alpha_2\Phi^*\varphi_2(v) \\ &= \left(\alpha_1\Phi^*\varphi_1 + \alpha_2\Phi^*\varphi_2\right)(v) \ . \end{aligned}$$

2. Es gilt für alle $\varphi \in Z^*$ wegen der Assoziativität der Verkettung von Abbildungen

$$(\Psi \circ \Phi)^*(\varphi) = \varphi \circ (\Psi \circ \Phi) = (\varphi \circ \Psi) \circ \Phi = (\Psi^* \varphi) \circ \Phi = \Phi^* (\Psi^*(\varphi)) = (\Phi^* \circ \Psi^*)(\varphi) \ .$$

□

Satz 6.1.12.
Seien V, W endlich–dimensionale K–Vektorräume mit geordneten Basen $\mathcal{A}$ und $\mathcal{B}$. Sei $\Phi : V \to W$ eine lineare Abbildung. Dann ist die darstellende Matrix der dualen Abbildung bezüglich der dualen Basen:

$$M_{\mathcal{A}^*}^{\mathcal{B}^*}(\Phi^*) = M_{\mathcal{B}}^{\mathcal{A}}(\Phi)^t \ .$$

Beweis.

- Es gilt mit $M_{\mathcal{B}}^{\mathcal{A}}(\Phi) = (c_{kj})$ und $M_{\mathcal{B}^*}^{\mathcal{A}^*}(\Phi^*) = (d_{\mu\nu})$ nach der Definition 3.2.17 der darstellenden Matrix

$$\Phi(a_j) = \sum_{k=1}^{m} c_{kj} b_k \quad \text{mit } m := \dim W, \qquad \text{für alle } j = 1, 2, \ldots, \dim V$$

$$\Phi^*\left(b_\mu^*\right) = \sum_{\nu=1}^{n} d_{\nu\mu} a_\nu^* \quad \text{für alle } \mu = 1, 2, \ldots, \dim W \qquad \text{mit } n := \dim V$$

- Wir vergleichen das Ergebnis der folgenden Rechnungen für festes $j = 1, \ldots, n$

$$b_\mu^*(\Phi(a_j)) = b_\mu^* \left(\sum_{k=1}^{m} c_{kj} b_k \right) = \sum_{k=1}^{m} c_{kj} b_\mu^*(b_k) = c_{\mu j}$$

$$b_\mu^*(\Phi(a_j)) = \Phi^*\left(b_\mu^*\right)(a_j) = \sum_{\nu=1}^{n} d_{\nu\mu} a_\nu^*(a_j) = d_{j\mu} \ .$$

□

Damit haben wir auch der Operation aus Definition 3.2.12.2, eine Matrix zu transponieren, einen basisunabhängigen Sinn im Rahmen von Vektorräumen gegeben.

Satz 6.1.13.
Seien V, W endlich-dimensionale K–Vektorräume und sei $\Phi : V \to W$ linear. Dann gilt für den Kern und das Bild der dualen Abbildung $\Phi^* : W^* \to V^*$, dass sie sich als Annulatoren ausdrücken lassen:

$$\begin{aligned} \text{Im } \Phi^* &= (\ker \Phi)^\circ \ \subset \ V^* \\ \ker \Phi^* &= (\text{Im } \Phi)^\circ \ \subset \ W^* \end{aligned}$$

Beweis.

- Wir zeigen zuerst die Inklusion Im $\Phi^* \subset (\ker \Phi)^\circ$. Sei $\varphi \in$ Im Φ^*, d.h. $\varphi = \Phi^*(\psi)$ für ein $\psi \in W^*$. Für jedes $v \in \ker \Phi$ gilt dann

$$\varphi(v) = \Phi^*(\psi)(v) \overset{\text{def von } \Phi^*}{=} \psi(\Phi(v)) = \psi(0) = 0 \ ,$$

  also gilt $\varphi \in (\ker \Phi)^\circ$.

- Sei umgekehrt $\varphi \in (\ker \Phi)^\circ$. Wir müssen ein Urbild von φ konstruieren. Finde dazu mit Hilfe von Lemma 3.6.9 geordnete Basen $\mathcal{A} = (v_1, \ldots, v_n)$ von V und $\mathcal{B} = (w_1, \ldots, w_m)$ von W, so dass die darstellende Matrix von Φ die Gestalt

$$M_{\mathcal{B}}^{\mathcal{A}}(\Phi) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

hat. Definiere $\psi \in W^*$ auf der Basis $\mathcal{B}$ von W durch

$$\psi(w_j) = \begin{cases} \varphi(v_j) \in K & j = 1, \ldots, r \\ 0 & \text{für } j \geq r+1 \, . \end{cases}$$

Dann gilt für alle $j = 1, \ldots, r$

$$\Phi^*(\psi)(v_j) = \psi(\Phi v_j) = \psi(w_j) = \varphi(v_j)$$

und für $j \geq r+1$

$$\Phi^*(\psi)(v_j) = \psi(\Phi v_j) = \psi(0) = 0 \ .$$

Andererseits folgt aus $\ker \Phi = \operatorname{span}_K(v_{r+1}, \ldots, v_n)$ und $\varphi \in (\ker \Phi)^\circ$ ebenfalls $\varphi(v_j) = 0$. Also ist $\Phi^*(\psi) = \varphi$, und daher ist $\varphi \in \operatorname{Im}\ \Phi^*$.

- Die andere Aussage folgt mit ähnlichen Argumenten.

□

Korollar 6.1.14.
Eine lineare Abbildung $\Phi : V \to W$ und ihre duale Abbildung $\Phi^* : W^* \to V^*$ haben gleichen Rang, es gilt $\operatorname{rg} \Phi^* = \operatorname{rg}(\Phi)$.

Beweis.

$$\begin{aligned} \operatorname{rg} \Phi^* &\overset{\text{def}}{=} \dim_K \operatorname{Im}\ \Phi^* = \dim_K (\ker \Phi)^\circ && \text{(nach Satz 6.1.13)} \\ &= \dim_K(V) - \dim_K \ker \Phi && \text{(nach Satz 6.1.8.3)} \\ &= \dim_K \operatorname{Im}\ \Phi && \text{(wegen der Dimensionsformel 3.1.7)} \\ &\overset{\text{def}}{=} \operatorname{rg} \Phi \, . \end{aligned}$$

□

Wegen

$$\operatorname{rg} \Phi = \operatorname{rg} M_{\mathcal{B}}^{\mathcal{A}}(\Phi)$$

und

$$\operatorname{rg}(\Phi^*) = \operatorname{rg}\left(M_{\mathcal{A}^*}^{\mathcal{B}^*}(\Phi^*)\right) \overset{6.1.12}{=} \operatorname{rg} M_{\mathcal{B}}^{\mathcal{A}}(\Phi)^t = \widetilde{\operatorname{rg}}\, M_{\mathcal{B}}^{\mathcal{A}}(\Phi)$$

folgt aus Korollar 6.1.14 erneut die Aussage, dass Zeilen- und Spaltenrang einer Matrix übereinstimmen, vergleiche Satz 3.6.15.

Definition 6.1.15
Sei V ein K–Vektorraum. Dann heißt der K-Vektorraum

$$V^{**} := (V^*)^* = \operatorname{Hom}_K(V^*, K)$$

der Bidualraum von V.

Betrachtung 6.1.16.
Betrachte zu einem Vektor $v \in V$ die Abbildung, die eine Linearform $\varphi \in V^*$ auf diesem Vektor auswertet:

$$\begin{aligned} \iota_V(v) : \ V^* &\to K \\ \varphi &\mapsto \varphi(v) \end{aligned}$$

Die Abbildung $\iota_V(v)$ ist linear, denn

$$\iota_V(v)(\alpha_1\varphi_1 + \alpha_2\varphi_2) = (\alpha_1\varphi_1 + \alpha_2\varphi_2)(v) = \alpha_1\varphi_1(v) + \alpha_2\varphi_2(v) = \alpha_1\iota_V(v)\varphi_1 + \alpha_2\iota_V(v)\varphi_2 \ ,$$

also ist $\iota_V(v) \in V^{**}$. Wir haben also für jeden K–Vektorraum V eine Abbildung

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

Satz 6.1.17.

1. ι_V ist eine injektive lineare Abbildung.
2. Ist $\dim_K V < \infty$, so ist ι_V ein Isomorphismus. Er heißt dann kanonischer Isomorphismus.
3. Diese Abbildungen sind funktoriell: seien V, W jeweils K–Vektorräume und sei $\Phi : \ V \to W$ linear. Dann kommutiert das folgende Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & W \\ {\scriptstyle \iota_V}\downarrow & & \downarrow{\scriptstyle \iota_W} \\ V^{**} & \xrightarrow[\Phi^{**}]{} & W^{**} \end{array}$$

Beweis.

1. Es gilt für $\alpha_1, \alpha_2 \in K$, $v_1, v_2 \in V$ und alle $\varphi \in V^*$

   $$\iota_V(\alpha_1 v_1 + \alpha_2 v_2)(\varphi) = \varphi(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1\varphi(v_1) + \alpha_2\varphi(v_2) = \left(\alpha_1\iota_V(v_1) + \alpha_2\iota_V(v_2)\right)(\varphi) \ .$$

   Also ist ι_V eine lineare Abbildung. Sei $v \in V$ mit $\iota_V(v) = 0$. Dann gilt für alle $\varphi \in V^*$: $0 = \iota_V(v)\varphi = \varphi(v)$. Nach Lemma 6.1.5 muss dann aber $v = 0$ gelten.

2. Ist $\dim_K V < \infty$, so gilt

   $$\dim_K V^{**} = \dim_K V^* = \dim_K V \ .$$

   Nach 1. ist ι_V injektiv, also auch bijektiv.

3. Sei $v \in V$ und $\varphi \in W^*$. Es gilt

   $$\left(\Phi^{**} \circ \iota_V(v)\right)(\varphi) = \Phi^{**}\left(\iota_V(v)\right)(\varphi) = \iota_V(v)\left(\Phi^*(\varphi)\right) = \Phi^*(\varphi)(v) = \varphi\left(\Phi(v)\right) \ .$$

   Andererseits ist

   $$\left(\iota_W \circ \Phi(v)\right)\varphi = \iota_W(\Phi(v)(\varphi) = \varphi\left(\Phi(v)\right) \ .$$

   Somit ist

   $$\Phi^{**} \circ \iota_V(v) = \iota_W \circ \Phi(v)$$

   für alle $v \in V$.

□

6.2 Bilinearformen

Definition 6.2.1

1. *Seien V, W zwei K–Vektorräume. Eine Abbildung $\beta : V \times W \to K$ heißt Bilinearform, wenn gilt*

$$\begin{aligned} \beta(v + v', w) &= \beta(v, w) + \beta(v', w) \\ \beta(v, w + w') &= \beta(v, w) + \beta(v, w') \\ \beta(\alpha v, w) &= \alpha\beta(v, w) = \beta(v, \alpha w) \end{aligned}$$

   für alle $v, v' \in V, w, w' \in W$ und $\alpha \in K$.

2. *Eine Bilinearform heißt nicht-ausgeartet im ersten (bzw. zweiten) Argument, falls gilt*

$$\begin{aligned} \beta(v, w) = 0 \quad &\text{für alle } w \in W \ \Rightarrow v = 0 \\ \text{bzw. } \beta(v, w) = 0 \quad &\text{für alle } v \in V \ \Rightarrow w = 0\,. \end{aligned}$$

   Mit $\mathrm{Bil}_K(V, W)$ bezeichnen wir die Menge aller K–Bilinearformen

$$\beta : V \times W \to K\,.$$

Beispiele 6.2.2.

1. Die Multiplikation auf dem Körper K

$$m : K \times K \to K$$

   liefert eine Bilinearform auf K.

2. Sei $B = (b_{ij}) \in M(m \times n, K)$. Dann definiert

$$\begin{aligned} \beta : K^m \times K^n &\to K \\ (x, y) &\mapsto x^t \cdot B \cdot y \end{aligned}$$

   eine Bilinearform.

3. Die Menge $\mathrm{Bil}_K(V, W)$ erhält durch Operationen auf den Abbildungswerten die Struktur eines K–Vektorraums.

4. Für jeden K–Vektorraum V liefert die Evaluationsabbildung

$$\begin{aligned} ev_V : V^* \times V &\to K \\ (\beta, v) &\mapsto \beta(v) =: \langle \beta, v \rangle \end{aligned}$$

   eine Bilinearform. Wegen Lemma 6.1.5 folgt aus $\langle \beta, v \rangle = 0$ für alle $\beta \in V^*$, dass $v = 0$. Also ist die Bilinearform nicht-ausgeartet im zweiten Argument. Gilt $\langle \beta, v \rangle = 0$ für alle $v \in V$, so ist nach Definition β der Nullvektor in V^*; also ist die Bilinearform $\langle \cdot, \cdot \rangle$ auch im ersten Argument nicht-ausgeartet.

5. Sei $\beta : V \times W \to K$ eine Bilinearform. Zu jedem festen $x \in V$ betrachte die Abbildung

$$\begin{aligned} \lambda_x : \ & W \to K \\ & y \mapsto \beta(x, y) \end{aligned}$$

Diese ist wegen der Linearität von β im zweiten Argument linear, also ist $\lambda_x \in W^*$. Wir haben also eine Abbildung

$$\begin{array}{rccc} \beta_1 : & V & \to & W^* \\ & x & \mapsto & \lambda_x \ . \end{array}$$

Wegen der Linearität von β im ersten Argument ist β_1 linear. Es gilt also

$$\beta(x, y) = \beta_1(x)(y) = \langle \beta_1(x), y \rangle$$

mit der kanonischen Paarung $\langle \cdot, \cdot \rangle$ von W und seinem Dualraum W^*.

Lemma 6.2.3.

1. Die Abbildung

$$\begin{aligned} \mathrm{Bil}_K(V, W) &\to \mathrm{Hom}_K(V, W^*) \\ \beta &\mapsto \beta_1 \end{aligned}$$

ist ein Isomorphismus von K–Vektorräumen. Wir können also Bilinearformen mit Hilfe linearer Abbildungen untersuchen.

2. Gilt $\dim_K V < \infty$ und $\dim_K W < \infty$, so ist $\dim_K \mathrm{Bil}_K(V, W) = \dim_K V \cdot \dim_K W$.

3. Die Bilinearform β ist genau dann nicht-ausgeartet im ersten Argument, falls die lineare Abbildung β_1 injektiv ist.

Beweis.

1. Wir geben eine Umkehrabbildung an: für eine lineare Abbildung $h : \ V \to W^*$, definiere eine Bilinearform $\beta_h : \ V \times W \to K$ mit Hilfe der Evaluationsabbildung für W

$$\beta_h(x, y) := \langle hx, y \rangle$$

2. ist dann klar. 3. folgt deswegen, weil β_1 genau dann nicht injektiv ist, wenn es $v \neq 0$ gibt mit $\beta_1 v = 0$. Genau dann gilt aber $\beta(v, w) = 0$ für alle $w \in W$, d.h. β ist im ersten Argument ausgeartet.

□

Betrachtung 6.2.4.
Sei V ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{A} = (v_1, \ldots, v_n)$ und W ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{B} = (w_1, \ldots, w_m)$. Setze für $\beta \in \mathrm{Bil}_K(V, W)$ $\beta_{ij} := \beta(v_i, w_j) \in K$. Die Matrix $B = (\beta_{ij}) \in M(n \times m, K)$ legt β eindeutig fest: ist $x = \sum_{i=1}^n x_i v_i$ und $y = \sum_{i=1}^m y_i w_i$, so ist

$$\beta(x, y) = \sum_{i,j} x_i \beta(v_i, w_j) y_j = x^t \cdot B \cdot y \ .$$

Matrizen können also sowohl lineare Abbildungen als auch Bilinearformen beschreiben. Man sollte sich bei einer Matrix – die ja eigentlich nur eine rechteckige Anordnung von Skalaren ist – also stets klarmachen, welches mathematische Objekt sie beschreibt.

Definition 6.2.5
Sei V ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{A} = (v_1, \ldots, v_n)$ und W ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{B} = (w_1, \ldots, w_m)$. Die Matrix $M_{\mathcal{A},\mathcal{B}}(\beta) \in M(n \times m, K)$ mit Einträgen $\beta_{ij} := \beta(v_i, w_j) \in K$ heißt darstellende Matrix der Bilinearform β bezüglich der geordneten Basen $\mathcal{A}, \mathcal{B}$.

Sei $\mathcal{A} = (v_1, \ldots, v_n)$ eine geordnete Basis von V, $\mathcal{B} = (w_1, \ldots, w_m)$ eine geordnete Basis von W und $\mathcal{B}^* = (w_1^*, \ldots, w_m^*)$ die duale Basis von W^*. Für die darstellende Matrix von $\beta_1 : V \to W^*$ gilt

$$\beta_1(v_i) = \sum_{j=1}^{m} M_{\mathcal{B}^*}^{\mathcal{A}}(\beta_1)_{ji} w_j^* \ .$$

Damit folgt:

$$M_{\mathcal{A},\mathcal{B}}(\beta)_{ij} = \beta(v_i, w_j) = \langle \beta_1(v_i), w_j \rangle = M_{\mathcal{B}^*}^{\mathcal{A}}(\beta_1)_{ji}$$

und somit $M_{\mathcal{A},\mathcal{B}}(\beta) = M_{\mathcal{B}^*}^{\mathcal{A}}(\beta_1)^t$.

Satz 6.2.6. Transformationsformel

1. Sei V ein endlich–dimensionaler K–Vektorraum mit geordneten Basen $\mathcal{A}$ und $\mathcal{A}'$ und W ein endlich–dimensionaler K–Vektorraum mit geordneten Basen $\mathcal{B}$ und $\mathcal{B}'$. Sei

   $$\beta : \ V \times W \to K$$

   eine Bilinearform. Dann gilt

   $$M_{\mathcal{A},\mathcal{B}}(\beta) = \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^t \cdot M_{\mathcal{A}'\mathcal{B}'}(\beta) \cdot T_{\mathcal{B}'}^{\mathcal{B}}$$

2. Ist insbesondere $V = W$ und wählt man $\mathcal{A} = \mathcal{B}$ und $\mathcal{A}' = \mathcal{B}'$, so gilt

   $$M_{\mathcal{A}}(\beta) = \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^t \cdot M_{\mathcal{A}'}(\beta) \cdot T_{\mathcal{A}'}^{\mathcal{A}}$$

Beweis.

- Wir rechnen mit $v_j = \sum_{j'} \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)_{j'j} v'_{j'}$ und $w_j = \sum_{j'} \left(T_{\mathcal{B}'}^{\mathcal{B}}\right)_{j'j} w'_{j'}$

$$\begin{aligned} M_{\mathcal{A},\mathcal{B}}(\beta)_{ij} = \beta(v_i, w_j) &= \sum_{i',j'} \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)_{i'i} \beta(v_{i'}, w_{j'}) \left(T_{\mathcal{B}'}^{\mathcal{B}}\right)_{j'j} \\ &= \sum_{i',j'} \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)_{i'i} M_{\mathcal{A}'\mathcal{B}'}(\beta)_{i'j'} \left(T_{\mathcal{B}'}^{\mathcal{B}}\right)_{j'j} \\ &= \left(\left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^t \cdot M_{\mathcal{A}'\mathcal{B}'}(\beta) \cdot T_{\mathcal{B}'}^{\mathcal{B}}\right)_{ij} \end{aligned}$$

- 2. folgt als Spezialfall.

□

Man mache sich sorgfältig den Unterschied zu den Transformationsformeln in Satz 3.6.5 (*) für lineare Abbildungen und (**) für Endomorphismen klar. Diese Transformationsformeln haben uns auf die Äquivalenzrelationen "äquivalent" und "ähnlich" auf Räumen von Matrizen geführt. Hier finden wir eine weitere Relation, die wir nur für quadratische Matrizen einführen, also nur für die Situation $V = W$ und $\mathcal{A} = \mathcal{B}$ und $\mathcal{A}' = \mathcal{B}'$.

Definition 6.2.7
Seien $B, C \in M(n \times n, K)$. Wir sagen, C sei kongruent zu B über K und schreiben

$$C \simeq B \ ,$$

wenn es eine invertible quadratische Matrix $S \in GL(n, K)$ gibt, so dass

$$C = S^t BS$$

gilt.

Bemerkungen 6.2.8.

1. Die Determinanten kongruenter Matrizen unterscheiden sich um ein Quadrat in $K^\times = K \setminus \{0\}$:
$$\det C = \det\left(S^t BS\right) = (\det S)^2 \det B \ .$$
2. Kongruenz ist eine Äquivalenzrelation auf $M(n \times n, K)$.
3. Zwei Matrizen sind genau dann kongruent, wenn sie die gleiche Bilinearform bezüglich verschiedener Basen beschrieben. Dies wird wie in Lemma 3.6.9 gezeigt.
4. Es gelte $\dim_k V = \dim_K W$. Dann ist die Bilinearform $\beta : V \times W \to K$ nach Lemma 6.2.3.3 genau dann nicht-ausgeartet, wenn die lineare Abbildung $\beta_1 : V \to W^*$ injektiv ist. Dies ist genau dann der Fall, wenn β_1 bijektiv ist und somit genau dann, wenn die darstellende Matrix $M_{\mathcal{A},\mathcal{B}}(\beta) = M^{\mathcal{A}}_{\mathcal{B}^*}(\beta_1)^t$ invertibel ist. Dann ist der Determinante der darstellenden Matrix $M_{\mathcal{A},\mathcal{B}}(\beta)$ ungleich Null.

Mit $\beta \in \mathrm{Bil}_K(V, W)$ und $\Phi \in \mathrm{End}(V)$ definiert offenbar auch

$$\beta_\Phi(x, y) := \beta(\Phi x, y)$$

eine Bilinearform $\beta_\Phi \in \mathrm{Bil}_K(V, W)$. Der folgende Satz gibt eine Umkehrung:

Satz 6.2.9.
Seien V, W endlich–dimensionale K–Vektorräume mit $\dim_K V = \dim_K W$. Sei $\beta \in \mathrm{Bil}_K(V, W)$ eine *nicht-ausgeartete* Bilinearform. Zu jeder beliebigen Bilinearform $\gamma \in \mathrm{Bil}_K(V, W)$ gibt es dann genau einen Endomorphismus $\Phi \in \mathrm{End}_K(V)$, so dass $\beta_\Phi = \gamma$ gilt, d.h.

$$\gamma(x, y) = \beta(\Phi x, y) \qquad \text{für alle } x \in V,\ y \in W \ .$$

Entsprechend gibt es auch genau ein $\Psi \in \mathrm{End}(W)$, so dass $\gamma(x, y) = \beta(x, \Psi y)$ für alle x, y gilt .

Eine *nicht ausgeartete* Bilinearform liefert uns also eine Bijektion zwischen Endomorphismen und Bilinearformen.

Beweis.

- Wegen Lemma 6.2.3 haben wir das Problem gelöst, wenn wir ein $\Phi \in \mathrm{End}(V)$ finden, so dass
$$\gamma_1 = (\beta_\Phi)_1$$
in $\mathrm{Hom}(V, W^*)$ gilt. Nun ist
$$\beta_\Phi(v, w) = \beta(\Phi v, w) = \langle \beta_1 \Phi v, w\rangle$$
für alle $v \in V$ und $w \in W$. Also muss für Φ die Gleichung
$$\beta_1 \circ \Phi = \gamma_1 \qquad (*)$$
gelten. Wegen Lemma 6.2.3.3 ist β_1 injektiv, wegen der Annahme $\dim_K V = \dim_K W$ ist β_1 sogar ein Isomorphismus. von V auf W^*. Also hat die Gleichung $(*)$ genau die Lösung
$$\Phi = (\beta_1)^{-1} \circ \gamma_1 \in \mathrm{End}(V) \ .$$

- Die Aussage für das zweite Argument folgt analog.

□

Korollar 6.2.10.
Es seien V, W endlich–dimensionale K–Vektorräume mit $\dim_K V = \dim_K W$ und

$$\beta : \ V \times W \to K$$

sei eine *nicht-ausgeartete* Bilinearform.

1. Zu $\Phi \in \mathrm{End}_K(V)$ gibt es genau ein $\Phi^\wedge \in \mathrm{End}_K(W)$ mit

   $$\beta(\Phi x, y) = \beta(x, \Phi^\wedge y) \qquad \text{für alle } x \in V,\ y \in W$$

   $\Phi^\wedge$ heißt Rechtsadjungierte von Φ bezüglich β.

2. Zu $\Psi \in \mathrm{End}_K(W)$ gibt es genau ein ${}^\wedge\Psi \in \mathrm{End}_K(V)$ mit

   $$\beta(x, \Psi y) = \beta({}^\wedge\Psi x, y) \qquad \text{für alle } x \in V,\ y \in W$$

   ${}^\wedge\Psi$ heißt die Linksadjungierte von Ψ bezüglich β.

Man beachte, dass die adjungierten Abbildungen $\Phi^\wedge$ und ${}^\wedge\Psi$ von der nicht-ausgearteten Bilinearform β abhängen, auch wenn wir dies in der Notation unterdrücken.

Beweis.
Wir zeigen nur 1.: wende die beiden Aussagen von Satz 6.2.9 auf die Bilinearform $\beta_\Phi(\cdot,\cdot) = \beta(\Phi\cdot,\cdot)$ an, um $\Psi : W \to W$ zu finden, so dass $\beta_\Phi(\cdot,\cdot) = \beta(\cdot, \Psi\cdot)$ gilt. □

Bemerkung 6.2.11.
Es gilt:

$$\begin{array}{rclcrcl} {}^\wedge(\Phi^\wedge) & = & \Phi & & (^\wedge\Psi)^\wedge & = & \Psi \\ (\Phi_1 \circ \Phi_2)^\wedge & = & \Phi_2^\wedge \circ \Phi_1^\wedge & & {}^\wedge(\Psi_1 \circ \Psi_2) & = & {}^\wedge\Psi_2 \circ^\wedge \Psi_1 \end{array}$$

Hierzu beachte man, dass für alle $v \in V$ und $w \in W$ gilt

$$\beta(v, \Psi w) = \beta({}^\wedge\Psi v, w) = \beta(v, ({}^\wedge\Psi)^\wedge w)$$

sowie

$$\beta({}^\wedge(\Psi_1 \circ \Psi_2)v, w) = \beta(v, \Psi_1 \circ \Psi_2 w) = \beta({}^\wedge\Psi_1 v, \Psi_2 w) = \beta({}^\wedge\Psi_2 \circ^\wedge \Psi_1 v, w)\ .$$

Lemma 6.2.12.
In der Situation von Korollar 6.2.10 gilt für jede geordnete Basis $\mathcal{A}$ von V und geordnete Basis $\mathcal{B}$ von W für die darstellenden Matrizen

$$M_{\mathcal{B}}\,(\Phi^\wedge) = M_{\mathcal{AB}}(\beta)^{-1} M_{\mathcal{A}}(\Phi)^t M_{\mathcal{AB}}(\beta)$$

Beweis.
Wir rechnen für alle Werte von i, j

$$\beta(\Phi v_i, w_j) = \sum_l M_{\mathcal{A}}(\Phi)_{li}\, M_{\mathcal{AB}}(\beta)_{lj} = \left(M_{\mathcal{A}}(\Phi)^t \cdot M_{\mathcal{AB}}(\beta)\right)_{ij}$$

$$\beta(v_i, \Phi^\wedge w_j) = \sum_k M_{\mathcal{B}}\,(\Phi^\wedge)_{kj}\, M_{\mathcal{AB}}(\beta)_{ik} = \left(M_{\mathcal{AB}}(\beta) \cdot M_{\mathcal{B}}(\Phi^\wedge)\right)_{ij}$$

Aus der Gleichheit der beiden Zeilen folgt die Behauptung. □

Definition 6.2.13
Es sei $\beta : V \times V \to K$ eine Bilinearform auf V. Wir sagen, ein Endomorphismus $\Phi \in \mathrm{End}_K(V)$ lasse β invariant oder sei eine Isometrie von V bezüglich β, wenn für alle $x, y \in V$ gilt

$$\beta(\Phi x, \Phi y) = \beta(x, y) \quad .$$

Satz 6.2.14.
Sei V ein n–dimensionaler K–Vektorraum und sei $\beta \in \mathrm{Bil}_K(V, V)$ eine *nicht-ausgeartete* Bilinearform. Dann sind für $\Phi \in \mathrm{End}(V)$ äquivalent:

1. Φ ist eine Isometrie bezüglich β.
2. Es gilt $\Phi^\wedge \circ \Phi = \mathrm{id}_V$.
3. Φ ist invertierbar und es gilt
$$\Phi^{-1} = \Phi^\wedge \ = \ {}^\wedge\Phi \ .$$
4. Sei $\mathcal{B}$ eine geordnete Basis von V. Sei $S := M_{\mathcal{B}}(\Phi)$ die darstellende Matrix des Endomorphismus Φ und $B := M_{\mathcal{BB}}(\beta)$ die darstellende Matrix der nicht-ausgearteten Bilinearform. Dann gilt
$$B = S^t B S \ .$$

Beweis.

1.⇔2. Aus der Definition der Rechtsadjungierten $\Phi^\wedge$ bezüglich β in Korollar 6.2.10 folgt, dass für alle $x, y \in V$ gilt
$$\beta\left(x, \Phi^\wedge \Phi y\right) = \beta(\Phi x, \Phi y) \ . \qquad (*)$$
Ist Φ eine Isometrie, so folgt $\beta\left(x, \Phi^\wedge \Phi y\right) = \beta(x, y)$. Weil β nicht-ausgeartet ist, folgt daraus $\Phi^\wedge \circ \Phi = \mathrm{id}_V$. Gilt umgekehrt 2. so folgt aus $(*)$, dass Φ eine Isometrie ist, also $\beta(x, y) = \beta(\Phi(x), \Phi(y))$.

