Lineare Algebra I und II*

Alexandru Constantinescu

31. Januar 2024

Freie Universität Berlin
Wintersemester 2023/2024

* Das ist kein Skript. Es sind nur meine Notizen zur Vorlesung. Fehler können und werden vorkommen.

Inhaltsverzeichnis

[1] Camille Jordan, französischer Mathematiker, 1838-1922.

Literaturverzeichnis

Die Reihenfolge ist "Vorlesungs-chronologisch".

[Webseite] http://userpage.fu-berlin.de/aconstant/LA1.html

[Hou12] Kevin Houston, *Wie man mathematisch denkt. Eine mathematische Einführung in die mathematische Arbeitstechnik für Studienanfänger*, Heidelberg: Springer Spektrum, **2012**.

[HA72] David Hilbert und Wilhelm, *Grundzüge der theoretischen Logik*, 6. Auflage, Springer-Verlag Berlin Heidelberg GmbH **1972**.

[Rau08] Wolfgang Rautenberg, *Grundkurs Mengenlehre*, Skript FU Berlin, http://page.mi.fu-berlin.de/raut/Mengenlehre/m.pdf **2008**.

[Bri85] Egbert Brieskorn, *Lineare Algebra und Analytische Geometrie I*, Vieweg&Sohn Verlag, **1985**.

[Fis01] Gerd Fischer, *Analytische Geometrie - Eine Einführung für Studienanfänger*, Vieweg+Teubner Verlag, **2001**.

[Fis09] Gerd Fischer, *Lehrbuch Lineare Algebra und Analytische Geometrie*, Vieweg+Teubner Verlag, **2011**.

[Bos08] Siegfried Bosch, *Lineare Algebra*, Springer-Lehrbuch, **2014**.

[Art91] Michael Artin, *Algebra*, Prentice Hall, **1991**.

[Rom08] Steven Roman, *Advanced Linear Algebra*, Springer, Graduate Texts in Mathematics Volume 135, **2008**.

Teil I

Lineare Algebra 1

Wichtige Zeichen

$\neg$ **oder** nicht()	Negation
$\Rightarrow$	impliziert
$\Longleftrightarrow$	äquivalent
$\forall$	für alle
$\exists$	es existiert
$\exists!$	es existiert und ist eindeutig
$\in$	ist ein Element von / ist in
$\emptyset$	die leere Menge
$\subseteq$	ist eine Teilmenge von (darf gleich sein)
$\subsetneq$	ist eine Teilmenge von, aber nicht gleich
$\cap$	Schnitt
$\cup$	Vereinigung
$\setminus$	Mengendifferenz
: **oder** |	mit der Eigenschaft, dass (wenn man eine Menge definiert, z.B. $\mathbb{N}_{>0} := \{n \in \mathbb{N} \ : \ n > 0\}$)
(a, b)	geordnetes Paar **oder** das offene Interval
$\sharp A$ oder $\|A\|$	die Kardinalität der Menge A
$A \to B$	eine Abbildung von A nach B
$A \hookrightarrow B$	eine injektive Abbildung von A nach B
$A \twoheadrightarrow B$	eine surjektive Abbildung von A nach B
$\mapsto$	bildet ab auf
f^{-1}	die inverse Abbildung **oder** das Urbild (wobei f eine Abbildung ist)
$\mathbb{N}$	die Menge der natürlichen Zahlen (samt 0)
$\mathbb{N}_{>0}$	die Menge der positiven natürlichen Zahlen
$\mathbb{Z}$	die Menge der ganzen Zahlen
$\mathbb{Q}$	die Menge der rationalen Zahlen
$\mathbb{R}$	die Menge der reellen Zahlen
$\mathbb{C}$	die Menge der komplexen Zahlen
i	die imaginäre Einheit. Es gilt $i^2 = 1$
$\overline{z}$	die konjugiert komplexe Zahl zu z
$\|z\|$	der Betrag von z
$\sum$	Summe
$\prod$	Produkt
$a \mid b$	a teilt b
$b \vdots a$	b ist teilbar durch a

Tabelle 1: Häufig beuntzte Symbole

α	A	Aplha	β	B	Beta	γ	Γ	Gamma
δ	Δ	Delta	ε	E	Epsilon	ζ	Z	Zeta
η	H	Eta	θ	Θ	Theta	ι	I	Iota
κ	K	Kappa	λ	Λ	Lambda	μ	M	My
ν	N	Ny	ξ	Ξ	Xi	o	O	Omikron
π	Π	Pi	ρ	R	Rho	σ	Σ	Sigma
τ	T	Tau	υ	Υ	Ypsilon	φ	Φ	Phi
χ	X	Chi	ψ	Ψ	Psi	ω	Ω	Omega

Tabelle 2: Griechisches Alphabet

Kapitel 1

Grundbegriffe

Wir werden folgende Zahlen-Mengen als bekannt annehmen[1]:

Die natürlichen Zahlen	$\mathbb{N}$	$=$	$\{0, 1, 2, \dots\}$
Die ganzen Zahlen	$\mathbb{Z}$	$=$	$\{\dots, -2, -1, 0, 1, 2, \dots\}$
Die rationalen Zahlen	$\mathbb{Q}$	$=$	$\{\frac{a}{b} \ : \ a, b \in \mathbb{Z} \text{ und } b \neq 0\}$
Die reellen Zahlen	$\mathbb{R}$	$=$	$\{\mathbf{x}, \mathbf{y} \ : \ \mathbf{x} = \text{endlich viele Ziffern}, \mathbf{y} = \text{unendlich viele Ziffern}\}$
Die komplexen Zahlen	$\mathbb{C}$	$=$	$\{a + i \cdot b \ : \ a, b \in \mathbb{R} \text{ und } i^2 = -1\}$.

Wenn a eine natürliche Zahl ist, dann schreiben wir $a \in \mathbb{N}$. Wenn a eine ganze Zahl ist, $a \in \mathbb{Z}$, usw. Mehr über dieser Notation finden Sie in Teil 1.2.1.

Wir werden hier die Konvention verwenden, dass die natürlichen Zahlen bei Null beginnen. Historisch gesehen war dies nicht immer der Fall und auch heute ist für manche Autoren[2] die kleinste natürliche Zahl die 1. Für unsere Zwecke wird jedoch Null als eine natürliche Zahl betrachtet.d aber Null eine natürliche Zahl sein.

