

Lineare Algebra
Studienjahr 2023/24
Julian Holstein
Fachbereich Mathematik

(Stand: 13. Februar 2024)

Inhaltsverzeichnis

1	Vorbereitung	2
1.1	Was ist lineare Algebra?	2
1.2	Geometrie von Geraden in der Ebene	2
1.3	Lineare Gleichungssysteme, Gauß'scher Algorithmus	10
1.4	Aussagen und Logik	14
1.5	Mengen	20
1.6	Relationen und Abbildungen	25
2	Algebraische Grundbegriffe	31
2.1	Gruppen	31
2.2	Homomorphismen	36
2.3	Körper	40
2.4	Die komplexen Zahlen	41
2.5	Ringe	44
3	Vektorräume	48
3.1	Vektorräume	48
3.2	Untervektorräume	50
3.3	Lineare Abbildungen	53
3.4	Quotientenvektorräume	58
4	Basen	62
4.1	Linearkombinationen	62
4.2	Basis und Dimension	65
4.3	Abbildungen und Dimension	71
5	Lineare Abbildungen und Basen	74
5.1	Lineare Abbildungen und Matrizen	74
5.2	Mehr über Matrizen	78
5.3	Der Gaußsche Algorithmus und Elementarmatrizen	81
5.4	Koordinatentransformationen	87
5.5	Äquivalenz und Ähnlichkeit von Matrizen	92
6	Intermezzo: Basen für beliebige Vektorräume	97

7	Determinanten	99
7.1	Vorüberlegungen	99
7.2	Die Determinantenabbildung	101
7.3	Existenz der Determinante	105
7.4	Eigenschaften der Determinante	110
A	Leere Summen und Produkte	113
B	Was bisher geschah: Kurzzusammenfassung der Kapitel 2 bis 5	115

Literatur:

Dieses Skript basiert auf Christoph Schweigerts Skript zur Linearen Algebra, erhältlich auf <https://www.math.uni-hamburg.de/home/schweigert/skripten/laskript.pdf>.

Nützliche ergänzende Literatur:

- Christian Bär: Lineare Algebra und analytische Geometrie Springer Fachmedien Wiesbaden, 2018. Volltextzugang Campus
<https://doi.org/10.1007/978-3-658-22620-6>
- Gerd Fischer: Lineare Algebra : eine Einführung für Studienanfänger. Springer Spektrum, 18. Auflage 2014. Volltextzugang Campus
<http://dx.doi.org/10.1007/978-3-658-03945-5>

Die aktuelle Version dieses Skriptes finden Sie unter

<http://www.math.uni-hamburg.de/home/holstein/lehre/lina23/LASkript.pdf>
als pdf-Datei.

Bitte schicken Sie Korrekturen und Bemerkungen an julian.holstein@uni-hamburg.de!

1 Vorbereitung

1.1 Was ist lineare Algebra?

Zur Einführung möchte ich etwas dazu sagen, womit wir uns in den nächsten beiden Semestern gemeinsam beschäftigen. Was ist lineare Algebra, warum studieren wir sie?

Ein paar Gedanken dazu.

1. Geometrisch: Lineare Algebra beschreibt und abstrahiert die Anschauung des linearen Raums: Gerade, Ebene, drei-dimensionaler Raum, mehr-dimensionaler Hyperraum. Keine komplizierten Kurven, nur Geraden, Ebenen usw. aber in beliebiger Dimension.¹
2. Lineare Algebra beschreibt und abstrahiert die Lösung von linearen Gleichungen wie $ax + b = c$, in der Praxis ist dies das Rechnen mit Reihen und Rechtecken voll von Zahlen.
3. Wenn ihnen diese etwas vagen Aussagen nicht allzu viel sagen ist das völlig in Ordnung, am Ende des Kurses werden Sie Ihre eigenen, konkreten Vorstellung haben, was lineare Algebra ist.
4. Lineare Algebra wird ihnen in nahezu jedem Kurs, den Sie während des Mathematikstudiums belegen wieder begegnen, Statistik, Differenzialgleichungen, komplexe Mannigfaltigkeiten, Quantenfeldtheorie ... Mathematik geht nicht ohne lineare Algebra.
5. Als ich in ihrem Alter war, war die coole Anwendung von Mathematik, mit der man Nichtmathematiker beeindrucken konnte der PageRank-Algorithmus von Google. Unter der Haube steckte darin eine Menge lineare Algebra. Heutzutage wirkt der PageRank-Algorithmus vielleicht schon etwas antiquiert, in den Nachrichten geht es um generative künstliche Intelligenz wie ChatGPT, Bard, Dall-E etc. Unter der Haube steckt wieder lineare Algebra. Wenn Sie in meinem Alter sind und vielleicht auch eine Anfängervorlesung halten, werden Sie vielleicht auch eine aktuelle, beeindruckende mathematische Anwendung präsentieren wollen. Niemand hier weiß, welche das sein wird. Aber höchstwahrscheinlich benutzt sie lineare Algebra.

In diesem Kapitel beschäftigen wir uns nun mit einigen vorbereitenden Betrachtungen: wir betrachten etwas elementare Geometrie der Ebene, dann werden wir lineare Gleichungssysteme systematische lösen. Anschließend führen wir etwas Handwerkszeug ein, nämlich die Sprache von Mengen und Abbildungen, Aussagen und deren Verknüpfungen.

1.2 Geometrie von Geraden in der Ebene

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

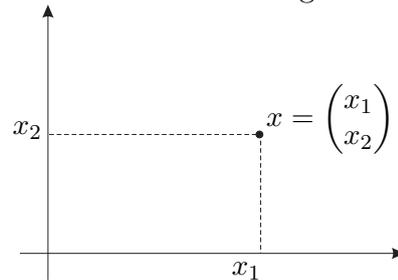
- Sie haben eine Vorstellung, was reelle Zahlen sind. In der Analysis wird dies noch einmal präzise eingeführt werden. Wir bezeichnen die Gesamtheit der reellen Zahlen mit \mathbb{R} . Wir sprechen auch von der *Menge* der reellen Zahlen.
- Reelle Zahlen können addiert und subtrahiert werden. Für Addition 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

¹Im zweiten Semester begegnen uns auch ein paar quadratische geometrische Objekte.

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.

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

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



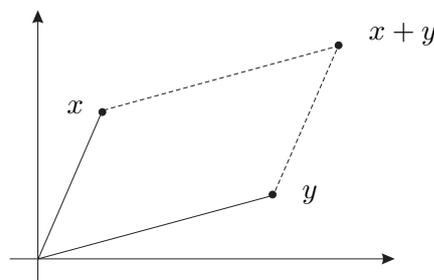
Wir wollen auf die Menge \mathbb{R}^2 noch andere Struktur aufprägen: Je zwei Elementen $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ können wir durch komponentenweise Addition ihre Summe

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

zuordnen. Wir nennen dies *Vektoraddition* oder einfach *Addition*.

Bemerkung 1.2.1. Doppelpunkte $:=$ werden immer anzeigen, dass ein Ausdruck auf der linken Seite durch den Ausdruck auf der rechten Seite definiert wird. Das Gleichheitszeichen $=$ dagegen macht eine Aussage über schon definierte Größen. (Wenn wir einfach eine Variable definieren benutzen wir in der Regel kein “:=”.)

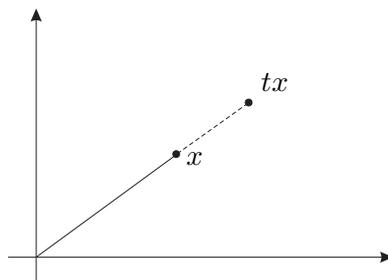
Wir stellen den Vektor $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ durch einen Pfeil vom Ursprung mit Spitze im Punkt mit Koordinaten (x_1, x_2) dar. Die Summe wird dann so veranschaulicht:



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 und bezeichnen dies als *Skalarmultiplikation*. Bildlich für $t > 0$:



Wie sieht das Bild für $t < 0$ aus?

In konkreten Beispielen sieht das so aus:

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

Es wird in der Vorlesung zahlreiche Beispiele für die neu eingeführten Konzepte geben, aber oft bleibt es auch Ihnen überlassen, sich Beispiele zu überlegen. Daher empfehle ich (wie fast alle meiner Kolleg*innen) mathematische Literatur immer mit Papier und Stift zur Hand zu lesen, so dass Sie einfache Beispiele ergänzen oder veranschaulichende Bilder selbst zeichnen können.

Bemerkung 1.2.2. Es gilt für alle $x, y, z \in \mathbb{R}^2$ und für alle $t, t' \in \mathbb{R}$:

1. $(x + y) + z = x + (y + z)$. [Assoziativität]
2. 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. Die Terme in mathematische Formeln erschließen sich oft nur aus dem Kontext.

3. 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]

4. $x + y = y + x$. [Kommutativität]

Die ersten vier Rechenregeln entsprechen genau denen für die Addition von reellen Zahlen.

5. $(tt')x = t(t'x)$. (Hier bezeichnet tt' die Multiplikation in \mathbb{R} und $t'x$ die Skalarmultiplikation, also komponentenweise Multiplikation in \mathbb{R}^2 !)
6. $1x = x$
7. $t(x + y) = tx + ty$
8. $(t + t')x = tx + t'x$. (Überlegen Sie sich genau, welche Verknüpfung bei $t + t'$ und welche bei $tx + t'x$ gemeint ist!)

All diese 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 im Skript oder in der Vorlesung gelegentlich übersprungen, aber fragen Sie ruhig nach, wenn etwas unklar ist!

Diese Rechnungen zeigen die Aussagen für alle möglichen Werte von x, y, z, t, t' . Zum Beispiel hängt die Aussage (a) von den drei Elementen $x, y, z \in \mathbb{R}^2$ ab, aber für jede Wahl der Elemente ist die Rechnung gültig und die Aussage wahr.

Bemerkung 1.2.3. Statt \mathbb{R}^2 können (und werden) wir allgemeiner für beliebiges $n \in \mathbb{N}$ die Menge aller geordneten n -Tupel reeller Zahlen \mathbb{R}^n . Insbesondere ist \mathbb{R}^1 einfach die Zahlengerade der reellen Zahlen und \mathbb{R}^3 ist ein Modell für den dreidimensionalen Raum. (Was ist \mathbb{R}^0 ?) Wir nennen auch ein Element $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ einen *Vektor* (in \mathbb{R}^n) und für $i = 1, 2, \dots, n$ die reelle Zahl v_i die i -te *Komponente* oder *Koordinate*.

Der \mathbb{R}^n ist zentrales Beispiel eines *Vektorraums* – die genaue Definition folgt später. \mathbb{R}^n tritt in vielen Anwendungen auf; unsere Methoden werden so beschaffen sein, dass sie nicht vom Wert von n abhängen, und auch nicht von der Wahl der reellen Zahlen als Komponenten. Dies ist ein Beispiel für Abstraktion in der Mathematik, die sehr oft zu größerer Anwendbarkeit in konkreten Problemen führt.

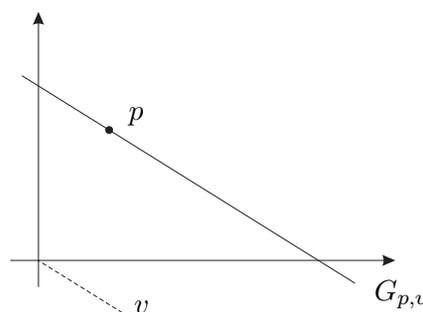
Durch *Definitionen* schafft man sich neue mathematische Begriffe. \mathbb{R}^2 ist eine Menge und hat eine Klasse interessanter Teilmengen. Die folgenden Überlegungen funktionieren für allgemeines \mathbb{R}^n genauso gut wie für \mathbb{R}^2 , also schreiben wir einfach \mathbb{R}^n .

Definition 1.2.4. Seien $p, v \in \mathbb{R}^n, v \neq 0$. Dies definiert für jedes $t \in \mathbb{R}$ einen Vektor $p + tv \in \mathbb{R}^n$. Dann heißt die Teilmenge $G_{p,v}$ all dieser $p + tv \in \mathbb{R}^n$ für verschiedene $t \in \mathbb{R}$ die (affine) *Gerade* durch den *Fußpunkt* p mit *Richtungsvektor* v . Wir schreiben kurz

$$G_{p,v} := \{p + tv \mid t \in \mathbb{R}\} \subset \mathbb{R}^n$$

oder

$$G_{p,v} = p + \mathbb{R}v.$$



In einer Definition heben wir die Begriffe hervor, die definiert werden.

Der Ausdruck $\{p + tv \mid t \in \mathbb{R}\}$ ist so zu lesen: wir betrachten die Teilmenge von \mathbb{R}^n , die aus den Elementen besteht, für die es ein $t \in \mathbb{R}$ gibt, so dass sich das Element als $p + tv$ schreiben lässt.

Den Ausdruck $p + \mathbb{R}v$ verstehen wir nur als verkürzte Schreibweise, wir definieren keine neue Additionsoperation.

Diese Darstellung einer Gerade heißt *Parameterdarstellung*, wir werden später andere Darstellungen treffen.

Beispiel 1.2.5. $G_{0,(1,1)}$ ist die Winkelhalbierende des ersten und dritten Quadranten.

Unsere Definition verlangt $v \neq 0$, und Sie sollten sich fragen, warum $v = 0$ nicht zugelassen wurde! (Wie sähe denn $G_{p,0}$ aus?) Jede Bedingung in einer mathematischen Definition oder Aussage hat eine Bedeutung.

Wir beweisen nun unsere erste Aussage, genauer gesagt beweisen wir ein *Lemma*, also eine Hilfsaussage.

Lemma 1.2.6. Seien $v, w, p, q \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$. Es gilt $G_{p,v} = G_{q,w}$ genau dann, wenn $q \in G_{p,v}$ und es ein $s \in \mathbb{R}$ gibt mit $w = sv$.

Beachten Sie, dass der Skalar s nicht 0 sein kann, denn sonst wäre $w = sv = 0 \cdot v = 0$, und wir nehmen ja gerade an, dass $w \neq 0$.

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.

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.

Dies ist der erste Beweis in unserem Kurs, und wir werden entsprechend sehr sorgfältig sein. In diesem Beweis geschieht einiges und wir werden schon einige Techniken treffen, die später eine Rolle spielen.

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” es gilt $G_{p,v} = G_{q,w}$, das heißt $G_{p,v}$ und $G_{q,w}$ enthalten die gleichen Vektoren in \mathbb{R}^n .

“Aussage 2” es ist $q \in G_{p,v}$ und es gibt ein $s \in \mathbb{R} \setminus \{0\}$ mit $w = sv$.

- Für eine gegebene Wahl von $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ ist jede der Aussagen entweder wahr oder falsch. 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 nur 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 schreiben dann

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

und sagen, die beiden Aussagen sind *äquivalent*.

Dazu zeigen wir (für jede v, w, p, q wie oben) zweierlei: zum einen: ist Aussage 1 wahr, dann ist auch Aussage 2 wahr. Man schreibt dann ‘

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

Zum anderen zeigen wir: ist Aussage 2 wahr, so ist auch Aussage 1 wahr. Von den vier möglichen Kombinationen

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.

⟨⟨Es erscheint, dass wir nun mehr Arbeit haben, weil wir zwei statt einer Aussage zeigen wollen. Aber jede der beiden Richtungen ist viel leichter, als direkt eine Äquivalenz zeigen zu wollen.⟩⟩

" \Rightarrow ". Es soll also für eine gewisse Wahl von $v, w \in \mathbb{R}^n$ mit $v \neq 0$ und $w \neq 0$ und $p, q \in \mathbb{R}^n$ Aussage 1 gelten. 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 $q \in G_{p,w}$. Also gibt es nach Definition von $G_{p,v}$ ein $s_1 \in \mathbb{R}$ mit $q = p + s_1v$. ⟨⟨Die Definition besagt, dass es eine Zahl gibt, die die Gleichung erfüllt, und wir geben dieser Zahl sofort einen Namen!⟩⟩

Ferner gilt auch $q + w \in G_{q,w} = G_{p,v}$, also gibt auch es ein $s_2 \in \mathbb{R}$ mit $q + w = p + s_2v$. Es folgt

$$w = (q + w) - q = (s_2 - s_1)v .$$

Beachten Sie: 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 Null sind. Damit 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.

" \Leftarrow " folgt am Mittwoch.

" \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.

Zwei Teilmengen T_1, T_2 einer Menge M , in Zeichen $T_1 \subset M$ und $T_2 \subset M$, sind genau dann gleich, wenn sie die gleichen Elemente von M enthalten. ⟨⟨Das ist die Definition von Gleichheit von Teilmengen einer Menge, und es entspricht unserer Anschauung.⟩⟩ Ist jedes Element von T_1 auch in T_2 , so gilt $T_1 \subset T_2$. 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 haben wir $q \in G_{p,v}$. Also gibt es ein $t_0 \in \mathbb{R}$, so dass $q = p + t_0v$ gilt. Sei $r \in G_{q,w}$ beliebig, also gilt $r = q + t_1w$ mit einem geeigneten $t_1 \in \mathbb{R}$.² Einsetzen liefert

$$r = p + t_0v + t_1sv = p + (t_0 + t_1s)v ,$$

woraus $r \in G_{p,v}$ folgt. Wir hatten über r keine weiteren Annahmen gemacht, als dass $s \in G_{q,w}$ gilt. Dann ist aber auch $r \in G_{p,v}$. 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 $u \in G_{p,v}$ beliebig. Nach Definition können wir also $u = p + t_2v$ schreiben mit einem geeigneten $t_2 \in \mathbb{R}$.

Wir haben immer noch $q = p + t_0v \in G_{p,v}$. Durch Umstellen folgt $p = q - t_0v$ mit $t_0 \in \mathbb{R}$. Weiterhin gilt $s \neq 0$. Denn wenn wir annehmen, dass $s = 0$ dann folgt aus $w = sv$ sofort $w = 0v = 0$. Aber wir betrachten ja nur $w \neq 0$ in diesem Lemma, dies ist also ein *Widerspruch*

²Das ist mathematischer Slang: "mit einem geeigneten..." ist eine Existenzaussage und gleichbedeutend mit "es gibt ein ...".

und die Annahme $s = 0$ muss falsch sein. $\langle\langle$ Dieses Argument ist ein einfaches Beispiel für einen *Widerspruchsbeweis* oder *indirekten Beweis*. $\rangle\rangle$ Also

$$s = p + t_2v = q - t_0v + t_2v = q + (t_2 - t_0)s^{-1}w.$$

Man beachte, dass wir hier die Annahme $s \neq 0$ ausgenutzt haben. Hieraus folgt $u \in G_{q,w}$. Jedes $u \in G_{p,v}$ liegt also auch in $G_{q,w}$ und wir haben die Inklusion $G_{p,v} \subset G_{q,w}$ gezeigt.

Zusammen mit der Inklusion $G_{q,w} \subset G_{p,v}$ folgt dass $G_{q,w} = G_{p,v}$. \square

Lemma 1.2.7. 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: Aus Lemma 1.2.6 folgt, dass wir einen beliebigen Punkt auf G , also insbesondere a , als Fußpunkt wählen können. (Denn G hat die Form $G_{p,v}$ für geeignete Vektoren p, v und wenn a in $G_{p,v}$ ist dann gilt $G_{p,v} = G_{a,v}$.)

Es ist also $G = G_{a,v}$ mit einem geeigneten Richtungsvektor $v \in \mathbb{R}^n \setminus \{0\}$, den wir noch nicht kennen. Aus $b \in G_{a,v}$ folgt, dass es $t_0 \in \mathbb{R}$ gibt mit $b = a + t_0v$.

Damit ist $b - a = t_0v$. Weiterhin gilt $b - a \neq 0$ nach Annahme, also können wir Lemma 1.2.6 anwenden und es gilt $G_{a,v} = G_{a,b-a}$. \square

Man kann Geraden auch anders darstellen:

Lemma 1.2.8. Sei $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2$ und sei $a_1x_1 + a_2x_2 = c$ eine lineare Gleichung in zwei Variablen. Wenn $a \neq 0$ dann ist die Menge aller Lösungen $L_{a,c} = \{(x_1, x_2) \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = c\}$ eine Gerade.

Beweis: Für den Vektor a bedeutet $a \neq 0$ dass $a_1 \neq 0$ oder $a_2 \neq 0$. Wir können annehmen, dass $a_1 \neq 0$, denn der Beweis für $a_2 \neq 0$ geht genauso, nur ein paar Indizes sind anders. $\langle\langle$ Wir sagen auch dass wir *ohne Beschränkung der Allgemeinheit* (o.B.d.A.) annehmen, dass $a_1 \neq 0$. Führen Sie den Beweis ruhig als Übung für den Fall $a_2 \neq 0!$ $\rangle\rangle$

Dann ist $b = \begin{pmatrix} \frac{c}{a_1} \\ 0 \end{pmatrix}$ ein Punkt in der Lösungsmenge $L_{a,c}$. Wir betrachten weiterhin den Vektor $v = \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$. $\langle\langle$ Wieso betrachten wir diesen Vektor? Wir könnten die Lösungsmenge unserer Gleichung in Beispielen bestimmen und versuchen, die Parametrisierung zu erraten. Oder wir könnten einen beliebigen Vektor herannehmen und dann sehen, ob unsere weiteren Rechnungen die Koordinaten einschränken. In einem mathematischen Beweis werden die Überlegungen, die zu solchen Annahmen führen oft ausgelassen, denn der Beweis funktioniert auch wenn die Form vom Himmel fällt. Ich werde aber oft versuchen, solche Ideen in Beweisen zu begründen. $\rangle\rangle$

Wir wollen zeigen $G_{b,v} = L_{a,c}$. Zuerst gilt dass $b + tv$ die Gleichung erfüllt, womit wir meinen, dass $(b_1 + tv_1, b_2 + tv_2)$ die Gleichung erfüllt. Wir rechnen nämlich: $a_1(b_1 + tv_1) + a_2(b_2 + tv_2) = a_1(\frac{c}{a_1} + ta_2) + a_2(0 + t(-a_1)) = c + ta_1a_2 - ta_1a_2 = c$. Also gilt $G_{b,v} \subset L_{a,c}$.

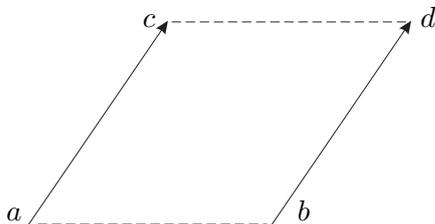
Wenn umgekehrt x die Gleichung erfüllt, dann betrachten wir $x - b$. Es gilt $x_1 = \frac{c - a_2x_2}{a_1}$ und damit $(x - b)_1 = \frac{c - a_2x_2}{a_1} - \frac{c}{a_1} = -\frac{a_2}{a_1}x_2$. Außerdem haben wir $(x - b)_2 = x_2 - 0$. Damit ist $x - b = \frac{x_2}{a_1} \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix} = \frac{x_2}{a_1}v$. Damit liegt $x \in G_{b,v}$ und da $x \in L_{a,c}$ beliebig war ist $L_{a,c} \subset G_{b,v}$. \square

Wir betrachten nun zwei einfache geometrische Anwendungen.

Definition 1.2.9. Gegeben zwei Punkte $a, b \in \mathbb{R}^2$, so heißt der Punkt $\frac{1}{2}(a+b)$ *Mittelpunkt* von a und b .

Der Mittelpunkt m erfüllt dass $m - a = b - m$ und damit ist $b - a = 2m$. So erklärt sich der Name.

Definition 1.2.10. Ein *Parallelogramm* ist ein 4-Tupel (a, b, c, d) von Punkten in \mathbb{R}^2 , so dass $c - a = d - b$ gilt:



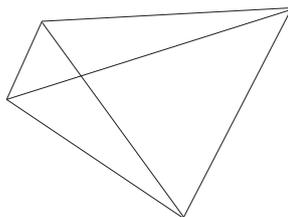
Ein Parallelogramm heißt *nicht-ausgeartet*, falls keine drei Punkte auf einer Geraden liegen.

In einem Parallelogramm gilt immer auch $b - a = d - c$. (Rechnen Sie das nach und veranschaulichen Sie es sich geometrisch.)

Satz 1.2.11 (*Diagonalensatz*). *In einem Parallelogramm treffen sich die Diagonalen in ihren Mittelpunkten.*

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

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



⟨⟨Im allgemeinen ist es eine gute Idee sich bei einem mathematischen Satz zu überlegen, wozu alle Annahmen gebraucht werden und sich gegebenenfalls ein Gegenbeispiel zu überlegen, falls die Annahmen verletzt sind. Das ist manchmal recht einfach, manchmal aber auch sehr kompliziert! Ich empfehle, es immer zu versuchen, aber nicht zu frustriert zu werden, wenn es nicht gelingt.⟩⟩

Auch für ausgeartete Parallelogramme treffen sich die Diagonalen in ihren Mittelpunkten, aber nicht nur dort!

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

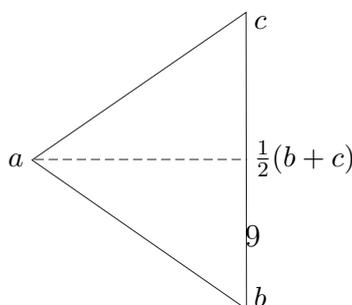
Wir rechnen:

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

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

Definition 1.2.12. 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 1.2.13 (*Schwerpunktsatz*). In einem nicht-ausgearteten Dreieck (a, b, c) schneiden sich die Seitenhalbierenden genau in dem Punkt $\frac{1}{3}(a + b + c)$, dem Schwerpunkt des Dreiecks.

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

Beweis: Die Seitenhalbierende durch a enthält a und den Seitenmittelpunkt $\frac{b+c}{2}$. Nach Lemma 1.2.6 ist sie in Parameterform wegen Lemma 1.2.7 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.

Wir müssen noch zeigen, dass es genau einen Schnittpunkt gibt. Wenn sich die Seitenhalbierenden in zwei Punkten schneiden, dann liegen alle Seitenhalbierenden nach Lemma 1.2.7 auf einer Geraden, und damit liegen alle Eckpunkte auf einer Geraden und das Dreieck ist ausgeartet, im Widerspruch zu unserer Annahme.

Der Ausdruck $\frac{a+b+c}{3}$ erklärt auch die Benennung des Schnittpunkts als Schwerpunkt. \square

1.3 Lineare Gleichungssysteme, Gauß'scher Algorithmus

Ein einzelnes lineare Gleichung ist einfach zu lösen, in zahllosen Anwendungen möchten wir Werte finden, die gleichzeitig mehrere lineare Gleichungen in mehreren Variablen lösen.

Definition 1.3.1. 1. Ein (reelles) *lineares Gleichungssystem* (LGS) ist ein System von Gleichungen der Form

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

mit gegebenen $a_{ij} \in \mathbb{R}$ und $b_i \in \mathbb{R}$. Gesucht sind alle reelle Lösungen x_1, \dots, x_n , also reelle Zahlen, die alle Gleichungen gleichzeitig erfüllen.

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

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

3. Wir nennen die rechteckige Anordnung reeller Zahlen

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

die *Koeffizientenmatrix* des linearen Gleichungssystems. Allgemein ist jede rechteckige Anordnung von Zahlen eine Matrix, wir werden in diesem Kurs noch viel Zeit mit ihnen verbringen.

In einer Matrix heißen die horizontalen Abschnitte $(a_{i1} \cdots a_{in})$ *Zeilen* und die vertikalen

Abschnitte $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ *Spalten*.

Die reellen Zahlen auf der rechten Seite fassen wir zu dem Vektor $b = (b_1 \dots, b_m) \in \mathbb{R}^m$ zusammen. Die Matrix

$$(A, b) := \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{mn} & \cdots & a_{mn} & b_m \end{array} \right)$$

heißt *erweiterte Koeffizientenmatrix* des inhomogenen linearen Gleichungssystems.

4. Die Lösungsmenge des Gleichungssystems bezeichnen wir $\text{Lsg}(A, b)$. Dies ist eine Teilmenge von \mathbb{R}^n .

Für Matrizen von gewisser Form ist der Lösungsraum einfach zu bestimmen:

- Definition 1.3.2.** 1. Eine Matrix A ist in *Zeilenstufenform*, falls für alle $i = 2, \dots, m$ 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 $k = 1, \dots, n$.
2. Eine Matrix ist in *spezieller Zeilenstufenform*, wenn sie in Zeilenstufenform ist und falls für alle $i = 1 \dots m$ gilt: ist $a_{i1} = a_{i2} = \dots = a_{i,k-1} = 0$ und $a_{ik} \neq 0$, so ist $a_{ik} = 1$. In Worten: der erste Eintrag ungleich 0 in jeder Zeile ist 1.

Wenn Sie eine neue abstrakte Definition treffen empfehle ich immer zuerst, sich Beispiele und Nichtbeispiele zu suchen:

Beispiele 1.3.3. • Matrizen in spezieller Zeilenstufenform: $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\begin{pmatrix} 1 & 5 & 9 \\ 0 & 0 & 0 \end{pmatrix}$

• Matrizen in Zeilenstufenform, aber nicht spezieller Zeilenstufenform: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

• Matrizen, die nicht in Zeilenstufenform sind:

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

Wir zeigen nun, wie man ein inhomogenes lineares Gleichungssystem löst, dessen Koeffizientenmatrix in Zeilenstufenform ist. Wir demonstrieren den *Algorithmus*, also das Rechenrezept, an einem Beispiel. Später im Kurs erfolgt eine allgemeine Betrachtung.

Beispiel 1.3.4. 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

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

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

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

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

Nicht jedes Gleichungssystem ist lösbar:

Bemerkung 1.3.5. Sei A eine Matrix in Zeilenstufenform. Wenn es einen Index $i \in \{1, \dots, m\}$ gibt, so dass $a_{ij} = 0$ für alle j , aber $b_i \neq 0$ gilt, dann ist die Lösungsmenge leer.

Denn dann ist die i -te Gleichung

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

was offensichtlich keine Lösung hat.

Wir wollen nun jedes lineare Gleichungssystem in Zeilenstufenform überführen. Dafür brauchen wir Umformungen unseres Gleichungssystems, die die Lösungsmenge nicht verändern.

Satz 1.3.6. *Es gibt die folgenden elementaren Zeilenumformungen:*

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 (\tilde{A}, \tilde{b}) aus (A, b) durch eine Folge elementarer Zeilenumformungen, so ändert sich die Lösungsmenge nicht, $\text{Lsg}(\tilde{A}, \tilde{b}) = \text{Lsg}(A, b)$.

Beweis:. Es kommt offensichtlich nicht auf 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

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \tag{1}$$

$$a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kn}x_n = b_k \tag{2}$$

und

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \tag{3}$$

$$(a_{k1} + \lambda a_{i1})x_1 + (a_{k2} + \lambda a_{i2})x_2 + \dots + (a_{kn} + \lambda a_{in})x_n = b_k + \lambda b_i \tag{4}$$

gleiche Lösungsräume haben. Erfüllt $x = (x_1, \dots, x_n)$ das erste Gleichungssystem, so auch die erste Gleichung des zweiten Systems. Wenn x Gleichung (2) erfüllt erfüllt es auch das λ -fache der Gleichung, und dann erfüllt x Gleichung (4) als Summe von zwei wahren Gleichungen.

Umgekehrt subtrahiert man das λ -fache von (3) von (4) und sieht, dass jede Lösung des zweiten Gleichungssystems auch (2) erfüllt (und (1) sowieso). \square

Betrachtung 1.3.7. 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 a)-c) auf die Untermatrix an, die durch Streichung der ersten Zeile entsteht.

Bemerkung 1.3.8. Wir können den Algorithmus etwas variieren, zum Beispiel haben wir in der Vorlesung in 2. direkt ein Vielfaches der ersten Zeile subtrahiert, anstatt die Zeilen erst zu skalieren.

Beispiel 1.3.9. Wir betrachten das inhomogene lineare Gleichungssystem von drei Gleichungen in drei Variablen mit erweiterter Koeffizientenmatrix

$\begin{pmatrix} 0 & 1 & 1 & | & 1 \\ 5 & 10 & -20 & | & 5 \\ 2 & 8 & 4 & | & 14 \end{pmatrix}$. Vertauschen

der ersten beiden Zeilen liefert das äquivalente System $\begin{pmatrix} 5 & 10 & -20 & | & 5 \\ 0 & 1 & 1 & | & 1 \\ 2 & 8 & 4 & | & 14 \end{pmatrix}$. Wir teilen die

erste Zeile durch 5 und die zweite 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 1.3.6, dass die beiden inhomogenen linearen Gleichungssysteme

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

die gleichen Lösungsmengen haben. Das rechte System lösen wir 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 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 1.3.6 in Zeilenstufenform (\tilde{A}, \tilde{b}) .

- Löse das lineare Gleichungssystem $\tilde{A}x = \tilde{b}$ in Zeilenstufenform sukzessive von unten nach oben. Wenn dabei eine unlösbare Gleichung auftritt ist das System nicht lösbar. Wenn dabei Variablen unbestimmt bleiben, dann parametrisieren sie eine unendliche Lösungsmenge.

1.4 Aussagen und Logik

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 1.4.1. Unter einer *Aussage* A verstehen wir ein sprachliches Gebilde, von dem es sinnvoll ist, zu fragen, ob es wahr (w) oder falsch (f) ist.

Beispiele 1.4.2. 1. Die Aussage: “Die Geraden G und G' im \mathbb{R}^n schneiden sich” ist entweder wahr oder falsch.

2. Die Aussage “Es gibt Tafeln im Hörsaal H2” hat Wahrheitswert w .

3. Die Aussage $3 \cdot 4 = 4$ hat Wahrheitswert f .

4. Die Ausdrücke “ $5 + 7$ ”, “Moin” und “Wie heißen Sie?” sind keine Aussagen.

5. 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 1.4.3. Für $n \in \{1, 2, 3, \dots\}$ ist eine n -stellige *Verknüpfung* von gegebenen Aussagen A_1, A_2, \dots, A_n eine Aussage $V(A_1, \dots, A_n)$, deren Wahrheitswert durch die Wahrheitswerte der gegebenen Aussagen A_1, \dots, 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”.

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

Dies ist für viele die verwirrendste Wahrheitstafel, deshalb ein paar Kommentare. Es ist einleuchtend, dass gilt, dass aus einer wahren Aussage eine wahre Aussage folgt, aber nicht gilt, dass aus einer wahren Aussage eine falsche Aussage folgt. (Das wäre verheerend.)

Aber wenn wir mit einer falschen Aussage beginnen, dann ist jedes Schlussfolgerung richtig. (Ex falso quodlibet.) “Wenn meine Oma Räder hat, ist sie ein Omnibus.” ist eine wahre Aussage ist, obwohl meine Großmutter keine Räder hat. Die Aussage “Wenn es regnet, ist die Straße nass.” wäre 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.

Es ist außerdem wichtig zu beachten, dass die Implikation nur von den Wahrheitswerten von A und B abhängt, nicht davon, ob irgendeine Kausalität oder sonst ein Zusammenhang besteht. Also: “Wenn St. Pauli deutscher Fußballmeister ist, gibt es unendlich viele Primzahlen.” oder “Wenn Bayern deutscher Fußballmeister ist, gibt es unendlich viele Primzahlen.” sind beide 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 1.2.6.

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

A	$\neg A$
w	f
f	w

Bemerkung 1.4.4. 1. Aus einer Wahrheitstafel folgt sofort, dass $A \vee \neg A$ immer wahr ist. Es gilt also der *Satz vom ausgeschlossenen Dritten*. Das ist mathematisch sehr nützlich und ermöglicht insbesondere Widerspruchsbeweise.

2. Manchmal sind auch die Symbole \top und \perp nützlich. Es hat \top immer den Wahrheitswert “wahr” und \perp das immer den Wahrheitswert “falsch”.

Wir können verschiedene Verknüpfungen miteinander verknüpfen und vereinbaren die Reihenfolge, ähnlich wie “Punkt vor Strich” in der normalen Arithmetik, \neg vor \wedge vor \vee vor \Rightarrow vor \Leftrightarrow .

Beispiel 1.4.5. 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.

Beispiel 1.4.6. Aus Wahrheitstafeln folgt auch, dass $A \Leftrightarrow B$ den gleichen Wahrheitswert hat wie $(A \Leftarrow B) \wedge (B \Rightarrow A)$. Das haben wir schon benutzt!

Definition 1.4.7. Gegeben seien mehrere Aussagen A, B, C, \dots 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, \dots den Wahrheitswert w annimmt, so sagt man, X und Y sind *logisch gleichwertig*. Die Aussage $X \Leftrightarrow Y$ heißt dann eine *Tautologie*.

Satz 1.4.8. 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 1.4.9. 1. Die Tautologie 1.4.8.6 liegt der Beweistechnik des “indirekten Beweises” zugrunde. Wir wollen zeigen, dass mit unserer Annahme A auch die Aussage B gilt. Nehmen wir also das Gegenteil von B an und zeigen, dass dann A nicht gilt. Da wir aber A annehmen, ist das ein Widerspruch, und B muss gelten.

Man beachte, dass sich die Richtung der Implikation umkehrt. Wir haben dies im Beweis von Lemma 1.2.7 gesehen. Um indirekte Beweise zu führen, sollte man also die relevanten Aussagen sorgfältig verneinen.

- Man kann alle n -stelligen Verknüpfungen als verschachtelte Verknüpfung dieser elementaren Verknüpfungen darstellen. Dafür reichen sogar die Negation \neg sowie die Konjunktion \wedge .

Wir wollen abschließend noch Aussagen betrachten, deren Wahrheitswert von einem Element (oder mehreren Elementen) einer gewissen Menge M abhängt. Solche Aussagen waren schon im Beweis von Lemma 1.2.6 aufgetreten. Sei zum Beispiel M die Menge aller Giraffen im Tierpark Hageback und für $x \in M$ die Aussage $C(x)$ “ x ist mehr als 4 Meter hoch”. Wir werden uns im nächsten Abschnitt näher mit dem Mengenbegriff befassen.

Definition 1.4.10. Eine *Aussageform* oder *Prädikat* ist eine Vorschrift P die jedem Element x der Menge M eine Aussage $P(x)$ und damit einen Wahrheitswert zuweist. Wir haben also eine Aussage für jedes $x \in M$. Wenn wir den Wahrheitswert der Aussagen betrachten erhalten wir eine Abbildung $P : M \rightarrow \{w, f\}$.

Bemerkungen 1.4.11. 1. Zum Beispiel sei $M = \mathbb{N}$ die Menge der natürlichen Zahlen $\{0, 1, 2, \dots\}$. Betrachte das Prädikat P , das $n \in \mathbb{N}$ die Aussage “Die Zahl n ist eine Primzahl.” zuordnet. $P(3)$ ist wahr und $P(6)$ ist falsch.

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

Seien M, N Mengen und seien $A(x), B(x)$ und $C(x, y)$ Prädikate, also Aussagen, deren Wahrheitswert davon abhängt, welche Elemente $x \in M$ bzw. $y \in N$ man einsetzt. Dann betrachten wir die folgenden Aussagen:

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

Man nennt auch \forall den *Allquantor* und \exists den *Existenzquantor*.