2.⇒3. Wegen $\dim V < \infty$ folgt aus $\Phi^\wedge \circ \Phi = \mathrm{id}_V$, dass Φ invertierbar mit Inverser $\Phi^\wedge$ ist. Die Aussage für ${}^\wedge\Phi$ folgt, indem wir die Linksadjungierte der Gleichung $\Phi^\wedge \circ \Phi = \mathrm{id}_V$ nehmen, also
$$\mathrm{id}_V = {}^\wedge \mathrm{id}_V = {}^\wedge \Phi \circ {}^\wedge \left(\Phi^\wedge\right) \overset{6.2.11}{=} {}^\wedge\Phi \circ \Phi \ .$$

3.⇒ 4. Wir überlegen uns zunächst, dass wegen Lemma 6.2.12 die darstellende Matrix von $\Phi^\wedge$ die Matrix $B^{-1} S^t B$ ist. Drückt man 3. durch die darstellenden Matrizen aus, erhält man $B^{-1} S^t B = S^{-1}$, was zu $S^t B S = B$ äquivalent ist.

4.⇒2. Wir rechnen wieder mit der darstellenden Matrix $B^{-1} S^t B = S^{-1}$ von $\Phi^\wedge$:
$$\left(B^{-1} S^t B\right) S = B^{-1} \left(S^t B S\right) \overset{4.}{=} B^{-1} \cdot B = E_n \ .$$
was 2. in darstellenden Matrizen ist.

□

Definition 6.2.15

1. *Es sei $\beta \in \mathrm{Bil}_K(V,W)$. Zwei Vektoren $x \in V$ und $y \in W$ heißen orthogonal bezüglich β, wenn $\beta(x,y) = 0$ gilt.*

2. *Sind $X \subset V$ bzw. $Y \subset W$ Teilmengen, so bezeichnen wir mit*

$$X^\perp := \{y \in W |\, \beta(x,y) = 0 \text{ für alle } x \in X\} \subset W$$
$$^\perp Y := \{x \in V |\, \beta(x,y) = 0 \text{ für alle } y \in Y\} \subset V$$

der orthogonale Raum bezüglich β.

Bemerkungen 6.2.16.

1. Offenbar ist
$$X^\perp = (\mathrm{span}_K X)^\perp \quad \text{und} \quad {}^\perp Y = {}^\perp (\mathrm{span}_K Y) \ ;$$
alle orthogonalen Räume sind Untervektorräume.

2. Seien V, W endlich–dimensionale K–Vektorräume und $\beta \in \mathrm{Bil}_K(V,W)$. Dann gilt für alle Untervektorräume $X \subset V$ bzw. $Y \subset W$:
$$\dim_K X + \dim_K X^\perp = \dim_K W + \dim_K \left({}^\perp W \cap X\right)$$
$$\dim_K Y + \dim_K {}^\perp Y = \dim_K V + \dim_K \left(V^\perp \cap\ Y\right)$$

Beweis.

- Die Dimensionsformel 3.1.7 für die lineare Abbildung
$$(\beta_1)_{|X} :\ X \to W^*$$
liefert zunächst die Beziehung $\dim_K \mathrm{Im}\,(\beta_1)_{|X} = \dim_K X - \dim_K \ker(\beta_1)_{|X}$.
- Es ist $\ker \beta_1 = \{v \in V |\, \beta(v,w) = 0 \ \text{ für alle } w \in W\} = {}^\perp W$, also
$$\ker(\beta_1)_{|X} = \ker(\beta_1) \cap X = {}^\perp W \cap X \ .$$
- Ferner ist nach Identifikation $\iota_w : W \xrightarrow{\simeq} W^{**}$, vgl. Satz 6.1.17:
$$\left(\mathrm{Im}\,(\beta_1)_{|X}\right)^0 = \{w \in W |\, \beta(x,w) = 0 \ \text{ für alle } x \in X\} = X^\perp$$
Zusammen mit Lemma 6.1.8.3 folgt
$$\begin{aligned} \dim_K X^\perp &= \dim_K (\mathrm{Im}\,\beta_{1|X})^\circ \overset{6.1.8.3}{=} \dim_K W - \dim_K \mathrm{Im}\,\beta_{1|X} \\ &= \dim_K W - \dim_K X + \dim_K({}^\perp W \cap X) \ . \end{aligned}$$

□

3. Speziell für $X = V$ folgt

$$\dim_K V + \dim_K V^\perp = \dim_K W + \dim_K({}^\perp W)$$

Ist $\dim_K V = \dim_K W < \infty$, so folgt $\dim_K(V^\perp) = \dim_K({}^\perp W)$. Für $\dim_K V = \dim_K W$ ist eine Bilinearform also genau dann nicht-ausgeartet im ersten Argument, wenn sie im zweiten Argument nicht-ausgeartet ist.

4. Gilt $\dim_K V = \dim_K W < \infty$ und ist β nicht-ausgeartet, so gilt ${}^\perp W = 0 = V^\perp$ und somit für alle Untervektorräume $X \subset V$ und $Y \subset W$:

$$\begin{aligned} \dim_K X + \dim_K X^\perp &= \dim_K V \\ \dim_K Y + \dim_K {}^\perp Y &= \dim_K W \ , \end{aligned}$$

sowie ${}^\perp(X^\perp) = X$ und $\left({}^\perp Y\right)^\perp = Y$.

Satz 6.2.17.
Sei V ein endlich-dimensionaler K–Vektorraum und $\beta \in \mathrm{Bil}_K(V,V)$ nicht-ausgeartet. Es gelte $\beta(x,y) = 0$ genau dann, wenn auch $\beta(y,x) = 0$ gilt. Dann ist β entweder symmetrisch, d.h. es gilt

$$\beta(x,y) = \beta(y,x) \ \text{ für alle } x, y \in V$$

oder schiefsymmetrisch, d.h. es gilt

$$\beta(x,y) = -\beta(y,x) \ \text{ für alle } x, y \in V \ .$$

Beweis.
Da β nicht ausgeartet ist, folgt aus Satz 6.2.9, dass es einen Endomorphismus $f : V \to V$ gibt, so dass für alle $x, y \in V$ gilt

$$\beta(y,x) = \beta(f(x), y) \, .$$

Betrachte nun ein beliebiges $x_0 \neq 0$ aus V und den zugehörigen eindimensionalen Vektorraum $M = \mathrm{span}_K(x_0)$ mit der Menge der dazu orthogonalen Vektoren $M^\perp = \{y \in V | \beta(x_0, y) = 0\}$.

Wegen der vorausgesetzten Verschwindenseigenschaft von β folgt aus $\beta(x_0, y) = 0$ für alle $y \in M^\perp$

$$0 = \beta(y, x_0) = \beta(f(x_0), y) \, ,$$

also, da β nicht entartet ist, dass $f(x_0) \in^\perp (M^\perp) = M$. Daher muss x_0 ein Eigenvektor von f sein. Da aber $x_0 \in V$ beliebig war, kann f nur ein Vielfaches der Identität sein, $f = \lambda \mathrm{id}_V$. Aus $\beta(x,y) = \beta(f(y), x) = \beta(f(x), f(y)) = \lambda^2 \beta(x,y)$ für alle $x, y \in V$ folgt schließlich $\lambda = \pm 1$. □

Definition 6.2.18

1. *Eine Bilinearform $\beta \in \mathrm{Bil}_K(V,V)$, für die*

$$\beta(x,x) = 0 \ \textit{für alle } x \in V$$

gilt, heißt alternierende Bilinearform.

2. *Eine nicht-ausgeartete alternierende Bilinearform heißt symplektische Bilinearform. Ein Vektorraum mit einer symplektischen Bilinearform heißt ein symplektischer Vektorraum.*

Bemerkungen 6.2.19.

1. Für jede alternierende Bilinearform gilt

$$0 = \beta(x+y, x+y) = \beta(x,x) + \beta(x,y) + \beta(y,x) + \beta(y,y) = \beta(x,y) + \beta(y,x) \; ;$$

sie ist also insbesondere schiefsymmetrisch. Für schiefsymmetrische Bilinearformen gilt nur $\beta(x,x) = -\beta(x,x)$, also $2 \cdot \beta(x,x) = 0$. Gilt im Körper K, dass $1+1 \neq 0$ ist, so ist jede schiefsymmetrische K-Bilinearform auch alternierend.

2. Eine Bilinearform ist schiefsymmetrisch, falls für jede Basis $\mathcal{B}$ für $B = M_{\mathcal{B}}(\beta)$ gilt $B^t = -B$. Eine solche Matrix heißt schiefsymmetrisch. Eine Matrix mit $B^t = B$ heißt symmetrisch. Aus $B^t = -B$ folgt für $B \in M(n \times n, K)$

$$\det B = \det B^t = \det(-B) = (-1)^n \det B \; .$$

Gilt im Körper K, dass $1+1 \neq 0$ ist, so muss n gerade oder $\det B = 0$ gelten. Ist die Form nicht ausgeartet, so ist $\det B \neq 0$. Über Körpern mit $\mathrm{char}(K) \neq 2$ haben symplektische Vektorräume also gerade Dimension.

Satz 6.2.20.
Sei β eine symplektische Bilinearform auf einem K–Vektorraum V; im Körper K gelte $1+1 \neq 0$. Dann besitzt V eine symplektische Basis, d.h. eine geordnete Basis

$$\mathcal{B} = (u_1, v_1, \ldots, u_m, v_m) \; ,$$

in der die darstellende Matrix die Block-diagonale Form

$$M_{\mathcal{B}}(\beta) = \begin{pmatrix} H & & & 0 \\ & H & & \\ & & \ddots & \\ 0 & & & H \end{pmatrix}$$

mit

$$H := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M(2 \times 2, K)$$

hat. Insbesondere ist $\dim_K V$ gerade.
Die Determinante jeder schiefsymmetrischen Matrix $B \in M(n \times n, K)$ ist ein Quadrat im Körper K.

Beweis.

- Sei $u_1 \neq 0$ beliebig, $u_1 \in V$. Da β nicht-ausgeartet ist, finde einen Vektor $v_1 \in V$ mit $\beta(u_1, v_1) = 1$ und somit auch $\beta(v_1, u_1) = -1$. Die Familie (u_1, v_1) ist linear unabhängig. Denn gilt $\lambda u_1 + \mu v_1 = 0$, so folgt

$$0 = \beta(\lambda u_1 + \mu v_1, v_1) = \lambda \quad \text{und} \quad 0 = \beta(\lambda u_1 + \mu v_1, u_1) = -\mu \; .$$

Auf dem zwei-dimensionalen Untervektorraum $\mathrm{span}_K(u_1, v_1) =: U_1$ wird die Einschränkung von β in der Basis (u_1, v_1) durch die Matrix H dargestellt.

- Betrachte das orthogonale Komplement $U_1^\perp$. Es ist $U_1 \cap U_1^\perp = \{0\}$, denn sei $v \in U_1 \cap U_1^\perp$. Wegen $v \in U_1$ schreiben wir $v = \lambda u_1 + \mu v_1$. Es folgt dann aus $v \in U_1^\perp$ mit ähnlicher Rechnung wie eben

$$0 = \beta(v, v_1) = \lambda \quad \text{und} \quad 0 = \beta(v, u_1) = -\mu \ .$$

  Da β nicht ausgeartet ist, folgt aus Bemerkung 6.2.16.4, dass $\dim_K V = \dim_K U_1 + \dim_K U_1^\perp$, und somit die Zerlegung von V als direkte Summe,

$$V = U_1 \oplus U_1^\perp \ .$$

  Die Einschränkung $\beta|_{U_1^\perp}$ auf den Untervektorraum $U_1^\perp$ ist alternierend. Sie ist auch nicht ausgeartet: angenommen, es gäbe $v_0 \in U_1^\perp, v_0 \neq 0$ mit $\beta(v_0, v'') = 0$ für alle $v'' \in U_1^\perp$. Schreibe dann ein beliebiges $v \in V$ wegen der direkten Summerzerlegung von V in der Form $v = v' + v''$ mit $v' \in U_1$ und $v'' \in U_1^\perp$. Dann gilt

$$\beta(v_0, v) = \beta(v_0, v') + \beta(v_0, v'') = 0 + 0 = 0$$

  im Widerspruch zu der Annahme, dass β auf ganz V nicht ausgeartet ist. Also trägt der Untervektorraum $\beta|_{U_1^\perp}$ eine symplektische Form, und wir können vollständige Induktion nach der Dimension von V anwenden.

- Für die letzte Aussage beachten, wir, dass, wenn β ausgeartet ist, $\det B = 0$ ohnehin ein Quadrat ist. Ist β ausgeartet, so gilt in der symplektischen Basis $\det M_{\mathcal{B}}(\beta) = 1$. Aus der Transformationsformel 6.2.8.1. folgt für die darstellende Matrix in einer beliebigen Basis $\det(B) = \det(S)^2 \det M_{\mathcal{B}}(\beta) = \det(S)^2$.

□

6.3 Tensorprodukte

Bevor wir symmetrische Bilinearformen studieren, wollen wir eine etwas abstraktere Begriffsbildung einführen. Dazu überlegen wir uns als Vorbereitung:

Definition 6.3.1
Sei K ein Körper und seien V, W und X gegebene K-Vektorräume. Dann ist eine K-bilineare Abbildung eine Abbildung

$$\alpha : V \times W \to X$$

die in beiden Argumenten K-linear ist, also

$$\alpha(\lambda v + \lambda' v', w) = \lambda\alpha(v, w) + \lambda'\alpha(v', w) \quad \text{und} \quad \alpha(v, \lambda w + \lambda' w') = \lambda\alpha(v, w) + \lambda'\alpha(v, w')$$

für alle $\lambda, \lambda' \in K$ und $v, v' \in V$, $w, w' \in W$.

Sei $\mathcal{B}$ eine Basis von V und $\mathcal{B}'$ eine Basis von W. Eine bilineare Abbildung α ist durch ihre Werte $\alpha(b, b')$ mit $b \in \mathcal{B}$ und $b' \in \mathcal{B}'$ festgelegt, vgl. Betrachtung 6.2.4 für Bilinearformen.

Offenbar ist dann für jede lineare Abbildung $\Phi : X \to X'$ auch die Abbildung $\Phi \circ \alpha : V \times W \to X'$ bilinear. Wir stellen uns die Frage, ob es für je zwei K-Vektorräume V, W einen "universellen" K-Vektorraum U mit einer "universellen" bilinearen Abbildung $\kappa : V \times W \to U$

gibt, so dass eine beliebige bilineare Abbildung $\beta : V \times W \to X$ dann so durch eine lineare Abbildungen $\Phi : U \to X$ in der Form $\beta = \Phi \circ \kappa$ beschrieben werden kann.

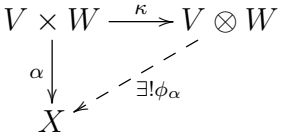
Definition 6.3.2
Das Tensorprodukt zweier K-Vektorräume V, W ist ein Paar, bestehend aus einem K-Vektorraum $V \otimes W$ und einer bilinearen Abbildung

$$\begin{array}{rccc} \kappa : & V \times W & \to & V \otimes W \\ & (v, w) & \mapsto & v \otimes w \end{array}$$

mit der folgenden universellen Eigenschaft: zu jeder bilinearen *Abbildung*

$$\alpha : V \times W \to X$$

gibt es genau eine lineare *Abbildung $\phi_\alpha : V \otimes W \to X$ mit $\alpha = \phi_\alpha \circ \kappa$. Als Diagramm:*



Betrachtung 6.3.3.

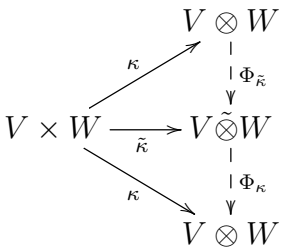
1. Damit ist die Theorie bilinearer Abbildungen mit der Theorie linearer Abbildungen verbunden.

2. Wir zeigen zunächst, dass das Tensorprodukt, wenn es denn existiert, bis auf eindeutige Isomorphie eindeutig ist. Angenommen, wir hätten zwei solche universelle Abbildungen

   $$\kappa : V \times W \to V \otimes W \qquad \tilde{\kappa} : V \times W \to V \tilde{\otimes} W \ .$$

   Man benutzt die universelle Eingeschaft von κ und findet für die spezielle bilineare Abbildung $\tilde{\kappa}$ eine eindeutige lineare Abbildung $\Phi_{\tilde{\kappa}} : V \otimes W \to V \tilde{\otimes} W$ mit $\Phi_{\tilde{\kappa}} \circ \kappa = \tilde{\kappa}$.

   Durch Vertauschen der Rollen erhält man ebenso eine lineare Abbildung $\Phi_\kappa : V \tilde{\otimes} W \to V \otimes W$ mit $\Phi_\kappa \circ \tilde{\kappa} = \kappa$. Die Abbildungen $\kappa = \mathrm{id}_{V \otimes W} \circ \kappa$ und $\Phi_\kappa \circ \Phi_{\tilde{\kappa}} \circ \kappa$ beschreiben die gleiche bilineare Abbildung $V \times W \to V \otimes W$. Wegen der Eindeutigkeitsaussage in der universellen Eigenschaft folgt $\Phi_\kappa \circ \Phi_{\tilde{\kappa}} = \mathrm{id}_{V \otimes W}$. Als Diagramm:

   Analog folgt $\Phi_{\tilde{\kappa}} \circ \Phi_\kappa = \mathrm{id}_{V \tilde{\otimes} W}$. (Man schaue sich jetzt noch einmal den Beweis von Satz 3.5.7 und Satz 3.5.13 an.)

3. Um die Existenz des Tensorprodukts zu zeigen, wählen wir eine Basis $\mathcal{B} := \{b_1, b_2, \ldots\}$ von V und $\mathcal{B}' := \{b'_1, b'_2, \ldots\}$ von W. Da eine bilineare Abbildung durch ihre Werte auf allen Paaren (b_i, b'_j) eindeutig festgelegt ist, brauchen wir uns nur einen Vektorraum zu

verschaffen, für den eine Basis durch diese Paare indiziert wird. Sei also $V \otimes W$ der Vektorraum der K-wertigen Funktionen auf der Menge $\mathcal{B} \times \mathcal{B}'$, die nur für endlich viele Elemente einen Wert ungleich Null annehmen. Wir bezeichnen mit $b_i \otimes b'_j$ die Funktion, die auf dem Paar (b_i, b'_j) den Wert Eins und sonst den Wert Null hat. Die bilineare Abbildung κ ist dann auf dem Paar (b_i, b'_j) vorgeschrieben durch $\kappa(b_i, b'_j) = b_i \otimes b'_j$. Sie erfüllt die universelle Eigenschaft, denn der bilinearen Abbildung $\alpha : V \times W \to X$ wird die eindeutig bestimmte lineare Abbildung $\Phi_\alpha : V \otimes W \to X$ mit $\Phi_\alpha(b_i \otimes b'_j) = \alpha(b_i, b'_j)$ zugeordnet.

4. Insbesondere ist für endlich-erzeugte Vektorräume V, W die Dimension des Tensorprodukts gleich $\dim_K V \otimes W = \dim_K V \cdot \dim_K W$.

5. Die Elemente des Vektorraums $V \otimes W$ heißen Tensoren. Wir schreiben auch $v \otimes w$ für $\kappa(v, w)$. Dann folgt aus der Bilinearität von κ sofort

   $$(\lambda_1 v_1 + \lambda_2 v_2) \otimes w = \lambda_1 v_1 \otimes w + \lambda_2 v_2 \otimes w \text{ und } v \otimes (\lambda_1 w_1 + \lambda_2 w_2) = v \otimes \lambda_1 w_1 + v \otimes \lambda_2 w_2 \ .$$

   Die Elemente der Form $v \otimes w$ mit $v \in V$ und $w \in W$ heißen Tensorprodukte. Die Tensorprodukte erzeugen $V \otimes W$, aber nicht jedes Element von $V \otimes W$ ist das Tensorprodukt eines Vektors $v \in V$ und $w \in W$. Tatsächlich bildet das Bild $\kappa(V \times W)$ in $V \otimes W$ einen $\dim_K V + \dim_K W$-dimensionalen Kegel in einem $\dim_K V \cdot \dim_K W$-dimensionalen Vektorraum.

Beispiele 6.3.4. .

1. Den Polynomring in mehreren Variablen definiert man induktiv. Insbesondere ist $K[t_1, t_2] := (K[t_1])[t_2]$ definiert als der Polynomring für den kommutativen Ring $K[t_1]$. Eine Basis sind die Monome $t_1^{n_1} t_2^{n_2}$ mit $n_1, n_2 \in \mathbb{N}_0$.

   Wir betrachten die Multiplikation von Polynomen

   $$\begin{array}{rccc} \xi : & K[t] \times K[t] & \to & K[t_1, t_2] \\ & (P(t), Q(t)) & \mapsto & P(t_1) \cdot Q(t_2) \end{array}$$

   Sie ist bilinear; nach der universellen Eigenschaft des Tensorprodukts induziert sie eine lineare Abbildung

   $$\xi_\otimes : \quad K[t] \otimes K[t] \to K[t_1, t_2]$$

   mit $t^i \otimes t^j \mapsto t_1^i \cdot t_2^j$. Da die Monome eine Basis der Polynomringe bilden, ist $\xi_\otimes$ bijektiv und wir haben den Polynomring in zwei Variablen als Tensorprodukt geschrieben.

2. Sei W ein beliebiger $\mathbb{R}$-Vektorraum. Betrachte $\mathbb{C}$ als zwei-dimensionalen $\mathbb{R}$-Vektorraum. Das Tensorprodukt $\mathbb{C} \otimes_{\mathbb{R}} W$ ist ein reeller Vektorraum, der durch die bilineare Abbildung

   $$\begin{array}{rcl} \mathbb{C} \times (\mathbb{C} \otimes_{\mathbb{R}} W) & \to & \mathbb{C} \otimes_{\mathbb{R}} W \\ (\lambda, \sum_j \lambda_j \otimes w_j) & \mapsto & \sum_j \lambda \cdot \lambda_j \otimes w_j \end{array}$$

   mit einer skalaren Multiplikation mit komplexen Zahlen so ausgestattet wird, dass er ein $\mathbb{C}$-Vektorraum wird. So kann man aus $\mathbb{R}$-Vektorräumen $\mathbb{C}$-Vektorräumen machen. Da dies mit algebraischen Zusatzstrukturen verträglich ist, kann man so auch zum Beispiel aus $\mathbb{R}$-Algebren $\mathbb{C}$-Algebren machen. Zum Beispiel gilt für den Polynomring $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}[X] \cong \mathbb{C}[X]$.

Betrachtung 6.3.5.

- Für zwei K-lineare Abbildungen

$$\Phi : V \to V' \qquad \text{und} \qquad \Psi : W \to W'$$

betrachten wir das Diagramm:

$$\begin{array}{ccc} V \times W & \xrightarrow{\kappa} & V \otimes W \\ {\scriptstyle \Phi\times\Psi}\downarrow & & \downarrow{\scriptstyle \exists!\Phi\otimes\Psi} \\ V' \times W' & \xrightarrow[\kappa']{} & V' \otimes W' \end{array}$$

Da die Abbildung $\kappa' \circ (\Phi \times \Psi)$ offenbar bilinear ist, existiert nach der universellen Eigenschaft der Tensorprodukts $V \otimes W$ eine lineare Abbildung

$$\Phi \otimes \Psi : \quad V \otimes W \to V' \otimes W' \ .$$

Diese erfüllt also

$$(\Phi \otimes \Psi)(v \otimes w) = \Phi(v) \otimes \Psi(w) \quad \text{ für alle } v \in V \text{ und } w \in W \ .$$

- Seien $\mathcal{B}^{(V)}, \mathcal{B}^{(V')}, \mathcal{B}^{(W)}, \mathcal{B}^{(W')}$ geordnete Basen. Dann gilt

$$(\Phi \otimes \Psi)(b_i^{(V)} \otimes b_j^{(W)}) = \Phi(b_i^{(V)}) \otimes \Psi(b_j^{(W)}) = \sum_{p,q} M_{\mathcal{B}^{(V')}}^{\mathcal{B}^{(V)}}(\Phi)_{pi} b_p^{(V')} \otimes M_{\mathcal{B}^{(W')}}^{\mathcal{B}^{(W)}}(\Psi)_{qj} b_q^{(W')}$$

Daher ist die darstellende Matrix von $\Phi \otimes \Psi$ bezüglich der Basen $b_i^{(V)} \otimes b_j^{(W)}$ von $V \otimes W$ und $b_p^{(V')} \otimes b_q^{(W')}$ von $V \otimes W$ gleich

$$M(\Phi \otimes \Psi)_{(p,q),(i,j)} = M(\Phi)_{p,i} \cdot M(\Psi)_{q,j} \ .$$

Bemerkungen 6.3.6.

1. Aus der Bilinearität von κ folgt, dass auch das Tensorprodukt von Abbildungen bilinear ist:

$$\begin{array}{rcl} (\lambda_1\Phi_1 + \lambda_2\Phi_2) \otimes \Psi & = & \lambda_1\Phi_1 \otimes \Psi + \lambda_2\Phi_2 \otimes \Psi \\ \Phi \otimes (\lambda_1\Psi_1 + \lambda_2\Psi_2) & = & \Phi \otimes \lambda_1\Psi_1 + \Phi \otimes \lambda_2\Psi_2 \end{array}$$

2. Ebenso folgt für Vektorräume die Verträglichkeit mit direkten Summen:

$$(V_1 \oplus V_2) \otimes W \cong (V_1 \otimes W) \oplus (V_2 \otimes W) \, ,$$

und analog im anderen Argument.

3. Man hat kanonische Isomorphismen

$$\begin{array}{rcccl} a_{U,V,W} & : & U \otimes (V \otimes W) & \to & (U \otimes V) \otimes W \\ & & u \otimes (v \otimes w) & \mapsto & (u \otimes v) \otimes w \end{array}$$

mit deren Hilfe man die K-Vektorräume $U \otimes (V \otimes W)$ und $(U \otimes V) \otimes W$ identifizieren kann. Das Tensorprodukt ist dann assoziativ.

4. Man hat kanonische Isomorphismen

$$\begin{array}{rcl} K \otimes V & \to & V \\ \lambda \otimes v & \mapsto & \lambda \cdot v \end{array}$$

mit Umkehrabbildung $v \mapsto 1 \otimes v$, mit deren Hilfe man den Grundkörper K als Eins unter dem Tensorprodukt auffassen kann.

5. Man hat kanonische Isomorphismen

$$\begin{array}{rccc} c_{U,V} : & U \otimes V & \to & V \otimes U \\ & u \otimes v & \mapsto & v \otimes u \ , \end{array}$$

mit deren Hilfe man die Faktoren vertauschen kann. Es gilt $c_{V,U} \circ c_{U,V} = \mathrm{id}_{U \otimes V}$.

6. Für endlich-dimensionale Vektorräume ist die kanonische Abbildung

$$\begin{array}{lcl} V^* \otimes W^* & \to & (V \otimes W)^* \\ \alpha \otimes \beta & \mapsto & (v \otimes w \mapsto \alpha(v) \cdot \beta(w)) \end{array}$$

ein Isomorphismus. Wir können insbesondere eine Bilinearform $\beta : V \times W \to K$ mit einem Element in $V^* \otimes W^*$ identifizieren.

7. Für endlich-dimensionale Vektorräume ist die kanonische Abbildung

$$\begin{array}{lcl} V^* \otimes W & \to & \mathrm{Hom}_K(V, W) \\ \alpha \otimes w & \mapsto & (v \mapsto \alpha(v)w) \end{array}$$

ein Isomorphismus.

6.4 Quadratische Formen

Definition 6.4.1

1. *Sei V ein K–Vektorraum. Eine quadratische Form auf V ist eine Abbildung*

$$q : V \to K$$

mit den beiden Eigenschaften:

(QF1) Für alle $\lambda \in K$ und $v \in V$ gilt $q(\lambda v) = \lambda^2 q(v)$.

(QF2) Die Abbildung

$$\begin{array}{l} \beta_q : V \times V \to K \\ \qquad (v, w) \mapsto q(v + w) - q(v) - q(w) \end{array}$$

ist eine Bilinearform auf V.

2. *Eine quadratische Form q heißt nicht-ausgeartet, wenn die zugehörige Bilinearform β_q nicht-ausgeartet ist.*

Bemerkungen 6.4.2.

1. Sei $V = K$; dann ist $q : K \to K$ mit $x \mapsto cx^2$ mit $c \in K$ eine quadratische Form. Denn es gilt

$$q(\lambda x) = c(\lambda x)^2 = \lambda^2 c x^2 = \lambda^2 q(x) \ ,$$

und

$$\beta_q(x, y) = c(x + y)^2 - cx^2 - cy^2 = 2cxy$$

ist eine Bilinearform auf K. Die Form ist genau dann nicht-ausgeartet, wenn $2c \in K \setminus \{0\}$ gilt.

2. Allgemeiner sei $V = K^n$ und $C = (c_{ij}) \in M(n \times n, K)$. Dann ist

$$q(x) = x^t C x = \sum_{i,j} x_i c_{ij} x_j$$

eine quadratische Form auf K^n.

3. Offensichtlich ist die zu einer quadratischen Form gehörige Bilinearform symmetrisch: $\beta_q(x, y) = \beta_q(y, x)$.

