Lineare Algebra

Studienjahr 2023/24

Julian Holstein
Fachbereich Mathematik

(Stand: 1. August 2024)

Inhaltsverzeichnis

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 beliebigier 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 ihnnen 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.

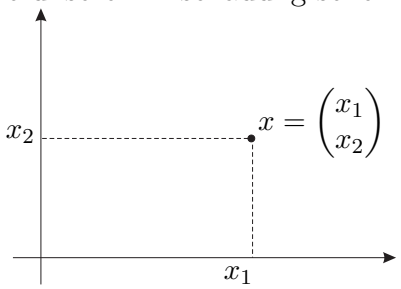
- Reelle Zahlen können addiert und subtrahiert werden. Für Addition und Multiplikation beliebiger reeller Zahlen gelten Assoziativ- und Kommutativgesetze. Es gibt eine reelle Zahl $0 \in \mathbb{R}$, so dass für alle reellen Zahlen, also alle $a \in \mathbb{R}$, gilt $0 + a = a + 0 = a$; für die

[1] Im zweiten Semester begegnen uns auch ein paar qudratische 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 $\mathbb{R}^2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\middle|\, x_1, x_2 \in \mathbb{R} \right\}$ von *geordneten* Paaren reeller Zahlen betrachten. Durch die Einführung kartesischer Koordinaten können wir diese als mathematisches Model für die Ebene unserer Anschauung sehen:



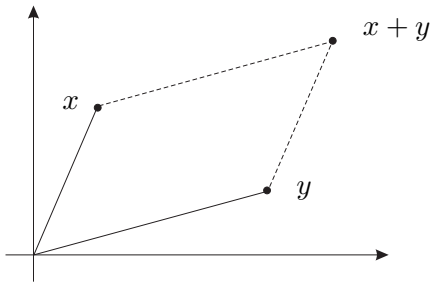
Wir wollen auf die Menge $\mathbb{R}^2$ noch andere Struktur aufprägen: Je zwei Elementen $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ können wir durch komponentenweise Addition ihre Summe

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

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 ":=".)

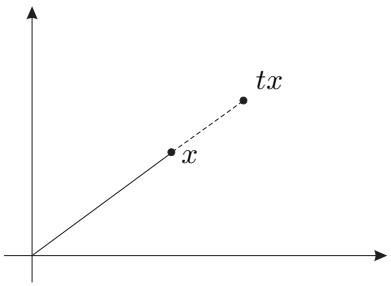
Wir stellen den Vektor $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ durch einen Pfeil vom Ursprung mit Spitze im Punkt mit Koordinaten (x_1, x_2) dar. Die Summe wird dann so veranschaulicht:



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:

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

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 veraunschaulichende 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 Multplikation in $\mathbb{R}$ und $t'x$ die Skalarmultiplikation, also komponenntenweise 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 $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ einen *Vektor* (in $\mathbb{R}^n$) und für $i = 1, 2, \ldots, n$ die reelle Zahl v_i die i-te *Komponente* oder *Koordinate*.

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.

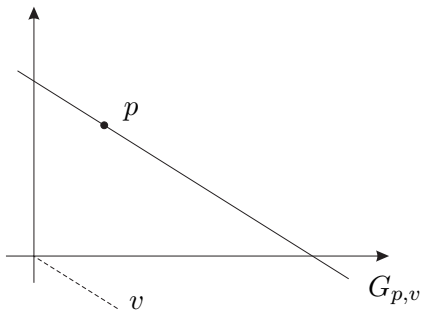
Durch *Definitionen* schafft man sich neue mathematische Begriffe. $\mathbb{R}^2$ ist eine Menge und hat eine Klasse interessanter Teilmengen. Die folgenden Überlegungen funktionieren für allgemeines $\mathbb{R}^n$ genauso gut wie für $\mathbb{R}^2$, also schreiben wir einfach $\mathbb{R}^n$.

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 $\{p + tv | \, t \in \mathbb{R}\}$ 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 nehemen 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.

$\langle\langle$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.$\rangle\rangle$

" $\Rightarrow$ ". 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 $q = p + s_1 v$. $\langle\langle$Die Definition besagt, dass es eine Zahl gibt, die die Gleichung erfüllt, und wir geben dieser Zahl sofort einen Namen!$\rangle\rangle$

Ferner gilt auch $q + w \in G_{q,w} = G_{p,v}$, also gibt auch es ein $s_2 \in \mathbb{R}$ mit $q + w = p + s_2 v$. 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. $\langle\langle$Das ist die Definition von Gleichheit von Teilmengen einer Menge, und es entspricht unserer Anschauung.$\rangle\rangle$ 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 $q = p + t_0 v$ gilt. Sei $r \in G_{q,w}$ beliebig, also gilt $r = q + t_1 w$ mit einem geeigneten $t_1 \in \mathbb{R}$.[2] Einsetzen liefert

$$r = p + t_0 v + t_1 s v = p + (t_0 + t_1 s)v \, ,$$

woraus $r \in G_{p,v}$ folgt. Wir hatten über r keine weiteren Annahmen gemacht, als dass $r \in G_{q,w}$ 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 $u = p + t_2 v$ schreiben mit einem geeigneten $t_0 \in \mathbb{R}$.

Wir haben immer noch $q = p + t_0 v \in G_{p,v}$. Durch Umstellen folgt $p = q - t_0 v$ 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*

[2] Das 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

$$u = p + t_2 v = q - t_0 v + t_2 v = q + (t_2 - t_0)s^{-1} w \, .$$

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 $G_{q,w} = G_{p,v}$. □

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 geeignetem 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 $b = a + t_0 v$.

Damit ist $b - a = t_0 v$. Weiterhin gilt $b - a \neq 0$ nach Annahme, also können wir Lemma 1.2.6 anwenden und es gilt $G_{a,v} = G_{a,b-a}$. □

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 $a_1 x_1 + a_2 x_2 = c$ eine lineare Gleichung in zwei Variablen. Wenn $a \neq 0$ dann ist die Menge aller Lösungen $L_{a,c} = \{(x_1, x_2) \in \mathbb{R}^2 \mid a_1 x_1 + a_2 x_2 = c\}$ 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: $a_1(b_1 + tv_1) + a_2(b_2 + tv_2) = a_1(\frac{c}{a_1} + ta_2) + a_2(0 + t(-a_1)) = c + ta_1 a_2 - ta_1 a_2 = c$. Also gilt $G_{b,v} \subset L_{a,c}$.

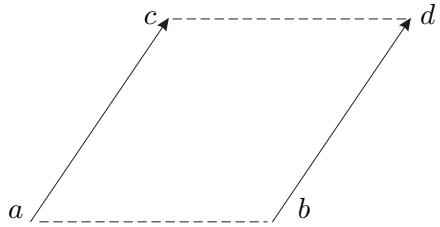
Wenn umgekehrt x die Gleichung erfüllt, dann betrachten wir $x - b$. Es gilt $x_1 = \frac{c - a_2 x_2}{a_1}$ und damit $(x - b)_1 = \frac{c - a_2 x_2}{a_1} - \frac{c}{a_1} = -\frac{a_2}{a_1} x_2$. Außerdem haben wir $(x - b)_2 = x_2 - 0$. Damit is $x - b = \frac{x_2}{a_1} \begin{pmatrix} -a_2 \\ a_1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

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:



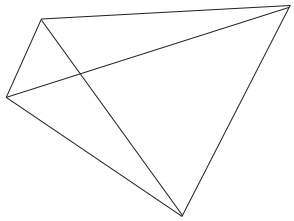
Ein Parallelogramm heißt *nicht-ausgeartet*, falls keine drei Punkte auf einer Geraden liegen.

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 frustiert 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:

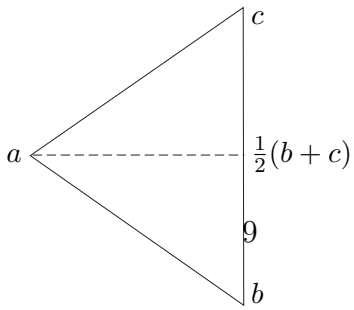
$$\frac{1}{2}(a + d) - \frac{1}{2}(b + c) = \frac{1}{2}(a + d - b - c) = 0\,.$$

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. □

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

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

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

2. Gilt $b_1 = \ldots = b_m = 0$, 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

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

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

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 $b = (b_1 \ldots, b_m) \in \mathbb{R}^m$ zusammen. Die Matrix

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

heißt *erweiterte Koeffizientenmatrix* des inhomogenen linearen Gleichungssystems.

4. Die Lösungsmenge des Gleichungssystems bezeichnen wir $\mathrm{Lsg}(A, b)$. 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 $i = 2, \ldots, m$ gilt: sind die ersten $(k-1)$ Einträge der $(i-1)$–ten Zeile gleich Null, so sind die ersten k Einträge der i–ten Zeile gleich Null, wobei $k = 1, \ldots, n$.

2. Eine Matrix ist in *spezieller Zeilenstufenform*, wenn sie in Zeilenstufenform ist und falls für alle $i = 1 \ldots m$ gilt: ist $a_{i1} = a_{i2} = \ldots = a_{i,k-1} = 0$ und $a_{ik} \neq 0$, so ist $a_{ik} = 1$. 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: $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\quad$ $\begin{pmatrix} 1 & 5 & 9 \\ 0 & 0 & 0 \end{pmatrix}$
- Matrizen in Zeilenstufenform, aber nicht spezieller Zeilenstufenform: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \mathbf{5} \end{pmatrix}$
- Matrizen, die nicht in Zeilenstufenform sind:

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

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

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

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

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 $i \in \{1, \ldots, m\}$ 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

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

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

Beweis:. Es kommt offensichtlich nicht auf die die Reihenfolge der Gleichungen an; damit ist 2. klar. Auch 1. sieht man sofort.

Beim dritten Typ kommt es auf nur zwei Zeilen i, k an. Es reicht also aus zu zeigen, dass die Gleichungssysteme

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i \tag{1}$$
$$a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n = b_k \tag{2}$$

und

$$a_{i1}x_1 + a_{i2}x_2 \qquad + \ldots + a_{in}x_n \qquad = b_i \tag{3}$$
$$(a_{k1} + \lambda a_{i1})x_1 + (a_{k2} + \lambda a_{i2})x_2 + \ldots + (a_{kn} + \lambda a_{in})x_n = b_k + \lambda b_i \tag{4}$$

gleiche Lösungsräume haben. Erfüllt $x = (x_1, \ldots, x_n)$ 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). □

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 $(\widetilde{A}, \widetilde{b})$.

3. Löse das lineare Gleichungssystem $\widetilde{A}x = \widetilde{b}$ 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 Ausssage $3 \cdot 4 = 4$ 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 $n \in \{1, 2, 3, \ldots\}$ ist eine n-stellige *Verknüpfung* von gegebenen Aussagen $A_1, A_2, \ldots, A_n$ eine Aussage $V(A_1, \ldots, A_n)$, deren Wahrheitswert durch die Wahrheitswerte der gegebenen Aussagen $A_1, \ldots A_n$ eindeutig bestimmt ist. Sie wird durch eine *Wahrheitstafel* beschrieben, die die Wahrheitswerte in Abhängigkeit der Wahrheitswerte der gegebenen Aussagen angibt.

Insbesondere definieren wir für zwei Aussagen A und B:

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

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

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

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

Es handelt sich also um ein nicht ausschließendes "oder".

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 $\bot$ nützlich. Es hat $\top$ immer den Wahrheitswert "wahr" und $\bot$ das immer den Wahrheitswert "falsch".

Wir können verschiedene Verknüpfugen 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 Wahrheistwert hat wie $(A \Leftarrow B) \wedge (B \Rightarrow A)$. Das haben wir schon benutzt!

Definition 1.4.7. Gegeben seien mehrere Aussagen $A, B, C, \ldots$ und zwei Aussagen X und Y, die durch die Verknüpfung dieser Aussagen entstanden sind. Wenn die Aussage

$$X \Leftrightarrow Y$$

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

Satz 1.4.8. *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.

2. 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 einen 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 $P: M \to \{w, f\}$.

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.

2. Man kann Prädikate mit Negationen, Konjunktionen, Disjunktionen, Implikationen und Äquivalenzen kombinieren und neue Prädikate erhalten. Zum Beispiel ordnet für ein gegebenes Prädikat $P : M \to \{w, f\}$ das Prädikat $\neg P : M \to \{w, f\}$ $m \in M$ 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:

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

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

Der Doppelpunkt gehört zur Aussage dazu, wird aber gelegetlich 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:

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

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

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.

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

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 y gilt $C(x,y)$ dann gibt es auch für alle y' ein x' (nämlich x!), 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. $\Big(\forall x \in M : A(x)\Big) \wedge \Big(\forall x \in M : B(x)\Big) \Leftrightarrow \forall x \in M : A(x) \wedge B(x)$.

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

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. $\Big(\forall x \in M : A(x)\Big) \vee \Big(\forall x \in M : B(x)\Big) \Rightarrow \forall x \in M : A(x) \vee B(x)$.

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

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.

[3] Marmite 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.

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

und schließlich verwenden wir die Verneinung von $A \Rightarrow B$: $\neg(A \to B) \Leftrightarrow \neg(\neg A \vee B) \Leftrightarrow A \wedge \not B$.

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

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. es kommt nicht mehrmals das gleiche Element in einer Menge vor. Wir vereinbaren daher $\{x, x, y\} = \{x, y\}$. Auch kommt es nicht 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 $a_1, a_2, \ldots, a_n$ enthält, schreibt man auch $M = \{a_1, a_2, \ldots, a_n\}$.
2. Es gibt eine ausgezeichnete Menge, die leere Menge $\emptyset$, die keine Elemente enthält.
3. Eine Menge ist durch ihre Elemente eindeutig bestimmt. 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 Axiomsystem 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. $\langle\langle$Sorgen Sie sich nicht zu sehr, wenn einige der Axiome schwer zugänglich sind, wir werden sie größtenteils nicht mehr gebrauchen.$\rangle\rangle$

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 Paarungsaxioms konstruieren wir die Mengen $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ ...

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

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

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

   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

   $$\forall x, y, z\colon (E(x, y) \wedge E(x, z) \Rightarrow y = z) \Rightarrow \forall A\colon \exists M\colon \forall y\colon (y \in M \iff \exists x\colon (x \in A \wedge E(x, y)))$$

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

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

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 $\cap_{i \in I} A_i$ als die Menge aller Elemente, die in jedem A_i enthalten sind, symbolisch

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

(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

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

(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. *Kommutativgesetze: $A \cap B = B \cap A$ und $A \cup B = B \cup A$*
2. *Assoziativitätsgesetze: $(A \cap B) \cap C = A \cap (B \cap C)$ und $A \cup (B \cup C) = (A \cup B) \cup C$*
3. *Distributivgesetze: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.*

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:

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

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 Axiomsystem 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 liefertein *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 Kommutativiät, Assoziativität und Destributivität 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 Indukton 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 $1+3+\ldots+(2n-1) = n^2$. (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 $n + 1$:

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

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

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

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

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

für die geordneten n-Tupel von Elementen in A.

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

Beispiel 1.6.2. Die Menge aller Karten in einem Kartenspiel erhalten wir als kartesischen Produkt der Farben $\{\diamondsuit, \spadesuit, \heartsuit, \clubsuit\}$ 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:

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

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

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

Beispiele 1.6.4.

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

2. $X = \mathbb{R}$ 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 $\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix} \sim \begin{pmatrix}y_1\\\vdots\\y_n\end{pmatrix} :\Leftrightarrow x_1^2 + \ldots + x_n^2 = y_1^2 + \ldots + y_n^2$. 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 $A = A'$. □

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 $X/R \to X$.

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 $0, 1, \ldots n-1$. Die kanonische Abbildung ordnet einer Zahl die Restklasse ihres Rests nach Division durch n zu, etwa für $n = 12$ haben wir $23 \mapsto [11]$ – 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, $\frac{a}{b} = [(1,2),(2,4),(-3,-6),\ldots]$. 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 $(a,b) \in R$ exisitert. Wir bezeichnen dieses b mit $f(a)$ und betrachten es als *Bild* von a.

Wir schreiben auch $f : A \to B$ 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 $f : A \longrightarrow B$ 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 $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Die *Verkettung*, *Verknüpfung* oder *Komposition* $g \circ f$ von g mit f ist die durch

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

definierte Abbildung.

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 $h : C \to D$ eine weitere Abbildung, dann gilt*

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

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

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

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 $f = \mathrm{id}_A : A \to A$ 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 $f \circ \mathrm{id} = f = \mathrm{id} \circ f$. (Dies ist die einzige Relation, die sowhol Äquivalenzrelation als auch Funktion ist.)

2. Für gegebene Mengen M, N bilden die Abbildungen $f : M \to N$ eine Menge $\mathrm{Abb}(M, N)$, 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 $f : \mathbb{R} \to \mathbb{R}^2$ 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 $f : \mathbb{Z} \to \mathbb{N}$ mit $x \mapsto x^2$, aber auch durch eine Fallunterscheidung, etwa
$$f : \mathbb{R} \to \mathbb{N} \qquad f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \notin \mathbb{Q} \end{cases} \quad \text{oder auch durch eine Tafel.}$$
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 $f : \mathbb{Q} \to \mathbb{Q}$, die eine Umkehrabbildung $g : \mathbb{Q} \to \mathbb{Q}$ besitzt, nämlich die Rechenvorschrift $x \mapsto \frac{1}{2}x$. Die Verkettungen $f \circ g$ und $g \circ f$ der beiden Abbildungen sind jeweils die Identität auf $\mathbb{Q}$. Die entsprechende Abbildung $\mathbb{Z} \to \mathbb{Z}$ hat aber keine Umkehrabbildung.

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

Definition 1.6.13. 1. Eine Abbildung $f : A \to B$ heißt *surjektiv*, falls es zu jedem $b \in B$ ein $a \in A$ gibt mit $f(a) = b$, d.h. falls für ihr Bild $f(A)$ gilt $f(A) = B$.

2. Eine Abbildung $f : A \to B$ heißt *injektiv*, falls aus $f(a_1) = f(a_2)$ folgt $a_1 = a_2$, d.h. aus $a_1 \neq a_2$ folgt stets $f(a_1) \neq f(a_2)$.

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

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

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

3. Ist M eine endliche Menge, so gilt für alle Abbildungen $f : M \to M$: f bijektiv $\Leftrightarrow$ f injektiv $\Leftrightarrow$ f surjektiv. Ist aber M eine unendliche Menge, so gibt es Abbildungen $f : M \to M$, die injektiv, aber nicht surjektiv sind, und Abbildungen, die surjektiv, aber nicht injektiv sind. Betrachte zum Beispiel die Abbildungen $f : \mathbb{N} \to \mathbb{N}$ und $g : \mathbb{N} \to \mathbb{N}$ 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 $f : A \to B$ hat eine eindeutige* Umkehrabbildung *oder* inverse Funktion *f^{-1}, die $f \circ f^{-1} = \mathrm{id}_B$ und $f^{-1} \circ f = \mathrm{id}_A$ 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 a ja $f(a) = f(a)$ erfüllt.

Wir zeigen Eindeutigkeit. Sei $g : B \to A$ 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 $b = f(f^{-1}(b))$. □

Definition 1.6.16.

1. Eine Menge M heißt *endlich*, wenn es ein $n \in \mathbb{N}_0$ und eine Bijektion
$$f : \{0, 1, \ldots, n-1\} \to M$$
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 $f : A \to B$ gibt.

Beispiele 1.6.17.

1. $\mathbb{N}$ und $\mathbb{Z}$ sind gleichmächtig, obgleich $\mathbb{N} \subsetneq \mathbb{Z}$. Eine Bijektion $f : \mathbb{N} \to \mathbb{Z}$ ist
$$f(a) = \begin{cases} \frac{a}{2} & \text{falls } a \text{ gerade} \\ -\frac{a+1}{2} & \text{falls } a \text{ ungerade} \end{cases}$$

2. Die Mengen $\mathbb{N}$ und $\mathbb{Q}$ sind gleichmächtig, nicht aber $\mathbb{N}$ und $\mathbb{R}$ (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
$$F : \mathrm{Abb}(A \times B, C) \to \mathrm{Abb}(A, \mathrm{Abb}(B, C))$$
mit Umkehrabbildung G. Hierbei wird $\phi : A \times B \to C$ auf die Abbildung $F(\phi) : A \to \mathrm{Abb}(B, C)$ abgebildet, die $a \in A$ die Abbildung $\phi(a, -) : b \mapsto \phi(a, b)$ zuordnet.

Umgekehrt wird einer Abbildung $\psi : A \to \mathrm{Abb}(B, C)$ die Abbildung $G(\psi) : A \times B \to C$ 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 Vektorraums abstruhiert 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

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

genannt *Verknüpfung*, so dass die folgenden Eigenschaften erfüllt sind:

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

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

a) Für alle $a \in G$ gilt $e \cdot a = a = a \cdot e$

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.

$\langle\langle$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".$\rangle\rangle$

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

• 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). □

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 $\{0, 1, \ldots, m-1\}$ lassen. Eine Äquivalenzklasse wird auch *Restklasse modulo m* genannt.

Für jedes $0 \leq r \leq m-1$ ist

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

eine Restklasse.

Wir schreiben $a = a' \mod m$ und sagen "a ist kongruent a' modulo m", wenn a und a' in der gleichen Restklasse liegen, also wenn $m \mid a - b$.

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 $a \in \overline{a}$ und $b \in \overline{b}$. Dann setzen wir als Wert der Verknüpfung die Restklasse von $a + b$, also

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

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 $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, so ist $a - a' = mk$ und $b - b' = ml$ mit $k, l \in \mathbb{Z}$.
Dann ist

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

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

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

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

Man mache sich in diesen Gleichungen zur Übung klar, welche Pluszeichen die Addition in $\mathbb{Z}$ und welche die Addition in $\mathbb{Z}/m\mathbb{Z}$ bezeichnen.

Das neutrale Element ist die Restklasse $\overline{0}$, denn

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

Das Inverse von $\overline{a}$ ist $\overline{-a} = \overline{m - a}$. 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

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

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

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

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

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

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

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

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

Definition 2.1.11. Sei G eine Gruppe. Wenn die zugrunde liegende Menge 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: Reflexivitiät 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 sinvvoll 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 bezichnen 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 $f : H \to Ha$ gegeben durch $h \mapsto ha$ ist eine Bijektion für jedes $a \in G$. Eine Umkehrfunktion $Ha \to H$ 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 $h \in H$. □

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

$$f : G \to H$$

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

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

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
$$\text{can}: \quad \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \quad ,$$
die jeder ganzen Zahl a ihre Äquivalenzklasse
$$\overline{a} = a + m\mathbb{Z} = a \bmod m$$
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: $can(a+b) = \overline{a+b} = \overline{a} + \overline{b}$.

3. Sei $G = (\mathbb{R}, +)$ und $H = (\mathbb{R}_{>0}, \cdot)$. Wähle festes ein $\alpha \in \mathbb{R}$ und betrachte die Exponentialabbildung
$$f_\alpha : \ \mathbb{R} \to \mathbb{R}_{>0}$$
mit $f_\alpha(x) = \mathrm{e}^{\alpha x}$. Dann gilt mit den Rechenregeln für Exponentialfunktionen:
$$f_\alpha(x+y) = \mathrm{e}^{\alpha(x+y)} = \mathrm{e}^{\alpha x}\mathrm{e}^{\alpha y} = f_\alpha(x) f_\alpha(y) \ .$$

4. Für eine beliebige Gruppe G und ein beliebiges Element $g \in G$ gibt es einen Homomorphismus $f_g : \mathbb{Z} \to G$ 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 $f : G \to H$ 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*
$$f^{-1}: \quad H \to G$$
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
$$f(g^{-1}) f(g) = f(g^{-1} g) = f(e_G) \underset{1.}{=} e_H \ .$$
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
$$\begin{aligned} f^{-1}(h) \cdot f^{-1}(h') &= (f^{-1} \circ f)\Big(f^{-1}(h) \cdot f^{-1}(h')\Big) \\ &= f^{-1}\left(f\Big(f^{-1}(h) \cdot f^{-1}(h')\Big)\right) \\ &\underset{(*)}{=} f^{-1}\left(f\Big(f^{-1}(h)\Big) \cdot f\Big(f^{-1}(h')\Big)\right) \\ &= f^{-1}(h \cdot h') \ . \end{aligned}$$
□

wobei wir für (*) verwendet haben, dass f ein Gruppenhomoromorphismus ist. $\langle\langle$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*inen wissen, dass $0 = 1 - 1$ ist".$\rangle\rangle$

Beispiele 2.2.4. Für den Gruppenhomomorphismus $\exp : \mathbb{R} \to R_{>0}$, $\exp(x) = e^x$, 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
$$\exp^{-1} =: \log \ : \ \mathbb{R}_{>0} \to \mathbb{R}$$
ist ein Gruppenhomomorphismus, d.h. es gilt
$$\log(xy) = \log(x) + \log(y) \quad \text{für} \quad x, y \in \mathbb{R}_{>0} \ .$$

Definition 2.2.5. Ein bijektiver Gruppenhomomorphismus heißt *Gruppenisomorphismus*. Zwei Gruppen G, H heißen *isomorph*, wenn es einen Gruppenisomorphismus $f : \ G \to H$ 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 $f : G \to H$ und $g : \ H \to K$ Gruppenisomorphismen, so ist auch $g \circ f : G \to K$ ein Gruppenisomorphismus (denn $g(f(a \cdot b))) = g(f(a)\dot{f}(b)) = gf(a) \cdot gf(b)$ und die Verkettung von Bijektionen ist bijektiv); daher ist die Relation transitiv.

$\langle\langle$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.$\rangle\rangle$

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 $\mathrm{e}^{\alpha x}$ 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 $\mathrm{Sym}\,(M) \cong \mathrm{Sym}\,(N)$. Zum Beispiel ist $\mathrm{Sym}\,(\{a, b, c\}) \cong \mathrm{Sym}\,(\{1, 2, 3\})$. 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 verketten wir auch die Abbildungen auf den Eckpunkten. Wir erhalten einen Homomorphismus $D_3 \to S_3$, der auch ein Isomorphismus ist.

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

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 $f : G \to H$ ein Gruppenhomomorphismus. Dann heißt das Urbild des neutralen Elements $e_H \in H$

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

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

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

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

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: $\operatorname{Im} f$ ist nicht leer, da $f(e_G) \in \operatorname{Im} f$ und für $f(a), f(b) \in \operatorname{Im} f$ gilt $f(a)f(b)^{-1} = f(ab^{-1}) \in \operatorname{Im} f$ mit Satz 2.2.3.

Bemerkung 2.2.9. Ein Wort noch zur (möglicherweise verwirrenden) Namensgebenug: Wir nennen allgemein die Funktion $\cdot : G \times G \to G$ 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 verschieden 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 $+, \cdot$

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

für die die folgenden Axiome gelten:

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

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

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

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

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

2. Auch die ganzes 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.

2. 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*.

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

$$a \cdot b =: ab$$

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

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

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

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

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

Beweis:. 1. 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.

2. 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 rellen 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 und für jedes $a + bi \neq 0$ ist

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

ein multiplikatives Inverses, denn wir rechnen

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

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*. .

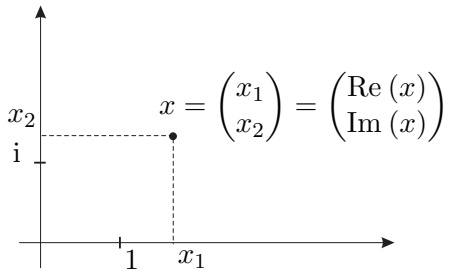
Die komplexe Zahl $i = 0 + 1 \cdot i$ heißt *imaginäre Einheit*.

Die injektive Abbildung

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

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 $x = x_1 + x_2 i \in \mathbb{C}$ nennen wir $x_1 = \mathrm{Re}\,(x) \in \mathbb{R}$ den *Realteil* und $x_2 = \mathrm{Im}\,(x) \in \mathbb{R}$ 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

$$\begin{array}{ll} \mathbb{C} & \to \mathbb{C} \\ x = x_1 + x_2 i & \mapsto \overline{x} = x_1 - i x_2 \end{array}$$

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 $z = z_1 + z_2 i \in \mathbb{C}$ 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.

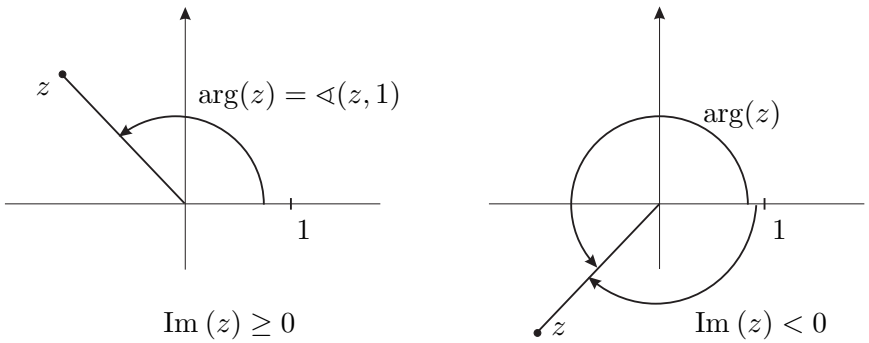
Bemerkungen 2.4.4. Seien $z, w \in \mathbb{C}$ beliebig.

1. Sofort aus der Definition folgt $\overline{1} = 1, \overline{0} = 0, \overline{\mathrm{i}} = -\mathrm{i}$ 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 $z \cdot \bar{z} = (z_1 + z_2 i)(z_1 - z_2 i) = z_1^2 + z_2^2 = |z|^2$.
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

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

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

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

heißt ein *Ring*, wenn gilt:

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

(R2) Die Multiplikation ist assoziativ.

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

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

2. Ein Element $1 \in R$ heißt *Einselement* 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
$$R := \{f : \ I \to \mathbb{R}\}$$
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 Elemente ist die konstante Funktion $x \mapsto 0$ und das multplikative 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 $\{f : M \to R\}$ 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
$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$
eine Multiplikation definieren. Sie ist wohldefiniert, denn für
$$a - a' = mk \qquad \text{und} \qquad b - b' = ml$$
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
$$\bar{a} \cdot (\bar{b} + \bar{c}) = \overline{a} \cdot \overline{b+c} = \overline{a \cdot (b+c)} = \overline{a \cdot b + a \cdot c} = \overline{a \cdot b} + \overline{a \cdot c} = \bar{a} \cdot \bar{b} + \bar{a} \cdot \bar{c}.$$
   Es liegt ein kommutativer Ring mit Eins $\overline{1}$ vor, den wir *Restklassenring* nennen.

   Wir rechnen zum Beispiel in $\mathbb{Z}/4\mathbb{Z}$:
$$\begin{aligned}&&\overline{2} \cdot \overline{2} = \overline{4} = \overline{0}\ .\\ \overline{2} \cdot \overline{1} = \overline{2} \qquad &&\overline{2} \cdot \overline{3} = \overline{6} = \overline{2}\ .\end{aligned}$$
   Die Menge $(\mathbb{Z}/m\mathbb{Z} \setminus \{\overline{0}\}, \cdot)$ ist also nicht immer eine Gruppe, denn für $m = 4$ hat die Restklasse $\overline{2}$ offenbar kein (multiplikatives) Inverses.

Der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ aus Beispiel 2.5.3.4 gibt uns unser letztes Beispiel für Körper, genau genommen eine unendliche Famile 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 $\overline{k}, \overline{l} \neq \overline{0}$, aber $\overline{0} = \overline{m} = \overline{k \cdot l} = \overline{k} \cdot \overline{l}$. Wegen Bemerkung 2.3.3.1 kann $\mathbb{Z}/m$ 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 $\ldots, -2k, -k, 0, k, 2k, 3k, \ldots$ (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 $\bar{k}$. □

Eine besondere Eigenschaft von $\mathbb{F}_p$ ist, dass die p-fache Summe der 1 mit sich selbst gleich 0 ist, also $1+1+\cdots+1 = 0$. 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 $m \cdot 1 = 1 + 1 + \cdots + 1 = 0$ 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

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

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 $e : \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}$, die jede Restklasse zu ihrem kleinsten nichtnegative 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 $\mathbb{C} \to \mathbb{C}$.

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 $\varphi : K \to E$ 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 $a_0 + a_1 t + a_2 t^2 + \dots a_n t^n$, 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 $f = a_0 + a_1 t + a_2 t^2 + \dots a_n t^n$ ist $-\infty$, falls $f = 0$, sonst gleich dem $\max_i \{i \mid a_i \neq 0\}$. Zum Beispiel gilt $\operatorname{grad}(t^3 + 2t + 1) = 3$.

2. Seien

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

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

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

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:

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

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 $K[t] \to K$.

So liefert jedes Polynom $f \in K[t]$ eine Funktion $K \to K$, 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 könen. 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

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

die den folgenden Axiomen genügen:

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

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

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

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

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

(V2d) $1 \cdot v = v$

2. Die Elemente eines Vektorraums heißen *Vektoren*. Das neutrale Element 0 von $(V, +)$ heißt *Nullvektor*. Im Zusammenhang mit K-Vektorräumen nennt man Elemente von K auch *Skalare*. Die Verknüpfung $\cdot$ heißt *Skalarmultiplikation*.

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 $\cdot : K \times V \to V$ durch komponentenweise Multiplikation

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

Dann ist $(K^n, +, \cdot)$ ein K–Vektorraum mit $0_{K^n} := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ 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 $\mathrm{Abb}(X, W)$ durch die Operationen auf den Funktionswerten zu einem K–Vektorraum.

   Genauer: Gegeben $f, g \in \mathrm{Abb}(X, W)$, 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 komponentweise 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:

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

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

4. Wir berechnen

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

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*

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

Sei $\cdot_W$ die Einschränkungen von

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

Dann ist $(W, +_W, \cdot_W)$ ein K–Vektorraum. 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 $K \times W \to W$ (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

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

ein Untervektorraum, ebenso

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

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

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 $W_{\Im} = \mathbb{R}\mathrm{i}$.

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 Famile von Untervektorräumen von V, d.h. für jedes $i \in I$ gibt es ein $W_i \leq V$. Dann ist auch $\cap_{i \in I} W_i = \{w \in V \mid \forall i \in I : w \in W_i\}$ ein Untervektorraum von V.*

Beweis:. Falls $I = \emptyset$ ist $\cap_\emptyset W_i = V$, 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 $0_W \in \cap_I W_i$.

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

(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

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

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

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

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

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

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

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

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

Beweis:. Wir prüfen:

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

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

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

Jedes W_i ist in $W_1 + \cdots + W_r$ enthalten, denn $w_1 = 0 + \cdots + w_i + \cdots + 0$ und 0 ist in jedem $W_{j\neq i}$ enhalten.

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

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 ⇒2. Wegen $V = W_1 + W_2$ lässt sich jedes $v \in V$ schreiben als $v = w_1 + w_2$ mit $w_i \in W_i$. Angenommen, es gäbe eine weitere Darstellung:

$$v = w_1 + w_2 = w_1' + w_2' \qquad w_i' \in W_i \ .$$

Daraus folgt

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

also $w_1 = w_1'$ und $w_2 = w_2'$.

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 $w = 0$. □

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 $\mathbb{C} = W_{\Im} + W_{\mathbb{R}}$ ist direkt da $W_{\Im} \cap W_{\mathbb{R}} = \{0\}$, also $\mathbb{C} = W_{\Im} \oplus W_{\Re}$.

Betrachten wir dagegen $W_{\Im} + W_{\Im} \subset \mathbb{C}$ dann gilt offensichtlich $W_{\Im} \cap W_{\Im} = W_{\Im}$ 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

$$f: \quad V \to W$$

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 $f : V \to W$ 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

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

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

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

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

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

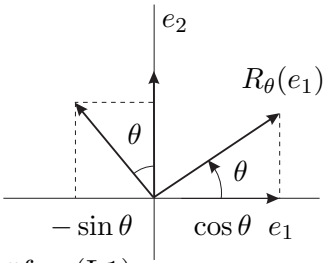
Beispiele 3.3.4. 1. Seien der Körper K und ein K-Vektorraum $V = W$ beliebig. Dann ist $f = \mathrm{id}_V$ eine lineare Abbildung. Allgemeiner ist für jedes $\lambda \in K$ die Abbildung $v \mapsto \lambda v$ K-linear.

Für allgemeine K-Vektorräume V, W ist $0 : V \to W$ 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

$$\begin{aligned} R_\theta : \quad & \mathbb{R}^2 \to \mathbb{R}^2 \\ R_\theta \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos\theta x - \sin\theta y \\ \sin\theta x + \cos\theta y \end{pmatrix} \end{aligned}$$

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



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:

$$R_\theta\left(\lambda\begin{pmatrix}x\\y\end{pmatrix}\right) = R_\theta\begin{pmatrix}\lambda x\\ \lambda y\end{pmatrix} = \begin{pmatrix}\cos\theta\lambda x - \sin\theta\lambda y\\ \sin\theta\lambda x + \cos\theta\lambda y\end{pmatrix}$$
$$= \lambda\begin{pmatrix}\cos\theta x - \sin\theta y\\ \sin\theta x + \cos\theta y\end{pmatrix} = \lambda R_\theta\begin{pmatrix}x\\y\end{pmatrix} .$$

3. Sei $V = W = K$; definiere eine lineare Abbildung $f : K \to K$ 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

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

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

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

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

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

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

1. *Ist $V' \subset V$ ein Untervektorraum, so ist das Bild $f(V')$ ein Untervektorraum von W. Insbesondere ist $f(V) = \operatorname{Im}(f)$ 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 U, V, W drei K–Vektorräume. Seien $f : U \to V$ und $g : V \to W$ zwei lineare Abbildungen, dann ist auch ihre Verknüpfung $g \circ f : U \to W$ eine lineare Abbildung.*

4. *Ist f linear und bijektiv, so ist die Umkehrabbildung $f^{-1} : W \to V$ 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

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

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

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

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

3. Wenn $f : U \to V$ und $g : V \to W$ (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

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

Betrachtung 3.3.6. Sei $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ und $f : V \to W$ 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 $f: \quad V \to W$ 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 $f: V \to W$ existiert.

Isomorphie ist eine Äquivalenzrelation[4]: 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 $\mathrm{Hom}_K(V, W)$. Wir schreiben $\mathrm{End}_K(V)$ für $\mathrm{Hom}_K(V, V)$.

Satz 3.3.9. $\mathrm{Hom}_K(V, W) \subset \mathrm{Abb}(V, W)$ *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 $\mathrm{Hom}_K(V, W)$, daher ist $\mathrm{Hom}_K(V, W) \neq \emptyset$. 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 Besipiel 3.1.2.9. Dies rechtfertigt unsere Notation*

Beweis:. Wir definieren eine Abbildung $a : V_1 \oplus V_2 \to V_1 \oplus' V_2$ 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 schreiebn 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.

[4]Wir 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*

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

Die K–Vektorraum–Struktur ist dabei komponentenweise definiert.

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

   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 $\oplus_{\mathbb{N}} K$. (Hier verkürzen wir die Schreibweise $\oplus_{i\in\mathbb{N}} K$, da die Summanden ja nicht von i abhängen.)

   Der Isomorphismus schickt $a_0+a_1t+a_2t^2+\dots+a_nt^n$ 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

$$\prod_{\lambda\in\Lambda} V_\lambda = \Big\{(v_\lambda)_{\lambda\in\Lambda} |\ v_\lambda \in V_\lambda\Big\}$$

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 Weiße, wie weit f davon entfernt ist, injektiv zu sein. Gibt es umgkehrt 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, das dies eine Äquivalenzerelation 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 $v\sim z$. □

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)*

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

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 $\ker(\pi) = U$. □

⟨⟨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' verschieden 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 $f : V \to W$ mit $f|_U = 0$ anzugeben oder eine K-lineare Abbildung $\tilde{f} : V/U \to W$.*

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

Beweis:. Gegeben $\tilde{f} : V/U \to W$ 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 f, dann erhalten wir aus $\tilde{f}$ die Abbildung $\tilde{f} \circ \pi : V \to W$ die v nach $\tilde{f}([v]) = f(v)$ schickt, also gleich f ist.

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

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

$$\bar{f} : V/\ker f \to \operatorname{Im} f$$

so dass gilt

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

wobei

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

die kanonische Surjektion und

$$\imath : \operatorname{Im} f \to W$$

die kanonische Einbettung von $\operatorname{Im} f$ *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*:

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ {\scriptstyle\pi}\downarrow & & \uparrow{\scriptstyle\imath} \\ V/\ker(f) & \underset{\exists!\bar{f}}{\xrightarrow{\sim}} & \operatorname{Im} f \end{array}$$

Dass das Diagramm "kommutiert", bedeutet einfach, dass $\imath \circ \bar{f} \circ \pi = f$, 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 Abbildugen verknüpfen und vergleichen wollen.⟩⟩

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

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

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

Nach Satz 3.4.3 gibt es nun $\tilde{f} : V/U \to W$ mit $f = \tilde{f} \circ \pi$ und per Definition faktorisiert $\tilde{f}$ als $\imath \circ \bar{f}$. Wegen $\operatorname{Im} \bar{f} = \operatorname{Im} f$ 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 $f : V \to W$ 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 $\operatorname{Im}(f) = W$ 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 $W \to V/\ker f$.

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

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

$$f : V \to W$$

mit $f(u,w) := w$ für alle $u \in U$ und $w \in W$. Dann ist $\ker f = \{(u,0)\} \cong U$ und $\mathrm{Im}\, f = W$. 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 $(a_i)_{i\in\mathbb{N}} \mapsto \lim_{i\to\infty} a_i$ 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 *endlichdimensionalen* 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 $v_1, \ldots, v_m \in V$ gegeben. Ein Element $w \in V$ der Form

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

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

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

$$\mathrm{span}_K(v_1, \ldots, v_m) := \{\lambda_1 v_1 + \ldots + \lambda_m v_m |\ \lambda_i \in K\}$$

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

3. Sei V ein K–Vektorraum und $M \subset V$ eine Teilmenge von V. Dann heißt

$$\mathrm{span}_K(M) := \{w = \lambda_1 v_1 + \ldots + \lambda_m v_m |\ \lambda_i \in K,\ v_i \in M,\ m \in \mathbb{N}\}$$

die *lineare Hülle* der M oder ihr *Erzeugnis* oder *Spann*. Falls M leer ist dann enthällt $\mathrm{span}_K(M)$ genau die leere Summe, also den Nullvektor, und es gilt $\mathrm{span}_K(M) = \{0\}$.

Also ist $\mathrm{span}_K(v_1, \ldots, v_m)$ nur eine andere Schreibweise für $\mathrm{span}_K(\{v_1, \ldots, v_m\})$.

Satz 4.1.2. *Sei V ein K–Vektorraum, $v_1, \ldots, v_m \in V$ und $M \subset V$ eine Teilmenge. Dann gilt:*

1. $\mathrm{span}_K(M)$ *ist ein Untervektorraum von V und es gilt $M \subset \mathrm{span}_K(M)$.*
2. *Ist $W \subset V$ ein Untervektorraum und $M \subset W$, so ist auch $\mathrm{span}_K(M) \subset W$.*
3. *Es gilt*

$$\mathrm{span}_K(M) = \bigcap_{M \subset W, W \leq V} W$$

hier wird der Schnitt über alle Untervektorräume von V betrachtet, die die Menge M enthalten.

Beweis:. 1. (UV1): Sei $m \in M \neq 0$

$$0 = 0m \in \mathrm{span}_K(M)$$

Falls M leer ist, dann ist 0 immer noch als leere Summe in $\mathrm{span}(M)$!

(UV2) Seien $x, y \in \mathrm{span}_K(M)$. Dann gibt es $\lambda_1, \ldots, \lambda_m \in K$, und $\mu_1, \ldots, \mu_k \in K$ sowie $v_1, \ldots, v_m \in M$ und $w_1, \ldots, w_k \in M$ so dass gilt

$$x = \lambda_1 v_1 + \ldots + \lambda_m v_m \qquad y = \mu_1 w_1 + \ldots + \mu_k w_K \, .$$

(Die v_i und w_i können gleich oder unterschiedlich sein.) Somit ist

$$x + y \in \mathrm{span}_K(M)$$

(UV3) folgt mit Notation wie in (UV2) sowie $\alpha \in K$:

$$\alpha x = (\alpha\lambda_1)v_1 + \ldots + (\alpha\lambda_m)v_m \in \mathrm{span}_K(M)$$

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 $\mathrm{span}_K(M)$.

3. $\mathrm{span}_K(M)$ 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

$$\bigcap_{M \subset W, W \leq V} W \subset \mathrm{span}_K(M).$$

Nach 3. ist aber $\mathrm{span}_K(M)$ in jedem der Untervektorräume, über die der Schnitt genommen wird, enthalten, also gilt auch die umgekehrte Inklusion

$$\mathrm{span}_K(M) \subset \bigcap_{M \subset W, W \leq V} W. \quad \square$$

Die letzte Aussage des Satzes können wir auch formulieren als: $\mathrm{span}_K(M)$ 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 $\mathrm{span}_K(0) = \{0\}$ der triviale Untervektorraum von K; für $v \neq 0$ ist $\mathrm{span}_K(v) = K$.

In einem beliebigen Vektorraum V mit einem Element $v \neq 0$ ist $\mathrm{span}_K(v) = Kv \cong K$, also eine Gerade.

2. Sei K ein beliebiger Körper und $V = K^n$. Setze wie in Betrachtung 3.3.6

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \ \ldots \ e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} ,$$

dann ist $\mathrm{span}_K(e_1, \ldots, e_n) = V$, denn für jedes $x \in K^n$ gilt

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \ldots + x_n e_n \in \mathrm{span}_K(e_1, \ldots, e_n)$$

3. Es gilt $W_1 + \ldots + W_r = \mathrm{span}_K(W_1 \cup \ldots \cup W_r)$.

   Jeder Vektor $v \in W_1 + W_2 + \ldots + W_r$ lässt sich als Summe $v = \sum_{i=1}^{r} w_i$ mit $w_i \in W_i$ schreiben und liegt daher in der linearen Hülle $\mathrm{span}_K(W_1 \cup \ldots \cup W_r)$. Damit ist die Inklusion "$\subset$" klar.

   Umgekehrt ist die Summe ein Untervektorraum, der $W_1 \cup \cdots \cup W_r$ 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 $\{v_1, \ldots, v_r\}$ von Vektoren aus V heißt *linear unabhängig*, falls gilt: sind $\lambda_1, \ldots, \lambda_r \in K$ und gilt

   $$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0 \ ,$$

   so folgt daraus $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$. Die Familie $\{v_1, \ldots, v_r\}$ heißt also genau dann linear unabhängig, wenn der Nullvektor sich nur trivial als Linearkombination von $v_1, \ldots, v_k$ darstellen lässt.

2. Eine beliebige 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

   $$0 = \lambda_1 v_1 + \ldots + \lambda_r v_r \ ,$$

   wobei nicht alle $\lambda_i \in K$ verschwinden.

$\langle\langle$Eine Menge ist nicht linear (un)abhängig *von* etwas, sondern einfach linear (un)abhängig. Gemeint ist: Die Elemente sind linear (un)abhängig voneinander.$\rangle\rangle$

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 $\{e_1, \ldots, e_n\}$ 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 $(v_1, \ldots, v_n)$ gilt: Wenn $\mu_1 v_1 + \cdots \mu_n v_n = 0$ dann ist $\mu_1 = \cdots = \mu_n = 0$.

   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 $\{v_1, \ldots, v_r\}$ von Vektoren eines K–Vektorraums sind die folgenden Bedingungen äquivalent:*

1. *Die Menge $\{v_i\}$ ist linear unabhängig.*

2. *Jeder Vektor $v \in \mathrm{span}_K(v_i)$ 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

$$v = \sum_i \lambda_i v_i = \sum_i \mu_i v_i \qquad \text{mit } \lambda_i, \mu_i \in K$$

folgt

$$0 = \sum_i (\lambda_i - \mu_i) v_i$$

Wegen der vorausgesetzten linearen Unabhängigkeit folgt $\lambda_i - \mu_i = 0$, also $\lambda_i = \mu_i$ für alle i. □

Definition 4.1.7. Eine Familie mit Indexmenge $\mathbb{N}$ oder $\underline{n} = \{1, 2, \ldots, n\}$ heißt *geordnete Familie*, wir sagen $v_i < v_j$ genau wenn $i < j$ in der

Lemma 4.1.8. *Eine geordnete Familie $(v_1, \ldots, v_n)$ von Vektoren in einem K-Vektorraum V definiert eine Abbildung $g : K^n \to V$. Die Abbildung g ist genau dann surjektiv, wenn $\mathrm{span}_K(v_1, \ldots, v_n) = V$, und genau dann injektiv ist, wenn $(v_1, \ldots, v_n)$ linear unabhängig ist.*

Beweis:. Um g zu definieren schicken wir einfach e_i nach v_i und damit ist $g(x) = g(\sum_{i=1}^n x_i e_i) = \sum_{i=1}^n x_i v$ eindeutig festgelegt. Die Abbildung ist linear.

Genau dann, wenn $\mathrm{span}_K(v_1, \ldots, v_n) = V$, 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 $(v_1, \ldots, v_n)$ 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 $x = 0$. □

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 $\mathrm{span}_K(M) = V$ 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 $(v_1, \ldots, v_n)$, die linear unabhängig ist und so dass die Menge $\{v_1, \ldots, v_n\}$ 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 $\{e_i\}_{i=1\ldots n}$ 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 $(v_1, v_2, \ldots, v_n)$. Die Standardbasis des K^n ist durch $(e_1, e_2, \ldots, e_n)$ eine geordnete Basis. Die Ordnung ist also eine *zusätzliche* Struktur, nämlich die Wahl einer Reihenfolge der Basisvektoren. (Eine Basis ist eine Teilmenge, da kommt es nicht auf eine Reihenfolge der Elemente an.)

4. $K = \mathbb{R}$ und $V = \mathbb{C}$, dann ist $M = \{1, \mathrm{i}\}$ eine $\mathbb{R}$–Basis von $\mathbb{C}$. Auch $\{1, -\mathrm{i}\}$ ist eine $\mathbb{R}$–Basis von $\mathbb{C}$, oder $\{1 + i, 1 - 2i\}$. Eine Basis von $\mathbb{C}$ als $\mathbb{C}$-Vekotrraum ist einfach von $\{1\}$ (oder jeder anderen komplexn 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 $\{1, X, \ldots, X^k, \ldots\}$ 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 $K^n \to V$ 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 Frage. 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.⇒ 1. Sei M minimales Erzeugendensystem. Angenommen, $M = \{v_i\}_{i \in I}$ ist linear abhängig. Wir finden also eine nicht-triviale Linearkombination

$$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$$

für die etwa $\lambda_r \neq 0$ ist. (Wir können die Vektoren immer umnummerieren, um diese Bedingung zu erfüllen.) Dann ist

$$v_r = -\frac{\lambda_1}{\lambda_r} v_1 - \ldots - \frac{\lambda_{r-1}}{\lambda_r} v_{r-1} \, .$$

(Beachten Sie, dass wir hier zum ersten Mal die Division in K benutzen!) Dann ist aber auch schon $(v_1, \ldots, v_{r-1})$ 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 $v \in V \setminus \mathrm{span}(M)$. 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 $v \in \mathrm{span}(M)$ 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 Ergzeugendensystem ist. Dann können wir $v = \sum_{w \in M \setminus \{v\}} \lambda_w w$ schreiben, und das widersprich der linearen Unabhängigkeit von M. □

Definition 4.2.5. Ein K–Vektorraum V heißt *endlich erzeugt*, falls V ein endliches Erzeugendsystem besitzt, d.h. falls es eine endliche Teilmenge $M = \{v_1, \ldots, v_m\}$ von V gibt, so dass $V = \mathrm{span}_K(M)$ gilt.

Lemma 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 $v_1, \ldots, v_n \in V$, so dass die Familie $(v_1, v_2, \ldots, v_n)$ linear unabhängig ist.*

Beweis:. Durch vollständige Induktion nach n.

- Induktionsanfang $n = 1$. Wähle $v_1 \in V$, $v_1 \neq 0$. Dies existiert, da sonst $V = \{0\}$ wäre und V somit endlich erzeugt wäre, nämlich von der leeren Menge $\emptyset$, vgl. Beispiel 4.1.3.1.

- Induktionsschritt:
  Sei $n \in \mathbb{N} \setminus \{0\}$ und seien $v_1, \ldots, v_n \in V$ linear unabhängig. Wähle $v_{n+1} \in V \setminus \mathrm{span}_K(v_1, \ldots, v_n)$. Solch ein v_{n+1} existiert, da andernfalls V von den $(v_1, \ldots, v_n)$, also endlich erzeugt wäre. Es bleibt zu zeigen, dass auch die Familie $(v_1, \ldots, v_n, v_{n+1})$ linear unabhängig ist. Sei

  $$0 = \sum_{j=1}^{n+1} \alpha_j v_j \qquad \alpha_j \in K$$

  eine Linearkombination des Nullvektors. Wäre $\alpha_{n+1} \neq 0$, so würden wir die Relation

  $$v_{n+1} = \sum_{j=1}^{n} \left(-\frac{\alpha_j}{\alpha_{n+1}} \right) v_j$$

  erhalten, die im im Widerspruch zu unserer Wahl von $v_{n+1} \notin \mathrm{span}_K(v_1, \ldots, v_n)$ steht. Also muss $\alpha_{n+1} = 0$ gelten; daraus folgt

  $$\sum_{j=1}^{n} \alpha_j v_j = 0$$

  und hieraus nach Induktionsannahme $\alpha_j = 0$ für alle $j = 1, \ldots, n$. □

Nicht jeder Vektorraum hat ein endliches Erzeugendensystem, zum Beispiel nicht der K–Vektorraum der Polynome $K[X]$ über einem beliebigen Körper K.

Satz 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 $M = \{v_1, \ldots, v_n\}$ 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 nach Satz 4.2.4.2 das Erzeugendensystem M nicht minimal. Es gibt also ein $v \in M$ mit $v \in \mathrm{span}_K(M \setminus \{v\})$ 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. □

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, $\mathcal{B} = \{v_1, \ldots, v_r\} \subset V$ eine Basis. Sei $w = \sum_{j=1}^{r} \alpha_j v_j \in V$ mit $\alpha_j \in K$. Dann gilt für jedes $k \in \{1, \ldots, r\}$ mit $\alpha_k \neq 0$:*

$$\mathcal{B}'_k := \{v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_r\}$$

ist eine Basis, d.h. das Basiselement v_k kann gegen w ausgetauscht werden.

Beweis:. • Nach Umnummerierung können wir $k = 1$ annehmen.

- Wir zeigen: $\mathcal{B}'$ ist ein Erzeugendensystem. Für jedes gegebene $v \in V$ existieren $\beta_j \in K$ für $j = 1, \ldots, r$, so dass

  $$v = \sum_{j=1}^{r} \beta_j v_j \qquad (*)$$

  gilt. Aus $\alpha_1 \neq 0$ folgt

  $$v_1 = \frac{1}{\alpha_1} w + \sum_{j=2}^{r} \Big(-\frac{\alpha_j}{\alpha_1}\Big) v_j \,.$$

  Dies setzen wir in (*) ein und erhalten

  $$v = \frac{\beta_1}{\alpha_1} w + \sum_{j=2}^{r} \Big(\beta_j - \alpha_j \frac{\beta_1}{\alpha_1}\Big) v_j \,.$$

  Somit ist $V \subset \mathrm{span}_K(\mathcal{B}')$, also ist $\mathcal{B}'$ 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:

  $$\beta\alpha_1 v_1 + \sum_{j=2}^{r} \Big(\beta\alpha_j + \beta_j\Big) v_j = 0$$

  Da die Familie $\{v_1, \ldots, v_r\}$ linear unabhängig ist, folgt

  $$\beta\alpha_1 = 0 \qquad \text{und} \qquad \beta\alpha_j + \beta_j = 0$$

  Aus $\alpha_1 \neq 0$ folgt $\beta = 0$ und daraus $\beta_j = 0$.

□

Satz 4.2.9 (*Austauschsatz*)**.** *Sei V ein K–Vektorraum, und $\mathcal{B} = \{v_1, \ldots, v_r\}$ eine Basis von V. Sei $\{w_1, \ldots, w_n\}$ eine linear unabhängige Teilmenge von V. Dann gilt $n \leq r$, und es gibt $i_1, \ldots, i_n \in \{1, \ldots, r\}$, so dass der Austausch*

$$\begin{aligned} &\text{von } v_{i_1} \text{ gegen } w_1, \ldots \\ &\text{von } v_{i_n} \text{ gegen } w_n \end{aligned}$$

eine neue Basis $\mathcal{B}^*$ von V liefert, die die vorgegebene linear unabhängige Menge $\{w_1, \ldots, w_n\}$ als Teilmenge enthält. Nach Umnummerierung zu $i_1 = 1, \ldots, i_n = n$ haben wir für die neue Basis

$$\mathcal{B}^* = \{w_1, \ldots, w_n, v_{n+1}, \ldots, v_r\} .$$

Beweis:. Vollständige Induktion nach n.

- Induktionsanfang: für $n = 0$ ist nichts zu zeigen. Sei die Aussage für $n - 1 \in \mathbb{N}$ gültig. Zu zeigen ist, dass die Aussage für n gültig ist.
  Sei also $\{w_1, \ldots, w_n\}$ linear unabhängig. Dann ist auch die Teilmenge $\{w_1, \ldots, w_{n-1}\}$ linear unabhängig. Nach Induktionsvoraussetzung gilt $n - 1 \leq r$ und (gegebenenfalls nach Umnummerierung) ist die Menge

  $$\bar{\mathcal{B}} := \{w_1, \ldots, w_{n-1}, v_n, \ldots, v_r\}$$

  eine Basis von V.

- Wir zeigen $n \leq r$. Nach Induktionsvoraussetzung ist $n - 1 \leq r$; wir müssen $n - 1 = r$ ausschließen. Dann wäre aber nach Induktionsvoraussetzung die Menge

  $$\bar{\mathcal{B}} := \{w_1, \ldots, w_{n-1}\}$$

  eine Basis von V, also eine maximale lineare unabhängige Teilmenge nachh Satz 4.2.4. Das steht im Widerspruch zur Voraussetzung, dass auch noch $\{w_1, \ldots, w_{n-1}, w_n\}$ linear unabhängig ist. Also ist $n - 1 < r$, also $n \leq r$.

- Zu zeigen ist, dass es ein $i_n \in \{n, \ldots, r\}$ gibt, so dass man v_{i_n} gegen w_n austauschen kann. Da $\bar{\mathcal{B}}$ eine Basis von V ist, finde mit $\alpha_k \in K$

  $$w_n = \sum_{j=1}^{n-1} \alpha_j w_j + \sum_{j=n}^{r} \alpha_j v_j .$$

  Wären alle $\alpha_n, \ldots, \alpha_r$ gleich Null, so wäre w_n Linearkombination der $\{w_1, \ldots, w_{n-1}\}$, im Widerspruch zur vorausgesetzten linearen Unabhängigkeit von $\{w_1, \ldots, w_n\}$. Also gibt es $i_n \in \{n, \ldots, r\}$ mit $\alpha_{i_n} \neq 0$. Wende nun das Austauschlemma 4.2.8 an und erhalte eine Basis $\mathcal{B}^* = \{w_1, \ldots, w_n, v_{n+1}, \ldots, v_r\}$.

  $\langle\langle$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 $\{w_1, \ldots, w_n\}$ linear unabhängig ist.$\rangle\rangle$ □

Korollar 4.2.10 (Basisergänzungssatz)**.** *Jede linear unabhängige Menge M in einem endlichdimensionalen 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 $\mathcal{B} = \{v_1, \ldots, v_r\}$ 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 $\{v_1, \ldots, v_r\}$ eine endliche Basis. Wäre eine weitere Basis $\mathcal{B}$ nicht endlich, gäbe es eine linear unabhängige Teilmenge $\{w_1, \ldots, w_{r+1}\} \subset \mathcal{B}$, im Widerspruch zum Austauschsatz 4.2.9.

2. Sind $\mathcal{B} = \{v_1, \ldots, v_r\}$ und $\mathcal{B}' = \{w_1, \ldots, w_k\}$ Basen von V, dann folgt aus dem Austauschsatz, da $\mathcal{B}'$ linear unabhängig und $\mathcal{B}$ Basis ist, $k \leq r$ und, indem man die Rollen von $\mathcal{B}'$ und $\mathcal{B}$ vertauscht, auch $r \leq k$, also $k = r$. □

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

$$\dim_K(V) \in \{0, 1, \ldots, \infty\}$$

heißt *Dimension* des Vektorraums V.

Beispiele 4.2.14. 1. Sei K ein beliebiger Körper und $V = K^n$. Dann hat die Standardbasis $e_1, \ldots, e_n$ die Länge n; und daher ist $\dim_K K^n = n$.