Der Doppelpunkt gehört zur Aussage dazu, wird aber gelegentlich weggelassen.

Sie können sich die Quantoren informell als unendliche Quantoren vorstellen: $\forall n \in \mathbb{N} : A(n)$ heißt $A(0) \wedge A(1) \wedge A(2) \dots$ und $\exists n \in \mathbb{N} : B(n)$ heißt $B(0) \vee B(1) \vee B(2) \dots$

Eine andere informelle Interpretation von $\forall x \in M : A(x)$ ist dass für alle x gilt $x \in M \Rightarrow A(x)$. Wir müssen aber immer über eine Menge quantifizieren, Aussagen über alle x sind gefährlich, das sehen wir im nächsten Abschnitt über Mengen.

Es gelten wieder “logische Rechenregeln”, insbesondere:

- $\neg(\forall x \in M : A(x)) \Leftrightarrow (\exists x \in M : \neg A(x))$.
- $\neg(\exists x \in M : A(x)) \Leftrightarrow (\forall x \in M : \neg A(x))$.

Die Verneinung der Aussage “Alle Schafe sind schwarz.” ist eben nicht “Kein einziges Schaf ist schwarz.” sondern “Es gibt (wenigstens) ein nicht-schwarzes Schaf.” Das bedeutet insbesondere: Wenn Sie eine Aussage der Form $\forall x : P(x)$ widerlegen wollen, benötigen Sie (nur) ein Gegenbeispiel.

Aussagen 1. und 2. sind äquivalent zueinander und äquivalent dazu, dass der Quantor $\forall x \in M : A(x)$ gleichbedeutend ist mit $\neg \exists x \in M : \neg A(x)$. In formaler Prädikatenlogik (die wir hier nicht betreiben), kann man dies als Definition von \exists verwenden.

- $\forall x \in M : \forall y \in N : C(x, y) \Leftrightarrow \forall y \in N : \forall x \in M : C(x, y)$

$$4. \exists x \in M : \exists y \in N : C(x, y) \Leftrightarrow \exists y \in N : \exists x \in M : C(x, y)$$

Gleiche Quantoren können wir also vertauschen und schreiben dann z.B. $\forall x, y \in M$.

$$5. \exists x \in M : \forall y \in N : C(x, y) \Rightarrow \forall y' \in N : \exists x' \in M : C(x', y') .$$

Die Reihenfolge von unterschiedlichen Quantoren ist wichtig! Wenn es ein y gibt so dass für alle x gilt $C(x, y)$ dann gibt es auch für alle x' ein y' (nämlich y !), sodass $C(x', y')$ erfüllt ist.

Aber es gilt nicht " \Leftarrow ": Um zu zeigen, dass dies nicht gilt müssen wir ein Beispiel finden, wo die rechte Seite wahr ist, aber die linke falsch, denn $\neg(A \Leftarrow B)$ ist äquivalent zu $\neg(\neg B \vee A) \Leftrightarrow B \wedge \neg A$.

Sei nun \mathbb{N} die Menge aller natürlichen Zahlen und $C(x, y)$ die Aussage $x > y$. Dann sagt die rechte Seite, dass es für jede Zahl eine größere Zahl gibt (wahr), aber die linke Seite, dass es eine Zahl gibt, die größer als alle anderen ist (falsch).

$$6. \left(\forall x \in M : A(x) \right) \wedge \left(\forall x \in M : B(x) \right) \Leftrightarrow \forall x \in M : A(x) \wedge B(x).$$

$$7. \exists x \in M : A(x) \vee B(x) \Leftrightarrow \left(\exists x \in M : A(x) \right) \vee \left(\exists x \in M : B(x) \right).$$

Man kann sich leicht vergewissern, dass Aussage 6. stimmt: Angenommen die rechte Seite gilt, dann gilt für alle $x \in M$ die Aussage $A(x) \wedge B(x)$, dann gilt auch für alle x $A(x)$, und damit gilt der erste Teil der linken Seite. Mit dem gleichen Argument gilt auch $\forall x \in M : B(x)$.

Sei umgekehrt $x \in M$ beliebig und wir nehmen die linke Seite an. Es gilt dann $x \in A(x)$ und es gilt $x \in B(x)$, was wir zeigen wollten.

Aussage 7. folgt genauso, wir können Sie aber auch aus Aussage 6. herleiten, indem wir beide Seite negieren und bedenken, dass $\neg\forall$ das gleiche ist wie $\exists\neg$.

$$8. \left(\forall x \in M : A(x) \right) \vee \left(\forall x \in M : B(x) \right) \Rightarrow \forall x \in M : A(x) \vee B(x).$$

$$9. \exists x \in M : A(x) \wedge B(x) \Rightarrow \left(\exists x \in M : A(x) \right) \wedge \left(\exists x \in M : B(x) \right)$$

Sie können sich vergewissern, dass die Richtung \Rightarrow jeweils gilt. (Es ist auch 9. äquivalent zu 8. indem wir das Kontrapositionsgesetz verwenden.)

Wir machen uns klar, warum " \Leftarrow " jeweils nicht gilt: sei M die Menge aller lebenden Menschen. Sei $A(x)$ die Aussage: " x mag Marmite." und $B(x)$ die Aussage: " x mag kein Marmite." Dann gilt die rechte Seite von 8. weil jeder Mensch Marmite mag oder nicht;³ die linke Seite würde aber nur gelten, wenn entweder alle Menschen Marmite mögen oder niemand, es gibt aber sowohl solche als auch solche.

Ähnlich gilt die rechte Seite von 9. weil mindestens ein Mensch Marmite mag und mindestens ein Mensch Marmite nicht mag. Die linke Seite würden nur gelten, wenn es einen Menschen gäbe, der Marmite gleichzeitig mag und nicht mag, was widersprüchlich ist.

Bemerkung 1.4.12. Was passiert mit unseren Quantoren, wenn M die leere Menge ist? $\exists x \in \emptyset : A(x)$ ist immer falsch, da \emptyset keine Elemente hat.

Dagegen ist $\forall x \in \emptyset : A(x)$ immer wahr, denn die Bedingung ist trivialerweise erfüllt: Die Aussage "wenn $x \in \emptyset$, dann $A(x)$ " stimmt, da die linke Seite falsch ist.

³Marmite ist Hefeextrakt, den man zum Beispiel aufs Frühstücksbrot schmieren kann. Wir arbeiten in einem vereinfachten Modell der Wirklichkeit, in dem niemand indifferent gegenüber Marmite ist. Anekdotisch ist dieses Modell nicht unrealistisch.

Etwas anschaulicher wird dies mit der Verneinung: Um zu zeigen dass $\forall x \in \emptyset : A(x)$ falsch ist, müssten wir ein $x \in \emptyset$ finden, das $A(x)$ nicht erfüllt. Aber es gibt gar kein $x \in \emptyset$.

Also: “Alle runden Dreiecke sind blau.” “Aber es gibt doch gar keine runden Dreiecke.” “Eben.”

Beispiel 1.4.13. Hier ist ein Beispiel zur Formalisierung einer Aussage mit Quantoren: “Jede zwei Geraden in \mathbb{R}^2 schneiden sich in genau einem Punkt.” Dafür legen wir zuerst M als die Menge aller Geraden in \mathbb{R}^2 fest.

$$\forall G, G' \in M : \exists x \in \mathbb{R}^2 : x \in G \wedge x \in G' \wedge (\forall y \in \mathbb{R}^2 : (y \in G \wedge y \in G') \Rightarrow x = y)$$

Aber diese Aussage ist falsch! Betrachten wir also die Verneinung

$$\neg(\forall G, G' \in M : \exists x \in \mathbb{R}^2 : x \in G \wedge x \in G' \wedge (\forall y \in \mathbb{R}^2 : (y \in G \wedge y \in G') \Rightarrow x = y))$$

die mithilfe von 1. und 2. äquivalent ist zu

$$\exists G, G' \in M : \forall x \in \mathbb{R}^2 : \neg(x \in G \wedge x \in G' \wedge (\forall y \in \mathbb{R}^2 : (y \in G \wedge y \in G') \Rightarrow x = y))$$

und jetzt verwenden wir unsere Rechenregeln aus Satz 1.4.8

$$\exists G, G' \in M : \forall x \in \mathbb{R}^2 : x \notin G \vee x \notin G' \vee \neg(\forall y \in \mathbb{R}^2 : y \in G \wedge y \in G' \Rightarrow x = y))$$

und noch einmal 1.

$$\exists p, q \in \mathbb{R}^2 : \exists v, w \in \mathbb{R}^2 \setminus \{0\} : \forall x \in \mathbb{R}^2 : x \notin G_{p,v} \vee x \notin G_{q,w} \vee \exists y \in \mathbb{R}^2 : \neg(y \in G_{p,v} \wedge y \in G_{q,w} \Rightarrow x = y)$$

und schließlich verwenden wir die Verneinung von $A \Rightarrow B$: $\neg(A \rightarrow B) \Leftrightarrow \neg(\neg A \vee B) \Leftrightarrow A \wedge \neg B$.

$$\exists p, q \in \mathbb{R}^2 : \exists v, w \in \mathbb{R}^2 \setminus \{0\} : \forall x \in \mathbb{R}^2 : x \notin G_{p,v} \vee x \notin G_{q,w} \vee \exists y \in \mathbb{R}^2 : y \in G_{p,v} \wedge y \in G_{q,w} \wedge x \neq y)$$

Wir interpretieren die Verneinung: Es gibt zwei Geraden in \mathbb{R}^2 , die sich entweder nicht schneiden oder in zwei verschiedenen Punkten schneiden.

Bemerkung 1.4.14. 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 $P(x)$.

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 $P(x)$.
- 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.

1.5 Mengen

“Definition” 1.5.1 (Cantor). “Eine *Menge* ist eine Zusammenfassung bestimmter, wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen.”

Dazu ein paar Anmerkungen

- Eine Menge besteht aus *Elementen* – den “Objekten unserer Anschauung oder unseres Denkens.” Das können erst einmal alle mögliche Dinge sein, insbesondere auch andere Mengen!
- Die Elemente sind wohlunterschieden, d.h. sie es kommt nicht mehrmals das gleiche Element in einer Menge vor. Wir vereinbaren daher $\{x, x, y\} = \{x, y\}$. Auch kommt es nicht auf die Reihenfolge der Elemente an, also $\{y, x\} = \{x, y\}$.
- Für jedes solche Objekt lässt sich entscheiden, ob es einer Menge ist oder nicht, und die Menge (“das Ganze”) ist durch ihre Elemente eindeutig charakterisiert.
- Es war schon Cantor klar, dass es eine leere Menge geben sollte, die kein Element enthält.

Wir formulieren unsere Definition etwas moderner zu unserer “Arbeitsdefinition”.

“Definition” 1.5.2. 1. Eine Menge ist etwas, das Elemente enthalten kann. Für ein Objekt a schreiben wir $a \in M$ wenn a Element von M ist. Für die Negation, dass a kein Element der Menge M ist, schreibt man $a \notin M$. Für eine endliche Menge M , die genau die endlich vielen Elemente a_1, a_2, \dots, a_n enthält, schreibt man auch $M = \{a_1, a_2, \dots, a_n\}$.

2. Es gibt eine ausgezeichnete Menge, die leere Menge \emptyset , die keine Elemente enthält.
3. Eine Menge ist durch ihre Elemente eindeutig bestimmt. Zwei Mengen M, N sind gleich, genau dann, wenn sie die gleichen Elemente enthalten, also wenn $a \in M \Leftrightarrow a \in N$. Man schreibt dann $M = N$.

Insbesondere ist die Ordnung der Elemente in einer Menge, oder eine mögliche Wiederholung von Elementen unerheblich.

Bemerkung 1.5.3. In diesem Kurs betreiben wir “naive Mengenlehre”, das heißt wir verwenden einen anschaulichen Begriff dessen, was eine Menge ist. Systematischer ist die sogenannte “axiomatische Mengenlehre”, in der die gesamte Mengenlehre auf einer Liste von Axiomen aufbaut, die Mengen erfüllen sollen. Insbesondere ist in diesem Axiomensystem nur die Existenz von Mengen garantiert, die sich aus den Axiomen ableiten lassen und keine informell beschriebenen Mengen wie die Menge aller Mengen oder die Menge aller Studierenden in diesem Raum.

Die Notwendigkeit für axiomatische Mengenlehre ergibt sich aus der *Russelschen Antinomie*: sie betrachtet die (sogenannte) Menge R aller Mengen x , die sich nicht selbst als Element enthalten, $R = \{x \mid x \notin x\}$ und fragt, ob R sich selbst als Element enthält. Jede Antwort führt zum Widerspruch.

Der Vollständigkeit halber gebe ich die üblichen Axiome der Mengenlehre. (Sorgen Sie sich nicht zu sehr, wenn einige der Axiome schwer zugänglich sind, wir werden sie größtenteils nicht mehr gebrauchen.)

Mengen und die Relation \in sind genau jene Objekte, die die folgenden Axiome erfüllen:

“Definition” 1.5.4 (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.

Das ist der entscheidende Grundsatz, dass die Menge durch ihre Elemente vollständig bestimmt ist. $\langle\langle$ Da die einzigen Objekte in der Theorie Mengen sind ist ein Element a einer Menge wieder eine Menge! $\rangle\rangle$

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

Beispiel: Wir haben schon die leere Menge und durch wiederholte Anwendung des Paarungssaxioms konstruieren wir die Mengen $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\} \dots$

4. Vereinigungssaxiom: Zu jeder Menge M von Mengen gibt es eine Menge X , die genau die Elemente der Elemente von M als Elemente enthält. Falls $M = \{A_1, A_2, \dots, A_N\}$ schreiben wir $X = A_1 \cup A_2 \cup \dots \cup A_n$.

Achtung: Das ist kein Tippfehler! Da die Elemente aller Mengen hier selbst Mengen sind ergibt es Sinn, die Elemente der Elemente zu betrachten.

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

Warnung: Dies ist genau genommen nicht ein einzelnes Axiom sondern eine Axiomschema, das für jedes Prädikat ein Axiom gibt.

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

Beispiel. Zum Beispiel ist die Potenzmenge von $\{\emptyset\}$ die Menge $\{\emptyset, \{\emptyset\}\}$.

8. Ersetzungsaxiom: Ist A eine Menge und f eine definierbare Funktion dann formen die Bilder aller Elemente von A wieder eine Menge.

Definierbare Funktionen lassen sich wieder über Prädikate definieren. Formal lautet unser Ersetzungsaxiom: Für jedes Prädikat $E(x, y)$, in dem die Variable M nicht vorkommt, gilt $\forall x, y, z : (E(x, y) \wedge E(x, z) \Rightarrow y = z) \Rightarrow \forall A : \exists M : \forall y : (y \in M \iff \exists x : (x \in A \wedge E(x, y)))$

Die erste Bedingung bedeutet das y von $E(x, y)$ eindeutig bestimmt ist, das Prädikat E definiert also eine Funktion.

Dieses Axiom ist wieder ein Axiomschema. Es enthält das Aussonderungssaxiom als Spezialfall.

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

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.

Warnung: Dieses Axiom hat einen besonderen Stellenwert, es scheint umstrittener zu sein, als andere. Es bedeutet, dass wir aus einer beliebigen Menge von Mengen gleichzeitig je ein Element auswählen können. Wenn diese Mengen keine Ordnung und keine bevorzugten Elemente haben, dann ist diese Aussage möglicherweise nicht offensichtlich.

Bemerkung 1.5.5. Die Zermelo-Fraenkel-Axiome schließen die Russellsche 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.

Wir werden die meisten Zermelo-Fraenkel Axiome im Verlauf der Vorlesung nicht explizit benutzen. Es wird meist ausreichen, zu wissen, dass eine Menge nie Element von sich selbst sein kann, und die grundlegenden Konstruktionen mit Mengen zu beherrschen.

Insbesondere verwenden wir die Konstruktion und Schreibweise $\{a \in A \mid P(a)\}$ für alle Elemente in A , die P erfüllen aus dem Aussonderungsaxiom. Wir verwenden auch $\mathcal{P}(M)$, die Menge aller Teilmengen in M , aus dem Potenzmengenaxiom.

Wir arbeiten also mit unserer “Definition” 1.5.2: “eine Menge ist etwas, das Elemente enthält”, und wenden uns nun einigen grundlegenden Definitionen und Konstruktionen zu.

Definition 1.5.6. 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 $\{x\}$ nie mit dem Element x verwechseln: eine Schachtel mit einem Hut ist eben etwas anderes als ein Hut.

Zwei Mengen sind genau dann gleich wenn sie Teilmengen voneinander sind:

$$A = B \Leftrightarrow A \subset B \wedge B \subset A .$$

2. Elementare Operationen auf Mengen können wir aus der Verknüpfung von Aussagen definieren. Seien A, B Mengen.

(a) Die *Schnittmenge* oder der Durchschnitt

$$A \cap B = \{a \mid a \in A \wedge a \in B\} .$$

ist die Menge aller Elemente die in A und B enthalten sind.

Wir können auch mehr als zwei Mengen schneiden. Sei I eine beliebige Menge und sei für jedes $i \in I$ genau eine Menge A_i gegeben. Wir definieren die Schnittmenge $\bigcap_{i \in I} A_i$ als die Menge aller Elemente, die in jedem A_i enthalten sind, symbolisch

$$\bigcap_{i \in I} A_i = \{a \mid \forall i \in I : a \in A_i\} .$$

(b) Wir definieren die *Vereinigung* als

$$A \cup B = \{a \mid a \in A \vee a \in B\}.$$

Dies ist die Menge aller Elemente die in A oder B liegen. Wir definieren auch $\cup_{i \in I} A_i$ als die Menge aller Elemente, die in mindestens einem A_i enthalten sind, symbolisch

$$\cup_{i \in I} A_i = \{a \mid \exists i \in I : a \in A_i\}.$$

(c) Die *Mengendifferenz* ist

$$A \setminus B = \{a \mid a \in A \wedge a \notin B\}.$$

die Menge aller Elemente von A , die nicht in B sind. Zum Beispiel ist $\mathbb{Z} \setminus \mathbb{N} = \{-1, -2, -3, \dots\}$ die Menge aller negative Zahlen.

Satz 1.5.7. *Seien A, B, C Mengen. Dann gilt*

1. *Kommutativgesetz:* $A \cap B = B \cap A$ 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. *Distributivgesetz:* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Beweis: Alle Aussagen folgen aus den entsprechenden Aussagen in Satz 1.4.8. Wir führen dies am Beispiel des ersten Distributivgesetzes 1.5.7.2 vor:

$$\begin{aligned} x \in A \cap (B \cup C) &\Leftrightarrow x \in A \wedge x \in (B \cup C) \\ &\Leftrightarrow x \in A \wedge (x \in B \vee x \in C) \\ &\stackrel{1.4.8.4}{\Leftrightarrow} (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C) \\ &\Leftrightarrow (x \in A \cap B) \vee (x \in A \cap C) \\ &\Leftrightarrow x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

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, auch wenn sie eng verwandt sind..

Beispiel 1.5.8. Die Menge \mathbb{N} der natürlichen Zahlen (mit 0) ist eine der wichtigsten Mengen in der Mathematik.

Wir können ihre Elemente durch die *Peano-Axiome* charakterisieren.

1. 0 ist eine natürliche Zahl.
2. Jede natürliche Zahl n hat eine natürliche Zahl n' als Nachfolger.
3. 0 ist kein Nachfolger einer natürlichen Zahl.
4. Natürliche Zahlen mit gleichem Nachfolger sind gleich.
5. Prinzip der *vollständigen Induktion*:

Sei $M \subset \mathbb{N}$ eine Teilmenge mit den beiden Eigenschaften, dass M die Null enthält, $0 \in M$, und dass mit n auch der Nachfolger in M liegt, also $n \in M \Rightarrow n' \in M$. Dann ist $M = \mathbb{N}$.

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

Wir können die natürlichen Zahlen auch axiomatisch im Zermelo-Frenkel Axiomensystem konstruieren. Da wir Mengen dort aus anderen Mengen konstruieren, identifizieren wir die natürliche Zahl 0 mit \emptyset und für jede natürliche Zahl n ihren Nachfolger mit $n \cup \{n\}$. Das ist die von Neumannsche Zahlenreihe. Wir erhalten $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, \{\emptyset\}\}$, $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, ...

Bemerkung 1.5.9. Das Prinzip der vollständigen Induktion liefert ein *Definitionsverfahren*, indem wir einen Begriff für 0 definieren und für jeden Nachfolger n' definieren, wenn wir ihn schon für n definiert haben, dann ist der Begriff für alle natürlichen Zahlen definiert.

1. Die *Fakultät* $n!$ von n ist induktiv definiert durch $0! = 1$ und $(n')! = n' \cdot n!$.
2. Auch Addition und Multiplikation in \mathbb{N} müssen erst definiert werden!
 - Addition: $m + 0 := m$ und $m + n' := (m + n)'$.
 - Multiplikation: $m \cdot 0 := 0$ und $m \cdot n' := (m \cdot n) + m$

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

Man kann nun mit viel Geduld Kommutativität, Assoziativität und Distributivität von dieser Definition ausgehend beweisen.

Das Prinzip der vollständigen Induktion liefert natürlich ein wichtiges *Beweisverfahren*, um eine Aussage der Form $\forall x \in \mathbb{N} : P(x)$ zu beweisen.

Dabei geht man wie folgt vor. Man zeigt zunächst, dass die Aussage für die Zahl $n = 0$ wahr ist.

Dies bezeichnet man als den *Induktionsanfang*. Danach zeigt man, dass aus $P(n)$ für eine Zahl n folgt, dass auch $P(n + 1)$ wahr ist. Dies nennt man den *Induktionsschritt*. Damit gilt die Aussage nach dem Prinzip der vollständigen Induktion für alle natürlichen Zahlen.

Man kann die Induktion auch statt 0 bei 1 beginnen. Dann zeigt man, dass die Aussage für alle $n \geq 1$ (wir schreiben auch \mathbb{N}^* für $\mathbb{N} \setminus \{0\}$).

Beispiel 1.5.10. Für alle natürlichen Zahlen $n \geq 1$ gilt die Aussage $1 + 3 + \dots + (2n - 1) = n^2$. (Die Aussage gilt mit der richtigen Definition auch für $n = 0$, aber der Einfachheit halber beginnen wir unsere Induktion bei 1 und beweisen unsere Aussage nur für $n \geq 1$.)

1. Induktionsanfang: Die Aussage gilt für $n = 1$, denn $1 = 1^2$. (Die Aussage gilt auch schon für $n = 0$.)
2. Induktionsschritt: Angenommen die Aussage gilt für die natürliche Zahl n . Dann ergibt sich für die Zahl $n + 1$:

$$\begin{aligned} 1 + 3 + \dots + (2n - 1) + (2(n + 1) - 1) &= (1 + 3 + \dots + (2n - 1)) + (2n + 1) \\ &= n^2 + 2n + 1 = (n + 1)^2 \end{aligned}$$

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 die Aussage 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 $n \in \mathbb{N}^*$ wahr ist.

Wir können im Induktionsschritt nicht nur $A(n-1)$ verwenden, sondern $A(k)$ für jedes $k < n$. Das heißt manchmal *starke Induktion*: Wenn für jedes n aus $\forall k < n : A(k)$ folgt, dass $A(n)$ gilt, dann gilt A für alle natürlichen Zahlen.

Der Induktionsanfang ist hier genau genommen ein Spezialfall des Induktionsschrittes: Sei $n = 0$, dann müssen wir zeigen $\forall k < 0 : A(k) \Rightarrow A(0)$. Aber die linke Seite ist tautologisch immer wahr, das heißt wir müssen $A(0)$ zeigen.

Wenn wir die Induktion mit einem Widerspruchsbeweis kombinieren erhalten wir die Beweistechnik des *kleinsten Gegenbeispiels*.

Wir verwenden also die Kontraposition der starken Induktion. Statt

$$(\forall k < n : A(k)) \Rightarrow A(n)$$

ist es äquivalent zu zeigen

$$\neg A(n) \Rightarrow (\exists k < n : \neg A(k)),$$

siehe Satz 1.4.8.6 und die Regeln zur Negation von Quantoren.

Das heißt, wenn wir für jedes Gegenbeispiel n , für das die Aussage $A(n)$ nicht gilt, ein kleineres Gegenbeispiel $k < n$ konstruieren können, so dass $A(k)$ auch nicht gilt, dann haben wir unsere Aussage für alle n gezeigt!

Anders ausgedrückt: Wir nehmen an, dass m das *kleinste* Gegenbeispiel ist, also gilt $A(k)$ für alle $k < m$. Aber dann gilt nach der starken Induktion auch $A(m)$, wir haben einen Widerspruch.

All diese Techniken (Induktion, starke Induktion, kleinstes Gegenbeispiel) sind logisch äquivalent, aber oft macht eine der Techniken die Beweisfindung einfacher.

Beispiel 1.5.11. Einige von Ihnen kennen das Beispiel aus dem Tutorium: Wir wollen zeigen, dass jede natürliche Zahl ein Produkt von endlich vielen Primfaktoren ist.

Angenommen das gilt nicht, dann gibt es eine kleinste Zahl $n > 1$, die nicht Produkt von Primfaktoren ist. Aber n kann nicht prim sein, sonst wäre $n = n$ ein Produkt von Primfaktoren. Wir können also $n = ab$ schreiben mit a, b natürliche Zahlen kleiner als n .

Da n das kleinste Gegenbeispiel war sind a und b Produkte von Primzahlen, $a = p_1 p_2 \cdots p_k$ und $b = q_1 q_2 \cdots q_\ell$. Aber dann gilt $n = p_1 p_2 \cdots p_k q_1 \cdots q_\ell$, und das ist in Widerspruch zu unserer Annahme.

1.6 Relationen und Abbildungen

Definition 1.6.1. Seien A_1, \dots, A_n Mengen. Dann ist das *kartesische Produkt* oder *direkte Produkt* $A_1 \times \dots \times A_n$ die Menge der *geordneten* n -Tupel mit Elementen in A_1, \dots, A_n , d.h.

$$A_1 \times \dots \times A_n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_1 \in A_1 \dots a_n \in A_n \right\}.$$

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

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

für die geordneten n -Tupel von Elementen in A .

Sie kennen das Konzept und die Schreibweise schon von $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$!

Beispiel 1.6.2. Die Menge aller Karten in einem Kartenspiel erhalten wir als kartesischen Produkt der Farben $\{\diamond, \spadesuit, \heartsuit, \clubsuit\}$ mit den Werten $\{2, 3, 4, 5, 6, 7, 8, 9, B, D, K, A\}$.

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.

Definition 1.6.3. 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:

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

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

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

Beispiele 1.6.4.

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

2. $X = \mathbb{R}$ und $x \sim y :\Leftrightarrow x \leq y$. Diese Relation ist reflexiv und transitiv, aber nicht symmetrisch (denn $1 \leq 2$, aber $2 \not\leq 1$), also keine Äquivalenzrelation.

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

4. 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. Für $m = 2$ formen die geraden und die ungeraden Zahlen jeweils eine Äquivalenzklasse.

Lemma 1.6.5. Ist R eine Äquivalenzrelation auf einer Menge X , so bilden die Äquivalenzklassen eine Partition von X , das heißt

- keine Äquivalenzklasse ist leer
- jedes a gehört zu einer Äquivalenzklasse.
- zwei beliebige Äquivalenzklassen A, A' sind entweder gleich oder disjunkt, es gilt also entweder $A = A'$ oder $A \cap A' = \emptyset$.

Beweis: Für beliebiges $a \in X$ definieren wir

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

Wir zeigen, dass dies eine Äquivalenzklasse ist.

- Wegen $a \in A_a$ ist sie nicht leer.
- Seien $x, y \in A_a$, so gilt $x \sim a$ und $y \sim a$, also wegen der Symmetrie auch $a \sim y$. Somit $x \sim y$ wegen Transitivität.
- Seien $x \in A_a$, $y \in X$ 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 $y \in A_a$.

Wir haben schon gezeigt, dass diese Äquivalenzklassen nicht leer sind, außerdem ist jedes $x \in X$ in A_x enthalten.

Es bleibt zu zeigen, dass zwei Äquivalenzklassen entweder gleich oder disjunkt sind. Angenommen $A \cap A' \neq \emptyset$ für zwei Äquivalenzklassen A, A' . 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 $A = A'$. \square

- Bemerkungen 1.6.6.**
1. 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 . $\langle\langle$ Verwechseln Sie das nicht mit der Mengendifferenz $X \setminus S!$ $\rangle\rangle$
 2. Es gibt eine *kanonische* (d.h. ausgezeichnete, ohne willkürliche Auswahl bestimmte) *Abbildung*, die jedem Element von X die Äquivalenzklasse zuweist, die es enthält.

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

3. Jedes $a \in A$ heißt ein *Repräsentant* der Äquivalenzklasse A . Im allgemeinen gibt es aber keine ausgezeichneten Repräsentanten und keine kanonische Abbildung $X/R \rightarrow X$.

Beispiele 1.6.7.

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. Wenn Sie einen Stundenplan erstellen wollen, dann arbeiten Sie lieber mit der Quotientenmenge als mit der Menge aller Schüler!

Die kanonische Abbildung ordnet einem Schüler seine Klasse zu. Sie wird bei der Beschriftung von Schulheften oft benutzt. Der Klassensprecher oder die Klassensprecherin ist ein Repräsentant der Klasse, aber es sind andere Repräsentanten denkbar, zum Beispiel der oder die erste im Alphabet.

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 $[z] := \{z + kn \mid k \in \mathbb{Z}\}$. Es gibt n Restklassen mit Repräsentanten $0, 1, \dots, n-1$. Die kanonische Abbildung ordnet einer Zahl die Restklasse ihres Rests nach Division durch n zu, etwa für $n = 12$ haben wir $23 \mapsto [11]$ – das kennen Sie von der Uhr!

3. Rationale Zahlen werden als Äquivalenzklassen auf der Menge $M := \{(m, n) \mid m, n \in \mathbb{Z}, n \neq 0\}$ 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, $\frac{a}{b} = [(1, 2), (2, 4), (-3, -6), \dots]$. Gekürzte Brüche mit positivem Nenner bilden ein System von Repräsentanten für die Äquivalenzklasse. Die rationalen Zahlen sind die Quotientenmenge.

Wir haben Abbildungen schon verwendet, jetzt können wir sie formal definieren:

Definition 1.6.8. 1. Seien A, B Mengen. Eine *Abbildung* oder *Funktion* 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 bezeichnen dieses b mit $f(a)$ und betrachten es als *Bild* von a .

Wir schreiben auch $f : A \rightarrow B$ oder $A \xrightarrow{f} B$ und $a \mapsto f(a)$.

Die Menge A heißt *Definitionsbereich* und B *Bildbereich* der Abbildung. Die Mengen A und B gehören zur Definition einer Abbildung.

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

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

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

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

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

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

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

definierte Abbildung.

Bemerkung 1.6.9. Formal definieren wir eine Abbildung also durch ihren *Graph* $\{(a, f(a)) \mid a \in A\} \subset A \times B$.

Der folgende Satz ist sehr einfach zu beweisen, aber sehr wichtig:

Satz 1.6.10. Die Verkettung von Abbildungen ist assoziativ ist: sei $h : C \rightarrow D$ eine weitere Abbildung, dann gilt

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

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

$$\begin{aligned} (h \circ g) \circ f(a) &= (h \circ g)(f(a)) = h(g(f(a))) \\ h \circ (g \circ f)(a) &= h((g \circ f)(a)) = h(g(f(a))) . \end{aligned} \quad \square$$

Außerdem sind Bildbereich und Definitionsbereich für $(h \circ g) \circ f$ und $h \circ (g \circ f)$ gleich.

Bemerkungen 1.6.11. 1. Eine besonders wichtige Abbildung ist $f = \text{id}_A : A \rightarrow A$ mit $a \mapsto a$, die *Identität* von A . Ihr Graph ist die Diagonale in $\Delta \subset A \times A$, i.e. $\Delta = \{(a, a) \mid a \in A\}$. Es gilt $f \circ \text{id} = f = \text{id} \circ f$. (Dies ist die einzige Relation, die sowohl Äquivalenzrelation als auch Funktion ist.)

2. Für gegebene Mengen M, N bilden die Abbildungen $f : M \rightarrow N$ eine Menge $\text{Abb}(M, N)$, denn sie sind eine Teilmenge der Potenzmenge $\mathcal{P}(M \times N)$, die aus allen Teilmengen von $M \times N$ besteht.

3. Zwei Vektoren $p \in \mathbb{R}^2$ und $v \in \mathbb{R}^2 \setminus \{0\}$ definieren eine Funktion der Funktion $f : \mathbb{R} \rightarrow \mathbb{R}^2$ mit $f(t) = p + tv$. Die Gerade $G_{p,v} \subset \mathbb{R}^2$ ist das Bild von f . Für verschiedene Werte von p und v bekommen wir möglicherweise das gleiche Bild, aber verschiedene Funktionen.

4. Eine Abbildung kann durch eine Rechenvorschrift gegeben sein, etwa $f : \mathbb{Z} \rightarrow \mathbb{N}$ mit $x \mapsto x^2$, aber auch durch eine Fallunterscheidung, etwa

$$f : \mathbb{R} \rightarrow \mathbb{N} \quad f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \notin \mathbb{Q} \end{cases} \quad \text{oder auch durch eine Tafel.}$$

Zum Beispiel gibt es eine Abbildung, die jedem Tag des Jahres 2022 die Tageshöchsttemperatur in Hamburg zuordnet, auch wenn es keine Rechenvorschrift gibt.

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 $f : \mathbb{Q} \rightarrow \mathbb{Q}$, die eine Umkehrabbildung $g : \mathbb{Q} \rightarrow \mathbb{Q}$ 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 $\mathbb{Z} \rightarrow \mathbb{Z}$ hat aber keine Umkehrabbildung.

Definition 1.6.12. Sei $f : A \rightarrow B$ eine Abbildung und $C \subset A$ eine Teilmenge. Wir schreiben $\iota_C : C \rightarrow A$ für die Inklusionsabbildung $a \mapsto a$. Die *Einschränkung* von f auf C ist Abbildung $f \circ \iota_C : C \rightarrow B$.

Definition 1.6.13. 1. Eine Abbildung $f : A \rightarrow B$ 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 $f(A) = B$.

2. Eine Abbildung $f : A \rightarrow B$ 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 $f(a_1) \neq f(a_2)$.

3. Eine Abbildung heißt *bijektiv*, wenn sie surjektiv und injektiv ist.

Bemerkungen 1.6.14. 1. Für jede Menge M ist die Identitätsabbildung $\text{id}_M : M \rightarrow M$ bijektiv.

2. Für jede Teilmenge $A \subset M$ ist die Inklusionsabbildung $\iota_A : A \rightarrow M$ injektiv. Für jede Äquivalenzrelation R ist die kanonische Abbildung $A \rightarrow A/R$ surjektiv.

3. Ist M eine endliche Menge, so gilt für alle Abbildungen $f : M \rightarrow M$: f bijektiv $\Leftrightarrow f$ injektiv $\Leftrightarrow f$ surjektiv. Ist aber M eine unendliche Menge, so gibt es Abbildungen $f : M \rightarrow M$, die injektiv, aber nicht surjektiv sind, und Abbildungen, die surjektiv, aber nicht injektiv sind. Betrachte zum Beispiel die Abbildungen $f : \mathbb{N} \rightarrow \mathbb{N}$ und $g : \mathbb{N} \rightarrow \mathbb{N}$ mit $f(n) = 2n$ und $g(m) = \lfloor \frac{m}{2} \rfloor$. (Die Symbole bedeuten, dass wir $\frac{m}{2}$ abrunden.) Dann ist f injektiv aber nicht surjektiv und g surjektiv, aber nicht injektiv.

Lemma 1.6.15. Eine bijektive Abbildung $f : A \rightarrow B$ hat eine eindeutige Umkehrabbildung oder inverse Funktion f^{-1} , die $f \circ f^{-1} = \text{id}_B$ und $f^{-1} \circ f = \text{id}_A$ erfüllt.

Umgekehrt ist jede Abbildung mit Umkehrabbildung bijektiv.

Beweis: Wir definieren f^{-1} wie folgt. Sei $b \in B$. Da f surjektiv ist gibt es mindestens ein a mit $f(a) = b$. Da f injektiv ist, gibt es nur ein solches a . Wir können also ohne Ambiguität $f^{-1}(b) = a$ definieren. Nach Konstruktion gilt $f(f^{-1}(b)) = b$. Es gilt auch $f^{-1}(f(a)) = a$, da ja $f(a) = f(a)$ erfüllt.

Wir zeigen Eindeutigkeit. Sei $g : B \rightarrow A$ eine weitere Umkehrabbildung. Es gilt $g(b) = g(f(f^{-1}(b))) = f^{-1}(b)$ wegen der Assoziativität, Satz 1.6.10.

Sei nun f eine Abbildung mit Umkehrabbildung f^{-1} . Es gilt f ist injektiv da aus $f(a) = f(a')$ folgt $f^{-1}(f(a)) = f^{-1}(f(a')) \Leftrightarrow a = a'$. Des weiteren ist f surjektiv, da $b = f(f^{-1}(b))$. \square

Definition 1.6.16. 1. Eine Menge M heißt *endlich*, wenn es ein $n \in \mathbb{N}_0$ und eine Bijektion

$$f : \{0, 1, \dots, n-1\} \rightarrow M$$

gibt. Die linke Seite ist formal definiert als Menge aller natürlichen Zahlen kleiner n . Für $n = 0$ erhalten wir die leere Menge!

2. Die natürliche Zahl

$$|M| := n$$

heißt die *Mächtigkeit* der Menge M . (Induktiv zeigt man, dass $n \in \mathbb{N}$ eindeutig bestimmt ist.)

3. Zwei Mengen (nicht unbedingt endlich) A, B heißen *gleichmächtig*, wenn es eine bijektive Abbildung $f : A \rightarrow B$ gibt.

Beispiele 1.6.17. 1. \mathbb{N} und \mathbb{Z} sind gleichmächtig, obgleich $\mathbb{N} \subsetneq \mathbb{Z}$. Eine Bijektion $f : \mathbb{N} \rightarrow \mathbb{Z}$ 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} (das lernen Sie in Analysis).

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

Bemerkung 1.6.18. Abbildungen sind eng mit dem kartesischen Produkt verbunden. Seien A, B und C Mengen. Dann gibt es eine Bijektion von Mengen von Abbildungen

$$F : \text{Abb}(A \times B, C) \rightarrow \text{Abb}(A, \text{Abb}(B, C))$$

mit Umkehrabbildung G . Hierbei wird $\phi : A \times B \rightarrow C$ auf die Abbildung $F(\phi) : A \rightarrow \text{Abb}(B, C)$ abgebildet, die $a \in A$ die Abbildung $\phi(a, -) : b \mapsto \phi(a, b)$ zuordnet.

Umgekehrt wird einer Abbildung $\psi : A \rightarrow \text{Abb}(B, C)$ die Abbildung $G(\psi) : A \times B \rightarrow C$ mit $G(\psi)(a, b) = \psi(a)(b)$ zugeordnet. Man rechne nach, dass man so eine Bijektion von Mengen von Abbildungen erhält.