Die oben gegebene Darstellung der reellen Zahlen hat einige Nachteile. Zum einen ist die Darstellung $\mathbf{x}, \mathbf{y}$ nicht eindeutig. Zum Beispiel gilt

$$0,99999\cdots = 1,00000\cdots .$$

Zum anderen vermittelt diese Darstellung oft wenig über die Eigenschaften der Zahl. Zum Beispiel sind Eigenschaften wie, dass $\sqrt{2}$ eine Nullstelle des Polynoms $x^2 - 2$ ist, oder dass die berühmte Kreiszahl π das Verhältnis zwischen dem Umfang und dem Durchmesser eines beliebigen Kreises beschreibt, nicht aus der Dezimalbruchentwicklung ablesbar. Der Vorteil dieser Darstellung liegt darin, dass alle reellen Zahlen eine solche Darstellung haben und dass es genau bestimmt werden kann, wenn zwei solche Darstellungen derselben reellen Zahl entsprechen. Diese und viele weitere Konzepte werden in der Analysis 1 Vorlesung ausführlich behandelt.

[1] Auch wenn die genauen und gründlichen Definitionen gar nicht einfach sind, für unsere Zwecke hier werden die "naiven" Definitionen reichen.

[2] Insbesondere in [Hou12], der Hauptquelle für den ersten Teil dieses Kurses, wird eine andere Konvention verwendet: Dort ist $\mathbb{N} = \{1, 2, \dots\}$.

1.1 Mathematische Aussagen

Die logischen Sachverhalte, die zwischen Urteilen, Begriffen usw. bestehen,
finden ihre Darstellung durch Formeln, deren Interpretation frei ist
von den Unklarheiten, die beim sprachlichen Ausdruck leicht auftreten können.

[HA72]

Im Zentrum jeder mathematischen Theorie stehen logische Zusammenhänge zwischen klar formulierte Aussagen. Wir werden uns hier nur ganz oberflächlich mit der Aussagenlogik und Aussagenkalkül beschäftigen. Eine gute Quelle mehr Tiefe in dieser Richtung ist [HA72].

1.1.1 Definition und Beispiele

Definition 1.1. Eine **mathematische Aussage** ist ein Satz, der entweder *wahr* oder *falsch* ist - **aber nicht beides**.

Der Wahrheitswert einer mathematischen Aussage darf nicht von den Umständen, dem Zeitpunkt oder der Person, die die Aussage ausspricht oder liest, abhängen. Also nicht alle Sätze sind Aussagen. In vielen Fällen kann das vom Kontext abhängig sein.

Beispiele:

(i) "Heute ist Montag."

(ii) "2+2=4."

(iii) "Alle Katzen sind grau."

(iv) "Alle Katzen sind nicht grau."

(v) "Es gibt Katzen die nicht grau sind."

(vi) "0=1."

(vii) "x ist eine ungerade Zahl."

(viii) "Die Wurzel einer natürlichen Zahl ist rational."

(ix) "Dieser Satz ist falsch."

Bemerkung 1.2. Ich möchte die abstrakte Schreibweise der Prädikatenlogik vermeiden. Ich finde Mathematik schöner wenn sie in einer natürlichen Sprache verfasst ist. Dabei sollte jedoch die Genauigkeit nicht geopfert werden. Daher ist es für Anfänger in der Mathematik immer eine gute Übung, mathematische Sätze in präzise Prädikate umzuformulieren.

Wir werden Aussagen mit großen Buchstaben bezeichnen: A, B, C usw. Nur in Abschnitt 1.1.7 werden wir kurz auf Aussagenvariablen eingehen und diese ähnlich bezeichnen.

1.1.2 Wahrheitstabellen und Negation

Wie in der Definition 1.1 festgelegt, ist eine Aussage entweder *wahr* oder *falsch*. Dies tragen wir in einer sogenannten Wahrheitstabelle als **Eingabe** ein:

A
wahr
falsch

Definition 1.3. Die **Negation** der Aussage A wird mit nicht(A) oder $\neg A$ bezeichnet und ist das kontradiktorische Gegenteil von A: Wenn A wahr ist, dann ist nicht(A) falsch und wenn A falsch ist, dann ist nicht(A) wahr.

Das heißt, dass wir für nicht(A) folgende Wahrheitstabelle (in Abhängigkeit von den Wahrheitswert von A) haben. In der ersten Spalte befindet sich die Eingabe und in der zweiten Spalte die Ausgabe[3].

A	nicht(A)
wahr	falsch
falsch	wahr

Negationen von einfachen Aussagen sind auch einfach. Es wird interessanter, wenn wird später Implikationen und Quantoren betrachten.

Beispiel 1.4. Die Negation von A : *Die natürliche Zahl n ist eine gerade Zahl.* ist

nicht(A): *Die natürliche Zahl n ist* **nicht** *eine gerade Zahl.*

Das kann auch als *Die natürliche Zahl n ist eine ungerade Zahl.* formuliert werden.

Bemerkung 1.5. Die Negation der Negation einer Aussage ist zu der ursprünglichen Aussage **logisch äquivalent**. Das heißt, dass die entsprechenden Spalten in der Wahrheitstabelle gleich sind:

A	nicht(A)	nicht(nicht(A))
wahr	falsch	wahr
falsch	wahr	falsch

1.1.3 Logische Verknüpfungen

Zwei Aussagen können mit Hilfe der logischen Verknüpfungen **und** und **oder** zu einer neuen Aussage verbunden werden.

Die Konjunktion: Und verhält sich wie der umgangssprachliche "und". Das heißt, die Aussage "A und B" ist nur dann wahr, wenn sowohl A als auch B wahr sind. Sonst ist "A und B" falsch.

Die Disjunktion: Oder - hier muss man mehr aufpassen. In der gewöhnlichen Sprache tritt "oder" in zweifacher Bedeutung auf:

- als "oder auch" - ein einschließliches "oder", wie im lateinischen *vel*[4],

[3] Allgemein sind die Spalten der kleinsten Bausteine die Input-Spalten und der Output sind die Spalten der Wahrheitstabelle für zusammengesetzte/abgeleitete Aussagen.

[4] Vom lateinischen *vel* kommt auch die Bezeichnung $\vee$.

- als “entweder oder” - ein ausschließliches “oder”, wie im lateinischen *aut-aut*.