2. $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, denn $\{1, \mathrm{i}\}$ ist eine $\mathbb{R}$–Basis. 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 $\{1, X, X^2, \ldots\}$ 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 $v_1, \ldots, v_{n+1} \in W \subset V$, im Widerspruch zum Austauschsatz 4.2.9.

Also besitzt W eine endliche Basis $\mathcal{B} = \{w_1, \ldots, w_r\}$; 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 $\mathcal{B} = \{w_1, \ldots, w_n\}$ eine Basis von W. Gäbe es $v \in V \setminus \mathrm{span}_K(\mathcal{B})$, so ware $\mathcal{B} \cup \{v\}$ linear unabhängig, im Widerspruch zum Austauschsatz 4.2.9. □

4.3 Abbildungen und Dimension

Ein Ziel unseres Studiums von Vektorräumen ist die Klassifizierung von lineearen Abbildung. Als ersten Schritt können wir lineare Abbildungen mit Dimension in Zusammenhang setzen.

Definition 4.3.1. Sei $f: V \to W$ eine lineare Abbildung. Dann heißt

$$\operatorname{rg}(f) := \dim_K \operatorname{Im} f$$

der *Rang* der linearen Abbildung f.

Beispiel 4.3.2. Die Abbildung $f: \mathbb{R}^2 \to \mathbb{R}^2$ mit $e_1 \mapsto e_1 + e_2$ und $e_2 \mapsto 2e_1 + 2e_2$ hat Bild $\operatorname{span}_{\mathbb{R}}(e_1 + e_2)$ mit Dimension 1. Also ist $\operatorname{rg}(f) = 1$.

Bemerkung 4.3.3. Sei $f: V \to W$ 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 $\lambda_1, \ldots, \lambda_n \in K$ sowie $v_1, \ldots, v_n \in E$ finden, so dass

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

gilt, woraus

$$w = f(v) = \sum \lambda_i f(v_i)$$

folgt.

Der nächste Satz ist extrem nützlich, um Dimensionen zu berechnen.

Satz 4.3.4 (*Dimensionsformel*)**.** *Sei $f: V \to W$ eine lineare Abbildung und $\dim_K V < \infty$. Dann gilt die* Dimensionsformel

$$\dim_K V = \dim_K \ker f + \operatorname{rg}(f) .$$

Beweis:. $\ker f$ ist ein endlich–dimensionaler Untervektorraum von V; wähle eine Basis $\{v_1, \ldots, v_k\}$ von $\ker f$ mit $k := \dim_K \ker f$ und ergänze diese nach dem Basisergänzungssatz Korollar 4.2.10 zu einer Basis $\{v_1, \ldots, v_n\}$ von V mit $n = \dim_K V$. Nach Bemerkung 4.3.3 ist $\{f(v_1), \ldots, f(v_n)\}$ ein Erzeugendensystem des Bildes $f(V)$. Aber $0 = f(v_1) = \ldots = f(v_k)$, also ist schon die Teilmenge

$$\{f(v_{k+1}), \ldots, f(v_n))\}$$

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 $\lambda_{k+1}, \ldots, \lambda_n \in K$ und gelte

$$\sum_{j=k+1}^{n} \lambda_i f(v_i) = 0 .$$

Hieraus folgt $f\left(\sum_{i=k+1} \lambda_i v_i\right) = 0$, und somit $\sum_{i=k+1}^{n} \lambda_i v_i \in \ker f = \operatorname{span}_K(v_1, \ldots, v_k)$. Aber da $\{v_1, \ldots, v_n\}$ linear unabhängig ist gilt $\operatorname{span}_K(v_1, \ldots, v_k) \cap \operatorname{span}_K(v_{k+1}, \ldots, v_n) = \{0\}$. □

Korollar 4.3.5. *Sei $f: V \to W$ 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. □

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 $\operatorname{diff} : \mathbb{R}[x] \to \mathbb{R}[x]$ 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 $\mathcal{B} = (v_1, \ldots, v_n)$. 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 $\mathcal{B} = (v_1, \ldots, v_n)$ eine lineare Abbildung f von K^n nach V und diese Abbildung ist injektiv genau wenn $\mathcal{B}$ linear unabhänging 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. □

Wir können nun alle endlich erzeugten (äquivalent: endlichdimensionalen) Vektorräume 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 rechtertigt 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 $\mathcal{B} = (v_1, \ldots, v_n)$. Dann ist nach Definition $n = \dim_K(V)$. Laut Lemma 4.1.8 definiert dies einen Isomorphismus $f : K^n \to V$. (Wir erinnern uns: f ist surjektiv, da $\mathcal{B}$ Erzeugendensystem ist, f is 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 Dimensionformel 4.3.4 $\dim V = \dim \ker(f) + \operatorname{rg}(f) = 0 + \dim W$. □

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 $\{e_1, \ldots, e_m\}$ für V und $\{f_1, \ldots, f_n\}$ für W. Wir behaupten, dass $\{(e_1, 0), \ldots, (e_m, 0), (0, f_1), \ldots, (0, f_n))\}$ 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 $(\sum_i \lambda e_i, \sum_j \mu_j f_j) = \sum_i \lambda(e_i, 0) + \sum \mu_j(0, f_j)$ schreiben lässt. □

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

$$\begin{aligned} W_1 = \operatorname{span}(e_1, e_2): &\quad x\text{-}y\text{-Ebene} \\ W_2 = \operatorname{span}(e_1, e_3): &\quad x\text{-}z\text{-Ebene} \\ W_1 \cap W_2 = \operatorname{span}(e_1): &\quad x\text{-Achse} \\ W_1 + W_2 = \mathbb{R}^3 & \end{aligned}$$

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 $\sigma : W_1 \oplus W_2 \to W_1 + W_2$ an, die definiert ist durch $(w_1, w_2) \mapsto w_1 + w_2$.

Nach Definiton ist σ surjektiv, also ist $\operatorname{rg}(\sigma) = \dim_K(W_1 + W_2)$. 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 $\operatorname{rg}(\sigma) = \dim_K(W_1 \oplus W_2) - \dim \ker(\sigma)$:

$$\dim_K(W_1 + W_2) = \dim_K W_1 + \dim_K W_2 - \dim_K(W_1 \cap W_2). \qquad \square$$

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 $(e_1, \ldots, e_n)$ 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 $v_1, \ldots, v_r \in V$ und $w_1, \ldots, w_r \in W$.*

1. *Ist $(v_1, \ldots, v_r)$ eine geordnete Basis von V, so gibt es genau eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$. Diese hat die beiden Eigenschaften:*
   (a) *$f(V) = \operatorname{Im} f = \operatorname{span}_K(w_1, \ldots, w_r)$*
   (b) *Die Abbildung f ist genau dann injektiv, wenn die Familie $(w_1, \ldots, w_r)$ in W linear unabhängig ist.*
2. *Ist die Familie $(v_1, \ldots, v_r)$ in V linear unabhängig (aber nicht unbedingt eine Basis)), so gibt es mindestens eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, r$.*

Beweis:. 1. Wir haben in der Vorlesung einen direketen Beweis geführt.

Hier folgt ein Beweis, in dem wir Lemma 4.1.8 verwenden. Die geordnete Basis $(v_1, \ldots, v_r)$ definiert einen Isomorphismus $g : K^n \to V$ nach. Die geordnete Familie $(w_1, \ldots, w_r)$ definiert d urch $e_i \mapsto w_i$ definiert eine lineare Abbildung $h : K^n \to W$ und das Bild von h ist $\operatorname{span}_K(w_1, \ldots, w_r)$; h ist injektiv genau wenn die (w_i) lineaer unabhängig sind.

Aber dann ist $h \circ g^{-1}$ die gesuchte Abbildung f und da g ein Isomorphismus ist gilt $\operatorname{Im}(f) = \operatorname{Im}(h) = \operatorname{span}_K(w_1, \ldots, w_r)$ und f ist injektiv, wenn h injektiv ist, also wenn die (w_j) linear unabhängig sind.

Die Abbildung ist eindeutig: Sei f' ein andere Abbildung mit $f'(v_i) = w_i$, dann folgt $f'(v) = f'(\sum \lambda_i v_i) = \sum_i \lambda_i f'(v_i) = \sum \lambda_i w_i = f(v)$ da (v_i) eine Basis ist.

2. Ist die Familie $(v_1, \ldots, v_r)$ 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:

$$(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$$

und ein f durch Vorgabe beliebiger Werte $w_{r+1}, \ldots, w_n \in W$ für $v_{r+1}, \ldots, v_n$ wie in 2. festlegen. □

Wir betrachten als nächstes lineare Abbildungen

$$f : K^n \to K^m.$$

Damit könen 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 $f: K^n \to K^m$ 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 $K \to K^n$, die λ nach λv schickt.

Der Zeilenvektor $w = (w_1, \dots, w_n)$ repräsentiert die lineare Abbildung $K^n \to K$, 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 $= (v_1, \dots, v_n)^T = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$.

Wir verwenden folgende Konvention: Abbildungen werden mit Kleinbuchstaben bezeuchnet 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, das $M(f)$ 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

$$f(v) = f\Big(\sum_{i=1}^{n} v_i e_i\Big) = \sum_{i=1}^{n} v_i f(e_i) = v_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + v_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} v_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} v_j \end{pmatrix} .$$

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

$$A \cdot v = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^{n} a_{1j} v_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} v_j \end{pmatrix}$$

Beispiel 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 $f : K^n \to K^m$ und $g : K^m \to K^l$ 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

$$M(g) = (a_{ij}) , \; M(f) = (b_{ij}) , \; M(g \circ f) = (c_{ij})$$

die darstellenden Matrizen. Dann ist die j-te Spalte

$$\begin{pmatrix} c_{1j} \\ \vdots \\ c_{lj} \end{pmatrix} = g \circ f(e_j) = g\Big(f(e_j)\Big) = g \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix} = g\Big(\sum_{k=1}^{m} b_{kj} e'_k\Big)$$
$$= \sum_{k=1}^{m} b_{kj} g(e'_k) = \sum_{k=1}^{m} b_{kj} \begin{pmatrix} a_{1k} \\ \vdots \\ a_{lk} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} a_{1k} b_{kj} \\ \vdots \\ \sum_{k=1}^{m} a_{lk} b_{kj} \end{pmatrix} ,$$

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 $w = (w_1, \ldots, w_n)$ als $(1 \times n)$-Matrix und einen Spaltenvektor $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ multiplizieren, erhalten wir $w.v = (\sum w_i v_i)$, 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 $(M(h) \cdot M(g)) \cdot M(f) = M((h \circ g) \circ f) = M(h \circ (g \circ f)) = M(h) \cdot (M(g) \cdot M(f))$. □

Definition 5.1.11. Wir setzen

$$E_n = M(\mathrm{id}_{K^n}) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

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. □

Bemerkung 5.1.13. Im Allgemeinen ist die Verkettung linearer Abbildungen und somit die Matrizenmultiplikation nicht kommutativ. Zum Beispiel finden wir für

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$
$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

5.2 Mehr über Matrizen

Wir wollen noch etwas mehr mit Matrizen rechnen. Nach Satz 3.3.9 ist $\mathrm{Hom}_K(K^n, K^m)$ 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 *Spaltenvekor*, $v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} = (v_1, \ldots, v_m)^T$.

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. □

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

$$B^T = (b'_{kj}) \quad \text{mit} \quad b'_{kj} = b_{jk}$$
$$A^T = (a'_{ji}) \quad \text{mit} \quad a'_{ji} = b_{ij}$$

Hieraus folgt

$$B^T \cdot A^T = (d_{ki}) \quad \text{mit}$$
$$d_{ki} = \sum_j b'_{kj} \cdot a'_{ji} = \sum_j b_{jk} \cdot a_{ij} = c_{ik} = c'_{ki} \ .$$

□

Unter der Entsprechung

$$\mathrm{Hom}(K^n, K^n) \to M(n \times n, K)$$

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*

$$GL(n, K) := \{A \in M(n \times n, K) : \ A \textit{ invertierbar}\}$$

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 $A^{-1}, B^{-1} \in M(n \times n, K)$

  $$AA^{-1} = A^{-1}A = E_n \qquad \text{und} \quad BB^{-1} = B^{-1}B = E_n \ ,$$

  so ist wegen der Assoziativität der Matrizenmultiplikation Lemma 5.1.10

  $$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = E_n \qquad \text{und} \quad AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = E_n \ .$$

  Damit haben wir eine Verknüpfung $GL(n, K) \times GL(n, K) \to GL(n, K)$.

- 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*

$$\begin{aligned} V \quad & \text{mit geordneter Basis } \mathcal{A} = (v_1, \ldots, v_n) \\ W \quad & \text{mit geordneter Basis } \mathcal{B} = (w_1, \ldots, w_m) \ . \end{aligned}$$

Dann gibt es zu jeder linearen Abbildung

$$f : \ V \to W$$

genau eine Matrix $A = (a_{ij}) \in M(m \times n, K)$, so dass

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad (*)$$

gilt. Die so erhaltene Abbildung

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{A}} : \quad & \mathrm{Hom}(V, W) \to M(m \times n, K) \\ & f \qquad\qquad \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(f) \end{aligned}$$

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

$\langle\langle$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 vertauschit: Bei $(*)$ summieren wir über die Zeilen der Matrix, bei der Multiplikationsformul über die Spalten! $\rangle\rangle$

Beweis:. Da $\mathcal{B} = (w_1, \ldots, w_m)$ 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 $M_{\mathcal{B}}^{\mathcal{A}}$ eine K–lineare Abbildung.

Schließlich ist $M_{\mathcal{B}}^{\mathcal{A}}$ 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 $M_{\mathcal{B}}^{\mathcal{A}}(f) = A$. □

Korollar 5.2.8. *Seien V, W zwei K-Vektorräume. Es gilt*

$$\dim_K \operatorname{Hom}(V, W) = \dim_K V \cdot \dim_K W \ .$$

Seien $\mathcal{A}$ und $\mathcal{B}$ Basen wir oben. Dann ist eine explizite Basis gegeben durch die linearne Abbildungen

$$\begin{aligned} & f_{ij} : && V \to W \\ \textit{mit} \quad & f_{ij}(v_k) := \begin{cases} w_i & \textit{für } k = j \\ 0 & \textit{sonst.} \end{cases} \end{aligned}$$

für jedes $i = 1 \ldots n$ und $j = 1 \ldots m$.

Beweis:. Es ist $M_{\mathcal{B}}^{\mathcal{A}}(f_{ij}) = E_{ij}$, wobei die Familie ($E_{ij}$ wie in Lemma 5.2.2 eine Basis des K–Vektorraums $M(m \times n, K)$ bilden. Da $M_{\mathcal{B}}^{\mathcal{A}}$ ein Isomorphismus ist, bildet auch (f_{ij}) eine Basis von $\operatorname{Hom}(V, W)$. □

Die naheliegende Frage, wie die Matrix $M_{\mathcal{B}}^{\mathcal{A}}(F)$ 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 $\mathcal{A} = \mathcal{B} = (v_1, v_2, \ldots, v_n)$ zu wählen. Man schreibt dann

$$M_{\mathcal{B}} := M_{\mathcal{B}}^{\mathcal{B}} \ .$$

Der Vektorraumisomorphismus

$$M_{\mathcal{B}} : \operatorname{End}(V) \to M(n \times n, K)$$

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, $M_{\mathcal{B}}(\mathrm{id}_V) = E_n$.

Die Frage, wie durch Wahl einer geeigneten Basis die darstellende Matrix eines Endomorphismus auf eine Standardform gebracht werden kann, werden wir erst im nächsten Semester beantworten können.

5.3 Der Gaußsche Algorithmus und Elementarmatrizen

Die folgenden Überlegungen hatten im im speziellen Fall des Körpers $\mathbb{R}$ der reellen Zahlen schon einmal gesehen. Wir wiederholen Sie nun für beliebigie 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

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

mit $a_{ij} \in K$ und $b_i \in K$. Gesucht sind $x_1, \ldots, x_n \in K$.

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

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

4. Wir nennen

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \in M(m \times n, K)$$

die *Koeffizientenmatrix* des linearen Gleichungssystems. Mit $b := (b_1 \ldots, b_m) \in K^m$ nennen wir die Matrix

$$(A, \mid b) := \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{mn} & \ldots & a_{mn} & b_m \end{array} \right) \in M(m \times (n+1), K)$$

die *erweiterte Koeffizientenmatrix* des inhomogenen linearen Gleichungssystems.

5. Die *Lösungsmenge* des linearen Gleichungssystems ist nun

$$\mathrm{Lsg}(A, b) := \{x \in K^n \mid Ax = b\}$$

Für ein gegebenes lineares Gleichungssystem $Ax = b$ führen wir wie in Definition 5.1.2.2 die lineare Abbildung

$$\alpha : \ K^n \to K^m \qquad \text{mit darstellender Matrix } \ M(\alpha) = A$$

ein. (Hier betrachten wir die Standardbasis für K^n und K^m.)

Satz 5.3.2. *Wenn* $x_0 \in \mathrm{Lsg}(A, b)$*, dann gilt* $\mathrm{Lsg}(A, b) = x_0 + \ker(\alpha) = \{x_0 + v \mid v \in \ker(\alpha)\}$.

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 $x_1 \in \mathrm{Lsg}(A, b)$. 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

$$\dim \mathrm{Lsg}(A,b) = \dim(x_0 + \ker\alpha) = \dim_K \ker\alpha = n - \mathrm{rg}\,\alpha.$$

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 $i = 2, \ldots, m$ 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 $k = 1, \ldots, n$.

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

Lemma 5.3.4. *Sei A eine Matrix in Zeilenstufenform. Dann ist die Lösungsmenge eines linearen Gleichungssystems mit Koeffizientenmatrix A genau denn leer, wenn es einen Index $i \in \{1, \ldots, m\}$ gibt, so dass $a_{ij} = 0$ für alle j, aber $b_i \neq 0$ gilt.*

Beweis:. "$\Leftarrow$" 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 Gleichungssstems 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 stehenen Zeilen bestimmt sind und die übrigen frei gewählt werden können. □

Satz 5.3.5. *Sei $A \in M(m \times n, K)$ eine Matrix in Zeilenstufenform und sei r die Anzahhl der Zeilen die nicht 0 sind. Sei f die lineare Abbildung, die von A dargestellt wird. Dann ist* $\mathrm{rg}\,(f) = r$ *und* $\dim\ker(f) = m - r$.

Wir veranschaulichen die Situation in einem Beispiel:

$$A = \begin{pmatrix} 2 & 1 & -4 & -1 & 0 & -2 \\ 0 & 0 & 3 & 0 & -7 & 7 \\ 0 & 0 & 0 & -1 & 11 & -11 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}(5 \times 6, \mathbb{C})$$

Dann wird $\mathrm{Im}\,(\alpha)$ 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 Linearkomibantionen $e_1 = \frac{1}{2}v_1, e_2 = \frac{1}{3}(v_2 + 2v_1)$ 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 den 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 $\mathrm{Im}\,(\alpha)$ ein Unterraum von $\mathrm{span}(e_1, \ldots, e_r)$ ist.

Wir wollen nun zeigen, dass alle e_i mit $i \leq r$ in $\mathrm{Im}\,(\alpha)$ liegen. Es reicht zu zeigen, dass wir diese e_i als Linearkombination der Spalten von A erhalten. Wir verwenden Induktion nach i.

Wir wählen k minimal, so dass $a_{1k} \neq 0$ und dann liegt $a_{1k}e_1$ und damit e_1 im Bild. Nehmen wir nun an, dass $e_1, \dots e_{i-1}$ in Im (α) sind. Sei die k-te Spalte von A die erste, in der der i-te Eintrag ungleich 0 ist, also $a_{ik} \neq 0$. Wir ziehen von diesem Spaltenvektor nun passende Vielfache von $e_1, \dots, e_{i-1}$ ab und erhalten $a_{ki}e_i$ und damit e_i in Im (α).

Die zweite Aussage folgt sofort aus der Dimensionsformel 4.3.4. □

Um eine Matrix in Zeilenstufenform zu bringen verwenden wir wieder den Gaußalgorithmus. Wir werden ihn allerdings aus einer neuen Perspektive betrachten.

Lemma 5.3.6. *Ist $T \in GL(m, K)$, so ist*

$$\mathrm{Lsg}(A, b) = \mathrm{Lsg}(TA, Tb)$$

Beweis:. Wir haben eine Lösung $x \in \mathrm{Lsg}(A, b)$ genau dann, wenn $Ax = b$ gilt. Da T invertibel sein soll, gilt dies, genau dann, wenn $TAx = Tb$ gilt, was aber heißt, dass $x \in \mathrm{Lsg}(Ta, Tb)$ liegt. □

Die entscheidende Beobachtung ist nun, dass jede Operation im Gaußalgorithmus genau einer Multiplikation mit einem passenden Element in der allgemeinen linearen Gruppe $GL(m, K)$ entspricht!

Lemma 5.3.7. *Die folgenden Matrizen liegen stets in $GL(n, K)$; sie heißen* Elementarmatrizen.

- *Die Diagonalmatrix $\Delta(1, 1, \dots, 1, \lambda, 1, \dots, 1)$ mit $\lambda \in K \setminus \{0\}$ an i-ter Stelle*

$$\begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

  hat offenbar die Diagonalmatrix $\Delta(1, 1, \dots, \lambda^{-1}, 1, \dots, 1)$ als Inverses. Dann ist $TAx = Tb$ das lineare Gleichungssystem, bei dem die i–te Zeile mit $\lambda \neq 0$ multipliziert wurde.

- *Die Matrix $\tau(i, j)$ für $i \neq j$ mit Einträgen*

$$\tau(i, j)_{kl} = \begin{cases} \delta_{kl} & \text{für } k, l \notin \{i, j\} \\ 0 & \text{für } k = l = i \text{ und } k = l = j \\ 1 & \text{sonst} \end{cases}$$

$$\begin{pmatrix}
1 & & & & & & & & & \\
& 1 & & & & & & & & \\
& & \ddots & & & & & & & \\
& & & 1 & & & & & & \\
& & & & 0 & 0 & 0 & 1 & & \\
& & & & 0 & 1 & 0 & 0 & & \\
& & & & & & \ddots & & & \\
& & & & 0 & 0 & 1 & 0 & & \\
& & & & 1 & 0 & 0 & 0 & & \\
& & & & & & & & 1 & \\
& & & & & & & & & \ddots \\
& & & & & & & & & & 1
\end{pmatrix}$$

Wegen $\tau(i,j)\tau(i,j) = E_n$ ist $\tau(i,j)$ in $GL(n,K)$. Die Matrix $\tau(i,j)A$ unterscheidet sich von A durch die Vertauschung der i–ten und j–ten Zeile.

- *Schließlich*

$$\delta(i,j,\lambda) = \begin{pmatrix}
1 & & & & \\
& 1 & & & \\
& & \ddots & & \\
& \lambda & & 1 & 0 \\
0 & 0 & \dots & 0 & 1
\end{pmatrix} \leftarrow j\text{–te Zeile} \qquad = E_n + \lambda E_{ji}$$

$$\uparrow\ i\text{–te Spalte}$$

Die Matrix $\delta(i,j,\lambda)A$ für $i \neq j$ entsteht aus A, indem das λ–fache der i–ten Zeile zur j–ten Zeile addiert wird.

Wegen $\delta(i,j,\lambda)\delta(i,j,-\lambda) = E_n$ ist dies in $GL(n,K)$.

$\langle\langle$Die Bedingung $i \neq j$ für $\delta(i,j,\lambda)$ fehlte in der Vorlesung, ist aber essentiell!$\rangle\rangle$

Satz 5.3.8. *Es gibt die folgenden elementaren* Zeilen*umformungen:*

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

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

Beweis:. Da alle Elementarmatrizen in $GL(n,K)$ liegen, folgt dies sofort aus Lemma 5.3.6. □

Damit haben wir nun für beliebige Körper das aus Betrachtung 1.3.7 bekannte Rezept hergeleitet, eine Matrix in spezielle Zeilenstufenform zu bringen.

Beispiel 5.3.9. Wir betrachten die folgende Matrix in $M(3 \times 3, \mathbb{C})$:

$$A = \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

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

$$\begin{pmatrix} \sqrt{2} & -2 & 0 \\ 0 & 1 & 1-i \\ 0 & -1+i & i-1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{i+1} & 0 \\ 0 & 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & i+1 & 2 \\ \sqrt{2} & -2 & 0 \\ -\sqrt{2} & 1+i & i-1 \end{pmatrix}$$

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 $T_s \cdot \ldots T_1 \cdot A$ 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 Elemetarmatrix entspricht, gibt es Elementarmatrizen T_i, so dass $T_n \cdot \ldots T_1 \cdot A = E_n$ gilt. Lösen wir nach A auf, so haben wir A als Produkt von Elementarmatrizen geschrieben, $A = T_1^{-1} \ldots T_n^{-1}$. □

Aus $A = T_1^{-1} \ldots T_n^{-1}$ 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 Famile 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:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 4 & 0 & 1 & 0 \\ 1 & 4 & 9 & 0 & 0 & 1 \end{array}\right) \overset{L3-L1, L2-L1}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 3 & 8 & -1 & 0 & 1 \end{array}\right) \overset{L3-3L2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & -1 & 2 & -3 & 1 \end{array}\right)$$

$$\overset{-1 \cdot L3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array}\right) \overset{L2-3L3, L1-L3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 3 & -3 & 1 \\ 0 & 1 & 0 & 5 & -8 & 3 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array}\right)$$

$$\overset{L1-L2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 5 & -2 \\ 0 & 1 & 0 & 5 & -8 & 3 \\ 0 & 0 & 1 & -2 & 3 & -1 \end{array}\right)$$

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 $f : K^n \to K^n$ invertierbar ist, wenn sie surjektiv oder injektiv ist. Angenommen f hat ein links-inverses, $f^{-1} \circ f = \mathrm{id}_{K^n}$. Dann muss f injektiv sein, und ist invertierbar nach Korollar 4.3.5. Habe umgekehrt f ein Rechtsinverses, $f \circ f^{-1} = \mathrm{id}_{K^n}$. 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 $\mathcal{B} = (v_1, \ldots, v_n)$ von V einen Isomorphismus von K-Vektorräumen

$$\begin{array}{rccc} \Phi_{\mathcal{B}} & : & K^n & \to & V \\ & & e_i & \mapsto & v_i \ . \end{array}$$

vom Standardvektorraum K^n auf V, den wir nun in Abhängigkeit von $\mathcal{B}$ mit $\Phi_{\mathcal{B}}$ bezeichnen.

Lemma 5.4.1. *Sei $f: V \to W$ eine lineare Abbildung und seien geordnete Basen $\mathcal{A} = (v_1, \ldots, v_n)$ von V und $\mathcal{B} = (w_1, \ldots, w_m)$ von W gegeben. Für die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$ 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

$$\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}} : \; K^n \to K^m$$

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

$$\begin{aligned} f(v_i) = f\Big(\Phi_{\mathcal{A}}(e_i)\Big) &= \Phi_{\mathcal{B}}\Big(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}(e_i)\Big) \\ &= \Phi_{\mathcal{B}}\left(\sum_{j=1}^{m} M\Big(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}\Big)_{ji} e_j\right) \\ &= \sum_{j=1}^{m} M\left(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}\right)_{ji} w_j \,. \end{aligned}$$

□

Beispiel 5.4.2. Sei $f: \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Achse $\mathbb{R}\begin{pmatrix}1\\1\end{pmatrix}$, der Winkelhalbierenden des ersten und dritten Quadranten.

Setze $b_1 = \begin{pmatrix}1\\1\end{pmatrix}$ und $b_2 = \begin{pmatrix}-1\\1\end{pmatrix}$. Dann ist $\mathcal{B} = (b_1, b_2)$ eine geordnete Basis von $\mathbb{R}^2$. Wegen

$$f(b_1) = b_1 \qquad f(b_2) = -b_2$$

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 $\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_B : \mathbb{R}^2 \to \mathbb{R}^2$. 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*

$$f:\ V \to W \quad \textit{und}\ g:\ W \to Z$$

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 Matri sein. Also: $M_{\mathcal{D}}^{\mathcal{C}}(g) \cdot M_{\mathcal{B}}^{A}(f)$ ist nur sinnvoll, wenn $\mathcal{C} = \mathcal{B}$!

Beweis:.

$$\begin{aligned} M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) &\overset{5.4.1}{=} M(\Phi_{\mathcal{C}}^{-1} \circ g \circ f \circ \Phi_{\mathcal{A}}) \\ &= M\left(\left(\Phi_{\mathcal{C}}^{-1} \circ g \circ \Phi_{\mathcal{B}}\right) \circ \left(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}\right)\right) \\ &= M\left(\Phi_{\mathcal{C}}^{-1} \circ g \circ \Phi_{\mathcal{B}}\right) \cdot M\left(\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}}\right) \\ &\overset{5.4.1}{=} M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \end{aligned}$$

Hierbei haben wir im dritten Schritt verwendet, dass M die Komposition von Abbildungen in eine Matrizenmultiplikation überführt, vgl. Betrachtung 5.1.6. □

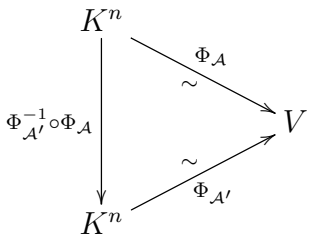
Jetzt können wir beantworten, wie die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{A}}$ sich bei Änderungen der Basen $\mathcal{A}$, $\mathcal{B}$ ändert.

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

$$T_{\mathcal{A}'}^{\mathcal{A}} := M_{\mathcal{A}'}^{\mathcal{A}}(\mathrm{id}_V) = M\left(\Phi_{\mathcal{A}'}^{-1} \circ \mathrm{id}_V \circ \Phi_{\mathcal{A}}\right) = M\left(\Phi_{\mathcal{A}'}^{-1} \circ \Phi_{\mathcal{A}}\right) \in GL(n, K)$$

Transformationsmatrix des Basiswechsels von $\mathcal{A}$ nach $\mathcal{A}'$.



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

$$T_{\mathcal{A}'}^{\mathcal{A}} \cdot T_{\mathcal{A}}^{\mathcal{A}'} = M_{\mathcal{A}'}^{\mathcal{A}}(\mathrm{id}_V) \cdot M_{\mathcal{A}}^{\mathcal{A}'}(\mathrm{id}_V) = M_{\mathcal{A}'}^{\mathcal{A}'}(\mathrm{id}_V \circ \mathrm{id}_V) = M_{\mathcal{A}'}^{\mathcal{A}'}(\mathrm{id}_V) = E_n\ .$$

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 $\mathcal{A} = (v_1, \ldots, v_n)$. 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 $\tilde{v} : \lambda \mapsto \lambda$ von K nach V betrachten dann ist $(v)_{\mathcal{A}}$ genau $M^{\mathcal{E}}_{\mathcal{A}}(\tilde{v})$, wobei wir $\mathcal{E}$ für die Standardbasis $\{(1)\}$ von K^1 schreiben.

Für eine zweite Basis $\mathcal{A}' = (v'_1, \ldots, v'_n)$ 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 $M^{\mathcal{E}}_{\mathcal{A}'}(\tilde{v}) = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V) \cdot M^{\mathcal{E}}_{\mathcal{A}}(\tilde{v})$ nach Satz 5.4.3.

Konkreter können wir überprüfen: die j-te Spalte von $T^{\mathcal{A}}_{\mathcal{A}'} = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)$ 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 $\mu_i = \sum c_{ij} \lambda_j$ ist. Insbesondere ist die Transformationsmatrix

$$T^{\mathcal{A}}_{\mathcal{A}'} = \begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \ldots & c_{nn} \end{pmatrix} .$$

3. Sei insbesondere $\mathcal{A}'$ die Standardbasis $\mathcal{E}$ und $\mathcal{A}$ eine beliebige Basis $(v_1, \ldots, v_n)$ und schreiben wir jeden Basisvektor v_i als $(v_{1i}, \ldots, v_{ni})^T$. Dann ist

$$T^{\mathcal{A}}_{\mathcal{E}} = \begin{pmatrix} | & \ldots & | \\ v_1 & \cdots & v_n \\ | & \ldots & | \end{pmatrix} = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nn} \end{pmatrix}$$

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 Standarbasis 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 Koffizient $\frac{x}{4}$, $x = x \cdot (1) = \frac{1}{4}x \cdot (4)$. 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 $M^{\mathcal{A}}_{\mathcal{B}}(f) = \frac{\mu}{4}$, denn $\mu(1) = \frac{\mu}{4}(4)$. Es ist auch $M^{\mathcal{B}}_{\mathcal{A}}(f) = 4\mu$, 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*

$$f: \quad V \to W$$

linear. Dann gilt

$$M_{\mathcal{B}'}^{\mathcal{A}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1} . \qquad (*)$$

2. Speziell für Endomorphismen $f: V \to V$ und zwei geordnete Basen $\mathcal{B}'$ und $\mathcal{B}$ von V ergibt sich

$$M_{\mathcal{B}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}(f) \cdot \left(T_{\mathcal{B}'}^{\mathcal{B}}\right)^{-1} . \qquad (**)$$

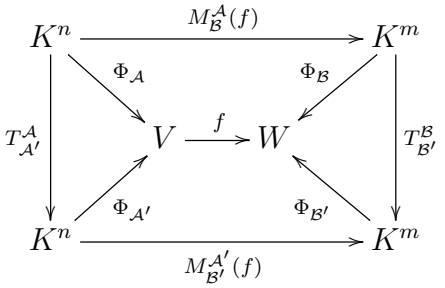
Wir können unsere Formel auch schreiben als

$$M_{\mathcal{B}'}^{\mathcal{A}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot T_{\mathcal{A}}^{\mathcal{A}'} .$$

.

$\langle\langle$ Seien Sie gewarnt: Es ist sehr leicht, die Rolle der beiden Basen beim Basiswechsel zu verwechseln! $\rangle\rangle$

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:



wobei hier Abbildungen zwischen Vektorräumen der Form K^n mit ihren darstellenden Matrizen identifiziert werden.

Beweis:. 1. Wir rechnen:

$$\begin{aligned} M_{\mathcal{B}'}^{\mathcal{A}'}(f) &= M_{\mathcal{B}'}^{\mathcal{A}'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) \\ &= M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_W) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot M_{\mathcal{A}}^{\mathcal{A}'}(\mathrm{id}_V) \qquad \text{[wegen Satz 5.4.3]} \\ &\overset{5.4.5.1}{=} T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot T_{\mathcal{A}}^{\mathcal{A}'} . \end{aligned}$$

2. ist als Spezialfall klar. □

Beispiel 5.4.7. Sei wie in Beipiel 5.4.2 $f: \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Achse $\mathbb{R}\begin{pmatrix}1\\1\end{pmatrix}$, der Winkelhalbierenden des ersten und dritten Quadranten. Setze $b_1 = \begin{pmatrix}1\\1\end{pmatrix}$ und $b_2 = \begin{pmatrix}-1\\1\end{pmatrix}$. Dann ist $\mathcal{B} = (b_1, b_2)$ eine geordnete Basis von $\mathbb{R}^2$. Wegen

$$f(b_1) = b_1 \qquad f(b_2) = -b_2$$

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

$$b_1 = e_1 + e_2 \qquad b_2 = -e_1 + e_2$$

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

$$Y = SXT^{-1} \qquad \text{gilt .}$$

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

$$Y = TXT^{-1} \qquad \text{gilt .}$$

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

$$Z = S_2(S_1XT_1^{-1})T_2^{-1} = (S_2S_1)X(T_2T_1)^{-1} .$$

$\langle\langle$ Diese Benennug 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. $\rangle\rangle$

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 $f : V \to W$ 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 $\mathrm{T} = T_{\mathcal{B}}^{\mathcal{E}_n}$.

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}. □

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, wobi die Nullen jeweils für Nullmatrizen verschiedener Größe stehen. Hier ist $r = \mathrm{rg}\,(f)$.

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 nennn 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

$$f: V \to W$$

eine lineare Abbildung. Wie im Beweis von Satz 4.3.4 ergänze eine geordnete Basis $(v_1, \ldots, v_k)$ des Untervektorraums $\ker f$ von V zu einer geordneten Basis $(v_1, \ldots, v_n)$ von V.

Nach Satz 5.1.1.1 ist dann

$$(f(v_{k+1}), \ldots, f(v_n)))$$

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 $\operatorname{span}_K(v_{k+1}, \ldots v_n) \subset V$ injektiv ist. Aber das gilt, da der Schnitt dieses Unterraums mit dem Kern von f trivial ist.

Ergänzt man die linear unabhängige Familie

$$(f(v_{k+1}), \ldots, f(v_n))$$

in W in irgendeiner Weise zu einer geordneten Basis

$$\mathcal{B} = (f(v_{k+1}), \ldots, f(v_n), w_1, \ldots, w_{m-r})$$

von W mit $m := \dim_K W$ und $r := \operatorname{rg} f$ und wählt bequemerweise als geordnete Basis von V die Familie

$$\mathcal{A} = (v_{k+1}, \ldots, v_n, v_1, \ldots, v_k)\ ,$$

so hat man für f die folgende Blockmatrix als darstellende Matrix:

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \begin{matrix} \Big\} m-r \\ \\ \underbrace{\qquad}_{n-r=k} \end{matrix}$$

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. □

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

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

anwenden und einen anderen Beweis des Homomorphiesatz 3.4.4 geben. Es gilt $\ker \pi = \ker f$; daher folgt, dass

$$\{[v_{k+1}], \ldots, [v_n]\}$$

eine Basis von $\operatorname{Im} \pi = V/\ker f$ ist.

Die Abbildung $\bar{f}: V/\ker f \to \operatorname{Im} f$ ist dann auf der Basis $\{[v_{k+1}], \ldots, [v_n]\}$ 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 Elemenntarmatrix beschieben werden.

Wir haben dann also $E_r = T_s \cdots T_1 A T_1' \cdots T_r'$ und die gewünschte Ähhnlickeit 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

$$\begin{array}{rr} \text{Spalten von } X \text{ heißt} & \textit{Spaltenrang } \operatorname{rg}(X) \\ \text{Zeilen von } X \text{ heißt} & \textit{Zeilenrang } \widetilde{\operatorname{rg}}(X) \end{array}$$

Als Beispiel betrachten wir mit $K = \mathbb{R}$ die linearen Abbildungen $f, g : \mathbb{R}^3 \to \mathbb{R}^2$ mit

$$M(f) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \qquad \operatorname{rg}(M(f)) = 1 \qquad \widetilde{\operatorname{rg}}(M(f)) = 1 \qquad \operatorname{rg}(f) = 1$$

$$M(g) = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 5 \end{pmatrix} \qquad \operatorname{rg}(M(g)) = 2 \qquad \widetilde{\operatorname{rg}}(M(g)) = 2 \qquad \operatorname{rg}(g) = 2$$

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 $\operatorname{rg}(f) = \operatorname{rg}(M) = \widetilde{\operatorname{rg}}(M)$, und unsere einfachen Beispiele sprechen schon einmal dafür.

Satz 5.5.8. *Wenn M darstellende Matrix für die lineare Abbildung $f : V \to W$ ist gilt* $\operatorname{rg}(f) = \operatorname{rg}(M)$.

Beweis:. Wir betrachten zurst den Fall, dass M eine linearen Abbildung $f : K^n \to K^m$ bezüglich der Standardbasen darstellt. Nach Definition 3.3.7.3) und Satz 5.1.1 (a):

$$\begin{aligned} \operatorname{rg}(f) &= \dim_K \operatorname{Im} f = \dim_K f(K^n) \\ &= \dim_K \operatorname{span}_K(f(e_1), \ldots, f(e_n)). \end{aligned}$$

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 $\Phi_{\mathcal{A}} : K^n \to V$ und $\Phi_{\mathcal{B}} : K^m \to W$ und es gilt $\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_A = M$ 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 $\operatorname{rg}(M) = \operatorname{rg}(g)$.

Aber da $\Phi_{\mathcal{B}}$ und $\Phi_{\mathcal{A}}$ Isomorphismen sind, ist der Rang von g gleich dem Rang von f und es gilt $\operatorname{rg}(M) = \operatorname{rg}(g) = \operatorname{rg}(f)$. □

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 $\operatorname{rg}(X) = \operatorname{rg}(f) = \operatorname{rg}(Y)$. □

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. $\operatorname{rg}(X) = \widetilde{\operatorname{rg}}(X^T)$ *und* $\widetilde{\operatorname{rg}}(X) = \operatorname{rg}(X^T)$.

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. □

Lemma 5.5.11. *Äquivalente Matrizen haben auch gleichen Zeilenrang. In Formeln: sind $X, Y \in M(m \times n, K)$ äquivalent, so ist*

$$\widetilde{\operatorname{rg}}(X) = \widetilde{\operatorname{rg}}(Y) .$$

Beweis:. Wir wissen also, dass es $S \in GL(m, K)$ und $T \in GL(n, K)$ gibt, so dass

$$Y = SXT^{-1}$$

gilt. Die Transposition dieser Matrixgleichung liefert

$$Y^T = (SXT^{-1})^T = (T^{-1})^T X^T S^T = (T^T)^{-1} X^T S^T$$

nach Lemma 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

$$\widetilde{\operatorname{rg}}(X) \overset{5.5.10.2}{=} \operatorname{rg}(X^T) \overset{5.5.9}{=} \operatorname{rg}(Y^T) \overset{5.5.10.2}{=} \widetilde{\operatorname{rg}}(Y) .$$

□

Satz 5.5.12. *Zeilenrang und Spaltenrang einer Matrix sind gleich: für $X \in M(m \times n, K)$ gilt*

$$\operatorname{rg}(X) = \widetilde{\operatorname{rg}}\ X .$$

Beweis:. Hat $X \in M(m \times n, K)$ Rang r, so ist X nach Bemerkung 5.5.2.4 äquivalent zu einer Matrix der Form $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Daher gilt

$$\operatorname{rg}(X) \underset{(5.5.9)}{=} \operatorname{rg} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r$$
$$\widetilde{\operatorname{rg}}(X) \underset{(5.5.11)}{=} \widetilde{\operatorname{rg}} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r$$

Hieraus folgt $\operatorname{rg}(X) = r = \widetilde{\operatorname{rg}}(X)$. □

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. □

Das Zornsche Lemma ist sogar äquivalent zum Auswahlaxiom, wir könnten in unserer Axiomatisierung also statt des Auswahlaxioms das Zornsche Lemm 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 $X(M, E) := \{A \,|\, \text{linear unabhängig}, M \subset A \subset E\}$. 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 Schrakne 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 $\lambda_1, \ldots, \lambda_n \in K$ und $v_1, \ldots v_n \in \cup\mathcal{K}$ mit $\sum_{j=1}^{n} \lambda_j v_j = 0$. Dann existieren $A_1 \ldots, A_n \in \mathcal{K}$ mit $v_j \in A_j$. Durch Umnummerieren können wir erreichen, dass $A_1 \subset A_2 \subset \ldots \subset A_n$ gilt. Daraus folgt $v_1, \ldots, v_n \in A_n$. Aus der linearen Unabhängigkeit von A_n ergibt sich $\lambda_1 = \ldots = \lambda_n = 0$. Also ist $\cup\mathcal{K}$ 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 $\lambda, \lambda_1, \ldots, \lambda_n \in K$, nicht alle Null und $b_1, \ldots, b_n \in B$, so dass $\lambda e + \sum_{j=1}^{n} \lambda_j b_j = 0$ gilt. Aus der linearen Unabhängigkeit von B folgt $\lambda \neq 0$. Damit gilt $e \in \mathrm{span}_K(B)$. Da dies für jedes $e \in E \setminus B$ gilt, folgt $E = B \cup (E \setminus B) \subset \mathrm{span}_K(B)$ und somit $V = \mathrm{span}(E) \subset \mathrm{span}(B)$. Also ist B eine Basis mit allen geforderten Eigenschaften. □

Wir haben insbesondere auch die folgenden Aussagen gezeigt, die wir für endlichdimensionale 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 $M = \emptyset$.

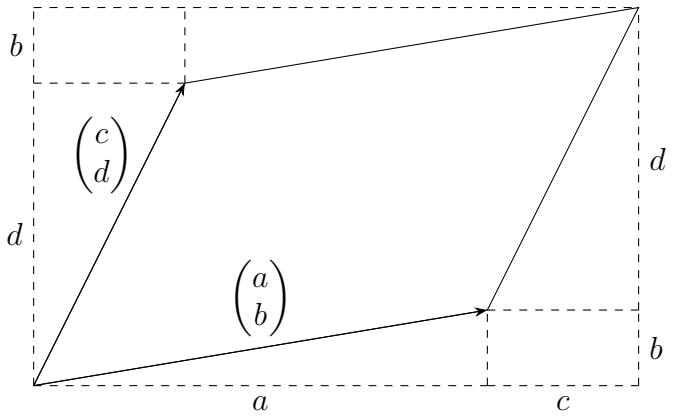
3. Basisergänzungssatz: Jede linear unabhängige Teilmenge $M \subset V$ lässt sich zu einer Basis von V ergänzen. Hier wählt man einfach $V = E$.

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

$$\mathsf{Area}(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}) = (a+c)(b+d) - \frac{1}{2}ab - \frac{1}{2}ab - \frac{1}{2}cd - \frac{1}{2}cd - bc - bc = ad - bc$$

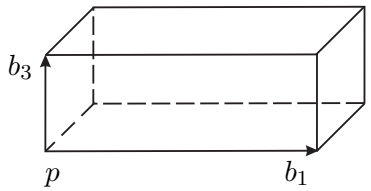
indem wir vom großen Rechteck vier Dreiecke und zwei kleiner Rechtecke abziehen.

Allgemeiner fragen wir uns nach dem Volumen des Parallelotops, das n Vektoren in $\mathbb{R}^n$ aufspannen.

Definition 7.1.1. Sei $(b_1, \ldots, b_n)$ eine Basis von $\mathbb{R}^n$. Dann heißt die Teilmenge

$$P = \{\alpha_1 b_1 + \cdots + \alpha_n b_n |\, 0 \leq \alpha_i \leq 1\} \subset \mathbb{R}^n$$

das von $b_1, \ldots b_3$ 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 $v = (v_1, \ldots, v_n)^T, w = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$. Dann ist das *Skalarprodukt* von v und w definiert als $v.w = \sum_i v_i w_i$.

Wir nennen $||v|| = \sqrt{v.v}$ die *(euklidische) Norm* des Vektors v, alltagssprachlich die *Länge*.

Zwei Vektoren v, w mit $v.w = 0$ heißen *orthogonal*.

Es gilt auch $v.w = v^T w$ und das Skalorprodukt (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 Skalarpodukt ist symmetrisch $v.w = w.v$ und das Vektorprodukt ist* antisymmetrisch*: $v \times w = -w \times v$. Insbesondere ist $v \times v = 0$.*

2. Skalarpodukt und Vektorprodukt sind bilinear, d.h. $(\lambda v + \mu w).u = \lambda v.u + \mu w.u$ und $(\lambda v + \mu w) \times u = \lambda v \times u + \mu w \times u$

3. Es gilt $(u \times v).w = u.(v \times w)$.

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 lineare 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 $(e_i \times e_j).e_k$. 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: $a.b = ||a|| \cdot ||b|| \cos(\angle(a, b))$. Aber was genau ist der Winkel zwischen zwei Vektoren, insbesondere in höheren Dimensionen? Wir *definieren* $\angle(a.b) = \cos^{-1}(\frac{a.b.}{||a|| \cdot ||b||})$. Dafür müssen wir wissen, dass $a.b \leq ||a|| \cdot ||b||$, 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 $||a||.||b|| \sin(\angle(a, b))$ und Orientierung entsprechend der "Rechte-Hand-Regel".

Es lässt sich nachrechnen, das 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 $a.(b \times c)$.

Lemma 7.1.7. *Das Spatprodukt erfüllt die folgenden Rechenregeln:*

1. $a.(b \times c)$ ist linear in jeder der drei Variablen.

2. $a.(b \times c)$ ist alternierend*: wenn zwei Vektoren gleich sind, dann ist das Spatprodukt 0.*

3. $e_1.(e_2 \times e_3) = 1$.

Beweis. 1. Dies folgt sofort aus Lemma 7.1.4.2.

2. Wir sehen aus Lemma 7.1.4.1, dass $a.(b \times b) = 0$.

   Es gilt $a.(a \times b) = 0$ denn mit Lemma 7.1.4.3 gilt $a.(a \times b) = (a \times a).b$.

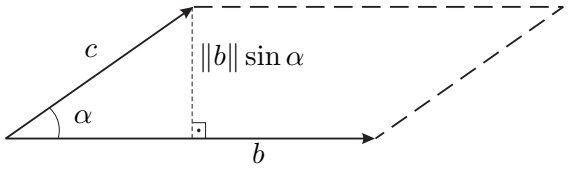
   Schließlich gilt $a.(b \times a) = -a.(a \times b)$ mit Lemma 7.1.4.1 und wir haben gerade gezeigt, dass dies 0 ist.

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 klassichen 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 $A \cdot h = a.(b \times c)$. 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 $a.(b \times c)$ 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 Polytop 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

$$M(n \times n, K) \to K$$

mit gewissen Eigenschaften verstehen. Dies ist die Determinantenabbildung:

Definition 7.2.1. Sei K ein Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\begin{aligned} \det : \quad M(n \times n, K) &\to K \\ A &\mapsto \det(A) \end{aligned}$$

heißt *Determinantenabbildung*, falls gilt:

(D1) det ist *linear* in jeder Zeile: es gilt

$$\det \begin{pmatrix} (a_1)^T \\ \vdots \\ (\lambda a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \lambda \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$

und

$$\det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i + a_i')^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i')^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$

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 Eigenschafen, 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
$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \mapsto a \cdot (b \times c) = (a \times b).c$$
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 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* $\det : M(n \times n, K) \to K$ *eine Determinantenabbildung. Dann gilt für alle $A, B \in M(n \times n, K)$ und alle $\lambda \in K$:*

1. $\det(\lambda A) = \lambda^n \det(A)$

2. *Ist eine Zeile von A gleich 0, so ist* $\det A = 0$.

3. *Entsteht B aus A durch Vertauschung zweier Zeilen, so ist* $\det B = -\det A$.

4. *Entsteht B aus A durch Addition des Vielfachen einer Zeile zu einer anderen, so ist* $\det B = \det A$.

5. *Sei A eine obere* Dreiecksmatrix *der Form*

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ 0 & & \lambda_n \end{pmatrix} .$$

Hierbei steht $*$ für beliebige Einträge oberhalb der Hauptdiagonalen. Dann ist

$$\det A = \lambda_1 \dots \lambda_n = \prod_{j=1}^{n} \lambda_j .$$

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 Zeitpuknt noch nicht, ob Determinantenabbildungen überhaupt existieren.

Beweis von Satz 7.2.3:. 1. Wir rechnen:

$$\det(\lambda A) = \det \left(\lambda \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_n)^T \end{pmatrix} \right) = \det \begin{pmatrix} \lambda(a_1)^T \\ \vdots \\ \lambda(a_n)^T \end{pmatrix} = \lambda \det \begin{pmatrix} (a_1)^T \\ \lambda(a_2)^T \\ \vdots \\ \lambda(a_n)^T \end{pmatrix}$$

$$= \lambda^2 \det \begin{pmatrix} (a_1)^T \\ (a_2)^T \\ \lambda(a_3)^T \\ \vdots \end{pmatrix} = \dots = \lambda^n \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_n)^T \end{pmatrix} .$$

2. Aus der Zeilenlinearität (D1) folgt sofort:

$$\det A = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ 0 \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ 0 \cdot 0 \\ \vdots \\ (a_n)^T \end{pmatrix} = 0 \cdot \det \begin{pmatrix} (a_1)^T \\ \vdots \\ 0 \\ \vdots \\ (a_n)^T \end{pmatrix} = 0 .$$

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):

$$0 = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i + a_j)^T \\ \vdots \\ (a_i + a_j)^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_n) \end{pmatrix} + \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_i)^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \det \begin{pmatrix} (a_1) \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_j)^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$
$$= 0 + \det A + \det B + 0 \, .$$

4. B entstehe aus A durch Addition des λ–fachen der j–ten Zeile zur i–ten Zeile, mit $i \neq j$. Dann ist

$$\det B = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_{i-1})^T \\ (a_i + \lambda a_j)^T \\ (a_{i+1})^T \\ \vdots \\ (a_n)^T \end{pmatrix} = \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_{i-1})^T \\ (a_i)^T \\ (a_{i+1})^T \\ \vdots \\ (a_n)^T \end{pmatrix} + \lambda \det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_{i-1})^T \\ (a_j)^T \\ (a_{i+1})^T \\ \vdots \\ (a_n)^T \end{pmatrix}$$
$$= \det A + \lambda \cdot 0 = \det A \, ,$$

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:

$$\det A = \det \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \cdot \det E_N \overset{\text{(D3)}}{=} \prod_{i=1}^{n} \lambda_i \, .$$

Sind nicht alle $\lambda_j \neq 0$, so sei i der größte Index, für den $\lambda_i = 0$ ist. Durch elementare Zeilenumformungen (Additon von vielfachen der Zeilen $i+1$ bis n) können wir die ganze i-te Zeile gleich 0 setzen. Es gilt

$$\det A = \det \begin{pmatrix} \lambda_1 & & & & & \\ 0 & \ddots & & & * & \\ & & \lambda_{i-1} & & & \\ 0 & 0 & 0 & 0 & \dots & 0 \\ & & & & \lambda_{i+1} & * \\ & 0 & & & & \ddots \end{pmatrix} = 0 = \prod_{j=1}^{n} \lambda_j \, .$$

□

Korollar 7.2.5. *Für jeden Körper K und jedes $n \geq 1$ gibt es* höchstens *eine Determinantenabbildung*

$$\det : M(n \times n, K) \to K .$$

Beweis:. Überführe eine Matrix A durch spezielle Zeilenumformung in eine obere Dreiecksmatrix A', wobei k Zeilenvertauschungen auftreten. Dann gilt

$$A' = \begin{pmatrix} \lambda_1 & & \\ & \ddots \quad * & \\ 0 & & \lambda_n \end{pmatrix}$$

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

$$A = \begin{pmatrix} 0 & 1 & \mathrm{i} \\ 1 & \mathrm{i} & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

Wir führen elementare Zeilenumformungen aus:

$$\begin{aligned} \det A &= -\det \begin{pmatrix} 1 & \mathrm{i} & 1 \\ 0 & 1 & \mathrm{i} \\ 2 & 3 & 4 \end{pmatrix} = -\det \begin{pmatrix} 1 & \mathrm{i} & 1 \\ 0 & 1 & \mathrm{i} \\ 0 & 3-2\mathrm{i} & 2 \end{pmatrix} \\ &= -\det \begin{pmatrix} 1 & \mathrm{i} & 0 \\ 0 & 1 & \mathrm{i} \\ 0 & 0 & 2-(3-2\mathrm{i})\mathrm{i} \end{pmatrix} = -(-3\mathrm{i}) = 3\mathrm{i} . \end{aligned}$$

7.3 Existenz der Determinante

Satz 7.3.1. *Für jeden Körper K und jedes $n \geq 1$ gibt es genau eine Determinantenabbildung*

$$\det_n : M(n \times n, K) \to K$$

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:

$$A_{ij}^{Str} := \left(\begin{array}{ccccc} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{array}\right) \in M\left((n-1) \times (n-1), K\right)$$

und definiere für beliebiges, aber festes $j \in \{1, \dots, n\}$

$$\det_n(A) := \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}\left(A_{ij}^{Str}\right) \ .$$

Wir müssen jetzt die Axiome (D1)-(D3) überprüfen.

(D1) Entstehe $\widetilde{A}$ aus A durch Multiplikation der k–ten Zeile mit $\lambda \in K$: $\tilde{a}_{kj} = \lambda a_{kj}$ und $\tilde{a}_{ij} = a_{ij}$ für $i \neq k$. Für die Streichungsmatrizen gilt:

- $\widetilde{A}_{kj}^{Str} = A_{kj}^{Str}$ da die einzige veränderte Zeile, nämlich die k-te Zeile, gestrichen wird.
- Für $i \neq k$ entsteht auch $\widetilde{A}_{ij}^{Str}$ aus A_{ij}^{Str} durch Multiplikation einer Zeile mit $\lambda \in K$, also gilt nach Induktionsannahme

$$\det_{n-1}\left(\widetilde{A}_{ij}^{Str}\right) = \lambda \det_{n-1}\left(A_{ij}^{Str}\right) \ .$$

Somit ist

$$\begin{aligned} \det_n \widetilde{A} &\stackrel{\text{def}}{=} \sum_{i \neq k} (-1)^{i+j} \widetilde{a}_{ij} \det_{n-1}\left(\widetilde{A}_{ij}^{Str}\right) + (-1)^{k+j} \widetilde{a}_{kj} \det_{n-1} \widetilde{A}_{kj}^{Str} \\ &= \sum_{i \neq k} (-1)^{i+j} a_{ij} \lambda \det_{n-1}\left(A_{ij}^{Str}\right) + (-1)^{k+j} \lambda a_{kj} \det_{n-1} A_{kj}^{Str} = \lambda \det_n(A) \ . \end{aligned}$$

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. □

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 $1 \leq j \leq n$.

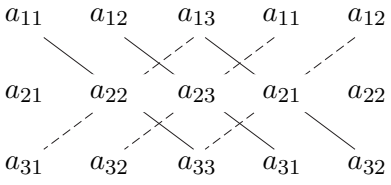
Die Vorzeichen im Laplace'schen Entwicklungssatz folgen einem Schachbrettmuster.

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.



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:

$$\det \begin{pmatrix} 0 & 1 & \mathrm{i} \\ 1 & \mathrm{i} & 1 \\ 2 & 3 & 4 \end{pmatrix} = 0 \cdot \det \begin{pmatrix} \mathrm{i} & 1 \\ 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & \mathrm{i} \\ 3 & 4 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$$
$$= -(4 - 3\mathrm{i}) + 2(1 - \mathrm{i}^2) = 3\mathrm{i} \; .$$

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* $a_1, \ldots, a_n$, $a'_j, a''_j \in K^n$ *und* $\lambda \in K$ *gilt*

$$\det(a_1 \ldots \lambda a_j \ldots a_n) = \lambda \det(a_1 \ldots a_j \ldots a_n)$$
$$\det(a_1 \ldots \; a'_j + a''_j \ldots a_n) = \det(a_1 \ldots a'_j \ldots a_n) + \det(a_1 \ldots a''_j \ldots a_n)$$

Beweis:. Aus der Entwicklung nach der j–ten Spalte folgt:

$$\det(a_1 \ldots \lambda a_j \ldots a_n) = \sum_{i=1}^{n} (-1)^{i+j} (\lambda a_{ij}) \det A_{ij}^{Str} = \lambda \det A \; .$$

Analog zeigt man die Additivität. □

Lemma 7.3.5. *Eine Determinantenabbildung schickt eine Matrix genau dann nach 0, wenn diese nicht maximalen Rang hat:* $\det A = 0 \Leftrightarrow \mathrm{rg}\ A < n$.

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 invertierbare 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 $\mathrm{rg}\,(A) < n$ 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. $\operatorname{rg}(A) = n$.
3. $\widetilde{\operatorname{rg}}(A) = \operatorname{rg}(A^T) = n$.
4. $\det(A) \neq 0$.

Beweis:. 1. $\Rightarrow$ 2.: Wenn A invertierbar ist stellt es eine surjektive Abbildung f dar und $\operatorname{rg}(f) = \operatorname{rg}(A) = n$.

2. $\Leftarrow$ 1.: Umgekehrt ist eine surjektive Abbildung $K^n \to K^n$ invertierbar (Korollar 4.3.5) und damit folgt aus $\operatorname{rg}(A) = n$, 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* $\det\left(A^T\right) = \det A$.

Beweis:. Wegen der Eindeutigkeit der Determinantenfunktion reicht es aus, zu zeigen, dass auch

$$\begin{aligned} \widetilde{\det}: \ M(n \times n, K) &\to K \\ A &\mapsto \det A^T \end{aligned}$$

eine Determinantenfunktion ist.

(D1) folgt aus der Spaltenlinearität in Satz 7.3.4.

(D2) Wenn A zwei gleiche Zeilen hat, so hat A^T zwei gleiche Spalten. Also ist $\operatorname{rg}(A^T) < n$, 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:

$$\begin{aligned}\det\begin{pmatrix}0 & 1 & \mathrm{i}\\ 1 & \mathrm{i} & 1\\ 2 & 3 & 4\end{pmatrix} &= 0\cdot\det\begin{pmatrix}\mathrm{i} & 1\\ 3 & 4\end{pmatrix} - \det\begin{pmatrix}1 & 1\\ 2 & 4\end{pmatrix} + \mathrm{i}\cdot\det\begin{pmatrix}1 & \mathrm{i}\\ 2 & 3\end{pmatrix}\\ &= -2 + \mathrm{i}(3-2\mathrm{i}) = 3\mathrm{i}\end{aligned}$$

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 $\operatorname{rg}(A) < n$ oder $\operatorname{rg}(B) < n$. 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 $\Delta(1, \ldots, \lambda, \ldots, 1)$ 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! Das folgt wieder aus Satz 7.2.3.