2 Algebraische Grundbegriffe

Die zentralen Begriffe der linearen Algebra sind die Begriffe des Vektorraums und der linearen Abbildung. Der Vektorraum abstruiert die Beispiele \mathbb{R}^2 und \mathbb{R}^3 , die wir schon betrachtet haben. Wir werden aber zunächst einige algebraische Grundbegriffe einführen.

Wir können uns zum Beispiel erinnern, dass die ersten 4 Eigenschaften von \mathbb{R}^n in Bemerkung 1.2.2 nur mit der Addition und nicht mit der Skalarmultiplikation zu tun haben. Ist das nicht auch schon eine interessante Struktur?

2.1 Gruppen

Wir betrachten zunächst die 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

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

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 $e \cdot a = a = a \cdot e$

b) Für jedes $a \in G$ gibt es ein $a^{-1} \in G$, so dass $a^{-1} \cdot a = e = a \cdot a^{-1}$ gilt.

Man nennt ein solches Element e auch ein *neutrales Element* und a^{-1} ein *inverses Element* zu a . Wir werden bald sehen, dass a^{-1} eindeutig ist, weshalb die Benennung a^{-1} legitim ist.

⟨⟨In der Vorlesung habe ich das neutrale Element “Einheit” genannt, das ist aber nicht üblich! “Einheit” hat eine andere Bedeutung. Gelegentlich nennt man das neutrale Element “Einheitselement”.⟩⟩

Beispiele 2.1.2. 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. Die Menge $\{e\}$ mit Verknüpfung $e \cdot e = e$ ist eine Gruppe, die sogenannte *triviale Gruppe* die nur aus einem neutralen Element besteht.

7. Wir betrachten ein gleichseitiges Dreieck. Es hat die folgenden Symmetrien:

- (a) Drei Spiegelungen S_1, S_2, S_3 in den drei Mittelsenkrechten, numeriert im Uhrzeigersinn.
- (b) Zwei Rotationen R_1, R_2 um $\frac{2\pi}{3}$ bzw. $\frac{4\pi}{3}$.
- (c) Die triviale Symmetrie oder Identität E .

Dann bildet die Menge $\{E, R_1, R_2, S_1, S_2, S_3\}$ eine Gruppe, denn wir können zwei Symmetrien durch Hintereinanderausführen \circ verknüpfen, E ist ein neutrales Element, und jedes Element hat ein Inverses: $S_i^{-1} = S_i$ und $R_1^{-1} = R_2$.

Man kann zum Beispiel berechnen: $R_1 \circ R_1 = R_2$, $S_1 \circ S_2 = R_1$ und $S_2 \circ S_1 = R_2$. Die Elemente der Gruppe kommutieren also nicht!

8. Die Symmetrien jedes mathematischen Objekts bilden eine Gruppe!

Bemerkungen 2.1.3. 1. Aus dem Assoziativgesetz (G1) folgt, dass alle möglichen Klammerungen eines mehrfachen Produkts das gleiche Ergebnis liefern. Daher können wir Klammern in mehrfachen Produkten weglassen. Wir können sie aber auch setzen, um die Lesbarkeit zu erhöhen.

2. Wir lassen auch den Multiplikationspunkt oft weg, also ab für $a \cdot b$. Weiterhin schreiben wir a^2 für $aa = a \cdot a$ und allgemein a^i für $a \cdot a \cdots a$ (i Faktoren) für $i > 0$. Für $i > 0$ definieren wir auch $a^{-i} = a^{-1} \cdots a^{-1}$ (i Faktoren). Schließlich gilt $a^0 := e$.

Damit gilt für alle $a, b \in \mathbb{Z}$ $g^a \cdot g^b = g^{a+b}$.

- 3. Wegen (G2) gibt es wenigstens ein Element $e \in G$; die Menge G kann also nicht leer sein.
- 4. Es reicht in einer Gruppe die Existenz eines linksneutralen Elementes und eines Linksinversen zu fordern, also nur $e \cdot a = a$ und $a^{-1}a = e$. Ein linksneutrales Element erfüllt automatisch auch $a \cdot e = a$ (wenn es Linksinverse gibt!) und ein Linksinverses erfüllt automatisch $aa^{-1} = e$ (wenn e linksneutral ist).

Satz 2.1.4. Sei (G, \cdot) eine Gruppe. Dann gilt:

- 1. Das neutrale Element ist eindeutig.
- 2. Für jedes gegebene $a \in G$ gibt es ein eindeutiges Inverses Element a^{-1} wie in (G2b).

Beweis: • Sei \tilde{e} ein weiteres neutrales Element, also gelte für alle $a \in G$ die Gleichung $\tilde{e} \cdot a = a$. Setze $a = e$ und erhalte $\tilde{e} \cdot e = e$. Aber da e neutral ist gilt auch $\tilde{e} \cdot e = \tilde{e}$.

• Seien a' und a'' zwei Inverse für a . Es gilt

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

durch Anwendung der Gruppenaxiome. □

Satz 2.1.5. [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 muss gelten

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

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. Damit ist die Lösung eindeutig.

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

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

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 ein Inverses von a^{-1} ; also gilt $(a^{-1})^{-1}a^{-1} = e$ per Definition. Andererseits gilt $a \cdot a^{-1} = e$, so dass auch a ein Inverses von a^{-1} ist. Da das Inverse nach Satz 2.1.4 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 $a \cdot b$. 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). \square

Wir betrachten noch zwei neue Beispiele:

Beispiel 2.1.6. Wir wählen ein festes $m \in \mathbb{N}$. Betrachte auf der Menge \mathbb{Z} wie in Beispiel 1.6.4.4 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 $\mathbb{Z}/m\mathbb{Z}$ oder \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 $\{0, 1, \dots, m-1\}$ lassen. Eine Äquivalenzklasse wird auch *Restklasse modulo m* genannt.

Für jedes $0 \leq r \leq m-1$ ist

$$\bar{r} := r + m\mathbb{Z} = \{x \in \mathbb{Z} \mid m \text{ teilt } x - r\}$$

eine Restklasse.

Wir schreiben $a = a' \pmod{m}$ und sagen “ a ist kongruent a' modulo m ”, wenn a und a' in der gleichen Restklasse liegen, also wenn $m \mid a - a'$.

Für $m = 0$ ist die Äquivalenzrelation die Gleichheit und $\mathbb{Z}/0\mathbb{Z} \cong \mathbb{Z}$.

Nun ist $(\mathbb{Z}, +)$ auch eine abelsche Gruppe. Wir wollen auch auf der Quotientenmenge eine Gruppenstruktur definieren, die möglichst nah an der Gruppenstruktur in \mathbb{Z} ist.

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

$$\bar{a} + \bar{b} := \overline{a + b}$$

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 $\bar{a} = \bar{a}'$ und $\bar{b} = \bar{b}'$, 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 $\overline{a + b} = \overline{a' + b'}$.

Die Assoziativität der Addition folgt aus der Assoziativität der Addition in \mathbb{Z} :

$$(\bar{a} + \bar{b}) + \bar{c} = \overline{a + b} + \bar{c} = \overline{(a + b) + c} = \overline{a + (b + c)} = \bar{a} + \overline{b + c} = \bar{a} + (\bar{b} + \bar{c}).$$

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 $\bar{0}$, denn

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

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

Also ist $(\mathbb{Z}/m\mathbb{Z}, +)$ eine Gruppe, auch *zyklische Gruppe* genannt.

Beispiel 2.1.7. Sei $M \neq \emptyset$ eine nichtleere Menge. Setze

$$\text{Sym}(M) = \{f : M \rightarrow M \mid f \text{ bijektiv}\}$$

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

(G1) Die Komposition von Abbildungen ist assoziativ wegen Satz 1.6.10.

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

(G2b) Das zu $f \in \text{Sym}(M)$ inverse Element ist die Umkehrabbildung aus Lemma 1.6.15.

$(\text{Sym}(M), \circ)$ heißt die *symmetrische Gruppe* von M . Wir schreiben S_n für $\text{Sym}(\{1, 2, \dots, n\})$ und nennen S_n die symmetrische Gruppe von Grad n . Wir nennen die Elemente von $\text{Sym}(M)$ *Permutationen*.

Bemerkung 2.1.8. Die symmetrischen Gruppen haben große praktische und theoretische Bedeutung. Gruppen sind die mathematische Formalisierung von Symmetrie, immer wenn wir ein symmetrisches Objekt betrachten dann können wir auch die Gruppe der Symmetrien betrachten. S_n ist die Symmetriegruppe einer endlichen Menge, und da Mengen grundlegende mathematische Objekte sind, sind dies grundlegende Gruppen.

Bemerkenswerterweise tritt jede endliche Gruppe als Untergruppe (siehe Definition 2.1.13) einer symmetrischen Gruppe auf!

Definition 2.1.9. Eine Gruppe (G, \cdot) heißt *abelsch* oder *kommutativ*, falls für alle $a, b \in G$ gilt $a \cdot b = b \cdot a$.

Abelsche Gruppen werden oft additiv geschrieben, also $a + b$ für die Verknüpfung von a und b , und 0 für das neutrale Element.

Beispiele 2.1.10. 1. Die Beispiele in 2.1.2.1-6 sowie Beispiel 2.1.6 sind abelsche Gruppen.

2. Wir haben gesehen, das Beispiel 2.1.2.7 nicht abelsch ist.

3. Die symmetrische Gruppe ist im allgemeinen (für $n \geq 3$) nicht abelsch. Um dies zu testen, müssen wir ein Gegenbeispiel finden, also zwei Elemente f, g mit $f \circ g \neq g \circ f$. 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

$$\begin{aligned} f \circ g(1) &= f(g(1)) = f(1) = 2 \\ g \circ f(1) &= g(f(1)) = g(2) = 3, \end{aligned}$$

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

Definition 2.1.11. Sei G eine Gruppe. Wenn $|G|$ endlich ist, heißt G *endliche Gruppe* und $|G|$ die *Ordnung* von G .

Beispiel 2.1.12. Man sieht leicht, dass die Ordnung von $\mathbb{Z}/m\mathbb{Z}$ gleich m ist und die Ordnung von S_n gleich $n!$, denn die Anzahl der Permutationen einer Menge mit n Elementen ist $n!$.

Definition 2.1.13. Sei (G, \cdot) eine Gruppe. Eine Teilmenge $H \subset G$ heißt *Untergruppe*, falls sie den folgenden Axiomen genügt:

(UG1) $e \in H$.

(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.

Wir haben die folgenden wichtigen Eigenschaften.

Satz 2.1.14. 1. Auf einer Untergruppe H von G definiert die Multiplikation von G wieder eine (assoziative) Multiplikation und H selbst ist eine Gruppe.

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

3. Untergruppen abelscher Gruppen sind abelsch.

Beweis: In jedem Fall lassen sich die Axiome leicht prüfen und wir überspringen den Beweis. □

Beispiele 2.1.15. 1. Sei $(G, \cdot) = (\mathbb{R}, +)$ die additive Gruppe der reellen Zahlen. Dann haben wir eine Kette von Untergruppen $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

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.

- $H = \{e\}$ ist eine Untergruppe, die sogenannte *triviale Untergruppe*.
- $H = G$ ist eine Untergruppe, in Worten: Jede Gruppe ist Untergruppe von sich selbst.
- Jedes Element $g \in G$ erzeugt eine Untergruppe $\{e, g, g^{-1}, g^2, g^{-2}, g^3, \dots\}$ wobei wir die Notation aus Bemerkung 2.1.3.1 verwenden.
Die Elemente g^i müssen nicht alle unterschiedlich sein! Da $e = g^0$ gilt (UG1), mit $g^a g^b = g^{a+b}$ gilt (UG2) und mit $(g^i)^{-1} = g^{-i}$ gilt (UG3).

4. Die Rotationen bilden eine Untergruppe der Symmetriegruppe D_3 denn die Verknüpfung von zwei Rotationen ist wieder eine Rotation. Die Spiegelungen sind keine Untergruppe, denn die Verknüpfung von zwei verschiedenen Spiegelungen ist keine Spiegelung.

Die algebraische Struktur von Gruppen führt zu vielen nicht-trivialen Eigenschaften. Wir geben ein Beispiel:

Satz 2.1.16 (Satz von Lagrange). *Sei H eine Untergruppe von G und G endlich. Dann teilt $|H|$ die Ordnung von G .*

Beweis: Wir führen auf G die folgende Äquivalenzrelation ein: $a \sim b$ wenn $ab^{-1} \in H$. Man prüft leicht, dass dies eine Äquivalenzrelation ist: Reflexivität folgt aus $e \in H$, Symmetrie folgt, da H die Inverse für alle Elemente von H enthält, und Transitivität folgt, da H abgeschlossen unter Verknüpfung ist.

Wir bezeichnen die Äquivalenzklasse von a mit Ha , das ist sinnvoll denn $b \sim a$ genau wenn $ba^{-1} \in H$, aber dann gilt $b = ha$ für $h = ba^{-1}$. Also ist die Äquivalenzklasse von a genau die Menge $\{ha \mid h \in H\} = Ha \subset G$. Die Menge aller Äquivalenzklassen bezeichnen wir mit $H \backslash G$.

Wenn wir Repräsentanten a_1, \dots, a_m für jedes Element von $H \backslash G$ wählen gilt $G = \cup_{i=1}^m Ha_i$ und alle Ha_i sind disjunkt, siehe Lemma 1.6.5. Dann gilt $|G| = \sum_i |Ha_i|$.

Es reicht also zu zeigen, dass alle Ha_i die gleiche Anzahl von Elementen haben, nämlich $|H|$. Aber die Abbildung $f : H \rightarrow Ha$ gegeben durch $h \mapsto ha$ ist eine Bijektion für jedes $a \in G$. Eine Umkehrfunktion $Ha \rightarrow H$ ist durch $x \mapsto xa^{-1}$ gegeben.

Alternativ ist die Abbildung injektiv, dann aus $ha = h'a$ folgt $h = haa^{-1} = h'aa^{-1} = h'$ und surjektiv denn jedes Element in Ha hat ja per Definition die Form ha für $h \in H$. \square

2.2 Homomorphismen

Wir wollen als nächstes Abbildungen von Gruppen betrachten, die kompatibel mit der Gruppenstruktur sind.

Definition 2.2.1. Seien (G, \cdot) und $(H, *)$ Gruppen. Eine Abbildung

$$f : G \rightarrow H$$

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

$$f(a \cdot b) = f(a) * f(b).$$

Beispiele 2.2.2. 1. $G = H = \mathbb{Z}$ mit Addition ganzer Zahlen. Wähle ein festes $m \in \mathbb{Z}$ und betrachte die Abbildung

$$\mu_m : \mathbb{Z} \rightarrow \mathbb{Z}$$

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

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

2. Betrachte die Abbildung

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

die jeder ganzen Zahl a ihre Äquivalenzklasse

$$\bar{a} = a + m\mathbb{Z} = a \bmod m$$

zuordnet. Dies heißt auch die *kanonische Abbildung*, weil es die natürlichste Abbildung in die Quotientenmenge ist. Wir haben die Addition auf $\mathbb{Z}/m\mathbb{Z}$ gerade so definiert, dass can ein (surjektiver) Gruppenhomomorphismus ist: $\text{can}(a + b) = \overline{a + b} = \bar{a} + \bar{b}$.

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

$$f_\alpha : \mathbb{R} \rightarrow \mathbb{R}_{>0}$$

mit $f_\alpha(x) = e^{\alpha x}$. Dann gilt mit den Rechenregeln für Exponentialfunktionen:

$$f_\alpha(x + y) = e^{\alpha(x+y)} = e^{\alpha x} e^{\alpha y} = f_\alpha(x) f_\alpha(y) \quad .$$

4. Für eine beliebige Gruppe G und ein beliebiges Element $g \in G$ gibt es einen Homomorphismus $f_g : \mathbb{Z} \rightarrow G$ definiert durch $i \mapsto g^i$. Da $f_g(i) f_g(j) = g^i g^j = g^{i+j} = f_g(i + j)$, vgl. Bemerkung 2.1.3.2, ist dies tatsächlich ein Homomorphismus.

Satz 2.2.3. *Seien G und H Gruppen mit neutralen Elementen e_G bzw. e_H . Sei $f : G \rightarrow H$ ein Gruppenhomomorphismus. Dann gilt:*

1. $f(e_G) = e_H$

2. Für alle $g \in G$ gilt $f(g^{-1}) = f(g)^{-1}$.

3. Ist f bijektiv, so ist

$$f^{-1} : H \rightarrow G$$

ebenfalls ein Gruppenhomomorphismus.

Beweis:. 1. Wir rechnen zunächst: $f(e_G) = f(e_G \cdot e_G) = f(e_G) \cdot f(e_G)$. Daraus folgt

$$\begin{aligned} e_H &= f(e_G)^{-1} f(e_G) = f(e_G)^{-1} [f(e_G) f(e_G)] \\ &= [f(e_G)^{-1} f(e_G)] \cdot f(e_G) = e_H \cdot f(e_G) = f(e_G). \end{aligned}$$

2. Für alle $g \in G$ gilt

$$f(g^{-1}) f(g) = f(g^{-1} g) = f(e_G) \stackrel{1.}{=} e_H \quad .$$

Aus der Eindeutigkeit der Inversen $f(g)^{-1}$ von $f(g)$ folgt $f(g^{-1}) = f(g)^{-1}$.

3. Seien $h, h' \in H$. Wir rechnen, da f bijektiv ist

$$\begin{aligned} f^{-1}(h) \cdot f^{-1}(h') &= (f^{-1} \circ f) \left(f^{-1}(h) \cdot f^{-1}(h') \right) \\ &= f^{-1} \left(f \left(f^{-1}(h) \cdot f^{-1}(h') \right) \right) \\ &\stackrel{(*)}{=} f^{-1} \left(f \left(f^{-1}(h) \right) \cdot f \left(f^{-1}(h') \right) \right) \\ &= f^{-1}(h \cdot h') \quad .\square \end{aligned}$$

wobei wir für (*) verwendet haben, dass f ein Gruppenhomomorphismus ist. (Wir verwenden auch den typischen Trick, dass wir eine Identität durch $g^{-1} \circ g$ ersetzen, was wir dann weiter umformen können. “Alle wissen, dass $1 - 1 = 0$ ist, Mathematiker*innen wissen, dass $0 = 1 - 1$ ist”.)

Beispiele 2.2.4. Für den Gruppenhomomorphismus $\exp : \mathbb{R} \rightarrow \mathbb{R}_{>0}$, $\exp(x) = e^x$, die Exponentialfunktion aus Beispiel 2.2.2.2, bedeuten die Aussagen von Satz 2.2.3 folgendes:

1. $e^0 = 1$
2. $e^{-x} = (e^x)^{-1} = \frac{1}{e^x}$.
3. Die Umkehrfunktion

$$\exp^{-1} =: \log : \mathbb{R}_{>0} \rightarrow \mathbb{R}$$

ist ein Gruppenhomomorphismus, d.h. es gilt

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

Definition 2.2.5. Ein bijektiver Gruppenhomomorphismus heißt *Gruppenisomorphismus*. Zwei Gruppen G, H heißen *isomorph*, wenn es einen Gruppenisomorphismus $f : G \rightarrow H$ gibt; in Zeichen: $G \cong H$.

“Isomorph sein” ist eine Äquivalenzrelation für Gruppen: die Relation ist reflexiv, weil die Identität ein Isomorphismus ist.

Wegen Satz 2.2.3.3 ist die Umkehrfunktion eines Isomorphismus auch ein Isomorphismus und die Relation ist symmetrisch.

Sind $f : G \rightarrow H$ und $g : H \rightarrow K$ Gruppenisomorphismen, so ist auch $g \circ f : G \rightarrow K$ ein Gruppenisomorphismus (denn $g(f(a \cdot b)) = g(f(a)f(b)) = gf(a) \cdot gf(b)$ und die Verkettung von Bijektionen ist bijektiv); daher ist die Relation transitiv.

(Da es keine Menge aller Gruppen gibt (so wie es keine Menge aller Mengen gibt) ist es technisch nicht ganz korrekt von einer Äquivalenzrelation zu sprechen. Es gilt aber, dass der Ausdruck “ G ist isomorph zu H ” reflexiv, symmetrisch und transitiv ist.)

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

Beispiel 2.2.6. 1. Da $e^{\alpha x}$ für $\alpha \neq 0$ eine Bijektion ist sind $(\mathbb{R}, +)$ und $(\mathbb{R}_{>0}, \cdot)$ isomorph.

Allerdings liefert jede der Abbildungen f_α 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 Gruppe $(\{1, -1\}, \cdot)$ mit der üblichen Multiplikation ist isomorph zu $\mathbb{Z}/2\mathbb{Z}$ indem wir $1 \mapsto [0]$ und $-1 \mapsto [1]$ abbilden.
3. Jede Bijektion von Mengen M und N induziert einen Isomorphismus $\text{Sym}(M) \cong \text{Sym}(N)$. Zum Beispiel ist $\text{Sym}(\{a, b, c\}) \cong \text{Sym}(\{1, 2, 3\})$. Die Gruppen unterscheiden sich nur in den Namen der Elemente.
4. Jede Symmetrie des gleichseitigen Dreiecks in D_3 permutiert die drei Eckpunkte. Wenn wir zwei Symmetrien verknüpfen, dann verketteten wir auch die Abbildungen auf den Eckpunkten. Wir erhalten einen Homomorphismus $D_3 \rightarrow S_3$, der auch ein Isomorphismus ist.

5. Die Homomorphismen $\mu_m : \mathbb{Z} \rightarrow \mathbb{Z} \quad k \mapsto mk$ aus Beispiel 2.2.2.1 sind für jedes $m \neq 0$ injektiv, aber nur für $m = \pm 1$ ein Gruppenisomorphismen.

Isomorphe Gruppen teilen all ihre wichtigen mathematischen Eigenschaften. Oft behandelt man sie daher, als wären Sie gleich, wir benennen nur die Elemente anders, vgl. Beispiel 2.2.6.3. Aber unsere Intuition macht dennoch Unterscheidungen zwischen verschiedenen isomorphen Gruppen, vgl. Beispiel 2.2.6.1.

Wenn wir Homomorphismen zwischen Gruppen betrachten wollen, ist es oft dennoch sinnvoll, isomorphe Gruppen nicht zu identifizieren, sondern ihnen ihren eigenen Namen zu belassen.

Definition 2.2.7. Seien G und H Gruppen und sei $f : G \rightarrow H$ ein Gruppenhomomorphismus. Dann heißt das Urbild des neutralen Elements $e_H \in H$

$$\ker(f) := f^{-1}(e_H) = \{g \in G \mid f(g) = e_H\}$$

der *Kern* von f und $\text{Im } f := f(G)$ das *Bild* von f .

Satz 2.2.8. Sei $f : G \rightarrow H$ ein Gruppenhomomorphismus. Dann gilt:

1. $\ker(f)$ ist eine Untergruppe von G .
2. f ist injektiv $\Leftrightarrow \ker(f) = \{e_G\}$.
3. $\text{Im } f$ ist eine Untergruppe von H .
4. f ist surjektiv $\Leftrightarrow \text{Im } f = H$.

Beweis: 4. Klar nach der Definition von Surjektivität. □

2. “ \Rightarrow ” Sei f injektiv. Sei $g \in \ker f$, d.h. $f(g) = e_H = f(e_G)$. Die Injektivität von f impliziert $g = e_G$, also $\ker f = \{e_G\}$.

“ \Leftarrow ” Sei $\ker f = \{e_G\}$. Seien $g, g' \in G$ mit $f(g) = f(g')$. Zu zeigen ist $g = g'$. Wir rechnen

$$f(g(g')^{-1}) = f(g)f(g')^{-1} = f(g)f(g)^{-1} = e_H .$$

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

1. Wir überprüfen die Untergruppenaxiome aus Definition 2.1.13:

(UG1) Wegen $f(e_G) = e_H$ ist $e_G \in \ker f$.

(UG2) Seien $g, g' \in \ker f$. Dann folgt

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

also ist $gg' \in \ker f$.

(UG3) Sei $g \in \ker f$. Dann folgt

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

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

3. Wir überprüfen die Kurzversion der Untergruppenaxiome aus Übung 5.2: $\text{Im } f$ ist nicht leer, da $f(e_G) \in \text{Im } f$ und für $f(a), f(b) \in \text{Im } f$ gilt $f(a)f(b)^{-1} = f(ab^{-1}) \in \text{Im } f$ mit Satz 2.2.3.

Bemerkung 2.2.9. Ein Wort noch zur (möglicherweise verwirrenden) Namensgebung: Wir nennen allgemein die Funktion $\cdot : G \times G \rightarrow G$ einer Gruppe *Verknüpfung* und das Element e *neutrales Element* oder *Einselement*.

Gelegentlich behandeln wir die Verknüpfung wie die Multiplikation und schreiben ab für $a \cdot b$ und a^3 für $a \cdot a \cdot a$. Wir sagen, dass wir die Gruppe *multiplikativ schreiben*. Ich spreche dann auch gelegentlich (etwas ungenau) von *Multiplikation* in der Gruppe und nenne das neutrale Element die *Eins* und lese gh als “ g mal h ” statt “ g verknüpft mit h ”. Im Gegensatz zur Multiplikation in \mathbb{Q} oder \mathbb{R} nehmen wir aber nicht an, dass $a \cdot b = b \cdot a$ ist.

In einer abelschen Gruppe, in der $a \cdot b = b \cdot a$ gilt, schreiben wir die Verknüpfung oft additiv als $a + b$ und nennen das neutrale Element *Null*.

Wir nenne auch die Hintereinanderausführung von Abbildungen *Verknüpfung*, alternativ *Verkettung* oder *Komposition*. In der symmetrischen Gruppe zum Beispiel ist die Verknüpfung von Gruppenelementen durch die Verknüpfung von Abbildungen gegeben. Aber im allgemeinen bilden Abbildungen keine Gruppe und die Verknüpfung von Gruppenelementen lässt sich nicht als Verknüpfung von Abbildungen definieren! Wir bezeichnen also mit dem gleichen Wort zwei verschiedenen Konzepte und müssen aus dem Kontext entscheiden, welches gemeint ist.

2.3 Körper

Es kommt oft vor, dass eine Menge zwei Operationen hat. Wir abstrahieren zuerst die Eigenschaften von \mathbb{R} und \mathbb{Q} . Sie formen eine kommutative Gruppe unter Addition und beinahe eine additive Gruppe unter Multiplikation.

Definition 2.3.1. Ein *Körper* ist eine Menge K mit zwei assoziativen Verknüpfungen $+, \cdot$,

$$+, \cdot : K \times K \rightarrow K,$$

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

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

Beispiele 2.3.2. 1. $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper

2. Auch die ganzen Zahlen haben Addition und Multiplikation, aber $(\mathbb{Z}, +, \cdot)$ ist kein Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist.

3. Ein Körper K muss mindestens zwei unterschiedliche Elemente haben, 0 und $1 \in K \setminus \{0\}$. In der Tat gibt es einen Körper mit genau zwei Elementen, überlegen Sie sich die einzig möglichen Regeln für Addition und Multiplikation!

Wir werden bald zwei weitere wichtige Beispiele von Körpern treffen.

Bemerkungen 2.3.3. 1. Es folgt aus der Rechnung

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

dass $0 \cdot a = a \cdot 0 = 0$ für alle $a \in K$.

Da $0 \cdot a = 0$ für alle $a \in K$ kann 0 kein Inverses haben, wenn wir zwei Verknüpfungen mit Distributivgesetz haben. Es ist also notwendig, $K \setminus \{0\}$ zu betrachten um eine Gruppenstruktur für \cdot zu finden.

- Es ist automatisch, dass die Multiplikation Werte in $K \setminus \{0\}$ annimmt: Aus $ab = 0$ folgt immer $a = 0$ oder $b = 0$, denn wenn a und b Inverse haben ist $1 = a^{-1}abb^{-1} = a^{-1}0b^{-1}$, was ein Widerspruch hat. Wir sagen ein Körper hat keine *Nullteiler*.
- Wir nennen die Verknüpfung $+$ Addition und die Verknüpfung \cdot Multiplikation. Wir lassen auch oft den Punkt “.” 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

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

Satz 2.3.4. *Sei K ein Körper. Dann gilt für alle $a, b, c \in K$:*

- $(-1) \cdot a = -a$ und $(-1) \cdot (-1) = 1$.
- 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. Aus der Distributivität folgt $a + (-1)a = (1 - 1)a = 0$, also ist $(-1) \cdot a$ das additive Inverse von a .

Insbesondere ist $(-1) \cdot (-1)$ das additive Inverse von -1 und damit 1 wegen Satz 2.1.5.3.

- Da $a \neq 0$ gilt, gibt es ein multiplikatives Inverses a^{-1} . Wir rechnen wie im Beweis von Satz 2.1.5

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

□

Mit dem Satz können wir zum Beispiel rechnen $a(-b) = a(-1)b = (-1)ab = -ab$.

2.4 Die komplexen Zahlen

Wir führen den Körper der komplexen Zahlen ein. Er hat zahllose theoretische und praktische Anwendungen.

Wir wissen schon, dass $(\mathbb{R}^2, +)$ eine abelsche Gruppe ist und wollen eine Multiplikation definieren.

Für die Multiplikation kann man allerdings *nicht* die komponentenweise Multiplikation benutzen, um einen Körper zu erhalten. Denn es gilt für die komponentenweise Multiplikation

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

wir haben also *Nullteiler* und das macht es unmöglich, Inverse zu finden.

Andererseits ist ein Ärgernis der reellen Zahlen, dass es für die Gleichung $x^2 = -1$ keine Lösung gibt. Wir führen also eine neue *imaginäre* Zahl i ein mit $i^2 = -1$.

Wir können nun ein Element $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ als $a + bi$ mit unserem neuen Symbol i schreiben.

Die Addition ist weiterhin $(a + bi) + (a' + b'i) = (a + a') + (b + b')i$. Ein Vektor in \mathbb{R}^2 ist nichts als ein Paar von reellen Zahlen, und wir schreiben unsere Zahlen einfach anders auf.

In dieser Schreibweise ist i nichts als ein Name für den Vektor $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ während wir den Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ einfach mit dem Symbol 1 schreiben (das wir sogleich als multiplikatives neutrales Element weglassen). Also

$$a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a \cdot 1 + b \cdot i = a + bi$$

Das ist eine eindeutige Schreibweise,

Definition 2.4.1. Wir bezeichnen mit $(\mathbb{C}, +, \cdot)$ die Menge der *komplexen Zahlen* $\{a + bi \mid a, b \in \mathbb{R}\}$ mit der komponentweisen Addition und der Multiplikation

$$(a + bi)(a' + b'i) = (aa' - bb') + (ab' + ba')i$$

Woher kommt die Multiplikationsregel? Einerseits gilt $(0 + 1 \cdot i)(0 + 1 \cdot i) = -1$, was wir wünschen.

Andererseits folgt notwendigerweise aus dem Distributivgesetz, dass

$$\begin{aligned} (a + bi)(a' + b'i) &= (aa' + (ab' + ba')i + bb' \cdot i^2 \\ &= (aa' - bb') + (ab' + ba')i \end{aligned}$$

und mit $i^2 = -1$ ist die Multiplikation nun festgelegt!

Satz 2.4.2. $(\mathbb{C}, +, \cdot)$ ist ein Körper.

Beweis: $(\mathbb{C}, +)$ ist eine abelsche Gruppe, wir haben ja nur die Elemente der abelschen Gruppe \mathbb{R}^2 anders benannt.

Das Assoziativitätsgesetz für die Multiplikation überlassen wir als (langwierige aber einfache) Übung.

Das neutrale Element der Multiplikation ist $1 = 1 + 0i$ und für jedes $a + bi \neq 0$ ist

$$(a + bi)^{-1} = \frac{a - bi}{a^2 + b^2}$$

ein multiplikatives Inverses, denn wir rechnen

$$a + bi \cdot \frac{a - bi}{a^2 + b^2} = \frac{(a + bi)(a - bi)}{a^2 + b^2} = \frac{a^2 + b^2 + 0}{a^2 + b^2} = 1$$

Auch Überprüfung des Distributivgesetzes überlassen wir als Übung. □

Fassen wir die Ebene \mathbb{R}^2 dermaßen als Körper der komplexen Zahlen auf, so sprechen wir auch von der *komplexen Zahlenebene*.

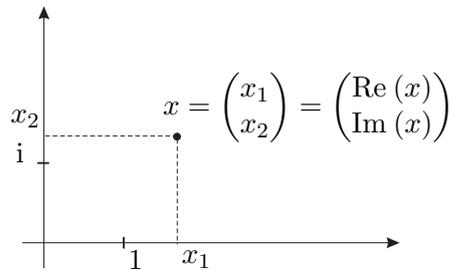
Die komplexe Zahl $i = 0 + 1 \cdot i$ heißt *imaginäre Einheit*.

Die injektive Abbildung

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{C} \\ \lambda &\mapsto \lambda \cdot 1 + 0 \cdot i \end{aligned}$$

erlaubt es, die reellen Zahlen mit einem *Unterkörper* der komplexen Zahlen zu identifizieren, also einer Teilmenge, die unter allen Körperoperationen abgeschlossen ist.

Gegeben eine beliebige komplexe Zahl $x = x_1 + x_2i \in \mathbb{C}$ nennen wir $x_1 = \operatorname{Re}(x) \in \mathbb{R}$ den *Realteil* und $x_2 = \operatorname{Im}(x) \in \mathbb{R}$ den *Imaginärteil* von x :



Die reellen Zahlen liegen dann auf der horizontalen Achse, der sogenannten *reellen Achse*. Vertikal durch 0 verläuft die *imaginäre Achse*.

Definition 2.4.3. 1. Die Abbildung

$$\begin{aligned} \mathbb{C} &\rightarrow \mathbb{C} \\ x = x_1 + x_2i &\mapsto \bar{x} = x_1 - ix_2 \end{aligned}$$

heißt *komplexe Konjugation*. Wir vertauschen i und $-i$, denn algebraisch sind diese beiden Lösungen von $i^2 = -1$ nicht zu unterscheiden!

Geometrisch ist komplexe Konjugation die Spiegelung an der reellen Achse.

2. Für eine komplexe Zahl $z = z_1 + z_2i \in \mathbb{C}$ heißt

$$|z| := \sqrt{z_1^2 + z_2^2}$$

der *Absolutbetrag* von z . Geometrisch betrachtet ist dies der euklidische Abstand von 0 zum Punkt z in der Ebene. .

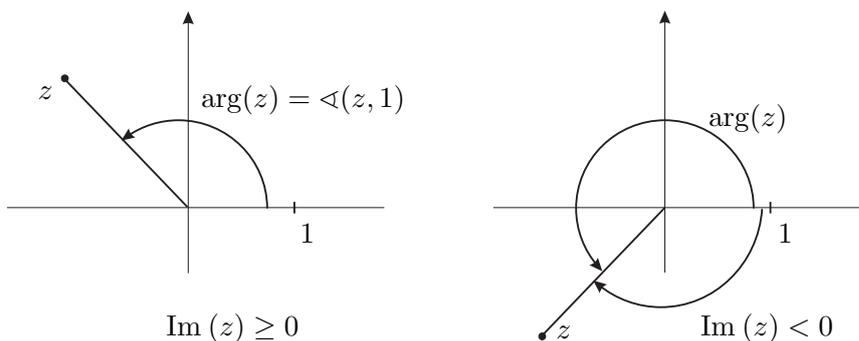
Bemerkungen 2.4.4. Seien $z, w \in \mathbb{C}$ beliebig.

1. Sofort aus der Definition folgt $\overline{\bar{z}} = z, \overline{0} = 0, \overline{i} = -i$ und $\overline{z + w} = \bar{z} + \bar{w}$ sowie $\overline{\bar{z}} = z$
2. Man kann auch nachrechnen $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$
3. Man rechnet $z \cdot \bar{z} = (z_1 + z_2i)(z_1 - z_2i) = z_1^2 + z_2^2 = |z|^2$.
4. Aus der Dreiecksungleichung in \mathbb{R}^2 folgt $|z + w| \leq |z| + |w|$.
5. Aus 3. und 4. folgt $|z \cdot w| = |z||w|$. Nun können wir $\frac{1}{z} = \frac{\bar{z}}{z\bar{z}}$ schreiben. Dies ist eine mögliche Herleitung unserer Formel aus dem Beweis von Satz 2.4.2.

Wir können eine komplexe Zahl auch geometrischer auffassen.

Definition 2.4.5. Sei $z \in \mathbb{C} \setminus \{0\}$; dann heißt der Winkel zwischen der reellen Achse und dem Vektor z das *Argument* von z .

Zeichnung:



Anders ausgedrückt ist das Argument von z die eindeutig bestimmte reelle Zahl $\varphi \in [0, 2\pi)$, für die gilt

$$z = |z|(\cos \varphi + i \sin \varphi) = |z|e^{i\varphi}$$

Für die Multiplikation von $z = |z|e^{i\varphi}$ und $w = |w|e^{i\theta}$ gilt

$$wz = |w||z|e^{i(\varphi+\theta)}$$

Mehr über diese Sichtweise lernen Sie in Analysis.

2.5 Ringe

Wir wären nun bereit Vektorräume über allgemeinen Körpern zu betrachten, aber ein weiteres algebraisches Konstrukt werden wir später brauchen. Wir haben festgestellt, dass die ganzen Zahlen mit Addition und Multiplikation kein Körper sind, sie sind aber ein hoch interessantes mathematisches Objekt.

Definition 2.5.1. 1. Eine Menge R zusammen mit zwei Verknüpfungen

$$\begin{aligned} + : R \times R &\rightarrow R & (a, b) &\mapsto a + b \\ \cdot : R \times R &\rightarrow R & (a, b) &\mapsto a \cdot b \end{aligned}$$

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

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

2. Ein Element $1 \in R$ heißt *Einselement* oder *Eins*, wenn für alle $a \in R$ gilt $1 \cdot a = a \cdot 1 = a$. Ein Ring mit Einselement heißt auch *unitärer Ring*, oder einfach *Ring*, wenn man annimmt, dass jeder Ring eine Eins hat. Das werden wir in diesem Kurs tun.

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

Bemerkungen 2.5.2. 1. Man beachte, dass die Addition in einem Ring immer kommutativ ist, die Multiplikation aber nicht kommutativ sein muss.

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 $a \in R$ wie in einem Körper

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

Beispiele 2.5.3. 1. Die ganzen Zahlen mit Addition und Multiplikation sind ein kommutativer Ring $(\mathbb{Z}, +, \cdot)$.

2. Jeder Körper ist ein kommutativer Ring, insbesondere \mathbb{Q} oder \mathbb{R} .

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

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

die Menge der reellwertigen Funktionen, so definieren wir Verknüpfungen durch Operationen auf den Funktionswerten, wir nennen dies *punktweise* Addition und Multiplikation:

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

Sie versehen R mit der Struktur eines kommutativen Rings. Wir prüfen zum Beispiel die Assoziativität der Multiplikation: $(f(gh))(x) = f(x)(gh(x)) = f(x)(g(x)h(x)) = (f(x)g(x))h(x) = ((fg)h)(x)$. Die Funktion $-f$ schickt x nach $-f(x)$, das additive neutrale Element ist die konstante Funktion $x \mapsto 0$ und das multiplikative neutrale Element ist die konstante Funktion $x \mapsto 1$.