Für uns ist “oder” allein immer ein **einschließliches oder**. Das heißt, damit die Aussage “A oder B” wahr ist muss mindestens eine der zwei Aussagen wahr sein; es dürfen jedoch auch beide wahr sein. Somit ist “A oder B” falsch genau dann, wenn sowohl A als auch B falsch sind. Hier ist das Ganze in Wahrheitstabellen zusammengefasst:

A	B	A und B	A oder B
w	w	w	w
w	f	f	w
f	w	f	w
f	f	f	f

In der Umgangssprache kann der Sinn von “oder” von dem Kontext und der angesprochenen Person abhängig sein. Zum Beispiel, wenn man einem Kind sagt:

Du räumst dein Zimmer auf, oder du bekommst kein Eis!

dann dürfen die zwei Aussagen nicht gleichzeitig wahr sein; es wäre gemein. In den wenigen Situationen, in denen wir das exklusive “oder” benötigen, werden wir es als “Entweder A oder B” ausdrücken. Wenn wir logische Äquivalenz mit $\Leftrightarrow$ bezeichnen, haben wir auch:

$$\text{Entweder } A \text{ oder } B \quad \Leftrightarrow \quad \text{nicht}(A \Leftrightarrow B).$$

Bemerkung 1.6. Es seien A, B und C beliebige Aussagen. Es gelten:

$$\begin{aligned}
\neg(A \text{ und } B) &\Leftrightarrow \neg(A) \text{ oder } \neg(B) \\
\neg(A \text{ oder } B) &\Leftrightarrow \neg(A) \text{ und } \neg(B) \\
A \text{ und } (B \text{ und } C) &\Leftrightarrow (A \text{ und } B) \text{ und } C \\
A \text{ oder } (B \text{ oder } C) &\Leftrightarrow (A \text{ oder } B) \text{ oder } C \\
A \text{ und } (B \text{ oder } C) &\Leftrightarrow (A \text{ und } B) \text{ oder } (A \text{ und } C) \\
A \text{ oder } (B \text{ und } C) &\Leftrightarrow (A \text{ oder } B) \text{ und } (A \text{ oder } C)
\end{aligned}$$

Beweis-Skizze: Übung. Q.E.D.

1.1.4 Implikationen

Eine **Implikation** (oder Subjunktion) ist eine Aussage der Form:

Wenn *Aussage A wahr ist,* **dann** *ist Aussage B wahr.* (1.1)

Man sagt dazu auch “A **impliziert** B” oder “Aus A folgt B” und wir schreiben dafür

$$A \Rightarrow B.$$

In diesem Kontext heißt die Aussage A **Voraussetzung** (oder Annahme, oder Bedingung) und die Aussage B heißt **Schlussfolgerung**. Implikationen sind oft nicht in der Form (1.1) gegeben. Zum Beispiel:

$$\textit{Die Summe zweier gerader Zahlen ist gerade.} \tag{1.2}$$

ist eine Implikation. Wenn wir folgende drei Aussagen betrachten:

A : *Die natürliche Zahl a ist gerade.*

B : *Die natürliche Zahl b ist gerade.*

C : *Die natürliche Zahl $a+b$ ist gerade.*

dann kann man die Aussage (1.2) so formulieren:

$$(A \text{ und } B) \Rightarrow C.$$

Eine äquivalente Formulierung der Implikation $A \Rightarrow B$ ist:

$$A \textit{ ist } \textbf{nur dann } \textit{wahr, } \textbf{wenn } B \textit{ wahr ist.}$$

Die Wahrheitstabelle für eine Implikation ist die folgende:

A	B	A⇒B
w	w	w
w	f	f
f	w	w
f	f	w

Hier ist wichtig zu bemerken, dass falsche Aussagen als Voraussetzung immer eine wahre Implikation geben. Also die Wahrheit der Implikation $A \Rightarrow B$ sagt nichts über die Wahrheit der Aussage B und genau so wenig über die Wahrheit von A. Nur wenn sowohl A als auch die Implikation $A \Rightarrow B$ wahr sind, dann folgt es, dass B wahr sein muss. Und wenn B falsch ist und $A \Rightarrow B$ wahr, dann muss auch A falsch sein.

Die Idee, dass eine falsche Voraussetzung die Implikation wahr macht scheint nicht eingängig zu sein. Hier ist mein Versuch diese falsche Intuition wieder gut zu machen. Weitere Motivation in dieser Richtung folgt gleich nach dem wir die Negation einer Implikation eingehen.

Beispiel 1.7. Die meisten Sätze in der Mathematik sagen, dass eine Implikation wahr ist:

$$\textit{Für alle } x \in \mathbb{R} \textit{ ist folgende Aussage wahr: } x > 3 \Rightarrow x^2 > 3.$$

Das ist nicht der klügste Satz, aber es ist wahr. Das heißt die obige Implikation ist wahr für alle $x \in \mathbb{R}$, auch wenn diese kleiner als 3 sind. Man kann hier auch bemerken, dass wenn $x > 3$ falsch ist, dann kann $x^2 > 3$ sowohl wahr als auch falsch sein. Also die logische Konvention aus der Wahrheitstabelle von $A \Rightarrow B$ erlaubt uns viele Theoreme als Implikation zu formulieren.

Das philosophische Problem ist, dass wahre Implikationen manchmal nichts mit dem gewöhnlichen Sinne von Folgerung zu tun haben. Zum Beispiel ist die Implikation

$$\textit{Wenn 22 durch 11 teilbar ist, dann ist die Erde rund.}$$

wahr, aber es gibt keine Folgebeziehungen zwischen den beiden Aussagen. Diese philosophische Aspekte werden wir nicht behandeln. Ich empfehle dafür [HA72, Kap.1, §11].

Die **Negation von** $A \Rightarrow B$ ist keine weitere Implikation, sondern:

$$A \text{ und } \neg(B).$$

Es ist eine gute Übung das mit Hilfe einer Wahrheitstabelle zu überprüfen. Das sollte auch intuitiv einfacher zu akzeptieren, und somit weitere Motivation für die Wahrheitstabelle einer Implikation. Zum Beispiel wenn jemand behauptet, dass

Wenn p eine Primzahl ist, dann ist p ungerade.

wahr für alle[5] p ist, dann sollte man antworten: *Nein, dass ist nicht wahr, weil* 2 *eine Primzahl ist und* 2 *ist nicht ungerade.* Man verwendet also als Argument, dass die Implikation falsch ist, weil die Negation davon wahr ist.

So kann man auch begründen, dass folgende Implikation wahr ist:

Wenn 9 eine Primzahl ist, dann ist 10 eine Primzahl. (1.3)

Die Negation davon ist: *9 ist eine Primzahl und 10 ist keine Primzahl.* Diese ist falsch, also muss die Implikation (1.3) wahr sein.

Die Umkehrung (oder Konversion) der Aussage $A \Rightarrow B$ ist die Aussage $B \Rightarrow A$. Wir sagen, dass zwei Aussagen **logisch äquivalent** sind wenn sowohl $A \Rightarrow B$ als auch deren Umkehrung, $B \Rightarrow A$, beide wahr sind. Wir schreiben dafür:

$$A \Leftrightarrow B.$$

Das A und B logisch äquivalent sind heißt auch, dass die Wahrheitswerte von A und B übereinstimmen:

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

Bemerkung 1.8. Die Operationen *und* und *oder* sind **kommutativ**. Das heißt, dass für alle Aussagen A und B gilt:

$$(A \text{ und } B) \Leftrightarrow (B \text{ und } A); \qquad (A \text{ oder } B) \Leftrightarrow (B \text{ oder } A).$$

Die Inversion der Implikation $A \Rightarrow B$ ist die Aussage:

$$\text{nicht}(A) \Rightarrow \text{nicht}(B).$$

Die Implikation und deren Inversion werden in der Umgangssprache manchmal als äquivalent gesehen:

Wenn du nicht aufräumst, dann bekommst du kein Eis.

wird mit Sicherheit von einem Kind als "Wenn ich aufräume, dann bekomme ich Eis!" interpretiert. Die kalte Logik sagt aber nichts über was nach dem Aufräumen passiert. Das heißt, als mathematische

[5] hier steckt eigentlich mehr als eine einfache Implikation, aber wir konzentrieren uns nur auf dem $A \Rightarrow B$ Teil.