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. *Ähnliche Matrizen haben die gleiche Determinante.*

Beweis:. Gilt

$$\widetilde{A} = T\cdot A\cdot T^{-1}$$

mit $T \in GL(n, K)$, so ist nach dem Determinantenmultiplikationssatz 7.4.1

$$\det\widetilde{A} = \det T\cdot\det A\cdot\det T^{-1} = \det A\,.$$

$\square$

Korollar 7.4.3. *Für eine invertierbare Matrix* A *gilt gilt* $\det A^{-1} = (\det(A))^{-1}$.

Beweis:. Es gilt

$$1 = \det(E_n) = \det\left(A\cdot A^{-1}\right) \overset{7.4.1}{=} \det A\cdot\det A^{-1}\,. \quad \square$$

Damit erhalten wir sofort:

Korollar 7.4.4. *Es ist* det *ein Gruppenhomomorphismus von* $GL(n, K)$ *nach* $K\setminus\{0\}$.

Insbesondere sehen wir wieder, dass invertierbare Matrizen Determinante ungleich 0 haben.
Wir können nun sogar eine Formel für das Inverse einer Matrix angeben.

Satz 7.4.5. *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

$$AB = BA = \det(A) E_n \ .$$

Wenn $\det(A) \neq 0$ dann hat A ein Inverses

$$A^{-1} = \frac{1}{\det A} B \ .$$

Insbesondere sehen wir wieder, dass eine Matrix mit Determinante ungleich 0 invertierbar ist.

Die Matrix B heißt auch die *Adjunkte* von A, nicht zu verwechseln mit der *Adjungierten*, die wir später treffen.

Beispiel 7.4.6. Sei $n = 2$, 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 von Satz 7.4.5:.

- 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 $i \neq k$:

$$\sum_{j=1}^{n} a_{ij} b_{jk} = \sum_{j=1}^{n} a_{ij} (-1)^{j+k} \det \left(A_{kj}^{Str} \right) = \sum_{j=1}^{n} \widetilde{a}_{kj} (-1)^{j+k} \det \left(\widetilde{A}_{kj}^{Str} \right) \overset{7.3.8}{=} \det \widetilde{A} \ ,$$

  wobei $\widetilde{A}$ aus A entsteht, indem man die k-te Zeile durch die i–te Zeile ersetzt. Dies ändert nicht die Streichungsmatrix, da die k-te Zeile dort ohnehin gestrichen wird. Beim Matrixelement $\widetilde{a}_{kj}$ haben wir eine entsprechende Änderung des Index vorgenommen. Wegen $i \neq k$ hat $\widetilde{A}$ zwei Zeilen mit identischen Einträgen, also ist $\det \widetilde{A} = 0$.

- Aus dem Spaltenentwicklungssatz folgen die analogen Aussagen für das Produkt $B \cdot A$.

□

Satz 7.4.7. Cramersche Regel *Seien $a_1, \ldots, a_n, b \in K^n$ und sei die Matrix $A = (a_1, \ldots, a_n)$ invertierbar. Dann ist die eindeutige Lösung $x \in K^n$ des inhomogenen linearen Gleichungssystems von n Gleichungen für n Variablen*

$$Ax = b$$

gegeben durch

$$x_i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A}$$

Beweis:. Siehe Übungsaufgabe 14.3. □

Definition 7.4.8. Für einen Endomorphismus $\Phi : V \to V$ eines endlich–dimensionalen Vektorraums V heißt

$$\det(\Phi) := \det\left(M_{\mathcal{B}}^{\mathcal{B}}(\Phi)\right)$$

Determinante von Φ, wobei $\mathcal{B}$ eine beliebige geordnete Basis von V ist.

Korollar 7.4.9. *Die Determinante eines linearen Endomorphismus f ist wohldefiniert und es gilt*

1. $\det f \neq 0 \Leftrightarrow f$ *ist Automorphismus.*
2. $\det(f^{-1}) = \frac{1}{\det f}$ *für alle Automorphismen* Φ.
3. $\det(f \circ g) = \det(f)\det(g)$

Beweis:. Ist V ein n–dimensionaler K–Vektorraum und $f : V \to V$ ein Endomorphismus, und sind $\mathcal{B}$ und $\mathcal{B}'$ zwei geordnete Basen von V, so sind die beiden darstellenden Matrizen

$$M_{\mathcal{B}}^{\mathcal{B}}(f) \quad \text{und} \quad M_{\mathcal{B}'}^{\mathcal{B}'}(f)$$

ähnlich und haben somit die gleiche Determinante nach Korollar 7.4.2.

Die Aussagen 1.-3. folgen sofort aus den entsprechenden Aussagen für Matrizen. □

Beispiel 7.4.10. Sei $V = \mathbb{R}^2$. Für eine Drehung $\Phi = R_\theta$ um den Ursprung ist

$$\det R_\theta = \det(M(R_\theta)) = \det\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \cos^2\theta + \sin^2\theta = 1 .$$

Für eine Spiegelung $\Phi = S_\theta$ an einer Ursprungsgeraden (im Winkel θ zur x-Achse) ist

$$\det S_\theta = \det(M(S_\theta)) = \det\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} = -\cos^2 2\theta - \sin^2 2\theta = -1 .$$

7.5 Permutationen und Determinanten

Wir erinnern an Beispiel 2.1.2.5: die Menge aller bijektiven Selbstabbildungen der Menge $\underline{n} := \{1, 2, \ldots, n\}$ bildet eine Gruppe, die *symmetrische Gruppe* S_n. Sie hat $|S_n| = n! = 1 \cdot 2 \cdot \ldots \cdot n$ Elemente und ist für $n \geq 3$ nicht abelsch.

Für $\sigma \in S_n$ ist die Schreibweise

$$\sigma = \begin{bmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{bmatrix}$$

bequem. Das Produkt von $\sigma, \tau \in S_n$ schreiben wir als

$$\tau \cdot \sigma = \begin{bmatrix} 1 & \ldots & n \\ \tau(1) & \ldots & \tau(n) \end{bmatrix} \begin{bmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{bmatrix} = \begin{bmatrix} 1 & \ldots & n \\ \tau(\sigma(1)) & \ldots & \tau(\sigma(n)) \end{bmatrix}$$

Beispiel:

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} .$$

Wir können dies auch graphisch darstellen, vgl. Übung 5.5.

Definition 7.5.1. 1. Ein Element $\sigma \in S_n$ heißt auch *Permutation*.

2. Eine Permutation $\tau \in S_n$ heißt *Transposition*, falls τ zwei Elemente der Menge $\underline{n}$ vertauscht und alle übrigen Elemente fest lässt.

Lemma 7.5.2. *1. Für jede Transposition $\tau \in S_n$ gilt $\tau^{-1} = \tau$.*

2. Jede Permutation lässt sich als Produkt von Transpositionen schreiben, $\sigma = \tau_1 \dots \tau_k$.

Die Darstellung in 2. ist keinesfalls eindeutig!

Beweis:. 1. folgt sofort aus der Definition

2. lässt sich zum Beispiel induktiv zeigen: Jede Permutation, die k Elemente fest lässt, ist Produkt von Transpositionen. Dies ist klar für $k = n - 2$. Angenommen es gilt für k, dann zeigen wir es auch für $k - 1$, indem wir durch eine Transposition eine zusätzliches Element fixieren: Wenn $\tau(i) = j$ ist dann lässt $\tau_{ij} \circ \tau$ das Element i fest, sowie alle Elemente, die von τ fixiert werden. (Denn τ kann ja j nicht fixieren, sonst wäre $\tau(i) = \tau(j)$, was ein Widerspruch ist.) □

Lemma 7.5.3. *Für eine Permutation $\sigma \in S_n$ sei*

$$E_\sigma = \begin{pmatrix} (e_{\sigma^{-1}(1)})^T \\ \vdots \\ (e_{\sigma^{-1}(n)})^T \end{pmatrix} \in M(n \times n, K) \ .$$

Die Abbildung $\sigma \mapsto E_\sigma$ ist ein Gruppenhomomorphismus.

Beweis:. Es gilt per Definition

$$(E_\sigma)_{ij} = \delta_{i,\sigma(j)}$$

Es folgt für $\sigma, \tau \in S_n$

$$(E_\sigma E_\tau)_{ij} = \sum_{l=1}^{n} \delta_{i\sigma(l)}\delta_{l\tau(j)} = \delta_{i,\sigma(\tau(j))} = E_{\sigma\cdot\tau}$$

□

Korollar 7.5.4. *Die Abbildung*

$$\begin{aligned} \text{sign} : \ S_n &\to \{\pm 1\} \\ \sigma &\mapsto \det E_\sigma =: \text{sign}(\sigma) \end{aligned}$$

ist ein Gruppenhomomorphismus. Sie heißt Signumsabbildung.

Das Signum einer Transposition $\tau \in S_n$ ist gleich -1 wegen Satz 7.2.3, denn E_τ ist genau die Elementare Matrix, die zwei Zeilen vertauscht.

Die Abbildung sign bestimmt also eindeutig, ob eine Permutation Produkt von gerade oder ungerade vielen Transpositionen ist.

Beweis:. Als Verknüpfung von Homomorphismen ist sign $= \det \circ E$ ein Homomorphismus. Da E_σ mit Lemma 7.5.2 durch Vertauschungen von Zeilen aus der Einheitsmatrix $E_{\text{id}} = E_n$ hervorgeht, ist

$$\det(E_\sigma) \in \{1, -1\} \ .$$

□

Definition 7.5.5. Der Kern der Signumsabbildung ist die *alternierende Gruppe* A_n:

$$A_n := \{\sigma \in S_n | \operatorname{sign}(\sigma) = +1\}$$

(Für $n \geq 1$ hat diese Gruppe $\frac{1}{2}n!$ Elemente und ist für $n \geq 4$ nicht abelsch.)

Bemerkung 7.5.6. Sie werden in einer Übungsaufgabe für das Signum einer beliebigen Permutation $\sigma \in S_n$ die Formel

$$\operatorname{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} \qquad (*)$$

zeigen.

Satz 7.5.7. Leibniz'sche Regel *Sei K ein Körper und $A \in M(n \times n, K)$ mit $A = (a_{ij})$. Dann gilt*

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)} \ .$$

Beweis:. Wir schreiben den i–ten Zeilenvektor von A als

$$(a_i)^T = \sum_{j=1}^{n} a_{ij} (e_j)^T$$

und rechnen

$$\begin{aligned}
\det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_n)^T \end{pmatrix} &\underset{(D1)}{=} \sum_{i_1=1}^{n} a_{1i_1} \det \begin{pmatrix} (e_{i_1})^T \\ (a_2)^T \\ \vdots \\ (a_n)^T \end{pmatrix} \\
&\underset{(D1)}{=} \sum_{i_1, i_2=1}^{n} a_{1i_1} a_{2i_2} \det \begin{pmatrix} (e_{i_1})^T \\ (e_{i_2})^T \\ \vdots \\ (a_n)^T \end{pmatrix} \underset{(D1)}{=} \sum_{i_1 \dots i_n=1}^{n} a_{1i_1} \cdot \dots \cdot a_{ni_n} \det \begin{pmatrix} (e_{i_1})^T \\ (e_{i_2})^T \\ \vdots \\ (e_{i_n})^T \end{pmatrix}
\end{aligned}$$

Von den n^n Termen der Summe sind aber nur die $n!$ Terme nicht null, für die alle Indizes $i_1, i_2, \dots, i_n$ paarweise verschieden sind. Dann gibt es $\sigma \in S_n$ mit $\sigma(j) = i_j$. Also gilt

$$\begin{aligned}
\det \begin{pmatrix} (a_1)^T \\ \vdots \\ (a_n)^T \end{pmatrix} &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \det \begin{pmatrix} (e_{\sigma(1)})^T \\ \vdots \\ (e_{\sigma(n)})^T \end{pmatrix} \\
&= \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} E_{\sigma^{-1}} \\
&= \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \operatorname{sign}(\sigma) \ ,
\end{aligned}$$

wobei wir außer der Definition des Signums noch verwendet haben, dass $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$, was sofort aus $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)^{-1}$ folgt. □

Korollar 7.5.8. *Sei K ein Körper, $n_i \geq 1$ und $n := n_1 + n_2$. Sei $A \in M(n \times n, K)$ von der Gestalt*

$$A = \begin{pmatrix} A^{(1)} & \vdots & * \\ \dots & . & \dots \\ 0 & \vdots & A^{(2)} \end{pmatrix}$$

mit $A^{(i)} \in M(n_i \times n_i, K)$. Dann ist

$$\det {}_{n_1+n_2} A = \det_{n_1} A^{(1)} \det_{n_2} A^{(2)} \ .$$

Beweis:. Mit dem Ausdruck aus der Leibnizschen Regel 7.5.7 gilt nun

$$\det A = \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)} \text{sign}(\sigma) = \sum_{\sigma \in S'} a_{1\sigma(1)} \dots a_{n\sigma(n)} \text{sign}(\sigma)$$

wobei S' aus den Permutationen σ besteht, für die gilt

$$\sigma(i) \in \{1, \dots, n_1\} \Leftrightarrow i \in \{1, \dots, n_1\}$$

Denn jeder Index $i\sigma(i)$ im unteren linken Quadranten der Matrix führt zu einem Faktor 0.

Aber eine solche Permutation muss auch $\{n_1 + 1, \dots, n\}$ auf sich selbst abbilden.

Damit lässt sich σ schreiben als (σ_1, σ_2), wobei σ_1 die ersten n_1 Einträge permutiert und σ_2 die letzten n_2 Einträge. Die Kommaschreibweise bedeutet, dass wir ein Element von S_{n_1} und ein Element von S_{n_2} zu einem Element von S_n machen, indem σ_1 die ersten n_1 Einträge permutiert und n_2 die letzten n_2 Einträge. (Formal gesehen definieren wir einen Homomorphismus $S_{n_1} \times S_{n_2} \to S_n$.)

Es gilt $\text{sign}(\sigma_1, \sigma_2) = \text{sign}(\sigma_1)\text{sign}(\sigma_2)$ Damit folgt

$$\begin{aligned} \det A &= \sum_{(\sigma_1,\sigma_2)} \text{sign}((\sigma_1, \sigma_2)) a^{(1)}_{1\sigma^{(1)}(1)} \dots a^{(1)}_{n_1\sigma^{(1)}(n_1)} \cdot a^{(2)}_{1\sigma^{(2)}(1)} \dots a^{(2)}_{n_2\sigma^{(2)}(n_2)} \\ &= \sum_{\sigma_1 \in S_{n_1}} \text{sign}\left(\sigma^{(1)}\right) a^{(1)}_{1\sigma^{(1)}(1)} \dots a^{(1)}_{n_1\sigma^{(1)}(n_1)} \cdot \sum_{\sigma_2 \in S_{n_2}} \text{sign}\left(\sigma^{(2)}\right) a^{(2)}_{1\sigma^{(2)}(1)} \dots a^{(2)}_{n_2\sigma^{(2)}(n_2)} \\ &= \det A^{(1)} \cdot \det A^{(2)} \ . \end{aligned}$$

Wobei wir in der zweiten Gleichung einfach ausmultiplizieren. □

Bemerkung 7.5.9. Wir wollen den Rechenaufwand für die Berechnung der Determinante einer $n \times n$-Matrix an Hand der Zahl der nötigen Multiplikationen einmal grob überschlagen. Bei der Leibnizschen Regel 7.5.7 müssen wir $n!$ Multiplikationen ausfüllen, und jede dieser Multiplikationen hat n Faktoren, also sind wir bei $(n-1)n!$ einfachen Multiplikationen. (Die Bestimmung von sign ist rechnerisch auch recht teuer, aber diese Kosten lassen sicht mit geschickter Enumerierung der Permutationen vermeiden.)

Bei den Entwicklungssätzen 7.3.2 bzw. 7.3.8 sehen wir induktiv, dass wir $n!$ Multiplikationen ausführen müssen.

Beim Gauß-Algorithmus braucht man für die Elimination unterhalb der i-ten Zeile für jede Zeile eine Division zur Berechnung des Eliminationsfaktors und dann $n - i$ Multiplikationen in der Zeile, die verändert wird. Bei $n - i$ Zeilen ist der Aufwand

$$(n-i)(n-i+c) = (n-i)^2 + c(n-i)$$

Multiplikationen, wobei wir annehmen, dass eine Division so teuer ist wie c Multiplikationen. Der genaue Faktor hängt von den Details der Implementierung der Rechnerarchitektur ab, wir

können aber davon ausgehen, dass c konstant ist, und nicht mit n wächst. Der Gesamtaufwand beträgt dann

$$\sum_{i=1}^{n-1} \big((n-i)^2 + c(n-i)\big) = \sum_{j=1}^{n-1} (j^2 + cj) \sim \frac{n^3}{3} + c\frac{n^2}{2} \sim \frac{n^3}{3}$$

Multiplikationen, und damit deutlich kleiner als $n!$.

7.6 Orientierungen

Wir wollen (im reellen Fall) das Vorzeichen der Determinante weiter betrachten. Dazu arbeiten wir in diesem Unterkapitel über dem Körper $\mathbb{R}$ der reellen Zahlen, und benutzen, dass es hier eine totale Ordnung gibt.

Definition 7.6.1. Sei V ein endlich-dimensionaler reeller Vektorraum. (In der Folge sei stets $\dim_{\mathbb{R}} V \geq 1$ angenommen.) Zwei geordnete Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$ von V heißen *gleich orientiert*, falls für die Transformationsmatrix

$$\det T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} > 0$$

gilt. Andernfalls heißen die geordneten Basen *entgegengesetzt orientiert*.

Bemerkung 7.6.2. Über $K = \mathbb{C}$ oder $\mathbb{F}_p$ gibt es keine natürliche Aufteilung der Elemente von $K = K \setminus \{0\}$ in positive und negative und entsprechend keine Orientierung.

Lemma 7.6.3. *Die Beziehung "gleich orientiert" liefert eine Äquivalenzrelation auf der Menge aller geordneten Basen eines gegebenen Vektorraums V.*

Beweis:.

- Reflexivität:
$$\det T_{\mathcal{B}}^{\mathcal{B}} = \det E_n = 1 > 0\,.$$

- Symmetrie: für zwei geordnete Basen $\mathcal{B}, \widetilde{\mathcal{B}}$ von V gilt
$$T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} \cdot T_{\mathcal{B}}^{\widetilde{\mathcal{B}}} = E_n$$
Insbesondere folgt
$$\det T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} \cdot \det T_{\mathcal{B}}^{\widetilde{\mathcal{B}}} = 1\,,$$
so dass beide Determinanten ungleich Null sind und gleiches Vorzeichen haben.

- Transitivität folgt aus Satz 5.4.3
$$\det T_{\mathcal{B}_3}^{\mathcal{B}_1} = \det\, \big(T_{\mathcal{B}_3}^{\mathcal{B}_2} T_{\mathcal{B}_2}^{\mathcal{B}_1}\big) = \det\, \big(T_{\mathcal{B}_3}^{\mathcal{B}_2}\big) \det\, \big(T_{\mathcal{B}_2}^{\mathcal{B}_1}\big) > 0\,,$$
wenn die geordenten Basen $\mathcal{B}_1$, $\mathcal{B}_2$ und $\mathcal{B}_2$, $\mathcal{B}_3$ von V jeweils paarweise gleich orientiert sind.

□

Definition 7.6.4. Sei V ein endlich dimensionaler reeller Vektorraum. Eine Äquivalenzklasse von geordneten Basen bezüglich der Äquivalenzrelation 7.6.3 heißt eine *Orientierung* von V.

Lemma 7.6.5. *Ein endlich–dimensionaler reeller Vektorraum besitzt genau zwei Orientierungen.*

Beweis:. • Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine geordnete Basis von V, setze $\widetilde{\mathcal{B}} = (-b_1, \ldots, b_n)$. Wegen

$$\det T_{\widetilde{\mathcal{B}}}^{\mathcal{B}} = \det \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & & & 1 \end{pmatrix} = -1$$

definieren die geordneten Basen $\mathcal{B}$ und $\widetilde{\mathcal{B}}$ unterschiedliche Orientierungen. Es gibt also mindestens zwei unterschiedliche Orientierungen auf V.

- Es mögen die geordneten Basen $\mathcal{B}_1, \mathcal{B}_2$ und $\mathcal{B}_2, \mathcal{B}_3$ unterschiedliche Orientierungen besitzen:

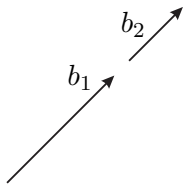
$$\det T_{\mathcal{B}_2}^{\mathcal{B}_1} < 0 \quad \text{und} \quad \det T_{\mathcal{B}_3}^{\mathcal{B}_2} < 0 \, .$$

Dann folgt, wiederum wegen Satz 5.4.3,

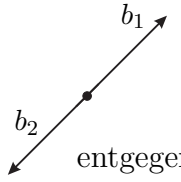
$$\det T_{\mathcal{B}_3}^{\mathcal{B}_1} = \det T_{\mathcal{B}_3}^{\mathcal{B}_2} \cdot \det T_{\mathcal{B}_2}^{\mathcal{B}_1} > 0 \, ,$$

also haben $\mathcal{B}_1$ und $\mathcal{B}_3$ gleiche Orientierung. □

Beispiele 7.6.6. 1. $\dim V = 1$

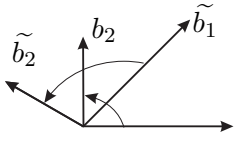


gleich orientiert



entgegengesetzt orientiert.

2. $\dim V = 2$: Hier kommt es auf den "Drehsinn der Basis" an: führt man den ersten Basisvektor durch die Drehung um einen betragsmäßig möglichst kleinen Winkel in den zweiten Basisvektor über, so sind die Basen genau dann gleich orientiert, wenn der Drehsinn dieser Drehungen gleich ist. Zum Beispiel haben



die gleiche Orientierung.

3. $\dim V = 3$: Analog ist hier die "Händigkeit der Basis" entscheidend. Eine Orientireung ist gegeben, wenn b_1, b_2, b_3 sich wie Daumen, Zeigefinger und Mittelfinger der gespreizten rechten Hand verhalten, die andere, wenn sie sich wie Daumen, Zeigefingur und Mittelfinger der linken Hand verhalten. (Es ist hier die anatomisch natürliche Spreizung der Finger gemeint, mit der drei zueinander senkrechte Richtungen bestimmt werden.)

Einen Automorphismus $\Phi : V \to V$ eines endlich–dimensionalen reellen Vektorraums nennen wir *orientierungserhaltend* oder *orientierungstreu*, falls $\det(\Phi) > 0$ gilt. Die orientierungstreuen Automorphismen von V bilden wegen des Determinantenmultiplikationssatze 7.4.1 eine Untergruppe der Automorphismengruppe von V.

7.7 Minoren

Wir wissen schon aus Lemma 7.3.5, dass eine quadratische Matrix von maximalem Rang dadurch charakterisiert werden kann, dass ihre Determinante nicht verschwindet. Man kann auch einen nicht-maximalen Rang durch die Berechnung von (mehreren) Determinanten bestimmen. Hierfür betrachten wir nicht nur quadratische Matrizen.

Definition 7.7.1. Ist $A \in M(m \times n, K)$ und $k \leq \min(m, n)$, so heißt eine $k \times k$-Matrix A', die durch Streichen von $m-k$ Zeilen und $n-k$ Spalten aus A hervorgeht, eine k-reihige *Untermatrix* von A. Ihre Determinante $\det A' \in K$ heißt ein k-reihiger *Minor* der Matrix A.

Bemerkung 7.7.2. Die $(n-1)$-reihigen Minoren einer quadratischen Matrix A wurden für die Spalten- und Zeilenentwicklung verwendet und bilden zusammen die Adjunkte, die bei der Bestimmung des Inversen in Satz 7.4.5 gebraucht wurde.

Satz 7.7.3. *Sei $A \in M(m \times n, K)$ und $r \in \mathbb{N}^*$. Dann sind die folgenden Bedingungen äquivalent:*

(i) $r = \operatorname{rg}(A)$.

(ii) Es gibt einen r-reihigen Minor ungleich Null, und für $k > r$ ist jeder k-reihige Minor gleich Null.

Der Satz schließt den Fall $0 = \operatorname{rg}(A)$ aus, aber das ist nicht besonders interessant, denn das ist nur der Fall, wenn A die Nullmatrix ist.

Beweis:. Wir zeigen zum Beweis, dass die folgenden beiden Bedingungen äquivalent sind:

(a) $\operatorname{rg}(A) \geq k$

(b) Es gibt eine k-reihige Untermatrix A' von A mit $\det A' \neq 0$.

Wir zeigen (b)$\Rightarrow$ (a): aus $\det A' \neq 0$ folgt nach Lemma 7.3.5 $\operatorname{rg}(A') = k$, und daraus $\operatorname{rg}(A) \geq k$, da der Rang einer Untermatrix durch den Rang der Matrix nach oben beschränkt ist. Denn die lineare Abhängigkeit von Zeilen (oder Spalten) der Matrix impliziert, dass auch die entsprechenden Zeilen (oder Spalten) der Untermatrix linear abhängig sind.

Um (a)$\Rightarrow$ (b) zu sehen, beachten wir, dass es wegen $\operatorname{rg}(A) \geq k$ sicher k linear unabhängige Zeilenvektoren von A gibt. Wir wählen k solche Zeilen aus; für die dadurch erhaltene rechteckige Matrix ist nach Satz 5.5.12 der Zeilenrang gleich dem Spaltenrang. Wir finden also k linear unabhängige Spalten dieser Untermatrix, die wir auswählen, so dass wir eine $k \times k$-Teilmatrix A' erhalten, die maximalen Rang k und somit nach Lemma 7.3.5 nicht-verschwindende Determinante hat. □

Bemerkung 7.7.4. Seien $A, B \in M(m \times n, K)$ mit $m > n$. Es gilt $\operatorname{rg} A \leq n$ und $\operatorname{rg} B = \operatorname{rg} B^T \leq n$, somit

$$\operatorname{rg}(AB^T) \leq \min\{\operatorname{rg} A, \operatorname{rg} B\} \leq n < m$$

Damit ist der Rang der $m \times m$-Matrix AB^T nicht maximal. Es gilt im Fall $m > n$ also immer $\det AB^T = 0$.

7.8 Spur

Die Determinante ist eine sehr natürliche Abbildung $M(n \times n, K) \to K$ mit viele guten Eigentschaften, insbesondere ist sie nicht nur für Matrizen, sondern für lineare Abbildungen definiert. Aber die Determinante ist nicht linear. Gibt es auch eine lineare Abbildung $M(n \times n, K)$. die für lineare Abbildungen definiert ist?

Definition 7.8.1. Die *Spur* einer Matrix $A = (a_{ij}) \in M(n \times n, K)$ ist definiert als $\text{Tr}\ (A) = \sum_{i=1}^{n} a_{ii}$.

Die Abkürzung kommt von englischen *trace*.

Satz 7.8.2. *Die Spur hat folgende Eigenschaften:*

1. $\text{Tr}\ : M(n \times n, K) \to K$ *ist linear.*
2. *Es gilt* $\text{Tr}\ (AB) = \text{Tr}\ (BA)$.
3. *Wenn* A *und* A' *ähnlich sind ist* $\text{Tr}\ (A) = \text{Tr}\ (A')$.
4. *Es gilt* $\text{Tr}\ (A) = \text{Tr}\ (A^T)$.

Beweis:. 1. Wir rechnen $\text{Tr}\ (\lambda A + \mu B) = \sum_i (\lambda A + \mu B)_{ii} = \sum_i (\lambda A_{ii} + \mu B_{ii}) = \lambda \text{Tr}\ (A) + \mu \text{Tr}\ (B)$.

2. Wir rechnen $\text{Tr}\ (AB) = \sum_i (\sum_j a_{ij} b_{ji}) = \sum_j \sum_i b_{ji} a_{ij} = \text{Tr}\ (BA)$.

3. Wenn $A' = SAS^{-1}$ gilt dann wenden wir einfach die erste Aussage auf die Matrizen S und AS^{-1} an.

4. Die diagonalen Einträge von A und A^T sind gleich. □

An dieser Stelle haben wir noch keine Interpretation der Spur, aber es ist bemerkenswert, dass die Summe der Diagonalen Einträge sich nicht verändert, wenn wir einen Basiswechsel vornehmen!

Definition 7.8.3. Sei V ein endlich-dimensionaler Vektorraum. Die Spur einer linearen Abbildung $f : V \to V$ ist definiert als $\text{Tr}\ M_{\mathcal{B}}(f)$ für irgendeine Basis $\mathcal{B}$ von V.

Dies ist nach Satz 7.8.2 wohldefiniert denn $\text{Tr}\ (SAS^{-1}) = \text{Tr}\ (S^{-1}SA) = \text{Tr}\ (S)$.

Wie schon bei der Determinante ist die Definition nicht sinnvoll, wenn wir es mit Endomorphismen von unendliche-dimensionalen Vektorräumen zu tun haben.

8 Intermezzo: Kodierungstheorie

Wir wenden uns einer konkreten Anwendung von linearer Algebra über dem endlichen Körper $\mathbb{F}_2$ zu.

Bei der Übertragung von Daten treten typischerweise Fehler auf. Dies führt zu den beiden Zielen der Fehlererkennung und Fehlerkorrektur.

Wir gehen davon aus, dass die Daten in einem *Binärkode* vorliegen, d.h. als eine Folge der Symbole 0 oder 1. Ein Datensatz fester Länge n ist also ein Vektor im $\mathbb{F}_2$-Vektorraum $V = (\mathbb{F}_2)^n$.

Definition 8.0.1. Sei $K = \mathbb{F}_2$ und $V = K^n$. Die Abbildung

$$d_H : V \times V \to \mathbb{N}$$
$$d_H(v, w) := |\{j \in \{1, \ldots, n\} |\, v_j \neq w_j\}|$$

heißt *Hamming-Abstand*. Sie gibt die Zahl der Komponenten an, in der sich die beiden Argumente v und w unterscheiden.

Lemma 8.0.2. *Der Hamming-Abstand hat für alle $u, v, w \in V$ die folgenden Eigenschaften:*

1. $d_H(v, w) \geq 0$ *und* $d_H(v, w) = 0$ *genau für* $v = w$
2. $d_H(v, w) = d_H(w, v)$ *(Symmetrie)*
3. $d_H(u, w) \leq d_H(u, v) + d_H(v, w)$ *(Dreiecksungleichung)*
4. $d_H(v, w) = d_H(v + u, w + u)$ *(Translationsinvarianz)*

Beweis:. 1,2 und 4 sind trivial. Für 3. beachten wir: nur für $u_j \neq w_j$ trägt die j–te Komponente den Wert 1 zum Hamming-Abstand $d(u, v)$ bei. Dann ist aber entweder $v_j \neq u_j$ oder $v_j \neq w_j$. □

Definition 8.0.3. Sei $\lambda \in \mathbb{N}$. Eine Teilmenge $C \subset (\mathbb{F}_2)^n$ heißt λ*–fehlerkorrigierender Kode*, falls für alle $u, v \in C$, $u \neq v$ gilt

$$d_H(u, v) \geq 2\lambda + 1$$

Zum Beispiel ist für $n = 3$ die zweielementige Teilmenge von K^3

$$C = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

wegen $d_H\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = 3$ ein 1–fehlerkorrigierender Kode, denn der Hamming-Abstand beträgt 3.

Die Benennung erklärst sich wie folgt: Wenn wir bei der Übertragung von $(0, 0, 0)^T$ nur einen Fehler machen, erhalten wir $(1, 0, 0)^T, (0, 1, 0)^T$ oder $(0, 0, 1)^T$ und keiner dieser Vektoren lässt sich durch einen einzigen Fehler aus $(1, 1, 1)^T$ erzeugen. Wenn wir also $(1, 0, 0)^T$ empfangen können wir den einen Fehler korrigieren, und wissen, dass $(0, 0, 0)^T$ gesendet wurde.

Das wird im folgenden einfachen Lemma verallgemeinert:

Lemma 8.0.4. *Sei $C \subset V$ ein λ–fehlerkorrigierender Kode. Dann gibt es zu jedem $v \in V$* höchstens *ein $w \in C$ mit $d_H(v, w) \leq \lambda$.*

Beweis:. Wei $v \in V$ gegeben und seien $w_1, w_2 \in C$ mit $d_H(v, w_i) \leq \lambda$. Dann gilt wegen der Dreiecksungleichung 8.0.2.3

$$d_H(w_1, w_2) \leq d_H(w_1, v) + d_H(v, w_2) \leq 2\lambda \, .$$

Da der Kode C als λ–fehlerkorrigierend vorausgesetzt wurde, folgt $w_1 = w_2$. □

Betrachtung 8.0.5. Wir verwenden zur Versendung unserer Nachricht nur die Elemente eines λ–fehlerkorrigierenden Kodes $C \subset (\mathbb{F}_2)^n$. Treten bei der Übermittlung des Elements λ oder weniger Fehler auf, so kann die Nachricht (nämlich das Element aus C) aus dem empfangenen Datum (nämlich ein Element in V) eindeutig rekonstruiert werden.

Es treten die folgenden *Probleme* auf:

- Wir müssen geeignete Mengen C finden, die gute fehlerkorrigiernde Eigenschaften haben.
- Wir müssen die Menge C speichern, und in vielen Anwendungen ist eine große Menge C von Vorteil, die a priori viel Speicherplatz benötigt.
- Zur Dekodierung müssen wir $w \in C$ wie in Lemma 8.0.4 finden. Dazu müssen wir a priori viele Vergleiche der empfangenen Nachricht mit Elementen in C anstellen.

Definition 8.0.6. Ein λ–korrigierender Kode $C \subset V$ heißt *linear*, falls C ein Untervektorraum von V ist.

Lineare Kodes bieten einen Vorteil beim Speicherplatz: ist $\dim_{\mathbb{F}_2} C = k$, so hat eine Basis k Elemente und mit Angabe von k Basisvektoren, das sind $n \cdot k$ Koordinaten, können wir die Menge C mit ihren 2^k Elementen vollständig beschreiben.

Bemerkenswerterweise ist auch die Fehlerkorrektur für lineare Codes einfacher: Sei $C \subset (\mathbb{F}_2)^n$ ein linearer λ–korrigierender Kode der Dimension k. Wir wählen eine surjektive lineare Abbildung:

$$\Phi : (\mathbb{F}_2)^n \to (\mathbb{F}_2)^{n-k}$$

mit $\ker \Phi = C$. Solche Surjektionen existieren: der Basisergänzungssatz 4.3.7 erlaubt es uns, eine Basis $b_1, \ldots, b_k$ von C zu einer Basis $b_1, \ldots, b_n$ von $(\mathbb{F}_2)^n$ zu ergänzen. Dann schicken wir b_j für $j > k$ auf den Standardbasisvektor b_{j-k} von $\mathbb{F}_2^{n-k}$.

Die darstellende Matrix $H = M(\Phi) \in M((n-k) \times n, \mathbb{F}_2)$ heißt eine *Kontrollmatrix* des Kodes. Sie ist nicht eindeutig bestimmt!

Ein Element $y \in \mathbb{F}_2^{n-k}$ heißt *zulässig*, wenn es ein $x_y \in \Phi^{-1}(y)$ gibt mit $d_H(x_y, 0) \leq \lambda$.

Wir überlegen uns, dass dieses $x_y \in (\mathbb{F}_2)^n$ für ein gegebenes zulässiges $y \in \mathbb{F}_2^{n-k}$ eindeutig ist:
Seien $x, x' \in \Phi^{-1}(y)$ mit $d_H(x, 0) \leq \lambda$ und $d_H(x', 0) \leq \lambda$. Dann ist $x - x' \in \ker \Phi = C$ und

$$d_H(x - x', 0) = d_H(x, x') \leq d_H(x, 0) + d_H(x', 0) \leq 2\lambda \, .$$

Da der Kode C aber λ-fehlerkorrigierend sein soll, folgt $x - x' = 0$.

Betrachtung 8.0.7. Die *Dekodierung* geschieht nun folgendermaßen: der Empfänger speichert eine Liste der zulässigen Elemente $y \in (\mathbb{F}_2)^{n-k}$ mit den zugehörigen eindeutig bestimmten $x_y \in \Phi^{-1}(y) \subset (\mathbb{F}_2)^n$ mit $d_H(x_y, 0) \leq \lambda$. Für jede empfangene Nachricht $v \in (\mathbb{F}_2)^n$ berechnet der Empfänger das Element

$$y = \Phi(v) \in (\mathbb{F}_2)^{n-k} \, .$$

Ist y nicht zulässig, so sind so viele Fehler bei der Übertragung aufgetreten, dass eine Korrektur nicht möglich ist. Ist y dagegen zulässig, so ist $w := v - x_y$ die ursprüngliche Nachricht. Um dies zu sehen, berechnen wir

$$\Phi(w) = \Phi(v - x_y) = \Phi(v) - \Phi(x_y) = y - y = 0 \ ;$$

wegen $\ker \Phi = C$ liegt also $w \in C$ und stellt wirklich eine mögliche Nachricht dar. Da gilt

$$d_H(w, v) = d_H(w - v, 0) = d_H(x_y, 0) \leq \lambda \ ,$$

ist nach Lemma 8.0.4 w eindeutig und somit die gesendete Nachricht.

Beispiel 8.0.8. Ein berühmte Beispiel ist der $(7, 4)$-Hamming Code $C \subset \mathbb{F}_2^7$. Dieser Code ist ist isomorph zu $\mathbb{F}_2^4$ und wird erzeugt von den Zeilen der Matrix

$$G = \begin{pmatrix} E_4 & B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} .$$

Wir zeigen, dass der Code 1-fehlerkorrigierend ist. Aufgrund der Transitivität der Hamming-Distanz reicht es, zu prüfen, dass alle Linearkombinationen der vier Zeilen (v_1, v_2, v_3, v_4) Abstand drei vom Nullvektor haben. Da wir über $\mathbb{F}_2$ arbeiten, müssen wir nicht allzuviele Linearkombinationen prüfen.

Es ist klar, dass alle v_i Abstand drei von 0 haben.

Man sieht auch relativ leicht, dass sich je zwei Zeilen in drei Koordinaten unterscheiden (zwei der ersten vier und einer der letzten drei Einträge sind unterschiedlich). Daraus folgt, dass $d(v_i, v_j) = d(v_i + v_j, 0) \geq 3$ für $i \neq j$.

Wir stellen noch fest, dass $v = v_1 + v_2 + v_3 + v_4 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ ist (mit $d(v, 0) \geq 3$). Damit sieht man dass $d(\sum_{j \neq i} v_j, 0) = d(v_i, v) \geq 3$ und damit hat die Summe von drei beliebigen v_j auch Abstand mindestens 3 von 0.

Also ist die ein 1-fehlerkorrigierender Kode. Wir können mit einem Codewort von 7 Bits ein Nachrichtenwort von 4 Bits so übermitteln, dass alle Nachrichtenwörter Abstand 3 voneinander haben. Wir nennen $4/7 \equiv 0.571$ die *Datenrate*.

Die Kontrollmatrix stellt $\Phi : \mathbb{F}_2^7 \to \mathbb{F}_2^3$ dar und $H = M(\Phi)$ muss $HG^T = 0$ erfüllen und Rang 3 haben.

$$H = \begin{pmatrix} B^T & E_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

erfüllt diese Bedingungen! (Wir profitieren hier davon, dass wir unsere Basis in einer besonders praktischen Form angegeben haben.)

Da der Code 1-fehlerkorrigierend ist, sind die zulässigen y genau die Bilder der 7 Einheitsvektoren, also genau die Spalten von H.

Der Hamming-Code lässt sich auf höhere Dimensionen erweitern: Mit einem Codewort mit $2^r - 1$ Bits lässt sich ein Nachrichtenwort von $2^r - 1 - r$ Bits 1-fehlerkorrigierend kodieren. Je größer r ist, desto höher ist der Anteil an Bits, die für die tatsächliche Nachricht benutzt werden. Da nur ein Fehler korrigiert wird, eignet sich der Hamming-Code nur dann, wenn die Fehlerrate positiv aber niedrig ist, zum Beispiel beim Arbeitsspeicher von Computern.

9 Eigenwerte

9.1 Ein Beispiel

Wir betrachten als nächstes ein Beispiel für eine Anwendung linearer Algebra in der reinen Mathematik.

Sie erinnern sich vielleicht an die Fibonacci-Zahlen $F_0 = 0$, $F_1 = 1$ und $F_n = F_{n-1} + F_{n-2}$. Die Reihe beginnt also

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

Es ist nicht sofort klar, wie wir die n-te Fibonacci-Zahl bestimmen können, ohne all $n-1$ vorherigen Fibonacci-Zahlen zu kennen.

In der Sprache der linearen Algebra können wir die Rekursion schreiben als:

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}$$

Wenn wir die Matrix als A schreiben, dann gilt also

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Wir können also F_{n+1} sofort ablesen, wenn wir A^n schnell berechnen können.

Das scheint nicht viel zu helfen, die n Matrixmultiplikationen sind genau die gleiche Arbeit wie die Bestimmung aller Fibonacci-Zahlen.

Es ist dagegen leicht, Potenzen einer diagonalen Matrix zu nehemn, denn es gilt $D^n = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}^n = \begin{pmatrix} d_1^n & 0 \\ 0 & d_2^n \end{pmatrix}$.

A ist nicht diagonal, aber unsere hauptsächlichen Objekte sind nicht feste Matrizen, sondern lineare Abbildungen. Sei also α die lineare Abbildung, die von A dargestellt wird. Möglicherweise gibt es eine andere Basis $\mathcal{B}$, für die $[\alpha]_{\mathcal{B}}$ diagonal ist.

Dazu müssten wir zwei linear unabhängige Vektoren v_1, v_2 und zwei Elemente $d_1, d_2 \in \mathbb{C}$ finden, so dass $Av_1 = d_1 v_1$ und $Av_2 = d_2 v_2$. Die Gleichung $Av = dv$ ist äquivalent zu $(A - dE_2)v = 0$. Das lässt sich genau dann lösen, wenn $\det(A - dE_2) = 0$ ist!

Wir betrachten also

$$\det \begin{pmatrix} 1-d & 1 \\ 1 & -d \end{pmatrix} = -(1-d)d - 1 = d^2 - d - 1$$

und diese quadratische Gleichung hat zwei Lösungen $d_{1/2} = \frac{1 \pm \sqrt{5}}{2}$.

Jetzt müssen wir v_i finden. Wir können für v_1 das Gleichungssystem

$$\begin{pmatrix} 1-d_1 & 1 \\ 1 & -d_1 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = 0$$

lösen, ein Blick auf die zweite Zeile zeigt sofort $v_{11} = d_1 v_{12}$.

Wir erhalten $v_1 = \begin{pmatrix} d_1 \\ 1 \end{pmatrix}$ und ebenso $v_2 = \begin{pmatrix} d_2 \\ 1 \end{pmatrix}$.

Der Basiswechsel von$\mathcal{B} = \{v_1, v_2\}$ auf $\mathcal{E}$ ist $T = T_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} d_1 & d_2 \\ 1 & 1 \end{pmatrix}$ mit Determinante $\sqrt{5}$.

Was haben wir gewonnen? Es gilt

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = T \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} T^{-1}$$

und

$$\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}^n = (T\begin{pmatrix}d_1 & 0\\ 0 & d_2\end{pmatrix}T^{-1})^n = T\begin{pmatrix}d_1 & 0\\ 0 & d_2\end{pmatrix}^n T^{-1}$$

Damit ist

$$\begin{pmatrix}F_{n+1}\\ F_n\end{pmatrix} = T\begin{pmatrix}d_1 & 0\\ 0 & d_2\end{pmatrix}^n T^{-1}\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}d_1 & d_2\\ 1 & 1\end{pmatrix}\begin{pmatrix}d_1^n & 0\\ 0 & d_2^n\end{pmatrix}\frac{1}{\sqrt{5}}\begin{pmatrix}1 & -d_2\\ -1 & d_1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}$$
$$= \frac{1}{\sqrt{5}}\begin{pmatrix}d_1 & d_2\\ 1 & 1\end{pmatrix}\begin{pmatrix}d_1^n & 0\\ 0 & d_2^n\end{pmatrix}\begin{pmatrix}1\\ -1\end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix}d_1 & d_2\\ 1 & 1\end{pmatrix}\begin{pmatrix}d_1^n\\ -d_2^n\end{pmatrix}$$

und wir lesen ab:

$$F_n = \frac{d_1^n - d_2^n}{\sqrt{5}} = \frac{(\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n}{\sqrt{5}}$$

Der Ausdruck auf der rechten Seite ist tatsächlich eine ganze Zahl! ⟨⟨Dies ist nicht die einzige Methode, die geschlossene Formel für die n-te Fibonacci-Zahl zu finden, aber diese Methode ein Problem linear darzustellen und dann durch Diagonalisierung zu vereinfachen ist sehr produktiv!⟩⟩

9.2 Definitionen

Die Klassen äquivalenter $m \times n$ Matrizen kennen wir aus Bemerkung 5.5.2.4: jede Matrix $A \in M(m \times n, K)$ ist äquivalent zu genau einer Matrix der Form

$$\begin{pmatrix}E_r & \vdots & 0\\ \dots & . & \dots\\ 0 & \vdots & 0\end{pmatrix} \qquad \text{mit } \; r = \operatorname{rg}(A)\ .$$

Dies erlaubt es uns, durch Wahl geeigneter Basen für V und für W eine besonders einfache Beschreibung einer gegebenen linearen Abbildung $\Phi : V \to W$ mit $\dim V = n$ und $\dim W = m$ zu finden.

Wir wollen in diesem Kapitel die Grundlagen für eine Beschreibung der Ähnlichkeitsklassen quadratischer Matrizen legen. Dies geht Hand in Hand mit dem Verständnis der Frage, auf welche Form die darstellende Matrix eines Endomorphismus eines endlich-dimensionalen Vektorraums durch geschickte Basiswahl gebracht werden kann. Entsprechend werden wir ab sofort frei zwischen der Sprache der (quadratischen) Matrizen und der linearen (Selbst-)Abbildungen endlich-dimensionaler Vektorräume wechseln.

Definition 9.2.1. Sei K ein Körper.

1. Sei V ein K–Vektorraum und $f : V \to V$ ein Endomorphismus. Ein Element $\lambda \in K$ heißt *Eigenwert* von f, falls es ein $v \in V \setminus \{0\}$ gibt, so dass $f(v) = \lambda v$ gilt. Dann heißt v *Eigenvektor* von f zum Eigenwert λ.

2. Sei $A \in M(n \times n, K)$. Ein Element $\lambda \in K$ heißt *Eigenwert* von A, falls es ein $v \in K^n \setminus \{0\}$ gibt mit $A \cdot v = \lambda v$. Dann heißt v *Eigenvektor* von A zum Eigenwert λ.

3. $f : V \to V$ ein Endomorphismus eines K–Vektorraums V. Für $\lambda \in K$ heißt die Menge aller Eigenvektoren mit Eigenwert λ

$$\operatorname{Eig}(f, \lambda) := \{v \in V |\ f(v) = \lambda v\}$$

der *Eigenraum* von f zum Wert λ. Die gleiche Definition gilt für eine Matrix.

Beispiele 9.2.2. 1. Wir haben gesehen, dass $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ Eigenwerte $\lambda_\pm = \frac{1\pm\sqrt{5}}{2}$ mit Eigenvektoren $v_\pm = \begin{pmatrix} \lambda_\pm \\ 1 \end{pmatrix}$ hat.

2. Eine Diagonalmatrix
$$\begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix}$$
hat Eigenwerte d_i und Eigenvektoren e_i!

3. Sei V der Vektorraum der beliebig oft differenzierbaren Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Die lineare Funktion $f \mapsto \frac{df}{dx}$ von V zu sich selbst hat als Eigenwerte alle $\theta \in \mathbb{R}$ mit Eigenvektoren $e^{i\theta x}$. Das folgt aus Ergebnissen der Analysis.

Definition 9.2.3. Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen Vektorraums. Die Funktion
$$\begin{aligned} P_f : \quad & K \to K \\ & \lambda \mapsto \det(\lambda \mathrm{id}_V - f) \end{aligned}$$
heißt das *charakteristisches Polynom* von f. Genauso heißt $P_A : A \mapsto \det(\lambda E_n - A)$ das charakteristische Polynom von A.

Hier betrachten wir P_f als Funktion, nicht als abstraktes Element eines Polynomrings (das kommt später).

Bemerkung 9.2.4. Mancherorts wird das charakteristische Polynom auch als $\det(f - \lambda \mathrm{id}_V) = (-1)^{\dim V} \det(\lambda \mathrm{id} V - f)$ definiert. Da wir uns nur für die Nullstellen des Polynoms interessieren, spielt es keine Rolle, welche Definition gewählt wird.

Satz 9.2.5. *Die Eigenwerte von $f : V \to V$ sind genau die Nullstellen von P_f. Der Eigenraum $\mathrm{Eig}(f, \lambda)$ ist genau der Kern von $f - \lambda \mathrm{id}$, insbesondere ist es ein Untervektorraum.*

Beweis:. Es gilt $f(v) = \lambda(v)$ genau wenn $v \in \ker(f - \lambda \mathrm{id})$. Diese Abbildung hat einen nichttrivialen Kern genau dann wenn $f - \lambda \mathrm{id}$ nicht bijektiv ist, siehe Korollar 4.3.5, d.h. wenn $\det(f - \lambda \mathrm{id}) = 0$, siehe Korollar 7.3.6. □

Bemerkung 9.2.6. Sei $\mathcal{B}$ eine endliche geordnete Basis von V und
$$A = M_{\mathcal{B}}^{\mathcal{B}}(f) \in M(n \times n, K) \ .$$
Es ist
$$M_{\mathcal{B}}^{\mathcal{B}}(\lambda \mathrm{id}_V - f) = \lambda E_n - A \ ,$$
woraus folgt:
$$P_f(\lambda) = \det(\lambda \mathrm{id}_V - f) = \det(\lambda E_n - A) =: P_A(\lambda) \ .$$
Insbesondere haben ähnliche Matrizen das gleiche charakteristische Polynom und die gleichen Eigenwerte, da sie den gleichen Endomorphismus bezüglich verschiedener Basen darstellen, vergleiche Bemerkung 5.5.2.3.

Beachten Sie, dass die Eigenvektoren von f auch unabhängig der Basis definiert sind, aber die Darstellung der Eigenvektoren von der Basis abhängt.

Bemerkung 9.2.7. Zur Bestimmung der Eigenwerte einer linearen Abbildung müssen wir also die Nullstellen von P_f finden. Es folgt zum Beispiel aus der Leibniz'schen Formel, dass P_f tatsächlich ein Polynom ist und Grad n hat.

Wenn $n = 2$ dann können wir einfach die p/q-Formel benutzen. Für größere Werte von n empfiehlt es sich in Übungsaufgaben einen Eigenwert durch Raten und Probieren zu bestimmen und dann durch Polynomdivision auf ein Polynom niedrigeren Grades zu reduzieren.

Es gibt auch für $n = 3, 4$ Formeln, die aber sehr unhandlich sind.

Wenn Sie eine Matrix außerhalb Ihrer Vorlesung treffen und nicht annehmen können, dass das die Eigenwerte leicht zu raten sind, dann bestimmen Sie die Nullstellen des charakteristischen Polynoms numerisch.

Sei zum Beispiel $p(x) = x^3 + x^2 - 11x - 12$ gegeben. Dann sehen wir $p(-1) = 0$ und rechnen durch Polynomodivision $x^3 + x^2 - 11x - 12 = (x+1)(x^2 + x - 12)$. Wir bestimen die Nullstellen von $x^2 + x - 12$ mit der quadratischen Formel und erhalten $\lambda_1 = -1$ $\lambda_2 = 3$, $\lambda_3 = -4$.

Wir betrachten noch einmal die rechnerische Bestimmung des Eigenraums zu.

Betrachtung 9.2.8. Gegeben sei ein Endomorphismus f eines endlichdimensionalen Vektorraums V mit Eigenwert λ. Wir wollen den Eigenraum $\mathrm{Eig}(f, \lambda) = \ker(f - \lambda \mathrm{id}_V)$ konkret bestimmen.

1. Wir wählen zuerst eine Basis $\mathcal{B}$ und schreiben V als K^n und $A = M_{\mathcal{B}}(f)$. In vielen Beispielen ist V schon K^n und wir können einfach die Standardbasis wählen.

2. Wir wollen nun das Gleichungssystem $(A - \lambda E_n).x = 0$. Dazu wenden wir den Gaußalgorithmus an und erhalten $A'.x = 0$ mit A' in reduzierter Zeilenstufenform.

3. Wir können sofort die Dimension des Eigenraums als Anzahl der Zeilen gleich 0 von A' ablesen.

4. Um eine Basis des Kerns zu finden lösen wir das Gleichungssystm durch sukzessives Einsetzen von unten nach oben. Wir fügen dabei jedes Mal eine freie Variable hinzu, wenn in einer Spalte der größte Index eines Eintrags ungleich Null höchstens so groß ist, wie in der Spalte zuvor, ander ausgedrückt "immer wenn die Nullen im unteren Teil des Vektors nicht weniger werden".

5. Wir schreiben unsere Lösung in freien Variablen als Linearkombination von Vektoren um, jede freie Variable korrespondiert zu einem Basisvektor des Kerns.

Hier die letzten Schritte einmal im Beispiel: Gegeben

$$A' = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Die letzte Zeile ist gleich 0 und ergibt die leere Gleichung. Die vorletzte Zeile ergibt $x_4 = 0$. Für die zweite Zeile dürfen wir eine neue freie Variable hinzufügen: $2x_2 + x_3 = 0$. Also gilt $x_3 = -2x_2$. Für die erste Zeile erhalten wir $2x_1 + x_2 + 3x_3 + 4x_4 = 0$, umgeschrieben erhalten wir $2x_1 + x_2 - 6x_2 + 0 = 0$ und damit ist $x_1 = -\frac{5}{2}x_2$.

Wir erhalten einen einzigen Basisvektor $(-\frac{5}{2}, 1, -2, 0)^T$.

Definition 9.2.9. Ein Endomorphismus $f : V \to V$ heißt *diagonalisierbar*, falls es eine Basis $\mathcal{B}$ von V gibt, die nur aus Eigenvektoren von f besteht. Ebenso heißt eine Matrix $A \in M(n \times n, K)$ diagonalisierbar, wenn es eine Basis $\mathcal{B}$ von K^n gibt so dass $T^{\mathcal{E}}_{\mathcal{B}} A T^{\mathcal{B}}_{\mathcal{E}}$ diagonal ist.

Ist V endlich-dimensional, so ist die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{B}}(f)$ bezüglich jeder Ordnung dieser Basis $\mathcal{B}$ eine Diagonalmatrix deren Einträge genau die Eigenwerte sind. Wenn wir die Basis umordnen, dann permutieren wir die Diagonaleinträge.

Wir sehen leicht, dass $\mathrm{Eig}(f, \lambda_1) \cap \mathrm{Eig}(f, \lambda_2) = \{0\}$ wenn $\lambda_1 \neq \lambda_2$. Denn für $v \in \mathrm{Eig}(f, \lambda_1 \cap \mathrm{Eig}(f, \lambda_2))$ folgt

$$\lambda_1 v = f(v) = \lambda_2 v \ ,$$

also $(\lambda_1 - \lambda_2)v = 0$, woraus $v = 0$ wegen $\lambda_1 \neq \lambda_2$ folgt. Dies verallgemeinert wie folgt:

Satz 9.2.10. *Sei $f : V \to V$ ein Endomorphismus. Seien $v_1, \ldots, v_m$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_m$. Dann ist die Familie $(v_1, \ldots, v_m)$ linear unabhängig.*

Beweis:. Wir benutzen vollständige Induktion nach m. Für $m = 1$ ist $v_1 \neq 0$ per Definition eines Eigenvektors klar. Gelte

$$0 = \sum_{i=1}^{m} \alpha_i v_i \qquad \text{mit} \qquad \alpha_i \in K \ . \qquad (*)$$

Es folgt

$$0 = f(0) = \sum_{i=1}^{m} \alpha_i f(v_i) = \sum_{i=1}^{m} \alpha_i \lambda_i v_i \ .$$

Die Multiplikation von $(*)$ mit λ_1 liefert:

$$0 = \sum_{i=1}^{m} \alpha_i \lambda_1 v_i \ .$$

Die Differenz der Gleichungen ist

$$\sum_{i=2}^{m} \alpha_i (\lambda_i - \lambda_1) v_i = 0 \ .$$

Nach Induktionsannahme ist die Familie $(v_2, \ldots, v_m)$ linear unabhängig, also $\alpha_i(\lambda_i - \lambda_1) = 0$ für alle $i = 2, \ldots, m$. Wegen $\lambda_i \neq \lambda_1$ folgt $\alpha_2 = \alpha_3 = \ldots = \alpha_m = 0$. Da $v_1 \neq 0$ ist, folgt auch $\alpha_1 = 0$. □

Korollar 9.2.11. *1. Ist $n := \dim_K V < \infty$, so hat jeder Endomorphismus $f : V \to V$ höchstens n verschiedene Eigenwerte.*

2. Ist $n := \dim_K V < \infty$ und hat f genau n verschiedene Eigenwerte, so ist f diagonalisierbar.

Beweis:. 1. ist klar, weil jede Familie von mehr als n Vektoren linear abhängig ist.

2. Wähle zu jedem Eigenwert λ_i einen Eigenvektor v_i. Die Familie $(v_i)_{i=1\ldots n}$ ist nach Satz 9.2.10 linear unabhängig und wegen $\dim_K V = n$ eine Basis. □

Definition 9.2.12. Wenn λ ein Eigenwert von f ist, dann ist

$$\mu_{geo}(f, \lambda) := \dim_K \mathrm{Eig}(f, \lambda)$$

ungleich 0 und heißt die *geometrische Vielfachheit* des Eigenwerts λ.

Satz 9.2.13. *Sei $A \in M(n \times n, K)$. Dann ist $P_A(\lambda)$ eine polynomiale Funktion vom Grad n,*

$$P_A(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0 \qquad \textit{mit } \; a_i \in K$$

und

$$a_n = 1, \quad a_{n-1} = -(a_{11} + a_{22} + \ldots + a_{nn}) = -\mathrm{Tr}\,(A) \quad \textit{und} \quad a_0 = (-1)^n \det A\,.$$

Beweis:. Wir betrachten die Lebnizsche Formel, die die Form

$$P_A(\lambda) = \det(\lambda E_n - A) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} \lambda\delta_{i\sigma(i)} - a_{i\sigma(i)}$$

annimmt. Es ist sofort klar, dass dies ein Polynom in λ ist, und der Term mit dem höchsten Grad t^n ist.

Als nächstes suchen wir den Koeffizienten von λ^{n-1}. Wir können im Produkt nur dann $n-1$ Faktoren gleich λ haben, wenn σ $n-1$ Stellen fixiert, das geht nur wenn $\sigma = \mathrm{id}$. Wir müssen nur noch den Koffizienten von λ^{n-1} in $\prod_{i=1}^{n}(t\delta_{ii} - a_{ii})$ bestimmen, und das ist $-\sum_{i=1}^{n} a_{ii}$.

Schließlich beachten wir

$$a_0 = P_A(0) = \det(0 \cdot \mathrm{id}_V - A) = (-1)^n \det A\,. \quad \square$$

Bemerkung 9.2.14. Um eine Matrix M zu diagonalisieren führen wir die folgenden Schritte durch:

1. Berechne das charakteristische Polynom $P_M(\lambda)$ und bestimme seine Nullstellen, vgl. Bemerkung 9.2.7. Dies sind die Eigenwerte.

2. Wenn alle Eigenwerte unterschiedlich sind, dann muss M zur Diagonalmatrix mit den Eigenwerten auf der Diagonale ähnlich sein! Wir können weitermachen um eine Basis zu finden, so dass der Basiswechsel M diagonalisiert.

3. Zu jedem Eigenwert λ bestimme den Eigenraum wie in Betrachtung 9.2.8.

   $$x \in \mathrm{Eig}(M, \lambda) = \ker(\lambda E_n - M) \Leftrightarrow (\lambda E_n - M)x = 0\,.$$

   Dieses lineare Gleichungssystem für x von n Gleichungen in n Unbestimmten kann zum Beispiel mit dem Gauß'schen Algorithmus gelöst werden.

4. Wähle Basen der Eigenräume. Wenn diese zusammen eine Basis $(v_1, \ldots, v_n)$ von K^n bilden, dann ist M diagonalisierbar. Wegen Satz 9.2.10 reicht es zu prüfen, dass die Summe der Dimensionen der Eigenräume, also $\sum_\lambda \mu_{geo}(f, \lambda)$, gleich $\dim(V)$ ist. Setze

   $$S^{-1} := \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}\,.$$

   Dann ist $D = SMS^{-1}$ eine Diagonalmatrix, denn es gilt für den i-ten Vektor der Standardbasis

   $$SMS^{-1}e_i = SMv_i = \lambda_i S v_i = \lambda_i e_i\,.$$

   Die diagonalen Einträge von D sind die Eigenwerte, und jeder Eigenwert kommt mit seiner geometrischen Vielfachheit $\mu_{geo}(f, \lambda)$ vor.