Satz 6.4.3.
Sei K ein Körper, in dem $1+1 \neq 0$ gilt. Dann entsprechen sich quadratische Formen über einem K–Vektorraum V und symmetrische Bilinearformen $\beta \in \mathrm{SBil}_K(V, V)$ umkehrbar eindeutig:

$$\begin{aligned} \mathrm{QF}(V) &\to \mathrm{SBil}_K(V, V) \\ q &\mapsto \beta_q \\ \mathrm{SBil}_K(V, V) &\to \mathrm{QF}(V) \\ \beta &\mapsto q_\beta \end{aligned}$$

mit $q_\beta(x) = 2^{-1}\beta(x, x)$.

Beweis.

1. Wir untersuchen für eine gegebene symmetrische Bilinearform β die Abbildung $q_\beta : V \to K$ mit $q_\beta(x) = 2^{-1}\beta(x, x)$ und deren zugehörige Bilinearform:

$$\begin{aligned} & q_\beta(x + y) - q_\beta(x) - q_\beta(y) \\ &= \frac{1}{2}\left(\beta(x + y, x + y) - \beta(x, x) - \beta(y, y)\right) \\ &= \frac{1}{2}\left(\beta(x, y) + \beta(y, x)\right) = \beta(x, y) \ . \end{aligned}$$

was tatsächlich eine Bilinearform ist, so dass (QF2) für q_β erfüllt ist. Ferner gilt

$$q_\beta(\lambda x) = \frac{1}{2}\beta(\lambda x, \lambda x) = \lambda^2 q_\beta(x) \ .$$

Also gilt auch (QF1) und q_β ist eine quadratische Form, zu der die Bilinearform β gehört. Die Gleichung

$$\beta(x, y) = q_\beta(x + y) - q_\beta(x) - q_\beta(y)$$

heißt Polarisierungsformel für die Bilinearform β.

2. Sei umgekehrt eine quadratische Form q gegeben, zu der die Bilinearform β_q gehört. Dieser wird nach 1. die folgende quadratische Form zugeordnet:

$$\frac{1}{2}\beta_q(x, x) = \frac{1}{2}\left(q(x + x) - q(x) - q(x)\right) = q(x) \ .$$

□

Bemerkung 6.4.4.
Die Situation ist für einen Körper K, in dem $1+1=0$ gilt, anders. Sei $K=\mathbb{F}_2$ und $q: K \to K$ eine quadratische Form. Dann gilt wegen (QF1) $q(0) = q(0\cdot 0) = 0^2 q(0) = 0$, wohingegen (QF1) den Wert von $q(1)$ nicht einschränkt, also $q(1) \in \{0,1\}$. Für die zugehörige symmetrische Bilinearform β gilt in beiden Fällen

$$\beta(0,0) = \beta(0,1) = \beta(1,0) = 0 \quad \text{und} \quad \beta(1,1) = q(0) - q(1) - q(1) = -2q(1) = 0 \ .$$

Hier führen also zwei verschiedene quadratische Formen auf die gleiche symmetrische Bilinearform. Die quadratische Form enthält also mehr Information als die symmetrische Bilinearform. Umgekehrt gibt es zur symmetrischen Bilinearform mit $\beta(1,1)=1$ keine quadratische Form.

Satz 6.4.5. Normalform für quadratische Formen
Sei K ein Körper, in dem $1+1 \neq 0$ gilt. Sei V ein endlich–dimensionaler K–Vektorraum und β eine symmetrische Bilinearform auf K. Dann existiert eine geordnete Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von V, in der die darstellende Matrix $M_{\mathcal{B}}(\beta)$ eine Diagonalmatrix ist, d.h. $\beta(b_i, b_j) = 0$ für $i \neq j$.

Beweis.
Durch vollständige Induktion nach $n := \dim V$. Für den Induktionsanfang $n=1$ ist nichts zu zeigen.

- Gilt $q_\beta(v) = 0$ für alle $v \in V$, so folgt aus Satz 6.4.3 $\beta = 0$, also $M_{\mathcal{B}}(\beta) = 0_n$ in jeder Basis $\mathcal{B}$ von V. In diesem Fall ist der Satz offensichtlich wahr.
- Sei also $b_1 \in V$ mit $q_\beta(b_1) \neq 0$. Die Linearform

$$\begin{aligned} \varphi: \ V &\to K \\ v &\mapsto \beta(b_1, v) \end{aligned}$$

ist wegen $\varphi(b_1) = \beta(b_1, b_1) \neq 0$ nicht die Nullform. Also ist die Dimension des Untervektorraums

$$U := \ker \varphi$$

gleich $\dim_K U = n-1$. Nach Induktionsannahme finde eine geordnete Basis $(b_2, \ldots, b_n)$ des Unterraums U mit $\beta(b_i, b_j) = 0$ für $i \neq j$ und $i, j \geq 2$. Wegen $b_1 \notin U$ ist $(b_1, \ldots, b_n)$ eine geordnete Basis von V. In ihr hat die symmetrische Bilinearform β wegen

$$\beta(b_1, b_i) = \varphi(b_i) = 0 \quad \text{für} \quad i = 2, \ldots, n$$

die gewünschte Diagonalgestalt.

□

Die diagonalisierende Basis $\mathcal{B}$ aus Satz 6.4.5 ist nicht eindeutig. Auch die Diagonalelemente sind nicht einmal bis auf die Reihenfolge eindeutig. Deshalb liegt eigentlich keine Normalform vor. Allerdings haben kongruente Diagonalmatrizen die gleiche Anzahl r_0 von Nullen auf der Diagonale.

Lemma 6.4.6.
Der Nullraum einer symmetrischen Bilinearform β auf einem K-Vektorraum V ist der Untervektorraum

$$N(\beta) := \{v \in V \mid \beta(v, w) = 0 \text{ für alle } w \in V\} \subset V \ .$$

Für jede diagonalisierende Basis $\mathcal{B} = (b_1, \ldots, b_n)$ wie in Satz 6.4.5 gilt

$$N(\beta) = \operatorname{span}_K\{b_i \mid \beta(b_i, b_i) = 0\} \ .$$

Beweis.

"$\supset$" Sei i so gewählt, dass für das Basiselement b_i die Gleichung $\beta(b_i, b_i) = 0$ gilt. Dann ist für jedes $w = \sum_{j=1}^{n} w_j b_j \in V$

$$\beta(b_i, w) = \sum_{j=1}^{n} w_j \beta(b_i, b_j) = w_i \beta(b_i, b_i) = 0 \ ,$$

also gilt $b_i \in N(\beta)$. Da $N(\beta)$ per Definition ein Untervektorraum ist, gilt auch $\mathrm{span}_K\{b_i \mid \alpha_i = 0\} \subset N(\beta)$.

"$\subset$" Sei $v \in N(\beta)$ beliebig. Schreibe $v = \sum_{j=1}^{n} v_j b_j$ und finde

$$0 = \beta(v, b_i) = \sum_{j=1}^{n} v_j \beta(b_j, b_i) = v_i \beta(b_i, b_i); \, .$$

Ist $\beta(b_i, b_i) \neq 0$, so muss $v_i = 0$ gelten. Also ist jedes $v \in N(\beta)$ eine Linearkombination derjenigen Basisvektoren b_i mit $\beta(b_i, b_i) = 0$.

□

Der Vergleich der Dimensionen zeigt dann die Gleichheit $r_0 = \dim_K N(\beta)$. Da der Nullraum nicht von der Wahl einer Basis von V abhängt, ist auch r_0 als Dimension der rechten Seite basisunabhängig. Ist $M_{\mathcal{B}}(\beta) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$, so ist β genau dann nicht-ausgeartet, wenn $\det M_{\mathcal{B}}(\beta) \neq 0$ gilt, also genau dann, wenn $\alpha_j \neq 0$ für alle j gilt, also wenn $r_0 = 0$ gilt.

Bemerkung 6.4.7.
Man kann den sogenannten Wittschen Relationensatz [4] zeigen:
Seien $\alpha_i, \beta_i \in K^\times$ und gelte die Kongruenz von Diagonalmatrizen

$$\mathrm{diag}(\alpha_1, \ldots, \alpha_n) \simeq \mathrm{diag}(\beta_1, \ldots, \beta_n) \ .$$

Dann lässt sich das n–Tupel $(\alpha_1, \ldots, \alpha_n)$ durch das fortgesetzte Anwenden der folgenden drei elementaren quadratischen Umformungen in das n–Tupel $(\beta_1, \ldots, \beta_n)$ überführen:

1. Vertauschen zweier Einträge
2. Multiplikation eines Eintrags mit einem Quadrat aus $K^\times$.
3. Ersetzen von (γ_1, γ_2) in $(\gamma_1, \ldots, \gamma_n)$ durch $(\gamma_1 + \gamma_2, (\gamma_1 + \gamma_2)\gamma_1\gamma_2)$, falls $\gamma_1 + \gamma_2 \neq 0$ (Wittsche Relation).

Für den Beweis verweisen wir etwa auf F. Lorenz, Lineare Algebra II, Kapitel VII §2 und eine Übungsaufgabe.

Wir wollen nun speziell quadratische Formen über den Körper $\mathbb{C}$ und $\mathbb{R}$ der komplexen bzw. reellen Zahlen untersuchen.

[4] Ernst Witt, 1911-1991, 1939-1991 Professor in Hamburg. Biographie unter http://www-history.mcs.st-andrews.ac.uk/Biographies/Witt.html

Satz 6.4.8.
Sei V ein endlich–dimensionaler $\mathbb{C}$–Vektorraum und

$$\beta : \ V \times V \to \mathbb{C}$$

eine symmetrische Bilinearform. Dann existiert eine geordnete Basis $\mathcal{B}$ von V, in der die darstellende Matrix die Gestalt

$$M_{\mathcal{B}}(\beta) = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{0, \ldots, 0}_{r_0})$$

hat. Die nicht-negativen ganzen Zahlen r und r_0 sind durch β eindeutig bestimmt und hängen nicht von der Wahl der geordneten Basis $\mathcal{B}$ ab.

Beweis.
Wegen Satz 6.4.5 gibt es eine geordnete Basis $\mathcal{B}' = (b'_1, \ldots, b'_n)$, in der gilt

$$M_{\mathcal{B}'}(\beta) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$$

Setze $b_i := \gamma_i b'_i$ mit

$$\gamma_i := \begin{cases} \frac{1}{\sqrt{\alpha_i}} & \text{falls } \alpha_i \neq 0 \\ 1 & \text{falls } \alpha_i = 0 \ . \end{cases}$$

Wir finden

$$\beta(b_i, b_j) = \gamma_i \gamma_j \beta(b'_i, b'_j) = \gamma_i \gamma_j \alpha_i \delta_{i,j} \ .$$

Es folgt für $\alpha_i = 0$ die Gleichung $\beta(b_i, b_i) = 0$ und für $\alpha_i \neq 0$ die Gleichung

$$\beta(b_i, b_j) = \frac{\alpha_i}{\alpha_i} = 1 \ .$$

Durch Umnummerierung erhält man die gewünschte Form.

Die Unabhängigkeit von r_0 und somit auch von r von der Wahl der geordneten Basis $\mathcal{B}$ folgt aus Lemma 6.4.6. □

Für quadratische Formen über $\mathbb{R}$ ist die Situation etwas komplizierter, weil nicht jede reelle Zahl das Quadrat einer reellen Zahl ist.

Satz 6.4.9. Sylvesterscher Trägheitssatz
Sei V ein endlich–dimensionaler $\mathbb{R}$–Vektorraum und

$$\beta : \ V \times V \to \mathbb{R}$$

eine symmetrische Bilinearform. Dann existiert eine geordnete Basis $\mathcal{B}$ von V, in der die darstellende Matrix die Gestalt

$$M_{\mathcal{B}}(\beta) = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r_+}, \underbrace{-1, -1, \ldots, -1}_{r_-}, \underbrace{0, \ldots, 0}_{r_0})$$

hat. Die nicht-negativen ganzen Zahlen r_+, r_- und r_0 sind durch die symmetrische Bilinearform β eindeutig bestimmt und hängen nicht von der Wahl der geordneten Basis $\mathcal{B}$ ab.

Beweis.

- Wegen Satz 6.4.5 gibt es eine geordnete Basis $\mathcal{B}' = (b'_1, \ldots, b'_n)$, in der gilt

$$M_{\mathcal{B}'}(\beta) = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$$

Setze $b_i := \gamma_i b'_i$ mit

$$\gamma_i := \begin{cases} \frac{1}{\sqrt{|\alpha_i|}} & \text{falls } \alpha_i \neq 0 \\ 1 & \text{falls } \alpha_i = 0 \, . \end{cases}$$

Wir finden

$$\beta(b_i, b_j) = \gamma_i \gamma_j \beta(b'_i, b'_j) = \gamma_i \gamma_j \alpha_i \delta_{i,j} \, .$$

Es folgt für $\alpha_i = 0$ die Gleichung $\beta(b_i, b_i) = 0$ und für $\alpha_i \neq 0$ die Gleichung

$$\beta(b_i, b_j) = \frac{\alpha_i}{|\alpha_i|} = \operatorname{sign}(\alpha_i) \in \{\pm 1\} \, .$$

Durch Umnummerierung erhält man die gewünschte Form.

- Wir wissen schon aus Lemma 6.4.6, dass r_0 als Dimension des Nullraums von β nicht von der Wahl der diagonalisierenden Basis abhängt.

- Wir zeigen:

$$r_+ = \max\{\dim_{\mathbb{R}} W \mid W \subset V \text{ Untervektorraum mit } q(v) > 0 \text{ für alle } v \in W \setminus \{0\}\} \, .$$

Die rechte Seite bezeichnen vorübergehend mit m. Aus dieser Darstellung von r_+ folgt sofort, dass r_+ und damit auch r_- nicht von der Wahl der Basis $\mathcal{B}$ abhängt. Zum Beweis betrachte zunächst den speziellen Untervektorraum

$$W_0 := \operatorname{span}_{\mathbb{R}}\{b_1, \ldots, b_{r_+}\} \, .$$

Für

$$v = \sum_{j=1}^{r_+} v_j b_j \in W_0 \setminus \{0\} \quad \text{mit} \quad v_j \in \mathbb{R}$$

gilt wegen $\beta(v, v) = 2q(v)$:

$$2q(v) = \beta(v, v) = \sum_{i,j=1}^{r_+} v_i v_j \beta(b_i, b_j) = \sum_{i=1}^{r_+} (v_i)^2 > 0 \, .$$

Daraus folgt für m als Maximum die Ungleichung

$$m \geq \dim_{\mathbb{R}} W_0 = r_+ \, .$$

Um die entgegengesetzte Ungleichung zu zeigen, zeigen wir, dass in jedem Untervektorraum W von V mit $\dim_{\mathbb{R}} W > r_+$ ein Vektor $v \in W \setminus \{0\}$ mit $q(v) \leq 0$ existiert. Sei also $\dim_{\mathbb{R}} W > r_+$. Wegen

$$\dim_{\mathbb{R}} W + \dim_{\mathbb{R}} \operatorname{span}_{\mathbb{R}}\{b_{r_+ + 1}, \ldots, b_n\} > r_+ + r_0 + r_- = \dim_{\mathbb{R}} V$$

finde mit Satz 2.5.3 ein $v \neq 0$ in

$$v \in W \cap \operatorname{span}_{\mathbb{R}}\{b_{r_+ + 1}, \ldots, b_n\} \ ,$$

Es gibt also ein $v \in W$, das die Darstellung

$$v = \sum_{j=r_+ +1}^{n} v_j b_j \quad \text{mit} \quad v_j \in \mathbb{R}$$

besitzt. Aus dieser Darstellung folgt die gewünschte Ungleichung:

$$2q(v) = \beta(v,v) = \sum_{j=r_+ +1}^{n} (v_j)^2 \, \beta(b_j, b_j) \leq 0 \; .$$

- Analog gilt

$$r_- = \max\{\dim_{\mathbb{R}} U \mid U \subset V \text{ mit } q(v) < 0 \text{ für alle } v \in U \setminus \{0\}\} \; .$$

□

Definition 6.4.10
Sei V ein $\mathbb{R}$-Vektorraum.

1. *Eine quadratische Form*
$$q : \; V \to \mathbb{R}$$
heißt
   - *positiv definit, falls $q(v) > 0$ für alle $v \in V$ mit $v \neq 0$ gilt, d.h. falls $r_- = r_0 = 0$ gilt.*
   - *negativ definit, falls $q(v) < 0$ für alle für alle $v \in V$ mit $v \neq 0$ gilt, d.h. falls $r_+ = r_0 = 0$ gilt..*
   - *positiv semidefinit, falls $q(v) \geq 0$ für alle $v \in V$ gilt, d.h. falls $r_- = 0$ gilt.*
   - *negativ semidefinit, falls $q(v) \leq 0$ für alle $v \in V$ gilt, d.h. falls $r_+ = 0$ gilt.*
   - *indefinit, falls es v_1 mit $q(v_1) > 0$ und v_2 mit $q(v_2) < 0$ gibt, d.h. falls $r_+ > 0$ und $r_- > 0$ gilt.*
2. *Sei q nicht-ausgeartet, d.h. es gelte $r_0 = 0$. Dann heißt die ganze Zahl*
$$\operatorname{sgn}(q) := r_+ - r_-$$
die Signatur von q und r_- auch der Trägheitsindex von q.

Betrachtung 6.4.11.
Wir betrachten nun eine quadratische Forme $q : \; \mathbb{R}^2 \to \mathbb{R}$. Wir versehen zusätzlich den Vektorraum $\mathbb{R}^2$ mit dem euklidischen Standardskalarprodukt $\langle \cdot, \cdot \rangle$, damit wir von Längen und Winkeln reden können. Wir interessieren uns für die euklidische Geometrie der Urbilder

$$q^{-1}(\rho) := \{x \in \mathbb{R}^2 \mid \; q(x) = \rho\} \quad \text{für} \quad \rho \in \mathbb{R} \; .$$

Die Elemente von $q^{-1}(\rho)$ sind also die Vektoren $(x_1, x_2) \in \mathbb{R}^2$, deren Koordinaten eine quadratische Gleichung in x_1, x_2 erfüllen, die durch q gegeben ist.

1. Sei die quadratische Form q positiv definit. Offenbar ist

$$\begin{aligned} &\text{für } \rho < 0 \quad \text{das Urbild } q^{-1}(\rho) = \emptyset \ , \\ &\text{für } \rho = 0 \quad \text{das Urbild } q^{-1}(0) = \{0\} \ . \end{aligned}$$

Für $\rho > 0$ nennen wir $q^{-1}(\rho)$ eine Ellipse.

Sei nun

$$\mathbb{S}^1 := \{x \in \mathbb{R}^2 | \ \|x\| = 1\} \ .$$

der Einheitskreis in der euklidischen Ebene. Auch der Einheitskreis ist durch eine quadratische Gleichung gegeben, nämlich

$$\|x\|^2 = (x_1)^2 + (x_2)^2 = 1 \ ,$$

die durch eine positiv definite quadratische Form gegeben ist. Wir zeigen: $E \subset \mathbb{R}^2$ ist genau dann eine Ellipse, wenn es eine invertible lineare Abbildung $T \in GL(2, \mathbb{R})$ gibt, so dass $E = T(\mathbb{S}^1)$ gilt.

Beweis.

$\Leftarrow$ Es ist $v \in T(\mathbb{S}^1)$ genau dann, wenn $T^{-1}v \in \mathbb{S}^1$ gilt, also wenn $\langle T^{-1}v, T^{-1}v\rangle = 1$. Betrachte also auf $\mathbb{R}^2$ die symmetrische Bilinearform $\beta(x, y) := \langle T^{-1}x, T^{-1}y\rangle$, die offensichtlich positiv definit ist und uns die positiv definite quadratische Form $q(x) = \frac{1}{2}\beta(x, x)$ liefert.

$\Rightarrow$ Sei q eine positiv definite quadratische Form und β die zugehörige symmetrische Bilinearform. Sei $\mathcal{B}$ die euklidische Standardbasis des $\mathbb{R}^2$. Nach dem Sylversterschen Trägheitssatz 6.4.9 existiert eine Basis $\mathcal{B}'$ des $\mathbb{R}^2$, in der die darstellende Matrix von β die folgende Diagonalgestalt hat:

$$M_{\mathcal{B}'}(\beta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 \ .$$

Setze $T := T^{\mathcal{B}}_{\mathcal{B}'} \in GL(2, \mathbb{R})$ und finde für die darstellende Matrix bezüglich der Standardbasis $\mathcal{B}$ mit der Transformationsformel aus Satz 6.2.6:

$$M_{\mathcal{B}}(\beta) = T^t M_{\mathcal{B}'}(\beta) T = T^t T \ .$$

Daher gilt für $v \in q^{-1}(\rho)$

$$2\rho = 2q(v) = \beta(v, v) = v^t M_{\mathcal{B}}(\beta) v = v^t T^t T v = (Tv)^t \, (Tv) = \|Tv\|^2 \ .$$

Daher ist $q(v) = \rho$ genau dann, wenn $\|Tv\|^2 = 2\rho$ gilt, also

$$q^{-1}(\rho) = \frac{1}{\sqrt{2\rho}} T^{-1} \left(\mathbb{S}^1\right) \ .$$

□

2. Ist die quadratische Form q negativ definit, so ist $-q$ positiv definit. Dieser Fall liefert also keine neuen Geometrien.

3. Für den indefiniten Fall finde eine geordnete Basis $\mathcal{B}' = (b_1, b_2)$ von $\mathbb{R}^2$, in der die darstellende Matrix der Bilinearform β zu q die Diagonalgestalt

$$M_{\mathcal{B}'}(\beta) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

hat. Es ist

$$q^{-1}(0) = \{v = v_1 b_1 + v_2 b_2 \in \mathbb{R}^2 |\, v_1^2 - v_2^2 = (v_1 + v_2)(v_1 - v_2) = 0\}$$

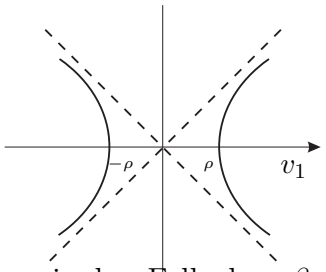
Wir finden als Urbild also ein Geradenkreuz aus zwei Ursprungsgeraden mit Richtungsvektoren $b_1 \pm b_2$ in der Ebene. Für $\rho \neq 0$ nennen wir $q^{-1}(\rho)$ eine Hyperbel. Wir führen die Standardhyperbel ein:

$$\mathbb{H}^1 := \{x \in \mathbb{R}^2 \,|\, (x_1)^2 - (x_2)^2 = 1\} \ .$$

Wie für Ellipsen zeigt man: $H \subset \mathbb{R}^2$ ist genau dann Hyperbel, wenn es ein $T \in GL(2, \mathbb{R})$ gibt, so dass $H = T(\mathbb{H}^1)$ gilt.

Eine Hyperbel hat zwei Zweige, etwa für $\rho > 0$:

$$q^{-1}(\rho) = \{v = v_1 b_1 + v_2 b_2 \in \mathbb{R}^2 | v_1 = \sqrt{2\rho + v_2^2}\} \cup \{v = v_1 b_1 + v_2 b_2 \in \mathbb{R}^2 | v_1 = -\sqrt{2\rho + v_2^2}\}$$



4. Schließlich betrachten wir den Fall, dass β positiv semi-definit, aber nicht positiv definit ist. In diesem Fall finden wir eine geordnete Basis $\mathcal{B}' = (b_1, b_2)$ von $\mathbb{R}^2$, in der die darstellende Matrix der Bilinearform die Diagonalgestalt

$$M_{\mathcal{B}'}(\beta) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

hat. Es ist

$$q^{-1}(\rho) = \{v = v_1 b_1 + v_2 b_2 \in \mathbb{R}^2 \mid (v_1)^2 = 2\rho\} \ .$$

Man findet also für $\rho < 0$ die leere Menge, für $\rho = 0$ die b_2-Achse und für $\rho > 0$ zwei affine Geraden parallel zur b_2-Achse.

Betrachtung 6.4.12 (Kegelschnitte)**.**

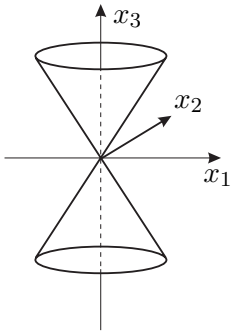
- Die quadratische Form auf $\mathbb{R}^n$ mit

$$Q\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n-1} (x_i)^2 - (x_n)^2$$

heißt Minkowski-Form. Sie ist nicht-ausgeartet mit $r_+ = n - 1$ und $r_- = 1$. Für $n = 4$ spielt sie eine zentrale Rolle in der speziellen Relativitätstheorie. Die dieser Theorie zu Grunde liegende Mathematik ist übrigens (weitgehend) lineare Algebra.

- Betrachte den Doppelkegel in $\mathbb{R}^3$

$$C = Q^{-1}(0) = \{x \in \mathbb{R}^3 |\, (x_3)^2 = (x_1)^2 + (x_2)^2\}$$
$$= \{x \in \mathbb{R}^3 |\, x_3 = \pm \| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \|\}$$



- Wir wollen die Schnittkurven des Doppelkegels C mit einer beliebigen affinen Ebene $E = p + E_0$ in $\mathbb{R}^3$ studieren, wobei $p \in \mathbb{R}^3$ und E_0 ein Untervektorraum der Dimension $\dim_{\mathbb{R}} E_0 = 2$ ist.

  Dazu betrachten wir die Einschränkung $q := Q_{|E_0}$ der Minkowski-Form Q auf die Ebene E_0. Sie versieht E_0 mit der Struktur eines zwei-dimensionalen quadratischen Raumes.

  Wir bezeichnen die zur Minkowski-Form Q gehörende Bilinearform auf $\mathbb{R}^3$ mit $B(\cdot,\cdot)$ und die zum quadratischen Raum (E_0, q) gehörende Bilinearform mit $\beta(\cdot,\cdot)$. Aus der Berechnung von $r_\pm$ in Satz 6.4.9 folgt, dass diese Indizes für die Einschränkung q von Q auf einen Untervektorraum kleiner sein müssen, also $r_+ \leq 2$ und $r_- \leq 1$.

  Andererseits gilt für den Index r_+ von q, dass $r_+ \geq 1$ sein muss. Denn Q ist positiv definit auf dem zwei-dimensionalen Untervektorraum $E_1 := \mathbb{R}e_1 + \mathbb{R}e_2 \subset \mathbb{R}^3$, und wegen

$$\dim_{\mathbb{R}}(E_0 \cap E_1) = \dim_{\mathbb{R}}(E_0) + \dim_{\mathbb{R}}(E_1) - \dim_{\mathbb{R}}(E_0 + E_1) \geq 2 + 2 - 3 = 1$$

  enthält der zwei-dimensionale Untervektorraum E_0 einen nicht-verschwindenden Vektor $v \in E_1$, für den also $q(v) = Q(v) > 0$ gilt. Also ist $r_+ \geq 1$.

  Wir müssen also für den zweidimensionalen quadratischen Raum (E_0, q) nur die folgenden drei Fälle unterscheiden:

Fall	r_+	r_-	r_0
I	2	0	0
II	1	1	0
III	1	0	1

- In den ersten beiden Fällen I,II ist der zweidimensionale quadratische Raum (E_0, q) nicht ausgeartet. Wir überlegen uns nun, dass dann der Fußpunkt $p \in E$ der affinen Ebene in $\mathbb{R}^3$ so gewählt werden kann, dass $B(x, p) = 0$ für alle $x \in E_0$ gilt.

  Daraus folgt sofort: für $x \in E_0$ gilt $x + p \in C$, genau für $0 = Q(x+p) = q(x) + Q(p)$, also ist die gesuchte Schnittkurve

$$C \cap E = q^{-1}(-Q(p)) + p\,. \qquad (*)$$

  Dann können wir die Schnittkurze $C \cap E$ in der affinen Ebene genauso untersuchen wie die Kurven in der Ebene in Betrachtung 6.5.10.

  Beweis:
  Sei $p_0 \in E$ beliebig. Da die Bilinearform β auf dem Untervektorraum E_0 in den Fällen I,II nicht ausgeartet ist, finde $y \in E_0$ mit $B(x, p_0) = \beta(x, y)$ für alle $x \in E_0$. Setze $p := p_0 - y \in E$. Dann gilt $B(x, p) = B(x, p_0) - B(x, y) = 0$ für alle $x \in E_0$.

- Wir sind weiterhin in den Fällen I und II: Wir zeigen, dass für diesen Fußpunkt $p \in E$ mit $B(x,p) = 0$ für alle $x \in E_0$ die folgenden drei Aussagen äquivalent sind:

  (a) $p = 0$,

  (b) $Q(p) = 0$,

  (c) $E = E_0$.

  Die Aussage (a) $\Rightarrow$ (b) ist trivial.
  (b) $\Rightarrow$ (c) Gelte also $Q(p) = 0$. Wir führen die Annahme $E \neq E_0$ zum Widerspruch. Dann liegt der Fußpunkt der affinen Ebene E nicht in dem Unvervektorraum E_0, also $p \notin E_0$. Es gilt $B(x,p) = 0$ für alle $x \in E_0$ sowieso und $B(p,p) = 2Q(p) = 0$ wegen (b) ebenfalls. Da p und E_0 ganz $\mathbb{R}^3$ erzeugen, folgt $B(v,p) = 0$ für alle $v \in \mathbb{R}^3$. Dies ist im Widerspruch dazu, dass die Minkowski-Form Q auf $\mathbb{R}^3$ nicht ausgeartet ist. Also folgt $E = E_0$, also (c).
  (c) $\Rightarrow$ (a) Gilt (c), ist also $E = E_0$, so ist $p \in E = E_0$, und wir können $\beta(p,x)$ für alle $x \in E_0$ betrachten. Es gilt $\beta(p,x) = B(p,x) = 0$ für alle $x \in E_0$. Da aber die quadratische Form q zu β auf E_0 in den Fällen I und II nicht ausgeartet ist, folgt daraus $p = 0$, also (a).