Aussagen sind diese **nicht logisch äquivalent**. Allgemein gibt es keinen Zusammenhang zwischen der Wahrheit der beiden. Zum Beispiel, die wahre Aussage:

Wenn die Zahl a größer als oder gleich mit 20 ist, dann ist a größer als oder gleich mit 10. (1.4)

hat die Inversion: *Wenn die Zahl a kleiner als 20 ist, dann ist a kleiner als 10* und diese ist falsch. Die Inversion folgender wahren Implikation ist aber wahr:

Wenn die Zahl a gerade ist, dann ist auch a^2 gerade.

Die Kontraposition der Implikation $A \Rightarrow B$ ist die Implikation "nicht(B)⇒nicht(A)". Zum Beispiel, die Kontraposition der obigen Implikation (1.4) ist:

Wenn a nicht größer als 10 ist, dann ist a nicht größer als 20.

Bemerkung 1.9. Eine Implikation ist zu ihrer Kontraposition logisch äquivalent.

Beweis-Skizze:

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

Q.E.D.

Diese Bemerkung ist sehr wichtig und wird uns eine Methode für mathematische Beweise liefern.

1.1.5 Notwendig und hinreichend

Eine **notwendige Bedingung** für die Aussage B ist eine Aussage A, die **gelten muss**, damit die Aussage B wahr sein kann. Eine Garantie für die Wahrheit von B ist sie aber nicht. Anders gesagt:

$$\big(A \text{ ist notwendig für } B\big) \quad \Longleftrightarrow \quad \big(B \Rightarrow A\big)$$

Zum Beispiel, für $a, b \in \mathbb{N}_{>0}$ haben wir

$$a \leqslant b \;\text{ ist notwendig für }\; a \,|\, b.$$

Eine **hinreichende Bedingung** für die Aussage B ist eine Aussage A, die **garantiert**, dass die Aussage B wahr ist. Allerdings kann B wahr sein auch wenn A nicht wahr ist. Anders gesagt:

$$\big(A \text{ ist hinreichend für } B\big) \quad \Longleftrightarrow \quad \big(A \Rightarrow B\big)$$

Zum Beispiel, für $a, b, c \in \mathbb{R}$ haben wir

"a, b, c sind die Seitenlängen eines Dreiecks" ist hinreichend für "$a + b \geqslant c$".

1.1.6 Quantoren

Der Ausdruck "für alle" wird als der **Allquantor** bezeichnet und wird als $\forall$ geschrieben. Zum Beispiel:

$$\forall\, a \in \mathbb{N} \textit{ ist } 2a+1 \textit{ eine ungerade Zahl}.$$

$$\forall\, a \in \mathbb{N} \textit{ gilt } 2a+3=5.$$

Um die Wahrheit einer Aussage, die mit $\forall\, x \ldots$ anfängt, zu überprüfen, muss man sich vorstellen dass man ein beliebiges Element x gegeben hat (und es nicht selber wählt!) und damit umgehen muss. Dieses x sollte mit jedem anderen Element mit der gegebenen Eigenschaft ersetzbar sein ohne unserer Argumentation zu schaden.

Der Ausdruck "es gibt" wird als der **Existenzquantor** bezeichnet und wird als $\exists$ geschrieben. Zum Beispiel:

$$\exists\, a \in \mathbb{N} \textit{ mit } 2a+1 \textit{ gerade}.$$

$$\exists\, a \in \mathbb{N},\ \textit{sodass } 2a+3=5.$$

Um die Wahrheit einer Aussage der Form "$\exists\, x$sodass $A(x)$ wahr ist" direkt zu beweisen, darf man sich selber das Element aussuchen oder erfinden, sodass $A(x)$ gilt.
Die Quantoren $\forall$ und $\exists$ kommutieren nicht! Zum Beispiel:

$$\forall\, a \in \mathbb{N}, \exists\, b \in \mathbb{N},\ \textit{sodass}\ \ b > a \textit{ gilt},$$

sagt uns, dass es keine größte natürliche Zahl gibt. Aber:

$$\exists\, b \in \mathbb{N},\ \textit{sodass}\ , \forall\, a \in \mathbb{N}\ \ b > a \textit{ gilt},$$

sagt uns, dass es eine größte natürliche Zahl gibt: b. Das ist falsch.

Die Negation einer Aussage die Quantoren enthält erfolgt nach dem Prinzip:

$$\begin{aligned} \text{nicht}\,(\forall\, x \textit{ gilt } A(x)) \quad &\Leftrightarrow \quad \exists\, x \textit{ sodass } \text{nicht}(A(x)) \\ \text{nicht}\,(\exists\, x \textit{ sodass } A(x)) \quad &\Leftrightarrow \quad \forall\, x \textit{ gilt } \text{nicht}(A(x)) \end{aligned}$$

Also die Negation von *Alle Katzen sind grau.* ist nicht *Alle Katzen sind nicht grau*, sondern

Es gibt eine Katze die nicht grau ist.

Bezeichnung. Wir werden oft neben der Existenz auch die Eindeutigkeit verlangen. Das drückt man als *es existiert und es ist eindeutig* aus und man bezeichnet es mit $\exists!$. Die Aussage "$\exists! x$ sodass $A(x)$ wahr ist" bedeutet:

$$\exists\, x \text{ sodass } A(x) \ \text{ und } \ ((A(x) \ \text{ und } A(y)) \Rightarrow x = y)\,.$$

1.1.7 Variablen, Aussagenformen, Tautologien, Kontradiktionen

Um den Begriff von Tautologie und Kontradiktion besser definieren zu können, müssen wir über logische Variablen und Ausdrücke sprechen. Eine Analogie für das, was wir vorhaben, ist die Einführung des Rechnens mit Buchstaben in der Schulmathematik. Zum Beispiel, Sie kennen bestimmt die Formel $a^2 - b^2 = (a - b)(a + b)$, oder allgemeiner:

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}) \qquad \forall\, n \geqslant 3.$$

Am Anfang verstehen wir a und b als Bezeichnung für irgendwelche beliebige reelle Zahlen. Irgendwann ist dann x nicht mehr als Zahl, sondern als abstraktes Symbol aufgetreten:

$$x^2 + 2x + 1.$$

Hier steht x nicht mehr für eine Zahl; man kann an der Stelle von x eine Zahl einsetzen, aber x selbst ist keine. Eine Variable ist ein selbständiges mathematisches Objekt, dass man ähnlich wie Zahlen manipulieren kann. Wir werden das in dieser Vorlesung genauer behandeln, wenn wir den Polynomring in einer Variable mit Hilfe von direkten Summen definieren werden.