Beispiele 9.2.15. 1. Wir betrachten eine Drehung des $\mathbb{R}^2$ um den Ursprung:

$$A = M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$P_A(\lambda) = \det \begin{pmatrix} \lambda - \cos\theta & \sin\theta \\ -\sin\theta & \lambda - \cos\theta \end{pmatrix} = \lambda^2 - 2\cos\theta\lambda + 1$$

Die komplexen Nullstellen des charakteristischen Polynoms sind

$$\lambda_{1,2} = \cos\theta \pm \mathrm{i}\sin\theta\ .$$

Diese sind nur für $\theta = 0, \pi$ reell. Nur dann gibt es reelle Eigenwerte und auch Eigenvektoren in $\mathbb{R}^2$. Bei $\theta = 0$ handelt es sich um die Identität, alle Vektoren sind Eigenvektoren zum Eigenwert 1; bei $\theta = \pi$ handelt es sich um die Punktspiegelung am Ursprung, alle Vektoren sind Eigenvektoren zum Eigenwert -1.

Es gibt immer komplexe Eigenwerte und komplexe Eigenvektoren und über $\mathbb{C}$ ist R_θ diagonalisierbar, vgl. Übung 12.3.

2. Wir betrachten eine Spiegelung des $\mathbb{R}^2$ an einer Ursprungsgeraden:

$$A = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

(a)

$$P_A(\lambda) = (\lambda - \cos 2\theta)(\lambda + \cos 2\theta) - \sin^2 2\theta = \lambda^2 - 1$$

Die beiden Eigenwerte sind $\lambda_{1,2} = \pm 1$.

(b) Eigenräume sind

$$\mathrm{Eig}(A, 1) = \left\{ t \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \mid\ t \in \mathbb{R} \right\}$$

$$\mathrm{Eig}(A, -1) = \left\{ t \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix} \mid\ t \in \mathbb{R} \right\}$$

Man beachte, dass der Eigenraum zum Eigenwert $+1$ die Spiegelachse ist und der Eigenraum zum Eigenwert -1 senkrecht bezüglich des euklidischen Standardskalarprodukts auf $\mathbb{R}^2$ auf der Spiegelachse steht.

(c) Wir erhalten

$$S^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

und finden, wie erwartet,

$$SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\ .$$

3. Als Beispiel bestrachten wir die Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Ihr charakteristisches Polynom ist

$$P_A(\lambda) = \det \begin{pmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{pmatrix} = (\lambda - 1)^2\ .$$

Also hat A nur den Eigenwert 1. Die Eigenvektoren bestimmen wir zu

$$x \in \mathrm{Eig}(A, 1) \Leftrightarrow \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} x = 0 \Leftrightarrow x_2 = 0\ .$$

Somit ist $\mathrm{Eig}(A, 1) = \mathrm{span}_K e_1$ nur eindimensional. Die Matrix A ist also nicht diagonalisierbar!

9.3 Polynome

In diesem Unterkapitel fixieren wir nicht einen Körper K, sondern einen kommutativen Ring R (mit Eins). Sie dürfen Sie sich unter R ruhig einen Körper vorstellen.

Definition 9.3.1. Gegeben sei ein kommutativer Ring R. Dann ist eine *R-Algebra* ein Ring A zusammen mit einem injektiven Ringhomomorphismus $i : R \to A$ so dass $i(r)a = ai(r)$ für alle $a \in A, r \in R$, d.h. die Elemente im Bild von R kommutieren mit allen Elementen von A.

Warnung. Üblicherweise wird eine R-Algebra einfach als Abbildung $i : R \to A$ definiert, ohne die Bedingung, dass i injektiv ist. Für uns ist R meistens ein Körper, dann ist i automatisch injektiv.

Wir schreiben auch einfach r für $i(r)$. Wenn R ein Körper ist, dann folgt aus der Multiplikation mit R, dass A ein R-Vektorraum ist.

Definition 9.3.2. Seien A, B zwei R–Algebren. Ein Ringhomomorphismus $\phi : A \to B$ heißt *R–Algebrenhomomorphismus*, wenn er mit den Injektionen $i_A : R \to A$ und $i_B : R \to B$ kommutiert, also gilt $\phi(i_A(r)) = i_B(r)$.

Es gilt dann insbesondere $\phi(i_A(r) \cdot a) = \phi(i_A(r))\phi(a) = i_B(r)\phi(a)$, oder kurz gesagt $\phi(ra) = r\phi(a)$.

Zum Beispiel ist $\mathbb{C}$ eine $\mathbb{R}$-Algebra und $\mathbb{Q}$ ist eine $\mathbb{Z}$-Algebra. Eine Algebra muss nicht kommutativ sein, z.B. ist $M(n \times n, K)$ eine K-Algebra, die Abbildung $K \to M(n \times n, K)$ ist durch $\lambda \mapsto \lambda E_n$ definiert.

Wir erinnern uns an Polynomringe aus Betrachtung 2.5.8:

Definition 9.3.3. Sei R ein kommutativer Ring. Wir definieren den *Polynomring* $R[X]$ in einer Unbestimmten X über R als die Menge aller beliebig langen Folgen $(r_0, r_1, r_2, \dots)$ in R, in denen nur endlich viele Koordinaten ungleich 0 sind. Wir schreiben $(r_0, r_1, \dots, r_n, 0, 0, \dots)$ als $\sum_{i=0}^{n} r_i X^i$.

Wir definieren eine Komponentenweise Addition und eine Mulitplikation durch

$$(\sum_{i=0}^{n} r_i X^i) \cdot (\sum_{i=0}^{m} s_i X^i) = \sum_{k=0}^{n+m} \sum_{i+j=k} r_i s_j X^{i+j}$$

Durch Nachrechnen von Assoziativität (siehe Übungsblatt) und Distributivität sehen wir, dass $R[X]$ ein Ring ist, der überdies kommutativ ist. Mit der natürlichen Injektion $r \mapsto r.X^0$ ist $R[X]$ zudem eine R-Algebra.

Die Polynomalgebra hat die folgende Eigenschaft, die wir *universelle Eigenschaft* nennen, zu diesem Begrif später mehr.

Satz 9.3.4. *Gegeben seien $R[X]$ und eine R–Algebra A, und ein Element $a \in A$. Dann gibt es genau einen R–Algebrenhomomorphismus $\varphi_a : R[X] \to A$ mit $\varphi_a(X) = a$.*

Beweis:. Als R-Algebrenhomomorphismus muss $\phi_a(r) = r$ für $r \in R$ sein, und außerdem gilt wegen der Multipliktivität $\phi_a(rX^i) = ra^i$ für $r \in R$ und $i = 1, 2, \dots$.

Zusammen mit der Linearität erhalten wir

$$\phi_a(\sum_i r_i X^i) = \sum_i r_i a^i$$

für $f = \sum_{i=0}^{n} r_i X^i$ als einzig möglichen R-Algebrahomomorphismus, und es ist leicht zu prüfen, dass dies tatsächlich ein R-Algebrenhomomorphismus ist.

Zum Beispiel prüfen wir Multiplikativität:

$$\begin{aligned}\phi_a(\sum_{i=0}^{n} r_i X^i) \cdot \phi_a(\sum_{i=0}^{m} s_i X^i) &= (\sum_{i=0}^{n} r_i a^i) \cdot (\sum_{i=0}^{m} s_i a^i) \\ &= \sum_{k=0}^{n+m} \sum_{i+j=k} r_i s_j a^{i+j} \\ &= \phi_a(\sum_{k=0}^{n+m} \sum_{i+j=k} r_i s_j X^{i+j})\end{aligned}$$

indem wir im zweiten Schritt die Distributivität in A verwenden. □

Bemerkungen 9.3.5. 1. Aus naheliegenden Gründen heißt φ_a auch der zum Element $a \in A$ gehörende *Einsetzungshomomorphismus*. Man schreibt auch für ein Polynom $f \in R[X]$ auch

$$\varphi_a(f) = f(a) \ .$$

2. Im Spezialfall der R-Algebra $A = R$ erhalten wir für jedes $\lambda \in R$ einen Wert

$$\varphi_\lambda(f) = f(\lambda) \in K \ .$$

Ein Polynom gibt also eine polynomiale Funktion $R \to R$. Dies ergibt wieder einen R-Algebrenhomomorphismus

$$\begin{array}{lcl} R[X] & \to & \mathrm{Abb}(R,R) \ , \\ f & \mapsto & \tilde{f} \end{array}$$

den *Auswertehomomorphismus*. Dieser ist im Allgemeinen *nicht* injektiv, d.h. ein Polynom kann nicht mit der induzierten polynomialen Funktion identifiziert werden: zum Beispiel gibt es für $R = \mathbb{F}_2$ nur 4 verschiedene Funktionen, aber unendlich viele verschiedene Polynome.

Wir können auch für andere Algebren A den Auswertehomomorphismus $R[X] \to \mathrm{Abb}(A, A)$ durch $f \mapsto (\tilde{f} : a \mapsto \phi_a(f))$. Ein Polynom definiert also nicht nur eine Abbildung von K nach K sondern für jede Algebra eine Abbildung A nach A. Dies ist ein weiterer Vorteil der Betrachtung abstrakter Polynome, den wir später gebrauchen werden.

Definition 9.3.6. Besitzt $f \in K[X]$ die Gestalt

$$f = a_0 + a_1 X + \ldots + a_n X^n$$

mit $a_n \neq 0$, so heißt a_n der *höchste Koeffizient* von f. Die nicht-negative Zahl n heißt dann der *Grad* von f. Dem Nullpolynom ordnen wir zu

$$\mathrm{grad}(0) = -\infty \ .$$

Ist der höchste Koeffizient $a_n = 1$, so heißt das Polynom *normiert*.

Bemerkungen 9.3.7. 1. Es gilt immer

$$\begin{aligned}&\mathrm{grad}(f+g) \leq \max\left(\mathrm{grad}(f), \mathrm{grad}(g)\right) \\ &\mathrm{grad}(fg) \leq \mathrm{grad}(f) + \mathrm{grad}(g) \ .\end{aligned}$$

2. Ist der kommutative Ring R *nullteilerfrei*, d.h. folgt aus $ab = 0$ mit $a, b \in K$, dass $a = 0$ oder $b = 0$ gilt, so gilt
$$\operatorname{grad}(fg) = \operatorname{grad}(f) + \operatorname{grad}(g) .$$
Umgekehrt muss R nullteilerfrei sein, wenn diese Beziehung genau für alle Polynome f, g gilt.

Konvention 9.3.8. Ab sofort soll R ein nullteilerfreier kommutativer Ring sein. Insbesondere kann R ein beliebiger Körper sein.

Satz 9.3.9 (*Division mit Rest* von Polynomen)**.** *Sei $f \neq 0$ ein Polynom in $R[X]$, dessen höchster Koeffizient in R ein multiplikatives Inverses hat. (Diese Bedingung ist automatisch erfüllt, wenn K ein Körper ist.) Zu jedem Polynom $g \in R[X]$ gibt es dann Polynome $q, r \in R[X]$ mit*
$$g = qf + r \quad \textit{und} \quad \operatorname{grad}(r) < \operatorname{grad}(f) .$$
Hierdurch sind q und r eindeutig bestimmt.

Ist $r = 0$, so sagen wir, dass f das Polynom g *teilt*, in Zeichen: $f|g$, bzw. dass g ein *Vielfaches* von f ist, genau wie wir das für ganze Zahlen sagen würden.

Beweis:. Den Beweis der Existenz von q und r für gegebenes f führen wir mit vollständiger Induktion nach $\operatorname{grad}(g)$.

- Für $\operatorname{grad}(g) < \operatorname{grad}(f)$ setze $q = 0$ und $r = g$.
- Gelte also $m := \operatorname{grad} g \geq n := \operatorname{grad} f$. Sei a der invertible höchste Koeffizient von f und c der höchste Koeffizient von g. Das Polynom
$$h := g - ca^{-1} f \cdot X^{m-n}$$
hat strikt kleineren Grad als g. Nach Induktionsannahme existieren $q_1, r \in K[X]$ mit
$$h = q_1 f + r \qquad \text{und} \quad \operatorname{grad} r < \operatorname{grad} f .$$
Daraus folgt eine Darstellung von g
$$g = h + ca^{-1} f \cdot X^{m-n} = \left(q_1 + ca^{-1} \cdot X^{m-n}\right) f + r \quad \text{und} \quad \operatorname{grad} r < \operatorname{grad} f .$$
- Um die Eindeutigkeit dieser Darstellung zu zeigen, nehmen wir an, es gelte
$$g = qf + r = \widetilde{q} f + \widetilde{r}$$
mit $q, \widetilde{q}, r, \widetilde{r} \in K[X]$ und $\operatorname{grad} \widetilde{r} < \operatorname{grad} f$, $\operatorname{grad} r < \operatorname{grad} f$. Hieraus folgt
$$r - \widetilde{r} = (\widetilde{q} - q) f$$
Wäre $\widetilde{q} \neq q$, so wäre
$$\operatorname{grad}(r - \widetilde{r}) = \operatorname{grad}(\widetilde{q} - q) + \operatorname{grad}(f) \geq \operatorname{grad}(f) ,$$
im Widerspruch zur Bedingung an den Grad eines Rests. Also muss $\widetilde{q} = q$ gelten und somit $\widetilde{r} = r$.

□

Beispiel 9.3.10. Es gilt: $(X^4 - 1) : (X - 1) = X^3 + X^2 + X + 1$.

Lemma 9.3.11. *Sei $f \in R[X]$. Ist $a \in R$ eine* Nullstelle *von f, d.h. gilt $\tilde{f}(a) = 0$, so gibt es genau ein Polynom $g \in K[X]$ mit*

$$f = (X - a)g$$

und $\text{grad}(g) = \text{grad}(f) - 1$.

Man sagt auch, dass man einen Linearfaktor abspalten kann.

Beweis:. Die Polynomdivision mit Rest nach Satz 9.3.9 liefert uns Polynome g, r, die

$$f = (X - a) \cdot g + r$$

erfüllen mit $\text{grad} r < \text{grad}(X - a) = 1$; also ist r ein konstantes Polynom. Wegen $0 = \tilde{f}(a) = \tilde{r}(a)$ folgt für das konstante Polynom $r = 0$. Es ist also $f = (X - a) \cdot g$ und außerdem gilt

$$\text{grad } f = \text{grad}(X - a) + \text{grad } g = 1 + \text{grad } g \quad \square$$

Korollar 9.3.12. *Sei $f \in R[X]$, $f \neq 0$. Hat f genau k paarweise verschiedene Nullstellen, so ist*

$$k \leq \text{grad}(f) \ .$$

Ein Polynom vom Grad n hat also höchstens n paarweise verschiedene Nullstellen.

Beweis:. Vollständige Induktion nach $n := \text{grad } f$.

- Für $n = 0$ ist f konstantes Polynom, hat also keine Nullstelle. Also ist $k = 0$.
- Induktionsschritt: hat f keine Nullstelle, so ist $k = 0$ und die Behauptung trivialerweise richtig. Hat f eine Nullstelle $\lambda \in R$, so gibt es nach Lemma 9.3.11 ein Polynom g vom Grad $\text{grad } g = \text{grad } f - 1$ mit
$$f = (X - \lambda)g$$
Jede von λ verschiedene Nullstelle ist dann auch Nullstelle von g, da R nullteilerfrei ist. Nach Induktionsvoraussetzung hat aber g höchstens $n - 1$ verschiedene Nullstellen. $\square$

Korollar 9.3.13. *Ist R ein Ring (weiterhin kommutativ ohne Nullteiler) mit unendlich vielen Elementen, so ist der Auswertehomomorphismus*

$$\begin{array}{lcl} R[X] & \to & \text{Abb}(R, R) \\ f & \mapsto & \tilde{f} \end{array}$$

injektiv.

Beweis:. Wäre f im Kern des Auswertehomomorphismus, aber nicht das Nullpolynom, so hätte f unendlich viele verschiedene Nullstellen. Dies ist im Widerspruch zu Korollar 9.3.12. $\square$

Definition 9.3.14. Sei $f \in R[X]$ und $f \neq 0$. Für $\lambda \in R$ heißt

$$\mu(\lambda, f) := \max\{r \in \mathbb{N} \mid \exists g \in R[X] \text{ mit } f = (X - \lambda)^r \cdot g\}$$

die *Vielfachheit* der Nullstelle λ von f.

Bemerkungen 9.3.15.

1. Es gilt $0 \leq \mu(\lambda, f) \leq \text{grad}(f)$
2. Gilt $f = (X - \lambda)^{\mu(\lambda, f)} g$, so ist λ keine Nullstelle von g.

3. $\mu(\lambda, f) = 0$ genau dann, wenn λ keine Nullstelle von f ist.

Lemma 9.3.16. *Sind $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Nullstellen von f mit Vielfachheiten $r_1, \ldots, r_k$, so ist*

$$f = (X - \lambda_1)^{r_1}(X - \lambda_2)^{r_2} \ldots (X - \lambda_k)^{r_k} g \ ,$$

wobei $g \in K[X]$ keine Nullstellen hat. Hierbei sind das Polynom g, die Nullstellen λ_i und ihre Vielfachheiten $, r_i$ bis auf Reihenfolge eindeutig.

Beweis:. Wir wenden induktiv Lemma 9.3.11 an. Die Nullstellen und Vielfachheiten sind wohldefiniert und nach Satz 9.3.9 ist damit auch g eindeutig bestimmt. □

Beispiele 9.3.17.

- Wir rechnen zunächst über dem Körper $\mathbb{R}$ der reellen Zahlen:

$$\begin{aligned} f &= X^5 - X^4 + X^3 - X^2 \in \mathbb{R}[X] \\ f &= X^2 \left(X^3 - X^2 + X - 1\right) \\ &= X^2 \left(X - 1\right) \left(X^2 + 1\right) \ . \end{aligned}$$

  Hier ist $\lambda_1 = 0, r_1 = 2$, $\lambda_2 = 1, r_2 = 1$, $g = X^2 + 1$ Man beachte, dass die Summe der Vielfachheiten kleiner als der Grad von f ist.

- Über dem Körper $\mathbb{C}$ der komplexen Zahlen dagegen erhalten wir

$$\begin{aligned} f &= X^5 - X^4 + X^3 - X^2 \in \mathbb{C}[X] \\ f &= X^2 \left(X - 1\right) \left(X + \mathrm{i}\right) \left(X - \mathrm{i}\right) \ . \end{aligned}$$

  Hier ist $\lambda_1 = 0, r_1 = 2$, $\lambda_2 = 1, r_2 = 1, \lambda_3 = +\mathrm{i}, r_3 = 1$, $\lambda_4 = -\mathrm{i}, r_4 = 1$ und $g = 1$. Man beachte, dass hier die Summer der Vielfachheiten gleich dem Grad von f ist.

Definition 9.3.18. Wir sagen, ein Polynom $f \in RK[X]$ zerfällt in *Linearfaktoren*, falls es sich in der Form

$$f = a(X - \lambda_1)^{r_1}(X - \lambda_2)^{r_2} \ldots (X - \lambda_n)^{r_n}$$

mit $a, \lambda_1, \ldots, \lambda_n \in K$ und $r_i \in \mathbb{N}$ schreiben lässt. Ein solcher Ausdruck heißt *Linearfaktorzerlegung* des Polynoms f.

Bemerkungen 9.3.19.

1. Ein Polynom f zerfällt genau dann in ein Produkt von Linearfaktoren, wenn $\sum_{\lambda \in K} \mu(\lambda, f) = \operatorname{grad} f$ gilt. (Man beachte, dass es nach Korollar 9.3.12 nur endlich viele Nullstellen gibt und daher die Summe endlich ist.)

2. Existiert eine Zerlegung in Linearfaktoren, so ist diese eindeutig bis auf die Reihenfolge der Faktoren.

Satz 9.3.20 (Fundamentalsatz der Algebra)**.** *Ist $f \in \mathbb{C}[X]$ mit $\operatorname{grad}(f) \geq 1$, so besitzt f wenigstens eine komplexe Nullstelle.*

Zum Beweis:. Dieser Satz ist trotz des Namens kein algebraisches sondern ein komplexanalytisches Resultat. Es kann mit Hilfsmitteln der Analysis, der komplexen Analysis, der algebraischen Topologie oder der Galoistheorie bewiesen werden. □

Korollar 9.3.21. *Jedes komplexe Polynom zerfällt in $\mathbb{C}[X]$ in Linearfaktoren.*

Beweis:. Wir schreiben nach Lemma 9.3.16 $f = (X - \lambda_1)^{r_1} \ldots (X - \lambda_n)^n g$ mit $\lambda_i \in \mathbb{C}$, wobei das komplexe Polynom g keine Nullstellen hat. Nach dem Fundamentalsatz der Algebra ist g konstant. □

9.4 Diagonalisierbarkeit

Wir arbeiten nun wieder über einem Körper K. Wir wollen untersuchen, welche Matrizen $A \in M(n \times n, K)$ diagonalisierbar sind, die also die Eigenschaft haben, dass K^n eine Basis aus Eigenvektoren von A besitzt. Wir wissen schon:

- Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar (Korollar 9.2.11.2).
- Ist A ähnlich zur Diagonalmatrix D, so ist

$$\begin{aligned} P_A(\lambda) &= P_D(\lambda) \\ &= \det \begin{pmatrix} \lambda - \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda - \lambda_n \end{pmatrix} \\ &= (\lambda - \lambda_1) \cdot \ldots \cdot (\lambda - \lambda_n) \end{aligned}$$

  d.h. das charakteristische Polynom zerfällt in Linearfaktoren.

Wir hatten das charakteristische Polynom eingeführt als polynomiale Abbildung

$$\begin{array}{rccl} P_A: & K & \to & K \\ & \lambda & \mapsto & \det(\lambda E_n - A). \end{array}$$

Aber das natürlichere Objekt ist das abstrakete charakteristische Polynom $P_A \in K[T]$. Wenn zum Beispiel $K = \mathbb{F}_p$ ist, dann gibt es nur endlich viele Abbildungen $\mathbb{F}_p \to \mathbb{F}_p$ aber unendlich viele Polynome.

Wir definieren das charakteristische Polynom also neu, als Determinante der Matrix $XE_n - A$, deren Einträge im Polynomring $K[X]$ liegen:

$$P_A \det(XE_n - A)$$

Determinanten lassen sich in der Tat wie in Kapitel 7 allgemeiner für Matrizen mit Einträgen in beliebigen kommutativen Ringen definieren.

Die Funktion $\lambda \mapsto \det(\lambda E_n - A)$ ist dann genau das Bild des Auswertehomomorphismus aus Bemerkung 9.3.5.3 und wir schreiben $\tilde{P}_A$ für diese Funktion.

Definition 9.4.1. Sei $A \in M(n \times n, R)$ für einen kommutativen Ring R. Dann ist die *Determinante* von A definiert als

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$$

Vgl. Satz 7.5.7. Die Determinante erfüllt Eigenschaften (D1-3) sowie die Eigenschaften aus Satz 7.2.3 und lässt sich auch durch Zeilen- oder Spaltenentwicklung bestimmen. ⟨⟨Es ist eine gute Übung (aber nicht auf dem Übungsblatt), sich zu vergewissern, das dies gilt.⟩⟩

Allerdings funktioniert der Gauß'sche Algorithmus ohne multiplikative Inverse nicht mehr ohne weiteres, und nicht jede Matrix mit Determinante ungleich 0 ist invertierbar (betrachte zum Beispiel $2 \in M(1 \times 1, \mathbb{Z})$).

Definition 9.4.2. Sei $A \in M(n \times n, K)$ Wir definieren das *charakteristische Polynom*

$$P_A(X) = \det(XE_N - A) \in K[X]$$

als Determinante der Matrix $XE_n - A \in M(n \times n, K[X])$, deren Einträge Elemente im Polynomring $K[X]$ sind.

Sei V ein endlichdimensionaler Vektorraum und sei $f \in \mathrm{End}_K(V)$. Wir definieren $P_f(X) = P_M(X)$ für eine darstellende Matrix von f.

Satz 9.4.3. *Das charakteristische Polynom eines Endomorphismus f ist wohldefiniert.*

Dies ist nicht klar, denn mit den oben diskutierten Einschränkungen der Determinante über R gilt unser Beweis der Multiplikativität der Determinantenfunktion nicht mehr!

Es gibt aber verschiedene Möglichkeiten, das gewünschte Ergebnis herzuleiten.

Beweis:. Sei S die Basiswechselmatrix so dass wir P_A und $P_{SAS^{-1}}$ vergleichen wollen.

Unser Beweis des Multiplikationssatzes 7.4.1 gilt (mit minimaler Abänderung), solange eine der beiden Faktoren eine Matrix mit Einträgen in einem Körper K ist, während die andere Matrix Eiträge in einer K-Algebra hat! Wir schauen hierzu den Beweis von Satz 7.4.1 noch einmal sorgfältig an: Angenommen $A \in M(n \times n, K)$ dann können wir A als Produkte von Elementarmatrizen schreiben und verfolgen, was diese mit der Determinante der Matrix B machen. Wir überspringen nur den letzten (überflüssigen!) Schritt, der B als Produkt von Elementarmatrizen schreibt, was mit $B \in M(n \times n, K[X]))$ nicht mehr gilt.

Außerdem rechnen wir leicht nach, dass $S^{-1}XE_nS = XS^{-1}E_nS = XE_n$ für $S \in M(n \times n, K)$. Daraus folgt

$$\begin{aligned}\det(XE_n - SAS^{-1}) &= \det(S(S^{-1}XE_nS - A)S^{-1}) \\ &= \det S \det(S^{-1}XE_nS - A) \det(S^{-1}) \\ &= \det(XE_n - A).\end{aligned}$$ □

Bemerkung 9.4.4. Es folgen noch einige Alternativen zu diesem Beweis:

1. $K[X]$ ist ein nullteilerfreier kommutativer Ring und wir können ihn in einen Quotientenkörper aus formalen Brüchen einbetten, genauso wie wir $\mathbb{Z}$ in $\mathbb{Q}$ einbetten, siehe Übung 5.1. Da nun $\det(XE_n - A)$ im Quotientenkörper für ähnliche Matrizen gleich ist, gilt das auch in $K[X]$.

2. Angenommen unser Körper K hat unendlich viele Elemente. Dann haben A und SAS^{-1} die gleiche charakteristische Funktion und wegen Korollar 9.3.13 auch das gleiche charakteristische Polynom.

   Jeder endliche Körper K lässt sich ohne allzu viel Mühe in einen unendlichen Körper einbetten, siehe Übung 5.1.

3. Sei R ein beliebiger kommutativer Ring, zum Beispiel $K[T]$. Dann gilt $\det(AB) = \det(A)\det(B)$ für $A, B \in M(n \times n, R)$. Für einen sorgfältigen Beweis müssten wir uns die Theorie der Matrizen über allgemeinen kommutativen Ringen genauer anschauen, dazu haben wir keine Zeit.

Definition 9.4.5. Ist $\lambda \in K$ Eigenwert der Matrix $A \in M(n \times n, K)$, so heißt die Vielfachheit der Nullstelle λ des charakteristischen Polynom die *algebraische Vielfachheit* des Eigenwerts λ:

$$\mu_{alg}(A, \lambda) := \mu(\lambda, P_A(X)) \ .$$

Für einen Endomorphisums $f \in \mathrm{End}_K(V)$ definieren wir genauso $\mu_{alg}(f, \lambda) = \mu(\lambda, P_f(X))$.

Bemerkung 9.4.6. Geometrische und algebraische Vielfachheit eines Eigenwerts einer Matrix können verschieden sein. Beispiel:

$$A = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix} \qquad \text{mit } \lambda_0 \in K$$

Dann ist das charakteristische Polynom $P_A(X) = (X - \lambda_0)^2$, also $\mu_{alg}(A, \lambda_0) = 2$, aber $\mu_{geo}(A, \lambda_0) = 1$, da A nur einen Eigenvektor hat.

Lemma 9.4.7. *Sei $A \in M(n \times n, K)$ und λ Eigenwert von A. Dann gilt*

$$1 \leq \mu_{geo}(A, \lambda) \leq \mu_{alg}(A, \lambda) \; .$$

Beweis:. Sei $k := \mu_{geo}(A, \lambda)$. Ergänze eine geordnete Basis $(v_1, \ldots, v_k)$ des Eigenraums $\mathrm{Eig}(A, \lambda)$ zu einer geordneten Basis $\mathcal{B} = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ von K^n. Für $S^{-1} = (v_1, \ldots, v_n)$ gilt $SAS^{-1}e_i = SAv_i = S(\lambda v_i) = \lambda e_i$ für $i = 1, \ldots, k$ und daher

$$SAS^{-1} = \begin{pmatrix} \lambda & & 0 & * \\ & \ddots & & * \\ 0 & & \lambda & * \\ 0 & \cdots & 0 & * \end{pmatrix}$$

wobei die Symbole * für Blockmatrizen stehen.

Nach Satz 9.4.3 haben ähnliche Matrizen das gleiche charakteristische Polynom. Da Korollar 7.5.8 direkt aus der Leibniz'schen Formel folgt, gilt es auch für Matrizen in $M(n \times n, R)$ und wir berechnen:

$$P_A(X) = P_{SAS^{-1}}(X) \overset{7.5.8}{=} (X - \lambda)^k \det\left(XE_{n-k} - D\right)$$

Daher ist $\mu_{alg}(A, \lambda) \geq k = \mu_{geo}(A, \lambda)$. □

Satz 9.4.8. *Sei V ein n–dimensionaler Vektorraum und sei $f : V \to V$ ein Endomorphismus. Seien $\lambda_1, \ldots, \lambda_N$ die paarweise verschiedenen Eigenwerte von f. Dann sind die folgenden Aussagen äquivalent:*

1. *f ist diagonalisierbar.*
2. *$P_f(X)$ zerfällt in Linearfaktoren und $\mu_{geo}(f, \lambda) = \mu_{alg}(f, \lambda)$ für alle Eigenwerte λ von f.*
3. $$\sum_{i=1}^{N} \mu_{geo}(f, \lambda_i) = n$$
4. *Es gilt die folgenden* Eigenraumzerlegung

$$V = \oplus_{i=1}^{N} \mathrm{Eig}(f, \lambda_i) \; ,$$

d.h. jedes $v \in V$ kann man eindeutig in der Form

$$v = v_1 + v_2 + \ldots + v_N$$

mit $v_i \in \mathrm{Eig}(f, \lambda_i)$ schreiben.

Beweis:. Wir müssen nur vier Implikationen zeigen.

1. $\Rightarrow$ 2. Wähle eine geordnete Basis $\mathcal{B}$ von Eigenvektoren:

$$\begin{aligned} v_1, \ldots, v_{k_1} \quad & \text{zum Eigenwert } \lambda_1, \text{ also } k_1 = \mu_{geo}(f, \lambda_1) \\ v_{k_1+1}, \ldots, v_{k_1+k_2} \quad & \text{zum Eigenwert } \lambda_2, \text{ also } k_2 = \mu_{geo}(f, \lambda_2) \\ \ldots & \end{aligned}$$

Dann ist

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & 0 & \\ & & & \lambda_2 & & \\ & & & & \ddots & \\ & & 0 & & \lambda_2 & \\ & & & & & \ddots \end{pmatrix}$$

Für das charakteristische Polynom folgt

$$P_f(X) = P_{M_{\mathcal{B}}^{\mathcal{B}}(f)}(X) = (X - \lambda_1)^{k_1} \ldots (X - \lambda_N)^{k_N} ,$$

also

$$\mu_{geo}(f, \lambda_i) = k_i = \mu_{alg}(f, \lambda_i) .$$

2. $\Rightarrow$3. Zerfällt P_f in Linearfaktoren, so gilt

$$n = \operatorname{grad} P_f = \sum_{i=1}^{N} \mu_{alg}(f, \lambda_i) .$$

Aus der weiteren Annahme $\mu_{geo}(f, \lambda_i) = \mu_{alg}(f, \lambda_i)$ folgt sofort 3.

3. $\Rightarrow$ 4. Setze $W := \operatorname{Eig}(f, \lambda_1) + \ldots + \operatorname{Eig}(f, \lambda_N)$.
Die Summe ist direkt: denn gelte

$$w = w_1 + \ldots + w_N = w_1' + \ldots + w_N'$$

mit $w_i, w_i' \in \operatorname{Eig}(f, \lambda_i)$, so liegt $w_i - w_i' \in \operatorname{Eig}(f, \lambda_i)$ und es gilt $0 = \sum_{i=1}^{N}(w_i - w_i')$. Aus Satz 9.2.10 folgt $w_i - w_i' = 0$. Wegen

$$\dim W = \sum_{i=1}^{N} \dim \operatorname{Eig}(f, \lambda_i) = \sum_{i=1}^{N} \mu_{geo}(f, \lambda_i) = n = \dim V$$

folgt auch $V = W$, also 4.

4. $\Rightarrow$ 1. Wähle Basen von $\operatorname{Eig}(f, \lambda_i)$, die zusammen eine Basis von V aus Eigenvektoren von f ergeben. □

Korollar 9.4.9. *Sei $A \in M(n \times n, K)$. Dann sind äquivalent:*

1. *A ist diagonalisierbar, d.h. ähnlich zu einer Diagonalmatrix.*
2. *$P_A(X)$ zerfällt in Linearfaktoren und $\mu_{alg}(A, \lambda) = \mu_{geo}(A, \lambda)$ für alle Eigenwerte λ von A.*

Beweis:. Wende Satz 9.4.8 auf die lineare Abbildung $f: K^n \to K^n$ mit $f(x) = Ax$ für $x \in K^n$ an. □

Auch nach diesem Kriterium ist also die Matrix aus Beispiel 9.4.6 nicht diagonalisierbar.

Bemerkung 9.4.10. Hier ist eine andere nützliche Anwendung des charakteristishen Polynoms: der Koeffizient a_{n-1} von X^{n-1} in P_A ist gleich der negativen Spur von A, $-\mathrm{Tr}\,(A)$. Wenn $P_A = \prod_i (X - \lambda_i)$ in Linearfaktoren zerfällt, ist er aber auch gleich der Summe der Eigenwerte (gezählt mit Ihrer algebraischen Multiplizität), $a_{n-1} = -\sum \lambda_i$ durch Ausmultiplizieren.

Das gibt uns einen einfachen Test, ob wir die Eigenwerte von A korrekt bestimmt haben: Wir summieren alle Eigenwerte mit ihren Multiplizitäten auf und vergleichen mit der Spur!

Wir werden später noch ein weiteres Kriterium für Diagonalisierbarkeit kennenlernen. Aber erst einmal behandel wir die folgende Frage: gegeben seien *zwei* Endomorphismen $f, g: V \to V$. Wann kann man sie *gleichzeitig* diagonalisieren, d.h. wann existiert *eine* Basis $\mathcal{B}$ von V, so dass die *beiden* darstellenden Matrizen $M_{\mathcal{B}}(f)$ und $M_{\mathcal{B}}(g)$ Diagonalmatrizen sind?

Der folgende Begriff wird im Beweis von Satz 9.4.12 und darüber hinaus nützlich sein:

Definition 9.4.11. Es sei f ein Endomorphismus eines K-Vektorraums V. Ein Untervektorraumraum $W \subset V$ mit $f(W) \subset W$ heißt ein *f–invarianter Untervektorraum.*

- Beispiele für f-invariante Unterräume sind die trivialen Untervektorräume $W = V$ und $W = \{0\}$ sowie $\mathrm{Eig}(f, \lambda)$ für jedes $\lambda \in K$.
- Die Summe f-invarianter Untervektorräume ist f-invariant: Für $w_1 + w_2$ mit $w_i \in W_i$ gilt

$$f(w_1 + w_2) = f(w_1) + f(w_2) \in W_1 + W_2 \,.$$

- Der Schnitt f-invarianter Untervektorräume ist ein f-invarianter Untervektorraum.

Satz 9.4.12. *Sei V ein endlich–dimensionaler K–Vektorraum und seien $f, g: V \to V$ diagonalisierbare Endomorphismen. Dann sind äquivalent:*

1. *f und g sind gleichzeitig diagonalisierbar.*
2. *f und g kommutieren, d.h. $f \circ g = g \circ f$.*

Beweis:1. $\Rightarrow$2. Sei $\mathcal{B}$ eine Basis von V, in der die beiden Matrizen $M_{\mathcal{B}}(f)$ und $M_{\mathcal{B}}(g)$ Diagonalmatrizen sind und rechnen mit Matrizen

$$M_{\mathcal{B}}(f \circ g) = M_{\mathcal{B}}(f) \cdot M_{\mathcal{B}}(g) = M_{\mathcal{B}}(g) \cdot M_{\mathcal{B}}(f) = M_{\mathcal{B}}(g \circ f) \,,$$

wobei die zweite Gleichung daraus folgt, dass die darstellenden Matrizen Diagonalmatrizen sind und Diagonalmatrizen kommutieren. Hieraus folgt $f \circ g = g \circ f$, da $M_{\mathcal{B}}: \mathrm{End}(V) \to M(n \times n, K)$ ein Isomorphismus von K-Vektorräumen ist (wir benötigen hier die Injektivität).

2.$\Rightarrow$1. Da f und g diagonalisierbar sind, gibt es die beiden Eigenraumzerlegungen

$$V = \bigoplus_{\lambda \in K} \mathrm{Eig}(f, \lambda) \,.$$
$$V = \bigoplus_{\mu \in K} \mathrm{Eig}(g, \mu) \,.$$

Aus der Kommutativität von f und g folgt, dass alle Eigenräume in diesen Zerlegungen unter f und g gleichzeitig invariant sind, denn wenn $v \in \mathrm{Eig}(f, \lambda)$, dann gilt $f(v) = \lambda v$.

Daraus folgt aber $f(g(v)) = g(f(v)) = g(\lambda v) = \lambda g(v)$. Also ist auch $g(v) \in \mathrm{Eig}(f, \lambda)$. Wir behaupten

$$V = \bigoplus_{\lambda,\mu} \mathrm{Eig}(f, \lambda) \cap \mathrm{Eig}(g, \mu) \ .$$

Da dies eine Zerlegung in gemeinsame Eigenräume von f und g ist, folgt die Behauptung.

Es gilt $V = \oplus_\lambda \mathrm{Eig}(f, \lambda)$, da f diagonalisierbar ist, und es reicht aus, für jeden festen Eigenwert $\lambda \in K$ von f die Summerzerlegung

$$\mathrm{Eig}(f, \lambda) = \sum_{\mu_i} \mathrm{Eig}(f, \lambda) \cap \mathrm{Eig}(g, \mu_i)$$

zu zeigen. Diese Summe ist automatisch direkt da Eigenvektoren für verschieden Eigenwerte linear unabhängig sind wegen Satz 9.2.10.

Sei also $v \in \mathrm{Eig}(f, \lambda)$; da g diagonalisierbar ist, können wir v wie jeden Vektor aus V als Summe $v = v'_1 + v'_2 + \ldots + v'_m$ mit $v'_i \in \mathrm{Eig}(g, \mu_i)$ schreiben. Dann gilt

$$\begin{aligned} f(v) &= f(v'_1) + f(v'_2) + \ldots + f(v'_m) \\ &= \lambda v = \lambda v'_1 + \lambda v'_2 + \ldots + \lambda v'_m \ . \end{aligned}$$

Da der Eigenraum $\mathrm{Eig}(g, \mu_i)$ f-invariant ist gilt $f(v'_i) \in \mathrm{Eig}(g, \mu_i)$. Da die Zerlegung von $f(v) = \lambda v$ in Komponenten bezüglich der direkten Summe $V = \oplus \mathrm{Eig}(g, \mu_i)$ eindeutig ist, gilt für jedes i, dass $f(v'_i) = \lambda v'_i$. Also ist $v'_i \in \mathrm{Eig}(f, \lambda) \cap \mathrm{Eig}(g, \mu_i)$, was zu zeigen war. □

9.5 Trigonalisierbarkeit

Wir wollen nun noch sehen, was sich über einen Endomorphismus aussagen lässt, wenn wir nur wissen, dass sein charakteristisches Polynom vollständig in Linearfaktoren zerfällt, aber die geometrischen Vielfachheiten nicht kennen.

Lemma 9.5.1. *Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen K–Vektorraums V. Ist W ein f–invarianter Untervektorraum, so teilt das charakteristische Polynom $P_{f|W}$ das charakteristische Polynom P_f.*

Beweis:. Ergänze eine geordnete Basis $\mathcal{B}'$ von W zu einer geordneten Basis $\mathcal{B}$ von V. Die darstellende Matrix ist in dieser Basis

$$M_{\mathcal{B}}(f) = \begin{pmatrix} M_{\mathcal{B}'}\left(f_{|W}\right) & \vdots & * \\ \ldots\ldots\ldots\ldots & & \\ 0 & \vdots & A \end{pmatrix}$$

Es folgt mit $n := \dim_K V$ und $k := \dim_K W$ mit Korollar 7.5.8 (was auch für Matrizen mit Einträgen in $K[X]$ gilt, vgl. den Beweis von Lemma 9.4.7)

$$P_f = \det\left(XE_n - M_{\mathcal{B}}(f)\right) = \det\left(XE_k - M_{\mathcal{B}'}(f_{|W})\right) \cdot \det\left(XE_{n-k} - A\right) \quad \square$$

Definition 9.5.2. Sei $f : V \to V$ ein Endomorphismus eines n–dimensionalen K–Vektorraums V.

1. Eine *Fahne* von V ist eine Kette von Untervektorräumen

   $$\{0\} = V_0 \subset V_1 \subset \ldots \subset V_n = V$$

   mit $\dim V_r = r$ für alle $r = 0, \ldots, n$.

2. Eine Fahne von V heißt f–invariant, falls alle Untervektorräume f-invariant sind, also für alle $r = 0, \ldots, n$ die Inklusion $f(V_i) \subset V_i$ gilt.

Beispiel 9.5.3. Jede geordnete Basis $(v_1, \ldots, v_n)$ eines Vektorraums V liefert mit $V_r := \operatorname{span}_K\{v_1, \ldots, v_r\}$ eine Fahne von V.

Satz 9.5.4. *Sei $f \in \operatorname{End}(V)$, $\dim_K V = n$. Dann sind äquivalent:*

1. *Es existiert eine f–invariante Fahne von V.*
2. *Es gibt eine geordnete Basis $\mathcal{B}$ von V, in der die darstellende Matrix $M_{\mathcal{B}}(f)$ eine obere Dreiecksmatrix ist.*

Beweis: 2. $\Rightarrow$1. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ und $A = (a_{ij}) = M_{\mathcal{B}}(f)$ eine obere Dreiecksmatrix. Betrachte die Fahne $V_i = \operatorname{span}_K\{v_1, \ldots, v_i\}$ wie in Beispiel 9.5.3. Wegen $a_{rs} = 0$ für $r > s$ gilt

$$f(v_i) = \sum_{j=1}^{n} a_{ji} v_j = \sum_{j=1}^{i} a_{ji} v_j$$

und damit ist $f(V_i) \subset V_i$ und die Fahne ist f–invariant.

1. $\Rightarrow$2. Wir konstruieren aus einer f–invarianten Fahne folgendermaßen eine geordnete Basis: sei $\{v_1\}$ eine Basis von V_1; wegen $V_1 \subset V_2$ ist es möglich, dies zu einer Basis $\{v_1, v_2\}$ von V_2 zu ergänzen (Basisergänzungssatz!). So fahren wir fort und ergänzen eine geordnete Basis $(v_1, v_2, \ldots, v_i)$ von V_i zu einer geordneten Basis $(v_1, v_2, \ldots, v_i, v_{i+1})$ von V_{i+1} für jedes $i < n$.

Für alle $i \in \{1, \ldots, n\}$ gilt $f(v_i) \in f(V_i) \subset V_i$; also gilt

$$f(v_i) = \sum_{j=1}^{i} a_{ji} v_j$$

für irgendwelche $a_{ji} \in K$ und (a_{ij}) ist die darstellende Matrix $M_{\mathcal{B}}(f)$. Sie erfüllt $a_{ij} = 0$ für $i > j$, ist also eine obere Dreiecksmatrix.

□

Definition 9.5.5. 1. Sei f Endomorphismus eines endlich–dimensionalen K–Vektorraums V. f heißt *trigonalisierbar*, falls es eine geordnete Basis $\mathcal{B}$ von V gibt, so dass $M_{\mathcal{B}}(f)$ eine obere Dreiecksmatrix ist.

2. Eine Matrix $A \in M(n \times n, K)$ heißt *trigonalisierbar*, falls sie zu einer oberen Dreiecksmatrix ähnlich ist.

Satz 9.5.6. *Sei f Endomorphismus eines endlich–dimensionalen K–Vektorraums V. Dann sind äquivalent:*

1. *f ist trigonalisierbar.*
2. *Das charakteristische Polynom P_f zerfällt vollständig in Linearfaktoren.*

Die diagonalen Einträge der oberen Dreiecksmatrix sind die Eigenwerte von A und λ kommt $\mu_{alg}(A, \lambda)$ mal vor.

Beweis:1. $\Rightarrow$2. Wir rechnen mit Satz 9.4.3 das charakteristische Polynom in einer geordneten Basis $\mathcal{B}$ aus, in der $M_{\mathcal{B}}(f)$ eine obere Dreiecksmatrix D ist:

$$P_f(X) = P_D(X) = \det \begin{pmatrix} X-\lambda_1 & & & \\ & X-\lambda_2 & * & \\ & & \ddots & \\ & & & X-\lambda_n \end{pmatrix} \overset{7.2.3.5}{=} \prod_{i=1}^{n}(X-\lambda_i)$$

2. $\Rightarrow$1. Vollständige Induktion nach $n = \dim_K V$. Für $n = 1$ ist nichts zu zeigen. Sei $n > 1$; das charakteristische Polynom zerfalle,

$$P_f(X) = \prod_{i=1}^{n}(X-\lambda_i)\ .$$

Wähle einen Eigenvektor $v_1 \in V$ zum Eigenwert λ_1 und ergänze zu einer Basis $\widetilde{\mathcal{B}} = (v_1, \ldots, v_n)$ von V. Dann ist

$$M_{\widetilde{\mathcal{B}}}(f) = \left(\begin{array}{c|c} \lambda_1 & a_2 \ldots a_n \\ \hline 0 & \\ \vdots & \tilde{A} \\ 0 & \end{array}\right)$$

Es folgt wie in Lemma 9.5.1

$$P_f(X) = (X-\lambda_1)P_{\tilde{A}}(X)\ .$$

Aus der Eindeutigkeit der Polynomdivision in Lemma 9.3.11 folgt, dass auch das charakteristische Polynom $P_{\tilde{A}}(X)$ in Linearfaktoren zerfällt. Nach Induktionsvoraussetzung ist die $(n-1)\times(n-1)$-Matrix $\widetilde{A}$ ähnlich zu einer oberen Dreiecksmatrix $\widetilde{D}$, d.h. es gibt eine invertible $(n-1)\times(n-1)$-Matrix $\tilde{S}$ mit:

$$\widetilde{S}\widetilde{A}\widetilde{S}^{-1} = \widetilde{D}$$

Setze

$$S := \left(\begin{array}{c|c} 1 & 0 \ldots 0 \\ \hline 0 & \\ \vdots & \widetilde{S} \\ 0 & \end{array}\right) \in M(n\times n, K)$$

mit inverser Matrix

$$S^{-1} = \left(\begin{array}{c|c} 1 & 0 \ldots 0 \\ \hline 0 & \\ \vdots & \widetilde{S}^{-1} \\ 0 & \end{array}\right) \in M(n\times n, K)$$

und rechne:

$$
\begin{aligned}
SM_{\widetilde{\mathcal{B}}}(f)S^{-1} &= \left(\begin{array}{c|c} 1 & 0\dots 0 \\ \hline 0 & \\ \vdots & \widetilde{S} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} \lambda_1 & a_2\dots a_n \\ \hline 0 & \\ \vdots & \widetilde{A} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0\dots 0 \\ \hline 0 & \\ \vdots & \widetilde{S}^{-1} \\ 0 & \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda_1 & a_2\dots a_n \\ \hline 0 & \\ \vdots & \widetilde{S}\widetilde{A} \\ 0 & \end{array}\right)\left(\begin{array}{c|c} 1 & 0\dots 0 \\ \hline 0 & \\ \vdots & \widetilde{S}^{-1} \\ 0 & \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda_1 & *\dots * \\ \hline 0 & \\ \vdots & \widetilde{S}\widetilde{A}\widetilde{S}^{-1} \\ 0 & \end{array}\right) .
\end{aligned}
$$

Dies ist eine obere Dreiecksmatrix, die gleich $M_{\mathcal{B}}(f)$ ist, wobei $\mathcal{B}$ aus den Zeilen der Matrix $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} S^{-1}$ besteht.

Es ist für den Beweis nicht notwendig, aber die Herleitung der Basis $\mathcal{B}$ geht wie folgt: Wir wissen aus Bemerkung 5.4.5.3, dass $T_{\mathcal{E}}^{\tilde{\mathcal{B}}}$ die Matrix $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ ist und die Elemente von $\mathcal{B}$ genau die Spalten von $T_{\mathcal{E}}^{\mathcal{B}} = T_{\mathcal{E}}^{\tilde{\mathcal{B}}} T_{\tilde{\mathcal{B}}}^{\mathcal{B}}$ sind. Aber die Basiswechselmatrix $T_{\tilde{\mathcal{B}}}^{\mathcal{B}}$ ist genau S^{-1}.

Die Charakterisierung der diagonalen Einträge folgt aus dem Beweis aus dem Beweis: Jeder Eigenwert taucht mit seiner algebraischen Vielfachheit auf. □

Korollar 9.5.7. *Jede Matrix $A \in M(n \times n, \mathbb{C})$ mit komplexen Einträgen ist trigonalisierbar.*

Beweis:. Nach dem Fundamentalsatz der Algebra 9.3.19.3 zerfällt das charakteristische Polynom in Linearfaktoren und das Ergebnis folgt aus Satz 9.5.6. □

Betrachtung 9.5.8. Der Beweis von Satz 9.5.6 liefert auch ein Rechenverfahren zur Triagonalisierung.

Sei $A \in M(n \times n, K)$ mit einem charakteristischem Polynom $P_A(X) = \prod_{i=1}^{n}(X - \lambda_i)$, das in Linearfaktoren zerfällt.

1. Bestimme, zum Beispiel mit dem Gauß-Algorithmus, einen Eigenvektor

$$v_1 = \begin{pmatrix} v_{11} \\ \vdots \\ v_{1n} \end{pmatrix}$$

zum Eigenwert λ_1. Weil für den Eigenvektor $v_1 \neq 0$ gilt, gibt es ein i, so dass die i-te Komponente $v_{1i} \neq 0$ von v_1 ungleich Null ist; wähle ein solches und setze

$$S_1 = (v_1, e_1, \dots, \widehat{e_i}, \dots, e_n)^{-1} \in GL(n, K) .$$

Hierbei zeigt das Symbol "Hut" an, dass der Vektor e_i der Standardbasis ausgelassen wird. Dann gilt $(S_1)^{-1}e_1 = v_1$, denn das Bild von e_1 ist der erste Spaltenvektor von $(S_1)^{-1}$. Dann ist

$$S_1 A S_1^{-1} e_1 = S_1 A v_1 = S_1 \lambda_1 v_1 = \lambda_1 e_1 ,$$

also

$$S_1 A S_1^{-1} = \left(\begin{array}{c|c} \lambda_1 & * \ldots * \\ \hline 0 & \\ \vdots & A_2 \\ 0 & \end{array}\right)$$

2. Berechne einen Eigenvektor $\widetilde{v}_2 = \begin{pmatrix} v_{22} \\ \vdots \\ v_{2n} \end{pmatrix} \in K^{n-1}$ der $(n-1) \times (n-1)$-Matrix A_2 zum Eigenwert λ_2. Ergänze den Vektor $\widetilde{v}_2 \in K^{n-1}$ durch Zufügen einer beliebigen ersten Komponente zu einem Vektor $v_2 \in K^n$. Wähle $j \geq 2$ so dass $v_{2,j} \neq 0$ und setze

$$S_2 = (e_1, v_2, e_2, \ldots, \widehat{e_j}, \ldots, e_n)^{-1} \in GL(n, K)$$

Dann ist

$$S_2 S_1 A S_1^{-1} S_2^{-1} e_1 = S_2 S_1 A S_1^{-1} e_1 = \lambda_1 S_2 e_1 = \lambda_1 e_1$$

$$S_2 S_1 A S_1^{-1} S_2^{-1} e_2 = S_2 S_1 A S_1^{-1} v_2 = S_2 \begin{pmatrix} * \\ | \\ \lambda_2 v_2 \\ | \end{pmatrix} = \begin{pmatrix} * \\ \lambda_2 \\ 0 \\ \vdots \end{pmatrix}$$

Also ist

$$S_2 S_1 A S_1^{-1} S_2^{-1} = \left(\begin{array}{cc|c} \lambda_1 & * & * \ldots * \\ 0 & \lambda_2 & * \ldots * \\ \hline & 0 & \\ & & A_3 \\ & & \end{array}\right)$$

Nach $n-1$ Schritten liefert der Algorithmus eine obere Dreiecksmatrix.

Beispiel: Die Matrix

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

hat, wie eine kurze Rechnung zeigt, das charakteristische Polynom $P_A(X) = (X-2)^3$, also den einzigen Eigenwert 2.

1. Wir bestimmen einen Eigenvektor zum Eigenwert 2:

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

und erhalten

$$S_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$S_1 A S_1^{-1} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 4 & 2 \\ 0 & -2 & 0 \end{pmatrix}$$

2. Die Eigenwerte von A_2 sind genau die Eigenwerte von A ohne λ_1, hier also wieder 2. Ein Eigenvektor zum Eigenwert 2 von $\begin{pmatrix} 4 & 2 \\ -2 & 0 \end{pmatrix}$ ist $\widetilde{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Setze

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{und} \quad S_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Man findet damit die zu A ähnliche obere Dreiecksmatrix

$$S_2 S_1 A S_1^{-1} S_2^{-1} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix} .$$

Diese ist natürlich nicht eindeutig bestimmt.

9.6 Minimalpolynom und Satz von Cayley-Hamilton

Theorem 9.6.1 (Cayley-Hamilton)**.** *Sei $A \in M(n \times n, K)$. Dann gilt für das charakteristische Polynom*

$$\widetilde{P_A}(A) = 0_n.$$

Wir erinnern uns an Bemerkung 9.3.5: $P_A \in K[X]$ wird vom Einsetzungshomomorphismus zum Element $A \in M(n \times n, K)$ auf einen eindeutig bestimmten Wert $\phi_A(p_A)$ abgebildet. Wir schreiben hier $\widetilde{P_A}$ für das charakteristishe Polynom betrachtet als Abbildung von $M(n \times n, K)$ nach $M(n \times n, K)$. (Wir setzen hier im letzten Teil von Bemerkung 9.3.5 $A = M(n \times n, K)$.

Konkret bedeutet dies alles, dass für $P_A = \sum_{i=0}^{n} c_i X^i$ die Matrix $\widetilde{P_A}(A) = \sum_{i=0}^{n} c_i A^i$ gleich 0 ist.

Auch die Schreibweise $P_A(A)$ statt $\widetilde{P}_A(A)$ ist gebräuchlich für die Anwendung des Einsetzungshomomorphismus auf P_A, vgl. wider Bemerkung 9.3.5. Wir betonen hier mit der Tilde den Wechsel vom abstrakten Polynom zur dazugehörigen Funktion.

Bemerkung 9.6.2. Der Satz von Cayley-Hamilton ist etwas subtiler, als er auf den ersten Blick aussieht. Ein naiver, aber falscher Beweis geht wie folgt: $\tilde{P}_A(A) = \det(A \cdot E_n - A) = \det(0) = 0$. Aber wir betrachten die Determinante von $XE_n - A$ als Element des Polynomrings $K[X]$, und wenn wir $X = A$ einsetzen dann erhalten wir einen Wert im Ring der Matrizen. Wir können uns eine Matrix mit Werten in Matrizen vorstellen, für die wir dann A einsetzen können.

Wir vergleichen also gewissermaßen nicht A mit $A \cdot E_n$ in $M(n \times n, K)$ sondern die Matrix mit Einträgen $a_{ij}E_n \in M(n \times n, M(n \times n, K))$ mit der diagonalen Matrix in der alle diagonalen Einträge gleich $A \in M(n \times n, K)$ sind.

$$\begin{pmatrix} a_{11}E_n & \cdots & a_{1n}E_n \\ \vdots & \ddots & \vdots \\ a_{n1}E_n & \cdots & a_{nn}E_n \end{pmatrix} \neq \begin{pmatrix} A & & 0 \\ & \ddots & \\ 0 & & A \end{pmatrix}$$

Deshalb haben wir die rechte Seite als 0_n geschrieben, um zu betonen, dass dies nicht die $0 \in K$ sondern die Nullmatrix in $M(n \times n, K)$ ist.

Vergleichen wir mit dem analogen Situation, wenn wir die Determinante durch die Spur ersetzen. Wir erhalten ein (lineares) Polynom $t \mapsto \mathrm{Tr}\,(tE_n - A) = nt - \mathrm{Tr}\,(A)$. Der naive "Beweis" $\mathrm{Tr}\,(A \cdot E_n - A) = \mathrm{Tr}\,(0) = 0$ scheint genauso zu funktionieren, aber wenn wir jetzt $t = A$ einsetzen erhalten wir $nA - \mathrm{Tr}\,(A)E_n = 0$, und das gilt nur, wenn A eine Skalarmatrix ist.

Beweis:. Wir beweisen den Satz nur für Unterkörper von $\mathbb{C}$, also insbesondere für $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ selbst.

Wir nehmen zuerst an, dass $K = \mathbb{C}$ gilt. Dann ist A trigonalisierbar und es gibt eine obere Dreiecksmatrix D und $S \in GL(n, K)$ mit $D = SAS^{-1}$. Es reicht nun, den Satz für D zu beweisen. Denn $P_A = P_D$ wegen Satz 9.4.3 und wenn $\sum_{i=0}^{n} c_i D^i = 0$ gilt dann haben wir auch

$$\sum_i c_i (SAS^{-1})^i = \sum_i c_i S A^i S^{-1} = S \sum_i c_i A^i) S^{-1} = 0$$

Wir nehmen also ohne Beschränkung der Allgemeinheit an, dass A eine obere Dreiecksmatrix mit diagonalen Einträgen λ_i ist und $P_A = (X - \lambda_1) \cdots (X - \lambda_n)$.

Wir betrachten $\tilde{P}_A(A) = (A - \lambda_1 E_n) \cdots (A - \lambda_n E_n)$. Der i-te Faktor hat Eintrag 0 an Position (i, i) und für alle (j, i) mit $j > i$.

$$P_A(A) = \begin{pmatrix} 0 & * & * \\ 0 & \lambda_2 - \lambda_1 & * \\ 0 & 0 & * \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 - \lambda_2 & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{pmatrix} \cdots \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 0 \end{pmatrix}$$

Wir sehen induktiv, dass die ersten i Zeilen des Produkts der ersten i Faktoren 0 sind. Damit ist das Produkt aller n Faktoren gleich 0.

Sei nun $K \subset \mathbb{C}$ ein Unterkörper (zum Beispiel $K = \mathbb{R}$). Dann können wir $A \in M(n \times n, K)$ als komplexe Matrix auffassen, und wir haben gerade gezeigt, dass $\tilde{P}_A(A) = 0 \in M(n \times n, \mathbb{C})$. Aber dann ist $\tilde{P}_A(A)$ auch 0 in $M(n \times n, K)$, denn wir können $\tilde{P}_A(A)$ ja über jedem Körper berechnen, der A und alle Koeffizienten von P_A enthält. □

Bemerkung 9.6.3. Um den Satz über einem beliebigen Körper zu zeigen, gibt es zwei Möglichkeiten: Wir beweisen direkt algebraisch, dass der Satz gilt (siehe zum Beispiel Satz 6.100 im Buch von Bär), oder wir erweitern unseren Beweis auf beliebige Körper. Dazu müssen wir zeigen, dass für jeden Körper K und jedes Polynom $f \in K[X]$ ein Körper K' existiert, so dass $K \subset K'$ ein Unterkörper ist und f in $K'[X]$ in Linearfaktoren zerfällt. Dies zeigen Sie in Algebra.

Es gilt sogar, dass K Unterkörper von einem Körper $\tilde{K}$ ist, in dem der Fundamentalsatz der Algebra gilt und jedes Polynom in Linearfaktoren zerfällt. So ein $\tilde{K}$ heißt *algebraischer Abschluss* von K . Die Existenz zeigt man auch in der Algebra (mit dem Zorschen Lemma).

Wir haben damit gezeigt, dass $A \in M(n \times n, K)$ ein Polynom erfüllt. Das ist allerdings nicht das einzige Polynom. Wenn zum Beispiel A einfacheine Diagonalmatirx λE_n ist, dann gilt schon $A - \lambda E_n = 0$, A erfüllt also das Polynom $X - \lambda$.