Dieser Ring ist kein Körper, denn eine Funktion f , die den Wert 0 annimmt, aber nicht konstant 0 ist, hat kein multiplikatives Inverses.

Allgemeiner sei M eine Menge und R ein Ring. Dann wird die Menge der Abbildungen $\{f : M \rightarrow R\}$ mit der punktweisen Addition und Multiplikation zu einem Ring.

4. Auf der abelschen Gruppe $\mathbb{Z}/m\mathbb{Z}$ mit $m \in \mathbb{N}$ aus Beispiel 2.1.6 kann man durch

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

eine Multiplikation definieren. Sie ist wohldefiniert, denn für

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

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} . Wir prüfen zum Beispiel

$$\bar{a} \cdot (\bar{b} + \bar{c}) = \bar{a} \cdot \overline{b + c} = \overline{a \cdot (b + c)} = \overline{a \cdot b + a \cdot c} = \overline{a \cdot b} + \overline{a \cdot c} = \bar{a} \cdot \bar{b} + \bar{a} \cdot \bar{c}.$$

Es liegt ein kommutativer Ring mit Eins $\bar{1}$ vor, den wir *Restklassenring* nennen.

Wir rechnen zum Beispiel in $\mathbb{Z}/4\mathbb{Z}$:

$$\begin{aligned}\bar{2} \cdot \bar{2} &= \bar{4} = \bar{0} . \\ \bar{2} \cdot \bar{1} &= \bar{2} \quad \bar{2} \cdot \bar{3} = \bar{6} = \bar{2} .\end{aligned}$$

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

Der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ aus Beispiel 2.5.3.4 gibt uns unser letztes Beispiel für Körper, genau genommen eine unendliche Familie von Körpern.

Satz 2.5.4. *Der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ ist genau dann ein Körper, wenn m eine Primzahl ist.*

Wir schreiben oft \mathbb{F}_p für eine Primzahl p um zu betonen, dass wir die Körperstruktur betrachten, während wir mit $\mathbb{Z}/p\mathbb{Z}$ oft nur die zyklische Gruppe unter Addition meinen.

Beweis: Ist m keine Primzahl, so gibt es $1 < k, l < m$ mit $m = k \cdot l$. Also ist $\bar{k}, \bar{l} \neq \bar{0}$, aber $\bar{0} = \overline{m} = \overline{k \cdot l} = \bar{k} \cdot \bar{l}$. Wegen Bemerkung 2.3.3.1 kann $\mathbb{Z}/m\mathbb{Z}$ kein Körper sein.

Sei umgekehrt m prim. Wir müssen zeigen, dass jedes $\bar{k} \in \mathbb{Z}/m\mathbb{Z}$ ein multiplikatives Inverses hat. Wir betrachten die additive Untergruppe H , die von k , erzeugt wird wie in Beispiel 2.1.15.3, also $\dots, -2k, -k, 0, k, 2k, 3k, \dots$ (additiv geschrieben).

Da $\mathbb{Z}/m\mathbb{Z}$ endlich ist, ist auch $|H|$ endlich. Nach dem Satz von Lagrange 2.1.16 gilt $|H| \mid |\mathbb{Z}/m\mathbb{Z}| = m$. Aber da m prim ist gilt $|H| = 1$ (unmöglich, da $0 \neq k$) oder $|H| = m$. Mit $|H| = m$ gilt $H = \mathbb{Z}/m\mathbb{Z}$ und insbesondere gibt es $\ell \in \mathbb{Z}$ mit $\ell \cdot \bar{k} = \bar{1} \in \mathbb{Z}/m\mathbb{Z}$ (hier schreiben wir $\ell \cdot \bar{k}$ für die ℓ -fache Summe von \bar{k} mit sich selbst. Dies ist keine Ringmultiplikation, da ℓ und \bar{k} in verschiedenen Ringen leben!)

Es gilt nun nach Definition $\ell \cdot \bar{k} = \overline{\ell \cdot k} = \bar{\ell} \cdot \bar{k}$ und damit ist $\bar{\ell}$ das gesuchte Inverse für \bar{k} . \square

Eine besondere Eigenschaft von \mathbb{F}_p ist, dass die p -fache Summe der 1 mit sich selbst gleich 0 ist, also $1 + 1 + \dots + 1 = 0$. Im Gegensatz dazu ist jede Summe von Einsen in \mathbb{Q} oder \mathbb{R} ungleich null. Die *Charakteristik* eines Körpers K ist das kleinste m so dass $m \cdot 1 = 1 + 1 + \dots + 1 = 0$ ist, oder 0 wenn es kein solches m gibt.

Definition 2.5.5. Seien $(R, +, \cdot)$ und (S, \oplus, \odot) Ringe, so heißt eine Abbildung

$$\varphi : R \rightarrow S$$

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

$$\begin{aligned} \varphi(a + b) &= \varphi(a) \oplus \varphi(b) \\ \varphi(a) \odot \varphi(b) &= \varphi(a \cdot b) \\ \varphi(1_R) &= 1_S \end{aligned}$$

Beispiele 2.5.6. 1. Die kanonische Abbildung von \mathbb{Z} auf $\mathbb{Z}/m\mathbb{Z}$ aus Beispiel 2.2.2.2, die durch $a \mapsto a + m\mathbb{Z}$ gegeben ist, ist nicht nur ein Gruppenhomomorphismus sondern auch ein Ringhomomorphismus. Dies folgt sofort aus der Definition.

Umgekehrt ist die Abbildung $e : \mathbb{Z}/m\mathbb{Z} \rightarrow \mathbb{Z}$, die jede Restklasse zu ihrem kleinsten nichtnegativen Element schickt kein Ringhomomorphismus, da sie nicht einmal ein additiver Gruppenhomomorphismus ist.

2. Wegen der Bemerkungen 2.4.4 respektiert die komplexe Konjugation Addition und Multiplikation in \mathbb{C} , wir haben also einen bijektiven Ringhomomorphismus $\mathbb{C} \rightarrow \mathbb{C}$.

Manchmal ist es aus dem Zusammenhang klar ob wir von einem Gruppenhomomorphismus oder Ringhomomorphismus sprechen und wir sagen nur *Homomorphismus*, aber im Zweifelsfall ist es besser, genauer zu sein.

Bemerkung 2.5.7. Wie schon erwähnt ist jeder Körper insbesondere ein kommutativer Ring.

Sei K ein Körper und E ein Ring, der nicht der triviale Ring $\{0\}$ ist. Dann ist jeder Ringhomomorphismus $\varphi : K \rightarrow E$ injektiv: Da φ insbesondere ein Gruppenhomomorphismus der additiven Gruppen ist, reicht es, 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 nur gelten kann, wenn $0 = 1$ in E , aber dann sind alle Elemente in E gleich 0 und $E = \{0\}$.

Wir brauchen später noch eine spezielle Familie von Ringen, die Polynomringe. Sei K ein Körper oder, allgemeiner, ein kommutativer Ring.

Betrachtung 2.5.8. 1. Wir legen einen Körper K fest (wir könnten auch allgemeiner einen kommutativen Ring R wählen).

Wir betrachten die Menge aller formalen Ausdrücke $a_0 + a_1t + a_2t^2 + \dots + a_nt^n$, wobei die *Koeffizienten* (a_0, a_1, \dots, a_n) Elemente von K sind und die t^i Symbole. Wir identifizieren t^0 mit 1 und wir identifizieren Ausdrücke wie $1 + t$ und $1 + t + 0t^2$, indem wir Terme $0 \cdot t^n$ weglassen. Diese formalen Ausdrücke formen die Menge der *Polynome*.

Der *Grad* eines Polynoms $f = a_0 + a_1t + a_2t^2 + \dots + a_nt^n$ ist $-\infty$, falls $f = 0$, sonst gleich dem $\max_i \{i \mid a_i \neq 0\}$. Zum Beispiel gilt $\text{grad}(t^3 + 2t + 1) = 3$.

2. Seien

$$f := a_0 + a_1t + a_2t^2 + \dots + a_nt^n \quad \text{und} \quad g := b_0 + b_1t + b_2t^2 + \dots + b_mt^m$$

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

$$f + g = (a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \dots + (a_N + b_N)t^N$$

wobei N das Maximum von m und n ist und wir fehlende Koeffizienten gleich 0 setzen können.

Zum Beispiel ist $(x^3 + 2x^2) + (x^2 + 4x) = x^3 + 3x^2 + 4x$.

Wir multiplizieren Polynome, indem wir "ausmultiplizieren", also formal das Distributivgesetz anwenden und die Regel $t^a \cdot t^b = t^{a+b}$ verwenden:

$$f \cdot g = a_0b_0 + (a_1b_0 + a_0b_1)t + (a_2b_0 + a_1b_1 + a_0b_2)t^2 + \dots$$

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 .

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 $\sqrt{2} \mapsto \sqrt{2}^4 - 3 = 4 - 3 = 1$. Diese Auswertung bei λ ist ein Ringhomomorphismus $K[t] \rightarrow K$.

So liefert jedes Polynom $f \in K[t]$ eine Funktion $K \rightarrow K$, in dieser Form haben Sie Polynome schon getroffen. Ein Polynom ist aber nicht gleich der Funktion, die es beschreibt! $t + t^2 \in \mathbb{F}_2[t]$ nimmt als Funktion immer den Wert 0 an, es ist aber nicht das gleiche Polynom wie $0 \in \mathbb{F}_2[t]$.

3 Vektorräume

3.1 Vektorräume

Nach unserer Vorarbeit in der Algebra kommen nun zum Begriff eines allgemeinen 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.

Wir haben abstrakte Körper eingeführt um über beliebige Vektorräume sprechen zu können. Es gibt sehr viele interessante Körper in der Mathematik, aber in diesem Skript kann K nur $\mathbb{R}, \mathbb{C}, \mathbb{Q}$ oder \mathbb{F}_p sein. Wenn Sie an irgendeiner Stelle feststecken und die Übersicht verlieren ist es durchaus ratsam erst einmal zu sehen, ob Sie die Situation für $K = \mathbb{R}$ verstehen.

Definition 3.1.1. 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

$$+ : V \times V \rightarrow V \quad \cdot : K \times V \rightarrow V ,$$

die den folgenden Axiomen genügen:

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

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

$$(V2a) \quad (\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$$

$$(V2b) \quad \alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$$

$$(V2c) \quad (\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$$

$$(V2d) \quad 1 \cdot v = v$$

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*.

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.

Um Verwechslungen zu vermeiden können wir 0_K für das neutrale Element von $(K, +)$ und 0_V für das neutrale Element von $(V, +)$ schreiben. Meist ist aber aus dem Zusammenhang klar, welche Null gemeint ist.

Beispiele 3.1.2. 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 $\cdot : K \times V \rightarrow V$ 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 $0_{K^n} := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ und $-x = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}$.

Da Addition und Skalarmultiplikation komponentenweise definiert sind können wir auch alle Axiome komponentenweise prüfen.

Zum Beispiel ist

$$\alpha \cdot \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) = \alpha \cdot \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} \alpha(x_1 + y_1) \\ \vdots \\ \alpha(x_n + y_n) \end{pmatrix} = \begin{pmatrix} \alpha x_1 + \alpha y_1 \\ \vdots \\ \alpha x_n + \alpha y_n \end{pmatrix} = \alpha \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \alpha \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

und das zeigt (V2b).

2. Setzen wir $K = \mathbb{R}$ erhalten wir unsere Beispiele \mathbb{R}^2 und \mathbb{R}^n vom Beginn des Kurses zurück.

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

4. $V = (\{0\}, +, \cdot)$ heißt der *Nullvektorraum*. Da jeder Vektorraum zumindest den Nullvektor enthält ist dies der kleinstmögliche Vektorraum. Wir schreiben den Nullvektorraum als $\{0\}$. Wir erhalten den Nullvektorraum auch als K^0 für beliebiges K !

5. $(\mathbb{C}, +, \cdot|_{\mathbb{R} \times \mathbb{C}}) = (\mathbb{R}^2, +, \cdot)$ ist ein \mathbb{R} -Vektorraum. Wenn wir \mathbb{C} als \mathbb{R} -Vektorraum betrachten, vergessen wir große Teile der Multiplikation.

6. $(\mathbb{R}, +, \cdot|_{\mathbb{Q} \times \mathbb{R}})$ ist ein \mathbb{Q} -Vektorraum!

7. Sei X eine Menge und W ein K -Vektorraum. Dann wird die Menge der Abbildungen $\text{Abb}(X, W)$ durch die Operationen auf den Funktionswerten zu einem K -Vektorraum.

Genauer: Gegeben $f, g \in \text{Abb}(X, W)$, setze $(f + g)(x) := f(x) + g(x)$. Dies ist eine abelsche Gruppe mit neutralem Element $0(x) = 0$ und Inversen $(-f)(x) = -f(x)$. Die skalare Multiplikation ist durch $(\lambda f)(x) := \lambda f(x)$ definiert.

Alle Axiome sind leicht zu prüfen (siehe Übungsblatt).

8. Die Menge aller unendlichen Folgen formt einen Vektorraum, denn wir können wieder komponentenweise skalar multiplizieren und addieren.

In der Tat sind Folgen mit Werten in K nur Abbildungen von \mathbb{N} nach K !

Auch die Teilmenge der Folgen, die nach 0 konvergieren bildet einen Vektorraum! Das folgt aus Resultaten, die Sie in Analysis zeigen.

9. Seien V, W zwei beliebige Vektorräume. Dann definieren wir die (*äußere*) *direkte Summe* $V \oplus W$ wie folgt: Elemente von $V \oplus W$ sind Paare $(v \in V, w \in W)$. Die Addition erfolgt durch $(v, w) + (v', w') = (v + v', w + w')$. Skalarmultiplikation ist gegeben durch $\lambda \cdot (v, w) = (\lambda v, \lambda w)$. Es gilt $-(v, w) = (-v, -w)$ und $0_{V \oplus W} = (0_V, 0_W)$. Wieder prüft man die Axiome Komponentenweise.

Satz 3.1.3. *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:

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

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

4. Wir berechnen

$$(-1)v + v \stackrel{(V2d)}{=} (-1)v + 1 \cdot v \stackrel{(V2a)}{=} (-1 + 1)v = 0_K \cdot v \stackrel{1.}{=} 0_V. \quad \square$$

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

3.2 Untervektorräume

Definition 3.2.1. Sei $(V, +, \cdot)$ ein K -Vektorraum. Eine Teilmenge $W \subset V$ heißt *Untervektorraum* und wir schreiben $W \leq V$, 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 3.2.2. Sei $(V, +, \cdot)$ ein K -Vektorraum. Sei $W \subset V$ ein Untervektorraum. Sei $+_W$ die Einschränkung von

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

Sei \cdot_W die Einschränkungen von

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

Dann ist $(W, +_W, \cdot_W)$ ein K -Vektorraum. Der Nullvektor in W stimmt mit dem Nullvektor in V überein.

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

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

so dass (UG3) erfüllt ist. Da W wegen (UV1) nicht leer ist wählen wir $w \in W$ und dann ist auch $0_V = w - w \in W$. (Alternativer Beweis: es $0_V = 0_k \cdot w \in W$.) Damit gilt (UG1).

2. Wegen (UV3) definiert \cdot_W tatsächlich ein Produkt $K \times W \rightarrow W$ (a priori nimmt die Einschränkung Werte in V an). Die Axiome (V2a-d) gelten sogar für alle Elemente in V , also erst recht für alle Elemente in der Teilmenge W . □

Beispiele 3.2.3. 1. $K = \mathbb{R}$ und $(\mathbb{C}, +, \cdot)$. Dann ist

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

ein Untervektorraum, ebenso

$$W_{\mathbb{C}} = \{0 + x_2 \cdot i \mid x_2 \in \mathbb{R}\}$$

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

$$\alpha \cdot v = (0 + i) \cdot (1 + 0 \cdot i) = 1 \cdot i \notin W_{\mathbb{R}}.$$

3. Für jeden Vektor $v \in \mathbb{R}^n \setminus \{0\}$ ist die Gerade $G_{0,v}$ durch 0 ein Untervektorraum von \mathbb{R}^n , aber für $p \neq 0$ ist $G_{p,v}$ im Allgemeinen kein Untervektorraum, es sei denn $0 = p + tv \in G_{p,v}$ für $t \in \mathbb{R}$.
4. Jeder Vektor v in einem Vektorraum V erzeugt einen Untervektorraum $Kv = \{\lambda v \mid \lambda \in K\}$. Zum Beispiel ist $W_{\mathbb{C}} = \mathbb{R}i$.
5. Im Vektorraum der reellwertigen Folgen sind die Nullfolgen ein Untervektorraum denn aus der Analysis wissen Sie, dass Summen und Vielfache von Nullfolgen wieder Nullfolgen sind.
6. In jedem Vektorraum V sind V selbst und $W := \{0\}$ Untervektorräume. $W = \{0\}$ heißt der *triviale Untervektorraum*.
7. Jeder Untervektorraum eines Untervektorraums von V ist selbst ein Untervektorraum von V .

Wir können aus Unterräumen neue Vektorräume konstruieren. Sei V ein Vektorraum und seien W_1, W_2 Untervektorräume von V . Dann ist auch $W_1 \cap W_2$ ein Untervektorraum. Wir haben sogar allgemeiner:

Satz 3.2.4. Sei I eine beliebige Menge und $\{W_i\}_{i \in I}$ eine Familie von Untervektorräumen von V , d.h. für jedes $i \in I$ gibt es ein $W_i \leq V$. Dann ist auch $\bigcap_{i \in I} W_i = \{w \in V \mid \forall i \in I : w \in W_i\}$ ein Untervektorraum von V .

Beweis: Falls $I = \emptyset$ ist $\bigcap_{\emptyset} W_i = V$, was trivialerweise ein Untervektorraum ist.

Sei also I nicht leer. Wir prüfen.

(UV1) Für alle $i \in I$ gilt $0_V \in W_i$, also gilt nach Definition $0_W \in \bigcap_I W_i$.

(UV2) folgt, da mit $v, w \in W_i$ auch $v + w \in W_i$ liegt, also aus $v, w \in \bigcap_I W_i$ folgt $v + w \in \bigcap_I W_i$.

(UV3) folgt analog. □

Bemerkung 3.2.5. $W_1 \cup W_2$ ist im Allgemeinen kein Untervektorraum.
 Als Gegenbeispiel betrachte \mathbb{C} als reellen Vektorraum, $K = \mathbb{R}$ und $V = \mathbb{C}$, und

$$W_{\mathbb{R}} \cup W_{\mathbb{S}} = \{x_1 + x_2 i \mid x_1 = 0 \text{ oder } x_2 = 0\}.$$

Dies ist das Achsenkreuz bestehend aus der reellen und imaginären Achse. Dann ist $x = 1 + 0 \cdot i \in W_{\mathbb{R}}$ auf der reellen Achse und $y = 0 + 1 \cdot i \in W_{\mathbb{S}}$ auf der imaginären Achse. Die Summe $x + y = 1 + 1 \cdot i \notin W_{\mathbb{R}} \cup W_{\mathbb{S}}$ ist aber nicht auf dem Achsenkreuz.

Da $W_1 \cup W_2$ selbst kein Untervektorraum ist definieren wir einen minimalen Unterraum, der W_1 und W_2 enthält:

Definition 3.2.6. Sei V ein K -Vektorraum, $r \in \mathbb{N}^*$ und seien $W_1, \dots, W_r \subset V$ Untervektorräume. Dann heißt

$$W_1 + \dots + W_r := \left\{ v \in V \mid \exists w_i \in W_i \text{ mit } v = \sum_{i=1}^r w_i \right\}$$

die (innere) *Summe der Untervektorräume* W_1, \dots, W_r .

Zum Beispiel ist \mathbb{C} wie in Beispiel 3.2.3.1 als Vektorraum die Summe von $W_{\mathbb{R}}$ und $W_{\mathbb{S}}$.

Lemma 3.2.7. Die Summe $W_1 + \dots + W_r$ aus Definition 3.2.6 ist ein Untervektorraum von V , der $W_1 \cup \dots \cup W_r$ enthält, und jeder andere Untervektorraum, der $W_1 \cup \dots \cup W_r$ als Teilmenge enthält, enthält auch $W_1 + \dots + W_r$ als Untervektorraum.

Beweis: Wir prüfen:

(UV1) gilt, da $0 = 0 + \dots + 0 \in W_1 + \dots + W_r$.

(UV2) gilt, da aus $v = w_1 + \dots + w_r$ und $v' = w'_1 + \dots + w'_r$ folgt, dass $v + w = (w_1 + w'_1) + \dots + (w_r + w'_r) \in W_1 + \dots + W_r$, da $w_i + w'_i \in W_i$.

(UV3) folgt da mit $v = w_1 + \dots + w_r$ gilt $\lambda v = \lambda(w_1 + \dots + w_r) = \lambda w_1 + \dots + \lambda w_r \in W_1 + \dots + W_r$.

Jedes W_i ist in $W_1 + \dots + W_r$ enthalten, denn $w_i = 0 + \dots + w_i + \dots + 0$ und 0 ist in jedem $W_{j \neq i}$ enthalten.

Sei weiterhin $U \leq V$ gegeben mit $W_i \subset U$ für alle i . Sei $v = w_1 + \dots + w_r \in W_1 + \dots + W_r$ beliebig. Dann gilt $v \in U$, da U unter Summen abgeschlossen ist. \square

Da jeder andere Untervektorraum, der alle W_i enthält auch die Summe enthält, ist die Summe notwendigerweise der kleinste Untervektorraum, der alle W_i enthält. Wir sagen die Summe ist *minimal* mit dieser Eigenschaft.

Lemma 3.2.8. Sei V ein 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.

Beweis: Wir zeigen zuerst $1 \Rightarrow 2$. 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:

$$v = w_1 + w_2 = w'_1 + w'_2 \quad w'_i \in W_i .$$

Daraus folgt

$$w_1 - w'_1 = w'_2 - w_2 \in W_1 \cap W_2 = \{0\} ,$$

also $w_1 = w'_1$ und $w_2 = w'_2$.

Umgekehrte nehmen wir an, dass $w \in W_1 \cap W_2$ liegt. Dann ist $w + 0 = 0 + w$ und nach der Eindeutigkeit muss gelten $w = 0$. \square

Definition 3.2.9. Ist $V = W_1 + W_2$ und gilt $W_1 \cap W_2 = \{0\}$, 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 3.2.10. Die Summe $\mathbb{C} = W_{\mathbb{Q}} + W_{\mathbb{R}}$ ist direkt da $W_{\mathbb{Q}} \cap W_{\mathbb{R}} = \{0\}$, also $\mathbb{C} = W_{\mathbb{Q}} \oplus W_{\mathbb{R}}$.

Betrachten wir dagegen $W_{\mathbb{Q}} + W_{\mathbb{Q}} \subset \mathbb{C}$ dann gilt offensichtlich $W_{\mathbb{Q}} \cap W_{\mathbb{Q}} = W_{\mathbb{Q}}$ und die Summe ist nicht direkt.

Wir haben die Notation \oplus schon einmal benutzt! Man beachte den leichten Unterschied im Kontext: Wenn wir zwei Unterräume von V betrachten, dann gibt es immer die Summe und manchmal ist diese Summe eine (innere) direkte Summe. Wenn wir dagegen zwei Vektorräume für sich betrachten, können wir immer die (äußere) direkte Summe bilden. Da es leicht ist, aus dem Auge zu verlieren, ob wir einen Vektorraum als Unterraum eines anderen Vektorraums auffassen oder für sich selbst, sollten wir sicherstellen dass keine Widersprüche entstehen.

Zu diesem Zweck wollen wir zwei verschiedene Vektorräume identifizieren, wie zuvor isomorphe Gruppen. Dafür brauchen wir erst einmal einen Begriff von Abbildung zwischen Vektorräumen.

3.3 Lineare Abbildungen

Vektorräume sind die zentralen mathematischen Objekte der linearen Algebra. So wie wir schon für Gruppen 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.

Definition 3.3.1. Sei K ein Körper. Seien V, W zwei K -Vektorräume. Eine Abbildung

$$f : V \rightarrow W$$

heißt K -linear oder K -Vektorraumhomomorphismus, falls gilt

(L1) f ist Gruppenhomomorphismus bezüglich der Addition, d.h. für alle $v, v' \in V$ gilt $f(v + v') = f(v) + f(v')$

(L2) f ist mit der skalaren Multiplikation verträglich, d.h. für alle $v \in V$ und für alle $\lambda \in K$ gilt $f(\lambda v) = \lambda f(v)$.

Lemma 3.3.2. Eine Abbildung $f : V \rightarrow W$ ist genau dann linear, wenn gilt

(L) für alle $v, v' \in V$ und für alle $\lambda, \lambda' \in K$ ist $f(\lambda v + \lambda' v') = \lambda f(v) + \lambda' f(v')$.

Beweis: Um zu sehen, dass aus (L) die Bedingung (L1) folgt, wähle in (L) 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

$$f(\lambda v + \lambda' v') \stackrel{(L1)}{=} f(\lambda v) + f(\lambda' v') \stackrel{(L2)}{=} \lambda f(v) + \lambda' f(v'). \quad \square$$

Bemerkungen 3.3.3. 1. Induktiv zeigt man aus Lemma 3.3.2, dass für alle $v_1, \dots, v_n \in V$ und für alle $\lambda_1, \dots, \lambda_n \in K$ gilt

$$f(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n).$$

2. $f(0) = 0$. Das gilt wegen Satz 2.2.3.1 für Gruppenhomomorphismen, angewandt auf die Gruppe $(V, +)$.

3. $f(v - v') = f(v) - f(v')$ für alle $v, v' \in V$. Wir rechnen

$$f(v - v') = f(1 \cdot v + (-1) \cdot v') \stackrel{(L)}{=} 1 \cdot f(v) + (-1)f(v') = f(v) - f(v').$$

Beispiele 3.3.4. 1. Seien der Körper K und ein K -Vektorraum $V = W$ beliebig. Dann ist $f = \text{id}_V$ eine lineare Abbildung. Allgemeiner ist für jedes $\lambda \in K$ die Abbildung $v \mapsto \lambda v$ K -linear.

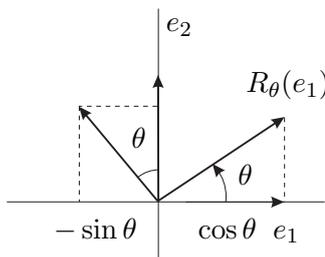
Für allgemeine K -Vektorräume V, W ist $0 : V \rightarrow W$ definiert durch $v \mapsto 0$ eine lineare Abbildung (Beweis siehe Beispiel 3).

2. $K = \mathbb{R}$ und $V = W = \mathbb{R}^2$. Wir wählen ein festes $\theta \in \mathbb{R}$ und betrachten

$$R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta x - \sin \theta y \\ \sin \theta x + \cos \theta y \end{pmatrix}$$

R_θ ist eine Drehung um den Winkel θ gegen den Uhrzeigersinn:



Wir prüfen (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. Sei $V = W = K$; definiere eine lineare Abbildung $f : K \rightarrow K$ durch $f(x) := \gamma x$ für ein festes $\gamma \in K$.

(L1) folgt aus $f(x + x') = \gamma(x + x') = \gamma x + \gamma x' = f(x) + f(x')$, wegen des Distributivgesetzes.

(L2) folgt aus $f(\lambda x) = (\lambda x)\gamma = \lambda(\gamma x) = \lambda f(x)$ mit dem Assoziativgesetz der Multiplikation.

Der gleiche Beweis zeigt auch, dass $v \mapsto \gamma v$ für jeden K -Vektorraum eine lineare Abbildung ist.

Wir berechnen

$$\ker f = \left\{ x \in K \mid \gamma x = 0 \right\} = \begin{cases} \{0\}, & \text{falls } \gamma \neq 0 \\ K, & \text{falls } \gamma = 0 \end{cases}$$

d.h. f ist injektiv genau für $\gamma \neq 0$.

$$\text{Im } f = \left\{ \gamma x \in K \mid x \in K \right\} = \begin{cases} \{0\}, & \text{falls } \gamma = 0 \\ K, & \text{falls } \gamma \neq 0 \end{cases}$$

d.h. f ist surjektiv genau für $\gamma \neq 0$.

Sei umgekehrt $f : K \rightarrow K$ linear und sei $\gamma = f(1)$. Dann gilt $f(x) = f(x \cdot 1) = x f(1) = \gamma x$. Das heißt jede lineare Abbildung $K \rightarrow K$ hat die Form $x \mapsto \gamma x$ für $\gamma \in K$.

Satz 3.3.5. Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann gilt

1. Ist $V' \subset V$ ein Untervektorraum, so ist das Bild $f(V')$ ein Untervektorraum von W . Insbesondere ist $f(V) = \text{Im}(f)$ ein Untervektorraum von W .
2. Sei $W' \subset W$ ein Untervektorraum; dann ist das Urbild $f^{-1}(W')$ ein Untervektorraum von V . Insbesondere ist der Kern $f^{-1}(0) = \ker f$ ein Untervektorraum von V .
3. Die Verknüpfung von zwei lineare Abbildungen ist linear: Sei K ein beliebiger Körper, und seien V, W, Z drei K -Vektorräume. Seien $f : V \rightarrow W$ und $g : W \rightarrow Z$ zwei lineare Abbildungen, dann ist auch ihre Verknüpfung $f \circ g$ eine lineare Abbildung.
4. Ist f linear und bijektiv, so ist die Umkehrabbildung $f^{-1} : W \rightarrow V$ ebenfalls linear.

Beweis:. 1. Da $V' \subset V$ ein Untervektorraum ist, folgt $0 \in V'$, also $f(0) = 0 \in f(V')$. Also $f(V') \neq \emptyset$.

Seien $w_1, w_2 \in f(V')$ und $\lambda_1, \lambda_2 \in K$. Zu zeigen ist:

$$\lambda_1 w_1 + \lambda_2 w_2 \in f(V') .$$

Wähle Urbilder $v_i \in V'$ mit $f(v_i) = w_i$. Dann ist

$$v := \lambda_1 v_1 + \lambda_2 v_2 \in V',$$

da V' ein Untervektorraum von V ist. Rechne

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 f(v_1) + \lambda_2 f(v_2) \stackrel{(L)}{=} f(\lambda_1 v_1 + \lambda_2 v_2) = f(v) \in f(V').$$

2. Sei $W' \subset W$ ein Untervektorraum. Also $0 \in W'$, daher $0 \in f^{-1}(\{0\}) \subset f^{-1}(W')$.
Seien $v_1, v_2 \in f^{-1}(W')$ und $\lambda_1, \lambda_2 \in K$. Es gilt

$$f(\lambda_1 v_1 + \lambda_2 v_2) \stackrel{(L)}{=} \lambda_1 f(v_1) + \lambda_2 f(v_2) \in W',$$

also $\lambda_1 v_1 + \lambda_2 v_2 \in f^{-1}(W')$.

3. Wenn $f : U \rightarrow V$ und $g : V \rightarrow W$ (L) erfüllen dann gilt $g \circ f(\lambda v + \lambda' v') = g(\lambda f(v) + \lambda' f(v')) = \lambda g(f(v)) + \lambda' g(f(v'))$ und $g \circ f$ erfüllt (L).

4. Nach Satz 2.2.3.3 ist f^{-1} Gruppenhomomorphismus bezüglich der Addition.
Für $w \in W$ und $\lambda \in K$ gilt

$$\lambda f^{-1}(w) = f^{-1} f(\lambda f^{-1}(w)) = f^{-1} \lambda (f f^{-1}(w)) = f^{-1}(\lambda w). \quad \square$$

Betrachtung 3.3.6. Sei $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ und $f : V \rightarrow W$ unsere lineare Abbildung. Wir schreiben einen Vektor $v \in \mathbb{R}^n$ als

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + v_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Dann gilt wegen (L)

$$f \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 f \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + v_n f \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Das heißt, f ist von den n Vektoren $f(e_i)$ bestimmt, wobei e_i den Vektor bezeichnet, der 1 in der i -ten Komponente und sonst gleich 0 ist.

Umgekehrt gibt jede Sammlung von n Vektoren in \mathbb{R}^m eine lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m .

Die $m \times n$ Komponenten dieser Vektoren können wir in eine Matrix schreiben, zum Beispiel erhalten wir für Beispiel 3.3.4.2 die Matrix

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

da gilt $R_\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ und $R_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$.

Die gleichen Überlegungen treffen zu, wenn wir \mathbb{R} durch einen beliebigen Körper K ersetzen!

Wir werden uns später damit beschäftigen, allgemeine lineare Abbildungen durch Matrizen zu repräsentieren.

Definition 3.3.7. 1. Eine lineare Abbildung $f : V \rightarrow W$ heißt

- *Isomorphismus* falls f bijektiv ist,
- *Endomorphismus* falls $V = W$,
- *Automorphismus* falls $V = W$ und f bijektiv ist.

2. Zwei K -Vektorräume V, W heißen *isomorph*, in Zeichen $V \cong W$, falls ein Vektorraumisomorphismus $f : V \rightarrow W$ existiert.

Isomorphie ist eine Äquivalenzrelation⁴: sie ist reflexiv, denn id ist wegen Beispiel 3.3.4.1 ein Isomorphismus. Transitivität folgt aus Satz 3.3.5.3. und Symmetrie, da nach Satz 3.3.5.4 das Inverse eines Isomorphismus ein Isomorphismus ist.

Definition 3.3.8. Seien V und W Vektorräume über K . Wir bezeichnen die Menge aller K -linearen Abbildungen von V nach W mit $\text{Hom}_K(V, W)$. Wir schreiben $\text{End}_K(V)$ für $\text{Hom}_K(V, V)$.

Satz 3.3.9. $\text{Hom}_K(V, W) \subset \text{Abb}(V, W)$ ist ein K -Untervektorraum des Abbildungsraums aus Beispiel 3.1.2.7.

Beweis: Sicher ist die Nullabbildung $f(v) = 0$ für alle $v \in V$ linear und daher in $\text{Hom}_K(V, W)$, daher ist $\text{Hom}_K(V, W) \neq \emptyset$. Die Summe $f + g : v \mapsto f(v) + g(v)$ linearer Abbildungen ist wieder linear:

$$\begin{aligned} (f + g)(\lambda v + \lambda' v') &= f(\lambda v + \lambda' v') + g(\lambda v + \lambda' v') = \lambda f(v) + \lambda' f(v') + \lambda g(v) + \lambda' g(v') \\ &= \lambda(f + g)(v) + \lambda'(f + g)(v') \end{aligned}$$

Ähnlich rechnet man nach, dass auch die Abbildung λf wieder linear ist. □

Wir kehren zurück zu unseren beiden direkten Summen:

Lemma 3.3.10. Seien $V_1, V_2 \subset W$ Unterräume mit $V_1 \cap V_2 = 0$. Dann ist die innere direkte Summe $V_1 \oplus' V_2$ isomorph zur äußeren direkten Summe $V_1 \oplus V_2$ aus Beispiel 3.1.2.9. Dies rechtfertigt unsere Notation

Beweis: Wir definieren eine Abbildung $a : V_1 \oplus V_2 \rightarrow V_1 \oplus' V_2$ durch $(v_1, v_2) \mapsto v_1 + v_2$. Es ist leicht zu sehen, dass a linear ist.

Per Definition von \oplus' ist a surjektiv.

Der Kern von a besteht aus allen (v_1, v_2) mit $v_1 - v_2 = 0 \in V$. Aber nach Annahme ist $V_1 \cap V_2 = 0$ und nach Lemma 3.2.8 folgt daraus, dass jedes Element in $V_1 + V_2$ sich eindeutig als Summe von Elementen in V_1 und V_2 schreiben lässt, dann folgt aus $v_1 + v_2 = 0 = 0 + 0$ dass $v_1 = v_2 = 0$ ist. Damit ist a injektiv. □

Bemerkung 3.3.11. Die beiden direkten Summen sind nicht nur isomorph, sondern der Isomorphismus ist auch kompatibel mit linearen Abbildungen von oder nach V_1 und V_2 , wir sagen, es ist ein *natürlicher Isomorphismus*.

Der Unterschied zwischen den beiden Konstruktionen ist subtil.

Wir können die (äußere) direkte Summe noch weiter verallgemeinern. Wir können nicht nur zwei Vektorräume "addieren".

Wir erinnern uns, dass eine Familie von K -Vektorräumen $(V_\lambda)_{\lambda \in \Lambda}$ gegeben ist durch eine Indexmenge Λ und für jedes $\lambda \in \Lambda$ ein K -Vektorraum V_λ . Die V_λ müssen nicht alle unterschiedlich sein.

⁴Wir ignorieren wieder die Subtilität, dass die Gesamtheit aller Vektorräume keine Menge ist

Definition 3.3.12. Gegeben sei eine nicht notwendigerweise endliche Familie $(V_\lambda)_{\lambda \in \Lambda}$ von K -Vektorräumen. Wir bilden die (äußere) direkte Summe

$$\bigoplus_{\lambda \in \Lambda} V_\lambda = \left\{ (v_\lambda)_{\lambda \in \Lambda}, \quad \text{nur endlich viele } v_\lambda \in V_\lambda \text{ ungleich Null} \right\}$$

Die K -Vektorraum-Struktur ist dabei komponentenweise definiert.

Beispiel 3.3.13. 1. Man sieht leicht, dass $K^n \cong \bigoplus_{i=1}^n K$ wobei wir auf der rechten Seite die konstante Familie $i \mapsto K$ betrachten. Wir schreiben statt $\bigoplus_{i \in \{1, 2, \dots, n\}}$ einfach $\bigoplus_{i=1}^n$.

Die Vektoren auf beiden Seiten sind gleich definiert, als n -Tupel von Elementen von K . Als Isomorphismus können wir einfach (x_1, \dots, x_n) nach (x_1, \dots, x_n) schicken

2. Der Vektorraum $K[t]$ aller Polynome über K ist isomorph zur unendlichen direkten Summe $\bigoplus_{i \in \mathbb{N}} K$. (Hier verkürzen wir die Schreibweise $\bigoplus_{i \in \mathbb{N}} K$, da die Summanden ja nicht von i abhängen.)

Der Isomorphismus schickt $a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$ nach $(a_0, a_1, a_2, \dots, a_n, 0, \dots)$. Diese Abbildung ist nach Definition linear, injektiv und surjektiv.

Bemerkung 3.3.14. Wir können für die Familie $(V_\lambda)_{\lambda \in \Lambda}$ auch den Vektorraum

$$\prod_{\lambda \in \Lambda} V_\lambda = \left\{ (v_\lambda)_{\lambda \in \Lambda} \mid v_\lambda \in V_\lambda \right\}$$

ohne die Endlichekeitsbedingung bilden.

Dies heißt das *Produkt* der Vektorräume V_λ .

Es ist klar, dass für endliche Indexmengen Produkt und direkte Summe übereinstimmen.

3.4 Quotientenvektorräume

Es stellen sich ein paar natürliche Fragen, nachdem wir Untervektorräume und lineare Abbildungen eingeführt haben.

Die einfachsten Injektionen zwischen Vektorräumen sind Inklusionen eines Untervektorraums. Gibt es ähnliche "natürliche Surjektionen"?