Was wir jetzt machen wollen, ist die Einführung von **Aussagenvariablen** auf ähnliche Weise. Das bedeutet, abstrakte logische Objekte einzuführen, die genauso manipuliert werden können wie Aussagen (also durch Negation, Konjunktion, Disjunktion, Implikation). Diese sind selbst keine Aussagen sind, aber können durch konkrete Aussagen ersetzt werden. Ausdrücke, die durch solche Manipulationen von logischen Variablen erhalten werden, nennen wir **Aussagenformen** oder **logische Ausdrücke**. Um dies präziser zu machen, sagen wir, dass Aussagenformen folgende Regeln erfüllen müssen:

1. Aussagenvariablen sind Aussagenformen.
2. Wenn X eine Aussagenform ist, dann ist auch $\text{nicht}(X)$ eine Aussagenform.
3. Wenn X und Y Aussagenformen sind, dann sind auch

   $$X \text{ und } Y, \qquad X \text{ oder } Y, \qquad X \Rightarrow Y, \qquad X \Leftrightarrow Y$$

   Aussagenformen.

Zum Beispiel, wenn A, B, C, D Aussagenariablen sind, dann sind folgende Ausdrücke Aussagenformen:

$$A \Rightarrow B, \qquad (A \text{ und } B) \Rightarrow C, \qquad (A \text{ oder } B) \Rightarrow (C \text{ und } \text{nicht}(D)).$$

Die Regeln sind so gemeint, dass eine Aussagenform durch *endlich* viele Operationen auf Aussagenvariablen erhalten wird.

Eine Aussagenform ist **allgemein gültig** wenn sie für jedes mögliche Einsetzen der Aussagenvariablen stets *wahr* ergibt. Eine **Tautologie** ist eine Aussage, die durch Einsetzen in eine allgemein gültige Aussagenform entsteht. Zum Beispiel

$$A \text{ oder } \text{nicht}(A), \quad A \Rightarrow A$$

sind allgemein gültig, also ist "$x > 0$ *oder* $x \leqslant 0$" eine Tautologie. Es gibt aber auch weniger offensichtliche Tautologien, wie zum Beispiel die, die aus den folgenden Aussagenformen entstehen.

$$(A \text{ und } (A \Rightarrow B) \text{ und } (B \Rightarrow C)) \Rightarrow C,$$
$$(A \text{ und } B) \Rightarrow (\text{nicht}(A) \Rightarrow B)\,.$$

Eine **Kontradiktion** (oder ein Widerspruch) ist eine Aussage, die durch Einsetzten in der Negation einer allgemein gültiger Aussage entsteht. Das heißt, es ist eine Aussage, die aus rein logischen Gründen falsch ist. Zwei Kontradiktionen, die häufig in *Beweise durch Widerspruch* vorkommen, entstehen durch Einsetzen in "A und nicht(A)" oder Einsetzen in "$A \Rightarrow$ nicht(A)".

[3] 23.10.'23

1.2 Mengen und Abbildungen

Ein Schema, das wir öfter in diesem Kapitel sehen werden, ist, dass bald nach der Einführung einer Kategorie von Objekten (in diesem ersten Fall werden es Mengen sein) folgt die Definition von "erlaubten Korrespondenzen" zwischen solchen Objekten (in diesem Fall werden es Abbildungen sein). Diese werden durch Pfeile von einem Objekt zur anderen Bezeichnet. Dieses Muster werden wir auch für Gruppen, Ringe, Körper und Vektorräume folgen. Solche Strukturen, die aus Objekten und Pfeile zwischen Objekten bestehen (und auch noch gewisse Axiome, die wir hier nicht nennen, erfüllen) heißen *Kategorien*. Zu der Definition von Kategorie kommen wir noch später vielleicht. Aber jetzt geht es los mit den elementarsten mathematischen Strukturen: Mengen.

1.2.1 Mengen

Wir geben hier Georg Cantors[6] Definition für Menge. Die saubere und genauere Definition ist wesentlich komplizierter. Cantors Definition ist einerseits ungenau, weil sie auf Intuition basiert, anderseits ist sie suggestiv und für unsere Zwecke ausreichend.

Definition 1.10 (Cantor 1895)**.** *Unter einer ‚**Menge**' verstehen wir jede* **wohldefinierte** *Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die ‚Elemente' von M genannt werden) zu einem Ganzen. In Zeichen drücken wir dies so aus:*

$$M = \{m\}.$$

Die Objekte in der Zusammenfassung werden **Elemente** der Menge genannt. Wenn x ein Element der Menge M ist, dann schreiben wir $x \in M$ und lesen das x *ist ein Element von* M oder x *ist in* M. Wenn x nicht ein Element der Menge M ist, dann schreiben wir $x \notin M$. Es existiert eine einzige Menge ohne Elemente; diese heißt die **leere Menge** und wird mit $\emptyset$ bezeichnet.

Wir haben alle Probleme die vorkommen könnten unter "wohldefiniert" versteckt. Ein wichtiger Punkt wäre, dass eine wohldefinierte Menge sich nicht selber als Element enthalten kann. Das führt zu Russels Paradox (siehe weiter unten). Um die Grundlagen der Theorie der Mengen ganz sauber aufzubauen, braucht man überraschend viel Zeit. Damit beschäftigt sich die Mengenlehre ([Rau08]). Ich werde hier nur bemerken, dass *die Mengenlehre verschafft sich ihre Mengen nicht aus de physikalischen Realität, sondern mittels eigens postulierter Existenzprinzipien.*[7] Hier ist Beispiel eines solches Axioms:

$$\exists \emptyset \forall\, yy \notin \emptyset$$

Das sagt uns, dass die leere Menge existiert. In menschlicher Sprache sagt das Axiom, dass es eine Menge $\emptyset$ existiert, für welche die Aussage $y \notin \emptyset$ wahr für alle möglichen y ist. Anders gesagt, für kein

[6] (1845-1918) Deutscher Mathematiker, Gründer der Mengenlehre: https://de.wikipedia.org/wiki/Georg_Cantor.
[7] [Rau08]

y ist die Aussage $y \in \emptyset$ wahr. Das heißt, dass $\emptyset$ kein einziges Element enthält.