Satz 9.6.4. *Gegeben $A \in M(n \times n, K)$ gibt es ein Polynom $\mu_A \in K[T]$, so dass für jedes Polynom q mit $\tilde{q}(A) = 0$ gilt $\mu_A \mid q$. Dieses Polynom ist das Polynom mit dem niedrigsten positiven Grad, das $\tilde{\mu}_A(A) = 0$ erfüllt. Es ist eindeutig, wenn wir zudem verlangen, dass der höchste Koeffizient 1 ist.*

Wir setzen den Grad von μ positiv um das Nullpolynom auszuschließen.

Beweis:. Wegen Satz 9.6.1 gibt es ein Polynom, das A erfüllt. Wir wählen nun μ, so dass es kein Polynom p mit kleinerem Grad gibt, das $\tilde{p}(A) = 0$ erfüllt.

Sei nun $\tilde{q}(A) = 0$ für irgendein Polynom q. Wir wenden die Division mit Rest aus Satz 9.3.9 an: Es gibt g, r so dass $q = g\mu + r$ und $\text{grad}(r) < \text{grad}(\mu)$. Aber $\tilde{r}(A) = \tilde{q}(A) - \tilde{g}(A)\tilde{\mu}(A) = 0$. Da μ minimal mit disere Eigenschaft ist, muss gelten $r = 0$ und $\mu \mid q$.

Wenn wir nun zwei Polynome μ, μ' mit der gewünschten Eigenschaft haben gilt $\mu = g\mu'$ und $\mu' = g'\mu$. Da μ und μ' den gleichen Grad haben, sind g und g' Polynome vom Grad 0, d.h. μ und μ' unterscheiden sich nur um einen Skalafaktor. □

Definition 9.6.5. Wir wählen das Polynom μ_A mit höchtem Koeffizienten 1 aus Satz 9.6.4 und nennen es das *Minimalpolynom* von A.

Bemerkung 9.6.6. Gegeben $S \in GL(n,K)$ und $p \in K[X]$ gilt $\tilde{p}(A) = 0$ genau dann wenn $S\tilde{p}(A)S^{-1} = p(SAS^{-1})$ gleich null ist.

Das Minimalpolynom ist also für ähnliche Matrizen gleich. Gegeben $f \in \mathrm{End}_K(V)$ können wir also das Minimalpolynom μ_f von f definieren als μ_A für irgendeine Matrix die f darstellt.

Alternativ können wir mit dem gleichen Beweis die folgende Variante von Satz 9.6.4 zeigen: Für $f \in \mathrm{End}_K(V)$, V endlichdimensional, gibt es ein Polynom $\mu_f \in K[X]$, so dass für jedes Polynom q mit $q(f) = 0$ gilt $\mu_f \mid -q = \mu_f$. Es ist eindeutig bis auf einen skalaren Faktor.

Im Beweis verwenden wir, dass Satz 9.6.1 auch für Endemorphismen gilt: Gegeben $f \in \mathrm{End}_K(V)$ gilt $\tilde{P}_f(f) = 0 \in \mathrm{End}_K(V)$.

Sie können sich vergewissern, dass beide Definitionen zum gleichen Mininmalpolynom für f führen.

Beispiele 9.6.7. Wir betrachten Minimalpolynom und charakteristisches Polynom in ein paar Beispielen.

A	μ_A	P_A
0_n	X	X^n
E_n	$X-1$	$(X-1)^n$
$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	X^2	X^2
$\begin{pmatrix} 0 & -7 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}$	$X^2(X-7)$	$X^2(X-7)^2$

Satz 9.6.8. *Sei $A \in M(n \times n, K)$. Dann haben das charakteristische Polynom $P_A(X)$ und das Minimalpolynom $\mu_A(X)$ dieselben Nullstellen. Die Vielfachheit einer Nullstelle $\lambda \in K$ als Nullstelle des Minimalpolynoms ist dabei kleiner als die (oder gleich der) Vielfachheit als Nullstelle des charakteristischen Polynoms.*

Beweis:.

- Nach Theorem 9.6.1 und Satz 9.6.4 gilt $P_A = g\mu_A$ mit $g \in K[X]$. Jede Nullstelle des Minimalpolynoms μ_A ist also auch Nullstelle des charakteristischen Polynoms P_A mit mindestens der gleichen Multiplizität.

- Sei $\lambda \in K$ eine Nullstelle von P_A, also ein Eigenwert. Wähle einen zugehörigen Eigenvektor v. Mit
$$\mu_A = X^m + \alpha_{m-1}X^{m-1} + \ldots + \alpha_0$$
rechnen wir mit der definierenden Eigenschaft $\widetilde{\mu}_A(A) = 0_n$ des Minimalpolynoms:
$$\begin{aligned} 0 = 0_n \cdot v = \widetilde{\mu_A}(A)v &= \left(A^m + \alpha_{m-1}A^{m-1} + \ldots + \alpha_0\right) v \\ &= \left(\lambda^m + \alpha_{m-1}\lambda^{m-1} + \ldots + \alpha_0\right) v = \widetilde{\mu_A}(\lambda)v\ . \end{aligned}$$
Hieraus folgt aber wegen $v \neq 0$, dass $\widetilde{\mu_A}(\lambda) = 0$. Also ist der Eigenwert λ auch Nullstelle des Minimalpolynoms. □

10 Die jordansche Normalform

10.1 Einführung

Wir widmen uns nun der Klassifizierung von Endomorphismen von endlich-dimensionalen Vektorräumen, bzw. der Klassifizierung von ähnlichen Matrizen.

Wenn A nicht diagonalsierbar ist wollen wir sie dennoch in einer aussagekräftige Form bringen.

Wir wissen schon, dass Matrizen wie $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ keine Basis aus Eigenvektoren haben. Aber wir können den Eigenvektore $v_1 = (1,0)^T$ mit dem Vektor $v_2 = (0,1)^T$ zu einer Basis ergänzen, der $(A - E_2)v_2 \neq 0$ aber $(A - E_2)^2 v_2 = 0$ erfüllt. Bemerkenswerterweise lassen sich alle nicht diagonalisierbaren Matrizen in eine solche Form bringen.

Definition 10.1.1. Eine Matrix

$$J = J(\lambda, k) = \begin{pmatrix} \lambda & 1 & 0 & & 0 \\ 0 & \lambda & 1 & & \\ & & \ddots & & \\ & & & \lambda & 1 \\ 0 & & & 0 & \lambda \end{pmatrix} \in M(k \times k, K)$$

heißt *Jordanblock* oder *Jordan-Matrix*. Es sind alle diagonalen Einträge gleich $\lambda \in K$, all Einträge in Position $(i, i+1)$ sind gleich 1 und alle anderen Einträge sind 0.

Satz 10.1.2. *Sei $f \in \mathrm{End}_K(V)$ und zerfalle P_f in Linearfaktoren. Dann gibt es eine Basis für V, so dass f von einer blockdiagonalen Matrix aus Jordan-Matrizen dargestellt wird, d.h.*

$$M(f)_{\mathcal{B}} = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix}$$

wobei $J_1, \ldots, J_m$ Jordan-Matrizen sind und die diagonalen Einträge von $M(f)_{\mathcal{B}}$ die Eigenwerte von f mit ihrer algebraischen Vielfachheit sind.

Diese Darstellung heißt jordansche Normalform*. Sie ist eindeutig bis auf Permutation der Jordan-Matrizen.*

Beispiel 10.1.3. Beispiel für eine Matrix in jordanscher Normalform:

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

Insbesondere müssen die verschiedenen Jordan-Matrizen keine unterschiedlichen Eigenwerte haben, und Jordanblöcke mit dem gleichen Eigenwert können unterschiedlich groß sein.

Die Jordan-Blöcke können auch Größe 1 haben! Wenn alle Jordan-Blöcke Größe 1 haben, dann gibt es keine Einträge oberhalm der Diagonalen und die Matrix ist diagonal.

Die Bedingung an das charakteristische Polynom ist aufgrund des Fundamentalsatzes der Algebra immer erfüllt, wenn wir über $\mathbb{C}$ arbeiten!

Damit ist auch jede Matrix (über $\mathbb{C}$) ähnlich zu einer Matrix in jordanscher Normalform, die Ähnlichkeitsklasse einer Matrix wird von ihren Eigenwerten mit algebraischer Vielfacht und der Größe der Jordan-Matrizen bestimmt.

Dieser Satz ist das aufwändigste Resultat des Kurses und wir werden uns in diesem Kapitel dem Beweis (und ein paar Beispielen) widmen.

10.2 Die Hauptraumzerlegung

Wir fixieren einen Körper K, einen n-dimensionalen Vektorraum V und einen Endomorphismus f.

Definition 10.2.1. Sei λ ein Eigenwert von f, dann heißt $\mathrm{Hau}(f,\lambda) = \ker((f-\lambda \mathrm{id}_V)^n)$ der *verallgemeinerte Eigenraum* oder *Hauptraum* von λ.

Die folgende Definition ist nun naheliegend:

Definition 10.2.2. Ein Endomorphismus $n \in \mathrm{End}_K(V)$ heißt *nilpotent*, wenn es $k \in \mathbb{N}$ gibt mit $n^k = 0$.

Wir wollen nun zeigen, dass sich V als direkte Summe der Haupträume schreiben lässt. Die Idee ist es, $g = (f - \lambda \mathrm{id}_V)$ wiederholt anzuwenden und die Kerne und Bilder von g^i zu betrachten (wie üblich ist $g^0 = \mathrm{id}_V$). Da V endlich-dimensional ist, stabilisieren sich Bild und Kern von g^i, wenn i groß genug ist.

Wir betrachten das folgende Lemma.

Lemma 10.2.3 (Lemma von Fitting)**.** *Sei $g \in \mathrm{End}_K(V)$. Wir definieren $r = \mu(P_g; 0)$ und $d = \min\{k \in \mathbb{N} \mid \ker(g^k) = \ker(g^{k+1})\}$. Es gilt:*

1. $d = \min\{k : \mathrm{Im}\,(g^k) = \mathrm{Im}\,(g^{k+1})\}$.
2. *Der Raum $W = \mathrm{Im}\,(g^d)$ ist g-invariant, $g|_W$ ist ein Isomorphismus und für alle $i \in \mathbb{N}$ gilt $\mathrm{Im}\,(g^{d+i}) = \mathrm{Im}\,(g^d)$.*
3. *Der Raum $U = \ker(g^d)$ ist g-invariant, $(g|_U)^d = 0$ und für alle $i \in \mathbb{N}$ gilt. $\ker(g^{d+i}) = \ker(g^d)$.*
4. $\mu_{g|_U} = t^d$.
5. $V = U \oplus W$.
6. $\dim(U) = r \geq d$ *und* $\dim W = n - r$.

Der Raum U ist nichts anderes, als der Hauptraum $\mathrm{Hau}(g;0) = \ker(g^{\mu_{alg}(g,0)}) = \ker(g^n)$ (beide Darstellungen sind mit Lemma gleich).

Der wichtigste Teil des Lemmas, ist dass wir einen g-invarianten Untervektorraum W finden mit $V = \mathrm{Hau}(g;0) \oplus W$.

Wenn $r = 0$ ist, dann reduziert sich das Lemm auf die triviale Zerlegung $V = 0 \oplus V$ mit $U = 0$ und $W = V$.

Beweis:. 1. Wir haben das Diagramm:

$$\begin{array}{ccccc} \ker g^k & \xrightarrow[\subset]{} & V & \longrightarrow & \mathrm{Im}\,(g^k) \\ \subset\downarrow & & =\downarrow & & \uparrow\subset \\ \ker(g^{k+1}) & \xrightarrow[\subset]{} & V & \longrightarrow & \mathrm{Im}\,(g^{k+1}) \end{array}$$

Es ist $\ker(g^k) = \ker(g^{k+1})$ genau wenn die Räume die gleiche Dimension haben, und das gilt mit der Dimensionsformel genau, wenn g^k und g^{k+1} den gleichen Rang haben, also wenn $\mathrm{Im}\,(g^k) = \mathrm{Im}\,(g^{k+1})$, siehe Satz 4.2.15.

2. $W = \mathrm{Im}\,(g^d)$ ist per Definition g-Invariand und da $\mathrm{Im}\,(g^{d+1}) = \mathrm{Im}\,(g^d)$ ist, sehen wir dass g eingeschränkt auf W surjektiv und damit ein Isomorphismus ist. Aber damit ist auch $\mathrm{Im}\,(g^d) \cong \mathrm{Im}\,(g^{d+i})$ für alle $i \in \mathbb{N}$. Das zeigt 2.

3. U ist auch g-invariant da $g^d(g(u)) = g(g^d(u)) = 0$ ist für $u \in U$.

   Per Definition ist $(g|_U)^d = 0$.

   Aus der Dimensionsformel und $\mathrm{Im}\,(g^d) \cong \mathrm{Im}\,(g^{d+i})$ folgt $\ker(g^k) \cong \ker(g^{k+i})$.

4. Es ist klar, dass $\mu_{g|_U}$ ein Teiler von t^d ist. Wenn aber $(g|_U)^{d-1} = 0$, dann ist $U = \ker g^d \subset \ker(g^{d-1})$ und das widerspricht der Minimalität von d.

5. Wir wählen $v \in U \cap W$. Es gilt $g^d(v) = 0$ und $v = g^d(w)$ für $w \in V$. Als ist $w \in \ker(g^{2d})$. Aber nach 2. ist dann $w \in \ker(g^d)$ und $v = g^d(w) = 0$.

   Aus der Dimensionsformel angewendet auf $g^d : V \to V$ folgt $\dim V = \dim U + \dim W$ und wir erhalten $V = U \oplus W$.

6. Es bleibt, die Dimension von U zu berechnen. Nach Definition gilt $\dim U \geq d$ (denn die Dimension des Kerns von g^i ist mindestens i).

   Wir schreiben $P_g = t^r \cdot Q$ für irgendein Polynom Q mit $Q(0) \neq 0$. Es gilt aber auch $P_g = P_{g|_U} \cdot P_{g|_W}$ wegen Lemma 10.2.4 und $P_{g|U} = t^m$ für $m = \dim(U)$. Da $g|_W$ ein Isomorphismus ist, gilt $P_{g|_W}(0) = \det(g|_W) \neq 0$ und es muss gelten $r = m$. □

Lemma 10.2.4. *Sei $f \in \mathrm{End}_K(V)$ und $V = U \oplus W$ eine Zerlegung in f-invariante Räume. Dann gilt $P_f = P_{f|_U} \cdot P_{f|_W}$.*

Beweis:. Siehe Übung 6.4. □

Mit diesem Lemma sind wir bereit für die Hauptraumzerlegung von V.

Satz 10.2.5. *Sei $f \in \mathrm{End}_K(V)$ und $P_f = \prod_{i=1}^k (t - \lambda_i)^{r_i}$ mit paarweise verschiedenen λ_i. Sei $V_i = \mathrm{Hau}(f, \lambda_i)$. Denn gilt:*

1. *V_i ist f-invariant und hat Dimension r_i für $i = 1, \ldots, k$.*
2. *$V = V_1 \oplus \cdots \oplus V_k$.*
3. *Wir können f als $f = f_d + f_n$ schreiben, so dass f_d diagonalisierbar ist, f_n nilpotent und $f_d \circ f_n = f_n \circ f_d$.*

Beweis:. Wir führen eine Induktion über die Zahl k der Eigenwerte. Für $k = 0$ ist die Aussage trivial.

Wir definieren $g = f - \lambda_1$, per Definition ist dann $P_g(X - \lambda_1) = P_f(Xt)$ und $\mu(P_g; 0) = \mu(P_f; \lambda_1) = r_1$.

Mit Lemma 10.2.5 ist $V = \mathrm{Hau}(g; 0) \oplus W = \mathrm{Hau}(f; \lambda_1) \oplus W$ für $W = \mathrm{Im}\,(g^{r_1})$. Die beiden Summanden sind g-invariant, aber da $f = g + \lambda_1 \mathrm{id}_V$ ist, sind sie auch f-invariant.

Aus Lemma 10.2.4 folgt nun $P_{f|_W} = \prod_{i=2}^{n}(t - \lambda_i)^{r_i}$ und wir können die Induktionsanname auf $f|_W$ anwenden. Damit gelten 1. und 2. per Induktion.

Wir wählen nun eine geordnete Basis $\mathcal{B}$ für V, bestehend aus Basen der Haupträume. Dann wählen wir

$$D = \begin{pmatrix} \lambda_1 E_{r_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_k E_{r_k} \end{pmatrix}$$

Da die Haupräume f-invariant sind, ist $N = M_{\mathcal{B}}(f) - D$ blockdiagonal

$$N = \begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_k \end{pmatrix}$$

und

$$D \cdot N = \begin{pmatrix} \lambda_1 N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k N_k \end{pmatrix} = N \cdot D$$

Wir definieren f_d als den Endomorphismus, der von D dargestellt wird (bezüglich $\mathcal{B}$) und f_n als Endomorphismus, der von N dargestellt wird. Mit dem Satz von Cayley-Hamilton 9.6.1 erfüllt $D + N$ das Polynom P_f. Wir könne den Satz aber auch auf $f|_{V_i}$ anwenden, da V_i f-invariant ist. Als muss jedes $\lambda_i E_{r_i} + N_i$ auch $P_{f|_{V_i}} = (X - \lambda_i)^{r_i}$ erfüllen. Aber dann ist $N_i^{r_i} = 0$ und N_i ist nilpotent. Da wir N^i Block für Block ausrechnen können ist dann auch $N^{\max\{r_i\}} = 0$ und N ist nilpotent. □

Bemerkung 10.2.6. In der Darstellung in Satz 10.2.5.3 ist die Zerlegung $f = f_d + f_n$ eindeutig mit den gegebenen Eigenschaften von f_d und f_n. Wir brauchen diese Eindeutigkeit nicht für unseren späteren Beweis der Eindeutigkeit der jordanschen Normalform.

Wir übersetzen die Hauptraumzerlegung nun in Matrix-Form:

Korollar 10.2.7. *Sei $A \in M(n \times n, K)$ mit $P_A(t) = \prod_{i=1}^{k}(t - \lambda_i)^{r_i}$ wobei die λ_i verschieden sind. Dann ist A ähnlich zu einer Matrix*

$$A' = \begin{pmatrix} \lambda_1 E_{r_1} + N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k E_{r_k} + N_k \end{pmatrix}$$

wobi alle $N_i \in M(r_i \times r_i, K)$ strikte obere Dreiecksmatrizen sind, d.h.

$$\lambda_i E_{r_i} + N_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}$$

Beweis:. Das Korollar folgt sofort, indem wir Satz 10.2.5 auf die lineare Abbildung anwenden, die A darstellt, wir müssen nur prüfen, dass die N_i tatsächlich als obere Dreiecksmatrizen gewählt werden können. Wir wählen hierzu eine geordnete Basis $\mathcal{B}_i$ des Hauptraums $V_i = \text{Hau}(f, \lambda_i)$, so dass $f|_{V_i}$ von einer oberen Dreiecksmatrix T_i dargestellt wird. Das ist nach Satz 9.5.6 möglich. Da λ_i der einzige Eigenwert von $f|_{V_i}$ ist, muss $N_i = T_i - \lambda_i E_{r_i}$ eine strikte obere Dreiecksmatrix sein. Wenn wir diese $\mathcal{B}_i$ nun zu einer Basis $\mathcal{B}$ von V zusammennsetzen hat $T^{\mathcal{E}}_{\mathcal{B}} A T^{\mathcal{B}}_{\mathcal{E}}$ die gewünschte Form. □

10.3 Nilpotente Endomorphismen und Beweis

Die Hauptraumzerlegung war der erste große Schritt zur jordanschen Normalform. Wir müssen als nächstes noch die nilpotenten Matrizen N_i in eine einheitliche Form bringen.

Wir schreiben der Einfachheit halber

$$J_k = J(0,k) = \begin{pmatrix} 0 & 1 & 0 & & 0 \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & & 1 \\ 0 & & & & 0 \end{pmatrix} \in M(k \times k, K)$$

und wollen beweisen:

Satz 10.3.1. *Sei g ein nilpotenter Endomorphismus eines K-Vektorraums V mit $\dim_K(V) = n$ und sei $d = \min\{k \mid g^k = 0\}$. Dann gibt es eindeutig bestimmte Zahlen $k_1, \ldots, k_d \in \mathbb{N}$ mit $\sum_{i=1}^{d} i k_i = n$ und eine Basin $\mathcal{B}$ von V so dass*

$$M_{\mathcal{B}}(g) = \begin{pmatrix} J_d & & & & 0 \\ & \ddots & & & \\ & & J_i & & \\ & & & \ddots & \\ 0 & & & & J_1 \end{pmatrix}$$

wobei der Block J_i k_i-mal vorkommt.

In dieser Darstellung kann k_i auch 0 sein, dann tritt der Block J_i gar nicht auf.

Die Idee des Beweises ist naheliegend: Sei g ein nilpotenter Endomorophismus, den wir in Jordan-Blöcke zerlegen wollen. In jedem Jordan-Block gibt es einen Eigenvektor im Kern von g (der erste Basisvektor). Der nächste Basisvektor ist kein Eigenvektor, wird aber von g auf einen Eigenvektor abgebildt, und liegt damit im Kern von g^2. Wir finden also erst den Eigenraum $\ker(g)$, dann sein Urbild $\ker(g^2)$, dann das Urbild des Urbildes $\ker(g^3)$ und so fort bis $\ker(g^d) = V$.

Beweis:. Mit $U_i = \ker(g^i)$ gilt

$$\{0\} = U_0 \leq U_1 \leq \cdots \leq U_{d-1} \leq U_d = V$$

Außerdem sind die Inklusionen echt, d.h. $U_i \neq U_{i+1}$. Dies folgt aus Lemma 10.2, oder wir zeigen direkt: Da d minimal ist mit $g^d = 0$, gibt es $v \in V$ mit $g^{d-1}(v) \neq 0$ aber $g^d(v) = 0$. Aber dann ist $g^{d-i-1} \in U_{i+1} \setminus U_i$.

Unser Ziel ist es nun, komplementäre Unterräume W_i zu finden, so dass $U_i = U_{i-1} \oplus W_i$ gilt. Die W_i sollen von g injektiv aufeinander abgebildet werden. Wir stellen fest, dass gilt:

1. $g^{-1}(U_{i-1}) = U_i$, denn $g(u) \in U_{i-1}$ genau wenn $g^{i-1}g(u) = 0 \Leftrightarrow u \in U_i$. Daraus folgt $g(U_i) \leq U_{i-1}$ (es muss keine Gleichheit gelten).

2. Für jedes $W \leq V$ mit $W \cap U_i = \{0\}$ für irgendein und $i > 0$ gilt $g|_W$ ist injektiv. Denn $\ker(g) = U_1 \leq U_i$ schneidet W trivial.

Nun wählen wir, z.B. durch Basisergänzung, irgendein W_d mit $V = U_d = U_{d-1} \oplus W_d$. Es gilt $g(W_d) \leq g(U_d) \leq U_{d-1}$ und außerdem $g(W_d) \cap U_{d-2} = 0$, denn für $w \in W_d$ mit $g(w) \in g(W_d) \cap U_{d-2}$ gilt $w \in W_d \cap g^{-1}U_{d-2} = W_d \cap U_{d-1} = \{0\}$.

Also können wir $U_{d-1} = U_{d-2} \oplus g(W_d) \oplus W'_{d-1}$ schreiben. Wir setzen $W_{d-1} = g(W_d) \oplus W'_{d-1}$. Wir erhalten also W_{d-1} erhalten, indem wir $g(W_d)$ zu einem komplementären Unterraum von U_{d-1} ergänzen.

Wir gehen induktiv vor und schreiben

$$U_i = U_{i-1} \oplus W_{i-1}$$

und schließlich

$$V = U_d = W_1 \oplus \cdots \oplus W_d$$

Es gilt hier $g(W_i) \leq W_{i-1}$ und mit 2. ist diese Abbildung injektiv.

Am Ende der Prozedur steht $U_1 = U_0 \oplus W_1$, das heißt $W_1 = U_1$.

Wir erhalten das folgende Diagramm

$$\begin{array}{ccccccccc}
U_d & & & & & & & & \\
\downarrow & & & & & & & & \\
U_{d-1} & \oplus & W_d & & & & & & \\
\downarrow & & \downarrow & & & & & & \\
U_{d-2} & \oplus & W_{d-1} & \oplus & W_d & & & & \\
\downarrow & & \downarrow & & \downarrow & & & & \\
\vdots & & \vdots & & \vdots & & & & \\
\downarrow & & \downarrow & & \downarrow & & & & \\
U_1 & \oplus & W_2 & \oplus & W_3 & \cdots & W_d & & \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
U_0 & \oplus & W_1 & \oplus & W_2 & \cdots & W_{d-1} & \oplus & W_d
\end{array}$$

in dem alle Zeilen isomorph zueinander und zu V sind und jeder Pfeil die Anwendung von g darstellt. Alle vertikalen Pfeile außer in der ersten Zeile sind Injektionen (dafür steht der kleine Haken).

Wir suchen nun eine Basis von V. Hierfür betrachten wir die direkte Summe aller Einträge in der zweiten Spalte unseres Diagramms. aus der letzten Zeile folgt, dass die direkte Summe aller Einträge in der zweiten Spalte gleich V ist.

Wir wählen zuerst eine geordnete Basis von $(w_1^{(d)}, \ldots, w_{k_d}^{(d)})$ von W_d.

Eine Basis von W_{d-1} erhalten wir, indem wir $(g(w_1^{(d)}), \ldots, g(w_{k_d}^{(d)}))$ um $(w_1^{(d-1)}, \ldots, w_{k_{d-1}}^{(d-1)}$ zu einer Basis ergänzen.

So gehen wir induktiv weiter vor, so dass wir eine Basis $\mathcal{B}$ von V erhalten, die aus $g^a(w_j^{(i)})$ besteht, mit $i = 1, \ldots, d$, $j = 1, \ldots, k_r$ und $a = 0, \ldots, i-1$. Für jedes $i = 1, \ldots, d$ und

$j = 1, \ldots, k_r$ bilden dann die Vektoren $(g^{i-1}(w_j^{(i)}), g^{i-2}(w_j^{(i)}, \ldots, g(w_j^{(i)}), w_j^{(i)})$ eine geordnete Basis $\mathcal{B}_j^i$ für einen r-dimensionalen g-invarianten Unterraum auf dem g von der Jordanmatrix J_i dargestellt wird.

Die drei Indizes beim Basisvektor $g^a(w_j^{(i)})$ haben die folgenden Interpretationen:

1. i besagt, wie groß der Jordan-Block ist, zu dem der Vektor gehört
2. j besagt, zu welchem der k_i Jordan-Blocks dieser Größe der Vektor gehört
3. a gibt an, welche Position innerhalb des Jordan-Blocks der Vektor einnimmt.

Damit ist $M_{\mathcal{B}}(g)$ in der gewünschten Form: Für jedes Paar i, j ist $M_{\mathcal{B}_j^i}(g|_{\mathrm{span}_K(\mathcal{B}_j^i)})$ genau ein Jordan-Block, denn g schickt den Eigenvektor nach 0 und verschiebt jeden anderen Basisvektor um eine Position.

Es verbleibt zu zeigen, dass die k_i eindeutig bestimmt sind.

Sicher sind die Unterräume U_i und ihre Dimensionen eindeutig aus g bestimmt. Es gilt aber

$$k_i = \dim(U_i/U_{i-1}) - \dim(U_{i+1}/U_i) \tag{5}$$

Die Anzahl der Jordan-Blocks ist nämlich immer $k_i = \dim(W_i) - \dim(g(W_{i+1}))$, und $\dim W_i = \dim U_i - \dim U_{i-1}$. Da außerdem $g|_{W_i+1}$ injektiv ist, folgt die Formel für k_i. □

Der Beweis von Satz 10.1.2 ist nun eine Formalität:

Satz 10.1.2. Sei $f \in \mathrm{End}_K(V)$ und zerfalle P_f in Linearfaktoren. Dann gibt es eine Basis für V, so dass f von einer blockdiagonalen Matrix aus Jordan-Matrizen dargestellt wird, d.h.

$$M(f)_{\mathcal{B}} = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix}$$

wobei $J_1, \ldots, J_m$ Jordan-Matrizen sind und die diagonalen Einträge von $M(f)_{\mathcal{B}}$ die Eigenwerte von f mit ihrer algebraischen Vielfachheit sind.

Die Darstellung ist eindeutig bis auf Permutation der Jordan-Matrizen.

Beweis:. Seien λ_i die verschiedenen Eigenwerte von f, r_i ihre algebraischien Vielfachheiten und $V_i = \mathrm{Hau}(f, \lambda_i)$ ihre Haupträume. Nach Korollar 10.2.7 existiert eine Basis $\mathcal{B}'$, so dass $M(f)$ block-diagonal ist und jeder Block die Form $\lambda_i E_{r_i} + N_i$ hat mit N_i eine strikte obere Dreiecksmatrix, die einen nilpotenten Endomorphismus g_i von V_i darstellt. Wir wenden nun Satz 10.3.1 auf jedes g_i an. So finden wir eine neue geordnete Basis $\mathcal{B}_i$ von V_i so dass $M(g_i)_{\mathcal{B}_i}$ eine block-diagonale Matrix aus Jordan-Matrizen ist. Sei $\mathcal{B}$ die Vereinigung dieser Basen $\mathcal{B}_i$. Da $\lambda_i E_{r_i}$ vom Basiswechsel unberührt bleibt, ist $M(f)_{\mathcal{B}}$ in jordanscher Normalform. (Sei T die Basiswechselmatrix von $\mathcal{B}' \cap V_i$ nach $\mathcal{B}_i$. Dann ist TN_iT^{-1} ein Jordan-Block J_i mit Eigenwert 0 und $T(\lambda_i E_{r_i} + N_i)T^{-1} = T\lambda_i E_{r_i} T^{-1} + TN_iT^{-1} = \lambda_i E_{r_i} + J_i$ ist ein Jordan-Block mit Eigenwert λ_i.)

Bis auf Umordnung der Jordan-Blöcke ist dies eindeutig, denn die Eigenwerte, die Dimensionen der Haupträume und die Vielfachheiten k_i^a des i-ten Jordanblocks für einen gegebenen Eigenwert λ_a sind alle eindeutig bestimmt.

Umgekehrt kann jede Umordnung der Jordan-Blöcke durch eine Umordnung der Elemente der geordneten Basis $\mathcal{B}$ erreicht werden. □

Korollar 10.3.2. *Sei $A \in M(n \times n, K)$ und zerfalle p_A in Linearfaktoren. Dann ist A ähnlich zu einer Matrix in jordanscher Normalform, die bis auf Permutation der Jordan-Matrizen eindeutig ist.*

Beweis:. Wir wenden Satz 10.1.2 auf den Endomorphismus an, den A bezüglich der Standardbases darstellt. □

10.4 Beispiele und Anwendungen

Der Beweis von Satz 10.3.1 gibt uns auch einen Algorithmus zur Bestimmung der jordanschen Normalform.

Betrachtung 10.4.1. Wir bestimmen die jordansche Normalform von $A \in M(n \times n, K)$.

1. Zuerst finden wir die Eigenwerte λ_a und ihr algebraischen Vielfachheiten r_a aus dem charakteristischen Polynom.

2. Dann berechnen wir für jedes $i = 1, \ldots, r_a$ die Dimension des verallgemeinerten Eigenraums $u_i^a = \dim_K(\ker((A - \lambda_a E_n)^i)) = n - \operatorname{rg}((A - \lambda_a E_n)^i)$ und setzen $u_i^a = \dim_K(U_i^a)$.

   Nicht jede Rechnung muss ausgeführt werden, wenn zum Beispiel $u_1^a = r_a - 1$ ist, dann muss $u_2^a = r_a$ sein und wir brauchen keine weiteren Rechnungen! (Überlegen Sie sich, warum das der Fall ist!)

3. Wir setzen

$$k_i^a = \dim(U_i^a/U_{i-1}^a) - \dim(U_{i+1}^a/U_i^a) \qquad = 2u_i^a - u_{i+1}^a - u_{i-1}^a$$

   nach Formel 5. Die jordansche Normalform von A hat nun für jedes a und i genau k_i^a Jordan-Blöcke der Größe i. Die Matrix A ist also ähnlich zu einer blockdiagonalen Matrix mit k_i Matrizen $J(\lambda_a, i)$ entlang der Diagonale für jedes a und i.

Betrachtung 10.4.2. Nun wollen wir nicht nur die Normalform finden, sondern auch die Basiswechselmatrix bestimmen. Wir fixieren $f \in \operatorname{End}_K(V)$ und gehen wie folgt vor (die Prozedur für $A \in M(n \times n, K$ geht genauso):

1. Wir finden wieder die Eigenwerte λ_a und ihr algebraischen Vielfachheiten r_a aus dem charakteristischen Polynom.

2. Dann berechnen wir für jedes $i = 1, \ldots, r_a$ den verallgemeinerten Eigenraum $U_i^a = \ker((A - \lambda_a E_n)^i)$ und setzen $u_i^a = \dim_K(U_i^a)$. Wir finden eine Basis $\mathcal{B}_a = (g^b(w_j^{(i)})_{b,i,j}$ von $\operatorname{Hau}(f, \lambda_a)$, indem wir Schritt für Schritt die Basis von U_i^a zu einer Basis von U_{i+1}^a ergänzen wie im Beweis von Satz 10.3.1.

3. Wir ordnen diese Basis indem wir zuerst einen Eigenvektor $g^{i-1}(w_j(i))$ und dann seine Urbilder $g^{i-2}(w_j(i)), g^{i-3}(w_j(i)), \ldots$ wählen, und dann denn nächsten Eigenvektor und all seine Urbilder.

4. Wie oben können wir die Größe der Jordan-Blocks aus der Formel 5 bestimmen: $k_i^a = 2u_i^a - u_{i+1}^a - u_{i-1}^a$.

5. Wir setzen die Basen $\mathcal{B}_a$ der Haupträume zu einer Basis $\mathcal{B}$ zusammen und $M_{\mathcal{B}}(f)$ ist in jordanscher Normalform.

Beispiel 10.4.3. Wir berechnen die jordansche Normalforem der reellen Matrix

$$A = \begin{pmatrix} 3 & -2 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 & 2 \\ 2 & -4 & 2 & 2 & 3 \\ -2 & 3 & 0 & -1 & -4 \\ 1 & -2 & 0 & 2 & 5 \end{pmatrix}$$

$\langle\langle$Beachten Sie dass diese Matrix im Skript und in der Vorlesung ursprünglich einen Tippfehler hatte!$\rangle\rangle$ Wir berechnun zuerst $P_A = (X-2)^5$. (Zum Beispiel geschicktesten durch Spaltenentwicklung nach der 3. Spalte und dann Zeilenentwicklung nach der 2. Zeile für die entstehende 4×4-Untermatrix.)

Damit ist $A = 2E_5 + N$ und $N = A - 2E_5$ ist nilpotent und wir berechnen

$$N = \begin{pmatrix} 1 & -2 & 0 & 2 & 3 \\ 0 & -1 & 0 & 1 & 2 \\ 2 & -4 & 0 & 2 & 3 \\ -2 & 3 & 0 & -3 & -4 \\ 1 & -2 & 0 & 2 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 0 & 2 & 3 \\ 0 & -1 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Damit ist $U_1 = \ker(N) = \operatorname{span}_{\mathbb{R}}(e_3, (2,1,0,-3,2)^T)$.

Als nächstes betrachten wir N^2:

$$N^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Und erhalten $U_2 = \ker(N^2) = \operatorname{span}_{\mathbb{R}}(e_2, e_3, e_4, e_1 + e_5)$.

Schließlich ist $N^3 = 0$ und $U_3 = \mathbb{R}^5$.

Damit gilt $\dim(U_3/U_2) = 1$, $\dim(U_2/U_1) = 2$ und $\dim(U_1) = 2$.

Die Formel 5 aus dem Beweis von Satz 10.3.1 gibt

$$\begin{aligned} k_3 &= 1 - 0 = 1 \\ k_2 &= 2 - 1 = 1 \\ k_1 &= 2 - 2 = 0 \end{aligned}$$

und damit hat die jordansche Normalform von A einen Jordan-Block der Größe 3 und einen der Größe 2. Aber wir wollen noch die passende Basis bestimmen. Sei g die lineare Abbildung mit $A = M_{\mathcal{E}}(g)$.

Wir schreiben also $\mathbb{R}^5 = U_3 = U_2 \oplus W_3$ mit $W_3 = \operatorname{span}_{\mathbb{R}}(e_1)$. Es gilt $g(W_3) = \operatorname{span}_{\mathbb{R}}((1,0,2,-2,1)^T)$.

Wir schreiben $U_2 = U_1 \oplus W_2$ wobei wir W_2 erhalten, indem wir $g(W_3)$ zu einem Komplement von U_1 ergänzen. Wir wählen $W_2 = g(W_3) \oplus \operatorname{span}_{\mathbb{R}}(e_2)$.

Als nächstes berechnen wir $g(W_2) = \operatorname{span}_{\mathbb{R}}(e_3, (-2,1,-4,3,-2)^T)$. Dies ist gleich U_1 und wir müssen nicht weiter ergänzen. Damit ist $g(W_2)$ auch gleich W_1.

Zusammenfassend erhalten wir

$$\mathbb{R}^5 = W_1 \oplus W_2 \oplus W_3$$

mit

$$\begin{array}{lll} & J_3 & J_2 \\ & & \\ w_1^{(3)} = & \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix} & & W_3 \\ & \downarrow & \\ g(w_1^{(3)}) = & \begin{pmatrix}1\\0\\2\\-2\\1\end{pmatrix} & w_1^{(2)} = \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix} & W_2 \\ & \downarrow & \downarrow & \\ g^2(w_1^{(3)}) = & \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} & g(w_1^{(2)}) = \begin{pmatrix}-2\\-1\\-4\\3\\-2\end{pmatrix} & W_1 \\ & \downarrow & \downarrow & \\ & 0 & 0 & \end{array}$$

Wobei die drei Zeilen für W_3, W_2 und W_1 stehen und die beiden Spalten für Jordan-Matrizen J_3 und J_2.

Für die Basis $\mathcal{B} = \text{span}_{\mathbb{R}}(\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\-2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}-2\\-1\\-4\\3\\-2\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix})$ gilt dann

$$M_{\mathcal{B}}(g) = \begin{pmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&1\\0&0&0&0&0\end{pmatrix}$$

und damit

$$M_{\mathcal{B}}(f) = \begin{pmatrix}2&1&0&0&0\\0&2&1&0&0\\0&0&2&0&0\\0&0&0&2&1\\0&0&0&0&2\end{pmatrix}.$$

Bemerkung 10.4.4. Sei M eine Matrix in $M(n \times n, K)$ in jordanscher Normalform, sei λ ein Eigenwert von λ und sei k_i die Anzahl der Jordan-Blöcke $J(\lambda, i)$. Dann können wir leicht die folgenden Informationen ablesen:

1. Die algebraische Vielfachheit von λ ist die Summe der Größen aller Jordan-Blocks: $\mu_{alg}(M,\lambda) = \sum_{i=1}^{n} ik_i$
2. Die geometrische Vielfachheit von λ ist die Anzahl der Jordan-Blocks: $\mu_{geo}(M,\lambda) = \sum_{i=1}^{n} k_i$.
3. Die Vielfachheit von λ im Minimalpolynom von A ist die Größe des größten Jordan-Block: $\mu(\lambda, \mu_M) = \max\{i \mid k_i \neq 0\}$.

All diese Aussagen folgen direkt aus den Definitionen.

Wir stellen noch das folgende Korollar der jordanschen Normalform fest:

Satz 10.4.5. *Sei $A \in M(n \times n, K)$ mit $K \subset \mathbb{C}$ Dann ist äquivalent:*

1. *Die Matrix A ist diagonalisierbar.*
2. *Das Minimalpolynom μ_A zerfällt in paarweise verschiedene Linearfaktoren.*

Beweis:. Wenn A diagonalisierbar ist, dann ist μ_A gleich dem Minimalpolynom einer Diagonalmatrix D. Seien $\lambda_1, \ldots, \lambda_k$ die unterschiedlichen Eigenwerte von A. Nach Satz 9.6.8 gilt $\prod_{i=1}^{k}(X - \lambda_i) \mid \mu_A$. Aber es ist leicht zu sehen, dass $\prod(D - \lambda_i) = 0$ da D dioganal ist. und damit ist $\prod_{i=1}^{k}(X - \lambda_i) = \mu_D$ nach Satz 9.6.4.

Umgekehrt zerfalle μ_A in paarweise verschiedene Linearfaktoren. Es zerfällt P_A in Linearfaktoren über $\mathbb{C}$ und nach Satz 9.6.8 sind alle Nullstellen auch Nullstellen von μ_A und daher Elemente von K. Also ist A nach Korollar 10.3.2 ähnlich zu einer Matrix A' in jordanscher Normalform. Wenn A nicht diagonalisierbar ist, dann gibt es einen Jordan-Block J_i für $i > 1$. Aber dann erfüllt A' nicht $\mu_A = \prod(X - \lambda_i)$. □

Der Satz gilt mit dem gleichen Beweis für jeden Körper, da jeder Körper eine Erweiterung hat, in der P_A in Linearfaktoren zerfällt, vgl. Bemerkung 9.6.3.

Betrachtung 10.4.6. Wir haben nun das folgende Schema zur Untersuchung der Diagonalisierbarkeit einer quadratischen Matrix $A \in M(n \times n, K)$, bei dem wir nicht explizit die Eigenräume berechnen müssen.

1. Berechne das charakteristische Polynom P_A und seine Zerlegung in Linearfaktoren.

   Zerfällt P_A nicht in Linearfaktoren, so ist A nach Satz 9.4.8 nicht diagonalisierbar. Zerfällt P_A in Linearfaktoren, gehe zu

2. Setze $f(X) := \prod_{i=1}^{m}(X - \lambda_i) \in K[X]$, wobei $\lambda_1, \ldots, \lambda_m$ die verschiedenen Nullstellen des charakteristischen Polynoms P_A sind.

   Gilt $\widetilde{f}(A) = 0_n$; so ist A nach Satz 10.4.5 diagonalisierbar.

   Gilt $\widetilde{f}(A) \neq 0_n$; so ist A nach Satz 10.4.5 nicht diagonalisierbar.

Wir betrachten noch eine Anwendung:

Beispiel 10.4.7. Eine *Differenzialgleichung* ist eine Gleichung für eine Funktion und ihre Ableitungen, z.B.

$$y'(t) = \lambda y(t).$$

Große Teilgebiete der angewandten Mathematik beschäftigen sich mit der Lösung bestimmter Differenzialgleichung.

Eine Lösung einer Differenzialgleichung ist eine differenzierbare Funktion $y(t)$, die (zusammen mit ihren Ableitungen) die Gleichung erfüllt. Im Beispiel also eine Funktion, deren Ableitung stets proportional zu ihrer eigenen Ableitung ist. Das ist in zahleichen Anwendungen interessant, zum Beispiel für Zerfalls- und Wachstumsprozesse. Aus der Analysis wissen Sie, dass $y(t) = ce^{\lambda t}$ für irgendeine Konstante c die einzigen Funktionen sind, die diese Gleichung erfüllen.

Wir betrachten nun ein lieares System von Differenzialgleichungen:

$$\begin{aligned} y_1'(t) &= a_{11}y_1(t) + a_{12}y_2(t) \\ y_2'(t) &= a_{21}y_1(t) + a_{22}y_2(t) \end{aligned}$$

Wenn wir (y_1, y_2) als vektorwertige Funktion auffassen, dann können wir dies als Matrixdifferenzialgleichung schreiben:

$$y' = Ay$$

Da Differenzierung ein linearer Endomorphismus der glatten Funktionen ist, bilden insbesondere die Lösungen einen Vektorraum!

Wenn wir annehmen, dass y Werte in $\mathbb{C}^2$ annimmt, dann können wir weiterhin A in jordansche Normalform bringen: Sei $J = SAS^{-1}$ und $z = Sy$. Dann können wir $z' = Jz$ lösen und anschließend $y = S^{-1}z$ substituieren.

Wir betrachten also entweder

$$\begin{aligned} z_1'(t) &= \lambda_1 z_1(t) \\ z_2'(t) &= \lambda_2 z_2(t) \end{aligned}$$

und finden leicht zwei linear unabhängige Lösungen:

$$\begin{pmatrix} e^{\lambda_1 t} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e^{\lambda_2 t} \end{pmatrix}$$

oder aber A ist nicht diagonalisierbar und wir erhalten

$$\begin{aligned} z_1'(t) &= \lambda z_1(t) + 1 z_2(t) \\ z_2'(t) &= \lambda z_2(t) \end{aligned}$$

und finden die folgenden Vektoren in der Lösungsmenge für z:

$$\begin{pmatrix} e^{\lambda t} \\ 0 \end{pmatrix}, \begin{pmatrix} te^{\lambda t} \\ e^{\lambda t} \end{pmatrix}$$

(Rechnen Sie nach, dass dies tatsächlich zwei Lösungen sind!)

In der Analysis lernen Sie, dass der Lösungsraum tatsächlich zweidimensional sind und wir somit alle Lösungen gefunden haben.

Diese Betrachtungen verallgemeinern selbstverständlich auf n Dimensionen!

11 Intermezzo: Universelle Eigenschaften

Wir haben schon an mehreren Stelle, oft ohne sie zu bennenen, *universelle Eigenschaften* getroffen.

Beispiel 11.0.1. Sei A eine Menge von reellen Zahlen. Dann ist das *Infimum* von A eine relle Zahl x mit der Eigenschaft

$$\forall a \in A \ : \ x \leq a \qquad (*)$$

und immer wenn für $y \in \mathbb{R}$ gilt $y \leq a$ für alle $a \in A$, dann gilt auch $y \leq x$.

Also hat x eine gewisse Eigenschaft $(*)$ und jedes andere Objekt mit der gleichen Eigenschaft ist kleiner als x.

Beachten Sie, dass das Infimum nicht für jede Menge existiert.

Beispiel 11.0.2. Sei A eine Menge von natürlichen Zahlen. Dann ist der *größte gemeinsame Teiler* der Mange A eine natürliche Zahl n mit der Eigenschaft

$$\forall a \in A \ : \ n \mid a \qquad (*)$$

und immer wenn für $m \in \mathbb{N}$ gilt $m \mid a$ für alle $a \in A$, dann gilt auch $m \mid n$.

Also hat x eine gewisse Eigenschaft $(*)$ und jedes andere Objekt mit der gleichen Eigenschaft teilt x.

Statt einer Menge von natürlichen Zahlen können wir auch eine Menge von Polynomen verwenden! Das Minmialpolynom von f ist der größte gemeinsame Teiler aller Polynome, die von f erfüllt werden.

Dies sind beides universelle Eigenschaften: Ein Objekt x hat eine bestimmte Eigenschaft und jedes andere Objekt mit der gleichen Eigenschaft steht in einer bestimmten Relation zu x. Einmal dreht es sich dabei um die Relation "$\leq$" und einmal um die Relation "teilt".

Wir betrachten ein komplizierteres Beispiel, wo unsere Relation "es gibt eine eindeutige lineare Abbildung" ist.

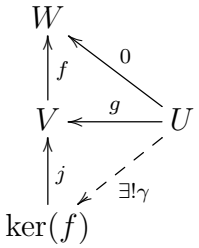
Dazu wenden wir uns wieder den Vektorräumen zu:

Beispiel 11.0.3. Sei $f : V \to W$ eine lineare Abbildung. Die Inklusion $j : \ker(f) \to V$ hat die folgende Eigenschaft: "Es gilt $f \circ j = 0$. "

Für jede andere lineare Abbildung $g : U \to V$ mit $f \circ g = 0 : U \to W$ gibt es eine eindeutige Abbildung $\gamma : U \to \ker(f)$ mit $g = j \circ \gamma$.

Zuerst prüfen wir, dass γ existiert: Per Definition ist $g(u)$ in $\ker(f) \subset V$, entsprechend können wir $\gamma : U \to \ker(f)$ durch $u \mapsto g(u)$ definieren. Der einzige Unterschied zwischen g und γ ist der Bildbereich der Abbildung!)

Sei nun $\gamma' : U \to \ker(f)$ eine andere Abbildung mit $j\gamma' = g$. Dann gilt $j \circ \gamma' = j \circ \gamma$. Aber j ist injektiv, das heißt für jedes u gilt $j \circ \gamma'(u) = j \circ \gamma(u) \Rightarrow \gamma'(u) = \gamma(u)$, also ist γ in der Tat eindeutig. Wir fassen die Situation im folgenden Diagramm zusammen:



Im Diagramm kommutieren beide Dreiecke, also $f \circ g = 0$ und $i \circ \gamma = g$. Dasy Symbol $\exists!$ bedeutet, dass es eine eindeutige gestrichtelte Abbildung gibt, für die beide Dreiecke kommutieren.

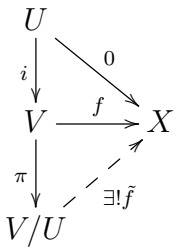
Eine analoge Eigenschaft gilt auch für den Quotientenvektorraum, wir müssen nur vorsichtig sein, da Quotientenvektorräume intuitiv schwerer zu fassen sind als Untervektorräume.

Der folgende Satz ist nichts als eine sorgfältigere Formulierung von Satz 3.4.3. Damals haben wir gesagt, es ist äquivalent eine Abbildung aus V/U anzugeben oder eine Abbildung aus V, die U nach 0 schickt.

Satz 11.0.4 (Universelle Eigenschaft des Quotientenvektorraums)**.** *Sei K ein Körper, V ein K-Vektorraum und $U \subset V$ ein Untervektorraum und $\pi : V \to V/U$ die kanonische Surjektion.*

Dann existiert für jeden K-Vektorraum X und jede lineare Abbildung $f : V \to X$ mit $f_{|U} = 0$ eine eindeutig bestimmte lineare Abbildung $\tilde{f} : V/U \to X$ mit $f = \tilde{f} \circ \pi$.

Man sagt, dass f über V/U eindeutig faktorisiert. Als Diagram können wir schreiben:



Beweis:. Sei $v \in V$. Für jede solche Abbildung $\tilde{f}$ muss gelten:

$$\tilde{f}([v]) = \tilde{f} \circ \pi(v) = f(v) \ ,$$

so dass $\tilde{f}$ eindeutig festgelegt ist. Dies ist wegen $f(v+u) = f(v) + f(u) = f(v)$ für alle $u \in U$ wohldefiniert, d.h. unabhängig vom Repräsentanten v von $[v]$. Aus der Linearität von f folgt wieder leicht die Linearität von $\tilde{f}$. □

⟨⟨Es war recht kompliziert, diesen Satz zu formulieren, aber der Beweis ist einfach!⟩⟩
Wir können die universelle Eigenschaft des Quotientenvektorraums auch so formulieren: sei $\mathrm{Hom}_U(V,X) := \{f : V \to X \,|\, \mathrm{linear} f_{|U} = 0\}$. Dann ist

$$\begin{array}{rcl} \mathrm{Hom}_K(V/U, X) & \to & \mathrm{Hom}_U(V,X) \\ \tilde{f} & \mapsto & \tilde{f} \circ \pi \end{array}$$

ein Isomorphismus von Vektorräumen. Die Surjektivität ist die Aussage, dass man zu jedem $f \in \mathrm{Hom}_U(V,X)$ ein $\tilde{f} \in \mathrm{Hom}_K(V/U,X)$ finden kann, also die Existenzaussage für $\tilde{f}$. Die Injektivität ist die Eindeutigkeitsaussage für $\tilde{f}$.

Wir nennen all diese Eigenschaften des Infimumus, des ggT, des Kerns und des Quotientenvektorrraums *universelle Eigenschaften*.

Ein mathematisches Objekt U eine *universelle Eigenschaft*, wenn es eine bestimmte Eigenschaft $(*)$ hat und jedes andere mathematische Objekt X, dass die gleiche Eigenschaft $(*)$ hat, in einer bestimmten eindeutigen Relation zu U steht (die sich aus $(*)$ ergibt). Diese Relation ist typischerweise eine eindeutige Abbildung von X nach U oder von U nach X.

⟨⟨Auch "$\leq$" und "kleiner gleich" lassen sich als Abbildungen auffassen! Aber dies jetzt zu formalisieren und eine formal korrekete Definiton der unviersellen Eigenschaften zu geben würde uns jetzt zu weit abseits des Weges führen.⟩⟩

Es gilt der folgende Meta-Satz: Ein Objekt mit einer universelle Eigenschaft ist immer eindeutig bestimmt (wenn es existiert)! Wir sehen das im folgenden Beispiel:

Beispiel 11.0.5. Die Polynomalgebra $K[X]$ mit dem Element X hat die triviale Eigenschaft eine Algebra mit einem ausgezeichneten Element zu sein.

Wir haben aber in Satz 9.3.4 gesehen: Für jede K-Algebra A mit einem markierten Objekt $a \in A$ gibt es eine eindeutige Abbildung $K[X] \to A$ mit $X \mapsto a$. Das ist die universelle Eigenschaft von $(K[X], X)$.

Aus der universellen Eigenschaft flogt auch, dass $K[X]$ die einzige K-Algebra mit dieser Eigenschaft ist. Denn habe B mit $b \in B$ die gleiche Eigenschaft, dann gibt es eindeutige Abbildungen $g : K[X] \to B$ mit $X \mapsto b$ und $h : B \to K[X]$ mit $b \mapsto X$. Diese Eigenschaften sind invers zueinander, denn hg ist die eindeutige Abbildung von $K[X]$ nach $K[X]$, die X nach X schickt. Das ist die Identität! Genauso ist $gh = \mathrm{id}_B$.

Sie kennen die Eindeutigkeit auch vom Infimum und vom ggT. Es ist nicht besonders schwierig, allgemein zu zeigen, dass Objekte mit universeller Eigenschaft eindeutig bestimmt sind – wenn man einmal die universelle Eigenschaft formal definiert hat.

Wir zeigen das Resultat nun noch am Beispiel des Quotientenvektorraums:

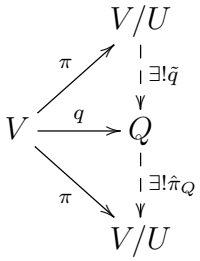
Satz 11.0.6. *Sei wieder $\pi : V \to V/U$ die kanonische Surjektion auf einen Quotientvektorraum. Sei Q ein weiterer K-Vektorraum und $q : V \to Q$ eine lineare Abbildung mit $q_{|U} = 0$, so dass für jede lineare Abbildung $f : V \to X$ mit $f_{|U} = 0$ eine eindeutig bestimmte lineare Abbildung $\tilde{f}_Q : Q \to X$ mit $f = \tilde{f}_Q \circ q$ existiert.*

Dann gibt es einen eindeutig bestimmten Isomorphismus $\tilde{\pi}_Q : Q \to V/U$, so dass $\pi = \tilde{\pi}_Q \circ q$ gilt.

Beweis:. Wenden wir die universelle Eigenschaft von V/U auf die lineare Abbildung $q : V \to Q$ an, so finden wir eine eindeutige lineare Abbildung $\tilde{q} : V/U \to Q$ mit $q = \tilde{q} \circ \pi$. Zum zweiten wenden wir die für Q geforderte Eigenschaft auf die lineare Abbildung $\pi : V \to V/U$ an und finden $\tilde{\pi}_Q : Q \to V/U$ mit $\pi = \tilde{\pi}_Q \circ q$. Es folgt

$$\tilde{\pi}_Q \circ \tilde{q} \circ \pi = \tilde{\pi}_Q \circ q = \pi \,.$$

Als Diagramm



Natürlich gilt auch $\mathrm{id}_{V/U} \circ \pi = \pi$. Aber die universelle Eigenschaft von V/U, angewandt auf $\pi : V \to V/U$ selbst, sagt, dass eine solche Abbildung eindeutig ist, also gilt $\tilde{\pi}_Q \circ \tilde{q} = \mathrm{id}_{V/U}$. Analog zeigt man auch $\tilde{q} \circ \tilde{\pi}_Q = \mathrm{id}_Q$. □

Auch die direkte Summe hat eine universelle Eigenschaft. Sei $(V_\lambda)_{\lambda \in \Lambda}$ eine Familie von Vektorräumen.

Lemma 11.0.7. *Definiere für jedes $\mu \in \Lambda$ die kanonische Injektion mit*

$$\begin{array}{rccl} \iota_\mu : & V_\mu & \hookrightarrow & \bigoplus_{\lambda \in \Lambda} V_\lambda \\ & v_\mu & \mapsto & (0, 0, \ldots, v_\mu, 0, \ldots) \end{array}$$

Ist W nun ein beliebiger K–Vektorraum, so gibt es für jede Familie von Abbildungen $g_\mu : V_\mu \to W$ $(\mu \in \Lambda)$ eine eindeutige lineare Abbildung $g : \oplus_{\lambda \in \Lambda} V_\lambda \to W$ mit $g_\mu = g \circ \iota_\mu$.

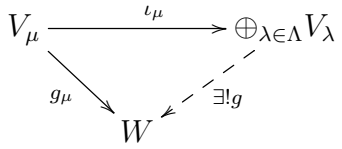
Äquivalent dazu gilt: Es gibt eine natürliche Bijektion von Mengen

$$\begin{array}{rl} \mathrm{Hom}_K\Big(\bigoplus_{\lambda\in\Lambda} V_\lambda, W\Big) & \tilde{\to} \prod_{\lambda\in\Lambda} \mathrm{Hom}_K(V_\lambda, W) \qquad (*) \\ f & \mapsto (f\circ \imath_\lambda)_{\lambda\in\Lambda} \end{array}$$

Wobei die rechte Seite das kartesische Produkt aller $\mathrm{Hom}_K(V_\lambda, W)$ *für* $\lambda \in \Lambda$ *ist, d.h. ein Element der rechten Seite ist gegeben durch je ein Element aus jeder Menge* $\mathrm{Hom}_K(V_\lambda, W)$*.*

Man kann also eine ganze Familie $(g_\mu)_{\mu\in\Lambda}$ von Abbildungen in ein und denselben K-Vektorraum W eindeutig durch eine einzige lineare Abbildung g aus der direkten Summe heraus beschreiben.

Dies sei noch einmal in dem folgenden kommutierenden Diagramm dargestellt:



für alle $\mu \in \Lambda$

Beweis:. Wir zeigen zuerst, dass die beiden Beschreibungen äquivalent sind: Ist für jedes $\mu \in \Lambda$ eine lineare Abbildung

$$g_\mu : V_\mu \to W$$

in einen beliebigen, aber festen Vektorraum W gegeben, dann ist die Surjektivität in $(*)$ genau die Existenz einer linearen Abbildung

$$g : \oplus V_\lambda \to W \ ,$$

so dass $g_\mu = g \circ \imath_\mu$ für alle $\mu \in \Lambda$ gilt.

Injektivität in $(*)$ bedeutet genau, dass diese Abbildung eindeutig ist.

Nun zeigen wir, dass $(*)$ gilt.

Jedes Element der direkten Summe lässt sich eindeutig schreiben in der Form

$$v = \sum_{\lambda\in\Lambda} \imath_\lambda(v_\lambda)$$

mit $v_\lambda \in V_\lambda$, nur endlich viele $v_\lambda \neq 0$. Deswegen ist insbesondere die Summe endlich und definiert.

Wir zeigen Injektivität von $(*)$: gelte $f \circ \imath_\lambda = 0$ für alle $\lambda \in \Lambda$. Dann gilt für ein beliebiges $v \in \oplus_{\lambda\in\Lambda} V_\lambda$

$$f(v) = f\left(\sum_{\lambda\in\Lambda} \imath_\lambda(v_\lambda)\right) = \sum_{\lambda\in\Lambda} f\circ\imath_\lambda(v_\lambda) = 0 \ ,$$

also ist f die Nullabbildung, der Nullvektor in $\mathrm{Hom}(\oplus_\Lambda V_\lambda, W)$.