- Im Fall I gilt mit dem eben gefundenen p die Ungleichung $Q(p) \leq 0$.
  Beweis:
  Wäre $Q(p) > 0$, so wäre $p \neq 0$ und, wie eben gezeigt, deswegen $p \notin E_0$.

  Jedes $v \in \mathbb{R}^3$ ließe sich dann in der Form $v = tp + x$ mit $t \in \mathbb{R}$ und $x \in E_0$ schreiben. Es gälte also für jedes $v \in \mathbb{R}^3$

  $$Q(v) = t^2 Q(p) + 2tB(p,x) + Q(x) = t^2 Q(p) + q(x) \geq 0 \ ,$$

  im Widerspruch zur Tatsache, dass die Minkowski-Form Q auf $\mathbb{R}^3$ indefinit ist.

- Damit gibt es nach Betrachtung 6.4.11.1 zwei Unterfälle bei der Diskussion von $(*)$ im Fall I:
  – ist $p = 0$, so ist E ein zweidimensionaler Untervektorraum und $C \cap E = q^{-1}(0)$ ein Punkt, der Ursprung.
  – ist $p \neq 0$, so ist E eine affine Ebene und $C \cap E = q^{-1}(-(Q(p)) + p$ eine Ellipse.

- Im Fall II ist q indefinit. Wie in Betrachtung 6.4.11.3 unterscheiden wir bei der Diskussion von $(*)$ zwei Unterfälle:
  – ist $p = 0$, so ist E ein Untervektorraum und $C \cap E = q^{-1}(0)$ ein Geradenkreuz mit Schnitt im Ursprung.
  – ist $p \neq 0$, so ist E eine affine Ebene und $C \cap E = q^{-1}(-(Q(p)) + p$ eine Hyperbel.

- Im Fall III können wir nicht mehr $p \in E$ so wählen, dass $B(x,p) = 0$ für alle $x \in E_0$ gilt. Allerdings gilt

  $$x + p \in C \Leftrightarrow B(x+p, x+p) = q(x) + l(x) + c = 0$$

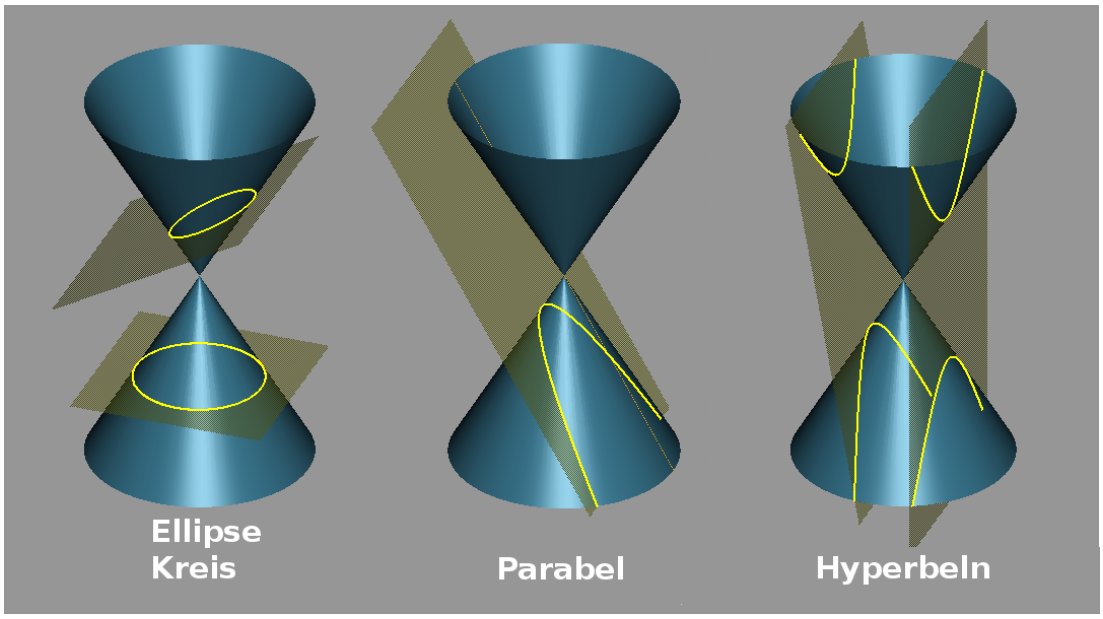
  mit $c := Q(p) \in \mathbb{R}$ und der Linearform $l(x) = B(x,p)$.
  – Für eine Ebene durch den Ursprung, $E = E_0$, wählen wir $p = 0$ als Fußpunkt und erhalten mit $c = 0$ und $l = 0$

  $$C \cap E = \{x \in E_0 \,|\, q(x) = 0\}$$

  was nach Betrachtung 6.4.11.4 eine Ursprungsgerade ist.
  – Ist E eine affine Ebene, so kann man zeigen, dass man eine Parabel erhält, also in einer geeigeneten Basis eine Menge der Form $P = \{v = v_1 b_1 + v_2 b_2 \,|\, v_2 = a(v_1 + b)^2 + d\}$.

- Wir fassen die Situation zusammen:

Typ	q	Raum E	$C \cap E$
I	nicht-ausgeartet: positiv definit	linear	Punkt $\{0\}$
		affin	Ellipse
II	nicht-ausgeartet: indefinit	linear	Geradenkreuz mit Schnitt im Ursprung
		affin	Hyperbel
III	positiv semidefinit	linear	Ursprungsgerade
		affin	Parabel.



6.5 Euklidische Vektorräume

Wir betrachten nun speziell reelle Vektorräume, um den Begriff des Skalarprodukts zu vertiefen.

Definition 6.5.1

1. *Sei V ein $\mathbb{R}$–Vektorraum. Eine positiv–definite symmetrische Bilinearform*

$$\langle \cdot, \cdot \rangle : \ V \times V \to \mathbb{R}$$

heißt ein (euklidisches) Skalarprodukt.

2. *Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt euklidischer Vektorraum.*

3. *Die Funktion $\| \cdot \| : V \to \mathbb{R}_{\geq 0}$*

$$\|x\| := \sqrt{\langle x, x \rangle}$$

auf einem euklidischen Vektorraum heißt Norm.

Beispiele 6.5.2.

1. Auf $V = \mathbb{R}^n$ heißt $\langle x, y \rangle := \sum_{i=1}^n x_i y_i = x^t E_n y$ das Standard-Skalarprodukt.

2. Der reelle Vektorraum $C^0([a,b],\mathbb{R})$ der stetigen Funktionen auf dem abgeschlossenen Interval $[a,b]$ wird durch
$$\langle f,g\rangle = \int_a^b f(x)g(x)dx$$
zum euklidischen Vektorraum. Der Konvergenzbegriff bezüglich der zugehörigen euklidischen Norm ist die Konvergenz im quadratischen Mittel. Sie tritt bei der Betrachtung der Fouriertransformation auf.

3. Untervektorräume euklidischer Vektorräume erhalten die Struktur eines eudklidischen Vektorraums durch Eindränkung des Skalarprodukts. Die direkte Summen euklidischer Vektorräume wird durch $\langle x_1+x_2, y_1+y_2\rangle := \langle x_1,y_1\rangle + \langle x_2,y_2\rangle$ zum euklidischen Vektorraum.

Satz 6.5.3 (Cauchy–Schwarz'sche Ungleichung)**.**
Sei $(V,\langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum. Dann gilt für alle $x,y\in V$
$$|\langle x,y\rangle| \le \|x\|\|y\| \,.$$
Gleichheit gilt genau dann, wenn die Familie (x,y) linear abhängig ist.

Beweis.

- Ohne Einschränkung sei $V=\operatorname{span}_{\mathbb{R}}\{x,y\}$. Ist $\dim_{\mathbb{R}} V \le 1$, so gilt das Gleichheitszeichen: dann ist etwa $y=\lambda x$ und somit
$$|\langle x,y\rangle| = |\lambda|\cdot|\langle x,x\rangle| = |\lambda|\cdot\|x\|^2 = \|x\|\cdot\|y\|\,.$$

- Sei also $\dim_{\mathbb{R}} V = 2$. Dann ist (x,y) eine geordnete Basis von V, in der das Skalarprodukt durch die reelle 2×2-Matrix
$$B = \begin{pmatrix} \langle x,x\rangle & \langle x,y\rangle \\ \langle y,x\rangle & \langle y,y\rangle \end{pmatrix}$$
dargestellt wird.

- Nach dem Sylversterschen Trägheitssatz 6.4.9 ist B als positiv definite Matrix kongruent zur Einheitsmatrix, $B=S^tS$, mithin $\det B = \det(S^t)\det S = \det(S)^2 > 0$. Das heißt aber:
$$\det B = \|x\|^2\|y\|^2 - |\langle x,y\rangle|^2 > 0\,.$$

□

Satz 6.5.4.
Sei V ein euklidischer Vektorraum. Dann gilt für $\|x\| := \sqrt{\langle x,x\rangle}$

1. $\|x\| \ge 0$ für alle $x\in V$
2. $\|x\| = 0 \;\Leftrightarrow\; x=0$
3. $\|\alpha x\| = |\alpha|\|x\|$ für alle $\alpha\in\mathbb{R}$ und $x\in V$.
4. $\|x+y\| \le \|x\| + \|y\|$ (Dreiecksungleichung)

Beweis.

1.,2. sind offensichtlich.

3. $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha^2\langle x, x\rangle} = |\alpha|\|x\|$.

4. Wegen der Cauchy–Schwarz'schen Ungleichung gilt

$$\|x+y\|^2 = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

□

Definition 6.5.5

1. *Für zwei vom Nullvektor verschiedene Vektoren x, y eines euklidischen Vektorraums heißt $\alpha \in [0, \pi]$ mit $\cos\alpha = \frac{\langle x,y\rangle}{\|x\|\|y\|}$ der Innenwinkel von x und y. Notation: $\alpha = \alpha(x, y)$.*

2. *Wir sagen, x stehe senkrecht auf y, wenn $\langle x, y\rangle = 0$ gilt. Notation: $x \perp y$. Dann ist $\alpha(x, y) = \frac{\pi}{2}$.*

Beispiel 6.5.6.
Der Vektorraum $V = C^\circ([0, \pi], \mathbb{R})$ der stetigen reellwertigen Funktionen auf dem abgeschlossenen Interval $[0, \pi]$ wird nach 6.5.2.2 durch das Skalarprodukt

$$\langle f, g\rangle = \int_0^\pi f(x)g(x)dx$$

zum euklidischen Vektorraum. Bezüglich dieses Skalarprodukts gilt

$$\langle \sin, \cos\rangle = \int_0^\pi \sin(x)\cos x dx = \left[\frac{1}{2}\sin^2 x\right]_0^\pi = 0\ ,$$

also $\sin \perp \cos$.

Aus dem Sylvesterschen Trägheitssatz 6.4.9 folgt sofort, dass ein n-dimensionaler euklidischer Vektorraum $(V, \langle\cdot,\cdot\rangle)$ eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ besitzt, in der

$$M_{\mathcal{B}}(\langle\cdot,\cdot\rangle) = E_n$$

gilt, d.h.

$$\langle b_i, b_j\rangle = \delta_{i,j} \quad \text{oder, was äquivalent ist:} \quad b_i \perp b_j \text{ für } i \neq j,\ \|b_i\| = 1\ .$$

Definition 6.5.7
Eine solche Basis eines endlich-dimensionalen euklidischen Vektorraums heißt Orthonormalbasis. Allgemeiner heißt eine Familie $(v_1, \ldots, v_k)$ von Vektoren $v_i \in V$ in einem euklidischen Vektorraum ein Orthonormalsystem, wenn $\langle v_i, v_j\rangle = \delta_{ij}$ gilt.

Lemma 6.5.8.
Jedes Orthonormalsystem in einem euklidischen Vektorraum ist linear unabhängig.

Beweis.
Sei $(v_1, \ldots, v_k)$ ein Orthonormalsystem und gelte

$$\sum_{i=1}^{k} \alpha_i v_i = 0 \qquad \text{mit} \quad \alpha_i \in \mathbb{R} \, .$$

Dann gilt

$$0 = \langle 0, v_i \rangle = \sum_{j=1}^{k} \alpha_j \langle v_j, v_i \rangle = \alpha_i \quad \text{für alle} \quad i = 1, \ldots, k \ .$$

□

Lemma 6.5.9.
Ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V, dann gilt für alle $v \in V$

$$v = \sum_{i=1}^{n} \langle v, b_i \rangle b_i$$

Beweis.
Da $\mathcal{B}$ eine Basis von V ist, gibt es Koeffizienten $\alpha_i \in \mathbb{R}$, so dass $v = \sum_{i=1}^{n} \alpha_i b_i$ gilt. Wir rechnen:

$$\langle v, b_j \rangle = \sum_{i=1}^{n} \alpha_i \langle b_i, b_j \rangle = \alpha_j \ .$$

□

Da euklidische Skalarprodukte symmetrische Bilinearformen sind, fallen das orthogonale Links- und Rechtskomplement zusammen; wir bezeichnen es mit $X^\perp$. Da das euklidische Skalarprodukt nicht ausgeartet ist, folgt nach Bemerkung 6.2.16.2 $\dim X + \dim X^\perp = \dim V$ für jeden Untervektorraum X von V. Ferner ist $X \cap X^\perp = \{0\}$, denn aus $v \in X \cap X^\perp$ folgt $\langle v, v \rangle = 0$. Da das euklidische Skalarprodukt positiv definit ist, folgt $v = 0$; somit ist die Summe direkt, es gilt $V = X \oplus X^\perp$.

Definition 6.5.10
Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich–dimensionaler euklidischer Vektorraum und U ein Untervektorraum von V. Dann heißt die lineare Abbildung

$$P_U : \ V \to V$$

mit $(P_U)_{|U} = \mathrm{id}_U$ und $(P_U)_{|U^\perp} = 0$ die orthogonale Projektion auf den Untervektorraum U.

Betrachtung 6.5.11.

- Ist $\mathcal{B}_1 = (b_1, \ldots, b_r)$ eine Orthonormalbasis von U und $\mathcal{B}_2 = (b_{r+1}, \ldots, b_n)$ eine Orthonormalbasis von $U^\perp$, dann ist $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V.

- Es ist nach Lemma 6.5.9 für $v \in V$

$$v = \sum_{j=1}^{n} \langle v, b_j \rangle b_j = \underbrace{\sum_{j=1}^{r} \langle v, b_j \rangle b_j}_{\in U} + \underbrace{\sum_{j=r+1}^{n} \langle v, b_j \rangle b_j}_{\in U^\perp} .$$

Damit erhalten wir die folgenden Formeln für die Projektionen:

$$P_U(v) = \sum_{j=1}^{r} \langle v, b_j \rangle b_j \quad \text{und} \quad P_{U^\perp}(v) = \sum_{j=r+1}^{n} \langle v, b_j \rangle b_j = (\mathrm{id}_V - P_U)(v) .$$

Es gilt

$$(P_U)^2 = P_U \quad \text{und} \quad (P_{U^\perp})^2 = P_{U^\perp} ;$$

also sind P_U und $P_{U^\perp}$ Idempotente. Ferner ist $\mathrm{id}_V = P_U + P_{U^\perp}$.

Betrachtung 6.5.12 (Schmidtsches Orthonormalisierungsverfahren)**.**

- Wir wollen aus einer beliebigen Basis $(v_1, \ldots, v_n)$ eines endlich–dimensionalen euklidischen Vektorraum eine Orthonormalbasis gewinnen.
1.Schritt: Setze $b_1 := \frac{v_1}{\|v_1\|}$. Dies ist wegen $v_1 \neq 0$ definiert.
$(k+1)$–ter Schritt: sei $b_1, \ldots, b_k$ ein Orthonormalsystem mit $\mathrm{span}_{\mathbb{R}}(b_1, \ldots, b_k) = \mathrm{span}_{\mathbb{R}}(v_1, \ldots, v_k)$. Dann ist

$$\widetilde{b}_{k+1} := v_{k+1} - P_{\mathrm{span}(b_1 \ldots b_k)} v_{k+1} = P_{\mathrm{span}(b_1 \ldots b_k)^\perp}(v_{k+1}) \neq 0 ,$$

da sonst v_{k+1} Linearkombination von $(b_1 \ldots b_k)$ und somit auch von $(v_1 \ldots v_k)$ wäre. Setze $b_{k+1} := \frac{\widetilde{b}_{k+1}}{\|\widetilde{b}_{k+1}\|}$ Dies ist ein normierter Vektor. Außerdem gilt für $1 \leq i \leq k$

$$\langle b_i, \widetilde{b}_{k+1} \rangle = \langle b_i, P_{\mathrm{span}(b_1 \ldots b_k)^\perp}(v_{k+1}) \rangle = 0$$

- Beispiel: Der drei-dimensionale reelle Vektorraum

$$V = \{ f \in \mathbb{R}[x] \mid \mathrm{grad}\, f \leq 2 \}$$

hat als eine Basis $v_1 = 1$, $v_2 = x$, $v_3 = x^2$. Als euklidisches Skalarprodukt betrachten wir wie in Beispiel 6.5.2.2

$$\langle f, g \rangle = \int_0^1 f(x) g(x) dx$$

1. Schritt: Setze $b_1 := v_1$, da

$$\|v_1\|^2 = \int_0^1 v_1^2 dx = 1 .$$

2. Schritt: Es gilt $\langle v_2, b_1 \rangle = \int_0^1 x \cdot 1 \cdot dx = \frac{1}{2}$; also setze

$$\widetilde{b}_2 := v_2 - P_{v_1} v_2 = v_2 - \frac{1}{2} b_1 = x - \frac{1}{2} .$$

Um den vom Nullvektor verschiedenen Vektor $\widetilde{b}_2$ zu normieren, berechnen wir

$$\left\| x - \frac{1}{2} \right\|^2 = \int_0^1 \left(x - \frac{1}{2} \right)^2 dx = \int_0^1 \left(x^2 - x + \frac{1}{4} \right) dx$$
$$= \left[\frac{x^3}{3} - \frac{x^2}{2} + \frac{1}{4} x \right]_0^1 = \frac{1}{12} .$$

Wir finden als Norm $\|\widetilde{b}_2\| = \|x - \frac{1}{2}\| = \frac{1}{2\sqrt{3}}$, somit als normierten Vektor $b_2 := 2\sqrt{3}(x - \frac{1}{2})$.

3. Schritt: Man findet durch ähnliche Rechnung $\langle v_3, b_1\rangle = \frac{1}{3}$ und $\langle v_3, b_2\rangle = \frac{1}{6}\sqrt{3}$. Somit ist

$$\widetilde{b}_3 = x^2 - \frac{1}{3}\cdot 1 - \frac{\sqrt{3}}{6}2\sqrt{3}(x - \frac{1}{2}) = x^2 - x + \frac{1}{6}\ .$$

Ferner ist $\|x^2 - x + \frac{1}{6}\|^2 = \frac{1}{180}$, also setze

$$b_3 := \sqrt{180}(x^2 - x + \frac{1}{6})\ .$$

- So erhält man wichtige Klassen von Polynomen. Wählt man statt des Intervalls $[0,1]$ das Intervall $[-1,1]$, so erhält man die Legendre-Polynome

$$P_0(x) = 1,\ P_1(x) = x,\ P_2(x) = \frac{1}{2}(3x^2 - 1), \ldots$$

die Lösungen der Differentialgleichung

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(1 - x^2\right)\ y'(x)\right] + n(n+1)\, y(x) = 0$$

sind.

Die Hermite-Polynome

$$H_0(x) = 1,\ H_1(x) = 2x,\ H_2(x) = (2x)^2 - 2,\ H_3(x) = (2x)^3 - 6(2x) = 8x^3 - 12x, \ldots$$

sind bezüglich der Gewichtsfunktion $e^{-x^2/2}$ orthogonal,

$$\int_{-\infty}^{\infty} e^{-x^2/2} H_n(x) H_m(x) dx = \sqrt{2\pi} n! \delta_{n,m}$$

und erfüllen die Hermitesche Differentialgleichung

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0\ .$$

Für die Laguerre-Polynome wird die Gewichtsfunktion e^{-x} auf der Halbachse $[0,\infty)$ betrachtet.

Betrachtung 6.5.13.
Die allgemeinen Überlegungen aus Kapitel 6.2 zeigen, dass für endlich–dimensionale euklidische Vektorräume gilt:

- Die lineare Abbildung $\tau_V :\ V \to V^*$
$$v \mapsto \langle v, \cdot\rangle$$
ist ein Isomorphismus, der freilich vom euklidischen Skalarprodukt abhängt. Es gilt also

$$\tau_V(v)(w) = \langle v, w\rangle \quad \text{(a1)} \quad \text{und} \quad \langle \tau_V^{-1}(\beta), w\rangle = \beta(w)\ \text{(a2)}\ .$$

- Sei (b_i) eine Orthonormalbasis. Der Vergleich der Identitäten $\delta_{ij} = b_i^*(b_j)$ und $\tau_V(b_i)(b_j) = \langle b_i, b_j\rangle = \delta_{i,j}$ für alle $i,j = 1,\ldots,n$ zeigt

$$\tau_V(b_i) = b_i^*\ .$$

- $\tau_V(U^\perp) = U^\circ$ für jede Teilmenge $U \subset V$. In der Tat liegt $\varphi \in U^\circ$ genau dann, wenn für alle $u \in U$ gilt $0 = \varphi(u) = \langle \tau_V^{-1}(\varphi), u\rangle$. Dies ist aber äquivalent zu $\tau_V^{-1}(\varphi) \in U^\perp$.

Satz 6.5.14.
Seien V, W endlich–dimensionale euklidische Vektorräume und τ_V, τ_W die zugehörigen Isomorphismen. Sei $F : V \to W$ eine lineare Abbildung. Dann hängt die adjungierte Abbildung $F^\wedge : W \to V$ mit der dualen Abbildung $F^* : W^* \to V^*$ folgendermaßen zusammen:

$$F^\wedge = \tau_V^{-1} \circ F^* \circ \tau_W \ ,$$

d.h. das Diagramm

$$\begin{array}{ccc} V & \xleftarrow{F^\wedge} & W \\ {\scriptstyle \tau_V}\downarrow & & \downarrow{\scriptstyle \tau_W} \\ V^* & \xleftarrow{F^*} & W^* \end{array}$$

kommutiert.

Beweis.
Es gilt für alle $w \in W$ und $v \in V$

$$\langle \tau_V^{-1} \circ F^* \circ \tau_W(w), v\rangle \overset{(a2)}{=} F^* \circ \tau_W(w)(v) \overset{\text{def. } F^*}{=} \tau_W(w)\,(Fv) \overset{(a1)}{=} \langle w, Fv\rangle \overset{\text{def. } F^\wedge}{=} \langle F^\wedge w, v\rangle \ .$$

□

Satz 6.5.15.
Seien V, W endlich–dimensionale euklidische Vektorräume und sei $F : V \to W$ eine lineare Abbildung.

1. Sind $\mathcal{A}, \mathcal{B}$ geordnete Orthonormalbasen von V bzw. W, so gilt
$$M_{\mathcal{A}}^{\mathcal{B}}(F^\wedge) = M_{\mathcal{B}}^{\mathcal{A}}(F)^t$$
2. Es gilt
$$\text{Im } F^\wedge = (\ker F)^\perp \qquad \text{und} \qquad \ker F^\wedge = (\text{Im } F)^\perp \ .$$
3. Ist F selbstadjungierter Endomorphismus, d.h. ist $V = W$ und gilt $F = F^\wedge$, so gilt
$$V = \ker F \oplus \text{Im } F$$
und die Untervektorräume sind orthogonal.

Die dritte Aussage wird in Satz 6.7.9 verallgemeinert werden zu der Aussage, dass selbstadjungierte Abbildungen diagonalisierbar mit orthogonalen Eigenräumen sind.

Beweis.
1. folgt mit der gleichen Rechnung wie im Beweis von Lemma 6.2.12.
2. Wir rechnen
$$\begin{aligned} \text{Im } F^\wedge &= \text{Im } \left(\tau_V^{-1} \circ F^* \circ \tau_W\right) \\ &= \tau_V^{-1}\text{Im } (F^* \circ \tau_W) \\ &= \tau_V^{-1}\text{Im } F^* && \text{da } \tau_W \text{ ein Isomorphismus ist.} \\ &= \tau_V^{-1}\,(\ker F)^\circ && \text{wegen Satz 6.1.13} \\ &= (\ker F)^\perp && \text{wegen Betrachtung 6.5.13} \end{aligned}$$
Die Aussage über $\ker F^\wedge$ folgt analog.

3. Wir rechnen mit orthogonalen Zerlegungen

$$V = \ker F \oplus (\ker F)^{\perp} \overset{2.}{=} \ker F \oplus \text{Im } F^{\wedge} \overset{F \text{ s.a.}}{=} \ker F \oplus \text{Im } F \ .$$

□

6.6 Orthogonale Abbildungen

In diesem Abschnitt sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich–dimensionaler euklidischer Vektorraum.

Definition 6.6.1
Die Isometrien Φ eines euklidischen Vektorraums V heißen orthogonale Abbildungen. Ein Endomorphismus $\Phi \in \text{End}(V)$ ist also genau dann eine orthogonale Abbildung, wenn für alle $v, w \in V$

$$\langle \Phi(v), \Phi(w) \rangle = \langle v, w \rangle \qquad \text{für alle} \quad v, w \in V$$

gilt.

Lemma 6.6.2.
Seien $\Phi, \Psi \in \text{End}\, V$ orthogonale Endomorphismen. Dann gilt für alle $v, w \in V$

1. $\|\Phi(v)\| = \|v\|$, d.h. Φ ist normerhaltend. Umgekehrt ist jede normerhaltende Abbildung orthogonal.
2. Für $v \neq 0$ und $w \neq 0$ gilt

$$\alpha(v, w) = \alpha(\Phi(v), \Phi(w)), \quad \text{d.h. } \Phi \text{ ist winkeltreu}$$

3. Φ ist ein Isomorphismus und die Umkehrabbildung Φ^{-1} ist ebenfalls orthogonal.
4. Die Verkettung $\Phi \circ \Psi$ orthogonaler Abbildungen ist orthogonal.
5. Ist $\lambda \in \mathbb{R}$ ein Eigenwert von Φ, so gilt $\lambda = \pm 1$.

Beweis.
Die meisten Aussagen folgen direkt aus der Definition.

1. Aus der Polarisierungsformel

$$\langle v, w \rangle = \frac{1}{2} \left(\|v + w\|^2 - \|v\|^2 - \|w\|^2 \right)$$

folgt aus $\|\Phi(v)\| = \|v\|$ für alle $v \in V$ sofort die Orthogonalität von Φ.

3. $v \in \ker \Phi \Rightarrow 0 = \|\Phi(v)\| = \|v\| \Rightarrow v = 0$, also ist Φ injektiv; wegen $\dim_{\mathbb{R}} V < \infty$ folgt, dass Φ ein Isomorphismus ist.
5. Ist $v \neq 0$ Eigenvektor zum Eigenwert $\lambda \in \mathbb{R}$, so gilt

$$\|v\| = \|\Phi v\| = \|\lambda v\| = |\lambda| \cdot \|v\| \ ,$$

also $|\lambda| = 1$.

□

Definition 6.6.3
Sei $(V, \langle\cdot,\cdot\rangle$ ein euklidischer Vektorraum.

1. *Die Gruppe*
$$O(V) := \{\Phi \in \mathrm{End}(V) |\ \Phi \ \textit{orthogonal}\}$$
heißt orthogonale Gruppe des euklidischen Vektorraum V, genauer von $(V, \langle\cdot,\cdot\rangle)$.
Die Untergruppe
$$SO(V) := \{\Phi \in O(V) |\ \det \Phi > 0\}$$
heißt spezielle orthogonale Gruppe von V. Beide Gruppen sind Untergruppen der allgemeinen linearen Gruppe $GL(V)$.

2. *Eine Matrix $A \in M(n \times n, \mathbb{R})$ heißt orthogonal, falls $A^t A = E_n$. gilt.*

3. *Wir führen die Gruppen*
$$O(n) := \{A \in M(n \times n, \mathbb{R}) |\ A^t A = E_n\}$$
$$SO(n) := \{A \in O(n) |\ \det A > 0\}$$
ein.

Für eine orthogonale Matrix gilt $A^t A = E_n$ und somit $A^t = A^{-1}$. Da das linksinverse Element auch ein rechtsinverses Element ist, folgt auch $AA^t = E_n$.

Satz 6.6.4.
Sei $\mathcal{B}$ eine geordnete Orthonormalbasis von V. Ein Endomorphismus $\Phi \in \mathrm{End}(V)$ ist genau dann orthogonal, wenn seine darstellende Matrix $A := M_{\mathcal{B}}(\Phi)$ bezüglich dieser Orthonormalbasis orthogonal ist.

Beweis.
Schreibe die Vektoren $v = \sum_{i=1}^n x_i b_i$ und $w = \sum_{i=1}^n y_i b_i$ als Linearkombination der Vektoren der Orthonormalbasis $\mathcal{B}$.
Setze $x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ und $\quad y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$.
Dann ist
$$\langle v, w\rangle = \sum_{i,j=1}^n x_i y_j \langle b_i, b_j\rangle = \sum_{i=1}^n x_i y_i = (x_1, \ldots, x_n) \begin{pmatrix} y_1 \\ \ldots \\ y_n \end{pmatrix} = x^t \cdot y\ .$$
Also gilt
$$\Phi \text{ orthogonal } \Leftrightarrow x^t y = \langle v, w\rangle = \langle \Phi v, \Phi w\rangle = (Ax)^t Ay = x^t (A^t A) y$$
für alle $x, y \in \mathbb{R}^n$. Hieraus folgt aber $A^t A = E_n$. □

Bemerkungen 6.6.5.