Eine Menge kann man beschreiben/definieren indem man:

1. die Elemente zwischen Kommas auflistet.
   (a) $\{1, 2, 3, 4, 5\}$.
   (b) $\{Haus, Hund, Katze, \clubsuit, \{1, 2, 3\}, 100, 0\%* :)\}$.
   (c) $\mathbb{N} = \{0, 1, 2, 3, \dots\}$. In dieser Vorlesung werden wir immer annehmen, dass $0 \in \mathbb{N}$.
   (d) $\mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$.
   (e) $X = \{1, 2, 4, 6, 10, 12, 16, 18, 22, 28, \dots\}$ - **Übung:** wie geht es hier weiter?
2. eine Eigenschaft angibt.

   $$\{x \ : \ \text{hat die Eigenschaft } E\} \text{ oder } \{x \in M \mid x \text{ hat die Eigenschaft } E\}.$$

   In diesem Fall liest man " | " oder " : " als *sodass* oder *für die gilt.*
   (a) $\emptyset = \{x \mid x \neq x\}$.
   (b) $\mathbb{R} = \{x \mid x \text{ ist eine reelle Zahl}\}$.
   (c) $2\mathbb{Z} = \{x \in \mathbb{Z} \mid x \text{ ist durch 2 teilbar}\}$.
   (d) **Achtung:** $\{x \mid 0 \leqslant x \leqslant 1\}$ ist unklar. Mögliche Aufklärungen sind:
      (i) $\{x \in \mathbb{Z} \mid 0 \leqslant x \leqslant 1\}$.
      (ii) $\{x \in \mathbb{R} \mid 0 \leqslant x \leqslant 1\}$.
      (iii) $\{x \in \mathbb{Q} \mid 0 \leqslant x \leqslant 1\}$.
3. eine Formel (oder einen Ausdruck) auf alle/einige Elemente einer Menge anwendet.

   $$\{\text{Formel/Ausdruck } F(x) \mid x, y, \dots \in M \text{ und } x, y, \dots \text{ haben die Eigenschaft } E\}.$$

   (a) $2\mathbb{Z} = \{2 \cdot x \mid x \in \mathbb{Z}\}$.
   (b) $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z} \text{ und } q \neq 0\}$.
   (c) $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R} \text{ mit } i^2 = -1\}$.

Ein Element gehört einer Menge genau dann, wenn es gleich mit einem aufgelisteten Element ist, wenn es die Eigenschaft hat, beziehungsweise wenn es durch die gegebene Formel erhalten wird.

Beispiele:

1. $3 \in \{1, 2, 3, 4, 5\}$ aber $6 \notin \{1, 2, 3, 4, 5\}$.
2. $2 \in \{2\}$ aber $2 \notin \{\{2\}\}$.
3. $2^{2^2} \in 2\mathbb{Z}$, $\frac{1}{2} \in \{x \in \mathbb{Q} \ : \ 0 \leqslant x \leqslant 1\}$, aber $\frac{1}{2} \in \{x \in \mathbb{Z} \ : \ 0 \leqslant x \leqslant 1\}$.
4. $123456^{101} \in 2\mathbb{Z}$ weil $123456^{101} = 2 \cdot (2^{100} \cdot 61728^{101})$.

5. $32749^2 \notin 71\mathbb{Z}$ weil $71 \not| 32749^2$.

6. $\sqrt{101} \notin \mathbb{Q}$ weil es keine $a, b \in \mathbb{Z}$ gibt, sodass $\frac{a}{b} = \sqrt{101}$. (Das braucht aber einen Beweis).

"Die nächste Definition und die folgende Bemerkung sind offensichtlich, aber sie werden sehr häufig angewendet. Deshalb formulieren wir sie explizit.

Definition 1.11. Zwei Mengen sind **gleich** (oder **identisch**) wenn sie dieselben Elemente enthalten.

Wenn die Menge M gleich der Menge N ist, schreiben wir $M = N$. Wenn nicht, schreiben wir $M \neq N$. Man sagt auch, dass zwei identische Mengen sind *umfangsgleich*.

Bemerkung 1.12. Um zu zeigen, dass $M = N$ muss man zeigen, dass

$$x \in M \Rightarrow x \in N \quad \textbf{und} \quad x \in N \Rightarrow x \in M.$$

Obwohl es nicht so kompakt ist, finde ich die Formulierung von $x \in M \Rightarrow x \in N$ als

$$\forall\, x \text{ mit } x \in M \text{ gilt } x \in N$$

besser für diese Vorlesung, weil es darauf hinweist wie der Beweis anfangen soll: *Sei $x \in M$ beliebig.*

Beispiele:

1. $\{2, 6, 13\} = \{6, 13, 2\}$.

2. $\{2, 2\} = \{2\} = \{z \in \mathbb{Z} \mid 1 < z < 3\}$.

3. $\{z \in \mathbb{Z} \mid z \text{ ist durch 2 teilbar}\} = \{2z \mid z \in \mathbb{Z}\}$.

4. $\{1, 2\} \neq \{1, 2, 3\}$.

5. $\{2, 3\} \neq \{\{2\}, 3\}$.

6. $\{x \in \mathbb{R} \mid 0 \leqslant x \leqslant 1\} \neq \{x \in \mathbb{Q} \mid 0 \leqslant x \leqslant 1\}$.

Definition 1.13. Eine Menge N heißt **Teilmenge** der Menge M, wenn jedes Element von N auch in M enthalten ist. Wir schreiben dafür $N \subseteq M$. Wenn $N \subseteq M$ und $N \neq M$, dann heißt N eine **echte Teilmenge** von M. Wir schreiben in diesen Fall $N \subsetneq M$ oder $N \subset M$.

Wir haben also:

$$N \subseteq M \quad \Leftrightarrow \quad x \in N \Rightarrow x \in M,$$

$$N \subsetneq M \quad \Leftrightarrow \quad (x \in N \Rightarrow x \in M) \text{ und } \exists\, x \in M, \text{ sodass } x \notin N.$$

Die Bezeichnung $N \subset M$ wird in der Literatur inkonsistent verwendet: manchmal für echte Teilmenge, manchmal an der Stelle von $\subseteq$. Es kann also verwirrend sein. Ich bevorzuge deswegen $N \subseteq M$ und $N \subsetneq M$.