Wir zeigen Surjektivität von $(*)$: für eine gegebene Familie von Abbildungen

$$g_\lambda : V_\lambda \to W \qquad \text{für alle } \lambda \in \Lambda$$

zu zeigen, setzen wir

$$g\left(\sum_{\lambda\in\Lambda} \imath_\lambda(v_\lambda)\right) := \sum_{\lambda\in\Lambda} g_\lambda(v_\lambda) \ .$$

Dies definiert eine K–lineare Abbildung g mit den gewünschten Eigenschaften. □

Beispiel 11.0.8. Eine analoge Aussage gilt auch für das direkte Produkt aus 3.3.14. Wir hatten für eine belibige Indexmenge Λ und Vektorräume $(V_\lambda)_{\lambda\in\Lambda}$ definiert:

$$\prod_{\lambda\in\Lambda} V_\lambda = \Big\{(v_\lambda)_{\lambda\in\Lambda} |\ v_\lambda \in V_\lambda\Big\}$$

Wir betrachten die kanonischen Surjektionen

$$pr_\mu :\ \prod_{\lambda\in\Lambda} V_\lambda \twoheadrightarrow V_\mu$$

auf die μ–te Komponente der Familie und erhalten einen Isomorphismus von K-Vektorräumen:

$$\begin{array}{rl} \mathrm{Hom}_K\Big(W, \prod_{\lambda\in\Lambda} V_\lambda\Big) & \xrightarrow{\sim} \prod_{\lambda\in\Lambda} \mathrm{Hom}_K(W, V_\lambda) \qquad (**) \\ f & \mapsto \Big(pr_\lambda \circ f\Big)_{\lambda\in\Lambda}\ . \end{array}$$

Wir reformulieren diesen Sachverhalt, den man die *universelle Eigenschaft* des Produkts nennt:

Ist für jedes $\mu \in \Lambda$ eine lineare Abbildung

$$g_\mu :\ W \to V_\mu$$

gegeben, folgt aus der der Surjektivität in $(**)$ die Existenz einer linearen Abbildung

$$g :\ W \to \prod_{\lambda\in\Lambda} V_\lambda$$

so dass $g_\mu = pr_\mu \circ g$ für alle $\mu \in \Lambda$ gilt. Diese ist wegen der Injektivität in $(**)$ eindeutig. Man kann also eine ganze Familie von Abbildungen aus einen K-Vektorraum W eindeutig durch eine einzige eindeutig bestimmte lineare Abbildung beschreiben.

Da direkte Summe und Produkt durch universelle Eigenschaften definiert sind, sind sie auch eindeutig bestimmt!

12 Bilineare Algebra

12.1 Der Dualraum

Definition 12.1.1. Sei K ein beliebiger Körper und V ein K–Vektorraum. Dann heißt der K-Vektorraum

$$V^* := \mathrm{Hom}_K(V, K) = \{\varphi : V \to K |\ \varphi \text{ linear}\}$$

der *Dualraum* von V. Die Elemente von V^* heißen *Linearformen* auf V.

Beispiele 12.1.2.

1. Sei $v \in K^n$ ein fester Vektor. Dann ist $w \mapsto v^T w$ eine lineare Abbildung $K^n \to K$ und damit ein Element von $(K^n)^*$.

   Allgemein können wir Elemente im Dualraum von K^n als Zeilenvektoren auffassen, denn lineare Abbildungen von K^n nach K werden, bezüglich der Standardbasis, genau von Zeilenvektoren in $M(1 \times n, K)$ dargestellt.

2. Sei $K = \mathbb{R}$ und $V = C^0([a, b], \mathbb{R})$ der Vektorraum der stetigen reellwertigen Funktionen auf dem abgeschlossenen Intervall $[a, b]$. Dann ist für jedes $x \in [a, b]$ die Abbildung

   $$\psi_a(f) = f(a)$$

   . eine Linearform.

3. Mit V wie in 2. ist

   $$\varphi(f) = \int_a^b f(x)\mathrm{d}x$$

   eine Linearform auf V.

   Wir können dies verallgemeinern: Für jedes $g \in V$ ist $\varphi_g : f \mapsto \int_a^b f(x)g(x)\mathrm{d}x$ eine Linearform!

Betrachtung 12.1.3. Sei $\dim_K V < \infty$. Aus Korollar 5.2.8 folgt

$$\dim_K V^* = \dim_K \mathrm{Hom}_K(V, K) = \dim_K V \cdot \dim_K K = \dim_K V \ .$$

Sei nun $\mathcal{B}$ eine Basis von V, $\mathcal{B} = \{b_1, \ldots, b_n\}$. Definiere Linearformen $b_j^* \in V^*$ durch ihre Werte auf den Basisvektoren b_i, vgl. Satz 5.1.1,

$$b_j^*(b_k) = \delta_{j,k} \ ,$$

woraus sofort folgt

$$b_j^*\left(\sum_{k=1}^n \alpha_k b_k\right) = \alpha_j \ .$$

Dann ist $\mathcal{B}^* = \{b_1^*, \ldots, b_n^*\}$ eine Basis von V^*. Das ist ein Spezialfall von Korrollar 5.2.8, in dem wir Basen für beliebige $\mathrm{Hom}(V, W)$ angegeben hatten. (Es lässt sich auch leicht direkt zeigen.)

Definition 12.1.4. Die Basis $\mathcal{B}^*$ von V^* heißt die zu $\mathcal{B}$ *duale Basis*. Ist $\mathcal{B}$ eine geordnete Basis, so ist auch die duale Basis $\mathcal{B}^*$ in natürlicher Weise eine geordnete Basis.

Beispiel 12.1.5. Sei $V = K^n$ und $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V.
Schreibe als Spaltenvektoren

$$b_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = \sum_{k=1}^{n} b_{kj} e_k \ .$$

Die zu $\mathcal{B}$ duale Basis $\mathcal{B}^* = (b_1^*, \ldots, b_n^*)$ schreibe als Zeilenvektoren

$$b_k^* = (v_{k1}, \ldots, v_{kn}) = \sum_{i=1}^{n} v_{ki} e_i^* \ .$$

(Wobei wir e_i^* mental mit e_i^T identifizieren können, vgl. Beispiel 12.1.2.1.) Dann gilt

$$\delta_{jk} = b_k^*(b_j) = \sum_{i=1}^{n} v_{ki} e_i^* (\sum_{\ell=1}^{n} b_{\ell j} e_\ell) = (v_{k1}, \ldots, v_{kn}) \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = v_{k1} b_{1j} + \ldots + v_{kn} b_{nj} \ .$$

Hieraus folgt die Matrizengleichung

$$E_N = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \ldots & v_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \ldots & b_{nn} \end{pmatrix}$$

Die Matrizen sind folglich Inverse.

Lemma 12.1.6. *Sei V ein beliebiger K–Vektorraum und sei $U \leq V$. und $\phi \in U^*$. Dann gibt es eine Linearform $\phi' \in V^*$, deren Einschränkung auf U gleich ϕ ist.*

Insbesondere gibt es für jedes $v \in V \setminus \{0\}$ eine Linearform $\varphi \in V^*$ mit $\varphi(v) \neq 0$.

Beweis:. Wir wählen eine Basis $\mathcal{B}$ von U und ergänzen zu einer Basis $\mathcal{B} \cup \mathcal{B}'$ von V mithilfe von Satz 6.0.3. Dann ist $V = U \oplus U'$ mit $U' = \mathrm{span}_K(\mathcal{B}')$. Dann definieren wir die lineare Abbildung $\phi' \in V^*$ durch $\phi : (u, u') \mapsto \phi(u)$.

Für den letzten Satz betrachten wir $U = \mathrm{span}_K(v)$ und $\phi(\lambda v) = \lambda$. □

Bemerkung 12.1.7. Sei $\dim_K V < \infty$. Da $\dim_K V = \dim_K V^*$ sind ein endlich-dimensionaler Vektorraum und sein Dualraum isomorph. Jede geordnete Basis $\mathcal{B}$ von V liefert einen Isomorphismus

$$\begin{array}{rccc} \Psi_{\mathcal{B}} : & V & \to & V^* \\ & b_i & \mapsto & b_i^*. \end{array}$$

Dieser Isomorphismus hängt jedoch von der Auswahl unserer Basis ab: Für verschiedene Basen $\mathcal{B}, \mathcal{B}'$ sind die Isomorphismen $\Psi_{\mathcal{B}}$ und $\Psi_{\mathcal{B}'}$ im allgemeinen verschieden. Wir sagen, es gibt keinen *kanonischen* Isomorphismus $V \cong V^*$.

Sei etwa $\alpha \in K \setminus \{0\}$, und betrachte eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ und ihre duale Bases $\mathcal{B}' = (b_1^*, \ldots, b_n^*)$ sowie die reskalierte Basis $\mathcal{B}' = (\alpha b_1, \ldots, \alpha b_n)$. Da gilt

$$b_i^*(\alpha b_j) = \alpha \delta_{ij}$$

ist die duale Basis der reskalierten Basis durch

$$(\mathcal{B}')^* = (\alpha^{-1} b_1^*, \ldots, \alpha^{-1} b_n^*)$$

gegeben.

Per Definition gilt für den von der Basis $\mathcal{B}'$ erzeugten Isomorphismus

$$\Psi_{\mathcal{B}'}(b'_i) = (b'_i)^*$$

und somit für alle $1 \leq i \leq n$

$$\Psi_{\mathcal{B}'}(b_i) = \Psi_{\mathcal{B}'}(\alpha^{-1}b'_i) = \alpha^{-1}\Psi_{\mathcal{B}'}(b'_i) = \alpha^{-1}(b'_i)^* = \alpha^{-2}b_i^* = \alpha^{-2}\Psi_{\mathcal{B}}(b_i) \ .$$

Es folgt $\Psi_{\mathcal{B}'} = \alpha^{-2}\Psi_{\mathcal{B}}$, so dass die Isomorphismen $\Psi_{\mathcal{B}'}$ und $\Psi_{\mathcal{B}}$ für $\alpha \neq \pm 1$ (und $n \geq 1$) verschieden sind.

Da es also für eine Identifizierung eines Vektorraums mit seinem Dualraum keinen *ausgezeichneten* Isomorphismus gibt, dürfen die beiden Vektorräume V und V^* selbst für endlichdimensionale Vektorräume nicht einfach miteinander identifiziert werden.

Dies zeigt sich zum Beispiel, wenn wir eine Abbildung $f : V \to W$ haben. Gibt es eine natürliche Abbildung $f' : V^* \to W^*$, die F entspricht? Mit der Auswahl von Basen $\mathcal{B}_V$ und $\mathcal{B}_W$ können wir $f' = \Psi_{\mathcal{B}_W} \circ f \circ \Psi_{\mathcal{B}_V}^{-1}$ definieren, aber für verschiedene Basen erhalten wir verschiedene Abbildungen!

Es gibt allerdings natürliche Abbildungen zwischen Dualräumen. Sie gehen allerdings in die andere Richtung!

Definition 12.1.8. Seien V, W zwei K–Vektorräume und sei $f : V \to W$ eine lineare Abbildung. Dann heißt die Abbildung

$$\begin{aligned} f^* : W^* &\to V^* \\ \varphi &\mapsto \varphi \circ f \end{aligned}$$

die zu f *duale Abbildung*.

Man beachte, dass

$$\varphi \circ f : \quad V \xrightarrow{f} W \xrightarrow{\varphi} K$$

tatsächlich eine Linearform auf V ist.

Lemma 12.1.9. *1. f^* ist eine lineare Abbildung.*

2. Seien V, W, Z jeweils K-Vektorräume und $f : V \to W$ und $g : W \to Z$ lineare Abbildungen. Dann gilt $(g \circ f)^* = f^* \circ g^*$.

Beweis:. 1. Seien $\varphi_1, \varphi_2 \in W^*$ und $\alpha_1, \alpha_2 \in K$. Wir rechnen für $v \in V$

$$\begin{aligned} f^*\left(\alpha_1\varphi_1 + \alpha_2\varphi_2\right)(v) &= \left(\alpha_1\varphi_1 + \alpha_2\varphi_2\right)(fv) \\ &= \alpha_1\varphi_1(fv) + \alpha_2\varphi_2(fv) \\ &= \alpha_1 f^*\varphi_1(v) + \alpha_2 f^*\varphi_2(v) \\ &= \left(\alpha_1 f^*\varphi_1 + \alpha_2 f^*\varphi_2\right)(v) \ . \end{aligned}$$

2. Es gilt für alle $\varphi \in Z^*$ wegen der Assoziativität der Verkettung von Abbildungen

$$(g \circ f)^*(\varphi) = \varphi \circ (g \circ f) = (\varphi \circ g) \circ f = (g^*\varphi) \circ f = f^*\left(g^*(\varphi)\right) = \left(f^* \circ g^*\right)(\varphi) \ .$$

□

Satz 12.1.10. *Seien V, W endlich–dimensionale K–Vektorräume mit geordneten Basen $\mathcal{A} = (a_1, \ldots, a_n)$ und $\mathcal{B} = (b_1, \ldots, b_m)$. Sei $f : V \to W$ eine lineare Abbildung. Dann ist die darstellende Matrix der dualen Abbildung bezüglich der dualen Basen:*

$$M_{\mathcal{A}^*}^{\mathcal{B}^*}(f^*) = M_{\mathcal{B}}^{\mathcal{A}}(f)^T \ .$$

Beweis:. Es gilt mit $M_{\mathcal{B}}^{\mathcal{A}}(f) = (c_{kj})$ und $M_{\mathcal{A}^*}^{\mathcal{B}^*}(f^*) = (d_{\mu\nu})$ nach der Definition 5.2.7 der darstellenden Matrix

$$f(a_j) = \sum_{k=1}^{m} c_{kj} b_k \quad \text{mit } m := \dim W, \qquad \text{für alle } j = 1, 2, \ldots, \dim V$$

$$f^*\left(b_\mu^*\right) = \sum_{\nu=1}^{n} d_{\nu\mu} a_\nu^* \quad \text{für alle } \mu = 1, 2, \ldots, \dim W \qquad \text{mit } n := \dim V$$

Wir vergleichen das Ergebnis der folgenden Rechnungen für festes $j = 1, \ldots, n$

$$b_\mu^*\left(f(a_j)\right) = b_\mu^*\left(\sum_{k=1}^{m} c_{kj} b_k\right) = \sum_{k=1}^{m} c_{kj} b_\mu^*(b_k) = c_{\mu j}$$

$$b_\mu^*\left(f(a_j)\right) = f^*\left(b_\mu^*\right)(a_j) = \sum_{\nu=1}^{n} d_{\nu\mu} a_\nu^*(a_j) = d_{j\mu} \ .$$

□

Damit haben wir auch der Operation aus Definition 5.2.1.2, eine Matrix zu transponieren, einen basisunabhängigen Sinn im Rahmen von Vektorräumen gegeben.

Korollar 12.1.11. *Eine lineare Abbildung $f : V \to W$ und ihre duale Abbildung $f^* : W^* \to V^*$ haben den gleichen Rang, es gilt* $\operatorname{rg} f^* = \operatorname{rg}(f)$.

Beweis:. Mit Satz 12.1.10 reicht es, den Rang einer Matrix mit dem der transponierten Matrix zu vergleichen. Aber das folgt aus Satz 5.5.12. □

Definition 12.1.12. Sei V ein K–Vektorraum und $M \subset V$ eine Teilmenge. Dann heißt

$$M^0 := \{\varphi \in V^* | \ \varphi(m) = 0 \quad \text{für alle } m \in M\} \subset V^*$$

der *Annullator* der Teilmenge M.

Lemma 12.1.13. *Es gilt:*

1. *Der Annullator M^0 ist ein Untervektorraum des Dualraums V^*.*
2. *Sei $U \leq V$ ein Untervektorraum und sei $\iota : U \to V$ die natürliche Injektion. Dann ist $U^0 = \ker(\iota^*)$.*
3. *Ist $\dim_K V < \infty$ und $U \leq V$ ein Untervektorraum, so gilt*

$$\dim_K U^0 = \dim_K V - \dim_K U \ .$$

Beweis:. 1. Lässt sich leicht aus der Definition prüfen.

2. Es gilt $U^0 = \{\phi \in V^* \mid \phi(\iota u) = 0 \quad \forall u \in U\} = \{\phi \in V^* \mid \iota^* \phi = 0\} = \ker(\iota^*)$.

3. Die Abbildung $\iota^* : V^* \to U^*$ ist surjektiv, das ist die Aussage von Lemma 12.1.6.

   Mit $U^0 = \ker(\iota^*)$ erhalten wir dann aus der Dimensionsformal und Betrachtung 12.1.3 $\dim_K U^0 = \dim V^* - \dim U^* = \dim V - \dim U$.

   Alternativ können wir den Bases per Hand führen: Ergänze eine geordnete Basis von U zu einer geordneten Basis $(u_1, \ldots, u_k, v_{k+1}, \ldots, v_n)$ von V. Sei $\mathcal{B}^* = (u_1^*, \ldots, u_k^*, v_{k+1}^*, \ldots, v_n^*)$ die duale Basis von V^*. Dann liegt

$$\varphi = \sum_{i=1}^{k} \alpha_i u_i^* + \sum_{j=k+1}^{n} \alpha_i v_i^* \ \in \ V^*$$

   genau dann in U^0, wenn für alle $j = 1, \ldots, k$ gilt

$$0 = \varphi(u_j) = \alpha_j$$

   Dies ist aber genau der Fall für $\varphi \in \operatorname{span}_K(v_{k+1}^*, \ldots, v_n^*)$. Somit ist $(v_{k+1}^*, \ldots, v_n^*)$ eine Basis von $U^0 \subset V^*$, und es gilt

$$\dim_K U^0 = n - k = \dim_K V - \dim_K U \ .$$

□

Beispiel 12.1.14. Sei $\dim_K V < \infty$ und $U \subset V$ eine *Hyperebene*, d.h. ein Untervektorraum der Dimension

$$\dim_K U = \dim_K V - 1 \ .$$

Nach Lemma 12.1.13.3 ist

$$\dim_K U^0 = n - (n-1) = 1 \ .$$

Also ist $U^0 = \operatorname{span}_K(\varphi)$ für jedes $\varphi \in U^0 \setminus \{0\}$. Für jede solche Linearform gilt aber φ

$$U = \{v \in V | \ \varphi(v) = 0\} \ .$$

Offensichtlich ist $U \leq \{v \in V \mid \phi(v) = 0\}$. Wenn die Inklusion echt ist, dann gibt es $u \in V \setminus U$ mit $\phi(u) = 0$. Aber dann verschwindet ϕ auf einem Untervektorraum, der U echt enthält und damit Dimension mindestens $\dim_K U + 1 = \dim_K V$ hat. Also ist $\phi = 0$, Widerspruch.

Wir können also eine Hyperebene durch eine Linearform beschreiben.

Betrachten wir nun allgemeiner eine *affine Hyperebene* $v_0 + U$, mit U wie oben und $v_0 \in V$, so gilt

$$\begin{aligned} v \in v_0 + U \ &\Leftrightarrow v - v_0 \in U \\ &\Leftrightarrow \varphi(v - v_0) = 0 \\ &\Leftrightarrow \varphi(v) = \varphi(v_0) =: c \end{aligned}$$

Wir finden also für jede affine Hyperebene eine Darstellung

$$U + v_0 = \{v \in V | \ \varphi(v) = c\}.$$

Satz 12.1.15. *Seien V, W endlich-dimensionale K–Vektorräume und sei $f : V \to W$ linear. Dann gilt für den Kern und das Bild der dualen Abbildung $f^* : W^* \to V^*$, dass sie sich als Annullatoren ausdrücken lassen:*

$$\begin{aligned} \ker f^* &= (\operatorname{Im} \ f)^0 \ \subset \ W^* \\ \operatorname{Im} \ f^* &= (\ker f)^0 \ \subset \ V^* \end{aligned}$$

Beweis:. Wir betrachten zuerst

$$\ker(f^*) = \{\phi \in W^* \mid \phi \circ f = 0\} = \{\phi \in W^* \mid \forall v \in V : \phi(f(v)) = 0\} = (\operatorname{Im} f)^0.$$

Für die zweite Gleichung beobachten wir dass $\phi \in \operatorname{Im} f^*$ die Form $v \mapsto \psi(f(v))$ für $\psi \in W^*$ hat. Aber damit verschwindet ϕ auf allen $v \in \ker(f)$ und liegt in $(\ker f)^0$. Wir haben also gezeigt, dass $\operatorname{Im} f^* \leq (\ker f)^0$. Gleichheit folgt, da beide Seiten die gleiche Dimension haben: Mit Satz 12.1.11 ist $\operatorname{rg}(f^*) = \operatorname{rg}(f)$ und mit Lemma 12.1.13 und der Dimensionsformel ist $\dim(\ker f)^0 = \dim V - \dim \ker(f) = \operatorname{rg}(f)$ □

Definition 12.1.16. Sei V ein K–Vektorraum. Dann heißt der K-Vektorraum

$$V^{**} := (V^*)^* = \operatorname{Hom}_K(V^*, K)$$

der *Bidualraum* von V.

Betrachtung 12.1.17. Betrachte zu einem Vektor $v \in V$ die Abbildung, die eine Linearform $\varphi \in V^*$ auf diesem Vektor auswertet:

$$\begin{aligned} \iota_V(v) : \ V^* &\to K \\ \varphi &\mapsto \varphi(v) \end{aligned}$$

Die Abbildung $\iota_V(v)$ ist linear, denn

$$\iota_V(v)(\alpha_1\varphi_1 + \alpha_2\varphi_2) = (\alpha_1\varphi_1 + \alpha_2\varphi_2)(v) = \alpha_1\varphi_1(v) + \alpha_2\varphi_2(v) = \alpha_1\iota_V(v)\varphi_1 + \alpha_2\iota_V(v)\varphi_2 \ ,$$

also ist $\iota_V(v) \in V^{**}$. Wir haben also für jeden K–Vektorraum V eine Abbildung

$$\iota_V : \ V \to V^{**}$$

Satz 12.1.18. *1. ι_V ist eine injektive lineare Abbildung.*

2. Ist $\dim_K V < \infty$, so ist ι_V ein Isomorphismus. Er heißt dann kanonischer Isomorphismus.

Beweis:. 1. Es gilt für $\alpha_1, \alpha_2 \in K$, $v_1, v_2 \in V$ und alle $\varphi \in V^*$

$$\iota_V(\alpha_1 v_1 + \alpha_2 v_2)(\varphi) = \varphi(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1\varphi(v_1) + \alpha_2\varphi(v_2) = (\alpha_1\iota_V(v_1) + \alpha_2\iota_V(v_2))(\varphi) \ .$$

Also ist ι_V eine lineare Abbildung. Sei $v \in V$ mit $\iota_V(v) = 0$. Dann gilt für alle $\varphi \in V^*$: $0 = \iota_V(v)\varphi = \varphi(v)$. Nach Lemma 12.1.6 muss dann aber $v = 0$ gelten.

2. Ist $\dim_K V < \infty$, so gilt

$$\dim_K V^{**} = \dim_K V^* = \dim_K V \ .$$

Nach 1. ist ι_V injektiv, also auch bijektiv.

□

Bemerkung 12.1.19. Die Abbildung ι_V hat eine besondere Eigenschaft: Sie ist *funktoriell*.

Seien V, W jeweils K–Vektorräume und sei $f : V \to W$ linear. Dann definieren wir $f^{**} = (f^*)^* : V^{**} \to W^{**}$ und es kommutiert das folgende Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & W \\ {\scriptstyle \iota_V}\downarrow & & \downarrow{\scriptstyle \iota_W} \\ V^{**} & \xrightarrow[\ f^{**}\ ]{} & W^{**} \end{array}$$

Sei nämlich $v \in V$ und $\varphi \in W^*$. Es gilt

$$\left(f^{**} \circ \iota_V(v)\right)(\varphi) = f^{**}\left(\iota_V(v)\right)(\varphi) = \iota_V(v)\left(f^*(\varphi)\right) = f^*(\varphi)(v) = \varphi\left(f(v)\right) \ .$$

Diese Gleichung folgen direkt aus den Definitionen, überlegen Sie sich bei jedem Term, in welchen Vektorräumen die verschiedenen Elemente leben!

Andererseits ist

$$\left(\iota_W \circ f(v)\right)\varphi = \iota_W(f(v)(\varphi) = \varphi\left(f(v)\right) \ .$$

Somit ist

$$f^{**} \circ \iota_V(v) = \iota_W \circ f(v)$$

für alle $v \in V$.

Die Isomorphismen $\Psi_{\mathcal{B}} : V \to V^*$, die von der Wahl einer Basis abhängen, sind dagegen nicht funktoriell.

12.2 Bilinearformen

Definition 12.2.1. 1. Seien V, W zwei K–Vektorräume. Eine Abbildung $\beta : \ V \times W \to K$ heißt *Bilinearform*, wenn gilt

$$\begin{aligned} \beta(v + v', w) &= \beta(v, w) + \beta(v', w) \\ \beta(v, w + w') &= \beta(v, w) + \beta(v, w') \\ \beta(\alpha v, w) &= \alpha\beta(v, w) = \beta(v, \alpha w) \end{aligned}$$

für alle $v, v' \in V, w, w' \in W$ und $\alpha \in K$.

2. Eine Bilinearform heißt *nicht-ausgeartet* im ersten (bzw. zweiten) Argument, falls gilt

$$\begin{aligned} \beta(v, w) = 0 \quad &\text{für alle } w \in W \ \Rightarrow v = 0 \\ \text{bzw. } \beta(v, w) = 0 \quad &\text{für alle } v \in V \ \Rightarrow w = 0 \ . \end{aligned}$$

Mit $\mathrm{Bil}_K(V, W)$ bezeichnen wir die Menge aller K–Bilinearformen

$$\beta : \ V \times W \to K \ .$$

Die Menge $\mathrm{Bil}_K(V, W)$ erhält durch punktweise Addition und Skalarmultiplikation die Struktur eines K–Vektorraums. Z.B. ist $(\beta + \beta')(v, w) = \beta(v, w) + \beta'(v, w)$.

Beispiele 12.2.2. 1. Die Multiplikation auf dem Körper K

$$m : \ K \times K \to K$$

liefert eine Bilinearform auf K.

2. Sei $B = (b_{ij}) \in M(m \times n, K)$. Dann definiert

$$\begin{aligned} \beta : K^m \times K^n &\to K \\ (x, y) &\mapsto x^T \cdot B \cdot y \end{aligned}$$

eine Bilinearform.

3. Für jeden K–Vektorraum V liefert die *Evaluationsabbildung*

$$\begin{aligned} ev_V : V^* \times V &\to K \\ (\beta, v) &\mapsto \beta(v) =: \langle \beta, v \rangle \end{aligned}$$

eine Bilinearform. Wegen Lemma 12.1.6 folgt aus $\langle \beta, v \rangle = 0$ für alle $\beta \in V^*$, dass $v = 0$. Also ist die Bilinearform nicht-ausgeartet im zweiten Argument. Gilt $\langle \beta, v \rangle = 0$ für alle $v \in V$, so ist nach Definition β der Nullvektor in V^*; also ist die Bilinearform $\langle \cdot, \cdot \rangle$ auch im ersten Argument nicht-ausgeartet.

4. Sei $V = W = C^0([a, b])$ der Raum aller stetigen Funktionen auf einem Intervall $[a, b] \subset \mathbb{R}$. Dann ist $(f, g) \mapsto \int_a^b f(x)g(x)dx$ eine Bilinearform.

Auf endlich-dimensionalen Vektorräumen sind Bilinarformen durch Ihre Werte auf Basisvektoren bestimmt.

Betrachtung 12.2.3. Sei V ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{A} = (v_1, \ldots, v_n)$ und W ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{B} = (w_1, \ldots, w_m)$. Setze für $\beta \in \mathrm{Bil}_K(V, W)$ $\beta_{ij} := \beta(v_i, w_j) \in K$. Die Matrix $B = (\beta_{ij}) \in M(n \times m, K)$ legt β eindeutig fest: ist $x = \sum_{i=1}^n x_i v_i$ und $y = \sum_{i=1}^m y_i w_i$, so ist

$$\beta(x, y) = \sum_{i,j} x_i \beta(v_i, w_j) y_j = x^T \cdot B \cdot y \; .$$

Matrizen können also sowohl lineare Abbildungen als auch Bilinearformen beschreiben. Man sollte sich bei einer Matrix – die ja eigentlich nur eine rechteckige Anordnung von Skalaren ist – also stets klarmachen, welches mathematische Objekt sie beschreibt.

Definition 12.2.4. Sei V ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{A} = (v_1, \ldots, v_n)$ und W ein endlich–dimensionaler K–Vektorraum mit geordneter Basis $\mathcal{B} = (w_1, \ldots, w_m)$. Die Matrix $M_{\mathcal{A},\mathcal{B}}(\beta) \in M(n \times m, K)$ mit Einträgen $\beta_{ij} := \beta(v_i, w_j) \in K$ heißt *darstellende Matrix* der Bilinearform β bezüglich der geordneten Basen $\mathcal{A}, \mathcal{B}$.

Beachten Sie, dass wir beide Basen im Subskript verwenden, anstatt eine ins Superskript zu setzen. Das soll uns helfen, Matrizen für lineare Abbildungen und bilinearformen auseinanderzuhalten.

Wir können Bilinearformen mit Hilfe linearer Abbildungen untersuchen.

Lemma 12.2.5. *1. Die Abbildung*

$$\begin{aligned} \mathrm{Bil}_K(V, W) &\to \mathrm{Hom}_K(V, W^*) \\ \beta &\mapsto \beta^\# \end{aligned}$$

mit $\beta^\# : x \mapsto \beta(x, -)$ ist ein Isomorphismus von K–Vektorräumen. Hier bezeichnet $\beta(x, -)$ die lineare Abbildung $w \mapsto \beta(x, w)$.

2. *Gilt $\dim_K V < \infty$ und $\dim_K W < \infty$, so ist $\dim_K \mathrm{Bil}_K(V, W) = \dim_K V \cdot \dim_K W$.*

3. *Die Bilinearform β ist genau dann nicht-ausgeartet im ersten Argument, falls die lineare Abbildung $\beta^\#$ injektiv ist.*

4. *Seien $\beta \in \mathrm{Bil}(V, W)$ und sei $\mathcal{A} = (v_1, \ldots, v_n)$ eine geordnete Basis von V, $\mathcal{B} = (w_1, \ldots, w_m)$ eine geordnete Basis von W, und $\mathcal{A}^*$, $\mathcal{B}^*$ die dualen Basen. Dann gilt $M_{\mathcal{A},\mathcal{B}}(\beta) = M^{\mathcal{A}}_{\mathcal{B}^*}(\beta^\#)^T$.*

Beweis:. 1. Da β linear im zweiten Argument ist, ist $\beta(x,-)$ für jedes x eine lineare Abbildung. Die Zuweisung $\beta \mapsto \beta^{\#}$ ist per Definition linear. Wir geben eine Umkehrabbildung an: für eine lineare Abbildung $h:\ V \to W^*$, definiere eine Bilinearform $\beta_h:\ V \times W \to K$ mit Hilfe der Evaluationsabbildung $\langle -,- \rangle : W^* \times W \to K$:

$$\beta_h(x,y) := \langle hx, y \rangle$$

Sie können sich vergewissern, dass diese Abbildungen inverse zueinander sind.

2. folgt sofort aus 1.

3. folgt deswegen, da $\beta^{\#}$ genau dann nicht injektiv ist, wenn es $v \neq 0$ gibt mit $\beta(v,-) = 0$, also $\beta(v,w) = 0$ für alle $w \in W$, d.h. β ist im ersten Argument ausgeartet.

4. Für die darstellende Matrix von $\beta^{\#} : V \to W^*$ gilt

$$\beta^{\#}(v_i) = \sum_{j=1}^{m} M^{\mathcal{A}}_{\mathcal{B}^*}(\beta^{\#})_{ji} w_j^* \ .$$

Damit folgt

$$M_{\mathcal{A},\mathcal{B}}(\beta)_{ij} = \beta(v_i, w_j) = \langle \beta^{\#}(v_i), w_j \rangle = M^{\mathcal{A}}_{\mathcal{B}^*}(\beta^{\#})_{ji} \ ,$$

was zu zeigen war. □

Satz 12.2.6. Transformationsformel

1. *Sei V ein endlich–dimensionaler K–Vektorraum mit geordneten Basen $\mathcal{A}$ und $\mathcal{A}'$ und W ein endlich–dimensionaler K–Vektorraum mit geordneten Basen $\mathcal{B}$ und $\mathcal{B}'$. Sei*

$$\beta:\ V \times W \to K$$

eine Bilinearform. Dann gilt

$$M_{\mathcal{A},\mathcal{B}}(\beta) = \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)^T \cdot M_{\mathcal{A}'\mathcal{B}'}(\beta) \cdot T^{\mathcal{B}}_{\mathcal{B}'}$$

2. *Ist insbesondere $V = W$ und wählt man $\mathcal{A} = \mathcal{B}$ und $\mathcal{A}' = \mathcal{B}'$, so gilt*

$$M_{\mathcal{A}}(\beta) = \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)^T \cdot M_{\mathcal{A}'}(\beta) \cdot T^{\mathcal{A}}_{\mathcal{A}'}$$

Beweis:. • Wir rechnen mit $v_j = \sum_{j'} \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)_{j'j} v'_{j'}$t, siehe Satz 5.2.7, dann ist $v_j = \sum_{j'} \left(T^{\mathcal{B}}_{\mathcal{B}'}\right)_{j'j} w'_{j'}$ und es gilt

$$\begin{aligned} M_{\mathcal{A},\mathcal{B}}(\beta)_{ij} = \beta(v_i, w_j) &= \sum_{i',j'} \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)_{i'i} \beta(v_{i'}, w_{j'}) \left(T^{\mathcal{B}}_{\mathcal{B}'}\right)_{j'j} \\ &= \sum_{i',j'} \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)_{i'i} M_{\mathcal{A}'\mathcal{B}'}(\beta)_{i'j'} \left(T^{\mathcal{B}}_{\mathcal{B}'}\right)_{j'j} \\ &= \left(\left(T^{\mathcal{A}}_{\mathcal{A}'}\right)^T \cdot M_{\mathcal{A}'\mathcal{B}'}(\beta) \cdot T^{\mathcal{B}}_{\mathcal{B}'}\right)_{ij} \end{aligned}$$

• 2. folgt als Spezialfall.

□

Man mache sich sorgfältig den Unterschied zu den Transformationsformeln in Satz 5.4.6 (*) für lineare Abbildungen und (**) für Endomorphismen klar. Diese Transformationsformeln haben uns auf die Äquivalenzrelationen "äquivalent" und "ähnlich" auf Räumen von Matrizen geführt. Hier finden wir eine weitere Relation, die wir nur für quadratische Matrizen einführen, also nur für die Situation $V = W$ und $\mathcal{A} = \mathcal{B}$ und $\mathcal{A}' = \mathcal{B}'$.

Definition 12.2.7. Seien $B, C \in M(n \times n, K)$. Wir sagen, C sei *kongruent* zu B über K und schreiben

$$C \simeq B \ ,$$

wenn es eine invertible quadratische Matrix $S \in GL(n, K)$ gibt, so dass

$$C = S^T B S$$

Dann stellen B und C bezüglich verschiedener Basen die gleiche Bilinearform dar.

Bemerkungen 12.2.8.

1. Die Determinanten kongruenter Matrizen unterscheiden sich um ein Quadrat in $K^\times = K \setminus \{0\}$:

$$\det C = \det\left(S^T B S\right) = (\det S)^2 \det B \ .$$

2. Kongruenz ist eine Äquivalenzrelation auf $M(n \times n, K)$.

3. Zwei Matrizen sind genau dann kongruent, wenn sie die gleiche Bilinearform bezüglich verschiedener Basen beschrieben. Dies wird wie in Lemma 5.5.3 gezeigt.

4. Um eine Matrix B in eine kongruente Matrix C zu überführen können wir simultane Zeilen- und Spaltenoperationen verwenden: Wenn T eine unserer Elementarmatrizen aus Lemma 5.3.7 ist, dann entsteht TBT^T indem wir die Zeilenumformung T_i und die Spaltenumformung T_i auf B anwenden. Konkret gilt:

   Sei $T = \Delta(1, \ldots, \lambda, \ldots, 1)$ mit λ an i-ter Stelle, dann multipzilizert $B \mapsto TBT^T = TBT$ die i-te Zeile und i-te Spalte mit λ.

   Sei $T = \tau(i, j)$, dann vertauscht $B \mapsto TBT^T = TBT$ die i-te und j-te Zeile und Spalte.

   Sei $T = \delta(i, j, \lambda)$ Dann addiert $B \mapsto TBT^T$ das λ-fache der i-ten Zeile zur j-ten Zeile und das λ-fache der i-ten Spalte zur j-ten Spalte.

   So lassen sich kongruente Matrizen oft durch eine Art Gaußverfahren finden.

Satz 12.2.9. *Sei $\beta \in \mathrm{Bil}(V, V)$ für einen endlich-dimensionalen Vektorraum V mit Basis $\mathcal{B}$. Es sind äquivalent:*

1. *β ist nicht-ausgeartet im ersten Argument.*
2. *$\beta^\#$ ist injektiv.*
3. *$\beta^\#$ ist ein Isomorphismus $V \cong V^*$*
4. *die darstellende Matrix $M_{\mathcal{B},\mathcal{B}}(\beta) = M^{\mathcal{B}}_{\mathcal{B}^*}(\beta^\#)^T$ ist invertibel*
5. *die Determinante der darstellenden Matrix $M_{\mathcal{B},\mathcal{B}}(\beta)$ ist ungleich Null*
6. *β ist nicht-ausgeartet im zweiten Argument.*

Wir sagen β ist nicht ausgeartet *wenn es eine der äquivalenten Bedingungen erfüllt.*

Beweis:. 1. und 2. sind äquivalent nach Lemma 12.2.5.3.

Die Äquivalenzen von 2. 3. 4. und 5. sind Standardergebnisse aus dem ersten Semester (in 4 verwenden wir auch Lemma 12.2.5.4).

Schließlich definieren wir die Bilinearform $\beta'(x,y) = \beta(y,x)$. Mit $B = M_{\mathcal{B},\mathcal{B}}(\beta)$ gilt $\beta(x,y) = x^T By$, aber da dies ein Skalar ist, gilt auch $\beta(y,x) = \beta(x,y) = \beta(x,y)^T = y^T B^T x$ und damit ist die darstellende Matrix von β' genau B^T. Es ist β nicht-ausgeartet im zweiten Argument, wenn β' nicht ausgeartet im ersten Argument ist, was genau dann gilt, wenn $\det(B^T) = \det(B) \neq 0$. Also ist 6. äquivalent zu 5. □

Korollar 12.2.10. *Eine nicht ausgeartete Bilinearform β in $\mathrm{Bil}_K(W,W)$ definiert einen Isomorphismus $\mathrm{Hom}(V,W) \cong \mathrm{Bil}_K(V,W)$, indem wir $f : V \to W$ nach $\gamma(v,w) = \beta(fv,w)$ schicken.*

Beweis:. Nach Satz 12.2.9 definiert β einen Isomorphismus $W \cong W^*$. Zusammen mit Lemma 12.2.5 erhalten wir den gewünschten Isomorphismus als Verknüpfung $\mathrm{Hom}(V,W) \cong \mathrm{Hom}(V,W^*) \cong \mathrm{Bil}_K(V,W)$. □

12.3 Symmetrische und Schiefsymmetrische Bilinearformen

Definition 12.3.1. 1. Eine Bilinearform $\beta \in \mathrm{Bil}_K(V,V)$, für die

$$\beta(x,y) = \beta(y,x) \text{ für alle } x,y \in V$$

gilt, heißt *symmetrische Bilinearform*.

2. Eine Bilinearform mit der Eigenschaft

$$\beta(x,y) = -\beta(y,x) \text{ für alle } x,y \in V$$

heißt *schief-symmetrisch*.

3. Eine Bilinearform $\beta \in \mathrm{Bil}_K(V,V)$, für die

$$\beta(x,x) = 0 \text{ für alle } x \in V$$

gilt, heißt *alternierende Bilinearform*.

4. Eine nicht-ausgeartete alternierende Bilinearform heißt *symplektische Bilinearform*. Ein Vektorraum mit einer symplektischen Bilinearform heißt ein *symplektischer Vektorraum*.

Bemerkung 12.3.2. Für jede alternierende Bilinearform gilt

$$0 = \beta(x+y,x+y) = \beta(x,x) + \beta(x,y) + \beta(y,x) + \beta(y,y) = \beta(x,y) + \beta(y,x)\ ;$$

sie ist also schief-symmetrisch.

Für eine schief-symmetrische Bilinearform gilt $\beta(x,x) = -\beta(x,x)$, also $2 \cdot \beta(x,x) = 0$. Gilt im Körper K, dass $1+1 \neq 0$ ist, so ist jede schiefsymmetrische K-Bilinearform auch alternierend. Gilt $1+1 = 0$ dann sind schief-symmetrische Bilinearformen genau die symmetrsichen Bilinearformen.

Bemerkung 12.3.3. Wir nehmen wieder an $1+1 \neq 0 \in K$. Dann ist $\frac{1}{2} \in K$. Dies gilt für all unsere üblichen Körper bis auf $\mathbb{F}_2$.

Dann ist jede Bilinearform β ist Summe einer symmetrischen und einer schief-symmetrischen Bilinaerform. Wir schreiben $\beta = \beta_s + \beta_a$ mit $\beta_s(x,y) = \frac{1}{2}(\beta(x,y) + \beta(y,x))$ und $\beta_a(x,y) = \frac{1}{2}(\beta(x,y) - \beta(y,x))$.

Lemma 12.3.4. *Eine Bilinearform β ist symmetrisch, falls für jede Basis $\mathcal{B}$ für $B = M_{\mathcal{B}}(\beta)$ gilt $B^T = B$. Eine Matrix mit $B^T = B$ heißt* symmetrisch.

Eine Bilinearform β ist schiefsymmetrisch, falls für jede Basis $\mathcal{B}$ für $B = M_{\mathcal{B}}(\beta)$ gilt $B^T = -B$. Eine solche Matrix heißt schiefsymmetrisch.

Beweis:. Für $\mathcal{B} = (b_1, \ldots, b_n)$ gilt für eine symmetrische Bilinearform $B_{ij} = \beta(b_i, b_j) = \beta(b_j, b_i) = B_{ji}$ und B ist symmetrisch. Ist umgekehrt $B = B^T$ dann gilt für $v, w \in V$ dass $\beta(v, w) = v^T B w = (v^T B w)^T = w^T B^T v = w^T B v = \beta(w, v)$. Hier haben wir verwendet, dass $\beta(v, w)$ als Skalar gleich seinem egienen Transponierten ist.

Der gleiche Beweis funktioniert für den zweiten Teli des Lemmas. □

Wir werden uns später noch ausführlicher mit den symmetrischen Bilinearformen beschäftigen. In diesem Abschnitt beweisen wir zwei Resultate über symplektische Bilinearformen.

Korollar 12.3.5. *Ein symplektischer K-Vektorraum hat gerade Dimension wenn in K gilt $1 + 1 \neq 0$.*

Beweis:. Aus $B^T = -B$ folgt für $B \in M(n \times n, K)$

$$\det B = \det B^T = \det(-B) = (-1)^n \det B \, .$$

Ist die Form nicht ausgeartet, so ist $\det B \neq 0$ und wenn $1 \neq -1$ muss n gerade sein. □

Satz 12.3.6. *Sei β eine symplektische Bilinearform auf einem K–Vektorraum V; im Körper K gelte $1 + 1 \neq 0$. Dann besitzt V eine* symplektische Basis*, d.h. eine geordnete Basis*

$$\mathcal{B} = (u_1, v_1, \ldots, u_m, v_m) \, ,$$

in der die darstellende Matrix die Block-diagonale Form

$$M_{\mathcal{B}}(\beta) = \begin{pmatrix} H & & & 0 \\ & H & & \\ & & \ddots & \\ 0 & & & H \end{pmatrix}$$

mit

$$H := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M(2 \times 2, K)$$

hat. Insbesondere ist $\dim_K V$ *gerade.*

Beweis:. Sei $u_1 \neq 0$ beliebig, $u_1 \in V$. Da β nicht-ausgeartet ist, finde einen Vektor $v_1 \in V$ mit $\beta(u_1, v_1) = 1$ und somit auch $\beta(v_1, u_1) = -1$. Die Familie (u_1, v_1) ist linear unabhängig. Denn gilt $\lambda u_1 + \mu v_1 = 0$, so folgt

$$0 = \beta(\lambda u_1 + \mu v_1, v_1) = \lambda \quad \text{und} \quad 0 = \beta(\lambda u_1 + \mu v_1, u_1) = -\mu \, .$$

Auf dem zwei-dimensionalen Untervektorraum $\operatorname{span}_K(u_1, v_1) =: U_1$ wird die Einschränkung von β in der Basis (u_1, v_1) durch die Matrix H dargestellt.

Wir betrachten nun das soganannte orthogonale Komplement ${}^{\perp}U_1 = \{v \in V \mid \beta(u, v) = 0 \text{ für alle } u \in U_1\}$.

Wir zeigen zuerst, dass $U_1 \cap {}^\perp U_1 = \{0\}$. Wir betrachten dafür $v \in U_1 \cap {}^\perp U_1$. Wegen $v \in U_1$ schreiben wir $v = \lambda u_1 + \mu v_1$. Es folgt dann aus $v \in {}^\perp U_1$ mit ähnlicher Rechnung wie eben

$$0 = \beta(v, v_1) = \lambda \quad \text{und} \quad 0 = \beta(v, u_1) = -\mu \ .$$

Da β nicht ausgeartet ist, folgt aus Lemma 12.3.7, das wir gleich zeigen werden, dass $\dim_K V = \dim_K U_1 + \dim_K {}^\perp U_1$, und somit erhalten wir die Zerlegung von V als direkte Summe,

$$V = U_1 \oplus {}^\perp U_1 \ .$$

Die Einschränkung $\beta|_{{}^\perp U_1}$ auf den Untervektorraum ${}^\perp U_1$ ist alternierend. Sie ist auch nicht ausgeartet: angenommen, es gäbe $v_0 \in {}^\perp U_1, v_0 \neq 0$ mit $\beta(v_0, v'') = 0$ für alle $v'' \in {}^\perp U_1$. Schreibe dann ein beliebiges $v \in V$ wegen der direkten Summerzerlegung von V in der Form $v = v' + v''$ mit $v' \in U_1$ und $v'' \in {}^\perp U_1$. Dann gilt

$$\beta(v_0, v) = \beta(v_0, v') + \beta(v_0, v'') = 0 + 0 = 0$$

im Widerspruch zu der Annahme, dass β auf ganz V nicht ausgeartet ist. Also trägt der Untervektorraum $\beta|_{{}^\perp U_1}$ eine symplektische Form, und wir können vollständige Induktion nach der Dimension von V anwenden. □

Lemma 12.3.7. *Sei V ein endlichdimensionaler Vektorraum mit nicht ausgearteter Bilinearform β. Sei $U \leq V$ und ${}^\perp U = \{v \in V \mid \beta(v, u) = 0$ für alle $u \in U\}$. Dann ist $\dim_K {}^\perp U = \dim_K(V) - \dim_K(U)$.*

Beweis:. Wir betrachten die Einschränkungsabbildung $r : V^* \to U^*$, die ϕ auf $\phi|_U$ abbildet. Nach Lemma 12.1.6 ist r surjektiv.

Nun stellen wir fest, dass ${}^\perp U$ per Definition der Kern von $r \circ \beta^\#$ ist, denn $v \in {}^\perp U$ genau wenn $\beta^\#(v)$ eingeschränkt auf U gleich 0 ist.

Da β nicht ausgeartet ist, ist $\beta^\#$ ein Isomorphismus, damit ist $\operatorname{rg}(r \circ \beta^\#) = \operatorname{rg}(r) = \dim(U)$. Aus der Dimensionsformel folgt nun $\dim {}^\perp U = \dim V - \dim U$. □

12.4 Quadratische Formen

Definition 12.4.1. 1. Sei V ein K–Vektorraum. Eine *quadratische Form* auf V ist eine Abbildung

$$q : V \to K$$

mit den beiden Eigenschaften:

(QF1) Für alle $\lambda \in K$ und $v \in V$ gilt $q(\lambda v) = \lambda^2 q(v)$.

(QF2) Die Abbildung

$$\begin{aligned} \beta_q : V \times V &\to K \\ (v, w) &\mapsto q(v + w) - q(v) - q(w) \end{aligned}$$

ist eine Bilinearform auf V.

2. Eine quadratische Form q heißt *nicht-ausgeartet*, wenn die zugehörige Bilinearform β_q nicht-ausgeartet ist.

Beispiele 12.4.2. 1. Sei $V = K$; dann ist $q : K \to K$ mit $x \mapsto cx^2$ mit $c \in K$ eine quadratische Form. Denn es gilt

$$q(\lambda x) = c(\lambda x)^2 = \lambda^2 cx^2 = \lambda^2 q(x) \ ,$$

und

$$\beta_q(x, y) = c(x + y)^2 - cx^2 - cy^2 = 2cxy$$

ist eine Bilinearform auf K. Die Form ist genau dann nicht-ausgeartet, wenn $2c \in K \setminus \{0\}$ gilt.

2. Sei $\beta \in \mathrm{Bil}(V, V)$ und $\frac{1}{2} \in K$ dann definiert $q_\beta : x \mapsto \frac{1}{2}\beta(x, x)$ eine quadratische Form auf V, denn $\beta(\lambda x, \lambda x) = \lambda^2\beta(x, x)$ (das ist QF1) und es gilt $\beta_{q_\beta}(x, y) = \frac{1}{2}\beta(x + y, x + y) - \frac{1}{2}\beta(x, x) - \frac{1}{2}\beta(y, y) = \frac{1}{2}\beta(x, y) + \frac{1}{2}\beta(y, x)$ (das ist eine Bilinearform und QF2 gilt). Wir erhalten auch eine quadratische Form wenn wir $\frac{1}{2}$ durch irgendein anderes $\lambda \in K$ ersetzen.

3. Für $a, b \in \mathbb{R}_{>0}$ sind die Formen $p(x, y) = ax^2 + by^2$ und $q(x, y) = ax^2 - by^2$ quadratische Formen. Für ein festes positives $c \in \mathbb{R}$ können wir die Gleichungen $p(x, y) = c$ und $q(x, y) = c$ betrachten.

   Die Gleichung $p(x, y) = c$ beschreibt eine Ellipse. Die Gelichung $q(x, y) = c$ beschreibt eine Hyperbel. Um dies zu sehen, betrachten wir den Koordinatenwechsel $(x, y) \mapsto (x', y') = (\sqrt{a}x - \sqrt{b}y, \sqrt{a} + \sqrt{b}y)$. Dann wird unsere Gleichung zu $x'y' = c$, oder $y' = \frac{c}{x'}$.

   Während also Linearformen affine Geraden beschreiben können wir durch quadratische Formen interessantere geometrische Objekte beschreiben.

Das letzte Beispiel legt nahe, dass quadratische Formen und symmetrische Bilinearformen sich entsprechen.

Satz 12.4.3. *Sei K ein Körper, in dem $1 + 1 \neq 0$ gilt und V ein K-Vektorraum. Dann sind der Vektorraum $\mathrm{QF}(V)$ aller quadratischen Formen und der Vektorraum $\mathrm{SBil}_K(V, V)$ der symmetrischen Bilinearformen isomorph:*

$$\begin{aligned} \mathrm{QF}(V) &\to \mathrm{SBil}_K(V, V) \\ q &\mapsto \beta_q \\ \mathrm{SBil}_K(V, V) &\to \mathrm{QF}(V) \\ \beta &\mapsto q_\beta \end{aligned}$$

mit $q_\beta(x) = \frac{1}{2}\beta(x, x)$.

Beweis:. Wir haben in Beispiel 12.4.2.2 gesehen, dass für eine gegebene symmetrische Bilinearform β die Abbildung q_β eine quadratische Form ist und $\beta_{q_\beta} = \beta$ gilt.

Per Definition ist β_q eine Bilinearform für $q \in QF(V)$ und offensichtlich symmetrisch. Schließlich gilt $q_{\beta_q}(v) = \frac{1}{2}\beta_q(v, v) = \frac{1}{2}q(v + v) - \frac{1}{2}q(v) - \frac{1}{2}q(v) = q(v)$ und die beiden Konstruktionen sind invers. Ferner gilt □

Die Gleichung

$$\beta(x, y) = q_\beta(x + y) - q_\beta(x) - q_\beta(y)$$

heißt *Polarisierungsformel* für die Bilinearform β.

Bemerkung 12.4.4. Die Situation ist für einen Körper K, in dem $1+1=0$ gilt, anders. Sei $K=\mathbb{F}_2$ und $q: K \to K$ eine quadratische Form. Dann gilt wegen (QF1) $q(0) = q(0\cdot 0) = 0^2 q(0) = 0$, wohingegen (QF1) den Wert von $q(1)$ nicht einschränkt, also $q(1) \in \{0,1\}$. Für die zugehörige symmetrische Bilinearform β gilt in beiden Fällen

$$\beta(0,0) = \beta(0,1) = \beta(1,0) = 0 \quad \text{und} \quad \beta(1,1) = q(0) - q(1) - q(1) = -2q(1) = 0 \ .$$

Hier führen also zwei verschiedene quadratische Formen auf die gleiche symmetrische Bilinearform. Die quadratische Form enthält also mehr Information als die symmetrische Bilinearform. Umgekehrt gibt es zur symmetrischen Bilinearform mit $\beta(1,1)=1$ keine quadratische Form.

Satz 12.4.5. *Sei K ein Körper, in dem $1+1\neq 0$ gilt. Sei V ein endlich–dimensionaler K–Vektorraum und β eine symmetrische Bilinearform auf K. Dann existiert eine geordnete Basis $\mathcal{B}=(b_1,\ldots,b_n)$ von V, in der die darstellende Matrix $M_{\mathcal{B}}(\beta)$ eine Diagonalmatrix ist, d.h. $\beta(b_i,b_j)=0$ für $i\neq j$.*

Beweis:. Durch vollständige Induktion nach $n := \dim V$. Für den Induktionsanfang $n=1$ ist nichts zu zeigen.

- Gilt $q_\beta(v)=0$ für alle $v\in V$, so folgt aus Satz 12.4.3 $\beta=0$, also $M_{\mathcal{B}}(\beta)=0_n$ in jeder Basis $\mathcal{B}$ von V. In diesem Fall ist der Satz offensichtlich wahr.
- Sei also $b_1\in V$ mit $q_\beta(b_1)\neq 0$. Die Linearform

$$\begin{aligned} \varphi: \ V &\to K \\ v &\mapsto \beta(b_1,v) \end{aligned}$$

  ist wegen $\varphi(b_1)=\beta(b_1,b_1)\neq 0$ nicht die Nullform. Also ist die Dimension des Untervektorraums

$$U := \ker\varphi$$

  gleich $\dim_K U = n-1$. (Wir haben hier verschieden Möglichkeiten, das gleiche auszudrücken: Die Form ϕ ist genau $\beta^{\#}(b_1)$, den Raum U könnten wir auch als $\{v \mid \beta(b_1,v)=0\} = \mathrm{span}_K(b_1)^\perp$ schreiben.)

  Nach Induktionsannahme finde eine geordnete Basis $(b_2,\ldots,b_n)$ des Unterraums U mit $\beta(b_i,b_j)=0$ für $i\neq j$ und $i,j\geq 2$. Wegen $b_1\notin U$ ist $(b_1,\ldots,b_n)$ eine geordnete Basis von V. In ihr hat die symmetrische Bilinearform β wegen

$$\beta(b_1,b_i) = \varphi(b_i) = 0 \quad \text{für} \quad i=2,\ldots,n$$

  die gewünschte Diagonalgestalt. □

Die diagonalisierende Basis $\mathcal{B}$ aus Satz 12.4.5 ist nicht eindeutig. Auch die Diagonalelemente sind nicht einmal bis auf die Reihenfolge eindeutig. Allerdings haben kongruente Diagonalmatrizen die gleiche Anzahl r_0 von Nullen auf der Diagonale.

Definition 12.4.6. Der *Nullraum* einer symmetrischen Bilinearform β auf einem K-Vektorraum V ist der Untervektorraum

$$N(\beta) := \{v\in V \mid \ \beta(v,w)=0 \text{ für alle } w\in V\} \subset V \ .$$

Lemma 12.4.7. *Für jede diagonalisierende Basis $\mathcal{B}=(b_1,\ldots,b_n)$ wie in Satz 12.4.5 gilt*

$$N(\beta) = \mathrm{span}_K\{b_i \mid \beta(b_i,b_i)=0\} \ .$$

Beweis:. "$\supset$" Sei i so gewählt, dass für das Basiselement b_i die Gleichung $\beta(b_i, b_i) = 0$ gilt. Dann ist für jedes $w = \sum_{j=1}^{n} w_j b_j \in V$

$$\beta(b_i, w) = \sum_{j=1}^{n} w_j \beta(b_i, b_j) = w_i \beta(b_i, b_i) = 0 \ ,$$

also gilt $b_i \in N(\beta)$. Da $N(\beta)$ per Definition ein Untervektorraum ist, gilt auch $\mathrm{span}_K\{b_i \mid \alpha_i = 0\} \subset N(\beta)$.

"$\subset$" Sei $v \in N(\beta)$ beliebig. Schreibe $v = \sum_{j=1}^{n} v_j b_j$ und finde

$$0 = \beta(v, b_i) = \sum_{j=1}^{n} v_j \beta(b_j, b_i) = v_i \beta(b_i, b_i); \, .$$

Ist $\beta(b_i, b_i) \neq 0$, so muss $v_i = 0$ gelten. Also ist jedes $v \in N(\beta)$ eine Linearkombination derjenigen Basisvektoren b_i mit $\beta(b_i, b_i) = 0$.

□

Der Vergleich der Dimensionen zeigt dann die Gleichheit $r_0 = \dim_K N(\beta)$. Da der Nullraum nicht von der Wahl einer Basis von V abhängt, ist auch r_0 als Dimension der rechten Seite basisunabhängig. Ist $M_{\mathcal{B}}(\beta) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$, so ist β genau dann nicht-ausgeartet, wenn $\det M_{\mathcal{B}}(\beta) \neq 0$ gilt, also genau dann, wenn $\alpha_j \neq 0$ für alle j gilt, also wenn $r_0 = 0$ gilt.

Wir wollen nun speziell quadratische Formen über den Körper $\mathbb{C}$ und $\mathbb{R}$ der komplexen bzw. reellen Zahlen untersuchen.

Satz 12.4.8. *Sei V ein endlich–dimensionaler $\mathbb{C}$–Vektorraum und*

$$\beta : \ V \times V \to \mathbb{C}$$

eine symmetrische Bilinearform. Dann existiert eine geordnete Basis $\mathcal{B}$ von V, in der die darstellende Matrix die Gestalt

$$M_{\mathcal{B}}(\beta) = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{0, \ldots, 0}_{r_0})$$

hat. Die nicht-negativen ganzen Zahlen r und r_0 sind durch β eindeutig bestimmt und hängen nicht von der Wahl der geordneten Basis $\mathcal{B}$ ab.

Beweis:. Wegen Satz 12.4.5 gibt es eine geordnete Basis $\mathcal{B}' = (b'_1, \ldots, b'_n)$, in der gilt

$$M_{\mathcal{B}'}(\beta) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$$

Setze $b_i := \gamma_i b'_i$ mit

$$\gamma_i := \begin{cases} \frac{1}{\sqrt{\alpha_i}} & \text{falls } \alpha_i \neq 0 \\ 1 & \text{falls } \alpha_i = 0 \ . \end{cases}$$

Wir finden

$$\beta(b_i, b_j) = \gamma_i \gamma_j \beta(b'_i, b'_j) = \gamma_i \gamma_j \alpha_i \delta_{i,j} \ .$$

Es folgt für $\alpha_i = 0$ die Gleichung $\beta(b_i, b_i) = 0$ und für $\alpha_i \neq 0$ die Gleichung

$$\beta(b_i, b_j) = \frac{\alpha_i}{\alpha_i} = 1 \ .$$

Durch Umnummerierung erhält man die gewünschte Form.

Die Unabhängigkeit von r_0 und somit auch von r von der Wahl der geordneten Basis $\mathcal{B}$ folgt aus Lemma 12.4.7. □

Für quadratische Formen über $\mathbb{R}$ ist die Situation etwas komplizierter, weil nicht jede reelle Zahl das Quadrat einer reellen Zahl ist.

Satz 12.4.9. Sylvesterscher Trägheitssatz *Sei V ein endlich–dimensionaler $\mathbb{R}$–Vektorraum und*

$$\beta : \ V \times V \to \mathbb{R}$$

eine symmetrische Bilinearform. Dann existiert eine geordnete Basis $\mathcal{B}$ von V, in der die darstellende Matrix die Gestalt

$$M_{\mathcal{B}}(\beta) = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r_+}, \underbrace{-1, -1, \ldots, -1}_{r_-}, \underbrace{0, \ldots, 0}_{r_0})$$

hat. Die nicht-negativen ganzen Zahlen r_+, r_- und r_0 sind durch die symmetrische Bilinearform β eindeutig bestimmt und hängen nicht von der Wahl der geordneten Basis $\mathcal{B}$ ab.

Definition 12.4.10. Das Tripel

$$(r_+, r_-, r_0)$$

heißt die *Signatur* einer symmetrischen Bilinearform (oder äquivalent einer quadratischen Form). Weiterhin heißt r_- auch der *Trägheitsindex* von q. Manchmal wird in der Literatur auch die ganze Zahl $r_+ - r_-$ als Signatur bezeichnet.

Beweis:.