1. Es folgt sofort für die Determinante einer orthogonalen Matrix

$$1 = \det E_n = \det\left(A^t A\right) = \left(\det\left(A\right)\right)^2 \quad ,$$

also $\det\left(A\right) \in \{\pm 1\}$. Daher gilt für die spezielle orthogonale Gruppe genauer

$$SO(V) := \{\Phi \in O(V) |\ \det \Phi = 1\} \quad \text{und} \quad SO(n) := \{A \in O(n) |\ \det A = 1\} \ .$$

2. Wir rechnen

$$\delta_{ik} = (A^t A)_{ik} = \sum_{j=1}^{n} a_{ji} a_{jk} \ ,$$

also bilden bei einer orthogonalen Matrix die Spaltenvektoren eine Orthonomalbasis des $\mathbb{R}^n$. Ähnlich folgt auch aus $AA^t = E_n$, dass

$$\delta_{ik} = (AA^t)_{ik} = \sum_{j=1}^{n} a_{ij} a_{kj} \ ,$$

also bilden auch die Zeilenvektoren eine orthogonalen Matrix einer Orthonormalbasis des $\mathbb{R}^n$.

Beispiele 6.6.6.

1. Eindimensionale euklidische Vektorräume, $n = 1$: eine Matrix $A = (a)$ ist genau dann orthogonal, wenn $(a^2) = A^t A = E_1 = (1)$ gilt, also für $a = \pm 1$. Es folgt

$$\begin{aligned} O(1) &= \{\pm 1\} \cong \mathbb{Z}_2 \\ SO(1) &= \{+1\} \quad \text{ist die triviale Gruppe.} \end{aligned}$$

2. Für $n = 2$ suchen wir Matrizen $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit reellen Einträgen und, wegen der Orthogonalität der Spaltenvektoren,

$$a^2 + c^2 = 1 \ , \quad b^2 + d^2 = 1 \ , \quad ab + cd = 0 \ .$$

Wegen der ersten beiden Gleichungen finden wir Winkel α, α' mit

$$a = \cos\alpha \, , \quad c = \sin\alpha \, , \quad b = \sin\alpha' \quad \text{und} \quad d = \cos\alpha' \ .$$

Ferner gilt

$$0 = ab + cd = \cos\alpha \sin\alpha' + \sin\alpha \cos\alpha' = \sin(\alpha + \alpha') \ .$$

Also ist $\alpha' = -\alpha \bmod 2\pi$ oder $\alpha' = \pi - \alpha \bmod 2\pi$. Also besteht

$$SO(2) = \left\{ M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$$

aus den Drehungen um den Ursprung und

$$O(2) = SO(2) \cup \left\{ M(S_\theta) = \begin{pmatrix} \cos 2\theta & +\sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$$

aus Drehungen um den Ursprung und Spiegelungen an Ursprungsgeraden, vgl. Beispiel 2.1.4.

Wir wollen nun noch die orthogonalen Matrizen für $n \geq 3$ untersuchen. Zentral wird das folgende Lemma sein:

Lemma 6.6.7.
Sei V ein endlich–dimensionaler euklidischer Vektorraum. Ist $\Phi \in O(V)$ und ist $W \subset V$ ein Φ–invarianter Untervektorraum, dann ist das orthogonale Komplement $W^\perp$ ebenfalls ein Φ–invarianter Untervektorraum.

Beweis.
Da V endlich–dimensional ist, ist Φ als injektive Abbildung ein Automorphismus. Ebenso ist $\Phi_{|W}$ ein Automorphismus des invarianten Untervektorraums W. Sei $v \in W^\perp$. Sei $w \in W$ beliebig; mit $w' := \Phi^{-1}(w) \in W$ finde:

$$\langle \Phi(v), w \rangle = \langle \Phi(v), \Phi(w') \rangle \overset{\Phi \text{ orth.}}{=} \langle v, w' \rangle = 0 \; .$$

Also ist $\Phi(v) \in W^\perp$. □

Lemma 6.6.8.
Ist $\Phi \in O(V)$, so besitzt V einen Φ–invarianten Untervektorraum W der Dimension 1 oder 2.

Beweis.

- Betrachte die "Symmetrisierung" $\Psi := \Phi + \Phi^{-1} \in \mathrm{End}(V)$ und finde

$$\begin{aligned} \langle \Psi(v), w \rangle &\overset{\text{def}}{=} \langle \Phi(v), w \rangle + \langle \Phi^{-1}(v), w \rangle \\ &= \langle \Phi^{-1}\Phi(v), \Phi^{-1}(w) \rangle + \langle \Phi\Phi^{-1}(v), \Phi(w) \rangle \qquad [\Phi \text{ und } \Phi^{-1} \text{ sind orthogonal}] \\ &= \langle v, \Psi(w) \rangle \; . \end{aligned}$$

  Eine solche Abbildung heißt selbstadjungiert, vgl. Satz 6.5.15. Wir werden in Satz 6.7.9 zeigen: für selbstadjungierte Abbildungen Ψ eines euklidischen Vektorraums gibt es eine Orthonormalbasis $\mathcal{B} = (b_1, \ldots, b_n)$ von Eigenvektoren:

$$\Psi(b_i) = \lambda_i b_i \quad \text{mit} \quad \lambda_i \in \mathbb{R} \; .$$

- Setze $W := \mathrm{span}_{\mathbb{R}}\{b_1, \Phi(b_1)\}$. Dann ist $\dim_{\mathbb{R}} W \in \{1, 2\}$. Wegen

$$\Phi(\mu b_1 + \nu\Phi(b_1)) = \mu\Phi(b_1) + \nu\Phi^2(b_1) \quad \text{mit} \;\; \mu, \nu \in \mathbb{R}$$

  kommt es darauf an, $\Phi^2(b_1) \in W$ zu zeigen; dann ist $\Phi(\mu b_1 + \nu\Phi(b_1)) \in W$ klar. Dies folgt aus

$$\Phi^2(b_1) = \Phi\left(\Phi(b_1) + \Phi^{-1}(b_1) - \Phi^{-1}(b_1)\right) = \Phi\Psi(b_1) - b_1 = \lambda_1\Phi(b_1) - b_1 \in W \; .$$

□

Satz 6.6.9.
Sei V ein endlich–dimensionaler euklidischer Vektorraum. Ist $\Phi \in O(V)$, so gibt es eine Orthogonalbasis $\mathcal{B}$ von V, in der die darstellende Matrix die folgende Blockdiagonalgestalt hat:

$$M_{\mathcal{B}}(\Phi) = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{A_1} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \boxed{A_x} \end{pmatrix}$$

wobei $A_j = M\left(R_{\theta_j}\right) \in M(2 \times 2, \mathbb{R})$ mit $\theta_j \in \mathbb{R}$ Drehmatrizen sind.

Beweis.
durch vollständige Induktion nach $n := \dim_{\mathbb{R}} V$.

- Der Induktionsanfang für $n = 1$ und $n = 2$ ist durch Beispiel 6.6.6.1 und 2 klar; beachte, dass man wie in Beispiel 3.6.6 für eine Spiegelung S_θ an einer Ursprungsgeraden stets eine Basis finden kann mit
$$M_{\mathcal{B}}(S_\theta) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$
- Lemma 6.6.8 garantiert die Existenz eines Φ–invarianten Untervektorraums W von V der Dimension Eins oder Zwei. Wegen Lemma 6.6.7 ist auch das orthogonale Komplement $W^\perp$ ein Φ–invarianter Untervektorraum. Man setzt nun eine Orthonormalbasis $\mathcal{B}_1$ von W und $\mathcal{B}_2$ von $W^\perp$, in der die jeweiligen Einschränkungen von Φ die gewünsche Blockdiagonalgestalt haben, zu einer Orthonormalbasis $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ von V zusammen: man findet so die gewünschte Blockdiagonalgestalt:
$$M_{\mathcal{B}}(\Phi) = \left(\begin{array}{c|c} \left(M_{\mathcal{B}_1}\left(\Phi|_W\right)\right. & 0 \\ \hline 0 & \left. M_{\mathcal{B}_2}\left(\Phi|_{W^\perp}\right)\right) \end{array} .$$

□

Beispiel 6.6.10.
Im Falle $\dim_{\mathbb{R}} V = 3$ treten nur die folgenden beiden Fälle auf:

- Drehungen um eine Ursprungsgerade. Hier hat man in einer geeigneten Orthonormalbasis von $\mathbb{R}^3$ die darstellende Matrix
$$M_{\mathcal{B}}(\Phi) = \begin{pmatrix} 1 & 0 \\ 0 & M\left(R_\theta\right) \end{pmatrix} \quad \text{falls} \quad \det \Phi = 1$$
- Spiegelungen an Ebenen durch den Ursprung, gefolgt durch eine Drehung mit Drehachse senkrecht zur Spiegelebene. Hier hat man in einer geeigneten Orthonormalbasis von $\mathbb{R}^3$ die darstellende Matrix
$$M_{\mathcal{B}}(\Phi) = \begin{pmatrix} -1 & 0 \\ 0 & M\left(R_\theta\right) \end{pmatrix} \quad \text{falls} \quad \det \Phi = -1 .$$

Im ersten Fall heißt $\mathbb{R}b_1 = \mathrm{Eig}(\Phi, 1)$ Drehachse von Φ und θ Drehwinkel von Φ.

Beweis.
Wir müssen nur noch Diagonalmatrizen betrachten, die Einträge ± 1 auf der Diagonale haben, und zeigen, dass sie Drehungen um Ursprungsgeraden oder Spiegelungen gefolgt von Drehungen beschreiben:

- $\mathrm{diag}(\pm 1, 1, 1) = \begin{pmatrix} \pm 1 & 0 \\ 0 & M(R_0) \end{pmatrix}$. Es liegt die Identität oder die Spiegelung an der $x_2 - x_3$-Ebene vor.
- $\mathrm{diag}(\pm 1, -1, -1) = \begin{pmatrix} \pm 1 & 0 \\ 0 & M(R_\pi) \end{pmatrix}$. Es liegt eine Punktspiegelung am Ursprung oder eine Drehung um π um die x_1-Achse vor.
- $\mathrm{diag}(1, 1, -1)$ und $\mathrm{diag}(1, -1, 1)$ sind ähnlich zu $\mathrm{diag}(-1, 1, 1)$. Es liegen Spiegelungen an den Koordinatenebenen vor.
- $\mathrm{diag}(-1, 1, -1)$ und $\mathrm{diag}(-1, -1, 1)$ sind ähnlich zu $\mathrm{diag}(1, -1, -1)$. Es liegen Drehungen um Koordinatenachsen um den Winkel π vor.

□

6.7 Selbstadjungierte und unitäre Endomorphismen

Wir führen zunächst das komplexe Analogon euklidischer Vektorräume ein. Hierbei spielt die komplexe Konjugation $z \mapsto \overline{z}$ eine wichtige Rolle.

Definition 6.7.1

1. *Sei V ein komplexer Vektorraum. Eine Abbildung*

$$h : V \times V \to \mathbb{C}$$

heißt Sesquilinearform, falls für alle $\alpha, \beta \in \mathbb{C}$ und $v, v', v'', w, w', w'' \in V$ gilt

$$h(\alpha v' + \beta v'', w) = \alpha h(v', w) + \beta h(v'', w)$$
$$h(v, \alpha w' + \beta w'') = \bar{\alpha} h(v, w') + \overline{\beta} h(v, w) \ .$$

Man sagt, die Sesquilinearform ist im zweiten Argument antilinear.

2. *Eine Sesquilinearform heißt hermitesch, falls gilt*

$$h(v, w) = \overline{h(w, v)} \quad \text{für alle } v, w \in V \ .$$

3. *Eine hermitesche Sesquilinearform heißt positiv definit, falls gilt*

$$h(v, v) > 0 \quad \text{für alle } v \in V \setminus \{0\} \ .$$

Bemerkungen 6.7.2.

- Auf $V = \mathbb{C}^n$ definiert $h(v,w) := \sum_{i=1}^n v_i \bar{w}_i$ eine hermitesche Sesquilinearform. Es gilt für alle $v \neq 0$

$$h(v,v) = \sum_{i=1}^n |v_i|^2 > 0 \ ,$$

also ist h positiv definit. h heißt Standard–Sesquilinearform auf $\mathbb{C}^n$.

- Wie im euklidischen Fall sieht man: ist $h : V \times V \to \mathbb{C}$ eine positiv–definite hermitesche Sesquilinearform, so ist

$$\|v\| := \sqrt{h(v,v)}$$

eine Norm auf V. Auch hier gilt die Cauchy–Schwarz'sche Ungleichung

$$|h(v,w)| \leq \|v\|\|w\| \ .$$

Wir sagen auch wie in Definition 6.5.5.2, $v \neq 0$ stehe senkrecht auf $w \neq 0$, in Zeichen $v \perp w$, wenn $\langle v, w\rangle = 0$ gilt. Da das Skalarprodukt nun aber komplexe Werte annimmt, führen wir keinen Begriff von Winkeln ein.

Für eine geordnete Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von V führen wir wie in Definition 6.2.5 die darstellende Matrix der Sesquilinearform ein:

$$M_{\mathcal{B}}(h) = (h(b_i, b_j))_{i,j=1\ldots n} \ .$$

Definition 6.7.3

1. *Eine positiv definite, hermitesche Sesquilinearform h auf einem $\mathbb{C}$–Vektorraum V heißt ein (komplexes) Skalarprodukt. Ein Paar (V,h), bestehend aus einem $\mathbb{C}$–Vektorraum V und einem komplexen Skalarprodukt h, heißt unitärer Vektorraum.*

2. *Eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ eines unitären Vektorraums (V,h) heißt Orthonormalbasis, falls gilt*

$$h(b_i, b_j) = \delta_{i,j} \quad \text{für alle} \quad i,j = 1, \ldots, n \ .$$

Jeder endlich–dimensionale unitäre Vektorraum besitzt eine Orthonormalbasis, vgl. Betrachtung 6.5.12.

3. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder ein unitärer Vektorraum. Ein Endomorphismus $\Phi \in \mathrm{End}(V)$ heißt selbstadjungiert, falls gilt*

$$\langle \Phi(v), w\rangle = \langle v, \Phi(w)\rangle \quad \text{für alle} \ \ v, w \in V \ .$$

Betrachtung 6.7.4.
Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine Orthonormalbasis eines unitären Vektorraums V, sei $\Phi \in \mathrm{End}(V)$ und $A := M_{\mathcal{B}}(\Phi) \in M(n \times n, \mathbb{C})$ seine darstellende Matrix. Gelte

$$v = \sum_{i=1}^n x_i b_i \qquad w = \sum_{i=1}^n y_i b_i$$

und seien $x, y \in \mathbb{C}^n$ die zugehörigen Koordinatenvektoren. Wir vergleichen die beiden Ausdrücke

$$\langle \Phi(v), w\rangle = \langle \sum_{i,j=1}^n A_{ij} x_j b_i, \sum_{k=1}^n y_k b_k\rangle \overset{\text{ONB}}{=} (Ax)^t \cdot \bar{y} = x^t A^t \bar{y}$$

$$\text{und} \quad \langle v, \Phi(w)\rangle = x^t \overline{Ay} = x^t \bar{A} \bar{y} \ .$$

Somit ist der Endomorphismus Φ genau dann selbstadjungiert, wenn für die darstellende Matrix bezüglich einer Orthonormalbasis $A^t = \bar{A}$ gilt. Hierbei bezeichnet $\overline{A}$ die Matrix, die aus A hervorgeht, indem man alle Einträge komplex konjugiert.

Definition 6.7.5

1. *Eine Matrix $A \in M(n \times n, \mathbb{C})$ heißt hermitesch, falls $A^t = \bar{A}$ gilt.*
2. *Für $A \in M(n \times n, \mathbb{C})$ nennt man $A^* := \bar{A}^t$ die Adjungierte von A. (Auch die Schreibweise $A^\dagger$ ist gebräuchlich.)*

Bemerkungen 6.7.6.

1. Sei V ein euklidischer Vektorraum und $\Phi \in \mathrm{End}(V)$. Es gilt für jede Orthonormalbasis $\mathcal{B}$ von V:
$$\Phi \text{ selbstadjungiert } \underset{6.5.15.1}{\Leftrightarrow} M_{\mathcal{B}}(\Phi)^t = M_{\mathcal{B}}(\Phi) \Leftrightarrow M_{\mathcal{B}}(\Phi) \text{ symmetrisch .}$$
2. Sei V ein unitärer Vektorraum und $\Phi \in \mathrm{End}(V)$. Es gilt für jede Orthonormalbasis $\mathcal{B}$ von V:
$$\Phi \text{ selbstadjungiert } \underset{6.7.4}{\Leftrightarrow} M_{\mathcal{B}}(\Phi)^* = M_{\mathcal{B}}(\Phi) \Leftrightarrow M_{\mathcal{B}}(\Phi) \text{ hermitesch .}$$

Unser nächstes Ziel ist nun, die Eigenräume von selbstadjungierten Endomorphismen unitärer (und euklidischer) Vektorräume zu verstehen.

Lemma 6.7.7.
Sei $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum und $\Phi \in \mathrm{End}(V)$ selbstadjungiert. Dann sind alle Eigenwerte von Φ reell.

Beweis.
Sei $v \in \mathrm{Eig}(\Phi, \lambda)$ und $v \neq 0$. Dann gilt wegen der Linearität im ersten Argument:
$$\langle \Phi(v), v \rangle = \langle \lambda v, v \rangle = \lambda \|v\|^2 .$$
Da Φ selbstadjungiert ist, ist dies wegen der Antilinearität im zweiten Argument gleich
$$\langle v, \Phi(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \|v\|^2 .$$
Also gilt $\bar{\lambda} = \lambda$. □

Lemma 6.7.8.
Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer oder ein unitärer Vektorraum und sei $\Phi \in \mathrm{End}(V)$ selbstadjungiert. Sind $\lambda \neq \mu$ zwei verschiedene Eigenwerte von Φ, so sind die Eigenräume orthogonal:
$$\mathrm{Eig}(\Phi, \lambda) \perp \mathrm{Eig}(\Phi, \mu)$$

Beweis.
Sei $v \in \mathrm{Eig}(\Phi, \lambda)$, $w \in \mathrm{Eig}(\Phi, \mu)$ mit $v, w \neq 0$. Dann gilt, da nach Lemma 6.7.7 der Eigenwert μ reell ist,
$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle \Phi(v), w \rangle \overset{\Phi \text{s.a.}}{=} \langle v, \Phi(w) \rangle = \langle v, \mu w \rangle = \mu \langle v, w \rangle ,$$
und somit $(\lambda - \mu)\langle v, w \rangle = 0$, also $v \perp w$. □

Satz 6.7.9.
Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich–dimensionaler euklidischer oder unitärer Vektorraum und sei $\Phi \in \mathrm{End}(V)$ selbstadjungiert. Dann besitzt V eine Orthonormalbasis, die aus Eigenvektoren von Φ besteht. Insbesondere ist Φ diagonalisierbar und es gilt die orthogonale Zerlegung

$$V = \mathrm{Eig}(\Phi, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(\Phi, \lambda_k) \quad \text{mit } \lambda_k \in \mathbb{R}\,.$$

Beweis.

- Sei zunächst $(V, \langle\cdot,\cdot\rangle)$ unitär. Nach dem Fundamentalsatz der Algebra 5.2.18 zerfällt das charakteristische Polynom P_Φ als komplexes Polynom vollständig in Linearfaktoren,

  $$P_\Phi(X) = (X - \lambda_1)\ldots(X - \lambda_n)\,.$$

  mit Eigenwerten $\lambda_i \in \mathbb{R}$ nach Lemma 6.7.7. Sei v_1' ein Eigenvektor zum Eigenwert λ_1; dann setze

  $$v_1 := \frac{v_1'}{\|v_1'\|}$$

  Setze $W := \mathrm{span}_{\mathbb{C}}(v_1)^\perp \subset V$. Der Untervektorraum W ist Φ–invariant: Sei $w \in W$, d.h. $\langle w, v_1\rangle = 0$. Es folgt $\langle\Phi(w), v_1\rangle = \langle w, \Phi(v_1)\rangle = \bar\lambda\langle w, v_1\rangle = 0$, also $\Phi(w) \in W$, vgl. Lemma 6.6.7 für euklidische Vektorräume.
  Offenbar ist auch $\Phi|_W$ ein selbstadjungierter Endomorphismus. Die Aussage folgt nun mit vollständiger Induktion nach $\dim_{\mathbb{C}} V$ wie im Beweis von Satz 6.6.9.

- Sei nun $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum. Es genügt zu zeigen, dass das charakteristische Polynom P_Φ schon über $\mathbb{R}$ in Linearfaktoren zerfällt. Dann können wir wie im komplexen Fall weiter schließen. Sei dazu $\mathcal{B}$ irgendeine Orthonormalbasis von V. Sei $A = M_{\mathcal{B}}(\Phi) \in M(n \times n, \mathbb{R})$. Da A selbstadjungiert und $\mathcal{B}$ eine Orthonormalbasis ist, ist A symmetrisch, $A^t = A$.
  Fasst man A als komplexe Matrix auf, also als Element in $M(n \times n, \mathbb{C})$, so ist $A^* = A$, d.h. A ist hermitesch. Alle Eigenwerte sind nach Lemma 6.7.7 reell, und nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom P_A vollständig über $\mathbb{C}$,

  $$P_A(X) = (X - \lambda_1)\ldots(X - \lambda_n)$$

  mit $\lambda_i \in \mathbb{R}$, also zerfällt P_A schon über $\mathbb{R}$.

□

Bemerkungen 6.7.10.

- Wir verwenden Satz 6.7.9, um symmetrische Matrizen besser zu verstehen. Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Wir können A als darstellende Matrix eines Endomorphismus $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ auffassen. Versieht man den $\mathbb{R}^n$ mit dem euklidischen Standardskalarprodukt, so ist dieser Endomorphismus nach Bemerkung 6.7.7 selbstadjungiert. Es gibt daher nach Satz 6.7.9 eine Basis $(v_1, \ldots, v_n)$ des $\mathbb{R}^n$, die bezüglich des euklidischen Standardskalarprodukts sogar eine Orthonormalbasis ist und die aus Eigenvektoren zu den Eigenwerten λ_i besteht. Also gilt $Av_i = \lambda_i v_i$. Insbesondere sind symmetrische Matrizen diagonalisierbar.

- Wir machen nun eine Bemerkung zu quadratischen Formen über dem Körper $\mathbb{R}$. Sei β eine symmetrische Bilinearform auf einem reellen n-dimensionalen Vektorraum V; ihre darstellende Matrix $A \in M(n \times n, \mathbb{R})$ bezüglich jeder Basis ist symmetrisch. Wegen des ersten Punkts können wir eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V aus Eigenvektoren finden. Wir finden für die Einträge der darstellenden Matrix $M_{\mathcal{B}}(\beta)$

  $$\beta(v_i, v_j) = v_j^t A v_i = \lambda_i v_j^t v_i = \lambda_i \delta_{ij}$$

  Durch Reskalierung, also Übergang zur Basis

  $$v_i' := v_i \text{ falls } \lambda_i = 0 \quad v_i' := \frac{1}{\sqrt{|\lambda_i|}} v_i \text{ falls } \lambda_i \neq 0$$

  erhalten wir eine Basis, in der die darstellende Matrix diagonal ist mit Diagonalelementen in $\{0, \pm 1\}$, also bis auf Permutation von der Form im Sylvesterschen Trägheitssatz 6.4.9 ist.

- Wir sehen also: n_+ ist gleich der Zahl der positiven Eigenwerte der symmetrischen Matrix A, n_- der negativen Eigenwerte von A, mit Vielfachheiten gezählt, und n_0 ist die Dimension des Kerns von A. Die quadratische Form ist genau dann nicht-ausgeartet, wenn $\det A \neq 0$ gilt. Für eine positiv definite Form ist $\det A > 0$ notwendig, aber nicht hinreichend, wie das Beispiel der Diagonalmatrix mit Einträgen $-1, -1$ zeigt.

Sei $(V, \langle \cdot, \cdot \rangle)$ von nun an ein endlich–dimensionaler unitärer Vektorraum.

Definition 6.7.11
Ein Endomorphismus $\Phi \in \mathrm{End}(V)$ heißt unitär, falls für alle $v, w \in V$ gilt $\langle \Phi(v), \Phi(w) \rangle = \langle v, w \rangle$.

Bemerkung 6.7.12.
Wie im euklidischen Fall (vergleiche Lemma 6.6.2) zeigt man für einen unitären Endomorphismus Φ

1. $\|\Phi(v)\| = \|v\| \quad \forall v \in V$
2. $v \perp w \Rightarrow \Phi(v) \perp \Phi(w)$
3. Φ ist unitär $\Rightarrow \Phi$ ist Automorphismus und Φ^{-1} ist unitär.
4. Es gilt $|\det(\Phi)| = 1$.
5. Mit Φ und Ψ ist auch $\Phi \circ \Psi$ ein unitärer Endomorphismus.
6. Die Eigenwerte $\lambda \in \mathbb{C}$ von Φ haben Betrag 1, also $|\lambda| = 1$.
7. Ist $\mathcal{B}$ eine Orthonormalbasis von V und $A = M_{\mathcal{B}}(\Phi)$, so ist der Endomorphismus Φ genau dann unitär, wenn für seine darstellende Matrix $A^* A = E_n$ gilt. Für unitäre Matrizen gilt also $A^* = A^{-1}$. Es folgt dann

   $$\det A^* = \det(\overline{A})^t = \det \overline{A} = \overline{\det A}$$

   und somit aus $A^* A = E_n$, dass

   $$1 = \det E_n = \det(A A^*) = \det A \cdot \overline{\det A} = |\det A|^2 \ .$$

Definition 6.7.13
Sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum.

1. *Die Menge*
$$U(V) := \{\Phi \in \mathrm{End}(V) \mid \Phi \text{ unitär } \}$$
ist eine Gruppe und heißt unitäre Gruppe von V.
$$SU(V) := \{\Phi \in \mathrm{End}(V) \mid \Phi \text{ unitär und } \det\Phi = 1\}$$
heißt spezielle unitäre Gruppe von V.
2. *$U(n) := \{A \in M(n\times n, \mathbb{C}) \mid A^*A = E_n\}$ bzw. $SU(n) := \{A \in M(n\times n, \mathbb{C}) \mid A^*A = E_n$ und $\det A = 1\}$ heißt (spezielle) unitäre Gruppe.*

Beispiel 6.7.14.

$$U(1) = \{A \in M(1\times 1, \mathbb{C}) \mid A^*A = 1\} = \{(a), a \in \mathbb{C} \text{ mit } |a| = 1\}$$
ist sogar abelsch. $SU(1)$ ist die triviale Gruppe, vgl. Beispiel 6.6.6.1 für den Fall orthogonaler Gruppen.

Satz 6.7.15.
Sei $\Phi \in U(V)$. Dann existiert eine Orthonormalbasis von V, die aus Eigenvektoren von Φ besteht.

Insbesondere ist jeder unitäre Endomorphismus Φ eines endlich-dimensionalen unitären Vektorraums diagonalisierbar. Jede unitäre Matrix ist also ähnlich zu einer Diagonalmatrix $D = (a_1, \ldots, a_n)$, deren Einträge komplexe Zahlen vom Betrag 1 sind, $|a_i| = 1$, also $A = SDS^{-1}$. Da die diagonalisierende Basis eine Orthonormalbasis ist, ist S unitär, also gilt auch $A = SDS^*$.

Beweis.
Nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom P_Φ über $\mathbb{C}$. Wir wählen einen Eigenvektor v_1 von Φ zum Eigenwert λ_1, und wie im Beweis von Satz 6.7.9 kommt es nur noch darauf an nachzuweisen, dass das orthogonale Komplemente $W := \mathrm{span}_{\mathbb{C}}(v_1)^\perp$ ein Φ–invarianter Untervektorraum ist. Es sei also $w \in W$, d.h. es gilt $\langle w, v_1\rangle = 0$. Es ist $|\lambda_1| = 1$, also $\lambda_1 \neq 0$, somit

$$\langle\Phi(w), v_1\rangle = \langle\Phi(w), \Phi(\Phi^{-1}v_1)\rangle \overset{\Phi \text{ unitär}}{=} \langle w, \Phi^{-1}(v_1)\rangle = \langle w, \frac{1}{\lambda_1} v_1\rangle = \frac{1}{\overline{\lambda_1}}\langle w, v_1\rangle = 0\ .$$

Also gilt $\Phi(w) \in W$. □

7 Allgemeine Klassifikation von Endomorphismen

7.1 Charakteristische Matrizen

Wir wollen nun endlich die Ähnlichkeitsklassen quadratischer Matrizen beschreiben. Wir wissen schon: jede *symmetrische* und jede hermitsche Matrix (vg. Satz 6.7.9), aber auch jede unitäre Matrix (vgl. Satz 6.7.15) $A \in M(n \times n, \mathbb{R})$ ist ähnlich zu einer Diagonalmatrix. Daraus folgt, dass diese Matrizen genau dann ähnlich sind, in Zeichen $A \simeq B$, wenn sie das gleiche charakteristische Polynom haben, $P_A = P_B$. Für nicht-symmetrische Matrizen kann dies nicht gelten, vgl. Betrachtung 5.5.10. Um Ähnlichkeitsklassen quadratischer Matrizen zu klassifizieren, brauchen wir also eine feinere Größe als das charakteristische Polynom. Dieses war als Determinante einer Matrix definiert. Die Idee ist, nicht nur die Determinante, sondern die Matrix selbst zu betrachten und aus ihr weitere Invarianten für die Ähnlichkeitsklassen zu gewinnen.