Ein Beispiel für Untervektorräume sind Kerne von linearen Abbildungen. Sind alle Untervektorräume Kerne einer linearen Abbildung?

Der Kern von f misst in gewisser Weise, wie weit f davon entfernt ist, injektiv zu sein. Gibt es umgekehrt einen Vektorraum, der misst, wie weit f davon entfernt ist, surjektiv zu sein? Dieser Vektorraum müsste genau dann verschwinden, wenn f surjektiv ist. (Das ist gerade nicht das Bild von f !)

Wir werden all diese Fragen nun beantworten, die Antwort ist allerdings eine konzeptionell recht anspruchsvolle Konstruktion.

Lemma 3.4.1. 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: Wir benutzen hier nur die Gruppenstruktur von V und U und haben schon im Beweis von Satz 2.1.16 gesehen, dass dies eine Äquivalenzrelation ist!

Hier noch einmal der 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 $v \sim z$. \square

Wir schreiben für die Äquivalenzklasse von v nun $[v] = v + U = \{v + u \mid u \in U\}$

Satz 3.4.2 (Definition). *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 3.4.1 wird durch die Verknüpfungen

$$[v] + [w] := [v + w] \quad \text{und} \quad \alpha[v] := [\alpha v]$$

zu einem K -Vektorraum. Er heißt Quotientenvektorraum von V nach U .

2. Die kanonische Surjektion (vgl. Bemerkung 1.6.6.3)

$$\pi : V \rightarrow V/U \quad \text{mit} \quad \pi(v) = [v]$$

ist linear. Es ist $\ker \pi = U$.

Beweis: 1. Wie in Beispiel 2.1.6 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

$$v \sim w \Rightarrow v - w \in U \Rightarrow \alpha(v - w) \in U \Rightarrow \alpha v \sim \alpha w.$$

Alle Vektorraumaxiome folgen nun durch Vererbung aus V , also $+_{V/U}$ ist assoziativ, da es durch $+_V$ definiert ist und so fort.

2. Wir zeigen, dass π linear ist: $\pi(\lambda v + \lambda' v') = [\lambda v + \lambda' v'] = \lambda[v] + \lambda'[v']$ nach der Definition von Addition und Skalarmultiplikation auf V/U . Vgl. Beispiel 2.2.2.2.

Die Surjektivität von π folgt direkt aus der Definition des Quotientenvektorraumes.

Es ist $v \in \ker(\pi)$ genau dann, wenn $[v] = 0$, was aber äquivalent zu $v \in U$ ist. Also ist $\ker(\pi) = U$. \square

⟨⟨Der nächste Satz ist extrem nützlich, um mit Quotientenvektorräumen zu arbeiten. Er erlaubt uns automatisch zu prüfen, dass eine Abbildung f , die wir auf einem Quotientenvektorraum definieren *wohldefiniert* ist, also dass $f([v]) = f([v'])$ gilt wenn v und v' verschiedenen Repräsentanten der Äquivalenzklasse $[v] = [v']$ sind.⟩⟩

Satz 3.4.3. *Seien $U \leq V$ und W K -Vektorräume. Dann ist es äquivalent eine K -lineare Abbildung $f : V \rightarrow W$ mit $f|_U = 0$ anzugeben oder eine K -lineare Abbildung $\tilde{f} : V/U \rightarrow W$.*

Anders ausgedrückt: Um eine K -lineare Abbildung $g : V/U \rightarrow W$ zu definieren, reicht es $f : V \rightarrow W$ zu definieren mit $f|_U = 0$. Dann setzen wir $g([v]) = f(v)$.

Beweis: Gegeben $\tilde{f} : V/U \rightarrow W$ definieren wir $f = \tilde{f} \circ \pi$ und dies ist linear und verschwindet auf U .

Gegeben f definieren wir $\tilde{f}([v]) = f(v)$. Dies ist wohldefiniert, denn aus $[v] = [v']$ folgt $v - v' \in U$ und dann ist $f(v') = f(v) + f(v' - v) = f(v)$. Weiterhin ist \tilde{f} linear, da f linear ist.

Wir prüfen, dass diese beiden Konstruktionen invers zueinander sind: Beginnen wir mit \tilde{f} dann erhalten wir aus f die Abbildung $\tilde{f} \circ \pi : V \rightarrow W$ die v nach $f([v]) = f(v)$ schickt, also gleich f ist.

Beginnen wir umgekehrt mit eine Abbildung $g : V/U \rightarrow W$ und betrachten $\widetilde{g \circ \pi}$. Diese Abbildung schickt $[v]$ nach $g \circ \pi(v) = g([v])$, ist also wieder g . \square

Theorem 3.4.4 (Homomorphiesatz/kanonische Faktorisierung). Seien V und W K -Vektorräume und sei $f : V \rightarrow W$ linear. Dann existiert ein eindeutiger Isomorphismus

$$\bar{f} : V/\ker f \rightarrow \text{Im } f$$

so dass gilt

$$f = \iota \circ \bar{f} \circ \pi ;$$

wobei

$$\pi : V \rightarrow V/\ker f$$

die kanonische Surjektion und

$$\iota : \text{Im } f \rightarrow W$$

die kanonische Einbettung von $\text{Im } f$ in W ist.

Man kann also jede lineare Abbildung zerlegen in eine kanonische Surjektion, einen Isomorphismus und eine Inklusion. Man schreibt dies auch als *kommutierendes Diagramm*:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \pi \downarrow & & \uparrow \iota \\ V/\ker(f) & \xrightarrow[\cong]{\bar{f}} & \text{Im } f \end{array}$$

Dass das Diagramm “kommutiert”, bedeutet einfach, dass $\iota \circ \bar{f} \circ \pi = f$, die beiden Wege um das Quadrat herum führen zum gleichen Ergebnis. (Zwar ist diese Schreibweise erst einmal deutlich aufwändiger als die Gleichheit von Abbildungen, aber auf gewisse Art ist sie übersichtlicher, insbesondere, wenn wir später mehr Abbildungen verknüpfen und vergleichen wollen.)

Beweis: Wenn solch ein \bar{f} existiert, ist es eindeutig. Sei nämlich f' eine andere Abbildung mit $f = \iota \circ f' \circ \pi$. Dann gilt für alle $v \in V$

$$f(v) = \iota \bar{f}(\pi(v)) = \iota f'(\pi(v))$$

und da ι injektiv ist, gilt $\bar{f} = f'$ auf der Äquivalenzklasse $[v]$.

Nach Satz 3.4.3 gibt es nun $\tilde{f} : V/U \rightarrow W$ mit $f = \tilde{f} \circ \pi$ und per Definition faktorisiert \tilde{f} als $\iota \circ \bar{f}$. Wegen $\text{Im } \bar{f} = \text{Im } f$ ist \bar{f} trivialerweise surjektiv. Die Abbildung \bar{f} ist auch injektiv: denn gilt für $v \in V$

$$0 = \bar{f}([v]) = f(v) ,$$

so ist $v \in \ker f$, also $[v] = 0$. □

Beispiel 3.4.5. Sei $V = \mathbb{R}^3$ und $W = \mathbb{R}$ mit $f : V \rightarrow W$ gegeben durch Projektion auf die letzte Koordinate, $(x_1, x_2, x_3) \mapsto (x_3)$. Dann ist der Kern der Untervektorraum aller $(x_1, x_2, 0)$.

Wir machen den Vergleich von $V/\ker(f)$ mit $\text{Im } (f) = W$ explizit.

Eine Abbildung aus einem Quotientenvektorraum können wir wie oben immer definieren, indem wir $[v]$ auf irgendeine Funktion von v schicken, wir müssen nur aufpassen, dass diese Abbildung wohldefiniert ist.

Hier setzen wir $\bar{f}([v]) = f(v) \in W$. Wenn $v = (v_1, v_2, v_3)$ und $v' = (v'_1, v'_2, v'_3)$ äquivalent sind, dann gibt es $(x_1, x_2, 0)$ mit $v'_1 = v_1 + x_1, v'_2 = v_2 + x_2, v'_3 = v_3$. Es gilt $\bar{f}([v']) = v'_3 = v_3 = \bar{f}([v])$, also ist \bar{f} wohl definiert.

Statt zu prüfen, dass \bar{f} injektiv und surjektiv ist finden wir eine Umkehrabbildung. Um eine Abbildung nach $V/\ker(f)$ zu definieren können wir einfach eine Abbildung nach V definieren

und mit der kanonischen Projektion verknüpfen. Wir wählen $g : x \mapsto [(0, 0, x)]$ als Abbildung $W \rightarrow V/\ker f$.

Dann ist $\bar{f} \circ g(x) = \bar{f}([(0, 0, x)]) = f((0, 0, x)) = x$, also ist $\bar{f} \circ g = \text{id}_W$. Wir rechnen $g \circ \bar{f}([(x_1, x_2, x_3)]) = g(x_3) = [(0, 0, x_3)]$. Aber $[(0, 0, x_3)] = [(x_1, x_2, x_3)]$ und damit ist $g \circ \bar{f} = \text{id}_{V/\ker(f)}$.

In diesem Fall ist es praktisch die Äquivalenzklassen, die die Elemente des Quotienten bilden, durch $[(0, 0, x_3)]$ zu repräsentieren. Die Repräsentation mit den zwei Nullen ist hier intuitiv, aber es gibt andere Möglichkeiten, und in komplizierteren Beispielen ist es nicht immer möglich solche bevorzugten Repräsentanten zu finden, deshalb verwenden wir die abstrakte Konstruktion.

Beispiel 3.4.6. Seien K -Vektorräume U, W gegeben und sei $V = U \oplus W$ die (äußere direkte Summe). Betrachte die lineare Abbildung

$$f : V \rightarrow W$$

mit $f(u, w) := w$ für alle $u \in U$ und $w \in W$. Dann ist $\ker f = \{(u, 0)\} \cong U$ und $\text{Im } f = W$. Nach Theorem 3.4.4 gibt es einen eindeutigen Isomorphismus

$$\bar{f} : V/U \xrightarrow{\sim} W$$

mit $\bar{f}([w + u]) = f(u + w) = w$, also gilt $V/U \cong W$.

Beispiel 3.4.7. Sei V der Vektorraum aller konvergierenden reellen Folgen. (Dies ist ein Vektorraum, da die Summe von zwei konvergierenden Folgen wieder konvergiert, und genauso das Vielfache einer konvergierenden Folge.) Wir betrachten den Unterraum V_0 aller Folgen, die nach 0 konvergieren.

Dann sind in V/V_0 alle Folgen identifiziert, deren Differenz nach 0 konvergiert, also alle Folgen mit dem gleichen Grenzwert. Die Abbildung $(a_i)_{i \in \mathbb{N}} \mapsto \lim_{i \rightarrow \infty} a_i$ induziert dann einen Isomorphismus von V/V_0 nach \mathbb{R} .

4 Basen

Als nächstes wollen wir unsere Vektorräume konkreter machen. Der Unterschied zwischen K^n und einem allgemeinen abstrakten Vektorraum ist noch recht groß, aber die meisten Vektorräume, die wir von nun an betrachten, werden isomorph zu einem K^n sein. Die natürliche Zahl n ist hierbei die *Dimension* des Vektorraums. Der Isomorphismus eines beliebigen *endlich-dimensionalen* Vektorraums ist durch die Wahl einer *Basis* gegeben. Wie so oft brauchen wir etwas Vorarbeit.

4.1 Linearkombinationen

Wir beginnen mit einer Verallgemeinerung der Summe von Untervektorräumen. Gegeben eine Menge von Vektoren, wir wollen wissen, wie viele andere Vektoren sich als Summen von Vielfachen dieser Vektoren schreiben lassen.

Definition 4.1.1. 1. Sei V ein K -Vektorraum und seien endlich viele Elemente $v_1, \dots, v_m \in V$ gegeben. Ein Element $w \in V$ der Form

$$w = \lambda_1 v_1 + \dots + \lambda_m v_m$$

mit $\lambda_1, \dots, \lambda_m \in K$ heißt *Linearkombination* der Vektoren v_1, \dots, v_m .

2. Die Menge aller Linearkombinationen der Vektoren v_1, \dots, v_m

$$\text{span}_K(v_1, \dots, v_m) := \{\lambda_1 v_1 + \dots + \lambda_m v_m \mid \lambda_i \in K\}$$

heißt der von den Vektoren v_1, \dots, v_m aufgespannte Raum oder das *Erzeugnis* dieser Vektoren.

3. Sei V ein K -Vektorraum und $M \subset V$ eine Teilmenge von V . Dann heißt

$$\text{span}_K(M) := \{w = \lambda_1 v_1 + \dots + \lambda_m v_m \mid \lambda_i \in K, v_i \in M, m \in \mathbb{N}\}$$

die *lineare Hülle* der M oder ihr *Erzeugnis* oder *Spann*. Falls M leer ist dann enthält $\text{span}_K(M)$ genau die leere Summe, also den Nullvektor, und es gilt $\text{span}_K(M) = \{0\}$.

Also ist $\text{span}_K(v_1, \dots, v_m)$ nur eine andere Schreibweise für $\text{span}_K(\{v_1, \dots, v_m\})$.

Satz 4.1.2. Sei V ein K -Vektorraum, $v_1, \dots, v_m \in V$ und $M \subset V$ eine Teilmenge. Dann gilt:

1. $\text{span}_K(M)$ ist ein Untervektorraum von V und es gilt $M \subset \text{span}_K(M)$.
2. Ist $W \subset V$ ein Untervektorraum und $M \subset W$, so ist auch $\text{span}_K(M) \subset W$.
3. Es gilt

$$\text{span}_K(M) = \bigcap_{M \subset W, W \leq V} W$$

hier wird der Schnitt über alle Untervektorräume von V betrachtet, die die Menge M enthalten.

Beweis:. 1. (UV1): Sei $m \in M \neq 0$

$$0 = 0m \in \text{span}_K(M)$$

Falls M leer ist, dann ist 0 immer noch als leere Summe in $\text{span}(M)$!

(UV2) Seien $x, y \in \text{span}_K(M)$. Dann gibt es $\lambda_1, \dots, \lambda_m \in K$, und $\mu_1, \dots, \mu_k \in K$ sowie $v_1, \dots, v_m \in M$ und $w_1, \dots, w_k \in M$ so dass gilt

$$x = \lambda_1 v_1 + \dots + \lambda_m v_m \quad y = \mu_1 w_1 + \dots + \mu_k w_k.$$

(Die v_i und w_i können gleich oder unterschiedlich sein.) Somit ist

$$x + y \in \text{span}_K(M)$$

(UV3) folgt mit Notation wie in (UV2) sowie $\alpha \in K$:

$$\alpha x = (\alpha \lambda_1) v_1 + \dots + (\alpha \lambda_m) v_m \in \text{span}_K(M)$$

2. Ist $M \subset W$ und W ein Untervektorraum, so liegen nach (UV2) und (UV3) auch alle Linearkombinationen von Elementen in M in W , und damit alle Elemente von $\text{span}_K(M)$.

3. $\text{span}_K(M)$ ist nach 2. selbst ein Untervektorraum von V , der die Teilmenge M enthält. Also ist er einer der Unterräume, über die der Schnitt genommen wird, und es gilt

Nach 3. ist aber $\text{span}_K(M)$ in jedem der Untervektorräume, über die der Schnitt genommen wird, enthalten, also gilt auch die umgekehrte Inklusion

$$\text{span}_K(M) \subset \bigcap_{M \subset W, W \leq V} W. \quad \square$$

Die letzte Aussage des Satzes können wir auch formulieren als: $\text{span}_K(M)$ ist bezüglich der Inklusion der kleinste Untervektorraum von V , der M enthält.

Beispiele 4.1.3. 1. Sei K ein beliebiger Körper, den wir als Vektorraum über sich selbst betrachten, und sei $v \in K$. Für $v = 0$ ist $\text{span}_K(0) = \{0\}$ der triviale Untervektorraum von K ; für $v \neq 0$ ist $\text{span}_K(v) = K$.

In einem beliebigen Vektorraum V mit einem Element $v \neq 0$ ist $\text{span}_K(v) = Kv \cong K$, also eine Gerade.

2. Sei K ein beliebiger Körper und $V = K^n$. Setze wie in Betrachtung 3.3.6

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

dann ist $\text{span}_K(e_1, \dots, e_n) = V$, denn für jedes $x \in K^n$ gilt

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \dots + x_n e_n \in \text{span}_K(e_1, \dots, e_n)$$

3. Es gilt $W_1 + \dots + W_r = \text{span}_K(W_1 \cup \dots \cup W_r)$.

Jeder Vektor $v \in W_1 + W_2 + \dots + W_r$ lässt sich als Summe $v = \sum_{i=1}^r w_i$ mit $w_i \in W_i$ schreiben und liegt daher in der linearen Hülle $\text{span}_K(W_1 \cup \dots \cup W_r)$. Damit ist die Inklusion “ \subset ” klar.

Umgekehrt ist die Summe ein Untervektorraum, der $W_1 \cup \dots \cup W_r$ enthält, und damit gilt “ \supset ” nach Satz 4.1.2.4.

Der Spann einer Menge M von Vektoren gibt an, welche anderen Vektoren wir mit Vektoren aus M schreiben können. Als nächstes können wir uns fragen, auf wie viele verschiedene Arten wir Vektoren mit Elementen aus M schreiben können. Da wir mit linearen Räumen arbeiten, reicht es zu bestimmen, ob wir den Nullvektor auf verschiedene Arten schreiben können.

Definition 4.1.4. Sei V ein K -Vektorraum.

1. Eine endliche nichtleere Menge $\{v_1, \dots, v_r\}$ von Vektoren aus V heißt *linear unabhängig*, falls gilt: sind $\lambda_1, \dots, \lambda_r \in K$ und gilt

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0,$$

so folgt daraus $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$. Die Familie $\{v_1, \dots, v_r\}$ heißt also genau dann linear unabhängig, wenn der Nullvektor sich nur trivial als Linearkombination von v_1, \dots, v_k darstellen lässt.

2. Eine beliebige Menge von Vektoren aus V heißt linear unabhängig, wenn jede *endliche nichtleere* Teilmenge linear unabhängig ist.

3. Andernfalls heißt die Menge *linear abhängig*; dann gibt es eine Darstellung des Nullvektors als nicht triviale Linearkombination, d.h. es gibt $\lambda_i \in K$ mit

$$0 = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

wobei nicht alle $\lambda_i \in K$ verschwinden.

⟨⟨Eine Menge ist nicht linear (un)abhängig *von* etwas, sondern einfach linear (un)abhängig. Gemeint ist: Die Elemente sind linear (un)abhängig voneinander.⟩⟩

Bemerkungen 4.1.5. 1. Es ist \emptyset *linear unabhängig* (denn die Bedingung ist trivialerweise erfüllt).

2. In K^n ist jede Teilmenge der Menge $\{e_1, \dots, e_n\}$ aus Beispiel 4.1.3.2 linear unabhängig.

3. Ein einziger 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 3.1.3.3 folgt nun $v = 0$.

4. Auch für eine Familie $(v_\lambda)_{\lambda \in \Lambda}$ definieren wir lineare Unabhängigkeit: $(v_\lambda)_{\lambda \in \Lambda}$ ist linear unabhängig wenn für jede endliche Unterfamilie (v_1, \dots, v_n) gilt: Wenn $\mu_1 v_1 + \dots + \mu_n v_n = 0$ dann ist $\mu_1 = \dots = \mu_n = 0$.

Man beachte, dass eine Familie mehrere gleiche Vektoren enthalten kann. Dann ist sie aber jedenfalls linear abhängig, denn gilt $v_1 = v_2$, so finden wir die nicht-triviale Linearkombination $1 \cdot v_1 + (-1)v_2 = 0$.

Lemma 4.1.6. Für eine Menge $\{v_1, \dots, v_r\}$ von Vektoren eines K -Vektorraums sind die folgenden Bedingungen äquivalent:

1. Die Menge $\{v_i\}$ ist linear unabhängig.
2. Jeder Vektor $v \in \text{span}_K(v_i)$ lässt sich in eindeutiger Weise als Linearkombination von Vektoren der Menge $\{v_i\}$ schreiben.

Beweis: 2. \Rightarrow 1. klar, denn bei einer linear abhängigen Menge hat der Nullvektor verschiedene Darstellungen.

1. \Rightarrow 2. Aus

$$v = \sum_i \lambda_i v_i = \sum_i \mu_i v_i \quad \text{mit } \lambda_i, \mu_i \in K$$

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

Definition 4.1.7. Eine Familie mit Indexmenge \mathbb{N} oder $\underline{n} = \{1, 2, \dots, n\}$ heißt *geordnete Familie*, wir sagen $v_i < v_j$ genau wenn $i < j$ in der

Lemma 4.1.8. Eine geordnete Familie (v_1, \dots, v_n) von Vektoren in einem K -Vektorraum V definiert eine Abbildung $g : K^n \rightarrow V$. Die Abbildung g ist genau dann surjektiv, wenn $\text{span}_K(v_1, \dots, v_n) = V$, und genau dann injektiv ist, wenn (v_1, \dots, v_n) linear unabhängig ist.

Beweis: Um g zu definieren schicken wir einfach e_i nach v_i und damit ist $g(x) = g(\sum_{i=1}^n x_i e_i) = \sum_{i=1}^n x_i v_i$ eindeutig festgelegt. Die Abbildung ist linear.

Genau dann, wenn $\text{span}_K(v_1, \dots, v_n) = V$, lässt sich jedes v schreiben als $\sum \lambda_i v_i = g(\sum \lambda_i e_i)$.

Sei nun g injektiv. Dann folgt aus $\sum x_i v_i = 0 \Leftrightarrow g(x) = 0$, dass $x = 0$ ist, also $x_i = 0$ für alle i . Ist umgekehrt (v_1, \dots, v_n) linear unabhängig und $g(x) = 0$, dann folgt aus $g(x) = \sum x_i v_i = 0$ genau, dass alle $x_i = 0$ und damit $x = 0$. \square

Zur Abwechslung formulieren wir das nächste Lemma für Familien, es gilt genauso für Mengen.

4.2 Basis und Dimension

Definition 4.2.1. Sei K ein Körper und V ein K -Vektorraum.

1. Eine Teilmenge $M \subset V$ heißt *Erzeugendensystem* von V , falls $\text{span}_K(M) = V$ gilt. Wir sagen auch M erzeugt V .
2. Eine Teilmenge $M \subset V$ heißt *Basis* von V , falls M ein linear unabhängiges Erzeugendensystem ist.

Manchmal ist es wichtig, eine Reihenfolge der Elemente der Basis festzulegen.

Definition 4.2.2. Eine *geordnete Basis* von V ist eine geordnete Familie (v_1, \dots, v_n) , die linear unabhängig ist und so dass die Menge $\{v_1, \dots, v_n\}$ ein Erzeugendensystem von V ist.

(Man kann auch für passende unendliche Indexmengen geordnete Basen definieren, aber das soll uns hier nicht interessieren.)

Beispiele 4.2.3. 1. Jeder Vektorraum V besitzt ein Erzeugendensystem, zum Beispiel sich selbst, $M = V$. Es ist nicht offensichtlich, ob jeder Vektorraum eine Basis besitzt.

2. $V = K^n$ besitzt die Basis $\{e_i\}_{i=1\dots n}$ aus Beispiel 4.1.3.2. Wir nennen diese Basis die *Standardbasis des K^n* .
3. Endliche geordnete Basen schreiben wir auch in der Form (v_1, v_2, \dots, v_n) . Die Standardbasis des K^n ist durch (e_1, e_2, \dots, 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.)
4. $K = \mathbb{R}$ und $V = \mathbb{C}$, dann ist $M = \{1, i\}$ eine \mathbb{R} -Basis von \mathbb{C} . Auch $\{1, -i\}$ ist eine \mathbb{R} -Basis von \mathbb{C} , oder $\{1 + i, 1 - 2i\}$. Eine Basis von \mathbb{C} als \mathbb{C} -Vektorraum ist einfach von $\{1\}$ (oder jeder anderen komplexen Zahl ungleich 0) gegeben.
5. $K = \mathbb{R}$ und $V = \mathbb{R}[X]$ der Vektorraum der Polynome mit reellen Koeffizienten, dann ist die Folge der Monome $\{1, X, \dots, X^k, \dots\}$ eine (unendliche!) Basis von V .

Wenn wir eine endliche geordnete Basis für einen Vektorraum V haben, dann haben wir nach Lemma 4.1.8 einen Isomorphismus $K^n \rightarrow V$ wobei n die Anzahl der Elemente in der Basis ist und *Dimension* von V heißt. Dann ist jeder Vektor eindeutig durch seine Komponenten bestimmt und lineare Abbildungen sind einfach Matrizen. Wir reduzieren lineare Algebra auf die Manipulation von Zeilen und Spalten von Zahlen.

Allerdings stellen sich ein paar offensichtliche Fragen. Ist die Dimension überhaupt wohldefiniert? Müssen alle Basen eines Vektorraums die gleiche Anzahl von Elementen haben? Ist die Dimension eines Untervektorraums von V immer kleiner oder gleich der Dimension von V ? Ist die Dimension eines Untervektorraums überhaupt endlich, wenn V endlich-dimensional ist?

Diese Fragen werden wir als nächstes beantworten, aber dazu brauchen wir wieder einmal Theorie.

Satz 4.2.4. *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: Wir müssen nur drei Implikationen zeigen:

2. \Rightarrow 1. Sei M minimales Erzeugendensystem. Angenommen, $M = \{v_i\}_{i \in I}$ ist linear abhängig. Wir finden also eine nicht-triviale Linearkombination

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0$$

für die etwa $\lambda_r \neq 0$ ist. (Wir können die Vektoren immer umnummerieren, um diese Bedingung zu erfüllen.) Dann ist

$$v_r = -\frac{\lambda_1}{\lambda_r} v_1 - \dots - \frac{\lambda_{r-1}}{\lambda_r} v_{r-1}.$$

(Beachten Sie, dass wir hier zum ersten Mal die Division in K benutzen!) Dann ist aber auch schon (v_1, \dots, v_{r-1}) 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 $M \cup \{v\}$ linear abhängig.

3. ⇒ 2. Sei M eine maximale linear unabhängige Teilmenge. Wir zeigen zuerst, dass M Erzeugendensystem ist. Wenn nicht, dann gibt es $v \in V \setminus \text{span}(M)$. Aber dann ist $M \cup \{v\}$ linear unabhängig, denn aus $\lambda v + \sum_{w \in M} \lambda_w w = 0$ mit $\lambda \neq 0$ würde $v \in \text{span}(M)$ folgen. Wenn aber $\lambda = 0$ ist sind auch alle $\lambda_w = 0$ da M linear unabhängig ist.

Wir zeigen noch, dass M minimal ist. Angenommen es gibt $v \in M$ so dass $M \setminus \{v\}$ ein Erzeugendensystem ist. Dann können wir $v = \sum_{w \in M \setminus \{v\}} \lambda_w w$ schreiben, und das widerspricht der linearen Unabhängigkeit von M . \square

Definition 4.2.5. Ein K -Vektorraum V heißt *endlich erzeugt*, falls V ein endliches Erzeugendensystem besitzt, d.h. falls es eine endliche Teilmenge $M = \{v_1, \dots, v_m\}$ von V gibt, so dass $V = \text{span}_K(M)$ gilt.

Lemma 4.2.6. Sei V ein K -Vektorraum, der nicht endlich erzeugt ist. Dann gibt es zu jeder natürlichen Zahl $n \in \mathbb{N} \setminus \{0\}$ Vektoren $v_1, \dots, v_n \in V$, so dass die Familie (v_1, v_2, \dots, 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 4.1.3.1.
- Induktionsschritt:
Sei $n \in \mathbb{N} \setminus \{0\}$ und seien $v_1, \dots, v_n \in V$ linear unabhängig. Wähle $v_{n+1} \in V \setminus \text{span}_K(v_1, \dots, v_n)$. Solch ein v_{n+1} existiert, da andernfalls V von den (v_1, \dots, v_n) , also endlich erzeugt wäre. Es bleibt zu zeigen, dass auch die Familie $(v_1, \dots, v_n, v_{n+1})$ linear unabhängig ist. Sei

$$0 = \sum_{j=1}^{n+1} \alpha_j v_j \quad \alpha_j \in K$$

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 Widerspruch zu unserer Wahl von $v_{n+1} \notin \text{span}_K(v_1, \dots, v_n)$ 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 $j = 1, \dots, n$. \square

Nicht jeder Vektorraum hat ein endliches Erzeugendensystem, zum Beispiel nicht der K -Vektorraum der Polynome $K[X]$ über einem beliebigen Körper K .

Satz 4.2.7 (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 $M = \{v_1, \dots, v_n\}$ ein Erzeugendensystem von V . Ist M auch linear unabhängig, so ist $\mathcal{B} = M$ und wir sind fertig.

Ist M nicht linear unabhängig, so ist nach Satz 4.2.4.2 das Erzeugendensystem M nicht minimal. Es gibt also ein $v \in M$ mit $v \in \text{span}_K(M \setminus \{v\})$ 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 4.2.4 dann eine Basis. \square

Dieser Beweis funktioniert nur für endlich erzeugte Vektorräume. Auch Vektorräume, die nicht endlich erzeugt sind, haben eine Basis. Um das zu zeigen, braucht man das Auswahlaxiom der Mengenlehre. Wir verschieben diesen Beweis auf später.

Satz 4.2.8 (Austauschlemma). Sei V ein K -Vektorraum, $\mathcal{B} = \{v_1, \dots, v_r\} \subset V$ eine Basis. Sei $w = \sum_{j=1}^r \alpha_j v_j \in V$ mit $\alpha_j \in K$. Dann gilt für jedes $k \in \{1, \dots, r\}$ mit $\alpha_k \neq 0$:

$$\mathcal{B}'_k := \{v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r\}$$

ist eine Basis, d.h. das Basiselement v_k kann gegen w ausgetauscht werden.

Beweis:. • Nach Umm Nummerierung können wir $k = 1$ annehmen.

- Wir zeigen: \mathcal{B}' ist ein Erzeugendensystem. Für jedes gegebene $v \in V$ existieren $\beta_j \in K$ für $j = 1, \dots, r$, so dass

$$v = \sum_{j=1}^r \beta_j v_j \quad (*)$$

gilt. Aus $\alpha_1 \neq 0$ folgt

$$v_1 = \frac{1}{\alpha_1} w + \sum_{j=2}^r \left(-\frac{\alpha_j}{\alpha_1}\right) v_j.$$

Dies setzen wir in (*) ein und erhalten

$$v = \frac{\beta_1}{\alpha_1} w + \sum_{j=2}^r \left(\beta_j - \alpha_j \frac{\beta_1}{\alpha_1}\right) v_j.$$

Somit ist $V \subset \text{span}_K(\mathcal{B}')$, also ist \mathcal{B}' Erzeugendensystem.

- Wir zeigen: \mathcal{B}' ist linear unabhängig. Seien $\beta, \beta_j \in K$ 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:

$$\beta \alpha_1 v_1 + \sum_{j=2}^r (\beta \alpha_j + \beta_j) v_j = 0$$

Da die Familie $\{v_1, \dots, v_r\}$ linear unabhängig ist, folgt

$$\beta \alpha_1 = 0 \quad \text{und} \quad \beta \alpha_j + \beta_j = 0$$

Aus $\alpha_1 \neq 0$ folgt $\beta = 0$ und daraus $\beta_j = 0$.

\square

Satz 4.2.9 (Austauschsatz). Sei V ein K -Vektorraum, und $\mathcal{B} = \{v_1, \dots, v_r\}$ eine Basis von V . Sei $\{w_1, \dots, w_n\}$ eine linear unabhängige Teilmenge von V . Dann gilt $n \leq r$, und es gibt $i_1, \dots, i_n \in \{1, \dots, r\}$, so dass der Austausch

$$\begin{aligned} & \text{von } v_{i_1} \text{ gegen } w_1, \dots \\ & \text{von } v_{i_n} \text{ gegen } w_n \end{aligned}$$

eine neue Basis \mathcal{B}^* von V liefert, die die vorgegebene linear unabhängige Menge $\{w_1, \dots, w_n\}$ als Teilmenge enthält. Nach Umnummerierung zu $i_1 = 1, \dots, i_n = n$ haben wir für die neue Basis

$$\mathcal{B}^* = \{w_1, \dots, w_n, v_{n+1}, \dots, v_r\}.$$

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 $\{w_1, \dots, w_n\}$ linear unabhängig. Dann ist auch die Teilmenge $\{w_1, \dots, w_{n-1}\}$ linear unabhängig. Nach Induktionsvoraussetzung gilt $n - 1 \leq r$ und (gegebenenfalls nach Umnummerierung) ist die Menge

$$\bar{\mathcal{B}} := \{w_1, \dots, w_{n-1}, v_n, \dots, v_r\}$$

eine Basis von V .

- Wir zeigen $n \leq r$. Nach Induktionsvoraussetzung ist $n - 1 \leq r$; wir müssen $n - 1 = r$ ausschließen. Dann wäre aber nach Induktionsvoraussetzung die Menge

$$\bar{\mathcal{B}} := \{w_1, \dots, w_{n-1}\}$$

eine Basis von V , also eine maximale lineare unabhängige Teilmenge nach Satz 4.2.4. Das steht im Widerspruch zur Voraussetzung, dass auch noch $\{w_1, \dots, w_{n-1}, w_n\}$ linear unabhängig ist. Also ist $n - 1 < r$, also $n \leq r$.

- Zu zeigen ist, dass es ein $i_n \in \{n, \dots, r\}$ 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 $\alpha_n, \dots, \alpha_r$ gleich Null, so wäre w_n Linearkombination der $\{w_1, \dots, w_{n-1}\}$, im Widerspruch zur vorausgesetzten linearen Unabhängigkeit von $\{w_1, \dots, w_n\}$. Also gibt es $i_n \in \{n, \dots, r\}$ mit $\alpha_{i_n} \neq 0$. Wende nun das Austauschlemma 4.2.8 an und erhalte eine Basis $\mathcal{B}^* = \{w_1, \dots, w_n, v_{n+1}, \dots, v_r\}$.

⟨⟨Wir können auch direkt vom ersten zum dritten Schritt gehen und die zweite Überlegung auslassen. Dann müssen wir sorgfältig den Fall $r = n - 1$ prüfen wenn die Summe $\sum_{j=n}^r \alpha_j v_j$ leer ist. Dies führt zum Widerspruch dazu, dass $\{w_1, \dots, w_n\}$ linear unabhängig ist.⟩⟩ □

Korollar 4.2.10 (Basisergänzungssatz). Jede linear unabhängige Menge M in einem endlich-dimensionalen Vektorraum V lässt sich zu einer Basis ergänzen.

Beweis: Siehe Übung 9.1.3. □

Definition 4.2.11. Sei V ein endlich erzeugter K -Vektorraum und $\mathcal{B} = \{v_1, \dots, v_r\}$ eine Basis. Die Zahl r heißt *Länge* der Basis \mathcal{B} .

Korollar 4.2.12. 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.

Beweis:. 1. Sei $\{v_1, \dots, v_r\}$ eine endliche Basis. Wäre eine weitere Basis \mathcal{B} nicht endlich, gäbe es eine linear unabhängige Teilmenge $\{w_1, \dots, w_{r+1}\} \subset \mathcal{B}$, im Widerspruch zum Austauschsatz 4.2.9.

2. Sind $\mathcal{B} = \{v_1, \dots, v_r\}$ und $\mathcal{B}' = \{w_1, \dots, w_k\}$ 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 $k = r$. \square

Definition 4.2.13. 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

$$\dim_K(V) \in \{0, 1, \dots, \infty\}$$

heißt *Dimension* des Vektorraums V .

Beispiele 4.2.14. 1. Sei K ein beliebiger Körper und $V = K^n$. Dann hat die Standardbasis e_1, \dots, e_n die Länge n ; und daher ist $\dim_K K^n = n$.

2. $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, denn $\{1, i\}$ ist eine \mathbb{R} -Basis. Es gilt $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, 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 $\{1, X, X^2, \dots\}$ ist eine Basis.

Satz 4.2.15. Sei V ein endlich erzeugter K -Vektorraum und $W \subset V$ ein Untervektorraum. Dann ist W endlich erzeugt und es gilt

$$\dim_K(W) \leq \dim_K(V).$$

Falls $\dim_K(W) = \dim_K(V)$, so ist $W = V$.

Beweis:. Setze $n := \dim_K(V) < \infty$. Wäre W nicht endlich erzeugt, so gäbe es nach Lemma 4.2.6 sicher $n + 1$ linear unabhängige Vektoren $v_1, \dots, v_{n+1} \in W \subset V$, im Widerspruch zum Austauschsatz 4.2.9.

Also besitzt W eine endliche Basis $\mathcal{B} = \{w_1, \dots, w_r\}$; da diese Familie linear unabhängig in V ist, folgt nach dem Austauschsatz $r = \dim_K(W) \leq \dim_K(V)$.

Sei nun $\dim_K(V) = \dim_K(W) = n$ und $\mathcal{B} = \{w_1, \dots, w_n\}$ eine Basis von W . Gäbe es $v \in V \setminus \text{span}_K(\mathcal{B})$, so wäre $\mathcal{B} \cup \{v\}$ linear unabhängig, im Widerspruch zum Austauschsatz 4.2.9. \square

4.3 Abbildungen und Dimension

Ein Ziel unseres Studiums von Vektorräumen ist die Klassifizierung von linearen Abbildung. Als ersten Schritt können wir lineare Abbildungen mit Dimension in Zusammenhang setzen.

Definition 4.3.1. Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann heißt

$$\text{rg}(f) := \dim_K \text{Im } f$$

der *Rang* der linearen Abbildung f .

Beispiel 4.3.2. Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $e_1 \mapsto e_1 + e_2$ und $e_2 \mapsto 2e_1 + 2e_2$ hat Bild $\text{span}_{\mathbb{R}}(e_1 + e_2)$ mit Dimension 1. Also ist $\text{rg}(f) = 1$.

Bemerkung 4.3.3. Sei $f : V \rightarrow W$ eine lineare Abbildung und $E \subset V$ ein Erzeugendensystem von V . Dann ist $f(E)$ ein Erzeugendensystem des Bildes $f(V)$. Denn für $w \in f(V)$ existiert $v \in V$ mit $f(v) = w$. Weil E ein Erzeugendensystem von V ist, können wir für dieses $v \in V$ Skalare $\lambda_1, \dots, \lambda_n \in K$ sowie $v_1, \dots, v_n \in E$ finden, so dass

$$v = \sum_{i=1}^n \lambda_i v_i$$

gilt, woraus

$$w = f(v) = \sum \lambda_i f(v_i)$$

folgt.

Der nächste Satz ist extrem nützlich, um Dimensionen zu berechnen.