- Wegen Satz 12.4.5 gibt es eine geordnete Basis $\mathcal{B}' = (b'_1, \ldots, b'_n)$, in der gilt

  $$M_{\mathcal{B}'}(\beta) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$$

  Setze $b_i := \gamma_i b'_i$ mit

  $$\gamma_i := \begin{cases} \frac{1}{\sqrt{|\alpha_i|}} & \text{falls } \alpha_i \neq 0 \\ 1 & \text{falls } \alpha_i = 0 \, . \end{cases}$$

  Wir finden

  $$\beta(b_i, b_j) = \gamma_i \gamma_j \beta(b'_i, b'_j) = \gamma_i \gamma_j \alpha_i \delta_{i,j} \, .$$

  Es folgt für $\alpha_i = 0$ die Gleichung $\beta(b_i, b_i) = 0$ und für $\alpha_i \neq 0$ die Gleichung

  $$\beta(b_i, b_j) = \frac{\alpha_i}{|\alpha_i|} = \mathrm{sign}(\alpha_i) \in \{\pm 1\} \, .$$

  Durch Umnummerierung erhält man die gewünschte Form.

- Wir wissen schon aus Lemma 12.4.7, dass r_0 als Dimension des Nullraums von β nicht von der Wahl der diagonalisierenden Basis abhängt.

- Wir zeigen:

  $$r_+ = \max\{\dim_{\mathbb{R}} W \mid W \subset V \text{ Untervektorraum mit } q(v) > 0 \text{ für alle } v \in W \setminus \{0\}\} \, .$$

  Die rechte Seite bezeichnen vorübergehend mit m. Aus dieser Darstellung von r_+ folgt sofort, dass r_+ und damit auch r_- nicht von der Wahl der Basis $\mathcal{B}$ abhängt. Zum Beweis betrachte zunächst den speziellen Untervektorraum

  $$W_0 := \mathrm{span}_{\mathbb{R}}\{b_1, \ldots, b_{r_+}\} \, .$$

Für

$$v = \sum_{j=1}^{r_+} v_j b_j \in W_0 \setminus \{0\} \quad \text{mit} \quad v_j \in \mathbb{R}$$

gilt wegen $\beta(v,v) = 2q(v)$:

$$2q(v) = \beta(v,v) = \sum_{i,j=1}^{r_+} v_i v_j \beta(b_i, b_j) = \sum_{i=1}^{r_+} (v_i)^2 > 0 \; .$$

Daraus folgt für m als Maximum die Ungleichung

$$m \geq \dim_{\mathbb{R}} W_0 = r_+ \; .$$

Um die entgegengesetzte Ungleichung zu zeigen, zeigen wir, dass in jedem Untervektorraum W von V mit $\dim_{\mathbb{R}} W > r_+$ ein Vektor $v \in W \setminus \{0\}$ mit $q(v) \leq 0$ existiert. Sei also $\dim_{\mathbb{R}} W > r_+$. Wegen

$$\dim_{\mathbb{R}} W + \dim_{\mathbb{R}} \operatorname{span}_{\mathbb{R}}\{b_{r_+ +1}, \ldots, b_n\} > r_+ + r_0 + r_- = \dim_{\mathbb{R}} V$$

finde mit Satz 4.3.10 ein $v \neq 0$ in

$$v \in W \cap \operatorname{span}_{\mathbb{R}}\{b_{r_+ +1}, \ldots, b_n\} \; ,$$

Es gibt also ein $v \in W$, das die Darstellung

$$v = \sum_{j=r_+ +1}^{n} v_j b_j \quad \text{mit} \quad v_j \in \mathbb{R}$$

besitzt. Aus dieser Darstellung folgt die gewünschte Ungleichung:

$$2q(v) = \beta(v,v) = \sum_{j=r_+ +1}^{n} (v_j)^2 \, \beta(b_j, b_j) \leq 0 \; .$$

- Analog gilt

$$r_- = \max\{\dim_{\mathbb{R}} U \mid U \subset V \text{ mit } q(v) < 0 \text{ für alle } v \in U \setminus \{0\}\} \; .$$

□

Definition 12.4.11. Sei V ein $\mathbb{R}$-Vektorraum.

1. Eine quadratische Form

$$q : \; V \to \mathbb{R}$$

(oder die dazugehörige Bilinearform) heißt

- *positiv definit*, falls $q(v) > 0$ für alle $v \in V$ mit $v \neq 0$ gilt, d.h. falls $r_- = r_0 = 0$ gilt.
- *negativ definit*, falls $q(v) < 0$ für alle für alle $v \in V$ mit $v \neq 0$ gilt, d.h. falls $r_+ = r_0 = 0$ gilt..
- *positiv semidefinit*, falls $q(v) \geq 0$ für alle $v \in V$ gilt, d.h. falls $r_- = 0$ gilt.
- *negativ semidefinit*, falls $q(v) \leq 0$ für alle $v \in V$ gilt, d.h. falls $r_+ = 0$ gilt.
- *indefinit*, falls es v_1 mit $q(v_1) > 0$ und v_2 mit $q(v_2) < 0$ gibt, d.h. falls $r_+ > 0$ und $r_- > 0$ gilt.

Beispiel 12.4.12. Die quadratische Form $m : (x,y,z,t) \mapsto x^2 + y^2 + z^2 - c^2 t^2$ auf $\mathbb{R}^4$ ist nicht ausgeartet und indefinit, mit Signatur $(3,1)$. Es ist die sogenannte *Minkowski-Metrik* und spielt eine wichtige Rolle in der (speziellen und allgemeinen) Relativitätstheorie. Der "Abstand" von zwei Punkten in der Raumzeit mit Koordinaten (x_1, y_1, z_2, t_1) und (x_2, y_2, z_2, t_2) ist $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2 (t_1 - t_2)^2}$.

12.5 Vektorräume mit innerem Produkt

Eine positive definite Bilinearform auf einem reellen Vektorraum heißt auch *inneres Produkt* oder *Skalarprodukt*. Sie gibt uns einen Abstandsbegriff für Vektoren und verallgemeinert das bekannte Skalarprodukt.

Definition 12.5.1. 1. Sei V ein $\mathbb{R}$–Vektorraum. Eine positiv–definite symmetrische Bilinearform

$$\langle \cdot, \cdot \rangle : \ V \times V \to \mathbb{R}$$

heißt *inneres Produkt* oder ein *(euklidisches) Skalarprodukt*.

2. Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt *euklidischer Vektorraum*.

3. Die Funktion $\| \cdot \| : V \to \mathbb{R}_{\geq 0}$
$$\|x\| := \sqrt{\langle x, x \rangle}$$
auf einem euklidischen Vektorraum heißt *Norm*.

Beispiele 12.5.2. 1. Auf $V = \mathbb{R}^n$ heißt $\langle x, y \rangle := \sum_{i=1}^n x_i y_i = x^T E_n y$ das *Standard-Skalarprodukt*. Die Norm der Differenz zwischen zwei Vektoren x, y, $||x - y||$, lässt sich als Abstand der Vektoren interpretieren.

2. Der reelle Vektorraum $C^0([a, b], \mathbb{R})$ der stetigen Funktionen auf dem abgeschlossenen Interval $[a, b]$ wird durch
$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$
zum euklidischen Vektorraum.

Wir führen nun das komplexe Analogon euklidischer Vektorräume ein.

Auch zwischen zwei komplexen Vektoren soll der Abstand eine reelle Zahl sein. Wir erinnern uns an den Absolutbetrag $|z| = \sqrt{z\overline{z}}$. Wir können also ein Skalarprodukt auf $\mathbb{C}^1$ durch $(z, w) \mapsto z\overline{w}$ definieren, und wollen dies verallgemeinern:

Definition 12.5.3. 1. Sei V ein komplexer Vektorraum. Eine Abbildung

$$h : V \times V \to \mathbb{C}$$

heißt *Sesquilinearform*, falls für alle $\alpha, \beta \in \mathbb{C}$ und $v, v', v'', w, w', w'' \in V$ gilt

$$h(\alpha v' + \beta v'', w) = \alpha h(v', w) + \beta h(v'', w)$$
$$h(v, \alpha w' + \beta w'') = \overline{\alpha} h(v, w') + \overline{\beta} h(v, w'') \ .$$

Man sagt, die Sesquilinearform ist im zweiten Argument antilinear oder semilinear.

2. Eine Sesquilinearform heißt *hermitesch*, falls gilt

$$h(v, w) = \overline{h(w, v)} \quad \text{für alle } \ v, w \in V \ .$$

Insbesondere ist $h(v, v) \in \mathbb{R}$.

3. Eine hermitesche Sesquilinearform heißt *positiv definit*, falls gilt

$$h(v, v) > 0 \quad \text{für alle } \ v \in V \setminus \{0\} \ .$$

4. Wir bezeichnen eine positive definite hermitesche Sesquilinearform auf einem $\mathbb{C}$-Vektorraum als *inneres Produkt* oder *(unitäres) Skalarprodukt* und das Paar (V, h) als *unitären Vektorraum.*

Wir haben also die Bedingung bilinear (zu sesquilinear) und symmetrisch (zu hermitesch) abgewandelt, um sicherzustellen, dass sich h wie der Absolutbetrag auf $\mathbb{C}$ verhält.

Warnung: In der Literatur findet sich auch die genau entgegengesetzte Konvention, dann ist h hermitesch genau wenn es linear im zweiten Argument und antilinear im ersten Argument ist.

Beispiel 12.5.4. Auf $\mathbb{C}^n$ definieren wir die hermitesche Form $h(z, w) = \sum_{i=1}^n z_i \overline{w_i}$.

Auch der Vektorraum $C^0([a, b], \mathbb{C})$ der stetigen komplexwertigen Funktionen auf dem abgeschlossenen Interval $[a, b]$ wird durch

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}dx$$

zum euklidischen Vektorraum.

Betrachtung 12.5.5. Für eine geordnete Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von V führen wir wie in Definition 12.2.4 die darstellende Matrix der Sesquilinearform ein:

$$M_{\mathcal{B}}(h) = (h(b_i, b_j))_{i,j=1\ldots n} \ .$$

Es gilt $h(\sum \alpha_i b_i, \sum \beta_j b_j) = \sum_{i,j} \alpha_i M_{\mathcal{B}}(h)_{ij}\overline{\beta_j}$.

Wenn wir nun einen Basiswechsel mit Basiswechselmatrix $T_{\mathcal{B}}^{\mathcal{B}'}$ vornehmen, dann ändert sich die darstellende Matrix M zu S^*MS, wobei wir zuerst S für $\overline{T_{\mathcal{B}}^{\mathcal{B}'}}$ schreiben und dann S^* als $\overline{S^T}$ definieren, also für die transponierte Matrix, in der wir jeden Eintrag komplex konjugieren.

Der Beweis von Satz 12.4.5 funktioniert auch für Sesquilinearformen und zeigt, dass jede hermitsche Sesquilinearform in einer geeigneten Basis von einer diagonalen Matrix dargestellt wird. (Zur Erinnerung: Wir wählen $b \in V$ mit $h(b, b) \neq 0$ und betrachten dann $V' = \ker(v \mapsto h(v, b)) \leq V$. Per Induktionsannahme gibt es eine Basis $\mathcal{B}$ von V', so dass $h|_{V'}$ diagonal dargestellt wir, und dann ist $M_{\mathcal{B}'\cup\{b\}}(h)$ diagonal.)

Betrachtung 12.5.6. Für einen euklidischen Vektorraum V liefert das Skalarprodukt nach Satz 12.2.9 einen Isomorphismus $\tau_V : V \cong V^*$, der v nach $v' \mapsto \langle v, v'\rangle$ schickt.

Für einen unitären Vektorraum V ist die Situation etwas subtiler: Die hermitesche Form h definiert eine Abbildung von V in den Raum der antilinearen Formen auf V, denn h ist ja im zweiten Argument nicht linear sondern antilinear.

Wir korrigieren das, indem wir für einen Vektorraum V den *konjugierten Vektorraum* $\overline{V}$ definieren. Er hat die gleichen Element und die gleiche Addition wie V, aber wir definieren $\lambda \cdot v = \overline{\lambda}v$ für $v \in \overline{V}$ (hier bezeichnet $\cdot$ die Skalamultiplikation in $\overline{V}$ und das ausgelassene Symbol die Skalarmultiplikation in V).

Es gilt $\dim_{\mathbb{C}} \overline{V}^* = \dim_{\mathbb{C}} V$. Wenn wir eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von V wählen erhalten wir einen Isomorphismus $\Theta_{\mathcal{B}} : V \cong \overline{V}$ durch $\sum \alpha_i b_i \mapsto \sum \overline{\alpha_i} b_i$. Dies ist tatsächlich eine lineare Abbildung, wir prüfuen $\Theta_{\mathcal{B}}(\lambda(\sum_i \alpha_i b_i)) = \overline{\lambda}\sum_i \overline{\alpha_i} b_i = \lambda \cdot \Theta_{\mathcal{B}}(\sum_i \alpha_i b_i)$.

Mit dieser Definition von $\overline{V}$ ist $h(v, -)$ eine lineare Abbildung von $\overline{V}$ nach $\mathbb{C}$, und damit induziert eine hermitesche Form h einen Isomorphismus $V \cong \overline{V}^*$, den wir auch mit τ_V bezeichnen.

Wenn Sie so wollen, ist eine Sesquilinearform auf V ein Element von $\mathrm{Bil}(V, \overline{V})$, und wir erhalten nach Lemma 12.2.5 eine Abbildung $\tau_V \in \mathrm{Hom}_{\mathbb{C}}(V, \overline{V}^*)$, die ein Isomorphismus ist, da die hermitesche Form nicht ausgeartet ist.

Definition 12.5.7. Wir sagen $(V, \langle -, - \rangle)$ ist ein *Vektorraum mit innerem Produkt* wenn entweder V ein reeller Vektorraum und $\langle -, - \rangle$ ein euklidisches Skalarpodukt ist, oder wenn V ein komplexer Vektorraum und $\langle -, - \rangle$ ein unitäres Skalarprodukt ist.

In der Folge schreiben wir (um Doppelungen zu vermeiden) $\overline{\lambda}$ für Skalare in $\mathbb{R}$ oder $\mathbb{C}$, wenn $\lambda \in \mathbb{R}$ dann ist dies einfach gleich λ. Genauso verstehen wir für einen reellen Vektorraum V den konjugierten Vektorraum $\overline{V}$ einfach als V.)

⟨⟨Wir werden zahlreiche Resultate für Vektorräume mit innerem Produkt beweisen, die über $\mathbb{R}$ und $\mathbb{C}$ gelten, wenn wir die Aussagen richtig interpretieren. Wenn Sie irgendwann den Faden verlieren, stellen Sie zuerst sicher, dass Sie die Aussage über $\mathbb{R}$ verstehen (indem Sie alle $^-$ ignorieren), und wenden sich dann dem komplexen Fall zu. ⟩⟩

Satz 12.5.8 (*Cauchy–Schwarz'sche Ungleichung*)**.** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit innerem Produkt. Dann gilt für alle $x, y \in V$*

$$|\langle x, y \rangle| \leq \|x\| \|y\| \, .$$

Gleichheit gilt genau dann, wenn die Familie (x, y) linear abhängig ist.

Dies ist eine der wichtigsten Ungleichungen der Mathematik. Der Beweis ist kurz und elementar (aber nicht unbedingt leicht zu finden)!

Beweis:. Sei $y \neq 0$ (sonst ist die Aussage trivialerweise wahr). Wir betrachten

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = ||x||^2 + \lambda\overline{\lambda}||y||^2 - \lambda\langle y, x \rangle - \overline{\lambda}\langle x, y \rangle$$

Wir setzen nun $\lambda = \frac{\langle x,y \rangle}{\langle y,y \rangle}$ oder äquivalent $\overline{\lambda} = \frac{\langle y,x \rangle}{\langle y,y \rangle}$ und erhalten

$$0 \leq ||x||^2 + \langle y, x \rangle \langle x, y \rangle ||y||^{-2} - \langle x, y \rangle \langle y, x \rangle ||y||^{-2} - \langle y, x \rangle \langle x, y \rangle ||y||^{-2}$$

oder

$$||x||^2 ||y||^2 \geq |\langle x, y \rangle|$$

durch Einsetzen der Definitionen. □

Satz 12.5.9. *Sei V ein Vektorraum mit innerem Produkt. Dann gilt für $\|x\| := \sqrt{\langle x, x \rangle}$*

1. $\|x\| \geq 0$ *für alle* $x \in V$
2. $\|x\| = 0 \Leftrightarrow x = 0$
3. $\|\alpha x\| = |\alpha| \|x\|$ *für alle* $\alpha \in K$ *und* $x \in V$.
4. $\|x + y\| \leq \|x\| + \|y\|$ *(*Dreiecksungleichung*)*

Beweis:. 1.,2. sind Teil der Definition.

3. $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha\overline{\alpha}\langle x, x \rangle} = |\alpha| \|x\|$.
4. Wegen der Cauchy–Schwarz'schen Ungleichung gilt

$$\|x + y\|^2 = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

□

Definition 12.5.10. 1. In einem Vektorraum mit innerem Produkt sagen wir x ist *orthogonal* zu y, oder x und y sind orthogonal, wenn $\langle x, y \rangle = 0$ gilt. Wir schreiben $x \perp y$.

2. Für zwei vom Nullvektor verschiedene Vektoren x, y eines euklidischen Vektorraums heißt $\alpha \in [0, \pi]$ mit $\cos\alpha = \frac{\langle x,y\rangle}{\|x\|\|y\|}$ der *Innenwinkel* $\angle(x, y)$ von x und y. Wir definieren keine Winkel in unitären Vektorräumen!

Beispiel 12.5.11. Der Vektorraum $V = C^\circ([0, \pi], \mathbb{R})$ der stetigen reellwertigen Funktionen auf dem abgeschlossenen Interval $[0, \pi]$ wird nach 12.5.2.2 durch das Skalarprodukt

$$\langle f, g\rangle = \int_0^\pi f(x)g(x)dx$$

zum euklidischen Vektorraum. Bezüglich dieses Skalarprodukts gilt

$$\langle \sin, \cos\rangle = \int_0^\pi \sin(x)\cos x dx = \left[\frac{1}{2}\sin^2 x\right]_0^\pi = 0 \ ,$$

also $\sin \perp \cos$.

Lemma 12.5.12. *Ein n-dimensionaler Vektorraum mit innerem Produkt $(V, \langle\cdot,\cdot\rangle)$ hat eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$, in der für $i, j = 1, \ldots, n$ gilt*

$$M_\mathcal{B}(\langle\cdot,\cdot\rangle) = E_n$$

d.h.

$$\langle b_i, b_j\rangle = \delta_{i,j}$$

oder, was äquivalent ist: $b_i \perp b_j$ für $i \neq j$ und $\|b_i\| = 1$.

Beweis:. Im euklidischen Fall folgt dies aus dem Sylvesterschen Trägheitssatz 12.4.9 (und der Annahme, dass $\langle -, -\rangle$ positiv definit ist). Im unitären Fall benutzen wir Betrachtung 12.5.5 um eine Basis $(b'_1, \ldots, b'_n)$ aus orthogonalen Vektoren zu finden und setzen dann $b_i = \frac{1}{||b'_i||}b'_i$. □

Definition 12.5.13. Eine solche Basis eines endlich-dimensionalen Vektorraums mit innerem Produkt heißt *Orthonormalbasis*. Allgemeiner heißt eine Familie $(v_1, \ldots, v_k)$ von Vektoren $v_i \in V$ in einem Vektorraum mit innerem Produkt ein *Orthonormalsystem*, wenn $\langle v_i, v_j\rangle = \delta_{ij}$ gilt.

Lemma 12.5.14. *Jedes Orthonormalsystem in einem Vektorraum mit innerem Produkt ist linear unabhängig.*

Beweis:. Sei $(v_1, \ldots, v_k)$ ein Orthonormalsystem und gelte

$$\sum_{i=1}^k \alpha_i v_i = 0 \qquad \text{mit} \quad \alpha_i \in K \ .$$

Dann gilt

$$0 = \langle 0, v_i\rangle = \sum_{j=1}^k \alpha_j\langle v_j, v_i\rangle = \alpha_i \quad \text{für alle} \quad i = 1, \ldots, k \ . \qquad \square$$

Lemma 12.5.15. *Ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V, dann gilt für alle $v \in V$*

$$v = \sum_{i=1}^n \langle v, b_i\rangle b_i$$

Beweis:. Da $\mathcal{B}$ eine Basis von V ist, gibt es Koeffizienten $\alpha_i \in K$, so dass $v = \sum_{i=1}^{n} \alpha_i b_i$ gilt. Wir rechnen:

$$\langle v, b_j \rangle = \sum_{i=1}^{n} \alpha_i \langle b_i, b_j \rangle = \alpha_j \ .$$

□

Definition 12.5.16. Sei $X \subset V$ eine Teilmenge, so ist der *orthogonale Raum* zu X (bezüglich $\langle , \rangle$)

$$X^{\perp} := \{y \in V | \ \langle x, y \rangle = 0 \text{ für alle } x \in X\} \subset V$$

Es ist leicht zu sehen, dass der orthogonale Raum stets ein Untervektorraum ist. Dies gilt auch, wenn V unitär ist und $y \mapsto \langle x, y \rangle$ nicht linear ist! Denn mit $\langle x, y \rangle = 0$ ist auch $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle = 0$.

Bemerkung 12.5.17. Der orthogonale Raum lässt sich für beliebige Bilinearformen definieren, vgl. auch Lemma 12.3.7, dann müssen wir allerdings gegebenenfalls zwischen links- und rechtsorthogonal unterscheiden.

Lemma 12.5.18. *Sei V ein Vektorraum mit innerem Produkt und $U \leq V$. Dann ist $V = U \oplus U^{\perp}$.*

Beweis:. Sei $v \in U \cap U^{\perp}$. Dann ist $\langle v, v \rangle = 0$ und da $\langle , \rangle$ positiv definit ist gilt $v = 0$. Also ist $U \cap U^{\perp} = 0$.

Im euklidischen Fall gilt nach Lemma 12.3.7 $\dim_{\mathbb{R}} U + \dim_{\mathbb{R}} U^{\perp} = \dim_{\mathbb{R}} V$, da das innere Produkt eine nicht ausgeartete Bilinearform ist.

Im unitären Fall gilt der gleiche Beweis wie für Lemma 12.3.7, mit dem einzigen Unterschied, dass $h^{\#}$ für eine Sesquilinearform eine Abbildung von V nach $\overline{V}^*$ ist, vgl. Betrachtung 12.5.6. Nun haben wir eine Einschränkung $r : \overline{V}^* \to \overline{U}^*$. Es gilt $\dim_{\mathbb{C}} \overline{U}^* = \dim_{\mathbb{C}} U$ und damit ist wieder $\dim_{\mathbb{C}} U^{\perp} = \dim_{\mathbb{C}} V - \dim_{\mathbb{C}} U$ aus der Dimensionsformel. □

Definition 12.5.19. Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich–dimensionaler Vektorraum mit innerem Produk und U ein Untervektorraum von V. Dann heißt die lineare Abbildung

$$P_U : \ V \to V$$

mit $(P_U)_{|_U} = \mathrm{id}_U$ und $(P_U)_{|U^{\perp}} = 0$ die *orthogonale Projektion* auf den Untervektorraum U.

Betrachtung 12.5.20. Ist $\mathcal{B}_1 = (b_1, \ldots, b_r)$ eine Orthonormalbasis von U und $\mathcal{B}_2 = (b_{r+1}, \ldots, b_n)$ eine Orthonormalbasis von $U^{\perp}$, dann ist $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V.

Es ist nach Lemma 12.5.15 für $v \in V$

$$v = \sum_{j=1}^{n} \langle v, b_j \rangle b_j = \underbrace{\sum_{j=1}^{r} \langle v, b_j \rangle b_j}_{\in U} + \underbrace{\sum_{j=r+1}^{n} \langle v, b_j \rangle b_j}_{\in U^{\perp}} \ .$$

Damit erhalten wir die folgenden Formeln für die Projektionen:

$$P_U(v) = \sum_{j=1}^{r} \langle v, b_j \rangle b_j \quad \text{und} \quad P_{U^{\perp}}(v) = \sum_{j=r+1}^{n} \langle v, b_j \rangle b_j = (\mathrm{id}_V - P_U)(v) \ .$$

Es gilt

$$(P_U)^2 = P_U \quad \text{und} \quad (P_{U^{\perp}})^2 = P_{U^{\perp}} \ .$$

Wir sagen, P_U und $P_{U^{\perp}}$ sind *idempotent*. Ferner ist $\mathrm{id}_V = P_U + P_{U^{\perp}}$.

Betrachtung 12.5.21 (Gram-Schmidt'sches Orthonormalisierungsverfahren)**.** Wir wollen aus einer beliebigen Basis $(v_1, \ldots, v_n)$ eines endlich–dimensionalen Vektorraums mit innerem PRodukt eine Orthonormalbasis gewinnen.

1.Schritt: Setze $b_1 := \frac{v_1}{\|v_1\|}$. Dies ist wegen $v_1 \neq 0$ definiert.

$(k+1)$–ter Schritt: sei $b_1, \ldots, b_k$ ein Orthonormalsystem mit $\mathrm{span}_K(b_1, \ldots, b_k) = \mathrm{span}_K(v_1, \ldots, v_k)$. Dann ist

$$\widetilde{b}_{k+1} := P_{\mathrm{span}(b_1 \ldots b_k)^\perp}(v_{k+1}) = v_{k+1} - P_{\mathrm{span}(b_1 \ldots b_k)} v_{k+1} = v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, b_i \rangle b_i$$

Dies ist ungleich 0, da sonst v_{k+1} Linearkombination von $(b_1 \ldots b_k)$ und somit auch von $(v_1 \ldots v_k)$ wäre. Setze $b_{k+1} := \frac{\widetilde{b}_{k+1}}{\|\widetilde{b}_{k+1}\|}$ Dies ist ein normierter Vektor. Außerdem gilt für $1 \leq i \leq k$

$$\langle b_i, \tilde{b}_{k+1} \rangle = \langle b_i, P_{\mathrm{span}(b_1 \ldots b_k)^\perp}(v_{k+1}) \rangle = 0$$

Beispiel 12.5.22. Sei $\mathbb{R}^3$ mit dem Standardskalarprodukt gegeben. Wir betrachten

$$\mathcal{B} = (v_1, v_2, v_3) = (\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix})$$

und folgen dem Schmidtschen Algorithmus: $b_1 = \frac{1}{\sqrt{3}} v_1$. Als nächstes ist $\widetilde{b_2} = v_2 - P_{\mathrm{span}(b_1, v_2)} = v_2 - \langle b_1, v_2 \rangle b_1 = \frac{1}{3}(1,1,-2)^T$ und $b_2 = \frac{1}{\sqrt{6}}(1,1,-2)^T$.

Schließlich ist $\widetilde{b_3} = v_3 - P_{\mathrm{span}(b_1, b_2)} v_3 = (-\frac{1}{2}, \frac{1}{2}, 0)^T$ und $b_3 = \frac{1}{\sqrt{2}}(1,-1,0)^T$.

Wenn wir den Algorithmus umgekehrt mit v_3 beginen erhalten wir $b_3 = v_3 = e_1$, $b_2 = v_2 - e_1 = e_2$ und $b_1 = v_1 - e_1 - e_2 = e_3$. Dies ist die Standardbasis!

12.6 Adjungierte Abbildungen

Als nächsten untersuchen wir Abbildungen zwischen Vektorräumen mit innerem Produkt. Bevor wir uns den naheliegenden Abbildungen zuwenden, die das innere Produkt bewahren, betrachten wir zuerst ein etwas subtileres Konzept:

Lemma 12.6.1. *Es seien V, W endlich–dimensionale K–Vektorräume mit innerem Produkt. Für jede lineare Abbildung $f : V \to W$ gibt es eine eindeutige lineare Abbildung $f^\dagger : W \to V$, so dass $\langle w, fv \rangle = \langle f^\dagger w, v \rangle$ für alle $v \in V$, $w \in W$.*

Definition 12.6.2. Wir nennen $f^\dagger$ wie im Lemma die *Adjungierte* Abbildung zu f

Beweis:. Die Zuordnung $\beta : (w, v) \mapsto \langle w, fv \rangle$ definiert ein Element in $\mathrm{Bil}(W, \overline{V})$. Dank Lemma 12.2.5.1 ist $\mathrm{Bil}(W, \overline{V}) \in \mathrm{Hom}(W, \overline{V}^*)$ und da V ein inneres Produkt hat ist $\overline{V}^* \cong V$ (siehe Betrachtung 12.5.6).

Zusammen erhalten wir einen Isomorphismus $\mathrm{Hom}(W, V) \cong \mathrm{Bil}(W, \overline{V})$, der g nach $(w, v) \mapsto \langle gw, v \rangle$ abbildet. Das Urbild von $\langle -, - \rangle$ ist dann die gesuchte Abbildung $f^\dagger$. □

Man beachte, dass die adjungierte Abbildungen $f^\dagger$ nicht nur von f sondern auch von den inneren Produkten auf V, W abhängt!

Beispiel 12.6.3. Sei K^n mit dem inneren Produkt $\langle v, w \rangle = v^T w$ ausgestattet. Dann gilt $\langle v, Aw \rangle = v^T A w = (A^T v)^T w = \langle A^T v, w \rangle$ und die Adjungierte zur linearen Abbildung, die A darstellt, wird von A^T dargestellt.

Bemerkungen 12.6.4. 1. Es gilt in der Situation wie im Lemma für eine lineare Abbildung $g : W \to V$, dass $\langle gw, v\rangle = \overline{\langle v, gw\rangle} = \overline{\langle g^\dagger(v), w\rangle} = \langle w, g^\dagger(v)\rangle$.

Wir müssen also nicht zwischen Links- und Rechtsadjungierten unterscheiden. (Wenn man Adjungierte für allgemeinere Bilinearformen definiert, ist dies notwendig!)

2. Es gilt auch
$$(f^\dagger)^\dagger = f$$
denn für beliebige $v \in V$, $w \in W$ gilt $\langle fv, w\rangle = \langle v, f^\dagger w\rangle = \langle (f^\dagger)^\dagger(v), w\rangle$. Da das innere Produkt nicht ausgeartet ist, muss $fv = (f^\dagger)^\dagger(v)$ gelten.

3. Es gilt $(g \circ f)^\dagger = f^\dagger \circ g^\dagger$ für $f \in \mathrm{Hom}_K(V, W)$, $g \in \mathrm{Hom}_K(W, U)$. Dies folgt aus der Betrachtung von $\langle (gf)^\dagger(v), w\rangle = \langle v, gf(w)\rangle = \langle g^\dagger(v), f(w)\rangle = \langle f^\dagger g^\dagger(v), w)\rangle$ für beliebige $v \in V$, $w \in W$.

4. Es gilt $(f+g)^\dagger = f^\dagger + g^\dagger$ und $(\lambda f)^\dagger = \overline{\lambda} f^*$ für $f, g \in \mathrm{Hom}_K(V, W)$ und $\lambda \in K$. Die zweite Aussage folgt aus $\langle \lambda f(v), w\rangle = \lambda\langle f(v), w\rangle = \lambda\langle v, f^\dagger(w)\rangle = \langle v, \overline{\lambda} f^\dagger(w)\rangle$.

Die Abbildung $f \to f^\dagger$ ist also antilinear.

Beispiel 12.6.3 verallgemeinert sich, allerdings brauchen wir dazu Orthonormalbasen.

Lemma 12.6.5. *Seien V, W endlich–dimensionale Vektorräume mit innerem Produkt und geordneten Orthonormalbasen $\mathcal{A}, \mathcal{B}$. Sei $f : V \to W$ eine lineare Abbildung. Dann gilt*
$$M_{\mathcal{A}}^{\mathcal{B}}\left(f^\dagger\right) = \overline{M_{\mathcal{B}}^{\mathcal{A}}(f)}^T$$

Beweis:. Wir berechnen mit den Orthonormalbasen $\mathcal{A} = (a_1, \ldots, a_n)$ und $\mathcal{B} = (a_1, \ldots, a_n)$
$$\langle f(a_i), b_j\rangle = \sum_k M_{\mathcal{A}}^{\mathcal{B}}(f)_{ki}\langle b_k, b_j\rangle = M_{\mathcal{A}}^{\mathcal{B}}(f)_{ji}$$
und
$$\langle a_i, f^\dagger(b_j)\rangle = \sum_k \overline{M_{\mathcal{B}}^{\mathcal{A}}(f^\dagger)_{kj}}\langle a_i, a_k\rangle = \overline{M_{\mathcal{B}}^{\mathcal{A}}(f^\dagger)_{ij}}$$
und vergleichen die beiden Ausdrücke. □

Definition 12.6.6. Für $A \in M(n \times n, \mathbb{C})$ nennt man $A^* := \bar{A}^T$ die *Adjungierte* von A. (Auch die Schreibweise $A^\dagger$ ist gebräuchlich.)

Diese Notation ist potenziell verwirrend: Für eine Matrix A bezeichnet A^* die konjugierte Transponierte, für eine lineare Abbildung f ist f^* die duale Abbildung. Aber die beiden Notationen sind kompatibel.

Wir betrachten hierzu, dass f nicht nur $f^* : W^* \to V^*$ induziert, sondern auch eine Abbildung zwischen den konjugierten Vektorräumen $\overline{W^*}$ und $\overline{V^*}$. Wir berechnen $\lambda.f^*(\phi) = \overline{\lambda} f^*(\phi) = \overline{\lambda}\phi \circ f = f^*(\lambda.\phi)$ wobei wir $\lambda.\phi$ für die konugierte Skalarmultiplikation schreiben. Aufgrund der konjugierten Skalarmultiplikation müssen wir die Einträge der Matrix, die f darstellt, konjugieren, es gilt also $M_{\mathcal{A}^*}^{\mathcal{B}^*}(f^*) = \overline{M_{\mathcal{B}}^{\mathcal{A}}(f)}^T$, was nach der Formel aus Lemma 12.6.5 genau $M_{\mathcal{A}}^{\mathcal{B}}(f^\dagger)$ ist!

Wir machen erst eine einleitende Betrachtung:

Betrachtung 12.6.7. Sei (b_i) eine Orthonormalbasis. Der Vergleich der Identitäten $\delta_{ij} = b_i^*(b_j)$ und $\tau_V(b_i)(b_j) = \langle b_i, b_j\rangle = \delta_{i,j}$ für alle $i, j = 1, \ldots, n$ zeigt
$$\tau_V(b_i) = b_i^* .$$

Satz 12.6.8. *Seien V, W endlich–dimensionale euklidische Vektorräume und τ_V, τ_W die zugehörigen Isomorphismen nach $\overline{V}^*$ und $\overline{W}^*$. Sei $f : V \to W$ eine lineare Abbildung. Dann hängt die adjungierte Abbildung $f^\dagger : W \to V$ mit der dualen Abbildung $f^* : \overline{W^*} \to \overline{V^*}$ folgendermaßen zusammen:*

$$f^\dagger = \tau_V^{-1} \circ f^* \circ \tau_W \ ,$$

d.h. das Diagramm

$$\begin{array}{ccc} V & \xleftarrow{f^\dagger} & W \\ {\scriptstyle \tau_V}\downarrow & & \downarrow{\scriptstyle \tau_W} \\ \overline{V^*} & \xleftarrow{f^*} & \overline{W^*} \end{array}$$

kommutiert.

Beweis:. Es gilt für alle $w \in W$ und $v \in V$

$$\tau_V f^\dagger(w)(v) = \langle f^\dagger(w), v\rangle = \langle w, f(v)\rangle = \tau_W(w)(fv) = f^*\tau_W(w)(v)$$

wobei wir nur die Definitionen verwendet haben. Also sind $\tau_V f^\dagger$ und $f^*\tau_W$ gleich, denn sie schicken jedes $w \in W$ zu Linearformen, die auf jedem $v \in V$ den gleichen Wert annehmen. □

Entsprechend wird die adjungierte Abbildung in der Literatur oft mit f^* bezeichnet, was wir aus Gründen der Klarheit vermieden haben. Wir haben daher die Notation $f^\dagger$ von den Physikern geliehen. (Manchmal sieht man auch das aunschauliche f^{ad}.)

Unter dem Isomorphismus τ_V entsprechen sich nicht nur duale Abbildungen und adjungierte Abbildungen, sondern auch Annulatoren und orthogonale Komplemente:

Satz 12.6.9. *Seien V, W endlich–dimensionale euklidische Vektorräume und sei $f : V \to W$ eine lineare Abbildung.*

1. *Für jede Teilmenge $U \subset V$ gilt $\tau_V(U^\perp) = U^\circ$.*
2. *Es gilt*
$$\mathrm{Im}\ f^\dagger = (\ker f)^\perp \qquad \textit{und} \qquad \ker f^\dagger = (\mathrm{Im}\ f)^\perp \ .$$

Beweis:. 1. In der Tat liegt φ genau dann in U°, wenn für alle $u \in U$ gilt $0 = \varphi(u) = \langle \tau_V^{-1}(\varphi), u\rangle$. Dies ist aber äquivalent zu $\tau_V^{-1}(\varphi) \in U^\perp$.

2. Der zweite Teil ist einfach die Übersetzung von Satz 12.1.15.

   Wir rechnen

$$\begin{aligned} \mathrm{Im}\ f^\dagger &= \mathrm{Im}\ \left(\tau_V^{-1} \circ f^* \circ \tau_W\right) \\ &= \tau_V^{-1}\mathrm{Im}\ (f^* \circ \tau_W) \\ &= \tau_V^{-1}\mathrm{Im}\ f^* && \text{da } \tau_W \text{ ein Isomorphismus ist.} \\ &= \tau_V^{-1}\,(\ker f)^\circ && \text{wegen Satz 12.1.15} \\ &= (\ker f)^\perp && \text{wegen Betrachtung Teil 1.} \end{aligned}$$

   Die Aussage über $\ker f^\dagger$ folgt analog.

□

12.7 Selbstadjungierte und normale Endomorphismen

Definition 12.7.1. Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder ein unitärer Vektorraum. Ein Endomorphismus $f \in \mathrm{End}(V)$ heißt *selbstadjungiert*, falls gilt $f = f^\dagger$, in anderen Worten

$$\langle f(v), w\rangle = \langle v, f(w)\rangle \quad \text{für alle } \ v, w \in V \ .$$

Aus Lemma 12.6.5 folgt sofort, dass $f \in \mathrm{End}(V)$ genau dann selbstadjungiert ist, wenn für eine Orthonormalbasis $\mathcal{B}$ gilt

$$M_{\mathcal{B}}(f) = M_{\mathcal{B}}(f)^* = \overline{M_{\mathcal{B}}(f)}^T \ .$$

Definition 12.7.2. Eine Matrix $A \in M(n\times n, \mathbb{C})$ heißt *hermitesch*, falls $A = A^*$ gilt (äquivalent falls $A^T = \overline{A}$ gilt).

Beispiele für hermitsche Matrizen sind

$$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1-i \\ 1+i & 1 \end{pmatrix}$$

sowie jede symmetrische Matrix.

Eine linear Abbildung ist also genau dann selbstadjungiert, wenn ihre darstellende Matrix in einer Orthonormalbasis hermitesch ist. Über den reellen Zahlen heißt das einfach, dass die Matrix symmetrisch ist.

Unser nächstes Ziel ist nun, die Eigenräume von selbstadjungierten Endomorphismen unitärer (und euklidischer) Vektorräume zu verstehen.

Lemma 12.7.3. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum und $f \in \mathrm{End}(V)$ selbstadjungiert. Dann sind alle Eigenwerte von f reell.*

Beweis:. Sei $v \in \mathrm{Eig}(f, \lambda)$ und $v \neq 0$. Dann gilt wegen der Linearität im ersten Argument:

$$\langle f(v), v\rangle = \langle \lambda v, v\rangle = \lambda \|v\|^2 \ .$$

Da f selbstadjungiert ist, ist dies wegen der Antilinearität im zweiten Argument gleich

$$\langle v, f(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda} \|v\|^2 \ .$$

Also gilt $\bar{\lambda} = \lambda$. □

Der nächste Satz gilt nicht nur für selbstadjungierte Endomorphismen, sondern für alle Endomorphismen, die mit ihrer Adungierten kommutieren.

Definition 12.7.4. Ein Endomorphismus $f \in \mathrm{End}_K(V)$ heißt *normal* wenn $f \circ f^\dagger = f^\dagger \circ f$.

Beispiel 12.7.5.

1. Wenn f seblstadjungiert ist, $f = f^\dagger$, ist f normal.
2. Wenn $f = -f^\dagger$ gilt, dann ist f normal.
3. Wenn $f^\dagger = f^{-1}$, dann ist f normal, denn Linksinverse zwischen endlichdimensionalen Vektorräumen sind auch rechtsinvers. Wir werden uns im nächsten Abschnitt mit diesem Beispiel beschäftigen.

Satz 12.7.6 (Spektralsatz)**.** *Sei $(V, \langle\cdot,\cdot\rangle)$ ein Vektorraum mit innerem Produkt und sei $f \in \mathrm{End}(V)$ normal. Sind $\lambda \neq \mu$ zwei verschiedene Eigenwerte von f, so sind die Eigenräume orthogonal.*

$$\mathrm{Eig}(f,\lambda) \perp \mathrm{Eig}(f,\mu)$$

Wenn überdies V endlich-dimensinal ist und das charkteritsche Polynom P_f in Linearfaktoren zerfällt, dann hat f eine Orthonormalbasis aus Eigenvektoren.

Umgekehrt ist jeder Endomorphismus, der eine Orthonormalbasis aus Eigenvektoren hat, normal.

Beweis:. Wir zeigen zuerst, dass $\mathrm{Eig}(f,\lambda) = \mathrm{Eig}(f^\dagger, \overline{\lambda})$ ist, wenn f normal ist.

Wir bestimmen $(f - \lambda\mathrm{id}_V)^\dagger = f^\dagger - \overline{\lambda}\mathrm{id}_V$ nach Bemerkung 12.6.4.4.

Dann ist auch $f - \lambda\mathrm{id}_V$ normal, denn

$$\begin{aligned}(f-\lambda\mathrm{id}_V)\circ(f^\dagger - \overline{\lambda}\mathrm{id}_V) &= ff^\dagger - \lambda f^\dagger - \overline{\lambda}f + \lambda\overline{\lambda}\mathrm{id}_V\\ &= f^\dagger f - \lambda f^\dagger - \overline{\lambda}f + \lambda\overline{\lambda}\mathrm{id}_V\\ &= (f^\dagger - \overline{\lambda}\mathrm{id}_V)\circ(f-\lambda\mathrm{id}_V)\end{aligned}$$

Es gilt $\mathrm{Eig}(f,\lambda) = \mathrm{Eig}(f-\lambda\mathrm{id}_V, 0)$ und $\mathrm{Eig}(f^\dagger,\overline{\lambda}) = \mathrm{Eig}(f^\dagger - \overline{\lambda}\mathrm{id}_V, 0)$, das heißt es reicht die 0-Eigenräume von einem normalen Endomorphismus und seiner Adjungiertem zu vergleichen. Sei also $g = f - \lambda\mathrm{id}_V$. Dann ist $v \in \mathrm{Eig}(g,0)$ genau wenn $||g(v)|| = 0$, was genau dann gilt, wenn $||g^\dagger(v)|| = 0$, denn für normale Abbildungen gilt für alle $v \in V$

$$\langle g(v), g(v)\rangle = \langle v, g^\dagger g(v)\rangle = \langle v, gg^\dagger(v)\rangle = \langle g^\dagger(v), g^\dagger(v)\rangle$$

Sei nun $v \in \mathrm{Eig}(f,\lambda)$, $w \in \mathrm{Eig}(f,\mu)$ mit $v, w \neq 0$. Dann gilt

$$\lambda\langle v,w\rangle = \langle \lambda v, w\rangle = \langle f(v), w\rangle = \langle v, f^\dagger(w)\rangle = \langle v, \overline{\mu}w\rangle = \mu\langle v,w\rangle\ ,$$

und somit $(\lambda-\mu)\langle v,w\rangle = 0$, also $v \perp w$ wenn $\lambda \neq \mu$. Dies zeigt die erste Aussage.

Zerfalle nun P_f in Linearfaktoren und sei $W \leq V$ die direkte Summe der Eigenräume. Wir wollen zeigen, dass $W = V$ ist. Wir betrachten dazu $V = W \oplus W^\perp$ und wählen $w \in W^\perp$. Wir bemerken, dass W per Konstruktion f-invariant ist, aber wegen $\mathrm{Eig}(f,\lambda) = \mathrm{Eig}(f^\dagger,\overline{\lambda})$ auch $f^\dagger$-invariant.

Aber damit ist $W^\perp$ auch f-invariant, denn für beliebiges $w \in W$ und $u \in W^\perp$ ist $\langle w, fu\rangle = \langle f^\dagger w, u\rangle = 0$. Damit gilt $P_f = P_{f|_W} \cdot P_{f|_{W^\perp}}$. Wenn $W^\perp \neq 0$, dann hat $P_{f|_{W^\perp}}$ Grad größer 0 und wir finden einen Eigenwert μ für $f|_{W^\perp}$. Aber dann trifft $\mathrm{Eig}(f,\mu)$ den $W^\perp$, im Widerspruch zur Definition von $W^\perp$.

Damit ist V direkte Summe der Eigenräume. Wir wählen eine Orthonormalbases für jeden Eigenraum, zum Beispiel durch das Gram-Schmidt'sche Verfahren aus Betrachtung 12.5.21, und die Vereinigung all dieser Basen ist eine Orthonormalbasis für V.

Für den Umkehrschluss sei (v_i) eine Orthonormalbasis aus Eigenvektoren mit Eigenwerten λ_i. Wir berechnen für jeden Eigenvektor

$$\begin{aligned}\langle f^\dagger f v_i, v_i\rangle &= \langle fv_i, fv_i\rangle = \lambda_i\overline{\lambda_i}\langle v_i, v_i\rangle\\ &= \langle\overline{\lambda_i}v_i, \overline{\lambda_i}v_i\rangle = \langle f^\dagger v_i, f^\dagger v_i\rangle\\ &= \langle ff^\dagger v_i, v_i\rangle\end{aligned}$$

wobei wir $\mathrm{Eig}(f,\lambda) = \mathrm{Eig}(f^\dagger,\overline{\lambda})$ benuzt haben, sowie

$$\begin{aligned}\langle f^\dagger f v_i, v_j\rangle &= \langle fv_i, fv_j\rangle = \lambda_i\overline{\lambda_j}\langle v_i, v_j\rangle\\ &= 0\\ &= \overline{\lambda_i}\lambda_j\langle v_i, v_j\rangle = \langle\overline{\lambda_i}v_i, \overline{\lambda_j}v_i\rangle = \langle f^\dagger v_i, f^\dagger v_j\rangle = \langle ff^\dagger v_i v_j\rangle\end{aligned}$$

Also gilt

$$\begin{aligned} ff^\dagger v_i &= \sum_j \langle ff^\dagger v_i, v_j\rangle v_j \\ &= \sum_j \langle f^\dagger f v_i, v_j\rangle v_j \\ &= f^\dagger f v_i \end{aligned}$$

und $ff^\dagger$ und $f^\dagger f$ stimmen auf der Basis (v_i) überein und sind die gleiche Abbildung. □

Bemerkung 12.7.7. Ist $f : V \to V$ ein selbstadjungierter Endomorphismus, so gilt

$$V = \ker f \oplus \text{Im } f$$

und die Untervektorräume sind orthogonal.

Wir können direkt mit orthogonaler Zerlegungen und Satz 12.6.9 rechnen:

$$V = \ker F \oplus (\ker F)^\perp = \ker F \oplus \text{Im } F^\dagger \overset{F \text{ s.a.}}{=} \ker F \oplus \text{Im } F \ .$$

Dies lässt sich als Spezialfall des nächsten Korollars verstehen: $\ker f$ ist der Eigenraum von 0 und $\Im f$ ist das Bild aller anderen Eigenvektoren, wenn f eine Basis aus Eigenvektoren hat.

Korollar 12.7.8. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlich–dimensionaler Vektorraum mit innerem Produkt und sei $f \in \text{End}(V)$ selbstadjungiert. Dann besitzt V eine Orthonormalbasis, die aus Eigenvektoren von f besteht. Insbesondere ist f diagonalisierbar und es gilt die orthogonale Zerlegung*

$$V = \text{Eig}(f, \lambda_1) \oplus \ldots \oplus \text{Eig}(f, \lambda_k) \quad \textit{mit } \lambda_k \in \mathbb{R} \ .$$

Beweis:. Sei zunächst $(V, \langle\cdot,\cdot\rangle)$ unitär. Nach dem Fundamentalsatz der Algebra 9.3.19 zerfällt das charakteristische Polynom P_f als komplexes Polynom vollständig in Linearfaktoren und Satz 12.7.6 angewendet auf den normalen Endomorphismus f beweist das Korollar.

Sei nun $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum. Es genügt zu zeigen, dass das charakteristische Polynom P_f schon über $\mathbb{R}$ in Linearfaktoren zerfällt. Dann können wir wie im komplexen Fall weiter schließen. Sei dazu $\mathcal{B}$ irgendeine Orthonormalbasis von V. Sei $A = M_{\mathcal{B}}(f) \in M(n\times n, \mathbb{R})$. Da A selbstadjungiert und $\mathcal{B}$ eine Orthonormalbasis ist, ist A symmetrisch, $A^T = A$.

Wir fassen nun A als komplexe Matrix auf, also als Element in $M(n \times n, \mathbb{C})$. Dann ist $A^* = A$, d.h. A ist hermitesch. Alle Eigenwerte sind nach Lemma 12.7.3 reell, und nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom P_A vollständig über $\mathbb{C}$,

$$P_A(X) = (X - \lambda_1) \ldots (X - \lambda_n)$$

mit $\lambda_i \in \mathbb{R}$, also zerfällt P_A schon über $\mathbb{R}$. □

Korollar 12.7.9. *Sei $A \in M(n \times n, K)$ hermitesch. Dann ist A diagonalisierbar über K. Insbesondere ist eine symmetrische reelle Matrix diagonalisierbar über $\mathbb{R}$.*

Beweis:. Wir können A als darstellende Matrix eines Endomorphismus $f : K^n \to K^n$ auffassen, $A = M_{\mathcal{E}}(f)$. Versieht man den K^n mit dem (euklidischen oder unitären) Standardskalarprodukt, so ist dieser Endomorphismus nach Bemerkung 12.7.3 selbstadjungiert, da die Standradbasis orthonormal ist. Es gibt daher nach Korollar 12.7.8 eine Basis $(v_1, \ldots, v_n)$ des $\mathbb{R}^n$, die bezüglich des euklidischen Standardskalarprodukts sogar eine Orthonormalbasis ist und die aus Eigenvektoren zu den Eigenwerten λ_i besteht. □

12.8 Isometrien

Definition 12.8.1. Es sei V ein K-Vektorraum mit innerem Produkt $\langle -, - \rangle$. Wir sagen, ein Endomorphismus $f \in \mathrm{End}_K(V)$ ist eine *(lineare) Isometrie* von V bezüglich $\langle -, - \rangle$, wenn für alle $x, y \in V$ gilt

$$\langle fx, fy \rangle = \langle x, y \rangle \quad .$$

Eine Isometrie eines euklidischen Vektorraums heißt *orthogonale Abbildungen*. Eine Isometrie eines unitären Vektorraums heißt *unitäre Abbildung*.

Satz 12.8.2. *Sei V ein n–dimensionaler K–Vektorraum mit innerem Produkt $\langle -, - \rangle$. Dann sind für $f \in \mathrm{End}(V)$ äquivalent:*

1. *f ist eine Isometrie bezüglich $\langle -, - \rangle$*
2. *Es gilt $f^\dagger \circ f = \mathrm{id}_V$.*
3. *f ist invertierbar und es gilt*
$$f^{-1} = f^\dagger \, .$$
4. *Für eine Orthonormalbasis $\mathcal{B}$ von V erfüllt $A = M_{\mathcal{B}}(f)$ die Bedingun $AA^* = E_n$.*

Beweis:.1.$\Leftrightarrow$2. Aus der Definition der Rechtsadjungierten $f^\dagger$ bezüglich β in Korollar 12.6.1 folgt, dass für alle $x, y \in V$ gilt

$$\langle x, f^\dagger fy \rangle = \langle fx, fy \rangle \, . \qquad (*)$$

Ist f eine Isometrie, so folgt $\langle x, f^\dagger fy \rangle = \langle x, y \rangle$. Weil $\langle -, - \rangle$ nicht-ausgeartet ist, folgt daraus $f^\dagger \circ f = \mathrm{id}_V$. Gilt umgekehrt 2. so folgt aus $(*)$, dass f eine Isometrie ist, also $\langle x, y \rangle = \langle f(x), f(y) \rangle$.

2.$\Leftrightarrow$3. s folgt aus $f^\dagger \circ f = \mathrm{id}_V$, dass f injektiv ist, und da $\dim V < \infty$ ist es damit invertierbar mit Inverser $f^\dagger$ ist. Der Umkehrschluss ist klar.

3.$\Leftrightarrow$ 4. Wir betrachten $M_{\mathcal{B}}(f^{-1})$ und $M_{\mathcal{B}}(f^\dagger)$ und wenden Lemma 12.6.5 an.

□

Definition 12.8.3. Eine Matrix $M \in M(n \times n, \mathbb{R})$ heißt *orthogonal*, wenn $MM^T = E_n$. Eine Matrix $M \in M(n \times n, \mathbb{C})$ heißt *unitär*, wenn $MM^* = E_n$.

Bemerkung 12.8.4. Die Basiswechselmatrix $T = T_{\mathcal{B}}^{\mathcal{A}}$ zwischen zwei Orthonormalbasen $\mathcal{A} = (a_1, \ldots, a_n)$, $\mathcal{B} = (b_1, \ldots, b_n)$ ist immer orthogonal (für $\mathbb{R}$-Vektorräume) bzw. unitär (für $\mathbb{C}$-Vektorräume). Denn es gilt mit $a_i = \sum_{j=1}^n T_{ji} b_j$

$$\begin{aligned} \delta_{ij} = \langle a_i, a_j \rangle &= \langle \sum_k T_{ki} b_k, \sum_\ell T_{\ell j} b_\ell \rangle \\ &= \sum_{k,\ell} T_{ki} \overline{T_{\ell j}} \langle b_k, b_\ell \rangle \\ &= \sum_k T_{ki} \overline{T_{kj}} = (T^* T)_{ij} \end{aligned}$$

Umgekehrt bilden die Spaltenvektoren einer orthogonalen bzw. unitären Matrix eine Orthonomalbasis des K^n. Wir rechnen

$$\delta_{ik} = (A^T A)_{ik} = \sum_{j=1}^{n} a_{ji} a_{jk} \ ,$$

Ähnlich folgt auch aus $AA^T = E_n$, dass

$$\delta_{ik} = (AA^T)_{ik} = \sum_{j=1}^{n} a_{ij} a_{kj} \ ,$$

also bilden auch die Zeilenvektoren eine Orthonormalbasis.

Lemma 12.8.5. *Seien $f, g \in \operatorname{End} V$ Isometrien. Dann gilt für alle $v, w \in V$*

1. *$\|f(v)\| = \|v\|$, d.h. f ist normerhaltend. Umgekehrt ist jede normerhaltende Abbildung eine Isometrie.*
2. *Es ist $v \perp w$ genau wenn $f(v) \perp f(w)$. Ist f orthogonal dann gilt für $v \neq 0$ und $w \neq 0$ auch $\angle(v, w) = \angle(f(v), f(w))$, d.h. f ist winkeltreu.*
3. *f ist ein Isomorphismus und die Umkehrabbildung f^{-1} ist ebenfalls eine Isometrie.*
4. *Die Verkettung $f \circ g$ von Isometrien ist eine Isometrie.*
5. *Ist $\lambda \in \mathbb{R}$ ein Eigenwert von f, so gilt $|\lambda| = 1$.*
6. *Es gilt $|\det(f)| = 1$.*

Beweis:. 1. Aus der Definition ist f normerhaltend. Umgekehrt folgt aus der Polarisierungsformel

$$\langle v, w \rangle = \frac{1}{2} \left(\|v + w\|^2 - \|v\|^2 - \|w\|^2 \right)$$

dass $\|f(v)\| = \|v\|$ für alle $v \in V$ sofort Orthogonalität von f impliziert. Für unitäre Abbildungen gilt die Polarisierungsgformel

$$\langle x, y \rangle = \frac{1}{4} ||x + y||^2 - ||x - y||^2 - i||x + iy||^2 + i||x - iy||^2$$

die Sie auf dem Übungsblatt überprüft haben. Damit ist f unitär, wenn es normerhaltend ist.

2. Dies folgt sofort aus der Definition.
3. Aus Satz **??** folgt, dass f ein Isomorphismus ist. Als nächstes ist $\langle f^{-1}v, f^{-1}w \rangle = \langle v, w \rangle$ indem wir die Isometrie f anwenden.
4. Direkt aus der Definition
5. Ist $v \neq 0$ Eigenvektor zum Eigenwert $\lambda \in \mathbb{R}$, so gilt

$$\|v\| = \|fv\| = \|\lambda v\| = |\lambda| \cdot \|v\| \ ,$$

also $|\lambda| = 1$.

6. Wir wählen orthonormale Basen und erhalten für die darstellende Matrix M dass $1 = \det(E_n) = \det(MM^*) = \det(M)\det(M^*)$. Aber es gilt auch $\det(M^*) = \overline{\det M^T} = \overline{\det M}$. Damit ist $|\det(M)| = \sqrt{\det(M)\overline{\det M}} = \sqrt{1} = 1$.

□

Definition 12.8.6. 1. Wir führen die *orthogonale Gruppe* und *spezielle orthogonale Gruppe*

$$O(n) := \{A \in M(n \times n, \mathbb{R}) |\ A^T A = E_n\}$$
$$SO(n) := \{A \in O(n) |\ \det A = 1\}$$

ein. $SO(n)$ ist Untergruppe von $O(n)$ und beide Gruppen sind Untergruppen der allgemeinen linearen Gruppe $GL(n, \mathbb{R})$.

2. Die Menge

$$U(n) := \{A \in M(n \times n, \mathbb{C}) \,|\, A^* A = E_n\}$$

ist eine Gruppe und heißt *unitäre Gruppe*

$$SU(n) := \{A \in M(n \times n, \mathbb{C}) \,|\, A^* A = E_n \text{ und } \det A = 1\}$$

heißt *spezielle unitäre Gruppe* von V.

Bemerkung 12.8.7. $SO(n)$ ist eine Untergruppe von $O(n)$ und $O(n)$ ist eine Untergruppe von $GL(n, \mathbb{R})$.

$SU(n)$ ist eine Untergruppe von $U(n)$ und $U(n)$ ist eine Untergruppe von $GL(n, \mathbb{C})$.

Wir können auch $O(n)$ als Untergruppe von $U(n)$ auffassen, nämlich als die unitären Matrizen mit reellen Einträgen, $O(n) = U(n) \cap GL(n, \mathbb{R}) \subset GL(n, \mathbb{C})$.

Auch für einen allgemeinen euklidischen oder unitären Vektorraum V können wir die Gruppe $O(V)$ und $SO(V)$ bzw. $U(V)$ und $SU(V)$ einführen. Sie sind isomorph zu $O(n), SO(n), U(n)$ und $SU(n)$.

Beispiele 12.8.8. 1. Eindimensionale euklidische Vektorräume, $n = 1$: eine Matrix $A = (a)$ ist genau dann orthogonal, wenn $(a^2) = A^T A = E_1 = (1)$ gilt, also für $a = \pm 1$. Es folgt

$$O(1) = \{\pm 1\} \cong \mathbb{Z}_2$$
$$SO(1) = \{+1\} \quad \text{ist die triviale Gruppe.}$$

2. Für $n = 2$ suchen wir Matrizen $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit reellen Einträgen und, wegen der Orthogonalität der Spaltenvektoren,

$$a^2 + c^2 = 1\ , \quad b^2 + d^2 = 1\ , \quad ab + cd = 0\ .$$

Wegen der ersten beiden Gleichungen finden wir Winkel α, α' mit

$$a = \cos\alpha\,, \quad c = \sin\alpha\,, \quad b = \sin\alpha' \quad \text{und} \quad d = \cos\alpha'\ .$$

Ferner gilt

$$0 = ab + cd = \cos\alpha \sin\alpha' + \sin\alpha \cos\alpha' = \sin(\alpha + \alpha')\ .$$

Also ist $\alpha' = -\alpha \bmod 2\pi$ oder $\alpha' = \pi - \alpha \bmod 2\pi$, wobei wir mit $\bmod 2\pi$ meinen, dass wir alle Winkel identifzieren, die sich um ein Vielfaches von 2π unterscheiden. Also besteht

$$SO(2) = \left\{ M(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \,|\, \theta \in \mathbb{R} \right\}$$

aus den Drehungen um den Ursprung und

$$O(2) = SO(2) \cup \left\{ M(S_\theta) = \begin{pmatrix} \cos\theta & +\sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$$

aus Drehungen um den Ursprung und Spiegelungen an Ursprungsgeraden (mit Winkel $\theta/2$), vgl. Beispiel 7.4.10.