Wir führen diese Betrachtungen gleich für einen beliebigen Körper K durch.

Definition 7.1.1
Sei $A \in M(n \times n, K)$. Dann heißt die Matrix

$$M_A(X) := XE_n - A \in M(n \times n, K[X])$$

die charakteristische Matrix von A.

Man beachte, dass die Einträge der charakteristischen Matrix Elemente eines kommutativen Rings, nämlich des Polynomrings $K[X]$, sind.

Satz 7.1.2 (Frobenius)**.**
Seien $A, B \in M(n \times n, K)$. Dann sind A und B genau dann ähnlich (über dem Körper K), wenn die zugehörigen charakteristischen Matrizen $M_A(X)$ und $M_B(X)$ äquivalent (über dem Polynomring $K[X]$) sind. In Zeichen:

$$A \simeq B \quad \text{über } K \Leftrightarrow M_A(X) \sim M_B(X) \quad \text{über } K[X]$$

Beweis.

- Ist $B = SAS^{-1}$ mit $S \in GL(n, K)$, so folgt

$$M_B(X) = XE_n - B = XSE_nS^{-1} - SAS^{-1} = SM_A(X)S^{-1} ;$$

  und natürlich ist S auch im Matrizenring $M(n \times n, K[X])$ invertierbar.

- Sei umgekehrt $M_A(X)$ äquivalent zu $M_B(X)$. Dann gibt es invertible Matrizen

$$P(X) = \sum_{i=0}^{m} X^i P_i \qquad \text{mit} \quad P_i \in M(n \times n, K)$$

$$Q(X) = \sum_{i=0}^{m} X^i Q_i \qquad \text{mit} \quad Q_i \in M(n \times n, K)$$

  so dass gilt

$$P(X)(XE_n - B) = (XE_n - A)Q(X) .$$

  Hieraus folgt durch Vergleich von Termen des gleichen Grades $1 \leq i \leq m$ in X

$$P_{i-1} - P_iB = Q_{i-1} - AQ_i \quad \text{für } 1 \leq i \leq m \qquad (1)$$

sowie in Grad 0 und $m+1$

$$P_0 B = AQ_0 \quad \text{und } P_m = Q_m \tag{2}$$

Wir multiplizieren (1) von links mit A^i und summieren über i:

$$\begin{aligned}\sum_{i=1}^{m} A^i P_{i-1} - \sum_{i=1}^{m} A^i P_i B &= \sum_{i=1}^{m} A^i Q_{i-1} - \sum_{i=1}^{m} A^{i+1} Q_i \\ &= AQ_0 - A^{m+1} Q_m \\ &= P_0 B - A^{m+1} P_m \qquad \text{wegen (2)} .\end{aligned}$$

Also gilt

$$A\left(\sum_{i=0}^{m} A^i P_i\right) - \left(\sum_{i=0}^{m} A^i P_i\right) B = 0 . \tag{3}$$

- Es kommt also darauf an, zu zeigen, dass die Matrix

$$S := \sum_{i=0}^{m} A^i P_i \in M(n \times n, K) \tag{4}$$

invertibel ist. Nach Voraussetzung ist $P(X)$ in $M(n \times n, K[X])$ invertibel. Also gibt es eine $n \times n$–Matrix mit Einträgen in $K[X]$,

$$R(X) = \sum_{i=0}^{m} X^i R_i \in M(n \times n, K[X]) \ ,$$

mit

$$P(X)R(X) = E_n X^0 = E_n . \tag{5}$$

- Wir behaupten, dass die quadratische Matrix

$$T := \sum_{j=0}^{m} B^j R_j \in M(n \times n, K) \tag{6}$$

mit Einträgen in K die Gleichung $ST = E_n$ erfüllt und somit das Inverse von S ist. Wir rechnen dies nach:

$$ST \underset{(6)}{=} \sum_{j=0}^{m} SB^j R_j \underset{(3)}{=} \sum_{j=0}^{m} A^j S R_j \underset{(4)}{=} \sum_{i,j=0}^{m} A^{i+j} P_i R_j \underset{(5)}{=} A^0 E_n = E_n .$$

□

Wir müssen also das Äquivalenzproblem für Matrizen mit Einträgen im Polynomring $K[X]$ untersuchen. Für Matrizen mit Einträgen in Körpern hatten wir dieses Problem in Korollar 3.6.12 gelöst. Hier müssen wir etwas mehr über die Struktur des Polynomrings wissen.

Der Polynomring $K[X]$ trägt die Struktur eines sogenannten euklidischen Rings:

Definition 7.1.3

Ein kommutativer Integritätsring R heißt euklidischer Ring, falls eine Abbildung

$$\nu : \ R \to \mathbb{N}_0$$

existiert mit $\nu(0) = 0$, so dass zu allen $a, b \in R$ mit $a \neq 0$ Elemente $q, r \in R$ existieren mit

$$b = qa + r \qquad \text{und} \quad \nu(r) < \nu(a) .$$

Die Abbildung ν heißt euklidische Normfunktion.

Beispiele 7.1.4.

1. Der Ring $\mathbb{Z}$ mit $\nu(a) = |a|$, die Division ist die Division mit Rest von ganzen Zahlen.

2. Ist K ein Körper, so ist der Polynomring $R = K[X]$ mit der Funktion $\nu(g) = 1 + \operatorname{grad}(g)$ für $g \neq 0$ nach Satz 5.2.10 ein euklidischer Ring. Die Division ist die Division mit Rest von Polynomen.

Betrachtung 7.1.5.

1. Wie in Beweis von Satz 5.5.4 folgt, dass alle Ideale eines euklidischen Rings R von der Form Ra mit $a \in R$ sind. Kommutative Ringe mit dieser Eigenschaft nennt man Hauptidealringe. Ein Ideal der Form Ra nennt man ein Hauptideal.

2. Sei R ein Ring mit Eins und $a, b \in R$. Gibt es ein $q \in R$ mit $b = qa$, so sagen wir, a teile b oder b sei ein Vielfaches von a, in Zeichen $a|b$. Man beachte, dass für die zugehörigen Hauptideale gilt:
$$a|b \Leftrightarrow Rb \subset Ra\,.$$

3. Seien $a_i \in R$ eine Familie von Elementen von R. Dann heißt $t \in R$ gemeinsamer Teiler der Elemente a_i falls $t|a_i$ für alle i gilt. Für die Hauptideale heißt dies
$$Ra_i \subset Rt \quad \text{für alle } i\ .$$
Ein gemeinsamer Teiler heißt größter gemeinsamer Teiler, falls er von jedem gemeinsamen Teiler geteilt wird.

4. Gegeben zwei Ideale J_1 und J_2 ist die Summe
$$J_1 + J_2 := \{x_1 + x_2 \mid x_i \in J_i\}$$
das kleinste Ideal, das J_1 und J_2 umfasst. Ist insbesondere R ein Hauptidealring, so ist für endlich viele Elemente $a_1, a_2, \ldots, a_n \in R$ das Ideal $Ra_1 + Ra_2 + \ldots + Ra_n$ ein Hauptideal, also
$$Ra_1 + Ra_2 + \ldots + Ra_n = Rd$$
für ein gewisses $d \in R$. Das Element d ist wegen $Ra_i \subset Rd$ gemeinsamer Teiler aller a_i. Da die Summe das kleinste Ideal ist, das alle Hauptideale Ra_i umfasst, ist d größter gemeinsamer Teiler.

5. In Hauptidealringen existieren also für je endlich viele Elemente größte gemeinsame Teiler (und kleinste gemeinsame Vielfache).

6. Ein Element $p \in R$, das nicht in R invertibel ist, heißt irreduzibel, wenn aus
$$p = ab$$
folgt, dass a oder b invertibel ist.

   Beispiele

   - Im Ring $R = \mathbb{Z}$ der ganzen Zahlen sind die irreduziblen Elemente die Primzahlen $\pm p$.
   - Im Polynomring $R = \mathbb{C}[X]$ sind die irreduziblen Elemente als Folge des Fundamentalsatzes der Algebra 5.2.18 die linearen Polynome $X - a$ mit $a \in \mathbb{C}$.

- Im Polynomring $R = \mathbb{R}[X]$ sind die irreduziblen Elemente die linearen Polynome $X - a$ und die quadratischen Polynome $aX^2 + bX + c$ mit $b^2 - 4ac < 0$.

7. In einem Hauptidealring kann man jedes Element $a \in R$ in der Form

$$a = \varepsilon p_1 \dots p_r \quad \text{mit } \varepsilon \text{ invertibel und } p_i \text{ irreduzibel}$$

schreiben. Gilt

$$a = \varepsilon' p_i' \dots p_s' \,,$$

so folgt $r = s$ und nach Umnummerierung gilt

$$p_i' = \varepsilon_i' p_i \qquad \text{mit } \varepsilon_i \in R \text{ invertibel} \,.$$

Da in einem Hauptidealring außerdem gilt, dass die irreduziblen Elemente genau diejenigen sind, für die

$$p|ab \Rightarrow p|a \text{ oder } p|b \,,$$

d.h. dass die irreduziblen Elemente gerade die Primelemente sind, haben wir in Hauptidealringen und somit euklidischen Ringen, die von $\mathbb{Z}$ vertraute Eindeutigkeit der Primzerlegung. (Einen Beweis dieser Tatsachen und mehr Details lernen Sie in der Vorlesung Algebra.) Wir nennen daher die irreduziblen Polynome in $K[X]$ auch Primpolynome.

7.2 Der Invariantenteilersatz

Der folgende Satz ist zentral; er verallgemeinert Bemerkung 3.6.8.3, die eine Normalform von Matrizen, die Einträge in Körpern haben, vorstellt auf Matrizen, deren Einträge Elemente eines euklidischen Rings sind.

Satz 7.2.1.
Sei R ein euklidischer Ring. Eine Matrix $A \in M(n \times n, R)$ lässt sich durch wiederholte Anwendung der elementaren Zeilen– *und* Spaltenumformungen aus Satz 3.4.10 vom

Typ 2) Addition des a–fachen, $a \in R$, einer Zeile/Spalte zu einer anderen Zeile/Spalte

Typ 3) Vertauschen zweier Zeilen bzw. Spalten

stets in eine Diagonalmatrix überführen

$$\begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix}$$

für deren Diagonalelemente überdies die Teilbarkeitsbedingungen $c_i | c_{i+1}$ für $i = 1, 2, \dots, n-1$ gelten.

Beweis.

- Wir dürfen $A \neq 0$ voraussetzen und können durch Vertauschungen von Zeilen und Spalten erreichen, dass

$$a_{11} \neq 0 \quad \text{und } \nu(a_{11}) \leq \nu(a_{ij}) \quad \text{für alle } i, j \text{ mit } a_{ij} \neq 0 \,.$$

Wir wollen alle Einträge a_{1i} und a_{i1} mit $i \neq 1$, also alle Einträge der ersten Zeile und der ersten Spalte außer dem Diagonalelement, durch Spalten- und Zeilenumformungen zum Verschwinden bringen.

- Sei etwa $a_{21} \neq 0$. Da der Ring R euklidisch ist, finde durch Division mit Rest von a_{21} durch $a_{11} \neq 0$ ein Element $q \in R$ mit

$$\nu(a_{21} - qa_{11}) < \nu(a_{11})$$

Addiere das $-q$–fache der 1.Zeile zur zweiten Zeile und finde eine Matrix A' mit

$$\nu(a'_{21}) < \nu(a_{11})$$

Sollte $a'_{21} \neq 0$ sein, so vertausche wieder Zeilen und Spalten und erhalte eine Matrix A'' mit

$$a''_{11} \neq 0 \quad \nu(a''_{11}) \leq \nu(a''_{ij}) \text{ für alle } a''_{ij} \neq 0 \ .$$

Da a'_{21} unter den Elementen von A'' ist, und $\nu(a'_{21}) < \nu(a_{11})$ gilt, folgt sicher $\nu(a''_{11}) < \nu(a_{11})$. Solange also ein Außerdiagonalelement in der ersten Zeile oder Spalte nicht verschwindet, kann man die euklidische Norm des ersten Diagonalemements verkleinern. Diese ist aber nach unten beschränkt; daher muss das Verfahren dadurch abbrechen, dass nach endlich vielen Schritten alle Elemente der ersten Zeile und ersten Spalte außer dem Diagonalelement verschwinden.

Nach endlich vielen Schritten - auch für Spalten - finde eine Matrix der Gestalt

$$D = \left(\begin{array}{c|ccc} d_{11} & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right) = (d_{ij})$$

mit $\nu(d_{11}) \leq \nu(a_{11}),\ d_{11} \neq 0$.

- Wir können außerdem noch $d_{11}|d_{ij}$ für alle i, j einrichten. Sei etwa das Matrixelement d_{ij} nicht durch d_{11} teilbar. Dann finde durch Division mit Rest von d_{ij} durch d_{11} ein Element $q \in R$ mit

$$\nu(d_{ij} - qd_{11}) < \nu(d_{11}) \quad \text{aber } \ d_{ij} - qd_{11} \neq 0 \ .$$

  - Addiere die erste Zeile von D zur i–ten Zeile.
  - Subtrahiere das q–fache der ersten Spalte der so entstandenen Matrix von der j–ten Spalte der Matrix und erhalte eine Matrix D' mit

$$d'_{ij} = d_{ij} - qd_{11} \ .$$

    Es ist

$$d'_{ij} \neq 0 \quad \text{und } \nu\left(d'_{ij}\right) < \nu\left(d_{ij}\right) \leq \nu\left(a_{11}\right) \ .$$

    Wende auf die nun erhaltene Matrix die gesamte bisherige Prozedur an. Dieser Prozess bricht irgendwann einmal ab. Dann erreicht man $d_{11}|d_{ij}$ für alle i, j.

- Induktiv folgt die Behauptung des Satzes.

□

Bemerkungen 7.2.2.

1. Sei R ein euklidischer Ring. Dann ist jede $n \times n$–Matrix C über R äquivalent zu einer Diagonalmatrix
$$\begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix}$$
mit Diagonaleinträgen, die die Bedingung $c_i | c_{i+1}$ für $i = 1, \ldots, n-1$ erfüllen. Denn Zeilen– bzw. Spaltentransformationen von Typ 2 und 3 können auch für Matrizen mit Einträgen in unitären Ringen wie in Lemma 3.4.9 durch die Multiplikation von links bzw. rechts mit invertiblen Elementarmatrizen, also durch Äquivalenzumformungen, realisiert werden. Man vergleiche dies mit Bemerkung 3.6.8.3, wo die Situation über Körpern behandelt wird.

2. Wir werden in Betrachtung 7.2.5 sehen, dass die Einträge c_i bis auf Multiplikation mit invertiblen Elementen (sog. Einheiten) eindeutig sind. Die invertiblen Elemente eines Polynomrings $K[X]$ über einem Körper sind aber gerade die nicht-verschwindenden konstanten Polynome.
Im Falle von Polynomringen, $R = K[X]$, kann man die Polynome c_i also dadurch eindeutig festlegen, dass man fordert, dass sie normierte Polynome sind. Man nennt das normierte Polynom $c_j \in K[X]$ den j–ten Invariantenteiler der Matrix $C = C(X) \in M(n \times n, K[X])$.

Von nun an beschränken wir uns auf den Fall $R = K[X]$. Die Verallgemeinerung auf allgemeine euklidische Ringe ist offensichtlich.

Definition 7.2.3
Sei $F \in M(n \times n, K[X])$. Wir bezeichnen mit
$$d_j(F) \in K[X]$$
den größten gemeinsamen normierten Teiler aller Unterdeterminanten j–ter Ordnung der Matrix F. Das normierte Polynom $d_j(F)$ heißt der j–te Determinantenteiler von F über $K[X]$.

Zum Beispiel ist der erste Determinantenteiler
$$d_1(F) = ggT(F_{11}(X) \ldots F_{nn}(X))$$
der ggT der Matrixelemente und das normierte Polynom $d_n(F)$ unterscheidet sich von $\det(F) \in K[X]$ höchstens durch die Multplikation mit einem invertiblen Element.

Lemma 7.2.4.
Äquivalente Matrizen $F_1, F_2 \in M(n \times n, K[X])$ haben gleiche Determinantenteiler.

Beweis.
Es reicht es aus zu zeigen, dass die Determinantenteiler von F_1 die von F_2 teilen. Aus dem gleichen Grund teilen dann auch die Determinantenteiler von F_2 die von F_1. Da Determinantenteiler normierte Polynome sind, folgt Gleichheit.

Für jede Matrix $P \in M(n \times n, K[X])$ sind die Zeilen der Matrix PF_1 Linearkombinationen der Zeilen von F_1. Die Minoren j-ter Ordnung von PF_1 sind daher Linearkombinationen der Minoren j-ter Ordnung von F_1. Für jede Matrix $Q \in M(n \times n, K[X])$ sind die Spalten von

$(PF_1)Q$ Linearkombinationen der Spalten von QF_1. Also sind die Minoren j-ter Ordnung von $F_2 = PF_1Q$ Linearkombinationen der Minoren j-ter Ordnung von PF_1. Die Minoren j-ter Ordnung von F_2 sind somit Linearkombinationen der Minoren j-ter Ordnung von F_1. Gemeinsame Teiler aller Minoren j-ter Ordnung von F_1 sind daher auch Teiler der Minoren j-ter Ordnung von F_2, woraus die Behauptung folgt. □

Betrachtung 7.2.5.
Für eine Matrix aus $M(n \times n, K[X])$ der Diagonalgestalt

$$\begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix}$$

mit den Teilbarkeitsbedingungen $c_i | c_{i+1}$ für $i = 1, \ldots, n-1$ aus Bemerkung 7.2.2.1 gilt für die Determinantenteiler

$$d_j(C) = c_1(C) \ldots c_j(C) \ ,$$

wobei wir auch die Invariantenteiler $c_i \in K[X]$ als normiert voraussetzen. Wegen

$$d_1(C) = c_1(C) \qquad \text{und} \quad d_{j+1}(C) = c_{j+1}(C) d_j(C)$$

legen die Invariantenteiler und die Determinantenteiler sich gegenseitig eindeutig fest. Also sind auch die Invariantenteiler in Satz 7.2.1 eindeutig.

Definition 7.2.6
Sei $A \in M(n \times n, K)$ eine quadratische Matrix und $M_A(X) \in M(n \times n, K[X])$ ihre charakteristische Matrix. Dann nennen wir die normierten Polynome

$$d_A^{(j)} = d_j\left(M_A(X)\right) \ , \quad c_A^{(j)} = c_j\left(M_A(X)\right) \in K[X]$$

die Determinanten– bzw. Invariantenteiler der Matrix $A \in M(n \times n, K)$. Es gilt $d_A^{(j)} = c_A^{(1)} \ldots c_A^{(j)}$ und $d_A^{(n)} = \det M_A(X) = P_A(X)$ ist gleich dem charakteristischen Polynom.

Satz 7.2.7.
Zwei Matrizen $A, B \in M(n \times n, K)$ über einem Körper K sind genau dann ähnlich, wenn sie dieselben Invariantenteiler (bzw. dieselben Determinantenteiler) besitzen.

Damit haben wir einen vollständigen Satz von Invarianten für das Ähnlichkeitsproblem für quadratische Matrizen gefunden.
Beweis.

- Wegen Satz 7.1.2 sind A und B genau dann ähnlich, wenn die charakteristischen Matrizen $M_A(X)$ und $M_B(X)$ äquivalent sind.
- Wegen Satz 7.2.1 ist jede der beiden charakteristischen Matrizen $M_A(X)$ und $M_B(X)$ äquivalent zu einer Diagonalmatrix, auf deren Diagonale die Invariantenteiler stehen. Diese sind aber eindeutig. Also sind die charakteristischen Matrizen $M_A(X)$ und $M_B(X)$ genau dann äquivalent, wenn die Invariantenteiler übereinstimmen. Aus Betrachtung 7.2.5 folgt, dass dies auch genau dann der Fall ist, wenn die Determinantenteiler übereinstimmen.

□

Beispiel 7.2.8.
Betrachte die Matrix

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 2 & 2 & 3 \end{pmatrix} \in M(3 \times 3, \mathbb{R})$$

Durch Vertauschung der ersten und zweiten Spalte sieht man, dass die charakteristische Matrix äquivalent ist zu

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

Wir eliminieren mit Spaltenoperationen die nicht-diagonalen Einträge der erste Zeile $\begin{pmatrix} 1 & 0 & 0 \\ X-2 & 1-(X-2)^2 & X-1 \\ -2 & -2+2(X-2) & X-5 \end{pmatrix}$
und dann mit Zeilenoperationen die nicht-diagonalen Einträge der ersten Spalte

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -(X-1)(X-3) & X-1 \\ 0 & 2(X-3) & X-5 \end{pmatrix}$$

Dann vertauschen wir die zweite und dritte Spalte $\begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & -(X-1)(X-3) \\ 0 & X-5 & 2(X-3) \end{pmatrix}$ und addieren das $(X-3)$–fache der zweiten Spalte zur dritten Spalte:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & X-5 & (X-3)^2 \end{pmatrix}$$

Subtraktion der zweiten Zeile von der dritten mit ausschließender Vertauschung der zweiten und dritten Zeile gibt $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & (X-3)^2 \\ 0 & X-1 & 0 \end{pmatrix}$ Weitere Eliminationen liefern

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & (X-3)^2 \\ 0 & 0 & \frac{1}{4}(X-1)(X-3)^2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \frac{1}{4}(X-1)(X-3)^2 \end{pmatrix} .$$

Wir finden die Invariantenteiler die normierten Polynome

$$c_A^{(1)} = c_A^{(2)} = 1 \qquad c_A^{(3)} = (X-1)(X-3)^2 .$$

und somit die Determinantenteiler

$$d_A^{(1)} = d_A^{(2)} = 1 \qquad d_A^{(3)} = (X-1)(X-3)^2 .$$

Jetzt können wir zum Beispiel einfach die Frage beantworten, ob die Matrix A ähnlich zur Matrix

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$$

ist. Dazu berechnen wir die Determinantenteiler von B: das charakteristische Polynom ist gleich $d_A^{(3)}$; $d_B^{(1)}$ ist der größte gemeinsame Teiler der Einträge der charakteristischen Matrix; da der Eintrag in der dritten Zeile und zweiten Spalte gleich 1 ist, ist dies gleich 1. Schließlich bestimmen wir $d_B^{(2)}$ durch die Betrachtung von Determinanten von 2×2 Streichungsmatrizen, etwa nach Streichung der ersten Zeile und ersten Zeile Spalte $(X-3)^2$ und nach Streichung der zweiten Zeile und dritten Spalte $-(X-1)$, und finden auch hier als größten gemeinsamen Teiler 1:

$$d_B^{(1)} = d_B^{(2)} = 1 \qquad d_B^{(3)} = (X-1)(X-3)^2 \; .$$

Dies sind dieselben Determinantenteiler wie die der Matrix A, also sind nach Satz 7.2.7 die Matrizen A und B ähnlich.

7.3 Normalformen für Matrizen

Wir haben durch die Begriffe der Determinanten- bzw. Invariantenteiler das Ähnlichkeitsproblem quadratischer Matrizen insofern gelöst, als wir vollständige Invarianten für die Ähnlichkeitsklassen angeben können. Man vergleiche dies mit der Situation beim Äquivalenzproblem aus Bemerkung 3.6.8.3: die Invariante war dort der Rang r. Wir konnten dort auch Repräsentanten für die Äquivalenzklassen angeben: eine Diagonalmatrix mit genau r Einsen auf der Diagonale.

Nun wollen wir Repräsentanten für die Ähnlichkeitsklassen angeben: Normalformen. Da die Invarianten nun normierte Polynome sind, müssen wir versuchen, normierten Polynomen Matrizen zuzuordnen. Wir erinnern uns an den Satz 5.5.7 von Cayley-Hamilton:

Definition 7.3.1
Sei

$$g = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0 \in K[X]$$

ein normiertes Polynom über einen Körper K vom Grad $n \geq 1$. Dann heißt die Matrix

$$B_g := \begin{pmatrix} 0 & 0 & & & -a_0 \\ 1 & 0 & & & \vdots \\ & 1 & \ddots & & \\ & & \ddots & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix} \in M(n \times n, K)$$

Begleitmatrix des Polynoms g. Für $n = \operatorname{grad}(g) = 1$ ist $g = X + a_0$. Wir setzen $B_g = (-a_0) \in M(1 \times 1, K)$.

Betrachtung 7.3.2.

1. Beachte, dass aus der Rechnung zum Beweis des Satzes von Cayley-Hamilton 5.5.7 sich das charakteristische Polynom der Begleitmatrix B_g ergibt:

$$P_{B_g}(X) = g(X) \; .$$

2. Streicht man in $M_{B_g}(X)$ die erste Zeile und letzte Spalte, so erhält man eine $(n-1)\times(n-1)$ Matrix mit Determinante $(-1)^{n-1}$. Also sind die Determinantenteiler von B_g die Polynome

$1, 1, \ldots, 1, g$. Aus Betrachtung 7.2.5 folgt sofort, dass dies auch die Invariantenteiler sind. Über dem Polynomring $K[X]$ sind also nach Betrachtung 7.2.7 äquivalent:

$$M_{B_g}(X) = XE_n - B_g \sim \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & g \end{pmatrix}$$

3. Seien allgemeiner $g_1, \ldots, g_r$ normierte Polynome vom Grad Eins oder größer, die der Teilbarkeitsbedingung $g_i | g_{i+1}$ für $1 \leq i \leq r-1$ genügen. Sei B die blockdiagonale Matrix aus Begleitmatrizen:

$$B = B_{g_1,\ldots,g_r} := \begin{pmatrix} B_{g_1} & & & \\ & B_{g_2} & & \\ & & \ddots & \\ & & & B_{g_r} \end{pmatrix} \in M(n \times n, K)$$

Die charakteristische Matrix von $B_{g_1,\ldots,g_r}$ ist äquivalent zu

$$XE_n - B \sim \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & g_1 & & \\ & & & & \ddots & \\ & & & & & g_r \end{pmatrix}.$$

Also besitzt B die Invariantenteiler

$$\underbrace{1, 1, \ldots, 1}_{n-r}, g_1, g_2, \ldots, g_r$$

mit $n := \sum_{i=1}^{r} \operatorname{grad} g_i$.

Satz 7.3.3 (Frobeniussche Normalform)**.**
Sei $A \in M(n \times n, K)$ beliebig. Dann ist A zu genau einer Matrix der Gestalt $B_{g_1,\ldots,g_r}$ ähnlich, wobei $g_1, \ldots, g_r$ normierte Polynome über K vom Grad ≥ 1 sind, für die

$$g_i | g_{i+1} \qquad i = 1, \ldots, i-1$$

gilt.

Beweis.

- Sei $1, \ldots, 1, g_1, \ldots, g_r$ das System der Invariantenteiler von A. Nach Betrachtung 7.3.2 ist dies auch das System der Invariantenteiler von $B_{g_1,\ldots,g_r}$. Aus Satz 7.2.7 folgt, dass A und $B_{g_1,\ldots,g_r}$ ähnlich sind.
- Die Eindeutigkeit der Frobeniusschen Normalform folgt aus der Eindeutigkeit der Invariantenteiler.

□

Invariantenteiler sind im allgemeinen keine Primpolynome, besitzen also eine (eindeutige) Zerlegung in Potenzen von Primpolynomen. (Hierbei muss man, anders als bei der Frobeniusschen Normalform, aufpassen, über welchem Körper man arbeitet: das Polynom X^2+1 ist im Polynomring $\mathbb{R}[X]$ irreduzibel, aber wegen $X^2+1=(X+\mathrm{i})(X-\mathrm{i})$ im Polynomring $\mathbb{C}[X]$ reduzibel.) Wir wollen nun diese Zerlegung ausnützen.

Die Frobeniussche Normalform hat aber im Gegensatz zu den weiteren Normalformen, die wir kennenlernen werden, einige Vorteile:
– die Weierstraß'sche und die Jordansche Normalform sind nur bis auf Wirkungen von Permutationsgruppen definiert. Dies führt bei gewissen Betrachtungen zu kombinatorischen Komplikationen.
– Im Falle reeller oder komplexer Matrizen hängt die Frobeniussche Normalform glatt von der Matrix ab.

Lemma 7.3.4.
Ist $g=h_1\dots h_k$ das Produkt paarweise *teilerfremder* normierter Polynome $h_1\dots h_k$ vom Grade ≥ 1, so gilt für die Begleitmatrizen

$$B_g \simeq \begin{pmatrix} B_{h_1} & & 0 \\ & \ddots & \\ 0 & & B_{h_k} \end{pmatrix} .$$

Beweis.
Die charakteristische Matrix der rechten Seite ist über $K[X]$ nach Betrachtung 7.3.2 äquivalent zu

$$H(X) := \begin{pmatrix} 1 & & & & & \\ & 1 & & & 0 & \\ & & \ddots & & & \\ & & & h_1 & & \\ & 0 & & & \ddots & \\ & & & & & h_k \end{pmatrix} .$$

Wir müssen zeigen, dass $H(X)$ und

$$G(X) := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & g \end{pmatrix}$$

die gleichen Determinantenteiler haben. Die von $G(X)$ sind nach Betrachtung 7.3.2 offensichtlich $1,1,,\dots,1,g$. Sei $n=\mathrm{grad}(g)$. Offenbar gilt:

$$d_n(H(X)) = h_1\dots h_k = g = d_n(G(X)) .$$

Ferner ist

$$d_{n-1}(G(X)) = 1 \ .$$

Bei der Berechnung von $d_{n-1}(H(X))$ tritt für jedes $i=1\dots r$ das Produkt $\prod_{j=1\ j\neq i}^{k} h_j$ als Unterdeterminante auf; als gemeinsamer Teiler aller dieser Produkte teilerfremder Matrizen ist der Determinantenteiler

$$d_{n-1}(H(X)) = 1 \ .$$

□

Satz 7.3.5 (Weierstraßsche Normalform)**.**
Sei $A \in M(n \times n, K)$ beliebig. Dann gibt es ein bis auf Reihenfolge eindeutig bestimmtes System $h_1 \dots h_m$ von Potenzen normierter *Primpolynome* aus $K[X]$, so dass A zu der Matrix

$$B_{h_1,\dots,h_m}$$

ähnlich ist. Die Polynome $h_i \in K[X]$ heißen Weierstraßsche Elementarteiler von A über K.