Satz 4.3.4 (*Dimensionsformel*). Sei $f : V \rightarrow W$ eine lineare Abbildung und $\dim_K V < \infty$. Dann gilt die Dimensionsformel

$$\dim_K V = \dim_K \ker f + \text{rg}(f).$$

Beweis: $\ker f$ ist ein endlich-dimensionaler Untervektorraum von V ; wähle eine Basis $\{v_1, \dots, v_k\}$ von $\ker f$ mit $k := \dim_K \ker f$ und ergänze diese nach dem Basisergänzungssatz Korollar 4.2.10 zu einer Basis $\{v_1, \dots, v_n\}$ von V mit $n = \dim_K V$. Nach Bemerkung 4.3.3 ist $\{f(v_1), \dots, f(v_n)\}$ ein Erzeugendensystem des Bildes $f(V)$. Aber $0 = f(v_1) = \dots = f(v_k)$, also ist schon die Teilmenge

$$\{f(v_{k+1}), \dots, f(v_n)\}$$

ein Erzeugendensystem von $f(V)$.

Wir zeigen nun, dass dieses Erzeugendensystem von $f(V)$ linear unabhängig und somit eine Basis von $f(V)$ ist. Seien $\lambda_{k+1}, \dots, \lambda_n \in K$ und gelte

$$\sum_{j=k+1}^n \lambda_j f(v_j) = 0.$$

Hieraus folgt $f\left(\sum_{i=k+1}^n \lambda_i v_i\right) = 0$, und somit $\sum_{i=k+1}^n \lambda_i v_i \in \ker f = \text{span}_K(v_1, \dots, v_k)$. Aber da $\{v_1, \dots, v_n\}$ linear unabhängig ist gilt $\text{span}_K(v_1, \dots, v_k) \cap \text{span}_K(v_{k+1}, \dots, v_n) = \{0\}$. \square

Korollar 4.3.5. Sei $f : V \rightarrow W$ eine lineare Abbildung und $\dim_K V = \dim_K W < \infty$. Dann sind äquivalent:

1. f ist injektiv
2. f ist surjektiv
3. f ist bijektiv

Beweis: Es genügt, die Äquivalenz von 1. und 2. zu zeigen. f ist genau dann injektiv, wenn der Kern trivial ist, $\ker f = \{0\}$. Dies ist genau dann der Fall, wenn $\dim_K \ker f = 0$ gilt. Wegen der Dimensionsformel 4.3.4 ist dies äquivalent zu $\dim_K W = \dim_K V = \dim_K \operatorname{Im} f$. Aus Satz 4.2.15 folgt, dass dies äquivalent zu $W = \operatorname{Im} f$ ist. Genau dann ist aber die lineare Abbildung f surjektiv. \square

Bemerkung 4.3.6. Achtung: Korollar 4.3.5 gilt *nicht* für unendlich-dimensionale Vektorräume! Ein Gegenbeispiel ist für $K = \mathbb{R}$ und $V = \mathbb{R}[x]$ die Abbildung $\operatorname{diff} : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ aus Übung 7.6, die als Kern die Polynome vom Grad Null hat. Sie ist also nicht injektiv; aber sie ist surjektiv.

Wir haben ein Korollar des Korollars:

Korollar 4.3.7. Sei V ein endlich erzeugter K -Vektorraum mit $n = \dim_K(V)$ und sei $\mathcal{B} = (v_1, \dots, v_n)$. Es sind äquivalent:

1. \mathcal{B} ist eine Basis.
2. \mathcal{B} ist linear unabhängig.
3. \mathcal{B} ist ein Erzeugendensystem.

Beweis: Nach Lemma 4.1.8 definiert $\mathcal{B} = (v_1, \dots, v_n)$ eine lineare Abbildung f von K^n nach V und diese Abbildung ist injektiv genau wenn \mathcal{B} linear unabhängig ist und surjektiv genau wenn \mathcal{B} ein Erzeugendensystem ist. Nach Korollar 4.3.5 sind die Bedingungen äquivalent.

Alternativ lässt sich der Beweis auch leicht mit den Techniken des letzten Abschnitts führen. \square

Wir können nun alle endlich erzeugten (äquivalent: endlichdimensionalen) Vektorraum bis auf Isomorphismus klassifizieren. Das heißt, wir haben eine Liste von Vektorräumen und jeder endlich-dimensionale Vektorraum ist isomorph zu genau einem Eintrag in der Liste.

Satz 4.3.8. Jeder endlich erzeugte K -Vektorraum V ist isomorph zu K^n , wobei $n = \dim(V)$. Zwei endlich erzeugte Vektorräume V und W sind genau dann isomorph, wenn sie die gleiche Dimension haben.

Es gibt also für jeden Körper und jedes n (bis auf Isomorphie) genau einen endlichdimensionalen Vektorraum der Dimension n ! Dies rechtfertigt unser besonderes Interesse an K^n .

Beweis: Wir wählen nach Satz 4.2.7 eine Basis von V und machen sie zu einer geordneten Basis $\mathcal{B} = (v_1, \dots, v_n)$. Dann ist nach Definition $n = \dim_K(V)$. Laut Lemma 4.1.8 definiert dies einen Isomorphismus $f : K^n \rightarrow V$. (Wir erinnern uns: f ist surjektiv, da \mathcal{B} Erzeugendensystem ist, f ist injektiv da \mathcal{B} linear unabhängig ist.)

Habe auch der Vektorraum W Dimension n , dann gilt nach Transitivität und Symmetrie des Isomorphismus $V \cong W$ denn $V \cong K^n$ und $K^n \cong W$.

Sei umgekehrt $f : V \cong W$ ein Isomorphismus. Dann gilt nach der Dimensionsformel 4.3.4 $\dim V = \dim \ker(f) + \operatorname{rg}(f) = 0 + \dim W$. \square

Für die nächste Anwendung brauchen wir ein kleines Lemma:

Lemma 4.3.9. *Seien V, W K -Vektorräume. Dann gilt $\dim_K(V \oplus W) = \dim_K V + \dim_K W$.*

Beweis: Wir betrachten Basen $\{e_1, \dots, e_m\}$ für V und $\{f_1, \dots, f_n\}$ für W . Wir behaupten, dass $\{(e_1, 0), \dots, (e_m, 0), (0, f_1), \dots, (0, f_n)\}$ eine Basis für $V \oplus W$ ist. Lineare Unabhängigkeit folgt, da $\{e_i\}$ und $\{f_j\}$ linear unabhängig sind. Weiterhin erzeugt diese Menge $V \oplus W$, da jeder Vektor (v, w) sich als $(\sum_i \lambda e_i, \sum_j \mu_j f_j) = \sum_i \lambda(e_i, 0) + \sum_j \mu_j(0, f_j)$ schreiben lässt. \square

Satz 4.3.10. *Sei V ein K -Vektorraum und $W_1, W_2 \leq V$ 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

$$\begin{aligned} W_1 &= \text{span}(e_1, e_2) : && x\text{-}y\text{-Ebene} \\ W_2 &= \text{span}(e_1, e_3) : && x\text{-}z\text{-Ebene} \\ W_1 \cap W_2 &= \text{span}(e_1) : && x\text{-Achse} \\ W_1 + W_2 &= \mathbb{R}^3 \end{aligned}$$

Die Dimensionsformel aus Satz 4.3.10 ergibt $3 = 2 + 2 - 1$.

Beweis: Dieser Satz folgt sofort aus Satz 4.3.4 indem wir unsere Vektorräume geschickt wählen. Wir wenden die Dimensionsformel auf die Summenabbildung $\sigma : W_1 \oplus W_2 \rightarrow W_1 + W_2$ an, die definiert ist durch $(w_1, w_2) \mapsto w_1 + w_2$.

Nach Definition ist σ surjektiv, also ist $\text{rg}(\sigma) = \dim_K(W_1 + W_2)$. Weiterhin ist $\dim_K(W_1 \oplus W_2) = \dim_K W_1 + \dim_K W_2$ nach Lemma 4.3.9. Der Kern von σ besteht genau aus $(w_1, -w_1)$ mit $w_1 \in W_1 \cap W_2$, also ist $\dim_K \ker(\sigma) = \dim_K(W_1 \cap W_2)$.

Wir nehmen all dies zusammen und erhalten für $\text{rg}(\sigma) = \dim_K(W_1 \oplus W_2) - \dim \ker(\sigma)$:

$$\dim_K(W_1 + W_2) = \dim_K W_1 + \dim_K W_2 - \dim_K(W_1 \cap W_2). \quad \square$$

5 Lineare Abbildungen und Basen

5.1 Lineare Abbildungen und Matrizen

Wir haben in Betrachtung 3.3.6 gesehen, dass eine lineare Abbildung auf K^n durch das Bild der Einheitsbasis (e_1, \dots, e_n) bestimmt wird und als Matrix geschrieben werden kann. Aber die Einheitsbasis spielt keine besondere Rolle: es gilt, dass jede lineare Abbildung mit der Wahl von Basen als Matrix geschrieben werden kann.

Die Matrix hängt dann aber nicht nur von der lineare Abbildung, sondern auch von der Wahl der Basen ab!

Wir haben einen vorbereitenden Satz.

Satz 5.1.1. *Gegeben seien endlich-dimensionale Vektorräume V und W sowie Vektoren $v_1, \dots, v_r \in V$ und $w_1, \dots, w_r \in W$.*

1. *Ist (v_1, \dots, v_r) eine geordnete Basis von V , so gibt es genau eine lineare Abbildung $f : V \rightarrow W$ mit $f(v_i) = w_i$. Diese hat die beiden Eigenschaften:*

(a) $f(V) = \text{Im } f = \text{span}_K(w_1, \dots, w_r)$

(b) *Die Abbildung f ist genau dann injektiv, wenn die Familie (w_1, \dots, w_r) in W linear unabhängig ist.*

2. *Ist die Familie (v_1, \dots, v_r) in V linear unabhängig (aber nicht unbedingt eine Basis), so gibt es mindestens eine lineare Abbildung $f : V \rightarrow W$ mit $f(v_i) = w_i$ für $i = 1, \dots, r$.*

Beweis:. 1. Wir haben in der Vorlesung einen direkten Beweis geführt.

Hier folgt ein Beweis, in dem wir Lemma 4.1.8 verwenden. Die geordnete Basis (v_1, \dots, v_r) definiert einen Isomorphismus $g : K^r \rightarrow V$ nach. Die geordnete Familie (w_1, \dots, w_r) definiert durch $e_i \mapsto w_i$ definiert eine lineare Abbildung $h : K^r \rightarrow W$ und das Bild von h ist $\text{span}_K(w_1, \dots, w_r)$; h ist injektiv genau wenn die (w_i) linear unabhängig sind.

Aber dann ist $h \circ g^{-1}$ die gesuchte Abbildung f und da g ein Isomorphismus ist gilt $\text{Im}(f) = \text{Im}(h) = \text{span}_K(w_1, \dots, w_r)$ und f ist injektiv, wenn h injektiv ist, also wenn die (w_j) linear unabhängig sind.

Die Abbildung ist eindeutig: Sei f' eine andere Abbildung mit $f'(v_i) = w_i$, dann folgt $f'(v) = f'(\sum \lambda_i v_i) = \sum \lambda_i f'(v_i) = \sum \lambda_i w_i = f(v)$ da (v_i) eine Basis ist.

2. Ist die Familie (v_1, \dots, v_r) nur linear unabhängig, aber keine Basis, so können wir mit Hilfe des Basisergänzungssatzes Korollar 4.2.10 die Familie zu einer Basis von V ergänzen:

$$(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$$

und ein f durch Vorgabe beliebiger Werte $w_{r+1}, \dots, w_n \in W$ für v_{r+1}, \dots, v_n wie in 2. festlegen. □

Wir betrachten als nächstes lineare Abbildungen

$$f : K^n \rightarrow K^m.$$

Damit können wir (bis auf Isomorphismen) alle linearen Abbildungen zwischen endlich-dimensionalen Vektorräumen verstehen.

Definition 5.1.2. Sei K ein Körper.

1. Ein rechteckiges Schema der Form

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

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 $M(m \times n, K)$. Wir schreiben oft eine Großbuchstaben A für eine Matrix und A_{ij} oder a_{ij} für den Eintrag in i -ter Zeilen und j -ter Spalte.

2. Sei $f : K^n \rightarrow K^m$ eine lineare Abbildung. Es seien

$$f(e_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, f(e_n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

mit $a_{ij} \in K$ die Bilder der Vektoren (e_1, \dots, e_n) der Standardbasis von K^n . Dann heißt

$$M(f) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

die *darstellende Matrix* von f (bezüglich der Standardbasen).

Beispiel 5.1.3. Vektoren in K^n sind Spezialfälle von Matrizen. Wir sollten unterscheiden zwischen (vertikalen) Spaltenvektoren, die $n \times 1$ -Matrizen sind, und (horizontalen) Zeilenvektoren, die $1 \times n$ -Matrizen sind.

Der Spaltenvektor $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ ist die darstellende Matrix für die lineare Abbildung $K \rightarrow K^n$,

die λ nach λv schickt.

Der Zeilenvektor $w = (w_1, \dots, w_n)$ repräsentiert die lineare Abbildung $K^n \rightarrow K$, die $\sum \lambda_i e_i$ nach $\sum \lambda_i w_i$ schickt.

Wenn wir einen Spaltenvektor als Zeilenvektor auffassen wollen und umgekehrt, schreiben

wir v^T , also $v^T = (v_1, \dots, v_n)^T = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$.

Wir verwenden folgende Konvention: Abbildungen werden mit Kleinbuchstaben bezeichnet und Matrizen mit Großbuchstaben. (Allerdings schreiben wir oft Kleinbuchstaben für Matrixeinträge.)

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 5.1.1 folgt, dass $M(f)$ und f 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

$$f(v) = f\left(\sum_{i=1}^n v_i e_i\right) = \sum_{i=1}^n v_i f(e_i) = v_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + v_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} v_j \\ \vdots \\ \sum_{j=1}^n a_{mj} v_j \end{pmatrix}.$$

Definition 5.1.4. 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

$$A \cdot v = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^n a_{1j} v_j \\ \vdots \\ \sum_{j=1}^n a_{mj} v_j \end{pmatrix}$$

Beispiel 5.1.5. Wir setzen $K = \mathbb{R}$ und $n = m = 2$ und betrachten Drehungen um den Ursprung, vgl. Beispiel 3.3.4.2. Mit Hilfe des Produkts einer Matrix mit einem Vektor erhält man mit $\theta \in \mathbb{R}$

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta x - \sin \theta y \\ \sin \theta x + \cos \theta y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

als darstellende Matrix

$$M(R_\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Betrachtung 5.1.6. Seien $f : K^n \rightarrow K^m$ und $g : K^m \rightarrow K^l$ lineare Abbildungen. Nach Bemerkung 3.3.4.5 ist dann $g \circ f$ wieder eine lineare Abbildung. Wir wollen deren darstellende Matrix $M(g \circ f)$ bestimmen. Dazu benutzen wir die Standardbasen $(e_1, \dots, e_n) \in K^n$ und $(e'_1, \dots, e'_m) \in K^m$. Seien

$$M(g) = (a_{ij}), \quad M(f) = (b_{ij}), \quad M(g \circ f) = (c_{ij})$$

die darstellenden Matrizen. Dann ist die j -te Spalte

$$\begin{aligned} \begin{pmatrix} c_{1j} \\ \vdots \\ c_{lj} \end{pmatrix} &= g \circ f(e_j) = g(f(e_j)) = g \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix} = g \left(\sum_{k=1}^m b_{kj} e'_k \right) \\ &= \sum_{k=1}^m b_{kj} g(e'_k) = \sum_{k=1}^m b_{kj} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{lk} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m a_{1k} b_{kj} \\ \vdots \\ \sum_{k=1}^m a_{lk} b_{kj} \end{pmatrix}, \end{aligned}$$

also

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

Definition 5.1.7. 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 $M(g \circ f) = M(g) \cdot M(f)$ gilt. Die Komposition linearer Abbildung wird also in die Multiplikation der darstellenden Matrizen überführt.

Beispiel 5.1.8. Wenn wir einen Spaltenvektor v als $m \times 1$ -Matrix auffassen, dann ist Matrixmultiplikation $M \cdot v$ genau gleich der Multiplikation aus Definition 5.1.4.

Wenn wir einen Zeilenvektor $w = (w_1, \dots, w_n)$ als $(1 \times n)$ -Matrix und einen Spaltenvektor $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ multiplizieren, erhalten wir $w \cdot v = (\sum w_i v_i)$, das *Skalarprodukt* von w^T und v , betrachtet als 1×1 -Matrix.

Beispiel 5.1.9. Wir berechnen $M(R_{\theta_1} \circ R_{\theta_2})$:

$$\begin{aligned} M(R_{\theta_1} \circ R_{\theta_2}) &= M(R_{\theta_1}) \cdot M(R_{\theta_2}) = \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix} \begin{pmatrix} \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 & -\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2 \\ \sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2 & -\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta_1 + \theta_2) & -\sin(\theta_1 + \theta_2) \\ \sin(\theta_1 + \theta_2) & \cos(\theta_1 + \theta_2) \end{pmatrix} = M(R_{\theta_1 + \theta_2}). \end{aligned}$$

Hier haben wir die Additionsformeln für trigonometrische Funktionen benutzt. Die Drehwinkel zweier Drehungen um den Ursprung addieren sich also.

Lemma 5.1.10. *Matrixmultiplikation ist assoziativ.*

Beweis:. Wir rechnen $((AB)C)_{i\ell} = \sum (\sum a_{ij} b_{jk}) c_{k\ell} = \sum_{j,k} a_{ij} b_{jk} c_{k\ell} = (A(BC))_{i\ell}$.

Alternativ folgt das Lemma aus dem Assoziativitätsgesetz für die Verkettung von Abbildungen: Jede Matrix hat die Form $M(f)$ für eine lineare Abbildung f und nach Definition ist $M(g \circ f) = M(g) \cdot M(f)$. Dann ist $(M(h) \cdot M(g)) \cdot M(f) = M((h \circ g) \circ f) = M(h \circ (g \circ f)) = M(h) \cdot (M(g) \cdot M(f))$. \square

Definition 5.1.11. Wir setzen

$$E_n = M(\text{id}_{K^n}) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

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 5.1.12. *Für eine $n \times m$ -Matrix M gilt $E_n \cdot M = M \cdot E_m = M$.*

Beweis:. Wir rechnen $(E_n \cdot M)_{ik} = \sum_j \delta_{ij} M_{jk} = M_{ik}$ denn Multiplikation mit δ_{ij} lässt nur den Summanden M_{ik} übrig. \square

Bemerkung 5.1.13. Im allgemeinen ist die Verkettung linearer Abbildungen und somit die Matrizenmultiplikation nicht kommutativ. Zum Beispiel finden wir für

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & B &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ AB &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & BA &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

5.2 Mehr über Matrizen

Wir wollen noch etwas mehr mit Matrizen rechnen. Nach Satz 3.3.9 ist $\text{Hom}_K(K^n, K^m)$ 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 5.2.1. 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:

$$\begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 1 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 0 & 1 \end{pmatrix}$$

Die Transponierte passt mit unserer Notation v^T zusammen, denn ein transponierter Zeilenvektor ist ein Spaltenvektor.

Soweit nicht anders vermerkt meinen wir mit einem Vektor von nun an einen *Spaltenvektor*,

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = (v_1, \dots, v_m)^T.$$

Lemma 5.2.2. $M(m \times n, K)$ ist ein Vektorraum der Dimension $m \cdot n$.

Beweis: Die Summe und Skalarmultiplikation erfüllen offensichtlich die Vektorraumaxiome. Die *Basismatrizen* E_{ij} , die genau eine 1 in der i -ten Spalte und j -ten Zeile haben und sonst 0 sind bilden eine Basis. \square

Lemma 5.2.3. *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. $A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$,
2. $A \cdot (B + B') = AB + A \cdot B'$ und $(A + A') \cdot B = AB + A'B$ (*Distributivgesetze*),
3. $(A \cdot B)^T = B^T \cdot A^T$.

Beweis: 1. und 2. zeigt man durch einfaches Hinschreiben.

3. 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

$$\begin{aligned} B^T &= (b'_{kj}) \quad \text{mit } b'_{kj} = b_{jk} \\ A^T &= (a'_{ji}) \quad \text{mit } a'_{ji} = a_{ij} \end{aligned}$$

Hieraus folgt

$$B^T \cdot A^T = (d_{ki}) \quad \text{mit}$$

$$d_{ki} = \sum_j b'_{kj} \cdot a'_{ji} = \sum_j b_{jk} \cdot a_{ij} = c_{ik} = c'_{ki}.$$

□

Unter der Entsprechung

$$\text{Hom}(K^n, K^n) \rightarrow M(n \times n, K)$$

entsprechen den Isomorphismen die folgenden Matrizen:

Definition 5.2.4. Eine Matrix $A \in M(n \times n, K)$ heißt *invertierbar*, wenn es eine Matrix $A^{-1} \in M(n \times n, K)$ gibt mit

$$A \cdot A^{-1} = A^{-1} \cdot A = E_n.$$

Beispiel 5.2.5. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $ad - bc \neq 0$. Auf dem Übungsblatt haben Sie nachgerechnet, dass A invertierbar ist mit $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Zum Beispiel ist

$$\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{-1}{6} \\ 0 & \frac{1}{3} \end{pmatrix}.$$

Korollar 5.2.6. 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 . Insbesondere ist A^{-1} eindeutig bestimmt, wenn es existiert.

$GL(n, K)$ 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 $A^{-1}, B^{-1} \in M(n \times n, K)$

$$AA^{-1} = A^{-1}A = E_n \quad \text{und} \quad BB^{-1} = B^{-1}B = E_n,$$

so ist wegen der Assoziativität der Matrizenmultiplikation Lemma 5.1.10

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = E_n \quad \text{und} \quad AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = E_n.$$

Damit haben wir eine Verknüpfung $GL(n, K) \times GL(n, K) \rightarrow GL(n, K)$.

- Assoziativität der Matrixmultiplikation ist bekannt, das neutrale Element ist die Einheitsmatrix E_n . Wir haben für jedes $A \in GL(n, K)$ die Existenz einer Matrix A^{-1} angenommen, und da A ein Inverses für A^{-1} ist, haben wir auch Inverse. Damit liegt eine Gruppe vor. □

Beachten Sie, dass $GL(n, K)$ keine Gruppe unter Addition ist, denn die Nullmatrix ist natürlich nicht invertierbar.

Satz 5.2.7. Gegeben seien K -Vektorräume

$$\begin{aligned} V & \text{ mit geordneter Basis } \mathcal{A} = (v_1, \dots, v_n) \\ W & \text{ mit geordneter Basis } \mathcal{B} = (w_1, \dots, w_m) . \end{aligned}$$

Dann gibt es zu jeder linearen Abbildung

$$f : V \rightarrow W$$

genau eine Matrix $A = (a_{ij}) \in M(m \times n, K)$, so dass

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

gilt. Die so erhaltene Abbildung

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{A}} : \text{Hom}(V, W) & \rightarrow M(m \times n, K) \\ f & \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(f) \end{aligned}$$

ist ein Isomorphismus von K -Vektorräumen. Insbesondere gilt

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{A}}(f + g) & = M_{\mathcal{B}}^{\mathcal{A}}(f) + M_{\mathcal{B}}^{\mathcal{A}}(g) \\ M_{\mathcal{B}}^{\mathcal{A}}(\lambda f) & = \lambda M_{\mathcal{B}}^{\mathcal{A}}(f) . \end{aligned}$$

Nach Wahl von geordneten Basen von V und von W kann man also lineare Abbildungen durch Matrizen beschreiben. Man sagt, die Matrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$ stelle die lineare Abbildung f bezüglich der geordneten Basen \mathcal{A} , \mathcal{B} von V und von W dar.

Wenn $V = K^n$, $W = K^m$ und \mathcal{A}, \mathcal{B} die Standardbasen sind, dann schreiben wir wie schon in Definition 5.1.2 .2 $M(f)$ für $M_{\mathcal{B}}^{\mathcal{A}}(f)$.

⟨⟨Die Formel (*) sieht der Formel für die Multiplikation einer Matrix mit einem Vektor sehr ähnlich, aber es sind verschiedene Formeln. Insbesondere haben wir einmal $\sum a_{ij} w_i$, wobei die $w_i \in W$ Basisvektoren sind und einmal $\sum a_{ij} x_j$ wobei die $x_j \in K$ Koordinaten für einen Vektor sind. Die Indizes sind gerade vertauscht: Bei (*) summieren wir über die Zeilen der Matrix, bei der Multiplikationsformel über die Spalten! ⟩⟩

Beweis: Da $\mathcal{B} = (w_1, \dots, w_m)$ eine geordnete Basis von W ist, sind für jedes v_j die Koeffizienten a_{ij} in (*) und somit die Spalten der Matrix eindeutig bestimmt. Somit ist die Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ wohldefiniert.

Da jede lineare Abbildung durch das Bild der Basisvektoren bestimmt ist, siehe Satz 5.1.1, ist $M_{\mathcal{B}}^{\mathcal{A}}$ injektiv.

Gehört zur Abbildung g die Matrix $B = (b_{ij})$, so rechnen wir:

$$\begin{aligned} (f + g)(v_j) & = f(v_j) + g(v_j) \\ & = \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m b_{ij} w_i \\ & = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i \end{aligned}$$

und für $\lambda \in K$

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

Also ist die Abbildung $M_{\mathcal{B}}^A$ eine K -lineare Abbildung.

Schließlich ist $M_{\mathcal{B}}^A$ surjektiv, denn für eine Matrix $A = (a_{ij})$ können wir die lineare Abbildung f definieren, die $\sum x_j v_j$ nach $\sum_{i,j} a_{ij} x_j w_i$ schickt. Es gilt $M_{\mathcal{B}}^A(f) = A$. \square

Korollar 5.2.8. Seien V, W zwei K -Vektorräume. Es gilt

$$\dim_K \text{Hom}(V, W) = \dim_K V \cdot \dim_K W .$$

Seien \mathcal{A} und \mathcal{B} Basen wie oben. Dann ist eine explizite Basis gegeben durch die lineare Abbildungen

$$f_{ij} : \quad V \rightarrow W$$

mit $f_{ij}(v_k) := \begin{cases} w_i & \text{für } k = j \\ 0 & \text{sonst.} \end{cases}$

für jedes $i = 1 \dots n$ und $j = 1 \dots m$.

Beweis: Es ist $M_{\mathcal{B}}^A(f_{ij}) = E_{ij}$, wobei die Familie (E_{ij}) wie in Lemma 5.2.2 eine Basis des K -Vektorraums $M(m \times n, K)$ bilden. Da $M_{\mathcal{B}}^A$ ein Isomorphismus ist, bildet auch (f_{ij}) eine Basis von $\text{Hom}(V, W)$. \square

Die naheliegende Frage, wie die Matrix $M_{\mathcal{B}}^A(F)$ sich ändert, wenn man die geordneten Basen \mathcal{A}, \mathcal{B} ändert, werden wir am Ende dieses Kapitels beantworten.

Bemerkungen 5.2.9. Ist $V = W$, d.h. liegt ein Endomorphismus vor, so ist es zweckmäßig, mit nur einer 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

$$M_{\mathcal{B}} : \text{End}(V) \rightarrow M(n \times n, K)$$

ist dann definiert durch die Gleichungen

$$f(v_j) = \sum_{i=1}^n a_{ij} v_i .$$

Die Einheitsmatrix $E_n = (\delta_{ij})$ beschreibt in jeder Basis \mathcal{B} von V die identische Abbildung, $M_{\mathcal{B}}(\text{id}_V) = E_n$.

Die Frage, wie durch Wahl einer geeigneten Basis die darstellende Matrix eines Endomorphismus auf eine Standardform gebracht werden kann, werden wir erst im nächsten Semester beantworten können.

5.3 Der Gaußsche Algorithmus und Elementarmatrizen

Die folgenden Überlegungen hatten im speziellen Fall des Körpers \mathbb{R} der reellen Zahlen schon einmal gesehen. Wir wiederholen Sie nun für beliebige Körper und mit einer neuen Perspektive. Wir werden zahlreiche nützliche Konsequenzen ziehen.

Definition 5.3.1. 1. Sei K ein Körper. Ein *lineares Gleichungssystem* ist ein System von Gleichungen der Form

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

mit $a_{ij} \in K$ und $b_i \in K$. Gesucht sind $x_1, \dots, x_n \in K$.

2. Gilt $b_1 = \dots = b_m = 0$, 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

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n, K)$$

die *Koeffizientenmatrix* des linearen Gleichungssystems. Mit $b := (b_1 \dots, b_m) \in K^m$ nennen wir die Matrix

$$(A, | b) := \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right) \in M(m \times (n+1), K)$$

die *erweiterte Koeffizientenmatrix* des inhomogenen linearen Gleichungssystems.

5. Die *Lösungsmenge* des linearen Gleichungssystems ist nun

$$\text{Lsg}(A, b) := \{x \in K^n \mid Ax = b\}$$

Für ein gegebenes lineares Gleichungssystem $Ax = b$ führen wir wie in Definition 5.1.2.2 die lineare Abbildung

$$\alpha : K^n \rightarrow K^m \quad \text{mit darstellender Matrix } M(\alpha) = A$$

ein. (Hier betrachten wir die Standardbasis für K^n und K^m .)

Satz 5.3.2. Wenn $x_0 \in \text{Lsg}(A, b)$ dann gilt $\text{Lsg}(A, b) = x_0 + \ker(\alpha) = \{x_0 + v \mid v \in \ker(\alpha)\}$.

Eine Teilmenge von K^n der Form $x_0 + V$ für einen Unterraum V heißt auch *affiner Unterraum* und wir nennen die Dimension von V auch die Dimension von $x_0 + V$.

Es gilt $x + V = y + W$ genau dann, wenn gilt $V = W$ und $x - y \in V$, siehe Übungsblatt.

Beweis:. Sei $x_1 \in \text{Lsg}(A, b)$. Dann gilt $\alpha(x_1) = \alpha(x_0)$ und $x_1 - x_0 \in \ker(\alpha)$, in anderen Worten $x_1 \in x_0 + \ker(\alpha)$.

Sei umgekehrt $x_1 \in x_0 + \ker(\alpha)$. Dann ist $\alpha(x_1) = \alpha(x_0 + v) = \alpha(x_0) = b$ mit $v \in \ker(\alpha)$. \square

Wir brauchen also nur eine Lösung des inhomogenen Gleichungssystems und erhalten dann alle anderen Lösungen durch Addition von Lösungen des zugehörigen homogenen Gleichungssystems.

Die Dimension der Lösungsmenge ist

$$\dim \text{Lsg}(A, b) = \dim(x_0 + \ker \alpha) = \dim_K \ker \alpha = n - \text{rg } \alpha.$$

wobei im vorletzten Schritt die Dimensionsformel 4.3.4 einging. Damit wir diese Formel benutzen können, brauchen wir eine Methode, um den Rang von α zu bestimmen. Wieder geht das am Besten, wenn eine Matrix in Zeilenstufenform ist.

Definition 5.3.3. 1. Eine Matrix $A \in M(m \times n, K)$ ist in *Zeilenstufenform*, falls für alle $i = 2, \dots, m$ gilt: ist $k = 1$ oder 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 $k = 1, \dots, n$.

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

Lemma 5.3.4. Sei A eine Matrix in Zeilenstufenform. Dann ist die Lösungsmenge eines linearen Gleichungssystems mit Koeffizientenmatrix A genau dann leer, wenn es einen Index $i \in \{1, \dots, m\}$ gibt, so dass $a_{ij} = 0$ für alle j , aber $b_i \neq 0$ gilt.

Beweis: “ \Leftarrow ” folgt wie in Bemerkung 1.3.5: in der i -ten Zeilen i erhalten wir die Gleichung $0 = \sum a_{ij}x_j = b_i$, die zum Widerspruch führt.

“ \Rightarrow ” folgt durch sukzessives Lösung des Gleichungssystems von unten nach oben: Gegeben $\sum a_{ij}x_j = b_i$ mit k minimal, so dass $a_{ik} \neq 0$ setzen wir $x_k = \frac{1}{a_{ik}}(b_i - \sum_{j>k} a_{ij}x_j)$, wobei einige x_j auf der rechten Seite von den weiter unten stehenden Zeilen bestimmt sind und die übrigen frei gewählt werden können. \square

Satz 5.3.5. Sei $A \in M(m \times n, K)$ eine Matrix in Zeilenstufenform und sei r die Anzahl der Zeilen die nicht 0 sind. Sei f die lineare Abbildung, die von A dargestellt wird. Dann ist $\text{rg}(f) = r$ und $\dim \ker(f) = n - r$.

Wir veranschaulichen die Situation in einem Beispiel:

$$A = \begin{pmatrix} 2 & 1 & -4 & -1 & 0 & -2 \\ 0 & 0 & 3 & 0 & -7 & 7 \\ 0 & 0 & 0 & -1 & 11 & -11 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \text{Mat}(5 \times 6, \mathbb{C})$$

Dann wird $\text{Im}(\alpha)$ von 6 Vektoren aufgespannt. Aber die letzten 2 Einträge sind jeweils 0, das heißt das Bild ist im Unterraum $\{x \in \mathbb{R}^5 \mid x_4 = x_5 = 0\} \cong \mathbb{R}^3$ enthalten. Umgekehrt finden wir die Vektoren $v_1 = 2e_1$, $v_2 = -4e_1 + 3e_2$ und $v_3 = -e_1 - e_3$ im Bild, aus denen wir leicht also Linearkombinationen $e_1 = \frac{1}{2}v_1$, $e_2 = \frac{1}{3}(v_2 + 2v_1)$ und $e_3 = -v_3 + \frac{1}{2}v_1$ erhalten. Das heißt, das Bild wird von den ersten 3 Standardbasisvektoren aufgespannt und hat Dimension 3.

Beweis: Wir müssen die Dimension des Bildes von f bestimmen. Das Bild wird per Definition von allen Spalten der Matrix A aufgespannt. Da die letzten $m - r$ Zeilen gleich 0 sind werden die letzten $m - r$ Koordinaten aller Spaltenvektoren 0 sein. Es folgt, dass $\text{Im}(\alpha)$ ein Unterraum von $\text{span}(e_1, \dots, e_r)$ ist.

Wir wollen nun zeigen, dass alle e_i mit $i \leq r$ in $\text{Im}(\alpha)$ liegen. Es reicht zu zeigen, dass wir diese e_i als Linearkombination der Spalten von A erhalten. Wir verwenden Induktion nach i .

und formen um:

$$\begin{pmatrix} \sqrt{2} & -2 & 0 \\ 0 & i+1 & 2 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

sowie

$$\begin{pmatrix} \sqrt{2} & -2 & 0 \\ 0 & i+1 & 2 \\ 0 & -1+i & i-1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

und

$$\begin{pmatrix} \sqrt{2} & -2 & 0 \\ 0 & 1 & 1-i \\ 0 & -1+i & i-1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{i+1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

und

$$\begin{pmatrix} \sqrt{2} & -2 & 0 \\ 0 & 1 & 1-i \\ 0 & 0 & -1-i \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1-i & 1 \end{pmatrix} \cdots \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

mit $i-1+(1-i)(1-i) = -1-i$. Schließlich erhalten wir mit zwei Zeilenmultiplikationen:

$$\begin{pmatrix} 1 & -\sqrt{2} & 0 \\ 0 & 1 & 1-i \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{1+i} \end{pmatrix} \cdots \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

Natürlich wenden wir die Zeilenumformungen in der Regel an, ohne uns Gedanken zu machen, was für Elementarmatrizen da gerade eine Rolle spielen. Aber es ist theoretisch wichtig, dass wir die intuitiven Umformungen formal als Multiplikation mit Elementarmatrizen verstehen können!

Wir haben zum Beispiel mit unserer Rechnung einen engen Zusammenhang zwischen dem Gauß-Algorithmus und der Struktur der Gruppe $GL(n, K)$ gefunden:

Korollar 5.3.10. *Jede Matrix $A \in GL(n, K)$ ist ein endliches Produkt von Elementarmatrizen.*

Beweis: Durch elementare Zeilenumformungen bringen wir A zunächst auf spezielle Zeilenstufenform, d.h. dass die Matrix $T_s \cdots T_1 \cdot A$ mit T_i Elementarmatrizen spezielle Zeilenstufenform hat. Diese Matrix ist invertibel und kann daher auf der Diagonale nur Einsen haben.

Die letzte Zeile hat daher nur eine 1 an letzter Stelle und ist sonst 0. Durch Abziehen von Vielfachen der letzten Zeile können wir die letzten Einträge aller anderen Zeilen auf 0 setzen. Nun hat die vorletzte Zeile eine 1 an vorletzter Stelle und ist sonst 0. Wir benutzen sie um alle anderen Zeilen auch an der vorletzten Zeile gleich 0 zu setzen. Durch weitere Zeilenumformungen erreichen wir die Einheitsmatrix.

Da jede Umformung der Multiplikation mit einer Elementarmatrix entspricht, gibt es Elementarmatrizen T_i , so dass $T_n \cdots T_1 \cdot A = E_n$ gilt. Lösen wir nach A auf, so haben wir A als Produkt von Elementarmatrizen geschrieben, $A = T_1^{-1} \cdots T_n^{-1}$. \square

Aus $A = T_1^{-1} \cdots T_n^{-1}$ folgt sofort, dass $T_n \cdot T_{n-1} \cdots T_1$ eine inverse Matrix zu A ist. Wir können den Gauss'schen Algorithmus benutzen, um eine Matrix zu invertieren. Dazu wenden wir die Zeilenumformungen gleichzeitig auf A und eine Einheitsmatrix an.

Betrachtung 5.3.11. Sei A eine Matrix und T_i eine Familie von Elementarmatrizen mit $T_s \cdot T_1 \cdot A = E_n$. Wir schreiben A und E_n nebeneinander und wenden die gleichen Zeilenumformungen an. Dann erhalten wir:

$$(A \mid E_n) \sim (T_s \cdot T_1 A \mid T_s \cdot T_1 E_n) = (E_n \mid A^{-1})$$

Dies ist eine effektive Methode um das Inverse einer invertierbaren Matrix zu bestimmen!

Der Algorithmus verrät uns auch, ob die Matrix überhaupt invertierbar ist: Wenn die Matrix nicht invertierbar ist dann erhalten wir eine Zeilenstufenform in der manche Zeilen 0 sind. Da die Multiplikation mit Elementarmatrizen den Rang nicht verändert, ist dann die Anzahl r der Zeilen, die nicht gleich 0 genau der Rang der zugehörigen linearen Abbildung, siehe Satz 5.3.5. Die Abbildung ist surjektiv (und damit bijektiv und invertierbar!) genau wenn alle Zeilen nicht 0 sind.

Beispiel 5.3.12. Wir wollen die Matrix A invertieren. Dazu betrachten wir die doppelte Koeffizientenmatrix $(A \mid E_n)$ und verwenden Zeilenumformungen:

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 4 & 0 & 1 & 0 \\ 1 & 4 & 9 & 0 & 0 & 1 \end{array} \right) &\stackrel{L3-L1, L2-L1}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 3 & 8 & -1 & 0 & 1 \end{array} \right) &\stackrel{L3-3L2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & -1 & 2 & -3 & 1 \end{array} \right) \\ &\stackrel{-1 \cdot L3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array} \right) &\stackrel{L2-3L3, L1-L3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 3 & -3 & 1 \\ 0 & 1 & 0 & 5 & -8 & 3 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array} \right) \\ &\stackrel{L1-L2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 5 & -2 \\ 0 & 1 & 0 & 5 & -8 & 3 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array} \right) \end{aligned}$$

Wir prüfen: $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 9 \end{pmatrix} \cdot \begin{pmatrix} -2 & 5 & -2 \\ 5 & -8 & 3 \\ -2 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

Bemerkung 5.3.13. Ein Nachtrag zu invertierbaren Matrizen: Solange A quadratisch ist, ist A invertierbar solange es ein Linksinverses *oder* Rechtsinverses gibt.