Beispiele:

1. $\{Haus, Hund, \{3,4\}\} \subseteq \{Hund, \{3,4\}, Katze, Haus\}$.
2. $\{3,4\} \not\subseteq \{Hund, \{3,4\}, Katze, Haus\}$.
3. $\{2z \mid z \in \mathbb{Z}\} \subseteq \mathbb{Z}$.
4. $\{2z \mid z \in \mathbb{Z}\} \not\subseteq \mathbb{N}$.
5. $\{1,2,3\} \not\subseteq \{2,3,4\}$.
6. $M \subseteq M$ gilt für alle Mengen M.
7. $\emptyset \subseteq M$ gilt für alle Mengen M.
8. $\emptyset \subsetneq M$ gilt für alle Mengen $M \neq \emptyset$.
9. Wir werden folgende Kette von Inklusionen als bekannt annehmen:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C},$$

   auch wenn ganz pedantisch gesehen, sind ganze Zahlen nicht Brüche $\frac{p}{q}$, sondern werden als Brüche betrachtet durch $z = \frac{z}{1}$. Analoges gilt für die anderen Inklusionen.

10. Sei M eine Menge, dann ist 2^M oder $\mathcal{P}(M) = \{N \ : \ N \subseteq M\}$. Insbesondere: $\emptyset \in 2^M \ \forall M$.

1.2.2 Vorsicht!

Eine essentielle Eigenschaft des axiomatischen Aufbaus der Mathematik ist die Widerspruchsfreiheit: eine Aussage und dessen Negation dürfen nicht gleichzeitig wahr oder gleichzeitig falsch sein. Anders gesagt, es soll keine Aussage A geben, sodass die Aussage "$A \Leftrightarrow \neg A$" wahr ist. Cantors Definition einer Menge kann zu folgender Antinomie führen.

Russels Paradox (1903) (nach dem Britischen Mathematiker Bertrand Russel[8] 1872-1970).
Sei $S = \{M \mid M \text{ ist eine Menge}\}$ die Zusammenfassung aller Mengen. Wäre S selbst eine Menge, dann hätte diese die Eigenschaft

$$\text{(E)} \qquad S \in S.$$

Es gibt aber auch Mengen die (E) nicht erfüllen, z.B. $\emptyset$, oder $\{1,2,3\}$. Man kann also folgende nichtleere Teilmenge von S definieren:

$$T = \{M \in S \ : \ M \notin M\}.$$

Dann haben wir aber, dass $T \in T \Rightarrow T \notin T$ und $T \notin T \Rightarrow T \in T$. Also $T \in T \Leftrightarrow T \notin T$ - und das bricht die Widerspruchsfreiheit.

Hier ist ein weiteres Paradox der naiven Mengenlehre. Wenn Russels Antinomie ein *logisches* Paradox ist, unser nächstes Beispiel ist ein *semantisches* Paradox. Das sollte hinweisen, dass eine präzise Methode die Theorie der Mengen aufzubauen sich nicht auf der Umgangssprache verlassen sollen.

[8] https://de.wikipedia.org/wiki/Bertrand_Russell

Berrys Paradox Sagen wir, dass alle Wörter der Deutschen Sprache in einem Wörterbuch aufgelistet sind. Betrachten wir folgende Menge

$T :=$ *Die Menge der natürlichen Zahlen, die mit weniger als fünfzehn Wörter beschreiben kann.*

Da es nur endlich viele Wörter gibt[9], gibt es auch nur endlich viele Kombinationen von höchstens 15 davon. Das heißt, dass die Menge T endlich ist und somit gibt es auch die kleinste natürliche Zahl mit $k \notin T$. Diese Zahl ist also *die kleinste natürliche Zahl die nicht mit weniger als fünfzehn Wörter beschrieben werden kann.* Wir haben diese aber gerade mit 14 Wörter beschrieben, also $k \in T$.

Es hat lange gedauert bis die Mengenlehre axiomatisch "gefestigt" wurde. Für unsere Zwecke hier reicht es nur zu wissen, dass es so etwas gibt. Wir werden aber weiter mit Cantors Definition arbeiten und als "wohldefinierte" Mengen die Mengen verstehen, die sich selber nicht enthalten. Die Zusammenfassung aller Mengen ist also keine Menge, sondern eine **Klasse**.

1.2.3 Das Kartesische Produkt

Nach Definition 1.11 spielt die Reihenfolge der Elementen einer Menge keine Rolle. Wir werden aber geordnete Auflistungen von Elementen brauchen. Es gibt mehrere Möglichkeiten diese mit Hilfe von bereits eingeführten Begriffe zu definieren. Wir werden hier zwei solche Varianten zeigen. Die erste Version braucht nur den Begriff von Menge, die zweite, die einfacher zu geordnetes n-Tupel verallgemeinerbar ist, braucht auch natürliche Zahlen.

Definition 1.14. (a) **[Kuratowski 1921]** Seien M und N zwei Mengen und seien $m \in M$ und $n \in N$. Das **geordnete Paar** mit erstem Element m und zweitem Element n ist definiert als

$$(m, n) := \{\{m\}, \{m, n\}\}.$$

(b) **[Hausdorff 1914]** Seien 1 und 2 zwei verschiedene Objekte, beispielsweise $\emptyset$ und $\{\emptyset\}$. Seien M_1 und M_2 zwei Mengen und seien $m_1 \in M_1$ und $m_2 \in M_2$. Das **geordnete Paar** mit erstem Element m_1 und zweitem Element m_2 ist die Menge

$$(m_1, m_2) := \{\{1, m_1\}, \{2, m_2\}\}.$$

Bemerkung 1.15. Ein geordnetes Paar ist also eine Zusammenfassung von zwei (nicht unbedingt verschiedene) Objekten, wobei die Reihenfolge eine essentielle Rolle spielt. Insbesondere

$$(m_1, m_2) = (n_1, n_2) \Leftrightarrow m_1 = n_1 \text{ und } m_2 = n_2.$$

Also, falls $m_1 \neq m_2$ haben wir $(m_1, m_2) \neq (m_2, m_1)$.

Definition 1.16. Seien M und N zwei Mengen. Das **kartesische Produkt** (oder die Produktmenge) von M und N ist die Menge aller Paare (m, n) die mit Elementen $m \in M$ und $n \in N$ möglich sind. Wir bezeichnen[10] diese Menge mit $M \times N$:

$$M \times N := \{(m, n) \mid m \in M \text{ und } n \in N\}.$$

Wenn $M = N$ schreiben wir für $M \times M =: M^2$.

Kartesisch heißt es nach dem Philosoph und Mathematiker René Descartes[11] (1596-1650).

[9] Auch wenn man auf Deutsch viele zusammengesetzte Wörter bauen kann, insbesondere um Zahlen zu beschreiben, werden Zahlen die größer als 1 Million nicht mehr durch ein einziges Wort beschrieben.

[10] Hier, in dieser Definition, bezeichnet (m, n) ein geordnetes Paar und nicht ein offenes Intervall.