Beweis.
Seien $g_1 \dots g_r$ die nicht–konstanten Invariantenteiler von A. Dann gilt wegen Satz 7.3.3

$$A \simeq \begin{pmatrix} B_{g_1} & & \\ & \ddots & \\ & & B_{g_r} \end{pmatrix}$$

Jedes der Polynome g_i besitzt im euklidischen Ring $K[X]$ eine Zerlegung

$$g_i = h_1 \dots h_k$$

in Potenzen teilerfremder normierter Primpolynome, die bis auf Reihenfolge eindeutig ist. Wende nun Lemma 7.3.4 an. □

Im Spezialfall $K = \mathbb{C}$ kommen wegen des Fundamentalsatzes der Algebra als Weierstraßsche Elementarteiler nur Polynome der Gestalt $(X - \alpha)^e$ mit $\alpha \in \mathbb{C}$ und $e \in \mathbb{N}$ in Frage. Wir wollen die Begleitmatrizen von Polynomen dieser Form über einem beliebigen Körper untersuchen.

Lemma 7.3.6.
Sei K ein beliebiger Körper und $\alpha \in K$, $e \in \mathbb{N}$. Dann ist die Begleitmatrix des Polynoms $(X - \alpha)^e$ ähnlich zu der $e \times e$–Matrix

$$J(\alpha, e) := \begin{pmatrix} \alpha & & & & \\ 1 & \alpha & & & 0 \\ & 1 & \alpha & & \\ & & & \ddots & \\ 0 & & & 1 & \alpha \end{pmatrix}$$

Eine Matrix der Form $J(\alpha, e)$ heißt eine Jordan-Matrix über K.

Beweis.
Wir vergleichen Determinantenteiler und wenden dann Satz 7.2.7 an. Die charakteristische Matrix von $J(\alpha, e)$ ist

$$XE_e - J(\alpha, e) = \begin{pmatrix} X-\alpha & & & & & \\ -1 & X-\alpha & & & & 0 \\ & -1 & X-\alpha & & & \\ & & & \ddots & & \\ 0 & & & & \ddots & \\ & & & & -1 & X-\alpha \end{pmatrix}$$

Der e–te Determinantenteiler von J ist $(X-\alpha)^e$. Streicht man die erste Zeile und letzte Spalte, so erhält man eine $(e-1) \times (e-1)$–Matrix mit Determinante $(-1)^{e-1}$. Die Matrix $J(\alpha, e)$ hat

also genau die gleichen Determinantenteiler wie die Begleitmatrix $B_{(X-\alpha)^e}$, vgl. Betrachtung 7.3.2.2. □

Insbesondere ist also das charakteristische Polynom von $J(\alpha, e)$ gleich $(X-\alpha)^e$, so dass es nur den Eigenwert α mit algebraische Vielfachheit e gibt. Seine geometrische Vielfachheit ist 1.

Satz 7.3.7. Jordansche Normalform
Es sei $A \in M(n \times n, K)$ und das charakteristische Polynom $P_A(X)$ zerfalle vollständig in Linearfaktoren. (Dies ist für $K = \mathbb{C}$ für jede Matrix der Fall, aber nicht für $K = \mathbb{R}$.) Dann gibt es ein bis auf Reihenfolge eindeutiges System von Jordan-Matrizen $J_1, \ldots, J_m$ über K, so dass A zu der blockdiagonalen Matrix

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & 0 \\ & & \ddots & \\ 0 & & & J_m \end{pmatrix}$$

ähnlich ist.

Beweis.

- Die Elementarteiler von A teilen die Invariantenteiler von A, die das charakteristische Polynom $P_A(X)$ teilen. Also zerfallen auch alle Elementarteiler vollständig in Linearfaktoren, sind also von der Form

$$h_i = (X-\alpha_i)^{e_i} \qquad \text{mit } \alpha_i \in K,\ e_i \in \mathbb{N}\,.$$

- Aus Satz 7.3.5 und Lemma 7.3.6 folgen die Ähnlichkeitsbeziehungen

$$A \simeq \begin{pmatrix} B_{h_1} & & 0 \\ & \ddots & \\ 0 & & B_{h_m} \end{pmatrix} \simeq \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}$$

  mit Jordanmatrizen $J_i = J(\alpha_i, e_i)$.

□

Bemerkungen 7.3.8.
Beispiel für eine Matrix in Jordanscher Normalform:

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

1. Da eine (untere) Dreiecksmatrix vorliegt, haben wir eine Verschärfung von Satz 5.4.6 über die Trigonalisierbarkeit beweisen.

2. Man hätte auch durch eine andere Ähnlichkeitstransformation eine obere Dreiecksmatrix erreichen können. Tatsächlich sind eine Matrix $A \in M(n \times n, K)$ und ihre Transponierte A^t immer ähnlich, da sie die gleichen Determinantenteiler haben (Übungsaufgabe).

3. Offenbar ist das charakteristische Polynom einer Matrix in Jordanscher Normalform

$$P_A(X) = \prod_{i=1}^{m} (X - \alpha_i)^{e_i} \ ;$$

auf der Hauptdiagonale einer Matrix in Jordanscher Normalform stehen also die Eigenwerte von A mit ihrer algebraischen Vielfachheit.

4. Man beachte, dass die Eigenwerte α_i in verschiedenen Blöcken nicht unbedingt verschieden sind.

5. Die geometrische Vielfachheit eines Eigenwerts ist gleich der Zahl der Jordan-Blöcke zu diesem Eigenwert.

6. Seien $\alpha_1, \ldots, \alpha_m$ die verschiedenen Eigenwerte von A. Sei m_i die Größe des größten Jordan-Blocks zum Eigenwert α_i. Betrachte

$$(\alpha E_e - J(\alpha, e))^k = \begin{pmatrix} 0 & 0 & & 0 & 0 \\ -1 & 0 & & & \\ 0 & -1 & & 0 & 0 \\ 0 & 0 & & 0 & 0 \\ .. & .. & ..... & .. & .. \\ 0 & 0 & & 0 & 0 \\ 0 & 0 & & -1 & 0 \end{pmatrix}^k$$

Aus den Regeln der Matrixmultiplikation folgt, dass dies für $k < e$ ungleich Null ist, für $k \geq e$ aber verschwindet. Daher ist das Minimalpolynom gleich

$$\mu_A(X) = \prod_{i=1}^{s} (X - \alpha_i)^{m_i} \ .$$

Insbesondere gilt:
Zerfällt das Minimalpolynom in paarweise verschiedene Linearfaktoren, so gilt $m_i = 1$ für alle Eigenwerte. Dann haben alle Jordan-Blöcke Länge 1, also ist A diagonalisierbar. Dies beweist die Implikation "2.⇒1." aus Satz 5.5.9

7. Indem man eine darstellende Matrix in Jordanscher Normalform als Summe einer Diagonalmatrix mit einer unteren Dreiecksmatrix mit Einträgen gleich Null auf der Diagonale schreibt, sieht man: jeder Endomorphismus $\Phi \in \mathrm{End}(V)$, dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt, lässt als Summe eines diagonalisierbaren Endomorphismus Φ_s und eines nilpotenten Endomorphismus Φ_n schreiben:

$$\Phi = \Phi_s + \Phi_n \ .$$

(Ein Endomorphismus Φ heißt nilpotent, wenn es ein $N \in \mathbb{N}$ gibt, so dass $\Phi^N = 0$ gilt.) Aus der blockdiagonalen Gestalt folgt sofort, dass die beiden Endomorphismen kommutieren:

$$\Phi_s \circ \Phi_n = \Phi_n \circ \Phi_s \ .$$

Man kann sogar Φ_s und Φ_n als Polynome in Φ ausdrücken; beide Polynome haben verschwindendes konstantes Glied.

8. Es gibt auch eine multiplikative Version der Jordan-Zerlegung:
   Sei Φ ein *Auto*morphismus eines endlich-dimensionalen K-Vektorraums, dessen charakteristisches Polynom über K vollständig in Linearfaktoren zerfällt. Dann besitzt Φ eine eindeutig bestimmte Produktzerlegung

$$\Phi = \Phi_s \circ \Phi_u \ ,$$

   wobei Φ_s ein diagonalisierbarer Automorphismus von V ist, Φ_u ein unipotenter Automorphismus ist (d.h. $\Phi - \mathrm{id}_V$ ist nilpotent) und Φ_s und Φ_u kommutieren.

   Um dies zu sehen, betrachte die Jordan-Zerlegung $\Phi = \Phi_s + \Phi_n$ wie oben und setze $\Phi_u := \mathrm{id}_V + \Phi_s^{-1}\Phi_n$, so dass folgt

$$\Phi = \Phi_s \circ (\mathrm{id}_V + \Phi_s^{-1} \circ \Phi_n) = \Phi_s \circ \Phi_u \, .$$

Beispiel 7.3.9.
Sei $A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 3 & 1 \\ -1 & 1 & 0 & 4 \end{pmatrix} \in M(4 \times 4, \mathbb{Q})$. Es ist $P_A(X) = (X-3)^4$. Offenbar hat die Matrix

$$A - 3E_4 = \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix}$$

Rang 1, also ist nach der Dimensionsformel $\mu_{geo}(A,3) \equiv \dim_K \ker(A-3E_4) = 4 - \mathrm{rg}\,(A-3E) = 3$. Die Jordansche Normalform ist

$$\begin{pmatrix} \boxed{3} & & & 0 \\ & \boxed{3} & & \\ & & \boxed{\begin{matrix} 3 & 0 \\ 1 & 3 \end{matrix}} & \\ 0 & & & \end{pmatrix}$$

und $\mu_A(X) = (X-3)^2$ ist das Minimalpolynom. In den ersten beiden Kästchen steht jeweils die Begleitmatrix des linearen Polynoms $X-3$. Nach Lemma 7.3.6 steht im dritten Kästchen eine Matrix, die ähnlich zur Begleitmatrix des Polynoms $(X-3)^2$ ist. Damit sind die Weiertraßschen Elementarteiler die drei Polynome $g_1 := (X-3), g_2 := (X-3), g_3 := (X-3)^2$. Da die Teilbarkeitsbedingung $g_1|g_2|g_3$ gilt, liegt haben wir auch die Invariantenteiler gefunden:

$$c_A^{(1)} = 1 \, , \ \ c_A^{(2)} = (X-3) \, , \ \ c_A^{(3)} = (X-3) \, , \ \ c_A^{(4)} = (X-3)^2$$

und die Determinantenteiler

$$d_A^{(1)} = 1 \, , \ \ d_A^{(2)} = (X-3) \, , \ \ d_A^{(3)} = (X-3)^2 \, , \ \ d_A^{(4)} = (X-3)^4$$

Die Frobeniussche Normalform ist

$$\begin{pmatrix} \boxed{3} & & & 0 \\ & \boxed{3} & & \\ & & \boxed{\begin{matrix} 0 & -9 \\ 1 & 6 \end{matrix}} & \\ 0 & & & \end{pmatrix} \ .$$

Sie ist hier gleich der Weierstraßschen Normalform, da in diesem Beispiel die Invariantenteiler gleich den Elementarteilern sind.

7.4 Normalformen für Endomorphismen, zyklische Teilräume

Die darstellenden Matrizen eines Endomorphismus Φ eines K-Vektorraums V bezüglich verschiedener geordneter Basen sind nach Satz 3.6.8.3 ähnlich. Ähnliche Matrizen haben nach Satz 7.2.7 die gleichen Determinanten- und Invariantenteiler. Daher können wir auch einem Endomorphismus unabhängig von der Wahl einer geordneten Basis Determinanten- und Invariantenteiler zuordnen. Um die bisher dargestelle Theorie für Matrizen auf Endomorphismen zu übertragen, benötigen wir die folgende Begriffsbildung:

Definition 7.4.1
Sei Φ ein Endomorphismus eines K-Vektorraums V. Wir nennen den Vektorraum V zyklisch bezüglich Φ oder Φ-zyklisch, wenn es einen Vektor $v \in V$ gibt, so dass $V = \operatorname{span}_K(v, \Phi(v), \Phi^2(v), \ldots)$ gilt. Jeder solcher Vektor $v \in V$ heißt Φ-zyklischer Vektor von V.

Satz 7.4.2.
Es sei Φ Endomorphismus eines endlich-dimensionalen Vektorraums V. Dann ist V genau dann zyklisch bezüglich Φ, wenn es eine geordnete Basis $\mathcal{B}$ von V gibt, in der die darstellende Matrix von Φ die Gestalt einer Begleitmatrix

$$\begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & -a_{n-2} \\ & & 1 & -a_{n-1} \end{pmatrix}$$

hat. Es ist dann

$$P_\Phi(X) = a_0 + a_1 X + \cdots + a_{n-1}X^{n-1} + X^n$$

das charakteristische Polynom von Φ.

Beweis.

- Sei V ein Φ-zyklischer Vektorraum und $v \in V$ ein Φ-zyklischer Vektor. Dann gibt es eine natürliche Zahl n, so dass die Familie

  $$\mathcal{B}' := \big(v, \Phi(v), \Phi^2(v), \ldots, \Phi^{n-1}(v)\big)$$

  linear unabhängig, die Familie

  $$\mathcal{B}'' := \big(v, \Phi(v), \Phi^2(v), \ldots, \Phi^{n}(v)\big)$$

  aber linear abhängig ist. Es gibt also Koeffizienten $a_i \in K$ mit

  $$a_0 v + a_1\Phi(v) + \cdots + a_{n-1}\Phi^{n-1}(v) + \Phi^n(v) = 0\,.$$

  Offenbar ist dann die Familie $\mathcal{B}'$ eine Basis eines Φ-zyklischen Unterraums von V, der v enthält. Da der Vektor v zyklisch sein soll, haben wir eine Basis von V gefunden. Man rechnet nach, dass die darstellende Matrix von Φ bezüglich dieser Basis die oben angegebene Gestalt hat.

- Das charakteristische Polynom berechnet man nun wie im Beweis des Satzes 5.5.7 von Cayley-Hamilton.
- Hat umgekehrt die darstellende Matrix in einer geordneten Basis $\mathcal{B}$ die angegebene Form, so sieht man sofort, dass der erste Basisvektor ein zyklischer Vektor ist.

□

Korollar 7.4.3.
Sei Φ ein Endomorphismus eines K-Vektorraums V, und sei V zyklisch bezüglich Φ. Dann stimmen Minimalpolynom und charakteristisches Polynom von Φ überein.

Beweis.
Aus der linearen Unabhängigkeit der Familie

$$\mathcal{B}' := \big(v, \Phi(v), \Phi^2(v), \ldots, \Phi^{n-1}(v)\big)$$

folgt, dass es kein Polynom $f \in K[X]$ vom Grade kleiner als n geben kann, das Φ als Nullstelle hat. Denn ist Φ Nullstelle von f, also $f(\Phi) = 0$, so folgt $f(\Phi(v)) = 0$ und wir finden eine nicht-triviale Relation von Vektoren in $\mathcal{B}'$. □

Korollar 7.4.4.
Für jeden Endomorphismus Φ eines endlich-dimensionalen Vektorraums V der Dimension n ist das Minimalpolynom μ_Φ gleich dem n-ten Invariantenteiler von Φ, also $\mu_\Phi = c_\Phi^{(n)}$.

Beweis.
Es seien $g_1, \ldots, g_r$ die nicht-konstanten Invariantenteiler von Φ mit Begleitmatrizen B_i. Da ähnliche Matrizen das gleiche Minimalpolynom haben, ist nur zu zeigen, dass die blockdiagonale Matrix

$$B = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_r \end{pmatrix}$$

aus Begleitmatrizen $B_i := B_{g_i}$ das Polynom g_r als Minimalpolynom besitzt. Nun gilt aber für jedes Polynom $g \in K[X]$

$$g(B) = \begin{pmatrix} g(B_1) & & & 0 \\ & g(B_2) & & \\ & & \ddots & \\ 0 & & & g(B_r) \end{pmatrix}.$$

Wir wissen aus Korollar 7.4.3, dass g_i das Minimalpolynom der Matrix B_i ist. Daher gilt $g(B) = 0$ genau dann, wenn alle g_i das Polynom g teilen. Wegen der Teilbarkeitsbeziehung $g_i | g_{i+1}$ für die Invariantenteiler ist dies aber äquivalent zu der Teilbarkeitsbeziehung $g_r | g$. Folglich ist g_r das Minimalpolynom der blockdiagonalen Matrix B. □

Bemerkungen 7.4.5.

1. Das charakteristische Polynom P_Φ ist als n-ter Determinantenteiler Produkt von Invariantenteilern,
$$P_\Phi = d_\Phi^{(n)} = c_\Phi^{(1)} \cdot c_\Phi^{(2)} \cdots c_\Phi^{(n)} .$$
Wegen Korollar 7.4.4 gilt $\mu_\Phi = c_\Phi^{(n)}$. Es folgt erneut die Teilbarkeitsbeziehung $\mu_\Phi \,|\, P_\Phi$, also der Satz von Cayley-Hamilton 5.5.7. Aus den Teilbarkeitsbeziehungen der Invariantenteiler folgt die weitere Teilbarkeitsbeziehung $P_\Phi \,|\, (\mu_\Phi)^n = (c_\Phi^{(n)})^n$.

2. Sei Φ ein Endomorphismus des endlich-dimensionalen Vektorraums V. Dann ist V genau dann zyklisch bezüglich Φ, wenn Minimalpolynom und charakteristisches Polynom übereinstimmen.
Eine Richtung war Korollar 7.4.3. Stimmen umgekehrt charakteristisches Polynom und Minimalpolynom überein, so sind die Invariantenteiler gleich $1, 1, \ldots, \mu_\Phi$, also die Invariantenteiler der Begleitmatrix des Minimalpolynoms. Begleitmatrizen aber sind darstellende Matrizen von Endomorphismen zyklischer Vektorräume.

Wir nennen schließlich einen Φ-invarianten Untervektorraum $U \subset V$ zyklisch bezüglich $\Phi \in \mathrm{End}_K(V)$, wenn U zyklisch bezüglich der Einschränkung von Φ auf U ist.

Satz 7.4.6.
Es sei Φ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V.

1. Dann kann man V als direkte Summe
$$V = V_1 \oplus \ldots V_2 \oplus \cdots \oplus V_r$$
Φ-zyklischer Untervektorräume $V_i \neq 0$ von V darstellen, wobei für die Minimalpolynome $g_i \in K[X]$ der Einschränkungen $\Phi_{|V_i}$ mit $i = 1, \ldots, r$ von Φ die Teilbarkeitsbeziehungen $g_i | g_{i+1}$ gelten.

2. Die Polynome g_i sind dann gleich den Invariantenteilern von Φ. Insbesondere ist g_r das Minimalpolynom von Φ.

Beweis.

- Wir zeigen zunächst die Eindeutigkeitsaussage für die Minimalpolynome der Einschränkungen $\Phi_{|V_i}$. Sei eine Zerlegung mit den beschriebenen Eigenschaften vorgegeben. Man wähle in jedem zyklischen Unterraum V_i eine Basis wie in Satz 7.4.2 und setzt sie zusammen zu einer Basis von V, in der die darstellende Matrix von Φ die blockdiagonale Gestalt aus Begleitmatrizen aus Korollar 7.4.4 besitzt. Aus der Eindeutigkeit der Invariantenteiler folgt dann, dass die auftretenden Polynome die nicht-konstanten Invariantenteiler von Φ sein müssen.

- Um die Existenz einer solchen Zerlegung zu zeigen, beachten wir, dass es nach Satz 7.3.3 eine Basis von V gibt, in der die darstellende Matrix von Φ blockdiagonale Gestalt $B_{g_1,\ldots,g_r}$ hat. Jedem der Kästchen entspricht dabei ein Φ-invarianter Unterraum. Da die Einschränkung von Φ auf diesen Unterraum V_i durch die Begleitmatrix B_{g_i} dargestellt wird, ist dieser Unterraum nach Satz 7.4.2 auch zyklisch.

□

Man beachte, dass im allgemeinen nur die auftretenden Minimalpolynome, nicht aber die Φ-zyklischen Untervektorräume eindeutig bestimmt sind. Dazu überlegen wir uns, dass die

Φ-zyklischen Unterräume der Abbildung $\Phi = \lambda \mathrm{id}_V$ mit $\lambda \in K$ genau die eindimensionalen Untervektorräume sind: denn wegen $\Phi(v) = \lambda v$ ist $\mathrm{span}_K(v, \Phi(v), \Phi^2(v), \ldots) = \mathrm{span}_K(v)$ für jedes $v \in V$. Damit ist auch schon die Zerlegung eines Eigenraums der Dimension größer als Eins in eindimensionale invariante Unterräume, die dann ja Φ-zyklische Untervektorräume sind, ist nicht eindeutig.

Wir halten schließlich auch noch ohne Beweis die Aussagen für Endomorphismen fest, die analog zur Weierstraß'schen Normalform (vgl. Satz 7.3.5) sind.

Satz 7.4.7.
Es sei Φ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V. Dann kann man V als direkte Summe

$$V = W_1 \oplus \ldots W_2 \oplus \cdots \oplus W_m$$

Φ-zyklischer Untervektorräume $W_i \neq 0$ von V darstellen, wobei das Minimalpolynom $h_i \in K[X]$ jeder Einschränkung $\Phi_{|W_i}$ mit $i = 1, \ldots, m$ von Φ Potenz eines normierten Primpolynoms in $K[X]$ ist. Die Polynome h_i sind gleich den Weierstraß'schen Elementarteilern von Φ und daher bis auf Reihenfolge eindeutig.

7.5 Eine Sprache: Kategorien und Funktoren

Definition 7.5.1
Eine Kategorie $\mathcal{C}$ besteht aus einer Klasse $Obj(\mathcal{C})$ von Objekten und zu je zwei Objekten X, Y von $\mathcal{C}$ einer Menge von Morphismen $\mathrm{Hom}_{\mathcal{C}}(X, Y)$. Wir bezeichnen ein Element $f \in \mathrm{Hom}_{\mathcal{C}}(X, Y)$ auch mit $X \xrightarrow{f} Y$. Als weitere Struktur fordern wir:

1. *Eine Kompositionsabbildung für jedes Tripel X, Y, Z von Objekten von $\mathcal{C}$*

$$\begin{aligned} \mathrm{Hom}_{\mathcal{C}}(Y, Z) \times \mathrm{Hom}_{\mathcal{C}}(X, Y) &\to \mathrm{Hom}(X, Z) \\ (Y \xrightarrow{g} Z,\ X \xrightarrow{f} Y) &\mapsto (X \xrightarrow{g \circ f} Z)\,, \end{aligned}$$

   die assoziativ ist, wo sie definiert ist.

2. *Zu jedem Objekt X von $\mathcal{C}$ eine Identitätsmorphismus $\mathrm{id}_X \in \mathrm{Hom}_{\mathcal{C}}(X, X)$, so dass für alle Morphismen $X \xrightarrow{f} Y$ die Gleichheiten $\mathrm{id}_Y \circ f = f$ und $f \circ \mathrm{id}_X = f$ gelten.*

Beispiele 7.5.2.

1. Set: die Objekte sind Mengen und die Morphismen sind beliebige Abbildungen. Diese Kategorie war Gegenstand von Kapitel 1.3 dieser Vorlesung.

2. *Grp*: die Objekte sind Gruppen und die Morphismen sind Gruppenhomomorphismen. Diese Kategorie war Gegenstand von Kapitel 2.1 dieser Vorlesung.

3. *Ab*: die Objekte sind abelsche Gruppen und die Morphismen wiederum Gruppenhomomorphismen.

4. K-Vekt: die Objekte sind K–Vektorräume und die Morphismen sind K–lineare Abbildungen. Die Objekte dieser Kategorie waren Gegenstand der Kapitel 2.3-2.5 dieser Vorlesung, die Morphismen wurden in Kapitel 3 dieser Vorlesung behandelt.

5. *Top*: die Objekte sind topologische Räume und die Morphismen sind stetige Abbildungen. Damit beschäftigt man sich in den Vorlesungen Analysis und Topologie.

Definition 7.5.3
Ein Morphismus $X \xrightarrow{f} Y$ heißt Isomorphismus, wenn es Morphismen $g, h : Y \to X$ gibt, so dass $f \circ g = \mathrm{id}_Y$ und $h \circ f = \mathrm{id}_X$ gilt.

Bemerkung: es folgt dann aus der Assoziativität der Komposition von Morphismen

$$h = h \circ \mathrm{id}_Y = h \circ (f \circ g) = (h \circ f) \circ g = \mathrm{id}_X \circ g = g \ .$$

Dieses Argument trat schon im dritten Teil des Beweises von Satz 1.4.22 auf.

Wir brauchen nun auch noch Abbildungen zwischen Kategorien:

Definition 7.5.4
Seien $\mathcal{C}, \mathcal{D}$ Kategorien. Ein kovarianter Funktor bzw. ein kontravarianter Funktor $F : \mathcal{C} \to \mathcal{D}$ besteht aus:

1. *einer "Abbildung" $F : Ob(\mathcal{C}) \to Ob(\mathcal{D})$*
2. *einer Abbildung für jedes Paar von Objekten A_1, A_2 in $\mathcal{C}$:*

$$F : \mathrm{Hom}_{\mathcal{C}}(A_1, A_2) \to \mathrm{Hom}_{\mathcal{D}}(F(A_1), F(A_2))$$
$$\text{bzw. } F : \mathrm{Hom}_{\mathcal{C}}(A_1, A_2) \to \mathrm{Hom}_{\mathcal{D}}(F(A_2), F(A_1))$$

so dass gilt

1. *für jedes Objekt A in $\mathcal{C}$ die Gleichung $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$.*
2. *$F(f \circ g) = F(f) \circ F(g)$ bzw. $F(f \circ g) = F(g) \circ F(f)$.*

Beispiele 7.5.5.

1. Vergissfunktoren, die zum Beispiel einen $\mathbb{C}$–Vektorraum durch Einschränkung der skalaren Multiplikation auf Skalare in $\mathbb{R} \subset \mathbb{C}$ als $\mathbb{R}$–Vektorraum ansehen. Natürlich kann man noch weitere Struktur vergessen und z.B. $\mathbb{R}$-Vektorräume als abelsche Gruppen oder gar nur als Mengen betrachten. Vergissfunktoren sind alle kovariant.

2. Der Funktor $F : K\text{-Vekt} \to K\text{-Vekt}$, der jedem Vektorraum seinen Dualraum zuordnet, $F(V) = V^*$, und jeder linearen Abbildung die duale Abbildung, $F(f) = f^*$, ist ein kontravarianter Funktor, vergleiche Lemma 6.1.11.

3. Der Funktor $F : K\text{-Vekt} \to K\text{-Vekt}$, der jedem Vektorraum seinen Bidualraum zuordnet, $F(V) = V^{**}$ und jeder Abbildung die biduale Abbildung, $F(f) = f^{**}$, ist ein kovarianter Funktor, vergleiche Betrachtung 6.1.15 und Satz 6.1.16.

4. Sei $\mathcal{C}$ eine beliebige Kategorie und W ein Objekt in $\mathcal{C}$. Dann definieren wir einen kovarianten Funktor

$$\mathrm{Hom}(W, ?) \ : \ \mathcal{C} \to \mathrm{Set}$$

durch

$$\mathrm{Hom}(W, ?) : X \to \mathrm{Hom}_{\mathcal{C}}(W, X)$$

auf Objekten und für einen Morphismus $X \xrightarrow{\varphi} Y$ durch die Abbildung

$$\begin{array}{rccc} \varphi_* : & \mathrm{Hom}_{\mathcal{C}}(W,X) & \to & \mathrm{Hom}_{\mathcal{C}}(W,Y) \\ & f & \mapsto & \varphi \circ f \quad . \end{array}$$

Die Funktoren φ_*, ψ_* sind in der Tat mit der Komposition verträglich, wo diese definiert ist. Es gilt

$$\varphi_* \circ \psi_*(f) = \varphi \circ (\psi \circ f) \overset{\text{ass}}{=} (\varphi \circ \psi) \circ f = (\varphi \circ \psi)_*(f) \ ,$$

also $\varphi_* \circ \psi_* = (\varphi \circ \psi)_*$. Analog wird durch

$$\mathrm{Hom}(?, W) : X \to \mathrm{Hom}_{\mathcal{C}}(X, W)$$

auf Objekten und die Abbildung

$$\begin{array}{rccc} \varphi^* : & \mathrm{Hom}_{\mathcal{C}}(Y,W) & \to & \mathrm{Hom}_{\mathcal{C}}(X,W) \\ & f & \mapsto & f \circ \varphi \quad . \end{array}$$

ein kontravarianter Funktor

$$\mathrm{Hom}(?, W) \ : \ \mathcal{C} \to \mathrm{Set}$$

definiert.