3.
$$U(1) = \{A \in M(1 \times 1, \mathbb{C}) \mid A^*A = 1\} = \{(a), a \in \mathbb{C} \text{ mit } |a| = 1\}$$

ist vermöge $x + iy \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$ isomorph zu $SO(2)$. Im Allgemeinen sind unitäre und orthogonale Gruppen nicht isomorph.

Satz 12.8.9. *Sei $f \in U(V)$. Dann existiert eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht.*

Beweis:. Dies ist ein direktes Korollar von Satz 12.7.6. □

Insbesondere ist jeder unitäre Endomorphismus f eines endlich-dimensionalen unitären Vektorraums diagonalisierbar. Jede unitäre Matrix ist also ähnlich zu einer Diagonalmatrix $D = (a_1, \ldots, a_n)$, deren Einträge komplexe Zahlen vom Betrag 1 sind, $|a_i| = 1$, also $A = SDS^{-1}$. Da die diagonalisierende Basis eine Orthonormalbasis ist, ist S unitär, also gilt auch $A = SDS^*$.

Betrachtung 12.8.10. Wir betrachten nun noch einmal eine allgemeine symmetrische Bilinearform oder hermitesche Sesquilinearform β auf K^n. (Allgemein heißt hier, dass β nicht unbedingt positiv definit ist.) Wir setzen $A = M_{\mathcal{E}}(\beta)$. (Ähnliche Überlegungen gelten für allgemeine endlichdimensionale Vektorräume.)

Da A hermitesch ist, finden wir nun eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ aus orthonormalen Eigenvektoren von A.

Wir betrachten zuerst den reellen Fall.

Wir finden für die Einträge der darstellenden Matrix $M_{\mathcal{B}}(\beta)$

$$\beta(v_i, v_j) = v_j^T A v_i = \lambda_i v_j^T v_i = \lambda_i \delta_{ij}$$

Durch Reskalierung, also Übergang zur Basis

$$v_i' := v_i \text{ falls } \lambda_i = 0 \quad v_i' := \frac{1}{\sqrt{|\lambda_i|}} v_i \text{ falls } \lambda_i \neq 0$$

erhalten wir eine Basis, in der die darstellende Matrix diagonal ist mit Diagonalelementen in $\{0, \pm 1\}$, also bis auf Permutation von der Form im Sylvesterschen Trägheitssatz 12.4.9 ist.

Wir sehen also wie in Übungsaufgabe 11.3: n_+ ist gleich der Zahl der positiven Eigenwerte der symmetrischen Matrix A, n_- der negativen Eigenwerte von A, mit Vielfachheiten gezählt, und n_0 ist die Dimension des Kerns von A. Die quadratische Form ist genau dann nichtausgeartet, wenn $\det A \neq 0$ gilt. Für eine positiv definite Form ist $\det A > 0$ notwendig, aber nicht hinreichend, wie das Beispiel der Diagonalmatrix mit Einträgen $-1, -1$ zeigt.

Die Basiswechselmatrix $T = T_{\mathcal{B}}^{\mathcal{E}}$ von $\mathcal{E}$ nach $\mathcal{B}$ ist nach Bemerkung 12.8.4 eine Isomotrie (denn sowohl $\mathcal{B}$ als auch $\mathcal{E}$ sind Orthonormalbasen für das Standardskalarproduk), d.h. $TT^T = E_n$, oder, anders ausgedrückt $T^{-1} = T^*$.

Damit ist die Kongruenzumformung $A \mapsto T^T A T$ auch eine Ähnlichkeitsumformung, und umgekehrt!

Der orthonormale Basiswechsel ändert also die Matrix auf gleiche Weise, egal ob wir sie als Bilinearform oder als lineare Abbildung betrachten.

Dies gilt nicht mehr, wenn wir die Basisvektoren skalieren. Dann ändert sich die Matrix ihrer Kongruenzklasse, aber nicht in ihrer Ähnlichkeitsklasse.

Wir haben also einerseits gesehen, dass Matrizen sowohl lineare Abbildungen als auch Bilinearformen darstellen können, und diese beiden Interpretationen sehr unterschiedlich sind. Andererseits führt die Analyse im symmetrischen Fall zu den gleichen Rechnungen!

Falls $K = \mathbb{C}$ ist wird die Situation etwas komplizierter. Nun verändert ein Basiswechsel mit Matrix $T = T_{\mathcal{B}}^{\mathcal{E}}$ die darstellende Matrix durch $A \mapsto T^T A\overline{T}$, siehe Betrachtung 12.5.5. Wenn wir eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ suchen sodass $\beta(v_i, V_j) = v_j^T \overline{A} v_i = \lambda_i \delta_{ij}$ gilt, dann müssen die v_i nicht Basisvektoren sondern konjugierte Basisvektoren sein (die wir aber immer noch orthonormal wählen können).

Es gelten aber die analogen Resultate zum reellen Fall: Es gilt $T^*T = E_n$, die Basiswechselmatrix ist unitär und A und T^*AT stellen sowohl die gleiche lineare Abbildung als auch die gleiche hermitesche Form dar.

Sei nun $(V, \langle\cdot,\cdot\rangle)$ ein endlich–dimensionaler euklidischer Vektorraum.

Wir wollen nun noch die orthogonalen Matrizen für $n \geq 3$ untersuchen. Zentral wird das folgende Lemma sein:

Lemma 12.8.11. *Sei V ein endlich–dimensionaler euklidischer Vektorraum. Ist $f \in O(V)$ und ist $W \subset V$ ein f–invarianter Untervektorraum, dann ist das orthogonale Komplement $W^\perp$ ebenfalls ein f–invarianter Untervektorraum.*

Beweis:. Wegen Satz 12.8.2 ist f ein Automorphismus und genauso ist $f_{|W}$ ein Automorphismus des invarianten Untervektorraums W. Sei $v \in W^\perp$. Sei $w \in W$ beliebig; da uch $w' := f^{-1}(w)$ in W ist gilt

$$\langle f(v), w\rangle = \langle f(v), f(w')\rangle \overset{f \text{ orth.}}{=} \langle v, w'\rangle = 0 \ .$$

Also ist $f(v) \in W^\perp$. □

Lemma 12.8.12. *Ist $f \in O(V)$, so besitzt V einen f–invarianten Untervektorraum W der Dimension 1 oder 2.*

Beweis:.

- Betrachte die "Symmetrisierung" $\Psi := f + f^{-1} \in \mathrm{End}(V)$ und finde

$$\begin{aligned}\langle \Psi(v), w\rangle &\overset{\text{def}}{=} \langle f(v), w\rangle + \langle f^{-1}(v), w\rangle \\ &= \langle f^{-1}f(v), f^{-1}(w)\rangle + \langle f f^{-1}(v), f(w)\rangle \qquad [f \text{ und } f^{-1} \text{ sind orthogonal}] \\ &= \langle v, \Psi(w)\rangle \ .\end{aligned}$$

  Also ist Ψ selstadjungiert. Mit Satz 12.7.8 gibt es also eine Orthonormalbasis $\mathcal{B} = (b_1, \ldots, b_n)$ von Eigenvektoren:

$$\Psi(b_i) = \lambda_i b_i \quad \text{mit} \quad \lambda_i \in \mathbb{R} \ .$$

- Setze $W := \mathrm{span}_{\mathbb{R}}\{b_1, f(b_1)\}$. Dann ist $\dim_{\mathbb{R}} W \in \{1, 2\}$. Wir müssen zeigen, dass W f-invariant ist. Da $f(b_1) \in W$ per Definition, reicht es, $f^2(b_1) \in W$ zu zeigen. Dies folgt aus

$$f^2(b_1) = f\left(f(b_1) + f^{-1}(b_1) - f^{-1}(b_1)\right) = f\Psi(b_1) - b_1 = \lambda_1 f(b_1) - b_1 \in W \ . \quad \square$$

Satz 12.8.13. *Sei V ein endlich–dimensionaler euklidischer Vektorraum. Ist $f \in O(V)$, so gibt es eine Orthogonalbasis $\mathcal{B}$ von V, in der die darstellende Matrix die folgende Blockdiagonalgestalt hat:*

$$M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{A_1} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \boxed{A_x} \end{pmatrix}$$

wobei $A_j = M\left(R_{\theta_j}\right) \in M(2 \times 2, \mathbb{R})$ mit $\theta_j \in \mathbb{R}$ Drehmatrizen sind.

Beweis:. Wir verwenden vollständige Induktion nach $n := \dim_{\mathbb{R}} V$.

- Der Induktionsanfang für $n = 1$ und $n = 2$ ist durch Beispiel 12.8.8.1 und 2 klar; beachte, dass man wie in Beispiel 9.2.15 für eine Spiegelung S_θ an einer Ursprungsgeraden stets eine Basis finden kann mit
$$M_{\mathcal{B}}(S_\theta) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$
- Lemma 12.8.12 garantiert die Existenz eines f–invarianten Untervektorraums W von V der Dimension 1 oder 2. Wegen Lemma 12.8.11 ist auch das orthogonale Komplement $W^\perp$ ein f–invarianter Untervektorraum. Man setzt nun eine Orthonormalbasis $\mathcal{B}_1$ von W und $\mathcal{B}_2$ von $W^\perp$, in der die jeweiligen Einschränkungen von f die gewünsche Blockdiagonalgestalt haben, zu einer Orthonormalbasis $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ von V zusammen: man findet so die gewünschte Blockdiagonalgestalt:
$$M_{\mathcal{B}}(f) = \left(\begin{array}{c|c} \left(M_{\mathcal{B}_1}\left(f|_W\right)\right) & 0 \\ \hline 0 & M_{\mathcal{B}_2}\left(f|_{W^\perp}\right) \end{array}\right) .$$

□

Beispiel 12.8.14. Im Falle $\dim_{\mathbb{R}} V = 3$ treten nur die folgenden beiden Fälle auf:

- Drehungen um eine Ursprungsgerade. Hier hat man in einer geeigneten Orthonormalbasis von $\mathbb{R}^3$ die darstellende Matrix
$$M_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & M\left(R_\theta\right) \end{pmatrix} \quad \text{falls} \quad \det f = 1$$
Es heißt $\mathbb{R}b_1 = \mathrm{Eig}(f, 1)$ die *Drehachse* von f und θ der *Drehwinkel* von f.
Der Basiswechsel identifiziert hier die Drehachse mit dem Raum $\mathrm{span}_{\mathbb{R}}(e_1)$.
- Spiegelungen an Ebenen durch den Ursprung, gefolgt von einer Drehung mit Drehachse senkrecht zur Spiegelebene. Hier hat man in einer geeigneten Orthonormalbasis von $\mathbb{R}^3$ die darstellende Matrix
$$M_{\mathcal{B}}(f) = \begin{pmatrix} -1 & 0 \\ 0 & M\left(R_\theta\right) \end{pmatrix} \quad \text{falls} \quad \det f = -1 .$$

Diese Charakterisierung folgt sofort aus dem Satz zusammen mit einer Analyse der Diagonalmatrizen mit Einträgen ± 1. Zum Beispiel ist $\mathrm{diag}(+1,-1,-1) = \begin{pmatrix} \pm 1 & 0 \\ 0 & M(R_\pi) \end{pmatrix}$ eine Drehung um π um die x_1-Achse vor.

Lemma 12.8.15. *Sei V ein euklidischer Vektorraum und $v, w \in V$ mit $||w|| = ||v||$. Dann gibt es eine Isometrie f mit $f(v) = w$.*

Beweis:. Wir wollen eine Spiegelung durch die Hyperebene definieren, die orthogonal zu $v - w$ steht. Daraus ergibt sich mit etwas geometrischer Überlegung die Formel

$$f : z \mapsto z - 2\frac{z.(w-v)}{(||w-v||^2}(w-v)$$

(Wir subtrahieren von z das Doppelte der Komponente von z entlang der Richtung $v - w$.) Dieses f ist linear und wir berechnen

$$\begin{aligned}\langle f(z), f(z)\rangle &= \langle z - 2\frac{z.(w-v)}{||w-v||^2}(w-v), z - 2\frac{z.(w-v)}{||w-v||^2)}(w-v)\rangle \\ &= ||z||^2 - 4\frac{(z.(w-v))^2}{||w-v||^2} + 4\frac{(z.(w-v))^2}{||w-v||^2} \\ &= ||z||^2\end{aligned}$$

damit ist f eine Isometrie nach Lemma 12.8.5.1.

Schließlich benutzen wir $||v|| = ||w||$ um zu berechnen

$$\begin{aligned}f(v) &= v - 2\frac{v.(w-v)}{(||w-v||^2}(w-v) \\ &= v - 2\frac{v.w - ||v||^2}{||w||^2 + ||v||^2 - 2w.v}(w-v) \\ &= v - 2\frac{-1}{2}(w-v) \\ &= v\,.\end{aligned}$$

□

Bemerkenswerterweise ist jede Abbildung, die das Skalarprodukt bewahrt automatisch linear!

Satz 12.8.16. *Sei V ein endlich-dimensionaler euklidischer Vektorraum und $f : V \to V$ eine beliebige (nicht unbendingt lineare!) Abbildung, so dass $\langle f(v), f(w)\rangle = \langle v, w\rangle$ für alle $v, w \in V$. Dann ist f linear.*

Beweis:. Wir wählen eine orthonormale Basis $(e_1, \ldots, e_n)$ von V. Wir werden zuerst per Induktion nach n zeigen, dass es ein lineare Isomotrie T gibt, so dass $T \circ f = \mathrm{id}_V$.

Wir wählen eine lineare Isometrie T_1 mit $T_1(f(e_1)) = e_1$ nach Lemma 12.8.15. Dann ist $T_1 \circ f$ eine Isometrie, die e_1 fest lässt. Außerdem bewahrt $T_1 \circ f$ den Raum $V' = \mathrm{span}_K(e_1)^\perp = \mathrm{span}_K(e_1, \ldots, e_n)$. Per Induktionsannahme gibt es eine Isometrie T' von V', so dass $T' \circ (T_1 \circ f)|_{V'}$ alle e_i fest lässt. Wir definieren T durch $\begin{pmatrix} 1 & 0 \\ 0 & T' \end{pmatrix} \circ T_1$.

Gegeben T gilt nun dass $Tf(e_i) = e_i$ für alle Basisvektoren. Gegeben ein beliebiges $v \in V$ schreiben wir $v = \sum_{i=1}^n \langle v, e_i\rangle e_i$ und $Tf(v) = \sum_{i=1}^n \langle Tf(v), e_i\rangle e_i$. Aber $\langle Tf(v), e_i\rangle = \langle Tf(v), Tf(e_i)\rangle = \langle v, e_i\rangle$ denn Tf erhält das innere Produkt. Damit stimmen alle Koeffizienten überein und $Tf(v) = v$. Wir haben nirgendwo angenommen, dass f linear ist, aber nun können wir schließen, dass $f = T^{-1}$ linear ist. □

Der Beweis zugt auch, dass jede Isometrie Produkt von höchstens n Spiegelungen in Hyperebenen ist.

Beispiel 12.8.17. Aus geometrischer Hinsicht ist es interesant quadratische Gleichungen bis auf Isometrie zu klassifizieren. Es bietet sich hier an auch die nichtlinearen *Tranlationen* (oder Paralleverschiebungen) $t_a : v \mapsto v + a$ für $a \in \mathbb{R}^n$ zu betrachten. Sie erhalten zwar nicht die Normen von Vektoren (was die Abstände zum Nullpunkt sind), aber die Normen von Differenzen von Vektoren: $||v - w|| = ||t_a(v) - t_a(w)||$, und damit Abstände von Vektoren. (Eine Warnung: In der Literatur wird der Begriff Isometrie oft auch für solche Abbildungen verwendet! Für uns ist jede Isometrie linear.)

Die Gruppe der *euklidischen Transformationen* E_n wird von $O(n)$ zusammen mit den $\{t_a \mid a \in \mathbb{R}^n\} \cong (\mathbb{R}^n, +)$ erzeugt, ihre Elemente sind Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$ der Form $v \mapsto Av+a$ für $A \in O(n)$ und $a \in \mathbb{R}^n$.

Gegeben sei nun eine quadratische Gleichung $Q(x) = \sum a_{ij}x_jx_j + \sum_i b_ix_i + c$, die wir als Summe aus einer quadratischen Form, einer linearen Form und einer konstanten betrachten können.

Die Lösungsmenge $\{x \in \mathbb{R}^n \mid Q(x) = 0\}$ beschreibt interessante geometrische Obejkte, zum Beispiel Parabeln und Hyperbeln (für $n = 2$), Ellipsoide wie unser Erde oder Hyperboloide ($n = 3$) und ähnliche Gebilde in höheren Dimensionen.

Euklidische Transformationen rotieren und verschieben diese Lösungsmengen, ohne ihre Die Klassifizierungen aller quadratischen Gleichungen bis auf eukldische Transforamtionen ist also ein interessantes Problem, das die lineare Algebra uns erheblich erleichtert. geometrischen Eigenschaften zu verändern. Man kann zeigen, dass es für jedes Q ein ein Element $f \in E_n$ gibt, so dass $Q(fx)$ eine der folgenden Formen hat, wobei die a_i und k positive Konstanten sind:

1. $\sum_{i=1}^{p} a_ix_i^2 - \sum_{i=1}^{q} a_ix_i^2$ mit $p + q \leq n$.
2. $\sum_{i=1}^{p} a_ix_i^2 - \sum_{i=1}^{q} a_ix_i^2 \pm 1$ mit $p + q \leq n$.
3. $\sum_{i=1}^{r} a_ix_i^2 - 2X_{p+q+1}$ mit $p + q < n$.

Hierbei sind $(p, q, n-p-q)$ die Signature der quadratischen Form in q. Für graphischen Beispiele in $\mathbb{R}^3$ betrachten wir Abbildung 12.8.17.

Um diese Klassifizierung zu beweisen schreiben wir $q(x) = x^TAx + bx + c$ für eine $(n \times n)$-Matrix A, einen Zeilenvektor b und einen Skalar c. Dann wenden wir zuerst einen orthonormalen Basiswechsel T_1 an, um A zu diagonalisieren, $T_1^TAT_1 = \text{diag}(d_1, \ldots, d_n)$, mit $d_i = 0$ für $i > p+q$. Anschließend können wir mit einer Verschiebung V_1 den neuen Linearterm $b'x$ reduzieren: Für $i \leq p + q$ (also $d_i \neq 0$) ersetzen wir x_i durch $x_i - \frac{b'i}{d_i}$. Wir haben nun eine neue quadratische Form $Q'(x) = \sum_{i=1}^{p+q} d_ix_i^2 + \sum_{i=p+q+1}^{n} b'_i + c'$. Durch Skalieren der gesamten Gleichung können wir $c' = \pm 1$ annehmen, es ändert nichts an der Lösungsmenge von $Q(x) = 0$ (diesen Schritt hatte ich in der Vorlesung ausgelassen)

Dies deckt die beiden ersten Fälle ab, wenn $p + q = n$ oder alle $b'_i = 0$ sind. Andernfalls finden wir eine Rotation T_2 der Ebene $\text{span}_{\mathbb{R}}(e_{p+q+1}, \ldots, e_n)$ (die $e_1, \ldots, p_{p+q}$ fix lässt), so dass $b'T_2 = e_{p+q+1}^T$. Anschließend verschieben wir mit einer weiter Translationn V_2 die i-Koordinate um die Konstante c', so dass wir (nach einer weiteren Skalierung) die dritte Form erreichen. Dann ist $Q(V_2T_2V_1T_1(x))$ in der gewünschten Form.

12.9 Tensorprodukte

Wir wollen nun noch einmal verallgemeinern, und einen ersten Einblick in *multilineare Algebra* geben.

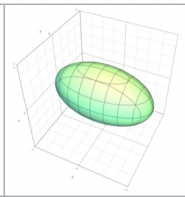
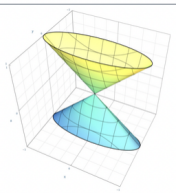
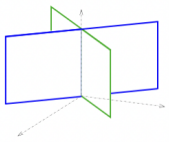
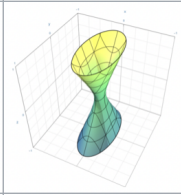
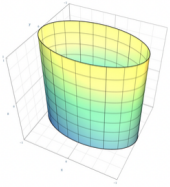
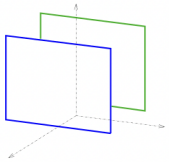
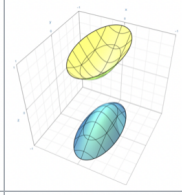
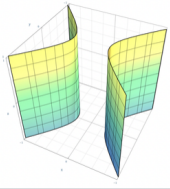
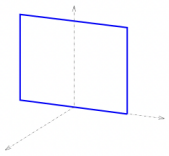
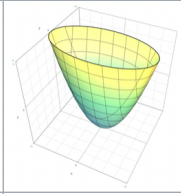
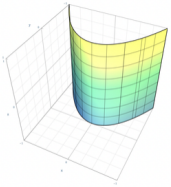
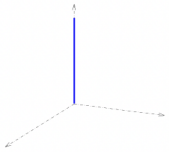
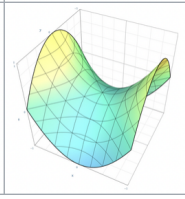
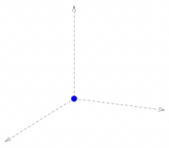
Nicht ausgeartete Quadriken		Ausgeartete Quadriken (gekrümmte Flächen)		Ausgeartete Quadriken (Ebenen u. a.)	
Ellipsoid $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}+\frac{z^2}{\gamma^2}=1$		Elliptischer Kegel $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}-\frac{z^2}{\gamma^2}=0$		Zwei schneidende Ebenen $\frac{x^2}{\alpha^2}-\frac{y^2}{\beta^2}=0$	
Einschaliges Hyperboloid $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}-\frac{z^2}{\gamma^2}=1$		Elliptischer Zylinder $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}=1$		Zwei parallele Ebenen $\frac{x^2}{\alpha^2}=1$	
Zweischaliges Hyperboloid $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}-\frac{z^2}{\gamma^2}=-1$		Hyperbolischer Zylinder $\frac{x^2}{\alpha^2}-\frac{y^2}{\beta^2}=1$		Eine Ebene $\frac{x^2}{\alpha^2}=0$	
Elliptisches Paraboloid $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}-2z=0$		Parabolischer Zylinder $\frac{x^2}{\alpha^2}-2y=0$		Eine Gerade $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}=0$	
Hyperbolisches Paraboloid $\frac{x^2}{\alpha^2}-\frac{y^2}{\beta^2}-2z=0$				Ein Punkt $\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}+\frac{z^2}{\gamma^2}=0$	

Abbildung 1: Quadriken in $\mathbb{R}^3$. Quelle: `de.wikipedia.org/wiki/Quadrik`

Definition 12.9.1. Sei K ein Körper und seien V, W und X gegebene K-Vektorräume. Dann ist eine *K-bilineare Abbildung* eine Abbildung

$$\alpha \ : \ V \times W \to X$$

die in beiden Argumenten K-linear ist, also

$$\alpha(\lambda v + \lambda' v', w) = \lambda\alpha(v, w) + \lambda'\alpha(v', w) \quad \text{und} \quad \alpha(v, \lambda w + \lambda' w') = \lambda\alpha(v, w) + \lambda'\alpha(v, w')$$

für alle $\lambda, \lambda' \in K$ und $v, v' \in V$, $w, w' \in W$.

Sei $\mathcal{B}$ eine Basis von V und $\mathcal{B}'$ eine Basis von W. Eine bilineare Abbildung α ist durch ihre Werte $\alpha(b, b')$ mit $b \in \mathcal{B}$ und $b' \in \mathcal{B}'$ festgelegt, vgl. Betrachtung 12.2.3 für Bilinearformen.

Offenbar ist dann für jede lineare Abbildung $f : X \to X'$ auch die Abbildung $f \circ \alpha : V \times W \to X'$ bilinear. Wir stellen uns die Frage, ob es für je zwei K-Vektorräume V, W einen "universellen" K-Vektorraum U mit einer "universellen" bilinearen Abbildung $\kappa : V \times W \to U$ gibt, so dass eine beliebige bilineare Abbildung $\beta : V \times W \to X$ dann so durch eine lineare Abbildungen $f : U \to X$ in der Form $\beta = f \circ \kappa$ beschrieben werden kann.

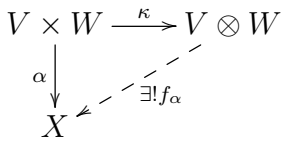
Definition 12.9.2. Das *Tensorprodukt* zweier K-Vektorräume V, W ist ein Paar, bestehend aus einem K-Vektorraum $V \otimes W$ und einer bilinearen Abbildung

$$\begin{array}{rccc} \kappa : & V \times W & \to & V \otimes W \\ & (v, w) & \mapsto & v \otimes w \end{array}$$

mit der folgenden universellen Eigenschaft: zu jeder *bilinearen* Abbildung

$$\alpha \,:\, V \times W \to X$$

gibt es genau eine *lineare* Abbildung $f_\alpha : V \otimes W \to X$ mit $\alpha = f_\alpha \circ \kappa$. Als Diagramm:



Satz 12.9.3. *Für zwei K-Vektorräume V, W existiert $V \otimes W$ und ist eindeutig bis auf eindeutige Isomorphie.*

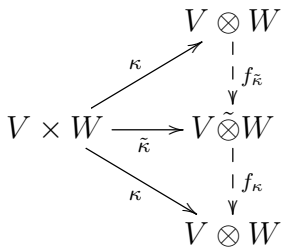
Beweis:. Da das Tensorpodukt durch eine universelle Eigenschaft gegeben ist, muss es eindeutig sein, wenn es existiert. Wir führen den Beweis noch einmal detailliert aus:

Angenommen, wir hätten zwei solche universelle Abbildungen

$$\kappa \,:\, V \times W \to V \otimes W \qquad \tilde{\kappa} \,:\, V \times W \to V \tilde{\otimes} W \;.$$

Man benutzt die universelle Eingeschaft von κ und findet für die spezielle bilineare Abbildung $\tilde{\kappa}$ eine eindeutige lineare Abbildung $f_{\tilde{\kappa}} : V \otimes W \to V \tilde{\otimes} W$ mit $f_{\tilde{\kappa}} \circ \kappa = \tilde{\kappa}$.

Durch Vertauschen der Rollen erhält man ebenso eine lineare Abbildung $f_\kappa : V \tilde{\otimes} W \to V \otimes W$ mit $f_\kappa \circ \tilde{\kappa} = \kappa$. Die Abbildungen $\kappa = \mathrm{id}_{V\otimes W} \circ \kappa$ und $f_\kappa \circ f_{\tilde{\kappa}} \circ \kappa$ beschreiben die gleiche bilineare Abbildung $V \times W \to V \otimes W$. Wegen der Eindeutigkeitsaussage in der universellen Eigenschaft folgt $f_\kappa \circ f_{\tilde{\kappa}} = \mathrm{id}_{V\otimes W}$. Als Diagramm:



Analog folgt $f_{\tilde{\kappa}} \circ f_\kappa = \mathrm{id}_{V\tilde{\otimes}W}$. (Dies ist genau das gleiche Argument wie im Beweis von Satz 11.0.6 an.)

Um die Existenz des Tensorprodukts zu zeigen, wählen wir eine Basis $\mathcal{B} := \{b_1, b_2, \ldots\}$ von V und $\mathcal{B}' := \{b'_1, b'_2, \ldots\}$ von W. Da eine bilineare Abbildung durch ihre Werte auf allen Paaren (b_i, b'_j) eindeutig festgelegt ist, brauchen wir uns nur einen Vektorraum zu verschaffen, für den eine Basis durch diese Paare indiziert wird. Sei also $V \otimes W$ der Vektorraum der K-wertigen Funktionen auf der Menge $\mathcal{B} \times \mathcal{B}'$, die nur für endlich viele Elemente einen Wert ungleich Null annehmen. Wir bezeichnen mit $b_i \otimes b'_j$ die Funktion, die auf dem Paar (b_i, b'_j) den Wert Eins und sonst den Wert Null hat. Dies ist offensichtlich eine Basis für $V \otimes W$.

Die bilineare Abbildung κ ist dann auf dem Paar (b_i, b'_j) vorgeschrieben durch $\kappa(b_i, b'_j) = b_i \otimes b'_j$. Sie erfüllt die universelle Eigenschaft, denn der bilinearen Abbildung $\alpha : V \times W \to X$ wird die eindeutig bestimmte lineare Abbildung $f_\alpha \,:\, V \otimes W \to X$ mit $f_\alpha(b_i \otimes b'_j) = \alpha(b_i, b'_j)$ zugeordnet. □

Insbesondere ist für endlich-erzeugte Vektorräume V, W die Dimension des Tensorprodukts gleich $\dim_K V \otimes W = \dim_K V \cdot \dim_K W$.

Die Elemente des Vektorraums $V \otimes W$ heißen *Tensoren*. Wir schreiben auch $v \otimes w$ für $\kappa(v, w)$. Dann folgt aus der Bilinearität von κ sofort

$$(\lambda_1 v_1 + \lambda_2 v_2) \otimes w = \lambda_1 v_1 \otimes w + \lambda_2 v_2 \otimes w \text{ und } v \otimes (\lambda_1 w_1 + \lambda_2 w_2) = v \otimes \lambda_1 w_1 + v \otimes \lambda_2 w_2 \;.$$

Die Elemente der Form $v \otimes w$ mit $v \in V$ und $w \in W$ heißen *Tensorprodukte*. Die Tensorprodukte erzeugen $V \otimes W$, aber nicht jedes Element von $V \otimes W$ ist das Tensorprodukt eines Vektors $v \in V$ und $w \in W$.

Beispiele 12.9.4.

1. Seien K^n und K^m mit Standardbasen (e_i) und (f_j) ausgestattet. Dann ist $(e_i \otimes f_j)_{i=1,\dots,n;j=1,\dots m}$ eine Basis für $K^n \otimes K^m \cong K^{nm}$.

2. Gegeben zwei Vektorräume V und W ist $(V \otimes W)^* \cong \mathrm{Bil}(V, W)$. Wenn also V und W endilich-dimensional sind ist $V \otimes W \cong (V \otimes W)^{**} \cong \mathrm{Bil}(V, W)^*$.

3. Den Polynomring in mehreren Variablen definiert man induktiv. Insbesondere ist $K[t_1, t_2] := (K[t_1])[t_2]$ definiert als der Polynomring für den kommutativen Ring $K[t_1]$. Eine Basis sind die Monome $t_1^{n_1} t_2^{n_2}$ mit $n_1, n_2 \in \mathbb{N}_0$.

   Wir betrachten die Multiplikation von Polynomen

   $$\begin{array}{rccc} \xi: & K[t] \times K[t] & \to & K[t_1, t_2] \\ & (P(t), Q(t)) & \mapsto & P(t_1) \cdot Q(t_2) \end{array}$$

   Sie ist bilinear; nach der universellen Eigenschaft des Tensorprodukts induziert sie eine lineare Abbildung

   $$\xi_\otimes: \quad K[t] \otimes K[t] \to K[t_1, t_2]$$

   mit $t^i \otimes t^j \mapsto t_1^i \cdot t_2^j$. Da die Monome eine Basis der Polynomringe bilden, ist $\xi_\otimes$ bijektiv und wir haben den Polynomring in zwei Variablen als Tensorprodukt geschrieben.

4. Sei W ein beliebiger $\mathbb{R}$-Vektorraum. Betrachte $\mathbb{C}$ als zwei-dimensionalen $\mathbb{R}$-Vektorraum. Das Tensorprodukt $\mathbb{C} \otimes_{\mathbb{R}} W$ ist ein reeller Vektorraum, der durch die bilineare Abbildung

   $$\begin{array}{rcl} \mathbb{C} \times (\mathbb{C} \otimes_{\mathbb{R}} W) & \to & \mathbb{C} \otimes_{\mathbb{R}} W \\ (\lambda, \sum_j \lambda_j \otimes w_j) & \mapsto & \sum_j \lambda \cdot \lambda_j \otimes w_j \end{array}$$

   mit einer skalaren Multiplikation mit komplexen Zahlen so ausgestattet wird, dass er ein $\mathbb{C}$-Vektorraum wird. Wir erhalten also einen $\mathbb{C}$-Vektorraum $\mathbb{C} \otimes_{\mathbb{R}} W$, indem wir von allen Elementen von W auch komplexe skalare Vielfache zulassen. (Dies ist die formale Operation, mit der Sie eine reelle Matrix als komplexe Matrix auffassen können!)

Betrachtung 12.9.5. Für zwei K-lineare Abbildungen

$$f: V \to V' \qquad \text{und} \qquad g: W \to W'$$

betrachten wir das Diagramm:

$$\begin{array}{ccc} V \times W & \xrightarrow{\kappa} & V \otimes W \\ {\scriptstyle f \times g}\downarrow & & \downarrow{\scriptstyle \exists! f \otimes g} \\ V' \times W' & \xrightarrow[\kappa']{} & V' \otimes W' \end{array}$$

Da die Abbildung $\kappa' \circ (f \times g)$ offenbar bilinear ist, existiert nach der universellen Eigenschaft der Tensorprodukts $V \otimes W$ eine lineare Abbildung

$$f \otimes g: \quad V \otimes W \to V' \otimes W' \ .$$

Diese erfüllt also

$$(f \otimes g)(v \otimes w) = f(v) \otimes g(w) \quad \text{für alle } v \in V \text{ und } w \in W \ .$$

Bemerkungen 12.9.6. 1. Aus der Bilinearität von κ folgt, dass auch das Tensorprodukt von Abbildungen bilinear ist:

$$\begin{aligned}(\lambda_1\Phi_1 + \lambda_2\Phi_2) \otimes \Psi &= \lambda_1\Phi_1 \otimes \Psi + \lambda_2\Phi_2 \otimes \Psi \\ \Phi \otimes (\lambda_1\Psi_1 + \lambda_2\Psi_2) &= \Phi \otimes \lambda_1\Psi_1 + \Phi \otimes \lambda_2\Psi_2\end{aligned}$$

2. Ebenso folgt für Vektorräume die Verträglichkeit mit direkten Summen:

$$(V_1 \oplus V_2) \otimes W \cong (V_1 \otimes W) \oplus (V_2 \otimes W)\,,$$

und analog im anderen Argument. Das Tensorproduk distribuirt also über der direketen Summe.

Man hat kanonische Isomorphismen

$$\begin{aligned}a_{U,V,W} \,:\, U \otimes (V \otimes W) &\to (U \otimes V) \otimes W \\ u \otimes (v \otimes w) &\mapsto (u \otimes v) \otimes w\end{aligned}$$

mit deren Hilfe man die K-Vektorräume $U \otimes (V \otimes W)$ und $(U \otimes V) \otimes W$ identifizieren kann. Das Tensorprodukt ist dann assoziativ.

3. Für endlich-dimensionale Vektorräume ist die kanonische Abbildung

$$\begin{aligned}V^* \otimes W &\to \mathrm{Hom}_K(V, W) \\ \alpha \otimes w &\mapsto (v \mapsto \alpha(v)w)\end{aligned}$$

ein Isomorphismus.

4. Für beliebige Vektorräume U, V, W gibt es einen kanonischen Isomorphismus

$$\mathrm{Hom}_K(U \otimes V, W) \cong \mathrm{Hom}_K(U, \mathrm{Hom}_K(V, W)$$

gegeben durch $f \mapsto f^{\#}$ mit $f^{\#}(u) : v \mapsto f(u \otimes v)$. Vergleichen Sie Bemerkung 1.6.18.

Wenn wir nun für gegebenes $k \in \mathbb{N}$ den Vektorraum $V^{\otimes k}$ betrachten, gibt es eine offensichtliche Symmetrie: Wir können die verschiedenen Faktoren vertauschen.

Das drückt sich aus, in dem wir für jedes Gruppenelement $\sigma \in S_k$ den Tensor $v = v_1 \otimes v_2 \otimes \cdots \otimes v_k$ nach

$$\sigma.v = v_{\sigma(1)}v_{\sigma(2)} \cdots v_{\sigma(k)}$$

schicken. Diese Zuordnung gibt eine Abbildung $S_k \times V^{\otimes k} \to V^{\otimes k}$ und man sieht leicht, dass $\tau.(\sigma.v) = (\tau \circ \sigma).v$ gilt sowie $e.v = v$.

Wir nennen eine solche Abbildung eine *Gruppenwirkung* von S_k auf $V^{\otimes k}$, wir können die gleichen Daten auch als Gruppenhomomorphismus $S_k \to GL(V^{\otimes k})$ betrachten.

Wir können nun verschiedene Untervektorräume von $V^{\otimes k}$ betrachten, in Abhängigkeit von der Wirkung von S_k.

Zum Beispiel schreiben wir eine Basis von $V \otimes V$ bestehend aus Vektoren der Form $v_i \otimes v_i, v_i \otimes v_j + v_j \otimes v_i$ und $v_i \otimes v_j - v_j \otimes v_i$ wenn (v_i) eine Basis von V ist. Dann spannen die antisymmetrischen Basisvektoren $v_i \otimes v_j - v_j \otimes v_i$ einen Unterraum auf, auf dem das nichttriviale Element τ von S_2 als -1 wirkt. Die anderen, symmetrischen Basisvektoren spannen einen Unterraum auf, auf dem τ als $+1$ wirkt.

Wir erhalten eine Zerlegung $V \otimes V = \mathrm{Sym}^2(V) \oplus \wedge^2(V)$ in symmetrische und antisymmetrische Tensoren. Vergleichen Sie dies mit der Zerlegung von Bilinearformen!

Wir definieren allgemein den Unterraum $\wedge^k V \subset V^{\otimes k}$ als den Unterraum aller Vektoren, auf denen jede Transposition in S_k als -1 wirkt.

Satz 12.9.7. *Sei V ein K-Vektorraum mit $\dim_K V = n$. Der Vektorraum $\wedge^n(V)$ ist eindimensional. Sei $f \in \mathrm{End}_K(V)$. Dann ist die induzierte Abbildung $\wedge^n(f) : \wedge^n(V) \to \wedge^n(V)$ die Mulitiplikation mit der Determinante von f.*

Beweisidee:. Durch Berechnung der Wirkung von Transpositionen auf allgemeinen Elementen von $V^{\otimes n}$ stellen wir fest, dass $\sum_{\sigma \in S_n} \mathrm{sign}(\sigma) v_{\sigma(1)} \otimes \cdots v_{\sigma(n)}$ genau $\wedge^n V$ aufspannt. Die induzierte Abbildung wird dann genau von der Leibnizformel für die Determinante gegeben. □

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

$$A = \sum_{i=1}^{n} a_i = \sum_{i=1}^{k} a_i + \sum_{i=k+1}^{n} a_i. \tag{6}$$

Für $k \in \{1, 2, \ldots, n\}$ ist das klar. Aber was ist mit $k = 0$ oder $k = n$? $\sum_{i=1}^{0} a_i$ ist eine leere Summe, da es keine ganze Zahl i mit $i \geq 1$ und $i \leq 0$ gibt, haben wir keinen Summanden. Aber 6 kann weiterhin gelten, solange wir diese leere Summe gleich 0 setzen.

Allgemein: Es soll immer gelten $\sum_{i \in I} a_i + \sum_{i \in J} a_i = \sum_{I \ J} a_i$ 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 $\{A_1, \ldots, A_n\}$ gegeben, die alle disjunkt sind, d.h. $A_i \cap A_j = \emptyset$ für alle i, j. Dann soll

$$|A_1 \cup \cdots \cup A_n| = |A_1| + \cdots + |A_n|$$

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 $\sum_{\emptyset} = |\bigcup_{\emptyset}| = 0$. 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:

$$\prod_{i=1}^{n} a_i = \prod_{i=1}^{k} a_i \cdot \prod_{i=k+1}^{n} a_i. \tag{7}$$

Aber damit dies stimmt, wenn $k = 0$ oder $k = n$ 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 $\{1, \ldots, n\} \to A$. 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 $\{1, \ldots, n\} \to \{1, \ldots, n\}$. 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 Alternativer Beweis für Cayley-Hamilton

Dieser Beweis funktioniert über einem beliebigen Körper.

Bewei von Satz 9.6.1:.

- Sei $v \in K^n \setminus \{0\}$ beliebig. Setze

$$v_i := A^i v \quad i = 0, 1, \ldots$$

Wähle m so, dass die Familie $(v_0, v_1, \ldots, v_{m-1})$ linear unabhängig, aber die Familie $(v_0, \ldots, v_m)$ linear abhängig ist. Es ist $m \leq n$, und es gilt

$$v_m = -\alpha_0 v_0 - \ldots - \alpha_{m-1} v_{m-1} \quad \text{mit gewissen } \alpha_i \in K\ . \qquad (*)$$

- Sei $W = \operatorname{span}_K(v_0, \ldots, v_{m-1})$. Dann ist offenbar W ein A–invarianter Untervektorraum, denn die darstellende Matrix von $A_{|W}$ bezüglich der Basis $(v_0, \ldots, v_{m-1})$ hat die Gestalt

$$\left(\begin{array}{ccccc|c} 0 & 0 & 0 & & & -\alpha_0 \\ 1 & 0 & 0 & & & -\alpha_1 \\ 0 & 1 & 0 & & & \vdots \\ 0 & 0 & 1 & & & \vdots \\ & & & \ddots & & \vdots \\ & & & & 1 & -\alpha_{m-1} \end{array}\right) .$$

- Wir berechnen das charakteristische Polynom der Matrix $A|_W$:

$$P_{A|_W}(X) = \det \left(\begin{array}{ccccc|c} X & & & & & \alpha_0 \\ -1 & X & & & & \\ & -1 & & & & \vdots \\ & & & \ddots & X & \alpha_{m-2} \\ & & & & -1 & X + \alpha_{m-1} \end{array}\right) .$$

Die Entwicklung nach der letzten Spalte liefert:

$$
\begin{aligned}
P_{A|W}(X) = (-1)^{m+1}\alpha_0 \cdot \det \begin{pmatrix} -1 & X & & & \\ & -1 & X & & \\ & & -1 & & \\ & & & \ddots & \\ & & & & X \\ & & & & -1 \end{pmatrix} \\
+(-1)^{m+2}\alpha_1 \cdot \det \begin{pmatrix} X & 0 & & & & \\ 0 & -1 & X & & & 0 \\ & 0 & -1 & X & & \\ & & & & \ddots & \\ & & & & & X \\ 0 & & & & & -1 \end{pmatrix} \\
+(-1)^{m+3}\alpha_2 \cdot \det \begin{pmatrix} X & 0 & & & & \\ -1 & X & 0 & & & 0 \\ & 0 & -1 & X & & \\ & & & & \ddots & \\ & & & & & X \\ 0 & & & & & -1 \end{pmatrix} \\
+\ldots+(-1)^{m+(m-1)}\alpha_{m-2} \cdot \det \begin{pmatrix} X & & & & 0 \\ -1 & X & & & \\ & -1 & X & & \\ & & & \cdots & \\ & & & & X \\ & & & & -1 \quad -1 \end{pmatrix} \\
+(-1)^{2m}\left(\alpha_{m-1}+X\right) \cdot \det \begin{pmatrix} X & & & 0 \\ -1 & X & & \\ & -1 & & \\ & & \ddots & \\ & & -1 & X \end{pmatrix} \\
= \alpha_0 \cdot 1 + \alpha_1 X + \ldots + \alpha_{m-2}X^{m-2} + \alpha_{m-1}X^{m-1} + X^m
\end{aligned}
$$

- Nach Lemma 9.5.1 existiert ein Polynom $g \in K[X]$ mit fix

$$P_A(X) = g(X)P_{A|W}(X) \ .$$

Es folgt

$$
\begin{aligned}
\widetilde{P_A}(A)v &= \widetilde{g}(A)\widetilde{P_{A|W}}(A)v \\
&= \widetilde{g}(A)\left(A^m + \alpha_{m-1}A^{m-1} + \ldots + \alpha_1 A + \alpha_0\right)v \\
&= \widetilde{g}(A)\left(v_m + \alpha_{m-1}v_{m-1} + \ldots + \alpha_1 v_1 + \alpha_0 v\right) = 0 \ ,
\end{aligned}
$$

wobei wir im letzten Schritt $(*)$ verwendet haben. Da dies für alle $v \in V$ gilt, ist die Behauptung gezeigt.

□

C 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 Quotienvektorraum V/U konstruieren, in dem wir den Unterraum U "gleich null setzen". Es gibt eine kanonische Surjektion $V \to V/U$.
   - 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 $\oplus_{i\in I} V_i$.

3. Wir haben K-lineare Abbildungen $f : V \to W$ 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 $\mathrm{Hom}_K(V, W)$ der K-linearen Abbildungen ist selbst ein K-Vektorraum.
   - Es gilt der Homomorphiesatz 3.4.4: für eine lineare Abbildung $f : V \to W$ erhalten wir einen eindeutigen Isomorphismus $\bar{f} : V/\ker f \xrightarrow{\sim} \mathrm{Im}\,(f)$ und eine Faktorisierung von $f = \iota \circ \bar{f} \circ \pi$ wie in diesem Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ {\scriptstyle\pi}\downarrow & & \uparrow{\scriptstyle\iota} \\ V/\ker f & \xrightarrow{\ \bar{f}\ } & \Phi(W) \end{array}$$

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 $\mathrm{span}_K(\mathcal{B}) = V$.
   - $\mathcal{B}$ 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 $\mathcal{B}$ Erzeugendensytem 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 erlaubte es uns, aus Erzeugendensystemen eine Basis auszuwählen.
- Der Basisergänzungssatz erlaubt es, linear unabhängige Familien zu Basen zu ergänzen.

Für jede lineare Abbildung $f : V \to W$ zwischen endlichdimensionalen Vektorräumen gilt die Dimensionsformel. Mit $\operatorname{rg}(f) = \dim(\operatorname{Im}(f))$ ist $\operatorname{rg}(f) + \dim \ker(f) = \dim(V)$.

Wir können auch die folgenden Dimensionen berechnen:

- $\dim_K(W_1 + W_2) = \dim_K(W_1) + \dim_K(W_2) - \dim_K(W_1 \cap W_2)$.
- $\dim(V/U) = \dim V - \dim U$
- $\dim \operatorname{Hom}_K(V, W) = \dim(V) \cdot \dim(W)$.

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 $f : V \to W$ 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:

     $$\Phi_{\mathcal{A}} : K^n \to V \text{ mit } \Phi_{\mathcal{A}}(e_i) = v_i \,.$$

   - 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

     $$M_{\mathcal{B}}^{\mathcal{A}} : \operatorname{Hom}_K(V, W) \to M(m \times n, K)$$

     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
$$T^{\mathcal{A}}_{\mathcal{A}'} = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)$$
beschrieben. Für lineare Abbildungen gilt die Transformationsformel 5.4.6
$$M^{\mathcal{A}'}_{\mathcal{B}'}(\Phi) = T^{\mathcal{B}}_{\mathcal{B}'} \cdot M^{\mathcal{A}}_{\mathcal{B}}(\Phi) \cdot \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)^{-1}.$$
- 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 $r = \mathrm{rg}\,(A)$. 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
$$\mathrm{Lsg}(A, b) := \{x \in K^n \mid Ax = b\}$$
eines inhomogenen linearen Gleichungssystems gilt:
   - $\mathrm{Lsg}(A, b) = \emptyset$ genau dann, wenn $b \notin \mathrm{Im}\, A$.
   - $\mathrm{Lsg}(A, b)$ ist entweder leer oder ein affiner Unterraum der Form $x_0 + \ker(\alpha)$ mit Dimension $n - \mathrm{rg}\, A$. 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.

D Was noch geschah: Kurzzusammenfassung der Kapitel 7-12

Sei K im Folgenden ein beliebiger Körper, V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$ ein Endomorphismus.

1. Für eine Matrix A definieren wir die *Determinante*.

   Das ist die eindeutig bestimmte Abbildung

   $$\det : M(n \times n, K) \to K$$

   die (D1) K-zeilenlinear, (D2) alternierend und (D3) normiert ist eingeführt.

   Zur *Berechnung* verwenden wir eine der folgenden Methoden:

   - Wir entwickeln induktiv nach Zeilen oder Spalten
   - Wir wenden die Leibniz'sche Regel an: $\det(A) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$, wobei $\mathrm{sign} : S_n \to \mathbb{Z}/2\mathbb{Z}$ ein surjektiver Gruppenhomomorphismus ist.
   - Für obere Dreiecksmatrizen gilt

     $$\det \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} \lambda_i$$

   - Für eine blockdiagonale Matrix gilt

     $$\det \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix} = \det A_1 \cdot \det A_2$$

   - Eine Determinante ändert sich nicht, wenn man ein Vielfaches einer Zeile zu einer anderen Zeile addiert. Man kann durch diese Umformungen eine Matrix in eine obere Dreiecksmatrix überführen und dann die Determinante als Produkt der Diagonalelemente berechnen.

   Die Determinante hat folgende *Eigenschaften :*

   - Multiplikativität: $\det(AB) = \det A \cdot \det B$. Deswegen sind die Determinante ähnlicher Matrizen gleich, ein Endomorphismus f eines endlich-dimensionalen Vektorraums hat eine wohldefiniert Determinant $\det(f) = \det(M_{\mathcal{B}}(f))$
   - $\det A \neq 0$ genau dann wenn A invertierbar ist, genau dann wenn $\mathrm{rg}\ A$ maximal ist.

   Die *Spur* $\mathrm{Tr}\ (A) = \sum_{i=1}^{r} a_{ii}$ ist eine lineare Funktion $M(n \times n, K) \to K$, die $\mathrm{Tr}\ (AB) = \mathrm{Tr}\ (BA)$ erfüllt und deshalb auch für Endomorphismen von endlich-dimensionalen Vektorräumen wohldefiniert ist.

2. Gilt $fv = \lambda v$ mit $v \neq 0$, so heißt $\lambda \in K$ *Eigenwert* und $v \in V \setminus \{0\}$ *Eigenvektor* von f.

   Der *Eigenraum* von f zum Eigenwert $\lambda \in K$ ist $\mathrm{Eig}(f, \lambda) := \ker(f - \lambda \mathrm{id}_V) \leq V$.

   Genauso definieren wir Eigenwerte und Eigenvektoren von Matrizen.

- Die geometrische Vielfachheit ist

$$\mu_{geo}(f,\lambda) := \dim_K \operatorname{Eig}(f,\lambda)$$

- Die Eigenwerte sind genau die Nullstellen des *charakteristischen Polynoms* $P_f(X) := \det(X\mathrm{id}_V - f)$. Die Vielfachheit von λ als Nullstelle von P_f heißt *algebraische Vielfachheit*.

  Determinante und Spur sind spezielle Koeffizienten von P_f:

$$a_n = 1 \qquad a_{n-1} = -\operatorname{Tr} f \qquad a_0 = (-1)^n \det f \ .$$

- Es gilt für jeden Eigenwert λ:

$$1 \le \mu_{geo}(f,\lambda) \le \mu_{alg}(f,\lambda) \ .$$

3. Polynome sind Elemente der *Polynomalgebra* $K[X]$.

   - Jedes Element a einer K-Algebra S definiert einen Einsetzungshomomorphismus, der X auf a abbildet, $\widetilde{P}(a)$ ist die Auswertung von P bei $X = a$.
   - In Polynomalgebren über Körpern gibt es eine Division mit Rest, a ist eine Nullstelle des Polynoms P genau wenn $(X - a) \mid P$.
   - Fundamentalsatz der Algebra: Polynome in $\mathbb{C}[X]$ zerfallen vollständig in Linearfaktoren.

4. Wir betrachten das Problem der Diagonalisierbarkeit: Eine lineare Abbildung bzw. Matrix ist diagonalisierbar, wenn Eigenvektoren eine Basis für V bzw. K^n bilden.

   - Das Minimalpolynom $\mu_f \in K[X]$ ist das normierte Polynom mit minimalem Grad, so dass $\widetilde{\mu_f}(f) = 0 \in \operatorname{End}_K(V)$.
   - Satz von Cayley–Hamilton: $\widetilde{P}_f(f) = 0 \in \operatorname{End}_K(V) \Leftrightarrow \mu_f \mid P_f$.
   - Minimalpolynom und charakteristisches Polynom haben die gleichen Nullstellen.
   - Wenn P_f vollständig in Linearfaktoren zerfällt, existiert eine geordnete Basis $\mathcal{B}$ von V, so dass die darstellende Matrix *Jordansche Normalform* hat:

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & 0 \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{matrix}} & & \\ & & \\ & & \ddots \end{pmatrix}$$

     Die diagonalen Einträge sind die Eigenwerte von f. Die jordansche Normalform ist nur bis auf Umordnung der Jordan-Blocks eindeutig.

   - Die Anzahl der Blöcke zum Eigenwert λ ist $\mu_{geo}(f;\lambda)$.

     Die Größe des größten Blocks zum Eigenwert λ ist die Vielfachheit von λ als Nullstelle in μ_f.
   - Zur Bestimmung der jordanschen Normalform bestimmen wir für jeden Eigenwert λ die Räume $U_i = \ker(f - \lambda\mathrm{id}_V)^i$.

     Dann ist die Anzahl der Jordan-Blocks mit Größe i gegeben als $2\dim U_i - \dim U_{i+1} - \dim U_i$.

- f ist genau dann diagonalisierbar wenn P_f vollständig in Linearfaktoren zerfällt und $\mu_{geo}(\lambda, f) = \mu_{alg}(\lambda, f)$ für alle Eigenwerte ist. Das gilt genau dann wenn μ_f vollständig in paarweise verschiedene Linearfaktoren zerfällt .
- Zwei diagonalisierbare Endomorphismen f, g sind genau dann *gleichzeitig* diagonalisierbar, wenn sie kommutieren, $f \circ g = g \circ f$.

5. Für jeden Vektorraum V betrachten wir den Dualraum $V^* = \mathrm{Hom}_K(V, K)$.

   - Einem Vektorraum V mit Basis $\mathcal{B} = \{b_1, \ldots, b_n\}\}$ ordnen wir seinen Dualraum $V^* := \mathrm{Hom}(V, K)$, wenn $\dim_K V < \infty$ mit dualer Basis $\mathcal{B}^* = \{b_1^*, \ldots, b_n^*\}$ mit $b_i^*(b_j) = \delta_{ij}$ zu.
   - Einer linearen Abbildung $V \xrightarrow{f} W$ ordnen wir zu die duale Abbildung $W^* \xrightarrow{f^*} V^*$. Es gilt $(f \circ g)^* = g^* \circ f^*$ und

     $$\mathrm{Im}\ f^* = (\ker\ f)^\circ \qquad \ker\ f^* = (\mathrm{Im}\ f)^\circ \qquad M_{\mathcal{A}^*}^{\mathcal{B}^*}(f^*) = M_{\mathcal{B}}^{\mathcal{A}}(f)^T\ ,$$

     wobei $U^\circ = \{f \in V^* \mid f|_U = 0\} \leq V^*$ der Annullator des Untervektorraums $U \subset V$ ist.
   - Ist $\dim V < \infty$, so gibt es einen kanonische Isomorphismen $i_V : V \xrightarrow{\sim} V^{**}$.

6. Wir betrachten *Bilinearformen* $\beta : V \times W \to K$

   - Die darstellende Matrix $M_{\mathcal{A},\mathcal{B}}(\beta)_{ij} = \beta(v_i, w_j)$erfüllt die *Transformationsformel* für Basiswechsel:

     $$M_{\mathcal{A},\mathcal{B}}(\beta) = \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^T M_{\mathcal{A}',\mathcal{B}'}(\beta) T_{\mathcal{B}'}^{\mathcal{B}}\ .$$

     Entsprechend heißen Matrizen M, N kongruent, wenn es $T \in GL(n, K)$ gibt mit $N = T^T M T$. Sie stellen die gleiche Bilinearform auf V für verschiedene Basen dar.
   - β ist *nicht-ausgeartet* wenn für jedes $v \in V \setminus \{0\}$ existiert $w \in V$ mit $\beta(v, w) \neq 0$. Das gilt genau wenn $\det M_{\mathcal{B}}(\beta) \neq 0$.

   Im Folgenden gelte im Körper K die Ungleichung $1 + 1 \neq 0$.

   Eine Bilinearform $\beta \in \mathrm{Bil}(V, V)$ heißt *symmetrisch* wenn $\beta(x, y) = \beta(y, x)$

   - Eine symmetrische Bilinearform ist äquivalent zu einer quadratischen Form $q : V \to K$ definiert durch $q(x) = \frac{1}{2}\beta(x, x)$, dann ist $\beta(x, y) = q(x + y) - q(x) - q(y)$ (Polarisierungsformel).
   - Ein symmetrische Bilinearform β hat eine Normalform: Es gibt eine Basis (b_i) mit $\beta(b_i, b_j) = \alpha_i \delta_{i,j}$ mit $\alpha_i \in K$.
   - Für $K = \mathbb{R}$ gilt gür eine passende Basis (Sylvesterscher Trägheitssatz):

     $$M_{\mathcal{B}}(\beta) = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r_+}, \underbrace{-1, \ldots, -1}_{r_-}, \underbrace{0, \ldots, 0}_{r_0})$$

     β ist nicht-ausgeartet genau wenn $r_0 = 0$.

     β ist positiv definit, d.h. $q(v) > 0$ für alle $v \in V \setminus \{0\}$, genau wenn $r_0 = r_- = 0$.
   - $M_{\mathcal{B}}(\beta)$ ist diagonaliserbar und r_+ (bzw. r_-, bzw. r_0) ist die Anzahl der Eignwerte (mit Vielfachheit) größer als 0 (bzw. kleiner als 0, bzw. gleich 0).

7. Vektorräume mit innerem Produkt sind euklidische oder unitäre Vektorräume, K ist $\mathbb{R}$ oder $\mathbb{C}$.

   - Ein *euklidischer Vektorraum* ist ein endlich-dimensionaler reeller Vektorraum mit positiv definiter symmetrischer Bilinearform $\langle \cdot, \cdot \rangle$.
   - Ein *unitärer Vektorraum* ist ein endlich-dimensionaler komplexer Vektorraum mit positiv definiter Sesquilinearform $\langle \cdot, \cdot \rangle \in \mathrm{Bil}(V, \overline{V})$.
   - Das innere Produkt wird von der Norm $\|x\| := \sqrt{\langle x, x \rangle}$ bestimmt
   - Es gilt die Cauchy-Schwarz'sche Ungleichung:

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

   - Jeder Vektorraum mit innerem Produkt hat *Orthonormalbasen*. Mit dem Gram-Schmidt Verfahren wird eine beliebige Basis zur Orthonormalbasis.
   - Ein Untervektorraum $X \leq V$ hat ein orthogonales Komplement $X^\perp = \{v \in V \mid \langle x, v \rangle = 0 \quad \forall x \in X\}$. Es gilt $V = X \oplus X^\perp$ und wir erhalten die orthogonale Projektion auf X mit Kern $X^\perp$.

   Für eine Abbildung $f : V \to W$ zwischen zwei Vekterräumen mit innerem Produkt gibt es eine eindeutige *adjungierte Abbildung* $f^\dagger : W \to V$ mit $\langle fv, w \rangle = \langle v, f^\dagger w \rangle$.

   Alle Eigenvektoren eines normalen Endomorphismus mit $f^\dagger f = f f^\dagger$ sind orthogonal zueinander.

   Die Abbildung f ist *selbstadjungiert*,wenn $f = f^\dagger$. Dann gilt:

   - Für eine Orthonormalbasis $\mathcal{B}$ ist $M_\mathcal{B}(f) = M_\mathcal{B}(f)^* := \overline{M_\mathcal{B}(f)^T}$, die darstellende Matrix ist symmetrisch ($K = \mathbb{R}$) bzw. hermitesch ($K = \mathbb{C}$) und gleich zur adjungierten Matrix.
   - Alle Eigenwerte von f sind reell.
   - f ist mit einer Orthonormalbasis diagonalisierbar.

   Die Abbildung f ist eine Isometrie wenn $f^\dagger = f^{-1}$ ader äquivalent $\langle fv, gw \rangle = \langle v, w \rangle$ für alle $v, w \in V$ oder äquivalent $||f(v)|| = ||v||$ für alle $v \in V$. Dann gilt

   - Isometrien für euklidische Vektorräume heißen *orthogonal*, für unitäre Vektorräume *unitär*.
   - Die Eigenwerte von f erfüllen $|\lambda| = 1$ und es gilt $|\det(f)| = 1$.
   - In einer Orthornomalbasis erfült $A = M_\mathcal{B}(f)$ die Gleichung $A^* A = E_n$. (Wir nennen A eine orthogonale bzw. unitäre Matrix.)
   - Orthogonale Matrizen in $M(n \times n, \mathbb{R})$ bilden eine Gruppe $O(n)$. Unitäre Matrizen in $M(n \times n, \mathbb{C})$ bilden eine Gruppe $U(n)$.
   - Isometrien sind normal. Unitäre Matrizen sind mit einer Orthonormalbasis diagonalisierbar.

   Wenn f normal und diagonalisierbar ist gibt es $T \in GL(n, K)$ mit $T^{-1} M T = T^* M T =: D$ diagonal, M ist ähnlich und (für $K = \mathbb{R}$) kongruent zur Diagonalmatrix D.