Wir wissen, dass eine lineare Abbildung $f : K^n \rightarrow K^n$ invertierbar ist, wenn sie surjektiv oder injektiv ist. Angenommen f hat ein links-inverses, $f^{-1} \circ f = \text{id}_{K^n}$. Dann muss f injektiv sein, und ist invertierbar nach Korollar 4.3.5. Habe umgekehrt f ein Rechtsinverses, $f \circ f^{-1} = \text{id}_{K^n}$. Dann ist f surjektiv und ist invertierbar nach Korollar 4.3.5. Es folgt, dass f invertierbar ist solange es ein Rechts- oder Linksinverses gibt. Genauso ist eine quadratische Matrix $A \in M(n \times n, K)$ invertierbar wenn es A^{-1} mit $A^{-1}A = E_n$ gibt oder wenn es A^{-1} mit $AA^{-1} = E_n$ gibt.

5.4 Koordinatentransformationen

Wir erinnern an Lemma 4.1.8: ist V ein n -dimensionaler K -Vektorraum, so liefert jede geordnete Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V einen Isomorphismus von K -Vektorräumen

$$\begin{aligned} \Phi_{\mathcal{B}} : K^n &\rightarrow V \\ e_i &\mapsto v_i. \end{aligned}$$

vom Standardvektorraum K^n auf V , den wir nun in Abhängigkeit von \mathcal{B} mit $\Phi_{\mathcal{B}}$ bezeichnen.

Lemma 5.4.1. Sei $f : V \rightarrow W$ eine lineare Abbildung und seien geordnete Basen $\mathcal{A} = (v_1, \dots, v_n)$ von V und $\mathcal{B} = (w_1, \dots, w_m)$ von W gegeben. Für die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$ mit

$$f(v_i) = \sum_{j=1}^m M_{\mathcal{B}}^{\mathcal{A}}(f)_{ji} w_j.$$

(vgl. Satz 5.2.7) gilt:

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = M(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}),$$

wobei M im Sinne von Definition 5.1.2 die Abbildung

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

bezüglich der Standardbasen von K^n und K^m darstellt.

Beweis: Wir finden unter Verwendung von Definition 5.1.2 im dritten und der Linearität von Φ im vierten Schritt

$$\begin{aligned} f(v_i) &= f(\Phi_{\mathcal{A}}(e_i)) = \Phi_{\mathcal{B}}(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}(e_i)) \\ &= \Phi_{\mathcal{B}}\left(\sum_{j=1}^m M(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}})_{ji} e_j\right) \\ &= \sum_{j=1}^m M(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}})_{ji} w_j. \end{aligned}$$

□

Beispiel 5.4.2. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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

$$f(b_1) = b_1 \quad f(b_2) = -b_2$$

wird der Endomorphismus f in der Basis \mathcal{B} durch die Matrix

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

dargestellt.

Wir vergleichen mit der Matrix der linearen Abbildung $\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{B}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Diese Verknüpfung schickt e_1 nach b_1 , dann b_1 nach b_1 und b_1 wieder nach e_1 . Sie schickt e_2 nach b_2 , dann nach $-b_2$ und dann nach $-e_2$. Wir erhalten also die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ bezüglich der Standardbasis, wie im Lemma.

Andererseits vertauscht Spiegelung in b_1 die beiden Standardbasisvektoren: $f(e_1) = e_2$ und $f(e_2) = e_1$. Also gilt in der Standardbasis $M(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Wir wollen als nächstes $M(f)$ und $M_{\mathcal{B}}(f)$ miteinander in Bezug setzen.

Satz 5.4.3. 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

$$f : V \rightarrow W \quad \text{und} \quad g : W \rightarrow Z$$

die Gleichung

$$M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) = M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) ,$$

wobei “ \cdot ” für die Matrizenmultiplikation steht.

Damit wir zwei darstellende Matrizen verknüpfen können, muss die Basis für den Wertebereich der rechten Matrix gleich der Basis für den Wertebereich der linken Matrix sein. Also: $M_{\mathcal{D}}^{\mathcal{C}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f)$ ist nur sinnvoll, wenn $\mathcal{C} = \mathcal{B}$!

Beweis:

$$\begin{aligned} M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) &\stackrel{5.4.1}{=} M(\Phi_{\mathcal{C}}^{-1} \circ g \circ f \circ \Phi_{\mathcal{A}}) \\ &= M\left(\left(\Phi_{\mathcal{C}}^{-1} \circ g \circ \Phi_{\mathcal{B}}\right) \circ \left(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}\right)\right) \\ &= M\left(\Phi_{\mathcal{C}}^{-1} \circ g \circ \Phi_{\mathcal{B}}\right) \cdot M\left(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}\right) \\ &\stackrel{5.4.1}{=} M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \end{aligned}$$

Hierbei haben wir im dritten Schritt verwendet, dass M die Komposition von Abbildungen in eine Matrizenmultiplikation überführt, vgl. Betrachtung 5.1.6. \square

Jetzt können wir beantworten, wie die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}$ sich bei Änderungen der Basen \mathcal{A} , \mathcal{B} ändert.

Wir überlegen uns zuerst, was mit der Identität passiert!

Definition 5.4.4. Seien \mathcal{A} und \mathcal{A}' zwei geordnete Basen von V . Dann heißt die quadratische Matrix

$$T_{\mathcal{A}'}^{\mathcal{A}} := M_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V) = M\left(\Phi_{\mathcal{A}'}^{-1} \circ \text{id}_V \circ \Phi_{\mathcal{A}}\right) = M\left(\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}}\right) \in GL(n, K)$$

Transformationsmatrix des Basiswechsels von \mathcal{A} nach \mathcal{A}' .

$$\begin{array}{ccc} K^n & & \\ & \searrow \Phi_{\mathcal{A}} & \\ & \sim & V \\ & \nearrow \Phi_{\mathcal{A}'} & \\ K^n & & \end{array}$$

$\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}}$

Da $\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}}$ eine Abbildung von K^n nach K^n ist können wir hier die Matrix $M(\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}})$ bezüglich der Standardbasis betrachten.

Bemerkungen 5.4.5. 1. 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

$$T_{\mathcal{A}'}^{\mathcal{A}} \cdot T_{\mathcal{A}}^{\mathcal{A}'} = M_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V) \cdot M_{\mathcal{A}}^{\mathcal{A}'}(\text{id}_V) = M_{\mathcal{A}'}^{\mathcal{A}'}(\text{id}_V \circ \text{id}_V) = M_{\mathcal{A}'}^{\mathcal{A}'}(\text{id}_V) = E_n .$$

Hier haben wir Satz 5.4.3 benutzt! Also sind Transformationsmatrizen invertierbar. Die Transformationsmatrizen $T_{\mathcal{A}'}^{\mathcal{A}}$ und $T_{\mathcal{A}}^{\mathcal{A}'}$ sind zueinander invers.

2. Gegeben ein Vektor $v \in V$ können wir ihn eindeutig schreiben als $v = \sum \lambda_i v_i$ für die Basis $\mathcal{A} = (v_1, \dots, v_n)$. Dann ist

$$(v)_{\mathcal{A}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

die *Koordinatendarstellung* von v bezüglich \mathcal{A} . Wenn wir $v \in V$ als lineare Abbildung $\tilde{v} : \lambda \mapsto \lambda v$ von K nach V betrachten dann ist $(v)_{\mathcal{A}}$ genau $M_{\mathcal{A}}^{\mathcal{E}}(\tilde{v})$, wobei wir \mathcal{E} für die Standardbasis $\{(1)\}$ von K^1 schreiben.

Für eine zweite Basis $\mathcal{A}' = (v'_1, \dots, v'_n)$ gilt $v = \sum \mu_i v'_i$. Die Transformationsmatrix erlaubt uns, die Koeffizienten μ_i aus den Koeffizienten λ_i zu bestimmen.

Es gilt

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = T_{\mathcal{A}'}^{\mathcal{A}} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

denn $M_{\mathcal{A}'}^{\mathcal{E}}(\tilde{v}) = M_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V) \cdot M_{\mathcal{A}}^{\mathcal{E}}(\tilde{v})$ nach Satz 5.4.3.

Konkreter können wir überprüfen: die j -te Spalte von $T_{\mathcal{A}'}^{\mathcal{A}} = M_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V)$ besteht nach Satz 5.2.7 genau aus den c_{ij} so dass $v_j = \sum_i c_{ij} v'_i$ ist. Also ist $\sum_j \lambda_j v_j = \sum_{i,j} \lambda_j c_{ij} v'_i$. Es folgt dass $\mu_i = \sum_j c_{ij} \lambda_j$ ist. Insbesondere ist die Transformationsmatrix

$$T_{\mathcal{A}'}^{\mathcal{A}} = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix}.$$

3. Sei insbesondere \mathcal{A}' die Standardbasis \mathcal{E} und \mathcal{A} eine beliebige Basis (v_1, \dots, v_n) und schreiben wir jeden Basisvektor v_i als $(v_{1i}, \dots, v_{ni})^T$. Dann ist

$$T_{\mathcal{E}}^{\mathcal{A}} = \begin{pmatrix} | & \dots & | \\ v_1 & \dots & v_n \\ | & \dots & | \end{pmatrix} = \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nn} \end{pmatrix}$$

4. Sei insbesondere $m = n = 1$. Eine Basis von K^1 wird von einem Element ungleich 0 gegeben, ein Element von $M(1 \times 1, K)$ ist einfach ein Element von K .

Zum Beispiel ist $\mathcal{A} = \{(4)\}$ eine Basis von K während $\mathcal{A}' = \{(1)\}$ die Standardbasis ist. Es gilt $(1) = \frac{1}{4}(4)$, also ist $T_{\mathcal{A}'}^{\mathcal{A}} = (\frac{1}{4})$ und mit $(4) = 4 \cdot (1)$ ist $T_{\mathcal{A}}^{\mathcal{A}'} = (4)$.

Wenn wir wie in 1. $x = x(1)$ in der Basis (4) schreiben wollen, ist der neue Koeffizient $\frac{x}{4}$, $x = x \cdot (1) = \frac{1}{4}x \cdot (4)$. In der neuen Basis \mathcal{B} ist der Basisvektor 4 mal so lang, also brauchen wir kleinere Koeffizienten.

Wir sehen hier gut, dass sich Koeffizienten und Basisvektoren gewissermaßen invers zueinander verhalten: Wenn ich den Basisvektor strecke dann schrumpft mein Koeffizient.

Wir betrachten als nächstes als Vorschau auf den nächsten Satz eine lineare Abbildung $f : x \mapsto \mu x$. Dann gilt $M_{\mathcal{A}}^{\mathcal{A}} = (\mu)$ und $M_{\mathcal{B}}^{\mathcal{B}} = (\mu)$, denn jeder Basisvektor wird ja von f mit μ multipliziert.

Es gilt aber $M_{\mathcal{B}}^{\mathcal{A}}(f) = \frac{\mu}{4}$, denn $\mu(1) = \frac{\mu}{4}(4)$. Es ist auch $M_{\mathcal{A}}^{\mathcal{B}}(f) = 4\mu$, denn $\mu(4) = 4\mu(1)$.

Satz 5.4.6 (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

$$f : V \rightarrow W$$

linear. Dann gilt

$$M_{\mathcal{B}'}^{\mathcal{A}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}. \quad (*)$$

2. Speziell für Endomorphismen $f : V \rightarrow V$ und zwei geordnete Basen \mathcal{B}' und \mathcal{B} von V ergibt sich

$$M_{\mathcal{B}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}(f) \cdot (T_{\mathcal{B}'}^{\mathcal{B}})^{-1}. \quad (**)$$

Wir können unsere Formel auch schreiben als

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

⟨⟨ Seien Sie gewarnt: Es ist sehr leicht, die Rolle der beiden Basen beim Basiswechsel zu verwechseln! ⟩⟩

Wir können die Zusammenhänge des Satzes auch als kommutatives Diagramm schreiben, d.h. alle Wege zwischen zwei fixen Ecken entlang der Kanten des folgenden Diagramms geben das gleiche Ergebnis. Es kommutiert also das folgende Diagramm:

$$\begin{array}{ccc}
 K^n & \xrightarrow{M_{\mathcal{B}}^{\mathcal{A}}(f)} & K^m \\
 \downarrow T_{\mathcal{A}'}^{\mathcal{A}} & \begin{array}{c} \searrow \Phi_{\mathcal{A}} \\ \nearrow \Phi_{\mathcal{A}'} \end{array} & \begin{array}{c} \nearrow \Phi_{\mathcal{B}} \\ \searrow \Phi_{\mathcal{B}'} \end{array} \\
 & V \xrightarrow{f} W & \\
 & \begin{array}{c} \nearrow \Phi_{\mathcal{A}'} \\ \searrow \Phi_{\mathcal{A}} \end{array} & \begin{array}{c} \searrow \Phi_{\mathcal{B}'} \\ \nearrow \Phi_{\mathcal{B}} \end{array} \\
 K^n & \xrightarrow{M_{\mathcal{B}'}^{\mathcal{A}'}(f)} & K^m \\
 & \downarrow T_{\mathcal{B}'}^{\mathcal{B}} & \\
 & &
 \end{array}$$

wobei hier Abbildungen zwischen Vektorräumen der Form K^n mit ihren darstellenden Matrizen identifiziert werden.

Beweis: 1. Wir rechnen:

$$\begin{aligned}
 M_{\mathcal{B}'}^{\mathcal{A}'}(f) &= M_{\mathcal{B}'}^{\mathcal{A}'}(\text{id}_W \circ f \circ \text{id}_V) \\
 &= M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}_W) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot M_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V) \quad [\text{wegen Satz 5.4.3}] \\
 &\stackrel{5.4.5.1}{=} T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot T_{\mathcal{A}'}^{\mathcal{A}}.
 \end{aligned}$$

2. ist als Spezialfall klar. □

Beispiel 5.4.7. Sei wie in Beispiel 5.4.2 $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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

$$f(b_1) = b_1 \quad f(b_2) = -b_2$$

wird der Endomorphismus f in der Basis \mathcal{B} durch die Matrix

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

dargestellt. Wir wollen f in der geordneten Standardbasis $\mathcal{E} = (e_1, e_2)$ des \mathbb{R}^2 ausdrücken. Aus

$$b_1 = e_1 + e_2 \quad b_2 = -e_1 + e_2$$

folgt

$$T_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} .$$

und

$$\left(T_{\mathcal{E}}^{\mathcal{B}}\right)^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} .$$

Es folgt

$$M_{\mathcal{E}}(f) = T_{\mathcal{E}}^{\mathcal{B}} M_{\mathcal{B}}(f) \left(T_{\mathcal{E}}^{\mathcal{B}}\right)^{-1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

Dies entspricht genau unserer direkten Berechnung aus Beispiel 5.4.2.

5.5 Äquivalenz und Ähnlichkeit von Matrizen

Die Form der Gleichung (*) beziehungsweise von Gleichung (**) aus der Transformationsformel gibt Anlass zu der folgenden Definition:

Definition 5.5.1. 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

$$Y = SXT^{-1} \quad \text{gilt .}$$

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

$$Y = TXT^{-1} \quad \text{gilt .}$$

Bemerkungen 5.5.2. 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

$$Z = S_2(S_1XT_1^{-1})T_2^{-1} = (S_2S_1)X(T_2T_1)^{-1}.$$

⟨⟨ Diese Benennung ist recht ungünstig. "Äquivalent" klingt wie eine stärkere Bedingung als "ähnlich". Möglicherweise sind Sie Ihrem Bruder ähnlich, aber sicher sind Sie nicht äquivalent zu Ihrem Bruder. Wir werden uns daran gewöhnen müssen, dass ähnliche Matrizen immer äquivalent und viele Matrizen äquivalent aber nicht ähnlich sind. ⟩⟩

Lemma 5.5.3. *Zwei Matrizen sind genau dann äquivalent, wenn sie dieselbe lineare Abbildung bezüglich verschiedener geordneter Basen beschreiben.*

Die Aussage des Lemmas bedeutet: Zwei Matrizen $X, Y \in M(m \times n, K)$ sind genau dann äquivalent, wenn es 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 $f : V \rightarrow W$ gibt, so dass gilt

$$X = M_{\mathcal{B}}^{\mathcal{A}}(f) \quad \text{und} \quad Y = M_{\mathcal{B}'}^{\mathcal{A}'}(f).$$

Beweis: " \Leftarrow " folgt sofort aus der Transformationsformel von Satz 5.4.6.

" \Rightarrow " Seien X, Y äquivalente $m \times n$ Matrizen:

$$Y = SXT^{-1}.$$

Wir betrachten K^m und K^n jeweils mit der Standardbasis \mathcal{E}_m und \mathcal{E}_n und finden f mit $X = M(f)$ nach Satz 5.2.7.

Wir wollen nun Basen \mathcal{A} und \mathcal{B} für K^m und K^n finden, so dass $Y = M_{\mathcal{A}}^{\mathcal{B}}(f)$ ist. Nach der Transformationsformel Satz 5.4.6 gilt das, wenn $S = T_{\mathcal{A}}^{\mathcal{E}_m}$ und $T^{-1} = T_{\mathcal{E}_n}^{\mathcal{B}}$, also $T = T_{\mathcal{B}}^{\mathcal{E}_n}$.

Nach Bemerkung 5.4.5.3 definieren wir die geordnete Basis \mathcal{A} durch die Spalten der Matrix S^{-1} , und \mathcal{B} durch die Spalten der Matrix T^{-1} . \square

Genauso kann man zeigen, dass zwei quadratische Matrizen genau dann ähnlich sind, wenn sie den gleichen Endomorphismus bezüglich verschiedener Basen beschreiben.

Satz 5.5.4. *Jede lineare Abbildung f zwischen endlich-dimensionalen Vektorräumen lässt sich von einer Matrix der Form*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

darstellen, wobei die Nullen jeweils für Nullmatrizen verschiedener Größe stehen. Hier ist $r = \text{rg}(f)$.

Jede Matrix X ist äquivalent zu einer Matrix dieser Form.

Ein *Blockmatrix* ist eine Matrix, deren Einträge nicht Elemente von K sondern wieder Matrizen, wir nennen die Einträge *Blöcke*, in unserem Beispiel gibt es einen Block $E_r \in M(r \times r, K)$ und die Größen der verschiedenen Blöcke sind wie folgt, wobei wir $0_{i,j}$ für die Nullmatrix in $M(i \times j, K)$ schreiben:

$$\begin{pmatrix} E_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}.$$

Wir können mit Blockmatrizen nur bedingt wie mit echten Matrizen rechnen, hier verwenden wir sie nur als nützliche abkürzende Notation.

Beweis: Wir betrachten die kanonische Faktorisierung und wählen geschickte Basen.

Seien V, W endlich-dimensionale K -Vektorräume und

$$f : V \rightarrow W$$

eine lineare Abbildung. Wie im Beweis von Satz 4.3.4 ergänze eine geordnete Basis (v_1, \dots, v_k) des Untervektorraums $\ker f$ von V zu einer geordneten Basis (v_1, \dots, v_n) von V .

Nach Satz 5.1.1.1 ist dann

$$(f(v_{k+1}), \dots, f(v_n))$$

ein geordnetes Erzeugendensystem von $\operatorname{Im} f$. Die Familie ist auch linear unabhängig. Nach Satz 5.1.1.1(b) reicht es zu zeigen, dass f , eingeschränkt auf den Unterraum $\operatorname{span}_K(v_{k+1}, \dots, v_n) \subset V$ injektiv ist. Aber das gilt, da der Schnitt dieses Unterraums mit dem Kern von f trivial ist.

Ergänzt man die linear unabhängige Familie

$$(f(v_{k+1}), \dots, f(v_n))$$

in W in irgendeiner Weise zu einer geordneten Basis

$$\mathcal{B} = (f(v_{k+1}), \dots, f(v_n), w_1, \dots, w_{m-r})$$

von W mit $m := \dim_K W$ und $r := \operatorname{rg} f$ und wählt bequemerweise als geordnete Basis von V die Familie

$$\mathcal{A} = (v_{k+1}, \dots, v_n, v_1, \dots, v_k),$$

so hat man für f die folgende Blockmatrix als darstellende Matrix:

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \left(\begin{array}{cc} E_r & 0 \\ 0 & 0 \end{array} \right) \left. \vphantom{\begin{array}{cc} E_r & 0 \\ 0 & 0 \end{array}} \right\} m-r$$

$$\underbrace{\hspace{10em}}_{n-r=k}$$

Der Rang der linearen Abbildung ist dann wegen Satz 5.3.5 gleich r .

Schließlich benutzen wir Lemma 5.5.3 um zu zeigen, dass jede Matrix äquivalent zu einer Matrix dieser Form ist. \square

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.

Bemerkung 5.5.5. Wir können die gleiche Betrachtung wie im Beweis des Satzes übrigens auch auf die kanonische Projektion

$$\pi : V \rightarrow V/\ker f$$

anwenden und einen anderen Beweis des Homomorphiesatz 3.4.4 geben. Es gilt $\ker \pi = \ker f$; daher folgt, dass

$$\{[v_{k+1}], \dots, [v_n]\}$$

eine Basis von $\operatorname{Im} \pi = V/\ker f$ ist.

Die Abbildung $\bar{f} : V/\ker f \rightarrow \operatorname{Im} f$ ist dann auf der Basis $\{[v_{k+1}], \dots, [v_n]\}$ von $V/\ker f$ durch

$$\bar{f}([v_i]) = f(v_i)$$

definiert und wie im Satz ist $f(v_i)$ linear unabhängig und erzeugt das Bild. Damit liefert \bar{f} nach Satz 5.1.1 als Bijektion von Basen einen Isomorphismus von $V/\ker f$ auf $\operatorname{Im} f$.

Bemerkung 5.5.6. Wir können eine gegebene Matrix $A \in M(m \times n, K)$ explizit in die Form $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ bringen. Dazu verwenden wir zuerst Zeilenumformungen um A in spezielle Zeilenstufenform zu bringen. Anschließend verwenden wir *Spaltenumformungen* um die erhaltene Matrix in die gewünschte Form zu bringen.

Genauso wie eine Zeilenumformung durch Multiplikation von links mit einer Elementarmatrix beschrieben werden kann, so kann eine Spaltenumformung durch Multiplikation von *rechts* mit einer Elementarmatrix beschrieben werden.

Wir haben dann also $E_r = T_s \cdots T_1 A T'_1 \cdots T'_r$ und die gewünschte Ähnlichkeit ist gezeigt.

Wenn wir mit Matrizen rechnen, würde wir gerne den Rang einer Abbildung aus der darstellenden Matrix definieren.

Definition 5.5.7. Sei $X \in M(m \times n, K)$. Die maximale Anzahl linear unabhängiger

Spalten von X heißt *Spaltenrang* $\text{rg}(X)$
 Zeilen von X heißt *Zeilenrang* $\widetilde{\text{rg}}(X)$

Als Beispiel betrachten wir mit $K = \mathbb{R}$ die linearen Abbildungen $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ mit

$$\begin{aligned} M(f) &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} & \text{rg}(M(f)) &= 1 & \widetilde{\text{rg}}(M(f)) &= 1 & \text{rg}(f) &= 1 \\ M(g) &= \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 5 \end{pmatrix} & \text{rg}(M(g)) &= 2 & \widetilde{\text{rg}}(M(g)) &= 2 & \text{rg}(g) &= 2 \end{aligned}$$

Für den Spaltenrang von $M(g)$ betrachten wir

$$10 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} -1 \\ 5 \end{pmatrix} = 0$$

Die anderen Rechnungen sind unmittelbar.

Wir haben nun also den Begriff Rang für Matrizen und lineare Abbildungen eingeführt und sollten sicherstellen, dass dieser Gebrauch kompatibel ist: Wenn M darstellende Matrix für f ist (bezüglich irgendeiner Basis!), soll gelten $\text{rg}(f) = \text{rg}(M) = \widetilde{\text{rg}}(M)$, und unsere einfachen Beispiele sprechen schon einmal dafür.

Satz 5.5.8. Wenn M darstellende Matrix für die lineare Abbildung $f : V \rightarrow W$ ist gilt $\text{rg}(f) = \text{rg}(M)$.

Beweis: Wir betrachten zuerst den Fall, dass M eine lineare Abbildung $f : K^n \rightarrow K^m$ bezüglich der Standardbasen darstellt. Nach Definition 3.3.7.3) und Satz 5.1.1 (a):

$$\begin{aligned} \text{rg}(f) &= \dim_K \text{Im } f = \dim_K f(K^n) \\ &= \dim_K \text{span}_K(f(e_1), \dots, f(e_n)). \end{aligned}$$

Rechts steht die lineare Hülle der Spalten der darstellenden Matrix $M(f)$. Dies ist wegen des Basisauswahlsatzes 4.2.7 gleich der maximalen Anzahl linear unabhängiger Spaltenvektoren der Matrix $M(f)$, also genau der Rang von $M(f)$.

Seien nun \mathcal{A} und \mathcal{B} beliebige Basen für V und W , sodass $M = M_{\mathcal{B}}^{\mathcal{A}}(f)$.

Die Basen definieren Abbildung $\Phi_{\mathcal{A}} : K^n \rightarrow V$ und $\Phi_{\mathcal{B}} : K^m \rightarrow W$ und es gilt $\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}} = M$ nach Lemma 5.4.1.

Die Matrix M stellt aber auch eine lineare Abbildung g von K^n nach K^m bezüglich der Standardbasen dar, und wir haben gezeigt $\text{rg}(M) = \text{rg}(g)$.

Aber da $\Phi_{\mathcal{B}}$ und $\Phi_{\mathcal{A}}$ Isomorphismen sind, ist der Rang von g gleich dem Rang von f und es gilt $\text{rg}(M) = \text{rg}(g) = \text{rg}(f)$. \square

Korollar 5.5.9. *Zwei äquivalente Matrizen haben den gleichen Spaltenrang.*

Beweis: Seien X, Y äquivalente Matrizen. Nach Lemma 5.5.3 gibt es eine lineare Abbildung f , so dass X und Y darstellende Matrizen für f sind (für unterschiedliche Basen). Es gilt $\text{rg}(X) = \text{rg}(f) = \text{rg}(Y)$. \square

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 im 2. Semester vollständig beschreiben können.

Lemma 5.5.10. *Sei $X \in M(m \times n, K)$ und $S \in GL(m, K)$. Dann gilt*

1. $(S^{-1})^T = (S^T)^{-1}$
2. $\text{rg}(X) = \widetilde{\text{rg}}(X^T)$ und $\widetilde{\text{rg}}(X) = \text{rg}(X^T)$.

Beweis: 1.: Aus $(S^{-1})^T \cdot S^T = (S \cdot S^{-1})^T = E_m^T = E_m$ folgt die Behauptung wegen der Eindeutigkeit der Inversen.

2. ist offensichtlich, da die Transposition Zeilen und Spalten vertauscht. \square

Lemma 5.5.11. *Äquivalente Matrizen haben auch gleichen Zeilenrang. In Formeln: sind $X, Y \in M(m \times n, K)$ äquivalent, so ist*

$$\widetilde{\text{rg}}(X) = \widetilde{\text{rg}}(Y).$$

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 5.5.10.1. Somit sind auch die transponierten Matrizen X^T und Y^T äquivalent. Wegen Lemma 5.5.10.2 und Korollar 5.5.9 erhält man die Gleichungen

$$\widetilde{\text{rg}}(X) \stackrel{5.5.10.2}{=} \text{rg}(X^T) \stackrel{5.5.9}{=} \text{rg}(Y^T) \stackrel{5.5.10.2}{=} \widetilde{\text{rg}}(Y).$$

\square

Satz 5.5.12. *Zeilenrang und Spaltenrang einer Matrix sind gleich: für $X \in M(m \times n, K)$ gilt*

$$\text{rg}(X) = \widetilde{\text{rg}}(X).$$

Beweis: Hat $X \in M(m \times n, K)$ Rang r , so ist X nach Bemerkung 5.5.2.4 äquivalent zu einer Matrix der Form $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Daher gilt

$$\begin{aligned} \text{rg}(X) &\stackrel{(5.5.9)}{=} \text{rg} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r \\ \widetilde{\text{rg}}(X) &\stackrel{(5.5.11)}{=} \widetilde{\text{rg}} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r \end{aligned}$$

Hieraus folgt $\text{rg}(X) = r = \widetilde{\text{rg}}(X)$. \square

Bemerkung 5.5.13. Um den Zeilen- oder Spaltenrang einer Matrix zu bestimmen verwenden wir elementare Zeilenumformungen, also wieder den gaußschen Algorithmus. Wir können auch Spaltenumformungen verwenden, wenn $n < m$ ist, dann bietet sich das an.

6 Intermezzo: Basen für beliebige Vektorräume

Um zu zeigen, dass jeder Vektorraum eine Basis hat, brauchen wir zuerst eine Definition.

Definition 6.0.1. 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 $\mathcal{K} \subset X$, so dass (\mathcal{K}, \leq) *total geordnet* ist, d.h. für alle $h, k \in \mathcal{K}$ gilt $h \leq k$ oder $k \leq h$.
- 3. Wir sagen eine Kette \mathcal{K} in X hat eine *obere Schranke*, wenn $m \in X$ mit $k \leq m$ für alle $k \in \mathcal{K}$ existiert.

Das folgende Lemma ist die entscheidende Zutat im Beweis.

Lemma 6.0.2 (Zornsches Lemma). *Jede nichtleere, partiell geordnete Menge X , in der jede Kette eine obere Schranke hat, 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.*

Beweis: Dieses folgt auf höchst nicht-triviale Weise aus dem Auswahlaxiom. Eine Beweisskizze können Sie auf dem Übungsblatt sehen und als herausfordernde Übungsaufgabe die Details selbst ausarbeiten. \square

Das Zornsche Lemma ist sogar äquivalent zum Auswahlaxiom, wir könnten in unserer Axiomatisierung also statt des Auswahlaxioms das zornsche Lemma annehmen. Wir zeigen nun:

Satz 6.0.3. *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 $X(M, E) := \{A \mid \text{linear unabhängig, } M \subset A \subset E\}$. Sie enthält M selbst und ist daher nicht leer. Sie ist durch Inklusion partiell geordnet.

- Wir zeigen, dass in $X(M, E)$ jede Kette eine obere Schranke hat, indem wir nachweisen, dass für jede Kette $\mathcal{K} \subset X(M, E)$ die Vereinigung $\cup \mathcal{K} := \cup_{k \in \mathcal{K}} k$ in $X(M, E)$ liegt. Es ist dann klar, dass für jedes $A \in \mathcal{K}$ gilt $A \subset \cup \mathcal{K}$, also $\cup \mathcal{K}$ eine obere Schranke ist.

Aus $M \subset A \subset E$ für alle $A \in \mathcal{K}$ folgt $M \subset \cup \mathcal{K} \subset E$. Zu zeigen ist noch, dass die Vereinigung $\cup \mathcal{K}$ linear unabhängig ist. Seien dazu $\lambda_1, \dots, \lambda_n \in K$ und $v_1, \dots, v_n \in \cup \mathcal{K}$ mit $\sum_{j=1}^n \lambda_j v_j = 0$. Dann existieren $A_1, \dots, A_n \in \mathcal{K}$ mit $v_j \in A_j$. Durch Umm nummerieren können wir erreichen, dass $A_1 \subset A_2 \subset \dots \subset A_n$ gilt. Daraus folgt $v_1, \dots, v_n \in A_n$. Aus der linearen Unabhängigkeit von A_n ergibt sich $\lambda_1 = \dots = \lambda_n = 0$. Also ist $\cup \mathcal{K}$ linear unabhängig und eine obere Schranke der Kette \mathcal{K} in $X(M, E)$. Somit ist $X(M, E)$ induktiv geordnet.

- Da $X(M, E)$ nicht leer und partiell geordnet ist, und jede Kette eine obere Schranke hat, 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 $\lambda, \lambda_1, \dots, \lambda_n \in K$, nicht alle Null und $b_1, \dots, b_n \in B$, 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 $e \in \text{span}_K(B)$. Da dies für jedes $e \in E \setminus B$ gilt, folgt $E = B \cup (E \setminus B) \subset \text{span}_K(B)$ und somit $V = \text{span}(E) \subset \text{span}(B)$. Also ist B eine Basis mit allen geforderten Eigenschaften. \square

Wir haben insbesondere auch die folgenden Aussagen gezeigt, die wir für endlich-dimensionale Vektorräume schon bewiesen hatten:

Korollar 6.0.4. 1. *Jeder K -Vektorraum hat eine Basis. Denn wähle einfach $E = V$ und $M = \emptyset$.*

2. *Basisauswahlsatz: Aus jedem Erzeugendensystem E eines Vektorraums kann man eine Basis auswählen. Hier wählt man einfach $M = \emptyset$.*

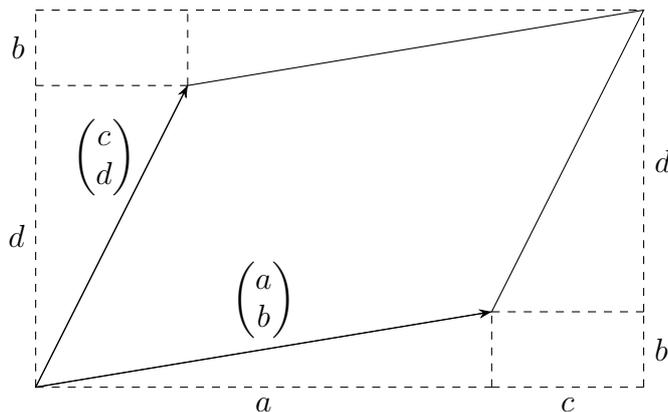
3. *Basisergänzungssatz: Jede linear unabhängige Teilmenge $M \subset V$ lässt sich zu einer Basis von V ergänzen. Hier wählt man einfach $V = E$.*

7 Determinanten

7.1 Vorüberlegungen

Wir betrachten Flächen und Volumen in reellen Vektorräumen.

Zwei Vektoren in \mathbb{R}^2 spannen ein Parallelogramm auf, und wir können uns überlegen, was die Fläche dieses Parallelograms ist.



Wir berechnen die Fläche als

$$\text{Area}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) = (a+c)(b+d) - \frac{1}{2}ab - \frac{1}{2}ab - \frac{1}{2}cd - \frac{1}{2}cd - bc - bc = ad - bc$$

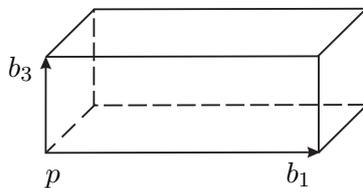
indem wir vom großen Rechteck vier Dreiecke und zwei kleiner Rechtecke abziehen.

Allgemeiner fragen wir uns nach dem Volumen des Parallelotops, das n Vektoren in \mathbb{R}^n aufspannen.

Definition 7.1.1. Sei (b_1, \dots, b_n) eine Basis von \mathbb{R}^n . Dann heißt die Teilmenge

$$P = \{\alpha_1 b_1 + \dots + \alpha_n b_n \mid 0 \leq \alpha_i \leq 1\} \subset \mathbb{R}^n$$

das von b_1, \dots, b_n aufgespannte *Parallelotop* oder *Spat*.



Wir führen zwei Konstruktionen ein, die uns helfen werden, das Volumen des Spats zu berechnen, aber auch unabhängig von hohem Interesse sind.

Definition 7.1.2. Seien $v = (v_1, \dots, v_n)^T, w = (w_1, \dots, w_n)^T \in \mathbb{R}^n$. Dann ist das *Skalarprodukt* von v und w definiert als $v \cdot w = \sum_i v_i w_i$.

Wir nennen $\|v\| = \sqrt{v \cdot v}$ die (*euklidische*) *Norm* des Vektors v , alltagssprachlich die *Länge*. Zwei Vektoren v, w mit $v \cdot w = 0$ heißen *orthogonal*.

Es gilt auch $v \cdot w = v^T w$ und das Skalarprodukt (nicht aber die Norm) kann über jedem Körper definiert werden.

Definition 7.1.3. Seien nun $v, w \in \mathbb{R}^3$. Dann ist das *Vektorprodukt* definiert als

$$v \times w = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Lemma 7.1.4. 1. Das Skalarprodukt ist symmetrisch $v \cdot w = w \cdot v$ und das Vektorprodukt ist antisymmetrisch: $v \times w = -w \times v$. Insbesondere ist $v \times v = 0$.

2. Skalarprodukt und Vektorprodukt sind bilinear, d.h. $(\lambda v + \mu w) \cdot u = \lambda v \cdot u + \mu w \cdot u$ und $(\lambda v + \mu w) \times u = \lambda v \times u + \mu w \times u$

3. Es gilt $(u \times v) \cdot w = u \cdot (v \times w)$.

Beweis: 1. Das folgt direkt aus der Definition. Für den letzten Satz betrachten wir $v \times v = -v \times v$ und erhalten $v \times v = 0$.

2. Direkt aus der Definition.

3. Dies ist eine direkte Rechnung. Wir können z.B. Koordinaten für unsere drei Vektoren wählen und beide Rechnungen ausführen.

Wir können alternativ betrachten, was mit Basisvektoren passiert, da beide Seiten der Gleichung linear in allen Einträgen sind.

Wir rechnen $e_1 \times e_2 = e_3$, $e_1 \times e_3 = -e_2$ und $e_2 \times e_3 = e_1$. Durch vertauschen der beiden Faktoren links ändert sich das Vorzeichen.

Wir schreiben jetzt ϵ_{ijk} für $(e_i \times e_j) \cdot e_k$. Es folgt aus unserer Rechnung dass $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ und $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$. Also gilt die Gleichung für alle Fälle, wenn die drei Indizes unterschiedlich sind.

Wenn zwei Indizes gleich sind, dann ist $\epsilon_{ijk} = 0$. Für die ersten beiden Indizes ist das klar wegen $e_i \times e_i = 0$. Aus der Berechnung von $e_i \times e_j$ folgt sofort, dass $\epsilon_{ijk} = 0$ auch gilt, wenn i oder j gleich k ist.

Damit gilt immer $\epsilon_{ijk} = \epsilon_{kij}$, und das zeigt 3. □

Bemerkung 7.1.5. Aus der Schule kennen Sie vielleicht geometrischere Definitionen: $a \cdot b = \|a\| \cdot \|b\| \cos(\angle(a, b))$. Aber was genau ist der Winkel zwischen zwei Vektoren, insbesondere in höheren Dimensionen? Wir definieren $\angle(a, b) = \cos^{-1}(\frac{a \cdot b}{\|a\| \cdot \|b\|})$. Dafür müssen wir wissen, dass $a \cdot b \leq \|a\| \cdot \|b\|$, das ist die *Cauchy-Schwarzsche Ungleichung*, dazu (wahrscheinlich) später mehr.