Wenn man für $\mathcal{C}$ die Kategorie der K-Vektorräume betrachtet, so hat die Morphismenmenge $\mathrm{Hom}(\cdot, \cdot)$ sogar die Struktur eines K-Vektorraums. In diesem Fall liefern $\mathrm{Hom}(?, W)$ und $\mathrm{Hom}(W, ?)$ sogar Funktoren in die Kategorie der K-Vektorräume.

5. Wir kommen noch einmal auf die Situation in 1. zurück: sei $K \subset L$ eine Körpererweiterung, etwa $K = \mathbb{R}$ und $L = \mathbb{C}$. Dann kann man jeden L-Vektorraum V auch als K-Vektorraum auffassen. Dies definierte einen Vergissfunktor. Notiert man diesen wieder mit V, so muss man bei der Tensorprodukten aufpassen, da für zwei L-Vektorräume die Tensorprodukte $V \otimes_L W$ und $V \otimes_K W$ nicht gleich sind. (Sie haben universelle Eigenschaften für K-bilineare bzw. für L-bilineare Abbildungen.)

   Insbesondere können wir L selbst als K-Vektorraum auffassen. Für einen K-Vektorraum V können wir das Tensorprodukt $L \otimes_K V$ mit der Struktur eines L-Vektorraums versehen:

$$\lambda(\mu \otimes_K v) := \lambda\mu \otimes_K v \quad \text{ für } \lambda, \mu \in K, v \in V \ .$$

   Dies liefert einen Funktor $\mathrm{Ind}_K^L : K\text{-Vekt} \to L\text{-Vekt}$, der Induktionsfunktor oder Skalarenerweiterung genannt wird, vgl. auch Beispiel 6.3.4.2.

Es stellt sich die Frage, wie der Vergissfunktor und der Induktionsfunktor zusammenhängen. Inverse können sie nicht sein: der Vergissfunktor macht aus einem komplex n-dimensionalen Vektorraum einen reell $2n$-dimensionalen, der vom Induktionsfunktor in einen komplex $2n$-dimensionalen Vektorraum überführt wird. Das folgende Konzept ist zentral in der Kategorientheorie und kodiert insbesondere universelle Eigenschaften:

Definition 7.5.6
Zwei Funktoren $F : \mathcal{C} \to \mathcal{D}$ und $G : \mathcal{D} \to \mathcal{C}$ heißen adjungiert, falls es für je zwei Objekte X in $\mathcal{C}$ und Y in $\mathcal{D}$ einen Isomorphismus

$$\Phi_{X,Y} : \ \mathrm{Hom}_{\mathcal{C}}(X, GY) \xrightarrow{\sim} \mathrm{Hom}_{\mathcal{D}}(FX, Y)$$

mit der folgenden natürlichen Eigenschaft gibt: für jeden Homomorphismus $X' \xrightarrow{f} X$ in $\mathcal{C}$ und $Y \xrightarrow{g} Y'$ in $\mathcal{D}$ kommutiert das Diagramm

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{C}}(X, GY) & \xrightarrow{\mathrm{Hom}(f,Gg)} & \mathrm{Hom}_{\mathcal{C}}(X', GY') \\ \Big\downarrow{\scriptstyle \Phi_{X,Y}} & & \Big\downarrow{\scriptstyle \Phi_{X',Y'}} \\ \mathrm{Hom}_{\mathcal{D}}(FX, Y) & \xrightarrow[\mathrm{Hom}(Ff,g)]{} & \mathrm{Hom}_{\mathcal{D}}(FX', Y') \end{array}$$

mit

$$\mathrm{Hom}(Ff, g)(\varphi) := g \circ \varphi \circ Ff : \quad FX' \xrightarrow{Ff} FX \xrightarrow{\varphi} Y \xrightarrow{g} Y'$$

und

$$\mathrm{Hom}(f, Gg)(\tilde{\varphi}) := Gg \circ \tilde{\varphi} \circ f : \quad X' \xrightarrow{f} X \xrightarrow{\tilde{\varphi}} GY \xrightarrow{Gg} GY' \ .$$

Man schreibt $F \dashv G$.

Beispiel 7.5.7.

1. Als Beispiel betrachten wir den Vergissfunktor

   $$G : \ K\text{-Vekt} \to \mathrm{Set} \ ,$$

   der einem Vektorraum die zugrunde liegende Menge zuordnet. Sein linksadjungierter Funktor muss ein Funktor

   $$F : \ \mathrm{Set} \to K\text{-Vekt} \, ,$$

   sein, der jeder Menge M einen Vektorraum zuordnet. Hierbei muss es einen Isomorphismus

   $$\Phi_{M,V} : \ \mathrm{Hom}_{\mathrm{Set}}(M, G(V)) \to \mathrm{Hom}_K(F(M), V) \qquad (*)$$

   geben.

2. Lineare Abbildungen aus dem gesuchten Vektorraum $F(M)$ heraus werden also eindeutig dadurch beschrieben, dass man jedem Element von M ein Bild im Zielvektorraum V zuordnet. Da man lineare Abbildungen genau auf Basen eindeutig vorschreiben kann, muss $F(M)$ ein Vektorraum sein, der eine Basis hat, die in Bijektion zu M steht. Einen solchen Vektorraum kennen wir: $F(M)$ ist der Vektorraum der Abbildungen $f : M \to K$, die nur für endlich viele Elemente von M einen Wert ungleich Null annehmen. Man beachte, dass wir eine Abbildung von Mengen haben

   $$M \xrightarrow{\eta_M} GF(M) \ ,$$

   die $m \in M$ auf ein Basiselement von $F(M)$ schickt, nämlich die Funktion, die nur auf $m \in M$ gleich Eins und sonst gleich Null ist. Der Vektorraum $F(M)$ heißt auch der von der Menge M erzeugte Vektorraum.

3. Wir müssen noch für jede Menge M und jeden K-Vektorraum V einen Isomorphismus

   $$\begin{array}{rccc} \Phi_{M,V} : & \mathrm{Hom}_{\mathrm{Set}}(M, G(V)) & \to & \mathrm{Hom}_K(F(M), V) \\ & \varphi & \mapsto & \Phi_{M,V}(\varphi) \end{array}$$

   angeben. Dieser ist

   $$\Phi_{M,V}(\varphi)(\sum_{m \in M} \lambda_m m) := \sum_{m \in M} \lambda_m \varphi(m) \ .$$

4. Insbesondere finden wir, dass für jeden K-Vektorraum V die einelementige Menge $\mathrm{Hom}_{\mathrm{Set}}(\emptyset, G(V))$ isomorph zu $\mathrm{Hom}_K(F(\emptyset), V)$ sein muss. Der einzige K-Vektorraum aber, von dem es genau eine lineare Abbildung in jeden K-Vektorraum V gibt, ist der Nullvektorraum, $F(\emptyset) = \{0\}$. Dies führt zu der Einsicht, dass Nullvektorraum von der leeren Mengen erzeugt wird, vergleiche Definition 2.4.16.

5. Kapitel 2.4 hatte dieses Paar adjungierter Funktoren zum Gegenstand. Man beachte, dass für eine beliebige Menge M der Vektorraum $F(M)$ eine universelle Eigenschaft hat: für jede Abbildung $\varphi : M \to V$ von Mengen existiert genau eine lineare Abbildung $F(M) \to V$, so dass das Diagramm

$$\begin{array}{ccc} M & \xrightarrow{\eta_M} & F(M) \\ & \searrow^{\varphi} & \downarrow{\scriptstyle \exists! \Phi_{M,V}(\varphi)} \\ & & V \end{array}$$

   kommutiert. Man beachte, dass in diesem Diagramm φ und η_M Abbildungen von Mengen sind, aber $\Phi_{M,V}(\varphi)$ eine lineare Abbildung. Man sollte eigentlich kommutierende Diagramme nur in einer Kategorie hinschreiben.

Beispiel 7.5.8.
Für je drei K-Vektorräume gilt

$$\mathrm{Hom}_K(V \otimes W, Z) \cong \mathrm{Hom}_K(V, \mathrm{Hom}_K(W, Z)) \ ,$$

was man zum Beispiel dadurch sieht, dass beide Seiten in Bijektion zu bilinearen Abbildungen $V \times W \to Z$ sind. Also ist für jeden K-Vektorraum W der Funktor

$$\begin{array}{rccc} - \otimes_K W : & K\text{-Vekt} & \to & K\text{-Vekt} \\ & V & \mapsto & V \otimes W \end{array}$$

der linksadjungierte Funktor zu dem Funktor

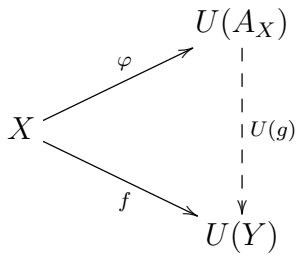
$$\begin{array}{rccc} \mathrm{Hom}_K(W, -) : & K\text{-Vekt} & \to & K\text{-Vekt} \\ & Z & \mapsto & \mathrm{Hom}_K(W, Z) \end{array}$$

Für eine universelle Eigenschaft braucht man tatsächlich weniger als ein Paar adjungierter Funktoren: im obigen Beispiel können wir schon die universelle Eigenschaft für *eine* einzige Menge M formulieren. Für eine genaue Formulierung brauchen wir aber zwei Kategorien $\mathcal{C}, \mathcal{D}$ und einen Funktor $U : \mathcal{D} \to \mathcal{C}$. In unserem Beispiel war $\mathcal{C}$ die Kategorie der Mengen, $\mathcal{D}$ die der K-Vektorräume und $U : \mathcal{D} = \mathrm{vect}_K \to \mathcal{C} = \mathrm{Set}$ der Vergissfunktor. Wir können uns vorstellen, dass wir dann nur das Objekt $X \in \mathcal{C}$ zu einem Objekt $A_M \in \mathcal{D}$ heben, in unserem Beispiel zum freien Vektorraum auf der Menge X.

Dies führt auf die folgende Definition universeller Eigenschaften in der Sprache von Kategorien und Funktoren:

Definition 7.5.9
Seien $\mathcal{C}, \mathcal{D}$ Kategorien und $U : \mathcal{D} \to \mathcal{C}$ ein Funktor. Ein universeller Morphismus von einem Objekt X in $\mathcal{C}$ nach U ist ein Paar (A_X, φ), bestehend aus einem Objekt A_X in $\mathcal{D}$ und einem Morphismus $\varphi : X \to U(A_X)$ in $\mathcal{C}$, so dass die folgende universelle Eigenschaft gilt:

Für jedes Objekt Y von $\mathcal{D}$ und jeden Morphismus $f : X \to U(Y)$ in $\mathcal{C}$ gibt es einen eindeutigen *Morphismus $g : A_X \to Y$ in $\mathcal{D}$, so dass das folgende Diagramm in $\mathcal{C}$ kommutiert:*



Wir haben also für jedes Objekt Y von $\mathcal{D}$ eine injektive Abbildung

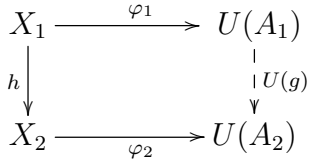
$$\mathrm{Hom}_{\mathcal{C}}(X, U(Y)) \to \mathrm{Hom}_{\mathcal{D}}(A, Y) \; . \qquad (*)$$

Diese Abbildung ist auch surjektiv: für $g \in \mathrm{Hom}_{\mathcal{D}}(A, Y)$ ist $U(g) \circ \varphi : X \to UY$ ein Urbild.

Bemerkungen 7.5.10.

1. Universelle Eigenschaften definieren wie gewohnt Objekte bis auf eindeutigen Isomorphismus. Gelingt es zu zeigen, dass zwei verschiedene Objekte die gleiche universelle Eigenschaft haben, so müssen diese insbesondere isomorph in $\mathcal{D}$ sein.

2. Universelle Konstruktionen sind funktoriell: sei (A_1, φ_1) ein universeller Morphismus von X_1 nach U und (A_2, φ_2) ein universeller Morphismus von X_2 nach U. Aus Definition 7.5.9, angewandt auf $\varphi_2 \circ h$, folgt, dass für jeden Morphismus $h : X_1 \to X_2$ in $\mathcal{C}$ ein eindeutig bestimmter Morphismus $g : A_1 \to A_2$ existiert, so dass das Diagramm



   kommutiert. So hatten wir auch das Tensorprodukt von linearen Abbildungen konstruiert.

3. Existiert ein universeller Morphismus für *jedes* Objekt X_i von $\mathcal{C}$, so definiert $X_i \mapsto A_i$ und $h \mapsto g$ einen kovarianten Funktor V von $\mathcal{C}$ nach $\mathcal{D}$. Es folgt dann aus dem Isomorphismus $(*)$, dass die Funktoren U und V adjungiert sind.

   Achtung: es liefert zwar jedes Paar adjungierter Funktoren universelle Morphismen; aber universelle Konstruktionen geben nur adjungierte Funktoren, wenn für jedes Objekt von $\mathcal{C}$ ein universeller Morphismus existiert.

Beispiele 7.5.11.

1. Um die universelle Eigenschaft 5.2.5 des Polynomrings zu verstehen, betrachten wir den Vergissfunktor

$$U : \; K - Alg \to \mathrm{Set} \; .$$

   Wir wollen nun den universellen Morphismus für die einelementige Menge $\bullet$ untersuchen. Dies ist ein Objekt in $K-Alg$, also eine K-Algebra R, mit einem Morphismus von Mengen $\bullet \to U(R)$, d.h. einem Element $X \in R$.

   Die universelle Eigenschaft ist nun die Forderung, dass es für jede K-Algebra S und einen Morphismus von Mengen $\bullet \to U(S)$, also für jedes Paar bestehend aus einer K-Algebra S und einem Element $a \in S$, einen eindeutigen Morphismus $f : R \to S$ von K-Algebren geben muss, so dass $\bullet \to U(R) \xrightarrow{f} U(S)$ gleich $\bullet \to S$ ist, also so dass $f(X) = a$ gilt. Wir haben so den Einsetzunghomomorphismus wieder erhalten vgl. Bemerkung 5.2.6.

2. Das Produkt einer Familie $(X_i)_{i\in I}$ von Vektorräumen. Hier ist $\mathcal{D} = K-\text{Vekt}$ und $\mathcal{C}$ ist die folgende Produktkategorie $(K-\text{Vekt})^I$: Objekte sind I-Tupel von K-Vektorräumen, ein Morphismus $(X_i)_{i\in I} \to (Y_i)_{i\in I}$ ist ein I-Tupel von linearen Abbildungen $f_i : X_i \to Y_i$. Als Funktor $U : \mathcal{D} \to \mathcal{C}$ betrachten wir den Diagonalenfunktor Δ, der einem K-Vektorraum X die konstante Familie zuordnet, $\Delta(X) = (X)_{i\in I}$, und einer linearen Abbildung $X \xrightarrow{\varphi} Y$ die konstante Familie linearer Abbildungen, $\Delta(f) = (f)_{i\in I}$.

   Der universelle Morphismus für ein Objekt $(X_i)_{i\in I}$ in $(K-\text{Vekt})^I$ ist ein K-Vektorraum P, mit einem Morphismus

$$\varphi : (X_i)_{i\in I} \to \Delta(P) = (P)_{i\in I}$$

   (dessen Komponenten φ_i gerade die kanonischen Injektionen $X_i \to P$ sind), so dass für jeden Morphismus der Produktkategorie

$$f : (X_i)_{i\in I} \to U(Y) = (Y)_{i\in I}$$

   (also für jede Familie linearer Abbildungen $f_i : X_i \to Y$ mit Y einem festen K-Vektorraum) eine lineare Abbildung $g : P \to Y$ existiert, so dass in der Produktkategorie

$$f = U(g) \circ \varphi$$

   gilt. Das heißt aber, dass für jedes $i \in I$ gilt $f_i = g \circ \varphi_i$, vergleiche Lemma 3.5.9

3. Für die direkte Summe muss man entsprechend den Begriff des kouniversellen Morphismus einführen, bei dem alle Pfeile umgedreht werden, vergleiche Lemma 3.5.13

7.6 Kurzzusammenfassung der Kapitel 4-7

Sei K im Folgenden ein beliebiger Körper.

7.6.1 Determinanten

Wir haben die Determinantenfunktion

$$\det : M(n \times n, K) \to K$$

als (D1) K-zeilenlineare, (D2) alternierende und (D3) normierte Funktion eingeführt.
Berechnung :

- Entwicklungssätze: nach Zeilen oder Spalten
- Leibniz'sche Regel mit $n!$ Termen (*nur* für 3×3-Matrizen: Sarrus'sche Regel mit 6 Termen)
- Obere Dreiecksmatrizen: $\det \begin{pmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} \lambda_i$
- Für eine blockdiagonale Matrix

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

  gilt $\det A = \det A_1 \cdot \det A_2$.

– Eine Determinante ändert sich nicht, wenn man ein Vielfaches einer Zeile zu einer anderen Zeile addiert. Man kann durch diese Umformungen eine Matrix in eine obere Dreiecksmatrix überführen und dann die Determinante als Produkt der Diagonalelemente berechnen.

Eigenschaften :

– Multiplikativität: $\det(AB) = \det A \cdot \det B$. Deswegen liefert die Determinante eine Funktion auf Ähnlichkeitsklassen von Matrizen und somit eine Funktion auf Endomorphismen endlich-dimensionaler Vektorräume.

– Beziehung zum Rang: $\det A \neq 0 \Leftrightarrow \operatorname{rg} A$ maximal $\Leftrightarrow A \in GL(n, K)$, also in der allgemeinen linearen Gruppe.

– Transformation der Volumensfunktion unter affinen Abbildungen $F : \mathbb{R}^n \to \mathbb{R}^n$ mit $F(x) = Ax + b$:

$$\operatorname{vol}_n F(X) = |\det A| \operatorname{vol}_n X \,.$$

Gleichungssysteme : $Ax = b$ mit A invertibel (also quadratisch).

– Inverse Matrix durch Streichungsmatrizen: $A_{ij}^{-1} = (\det A)^{-1}(-1)^{i+j} \det\left(A_{ji}^{str}\right)$.

– Cramersche Regel: das inhomogene lineare Gleichungssystem $Ax = b$ mit A quadratisch und invertibel hat die eindeutige Lösung

$$x_i = \frac{\det(a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_n)}{\det A}$$

7.6.2 Endomorphismen $\Phi \in \operatorname{End}_K(V)$

Gilt $\Phi v = \lambda v$ mit $v \neq 0$, so heißt $\lambda \in K$ Eigenwert und $v \in V \setminus \{0\}$ Eigenvektor. Der Eigenraum von Φ zum Eigenwert $\lambda \in K$ ist $\operatorname{Eig}(\Phi, \lambda) := \ker(\Phi - \lambda \mathrm{id}_V)$, ein Untervektorraum.

– Geometrische Vielfachheit:

$$\mu_{geo}(\Phi, \lambda) := \dim_K \operatorname{Eig}(\Phi, \lambda)$$

– Charakteristisches Polynom: $P_\Phi(X) := \det(X\mathrm{id}_V - \Phi)$ mit Determinante und Spur als speziellen Koeffizienten:

$$a_n = 1 \qquad a_{n-1} = -\operatorname{Tr} \Phi \qquad a_0 = (-1)^n \det \Phi \,.$$

– Vielfachheiten: Eigenwerte sind Nullstellen des charakteristischen Polynoms; die Vielfachheit als Nullstelle von P_Φ heißt algebraische Vielfachheit des Eigenwerts. Es gilt für jeden Eigenwert λ:

$$1 \leq \mu_{geo}(\Phi, \lambda) \leq \mu_{alg}(\Phi, \lambda) \,.$$

7.6.3 Polynomalgebra

– Universelle Eigenschaft ist Einsetzen: für jede K–Algebra S und jedes Element $a \in S$ existiert ein eindeutiger Algebrenhomomorphismus, der Einsetzunghomomorphismus

$$\begin{aligned} \varphi_a : \; K[X] &\to S \\ X &\mapsto a \end{aligned}$$

– In Polynomalgebren über Körpern gibt es eine Division mit Rest: sie sind euklidische Ringe ⇒ Hauptidealringe (wie der Ring $\mathbb{Z}$ der ganzen Zahlen) ⇒ eindeutige Primzerlegung

– Fundamentalsatz der Algebra: Polynome in $\mathbb{C}[X]$ zerfallen vollständig in Linearfaktoren. Primpolynome in $\mathbb{C}[X]$ sind von der Form $(X-a)$ mit $a \in \mathbb{C}$, Primpolynome in $\mathbb{R}[X]$ sind von der Form $(X-a)$ mit $a \in \mathbb{R}$ oder X^2+pX+q mit $p^2-4q<0$.

7.6.4 Diagonalisierbarkeit, Trigonalisierbarkeit

– Das Minimalpolynom $\mu_\Phi \in K[X]$ ist der normierte Erzeuger des Annihilatorideals von $\Phi : V \to V$ in $K[X]$:

$$\mathrm{Ann}(\Phi) = \{f \in K[X] \mid f(\Phi) = 0\} = \mu_\Phi \cdot K[X]$$

– Satz von Cayley–Hamilton: $\widetilde{P}_\Phi(\Phi) = 0 \Leftrightarrow \mu_\Phi \mid P_\Phi$.
Bemerkung: umgekehrt gilt $P_\Phi \mid \mu_\Phi^n$ mit $n := \dim_K V$.

– Minimalpolynom und charakteristisches Polynom haben die gleichen Nullstellen.

– Minimalpolynom und charakteristisches Polynom sind bezüglich der Teilbarkeit größte Invarianten- bzw. Determinantenteiler

– Charakteristisches Polynom P_Φ zerfällt vollständig in Linearfaktoren $\Leftrightarrow$ Φ trigonalisierbar. Für solche Φ existiert sogar eine geordnete Basis, so dass

$$M_{\mathcal{B}}(\Phi) = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & 0 \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{matrix}} & & \\ & & \\ & & \ddots \end{pmatrix}$$

(Jordansche Normalform) Weitere Normalformen für beliebige Matrizen: Frobenius'sche und Weierstraß'sche Normalform: blockdiagonale Matrizen mit Begleitmatrizen als Blöcken.

– Φ diagonalisierbar $\Leftrightarrow$ Charakteristisches Polynom P_Φ zerfällt vollständig in Linearfaktoren und $\mu_{geo}(\lambda, \Phi) = \mu_{alg}(\lambda, \Phi)$ für alle Eigenwerte.
$\Leftrightarrow$ Minimalpolynom μ_Φ zerfällt vollständig in paarweise verschiedene Linearfaktoren.

– Zwei diagonalisierbare Endomorphismen Φ, Ψ sind genau dann *gleichzeitig* diagonalisierbar, wenn sie vertauschen, $\Phi \circ \Psi = \Psi \circ \Phi$.

– V ist Φ-zyklisch $\Leftrightarrow$ Minimalpolynom und charakteristisches Polynom sind gleich $\Leftrightarrow$ Es gibt eine Basis $\mathcal{B}$, in der die darstellende Matrix $M_{\mathcal{B}}(\Phi)$ eine Begleitmatrix eines Polynoms ist.

7.6.5 Dualraum

– Einem Vektorraum V mit Basis $\mathcal{B}$ ordnen wir seinen Dualraum $V^* := \mathrm{Hom}(V, K)$, wenn $\dim_K V < \infty$ mit dualer Basis $\mathcal{B}^*$ zu.

– Einer linearen Abbildung $V \xrightarrow{f} W$ ordnen wir zu die duale Abbildung $W^* \xrightarrow{f^*} V^*$. Es gilt $(f \circ g)^* = g^* \circ f^*$ und

$$\text{Im } f^* = (\ker f)^\circ \qquad \ker f^* = (\text{Im } f)^\circ \qquad M_{\mathcal{A}^*}^{\mathcal{B}^*}(f^*) = M_{\mathcal{B}}^{\mathcal{A}}(f)^t \ ,$$

wobei $U^\circ \in V^*$ der Annulator des Untervektorraums $U \subset V$ ist.

– Ist $\dim V < \infty$, so gibt es kanonische Isomorphismen $i_V : V \xrightarrow{\sim} V^{**}$ mit $f^{**} = i_W \circ f \circ i_V^{-1}$.

7.6.6 Bilinearformen $\beta : V \times W \to K$

– Transformationsformel für darstellende Matrix $M_{\mathcal{A},\mathcal{B}}(\beta)_{ij} = \beta(v_i, w_j)$:

$$M_{\mathcal{A},\mathcal{B}}(\beta) = \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^t M_{\mathcal{A}',\mathcal{B}'}(\beta) T_{\mathcal{B}'}^{\mathcal{B}} \ .$$

$\Rightarrow$ Begriff der Kongruenz von Matrizen

Im Folgenden gelte im Körper K die Ungleichung $1 + 1 \neq 0$.

1. <u>alternierend</u>: $\beta(x, x) = 0$ für alle $x \in V$.

   – Ist β alternierend und nicht-ausgeartet, so heißt β symplektisch. $\Rightarrow \dim_K V$ gerade.
   – Normalform für symplektische Bilinearformen:

$$M_{\mathcal{B}}(\beta) = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & \\ & & & \\ 0 & & & \ddots \end{pmatrix}$$

2. <u>symmetrisch</u>

   – Festgelegt durch quadratische Form $q : V \to K \quad q(x) = \frac{1}{2}\beta(x, x)$.
   – Polarisierungsformel: $\beta(x, y) = q(x + y) - q(x) - q(y)$ ist Bilinearform.
   – Normalform: $\beta(b_i, b_j) = \alpha_i \delta_{i,j}$ mit $\alpha_i \in K$. Kongruenzproblem $\to$ Wittsche Relationen.
   – Insbesondere für reelle Vektorräume, $K = \mathbb{R}$ (Sylvesterscher Trägheitssatz):

$$M_{\mathcal{B}}(\beta) = \text{diag}(\underbrace{1, \ldots, 1}_{r_+}, \underbrace{-1, \ldots, -1}_{r_-}, \underbrace{0, \ldots, 0}_{r_0})$$

   β ist nicht-ausgeartet $\Leftrightarrow r_0 = 0 \Leftrightarrow$ für jedes $v \in V \setminus \{0\}$ existiert $w \in V$ mit $\beta(v, w) \neq 0 \Leftrightarrow \det M_{\mathcal{B}}(\beta) \neq 0$.
   β ist positiv definit $\Leftrightarrow r_0 = r_- = 0 \Leftrightarrow q(v) > 0$ für alle $v \in V \setminus \{0\}$.
   – Orthogonales Komplement: Im Spezialfall $V = W$ und symmetrischer positiv definiter Bilinearformen gilt $V = X \oplus X^\perp$.

7.6.7 Euklidische und unitäre Vektorräume

1. Euklidischer Vektorraum: endlich-dimensionaler reeller Vektorraum mit positiv definiter symmetrischer Bilinearform $\langle \cdot, \cdot \rangle$. Norm $\|x\| := \sqrt{\langle x, x \rangle}$.

   – Cauchy-Schwarz'sche Ungleichung:

   $$|\langle x, y \rangle| \leq \|x\| \|y\| \quad \cos \alpha = \frac{\langle x, y \rangle}{\|x\| \|y\|} \text{ Innenwinkel}$$

   – Orthonormalbasen, Orthonormierungsverfahren

   – Orthogonale Projektion auf Untervektorräume

2. Unitärer Vektorraum: endlich-dimensionaler komplexer Vektorraum mit positiv definiter Sesquilinearform $\langle \cdot, \cdot \rangle$. Norm $\|x\| := \sqrt{\langle x, x \rangle}$.

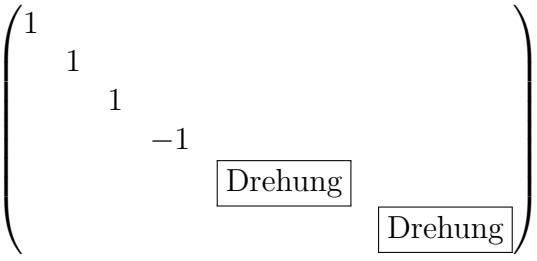
3. Selbstadjungierte Endomorphismen $\Phi : V \to W$ mit $\langle \Phi v, w \rangle = \langle v, \Phi w \rangle$ für alle $v \in V$ und $w \in W$

   – Euklidischer Fall: in einer Orthonormalbasis $\mathcal{B}$ ist $M_{\mathcal{B}}(\Phi) = M_{\mathcal{B}}(\Phi)^t$, also symmetrische Matrix.

   – Unitärer Fall: in einer Orthonormalbasis $\mathcal{B}$ ist $M_{\mathcal{B}}(\Phi) = M_{\mathcal{B}}(\Phi)^*$ also hermitesche Matrix.

   – Selbstadjungierte Endomorphismen sind mit einer Orthonormalbasis diagonalisierbar, alle ihre Eigenwerte sind reell. Insebsondere sind Eigenräume zu verschiedenen Eigenwerten orthogonal.

4. Isometrien $\Phi : V \to V$ mit $\langle \Phi v, \Phi w \rangle = \langle v, w \rangle$ für alle $v, w \in V$

euklidischer Vektorraum	unitärer Vektorraum
orthogonale Abbildungen	unitäre Abbildungen
bilden eine Gruppe	
$O(V)$	$U(V)$
sind normerhaltend	
winkeltreu	erhalten Orthogonalität
Eigenwerte:	
$\lambda \in \{\pm 1\}$	$\lvert\lambda\rvert = 1$
Determinante:	
± 1	$\lvert\det A\rvert = 1$
Untergruppe mit Determinante 1:	
$SO(V)$	$SU(V)$
$A := M_{\mathcal{B}}(\Phi)$ in Orthonormalbasis:	
$A^t A = E_n$	$A^* A = E_n$
orthogonale Matrix	unitäre Matrix
Normalform:	
	diagonalisierbar.