Das Vektorprodukt $a \times b$ ist der Vektor orthogonal zu a und b mit Norm $\|a\| \cdot \|b\| \sin(\angle(a, b))$ und Orientierung entsprechend der "Rechte-Hand-Regel".

Es lässt sich nachrechnen, dass dies aus unserer Definition folgt, die Details können Sie auf dem Übungsblatt ausarbeiten.

Das Vektorprodukt hat zahlreiche Anwendungen (in Ingenieurwissenschaft und Physik), wir werden es hier nicht im Detail studieren.

Definition 7.1.6. Das *Spatprodukt* von drei Vektoren $a, b, c \in \mathbb{R}^3$ ist $a \cdot (b \times c)$.

Lemma 7.1.7. Das Spatprodukt erfüllt die folgenden Rechenregeln:

1. $a \cdot (b \times c)$ ist linear in jeder der drei Variablen.

2. $a \cdot (b \times c)$ ist alternierend: wenn zwei Vektoren gleich sind, dann ist das Spatprodukt 0.

3. $e_1 \cdot (e_2 \times e_3) = 1$.

Beweis. 1. Dies folgt sofort aus Lemma 7.1.4.2.

2. Wir sehen aus Lemma 7.1.4.1, dass $a \cdot (b \times b) = 0$.

Es gilt $a \cdot (a \times b) = 0$ denn mit Lemma 7.1.4.3 gilt $a \cdot (a \times b) = (a \times a) \cdot b$.

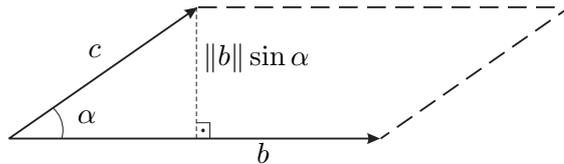
Schließlich gilt $a \cdot (b \times a) = -a \cdot (a \times b)$ mit Lemma 7.1.4.1 und wir haben gerade gezeigt, dass dies 0 ist.

3. Dies folgt sofort aus der Definition. □

Bemerkung 7.1.8. Die Eigenschaften aus Lemma 7.1.7 gelten auch für das Volumen des Parallelotops, das a, b, c aufspannen. (Wenn wir einen Vektor skalieren dann skaliert sich das Volumen. Wenn $a = b$ (oder $b = c$ oder $a = c$) dann ist das Parallelogramm entartet und hat kein Volumen. Das Volumen des Einheitswürfels soll 1 sein.)

Das Spatprodukt ist in der Tat nichts anderes als das Volumen des Spats, den a, b, c aufspannen. Mit den klassischen Beschreibungen aus Bemerkung 7.1.5 lässt sich das zeigen:

Wir betrachten das Parallelogramm, das b und c aufspannen als Basis, und mit $\alpha = \angle(b, c)$ haben wir die Grundfläche $A = \|b\| \cdot \|c\| \sin(\alpha)$.



Als nächstens bestimmen wir die Höhe des Spats. Dies ist $h = \|a\| \cos(\beta)$, wobei β der Winkel zwischen a und der Orthogonalen zu a und b ist. Also ist $\beta = \angle(a, b \times c)$.

Zusammen erhalten wir das Volumen $A \cdot h = a \cdot (b \times c)$. Allerdings haben wir einen Punkt vernachlässigt: Das Vorzeichen von $b \times c$ hängt von der Reihenfolge der Vektoren b und c ab. Der Absolutbetrag von $a \cdot (b \times c)$ ist also das Volumen des Spats, aber das Vorzeichen gibt die *Orientierung* der drei Vektoren a, b, c an!

Wir haben diese Vorüberlegungen nicht im Detail ausgeführt, da wir uns als nächstes mit einem abstrakten Volumenbegriff beschäftigen wollen, der in jedem endlichdimensionalen Vektorraum sinnvoll ist.

Es wird sich zeigen, dass die Eigenschaften aus Lemma 7.1.7 ausreichen, um einen Volumensbegriff einzuführen!

7.2 Die Determinantenabbildung

Wir wollen den Volumensbegriff für ein Polypop aus \mathbb{R}^2 und \mathbb{R}^3 auf allgemeine endlichdimensionale Vektorräume erweitern. Dazu betrachten wir eine Matrix, deren Zeilenvektoren die Transponierten der Basisvektoren sind, die das Polytop aufspannen.

Wir möchten dann das Volumen als Abbildung

$$M(n \times n, K) \rightarrow K$$

mit gewissen Eigenschaften verstehen. Dies ist die Determinantenabbildung:

Definition 7.2.1. Sei K ein Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\det : M(n \times n, K) \rightarrow K$$

$$A \mapsto \det(A)$$

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

$$\det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i + a_i')^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i')^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$

Da all unsere Vektoren als Spaltenvektoren zu betrachten sind schreiben wir den i ten Zeilenvektor von A als transponierte $(a_i)^T$ für einen Spaltenvektoren $a_i \in K^n$.

(D2) \det ist alternierend, d.h. stimmen zwei Zeilen überein, so ist $\det(A) = 0$.

(D3) \det ist normiert, d.h. $\det(E_n) = 1$.

Dies sind alles Eigenschaften, die wir vom Volumen eines Spats erwarten würde.

Beispiele 7.2.2. 1. Die Identität erfüllt (D1-3) für K^1 .

2. Man rechnet leicht nach, dass die Zuordnung $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$ (D1-3) für K^2 erfüllt.

3. Die Zuordnung

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \mapsto a \cdot (b \times c) = (a \times b) \cdot c$$

erfüllt nach Lemma 7.1.7 (D1-3) für \mathbb{R}^3 . Hier haben wir eine 3×3 -Matrix durch 3 Zeilenvektoren (transponierte Spaltenvektoren) dargestellt.

Gegeben eine Determinantenabbildung \det nennen wir $\det(A)$ die *Determinante von A*.

Diese Namensgebung legt schon nahe, dass wir zeigen wollen, dass es eine eindeutige Determinantenabbildung gibt. Wir werden nun Eigenschaften einer Determinantenabbildung zeigen, sowie ihre Existenz und Eindeutigkeit beweisen, und sie berechnen.

Satz 7.2.3. Sei K ein Körper und $n \in \mathbb{N}$. Sei $\det : M(n \times n, K) \rightarrow K$ 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

$$\det A = \lambda_1 \dots \lambda_n = \prod_{j=1}^n \lambda_j.$$

- Bemerkung 7.2.4.**
1. Eine Determinantenabbildung ist wegen 1. für $n \geq 2$ keine lineare Funktion!
 2. Sie ist auch nicht additiv: zum Beispiel gilt für $A = B = E_2$, dass $\det A + \det B = 2 \det E_2 = 2$, aber $\det(A + B) = \det(2E_2) = 4 \det E_2 = 4$.
 3. Wenn wir elementare Zeilenumformungen auf eine Matrix A anwenden, verändern wir $\det(A)$ in einer genau kontrollierten Weise. Wir können also den Gaußschen Algorithmus verwenden um Determinanten zu berechnen! Dazu folgt gleich ein Beispiel. Allerdings wissen wir zu diesem Zeitpunkt noch nicht, ob Determinantenabbildungen überhaupt existieren.

Beweis von Satz 7.2.3:. 1. Wir rechnen:

$$\begin{aligned} \det(\lambda A) &= \det \left(\lambda \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_n)^T \end{pmatrix} \right) = \det \begin{pmatrix} \lambda(a_1)^T \\ \vdots \\ \lambda(a_n)^T \end{pmatrix} = \lambda \det \begin{pmatrix} (a_1)^T \\ \lambda(a_2)^T \\ \vdots \\ \lambda(a_n)^T \end{pmatrix} \\ &= \lambda^2 \det \begin{pmatrix} (a_1)^T \\ (a_2)^T \\ \lambda(a_3)^T \\ \vdots \end{pmatrix} = \dots = \lambda^n \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_n)^T \end{pmatrix}. \end{aligned}$$

2. Aus der Zeilenlinearität (D1) folgt sofort:

$$\det A = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ 0 \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ 0 \cdot 0 \\ \vdots \\ (a_n)^T \end{pmatrix} = 0 \cdot \det \begin{pmatrix} (a_1)^T \\ \vdots \\ 0 \\ \vdots \\ (a_n)^T \end{pmatrix} = 0.$$

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):

$$0 = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i + a_j)^T \\ \vdots \\ (a_i + a_j)^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$

$$= 0 + \det A + \det B + 0.$$

4. B entstehe aus A durch Addition des λ -fachen der j -ten Zeile zur i -ten Zeile, mit $i \neq j$. Dann ist

$$\det B = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_{i-1})^T \\ (a_i + \lambda a_j)^T \\ (a_{i+1})^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_{i-1})^T \\ (a_i)^T \\ (a_{i+1})^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \lambda \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_{i-1})^T \\ (a_j)^T \\ (a_{i+1})^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$

$$= \det A + \lambda \cdot 0 = \det A,$$

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 5.3.8 verhält.

5. Sind alle $\lambda_i \neq 0$, so kann man A durch Zeilenumformungen, die die Determinante nicht ändern (Addition von Vielfachen einer Zeile zu einer anderen Zeile) in eine Diagonalmatrix überführen mit gleichen Diagonalelementen:

$$\det A = \det \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \cdot \det E_N \stackrel{(D3)}{=} \prod_{i=1}^n \lambda_i.$$

Sind nicht alle $\lambda_j \neq 0$, so sei i der größte Index, für den $\lambda_i = 0$ ist. Durch elementare Zeilenumformungen (Addition von Vielfachen der Zeilen $i+1$ bis n) können wir die ganze i -te Zeile gleich 0 setzen. Es gilt

$$\det A = \det \begin{pmatrix} \lambda_1 & & & & & \\ 0 & \ddots & & & * & \\ & & \lambda_{i-1} & & & \\ 0 & 0 & 0 & 0 & \dots & 0 \\ & & & & \lambda_{i+1} & * \\ & & & & & \ddots \\ & & & 0 & & \end{pmatrix} = 0 = \prod_{j=1}^n \lambda_j.$$

□

Korollar 7.2.5. Für jeden Körper K und jedes $n \geq 1$ gibt es höchstens eine Determinantenabbildung

$$\det : M(n \times n, K) \rightarrow K .$$

Beweis: Überführe eine Matrix A durch spezielle Zeilenumformung in eine obere Dreiecksmatrix A' , wobei k Zeilenvertauschungen auftreten. Dann gilt

$$A' = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & * & \\ 0 & & & \lambda_n \end{pmatrix}$$

Eine Matrix in Zeilenstufenform ist ein Spezialfall einer oberen Dreiecksmatrix, wir können also zum Beispiel unsere Matrix in Zeilenstufenform bringen. Dazu müssen wir keine Skalierungen von Zeilen vornehmen, die Addition von Vielfachen von Zeilen und die Vertauschung von Zeilen reichen aus.

Für jede Determinantenabbildung muss also gelten

$$\det A = (-1)^k \det A' = (-1)^k \prod_{j=1}^n \lambda_j . \quad \square$$

Bemerkung 7.2.6. Dies zeigt tatsächlich nur, dass es höchstens eine Determinantenabbildung gibt, es ist also gerechtfertigt $\det(A)$ die Determinante von A zu nennen.

Wir haben aber noch nicht gezeigt, dass es überhaupt eine Determinantenabbildung gibt! Insbesondere ist nicht klar, dass zwei verschiedene Abfolgen von Zeilenumformungen das gleiche Ergebnis liefern!

Beispiel 7.2.7. Wir arbeiten über dem Körper $K = \mathbb{C}$ und betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

Wir führen elementare Zeilenumformungen aus:

$$\begin{aligned} \det A &= -\det \begin{pmatrix} 1 & i & 1 \\ 0 & 1 & i \\ 2 & 3 & 4 \end{pmatrix} = -\det \begin{pmatrix} 1 & i & 1 \\ 0 & 1 & i \\ 0 & 3 - 2i & 2 \end{pmatrix} \\ &= -\det \begin{pmatrix} 1 & i & 0 \\ 0 & 1 & i \\ 0 & 0 & 2 - (3 - 2i)i \end{pmatrix} = -(-3i) = 3i . \end{aligned}$$

7.3 Existenz der Determinante

Satz 7.3.1. Für jeden Körper K und jedes $n \geq 1$ gibt es genau eine Determinantenabbildung

$$\det_n : M(n \times n, K) \rightarrow K$$

Beweis: Wir beweisen die 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:

$$A_{ij}^{Str} := \left(\begin{array}{cccc|cccc} a_{11} & \dots & a_{1j} & \dots & a_{1n} & & & \\ \vdots & & \vdots & & \vdots & & & \\ \hline a_{i1} & \dots & a_{ij} & \dots & a_{in} & & & \\ \vdots & & \vdots & & \vdots & & & \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} & & & \end{array} \right) \in M((n-1) \times (n-1), K)$$

und definiere für beliebiges, aber festes $j \in \{1, \dots, n\}$

$$\det_n(A) := \sum_{i=1}^n (-1)^{i+j} a_{ij} \det_{n-1}(A_{ij}^{Str}) .$$

Wir müssen jetzt die Axiome (D1)-(D3) überprüfen.

- (D1) Entstehe \tilde{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:

- $\tilde{A}_{kj}^{Str} = A_{kj}^{Str}$ da die einzige veränderte Zeile, nämlich die k -te Zeile, gestrichen wird.
- Für $i \neq k$ entsteht auch \tilde{A}_{ij}^{Str} aus A_{ij}^{Str} durch Multiplikation einer Zeile mit $\lambda \in K$, also gilt nach Induktionsannahme

$$\det_{n-1}(\tilde{A}_{ij}^{Str}) = \lambda \det_{n-1}(A_{ij}^{Str}) .$$

Somit ist

$$\begin{aligned} \det_n \tilde{A} &\stackrel{\text{def}}{=} \sum_{i \neq k} (-1)^{i+j} \tilde{a}_{ij} \det_{n-1}(\tilde{A}_{ij}^{Str}) + (-1)^{k+j} \tilde{a}_{kj} \det_{n-1} \tilde{A}_{kj}^{Str} \\ &= \sum_{i \neq k} (-1)^{i+j} a_{ij} \lambda \det_{n-1}(A_{ij}^{Str}) + (-1)^{k+j} \lambda a_{kj} \det_{n-1} A_{kj}^{Str} = \lambda \det_n(A) . \end{aligned}$$

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, wir vertauschen der Reihe nach die l -te Zeile mit der $(l - 1)$ -ten, mit der $(l - 2)$ -ten und so fort, bis die l -te Zeile an k -ter Stelle steht. 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 7.2.3.2 ihre Determinante. \square

Aus dem Satz folgt sofort der

Satz 7.3.2. 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 7.3.3. 1. K beliebig, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, indem wir nach der ersten Spalte entwickeln.

Nach der zweiten Spalte entwickelt erhalten wir $-bc + da$. 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.

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ & \diagdown & & \diagup & \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ & \diagup & & \diagdown & \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

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:

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix} &= 0 \cdot \det \begin{pmatrix} i & 1 \\ 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & i \\ 3 & 4 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \\ &= -(4 - 3i) + 2(1 - i^2) = 3i. \end{aligned}$$

Rechnerisch ist die Berechnung von Determinanten mit Hilfe von Entwicklungssätzen weniger effizient als mit dem Gaußalgorithmus.

Satz 7.3.4. Die Determinante ist auch linear in jeder Spalte; d.h. für alle Spaltenvektoren $a_1, \dots, a_n, a'_j, a''_j \in K^n$ und $\lambda \in K$ gilt

$$\begin{aligned} \det(a_1 \dots \lambda a_j \dots a_n) &= \lambda \det(a_1 \dots a_j \dots a_n) \\ \det(a_1 \dots a'_j + a''_j \dots a_n) &= \det(a_1 \dots a'_j \dots a_n) + \det(a_1 \dots a''_j \dots a_n) \end{aligned}$$

Beweis: Aus der Entwicklung nach der j -ten Spalte folgt:

$$\det(a_1 \dots \lambda a_j \dots a_n) = \sum_{i=1}^n (-1)^{i+j} (\lambda a_{ij}) \det A_{ij}^{Str} = \lambda \det A.$$

Analog zeigt man die Additivität. □

Lemma 7.3.5. Eine Determinantenabbildung schickt eine Matrix genau dann nach 0, wenn diese nicht maximalen Rang hat: $\det A = 0 \Leftrightarrow \text{rg } A < n$.

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' , vergleiche Betrachtungen 1.3.7 und 1.3.8.

Die Matrizen A und A' haben gleichen Rang, denn sie gehen auseinander hervor durch Multiplikation mit einem Produkt von Elementarmatrizen, vgl. Lemma 5.3.7, das eine invertible Matrix ist.

Aus Lemma ??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 Determinantenabbildung, $\det A = 0$ genau dann, wenn wenigstens ein λ_j verschwindet, $\lambda_j = 0$.

Wenn j maximal ist, so dass λ_j verschwindet, dann ist A' durch weitere Zeilenumformungen in eine Matrix zu überführen, in der die j -te Zeile 0 ist, vergleiche den Beweis von Satz 7.2.3.5. Aber dann ist der (Zeilen)Rang von $A < n$.

Wenn umgekehrt $\text{rg}(A) < n$ ist, dann muss A' in Zeilenstufenform eine Zeile haben, die gleich 0 ist, siehe Satz 5.3.5, aber das ist nur möglich, wenn ein $\lambda_j = 0$ ist. □

Wir sammeln ein paar Ergebnisse im folgenden Korollar:

Korollar 7.3.6. *Es sind für eine $n \times n$ -Matrix A äquivalent:*

1. A ist invertierbar.
2. $\text{rg}(A) = n$.
3. $\widetilde{\text{rg}}(A) = \text{rg}(A^T) = n$.
4. $\det(A) \neq 0$.

Beweis: 1. \Rightarrow 2.: Wenn A invertierbar ist stellt es eine surjektive Abbildung f dar und $\text{rg}(f) = \text{rg}(A) = n$.

2. \Leftarrow 1.: Umgekehrt ist eine surjektive Abbildung $K^n \rightarrow K^n$ invertierbar (Korollar 4.3.5) und damit folgt aus $\text{rg}(A) = n$, dass A invertierbar ist.

2. \Leftrightarrow 3. wegen Satz 5.5.12.

2. \Leftrightarrow 4. ist Lemma 7.3.5. □

Satz 7.3.7. *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

$$\begin{aligned} \widetilde{\det} : M(n \times n, K) &\rightarrow K \\ A &\mapsto \det A^T \end{aligned}$$

eine Determinantenfunktion ist.

(D1) folgt aus der Spaltenlinearität in Satz 7.3.4.

(D2) Wenn A zwei gleiche Zeilen hat, so hat A^T zwei gleiche Spalten. Also ist $\text{rg}(A^T) < n$, nach Lemma 7.3.5 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 7.3.8. 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.

2. Dann entsteht \widetilde{A}^T aus A^T durch Vertauschen zweier Zeilen. Aus Satz 7.2.3.3 folgt

$$\det \widetilde{A} = \det \widetilde{A}^T = -\det A^T = -\det A.$$

□

Als Beispiel verwenden wir die Entwicklung nach der ersten Zeile zur Berechnung der folgenden Determinante:

$$\begin{aligned} \det \begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix} &= 0 \cdot \det \begin{pmatrix} i & 1 \\ 3 & 4 \end{pmatrix} - \det \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} + i \cdot \det \begin{pmatrix} 1 & i \\ 2 & 3 \end{pmatrix} \\ &= -2 + i(3 - 2i) = 3i \end{aligned}$$

Vorlesungsskript bis hier. Dies ist das Ende des ersten Semesters. Es folgt das provisorische Skript für die erste Vorlesung des zweiten Semesters. Kleinere Änderungen in der Vorlesung sind möglich.

7.4 Eigenschaften der Determinante

Satz 7.4.1. Determinantenmultiplikationssatz *Für zwei Matrizen $A, B \in M(n \times n, K)$ gilt $\det(A \cdot B) = \det A \det B$.*

Beweis:. Wenn A oder B nicht invertierbar ist, dann gilt $\text{rg}(A) < n$ oder $\text{rg}(B) < n$. Damit ist $\det(A)\det(B) = 0$ wegen Lemma 7.3.5. Aber dann ist auch AB nicht invertierbar (die zugehörige Abbildung ist nicht surjektiv, wenn A nicht surjektiv ist und nicht injektiv, wenn B nicht injektiv ist), und $\det(AB) = 0$.

Wir nehmen also an $A, B \in GL(n, K)$ und schreiben $A = T_1 \cdots T_t$ und $B = S_1 \cdots S_s$ als Produkt von Elementarmatrizen mit Korollar 5.3.10. Wir berechnen nun die Determinanten der Elementarmatrizen:

- Die Determinante von $\tau(i, j)$ ist -1 wegen Satz 7.2.3.2 und $\det(E_n) = 1$.
- Die Determinante von $\delta(i, j, \lambda)$ ist 1 wegen Satz 7.2.3.4 und $\det(E_n) = 1$.
- Die Determinante der Diagonalmatrix $\Delta(1, \dots, \lambda, \dots, 1)$ ist λ wegen Satz 7.2.3.5.

Dann gilt $\det(A) = \det(T_1) \cdots \det(T_t) \cdots \det(E - n)$, denn jedes Produkt mit einer Elementarmatrix entspricht einer Zeilenumformung, die die Determinante genau um $\det(T_i)$ ändert!

Nach der gleichen Rechnung gilt $\det(B) = \det(S_1) \cdots \det(S_s)$ und

$$\det(A \cdot B) = \det(T_1) \cdots \det(T_t) \det(S_1) \cdots \det(S_s) = \det(A) \det(B). \quad \square$$

Korollar 7.4.2. *Ist $A \in GL(n, K)$, so ist $\det A \neq 0$ und es gilt $\det A^{-1} = (\det(A))^{-1}$.*

Beweis:. Es gilt

$$1 = \det(E_n) = \det(A \cdot A^{-1}) \stackrel{7.4.1}{=} \det A \cdot \det A^{-1}. \quad \square$$

Wir können nun sogar eine Formel für das Inverse einer Matrix angeben.

Satz 7.4.3. *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

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Sie haben auf dem Übungsblatt verifiziert, dass $E_2 = AA^{-1}$.

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 7.3.8 für die i -te Zeile.

• Der (i, k) -te Eintrag von AB ist für $i \neq k$:

$$\sum_{j=1}^n a_{ij} b_{jk} = \sum_{j=1}^n a_{ij} (-1)^{j+k} \det (A_{kj}^{Str}) = \sum_{j=1}^n \tilde{a}_{kj} (-1)^{j+k} \det (\tilde{A}_{kj}^{Str}) \stackrel{7.3.8}{=} \det \tilde{A},$$

wobei \tilde{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 \tilde{a}_{kj} haben wir eine entsprechende Änderung des Index vorgenommen. Wegen $i \neq k$ hat \tilde{A} zwei Zeilen mit identischen Einträgen, also ist $\det \tilde{A} = 0$.

• Aus dem Spaltenentwicklungssatz folgen die analogen Aussagen für das Produkt $B \cdot A$. □

Satz 7.4.4. Cramersche Regel Seien $a_1, \dots, a_n, b \in K^n$ und sei die Matrix $A = (a_1, \dots, a_n)$ invertierbar. Dann ist die eindeutige Lösung $x \in K^n$ des inhomogenen linearen Gleichungssystems von n Gleichungen für n Variablen

$$Ax = b$$

gegeben durch

$$x_i = \frac{\det (a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)}{\det A}$$

Beweis: • Das lineare Gleichungssystem $Ax = b$ hat die eindeutige Lösung

$$x = A^{-1}b$$

• Mit dem Ergebnis von Satz 7.4.3

$$(A^{-1})_{ij} = \frac{1}{\det A} (-1)^{i+j} \det A_{ji}^{Str}$$

folgt

$$x_i = \sum_{j=1}^n (A^{-1})_{ij} b_j = \frac{1}{\det A} \sum_{j=1}^n (-1)^{i+j} \det (A_{ji}^{Str}) \cdot b_j.$$

Wir zeigen nun die Identität

$$(-1)^{i+j} \det A_{ji}^{Str} = \det (a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n).$$

Sie folgt aus

$$(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n) = \begin{pmatrix} a_{1,1} & \dots & a_{i-1,1} & 0 & a_{i+1,1} & \dots & a_{n,1} \\ \vdots & & & \vdots & & & \vdots \\ a_{1,j} & \dots & a_{i-1,j} & 1 & a_{i+1,j} & \dots & a_{n,j} \\ & & & 0 & & & \\ & & & \vdots & & & \\ a_{1,n} & \dots & a_{i-1,n} & 0 & a_{i+1,n} & \dots & a_{n,n} \end{pmatrix}$$

und Entwicklungssatz 7.3.2 nach der i -ten Spalte. Aus der Spalten-Linearität der Determinante und

$$b = \sum_{j=1}^n b_j e_j$$

folgt sofort

$$x_i = \frac{1}{\det A} \sum_{j=1}^n \det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n) b_j = \frac{1}{\det A} \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n).$$

□

Betrachtung 7.4.5. 1. Ähnliche Matrizen haben die gleiche Determinante. Denn gilt

$$\tilde{A} = T \cdot A \cdot T^{-1}$$

mit $T \in GL(n, K)$, so ist nach dem Determinantenmultiplikationssatz 7.4.1

$$\det \tilde{A} = \det T \cdot \det A \cdot \det T^{-1} = \det A.$$

2. Ist V ein n -dimensionaler K -Vektorraum und $\Phi : V \rightarrow V$ 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 7.4.6. Für einen Endomorphismus $\Phi : V \rightarrow V$ eines endlich-dimensionalen Vektorraums V heißt

$$\det(\Phi) := \det(M_{\mathcal{B}}^{\mathcal{B}}(\Phi))$$

Determinante von Φ , wobei \mathcal{B} eine beliebige geordnete Basis von V ist.

Bemerkungen 7.4.7. 1. Für Endomorphismen $\Phi, \Psi : V \rightarrow V$ 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

$$\det R_\theta = \det(M(R_\theta)) = \det \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \cos^2 \theta + \sin^2 \theta = 1.$$

A Leere Summen und Produkte

An verschiedenen Stellen betrachten wir leere Summen und Produkte und ähnliche Konstruktionen. Wir können diese Konstruktionen per Konvention definieren, aber es stellt sich heraus, dass es jeweils nur eine sinnvolle Antwort gibt.

Was ist eine leere Summe? Wenn wir eine endliche Summe betrachten, z.B. $A = \sum_{i=1}^n a_i$ dann soll natürlich für jedes k gelten

$$A = \sum_{i=1}^n a_i = \sum_{i=1}^k a_i + \sum_{i=k+1}^n a_i. \quad (5)$$

Für $k \in \{1, 2, \dots, n\}$ ist das klar. Aber was ist mit $k = 0$ oder $k = n + 1$? $\sum_{i=1}^0 a_i$ ist eine leere Summe, da es keine ganze Zahl i mit $i \geq 1$ und $i \leq 0$ gibt, haben wir keinen Summanden. Aber 5 kann weiterhin gelten, solange wir diese leere Summe gleich 0 setzen.

Allgemein: Es soll immer gelten $\sum_{i \in I} a_i + \sum_{i \in J} a_i = \sum_{i \in I \cup J} a_i$ für Indexmengen I und J mit $I \cap J = \emptyset$. Aber wenn J leer ist dann kann dies nur für $\sum_{i \in J} a_i = 0$ gelten.

Wir können auch einen Schritt zurück gehen und uns fragen, was Addition eigentlich bedeuten soll. Wenn wir uns erinnern, wie wir Addition in der Grundschule gelernt haben, ("ich habe 3 Äpfel und jemand gibt mir noch 5 Äpfel dazu") und das mit unseren neuen abstrakten Sichtweise formalisieren, dann erhalten wir: Sei eine endliche Menge von endlichen Mengen $\{A_1, \dots, A_n\}$ gegeben, die alle disjunkt sind, d.h. $A_i \cap A_j = \emptyset$ für alle i, j . Dann soll

$$|A_1 \cup \dots \cup A_n| = |A_1| + \dots + |A_n|$$

gelten, hier ist $|A_i|$ die Mächtigkeit von A_i , also die Anzahl der Elemente.

Dann ist die leere Summe genau die Mächtigkeit der leeren Vereinigung: $\sum_{\emptyset} |A_i| = |\cup_{\emptyset} A_i|$. Aber die leere Vereinigung muss die leere Menge sein, denn sie enthält genau die Elemente, die in einer der Mengen enthalten sind, die wir vereinigen. Aber es gibt gar keine Mengen, die wir vereinigen! Also gilt $\sum_{\emptyset} |A_i| = |\cup_{\emptyset} A_i| = 0$. Wenn ihnen die Leere rechts vom Summenzeichen nicht gefällt, können wir auch schreiben $\sum_{\emptyset} |A_i| = |\cup_{\emptyset} A_i|$. Aber das A_i ist nur ein Platzhalter, wir haben ja gar keine Summanden/Mengen.

Wir haben bisher nicht spezifiziert, was für Summen wir betrachten, aber das erste Argument gilt in beliebigen abelschen Gruppen, insbesondere für ganze Zahlen, reelle Zahlen oder Vektoren in einem beliebigen Vektorraum!

Insbesondere folgt, dass der Nullvektor als leere Summe in jedem Spann enthalten ist, sogar im Spann der leeren Menge!

Betrachten wir als Nächstes noch das leere Produkt. Wieder soll gelten:

$$\prod_{i=1}^n a_i = \prod_{i=1}^k a_i \cdot \prod_{i=k+1}^n a_i. \quad (6)$$

Aber damit dies stimmt, wenn $k = 0$ oder $k = n + 1$ ist muss das leere Produkt nun 1 sein.

Die leere Summe gibt das neutrale Element der Addition, das leere Produkt ergibt das neutrale Element der Multiplikation.

Allgemein gilt in jeder Gruppe, dass die "leere Verknüpfung" das neutrale Element sein muss!

Wenn wir diese Aussage mit Mengen interpretieren wollen, dann ist das Ergebnis vielleicht etwas überraschend: Das leere Produkt von Mengen ist die Menge mit einem Element! Wie kann das sein? Was sind die Elemente in einem kartesischen Produkt $\prod_{i=1}^n A$ bei dem alle

Faktoren gleich sind? Ein Element $a_i \in A$ für jedes i . Aber das ist nichts anderes als eine Funktion $\{1, \dots, n\} \rightarrow A$. Also ist $|\prod_{i=1}^n A|$ die Anzahl dieser Funktionen.

Und für jede Menge es gibt genau eine Funktion von der leeren Menge nach A ! Nämlich die leere Funktion, das hatten wir in der Übungsaufgabe 4.2.2 gezeigt.

Da das leere Produkt 1 muss insbesondere auch gelten $0! = 1$. Wir können auch dies aus der Perspektive von Abbildungen von Mengen betrachten: $n!$ ist die Anzahl aller Permutationen von n Elementen, also aller Bijektionen $\{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Dann ist $0!$ die Anzahl der Bijektionen von \emptyset nach \emptyset . Wie wir uns gerade erinnert haben, gibt es genau eine Abbildung von \emptyset nach \emptyset . Diese leere Abbildung ist trivialerweise injektiv und surjektiv, also ist es eine Bijektion und $0! = 1$.

B Was bisher geschah: Kurzzusammenfassung der Kapitel 2 bis 5

1. Wir haben zuerst unser algebraisches Grundwerkzeug entwickelt:

- Gruppen, z.B.. \mathbb{Z}/n oder S_n
- Körper, insbesondere \mathbb{R} , \mathbb{Q} , die komplexen Zahlen \mathbb{C} und den endlichen Körper \mathbb{F}_p mit p Elementen, wobei p prim ist.
- Ringe wie \mathbb{Z} oder $(\mathbb{Z}/n, +, \cdot)$ oder den Polynomring $K[T]$ für einen Körper K getroffen.

Abbildungen, die algebraische Struktur erhalten heißen Homomorphismen.

2. Für einen gegebenen Körper K haben wir den Begriff des K -Vektorraums kennengelernt. Wir betrachten auch:

- Untervektorräume $U \leq V$.
- Zu einem Untervektorraum $U \subset V$ können wir den Quotientenvektorraum V/U konstruieren, in dem wir den Unterraum U "gleich null setzen". Es gibt eine kanonische Surjektion $V \rightarrow V/U$.
- Zu zwei Untervektorräumen bilden wir ihre Summe $W_1 + W_2$.
- Für eine Familie von Vektorräumen $(V_i)_{i \in I}$ bilden wir die äußere direkte Summe $\bigoplus_{i \in I} V_i$.

3. Wir haben K -lineare Abbildungen $f : V \rightarrow W$ definiert. Eine K -lineare Abbildung mit K -linearem Inversen ist ein Isomorphismus, geschrieben $V \cong W$.

- Urbilder von Untervektorräumen in W sind Untervektorräume von V . Insbesondere ist der Kern von f als Urbild von $0 \in W$ ein Untervektorraum von V . Die lineare Abbildung f ist genau dann injektiv, wenn $\ker f = \{0\}$ gilt.
- Bilder von Untervektorräumen in V sind Untervektorräume in W , insbesondere ist $f(V) \leq W$.
- Der Raum $\text{Hom}_K(V, W)$ der K -linearen Abbildungen ist selbst ein K -Vektorraum.
- Es gilt der Homomorphiesatz 3.4.4: für eine lineare Abbildung $f : V \rightarrow W$ erhalten wir einen eindeutigen Isomorphismus $\bar{f} : V/\ker f \xrightarrow{\sim} \text{Im}(f)$ und eine Faktorisierung von $f = \iota \circ \bar{f} \circ \pi$ wie in diesem Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \pi \downarrow & & \uparrow \iota \\ V/\ker f & \xrightarrow{\bar{f}} & \text{Im}(f) \end{array}$$

4. Aus den Vektoren eines Vektorraums kann man Linearkombinationen bilden. Dies führt auf zwei wichtige Begriffe

- $\mathcal{B} \subset V$ ist ein Erzeugendensystem wenn gilt $\text{span}_K(\mathcal{B}) = V$.
- \mathcal{B} ist linear unabhängig, wenn nur triviale Linearkombinationen von Elementen in \mathcal{B} gleich null sind.

- Eine Familie \mathcal{B} von n Vektoren in V definiert eine lineare Abbildung von K^n nach V , die injektiv ist, wenn \mathcal{B} linear unabhängig ist und surjektiv, wenn \mathcal{B} Erzeugendensystem ist.
- Linear unabhängige Erzeugendensysteme heißen Basen. Basen sind maximale linear unabhängige Familien und minimale Erzeugendensysteme. Sie existieren für jeden Vektorraum.
- Die Anzahl der Basiselemente ist eindeutig und heißt die Dimension des Vektorraums.
- Der Basisauswahlsatz erlaubt es uns, aus Erzeugendensystemen eine Basis auszuwählen.
- Der Basisergänzungssatz erlaubt es, linear unabhängige Familien zu Basen zu ergänzen.

Für jede lineare Abbildung $f : V \rightarrow W$ zwischen endlichdimensionalen Vektorräumen gilt die Dimensionsformel. Mit $\text{rg}(f) = \dim(\text{Im}(f))$ ist $\text{rg}(f) + \dim \ker(f) = \dim(V)$.

Wir können auch die folgenden Dimensionen berechnen:

- $\dim_K(W_1 + W_2) = \dim_K(W_1) + \dim_K(W_2) - \dim_K(W_1 \cap W_2)$.
- $\dim(V/U) = \dim V - \dim U$
- $\dim \text{Hom}_K(V, W) = \dim(V) \cdot \dim(W)$.

5. Wir haben den Vektorraum $M(m \times n, K)$ der $m \times n$ -Matrizen betrachtet.

- Wir können Matrizen multiplizieren: $(A \cdot B)_{ik} = \sum_j A_{ij} B_{jk}$. Vektoren lassen sich als Matrizen auffassen und insbesondere gilt $(A \cdot v)_i = \sum_j A_{ij} v_j$
- Wir können Matrizen transponieren: $(A^T)_{ji} = A_{ij}$. Es gilt $(AB)^T = B^T A^T$.
- Invertierbare Matrizen bilden eine Gruppe $GL(n, K)$.
- Matrizen in $M(m \times n, K)$ stellen lineare Abbildungen von K^n nach K^m dar.

Wir haben explizite Beschreibungen für jede lineare Abbildungen $f : V \rightarrow W$ zwischen endlich-dimensionalen Vektorräume. Dazu müssen wir geordnete Basen $\mathcal{A} = (v_1, \dots, v_n)$ für V und $\mathcal{B} = (w_1, \dots, w_m)$ für W wählen.

- Jede geordnete Basis \mathcal{A} von V gibt einen Isomorphismus vom Standardvektorraum K^n mit Standardbasis auf V :

$$\Phi_{\mathcal{A}} : K^n \rightarrow V \text{ mit } \Phi_{\mathcal{A}}(e_i) = v_i .$$

- Es gibt eindeutige Elemente (m_{ij}) von K so dass $f(v_i) = \sum_j m_{ji} w_j$. Dann ist $M_{\mathcal{B}}^{\mathcal{A}}(f) = (m_{ij})$ die darstellende Matrix von f bezüglich der Basen \mathcal{A} und \mathcal{B} . Dies gibt einen Isomorphismus

$$M_{\mathcal{B}}^{\mathcal{A}} : \text{Hom}_K(V, W) \rightarrow M(m \times n, K)$$

von K Vektorräumen.

Es gilt $M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) = M_{\mathcal{C}}^{\mathcal{A}}(g \circ f)$.

- Ein Basiswechsel zwischen Basen \mathcal{A} und \mathcal{A}' wird durch invertierbare Transformationsmatrizen

$$T_{\mathcal{A}'}^{\mathcal{A}} = M_{\mathcal{A}'}^{\mathcal{A}}(\text{id}_V)$$

beschrieben. Für lineare Abbildungen gilt die Transformationsformel 5.4.6

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\Phi) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(\Phi) \cdot \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1}.$$

- Zwei (nicht notwendigerweise quadratische) Matrizen X, Y heißen äquivalent, wenn es invertible 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 $r = \text{rg}(A)$. Zeilenrang und Spaltenrang einer Matrix sind gleich dem Rang der zugehörigen linearen Abbildungen.

- Zwei *quadratische* Matrizen X, Y heißen ähnlich, wenn es *eine* invertible Matrix S gibt mit $Y = SXS^{-1}$.

6. Für die Lösungsmenge

$$\text{Lsg}(A, b) := \{x \in K^n \mid Ax = b\}$$

eines inhomogenen linearen Gleichungssystems gilt:

- $\text{Lsg}(A, b) = \emptyset$ genau dann, wenn $b \notin \text{Im } A$.
- $\text{Lsg}(A, b)$ ist entweder leer oder ein affiner Unterraum der Form $x_0 + \ker(\alpha)$ mit Dimension $n - \text{rg } A$. Hier ist $A = M(\alpha)$ und $\ker(\alpha)$ sind die Lösungen des zugehörigen homogenen Gleichungssystems.

Mit dem Gaußschen Algorithmus können wir durch elementare Zeilenumformungen

- lineare Gleichungssysteme lösen,
- Matrizen invertieren,
- den Rang einer Matrix berechnen.