[11] https://de.wikipedia.org/wiki/Ren%c3%a9_Descartes

Beispiele:

1. $M = \{0, 1\}$ $N = \{1, 2, 3\}$

$$M \times N = \{(0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3)\}.$$

2. Die Punkte der Ebene in der euklidischen Geometrie sind die Elemente von $\mathbb{R}^2$.

3. Die Menge $\mathbb{Z}^2$ ist eine *diskrete*[12] Teilmenge von $\mathbb{R}^2$.

4. Die Teilmenge $\mathbb{Q}^2$ ist im Gegenteil dicht in $\mathbb{R}^2$. Das heißt, dass zu jedem Punkt in $\mathbb{R}^2$ und für jeden Abstand ε gibt es einen Punkt in $\mathbb{Q}^2$ der näher als ε ist.

Wir werden später Tupel allgemeiner definieren (cf. Definition 1.20). Um Abbildungen zu definieren werden wir Tripel brauchen, deswegen definieren wir hier n-Tupel erstmals induktiv. Sei $n \in \mathbb{N}$, mit $n \geqslant 1$. Wenn $n = 1$, dann ist ein 1-Tupel einfach eine Menge die ein einziges Element enthält. Wenn $n = 2$, dann ist ein 2-Tupel ein geordnetes Paar. Für $n \geqslant 3$ ist ein n-Tupel ein geordnetes Paar, welches an der ersten Stelle ein $n-1$-Tupel hat:

$$(a_1, \ldots, a_n) := ((a_1, \ldots, a_{n-1}), a_n).$$

Wicht ist hier wieder, dass n-Tupel durch folgende Bemerkung vollständig charakterisiert sind:

$$(a_1, \ldots, a_n) = (b_1, \ldots, b_n) \iff a_i = b_i \quad \forall\, i = 1, \ldots, n.$$

1.2.4 Abbildungen

Wir fangen mit der intuitiven Definition von Abbildung an. Das einzige Problem ist, dass sich diese Definition zu stark auf der Umgangssprache verlässt.

Es seien A und B zwei Mengen. Eine Abbildung f von A nach B ist eine Vorschrift, die jedem Element $x \in A$ auf eindeutiger Weise ein Element $y \in B$ zuordnet. Wir schreiben dafür $f : A \longrightarrow B$ und bezeichnen für alle $x \in A$ das entsprechende eindeutig bestimmte Element aus B mit $f(x)$:

$$\forall\, x \in A, \exists!\ f(x) \in B.$$

Da "Vorschrift" nicht mathematisch definiert ist, sollten wir versuchen uns auf schon eingeführte Begriffe verlassen. Das heißt wir sollten Abbildungen allein mit Hilfe von Mengen definieren. Man kann das machen, indem man "Vorschrift" als eine Teilmenge der Produktmenge $A \times B$, die eine zusätzliche Eigenschaft erfüllen muss, definiert.

Definition 1.17. Eine **Abbildung** f von einer Menge A in einer Menge B ist ein Tripel $f = (A, B, \Gamma_f)$, wobei Γ_f eine Teilmenge der Produktmenge $A \times B$ ist und folgende Eigenschaft hat:

$$\forall\, x \in A, \quad \exists!\ y \in B \text{ sodass } (x, y) \in \Gamma_f.$$

[12] Es heißt so, weil die Paare $(a, b) \in \mathbb{Z}^2$ einen Mindestabstand halten.

Die Bezeichnung Γ steht für Graph und so heißt auch die Menge Γ_f: der **Graph von f**. Wir werden die Schreibweise $f = (A, B, \Gamma_f)$ sehr selten verwenden. Wir werden stattdessen die natürlichere und suggestive Schreibweise

$$f : A \longrightarrow B$$

für Abbildung von A nach B benutzen. Wir schreiben dafür auch

$$A \xrightarrow{f} B.$$

[4] 25.10.'23

Wir bezeichnen mit $f(x)$ das durch f eindeutig bestimmte y aus B, sodass $(x, y) \in \Gamma_f$. Wir bezeichnen diese Zuordnung auch mit

$$x \longmapsto f(x) \qquad \text{oder} \qquad x \xmapsto{f} f(x)$$

Die Menge A heißt der **Definitionsbereich** von f. Die Menge B heißt der **Wertebereich** von f. Wenn $A' \subseteq A$, ist das **Bild von A' unter f** die Menge

$$f(A') := \{f(x) \mid x \in A'\}.$$

Die Menge $f(A)$ heißt das **Bild** von f. Wenn $B' \subseteq B$ ist das **Urbild** von B' unter f die Menge

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

Wenn $B' = \{y\}$, dann schreiben wir $f^{-1}(y) := f^{-1}(\{y\})$ und nennen diese Menge die **Faser** von f über y.

Obwohl die Definition 1.17 weniger intuitiv ist, hat diese einen wichtigen Vorteil: Es bestimmt ganz genau was es bedeutet, dass zwei Abbildungen gleich sind. Zwei Abbildungen $f = (A, B, \Gamma_f)$ und $g = (C, D, \Gamma_g)$ sind genau dann gleich, wenn

$$A = C, \quad B = D, \quad \text{und } \Gamma_f = \Gamma_g.$$

Alle drei sind Gleichheiten von Mengen und das wurde auch schon eindeutig festgelegt. Wenn wir die Abbildungen als $f : A \longrightarrow B$ und $g : C \longrightarrow D$ schreiben, dann haben wir aus der Definition von geordnetem Paar, dass

$$f = g \quad \Leftrightarrow \quad A = C, \quad B = D, \quad \text{und } f(x) = g(x) \quad \forall\, x \in A.$$

Zum Beispiel, $\lfloor \cdot \rfloor : \mathbb{R} \longrightarrow \mathbb{R}$, definiert durch

$$\lfloor x \rfloor := \max\{z \in \mathbb{Z} \ : \ z \leqslant x\},$$

und $\lfloor \cdot \rfloor : \mathbb{R} \longrightarrow \mathbb{Z}$ (genauso definiert) sind nicht gleich, obwohl beide denselben Effekt auf die reellen Zahlen haben.

Sei $f : A \longrightarrow B$ eine Abbildung und sei $A' \subseteq A$. Eine **Einschränkung** (des Definitionsbereichs) von f (auf A'), ist eine Abbildung $f|_{A'} : A' \longrightarrow B$ definiert durch $f_{A'}(a) := f(a)$ für alle $a \in A'$. Also wenn A' eine echte Teilmenge von A ist, dann ist die Einschränkung nicht gleich mit der ursprünglichen Abbildung. Mann kann auch den Wertebereich einer Abbildung einschränken oder sogar erweitern, aber das geht nur für Mengen B' mit $\text{Bild}(f) \subseteq B'$.

Abbildungen zwischen Mengen gibt es fast immer. Die einzige Ausnahme ist wenn der Definitionsbereich nicht die leere Menge ist, aber der Wertebereich leer ist. Insbesondere, es gibt genau eine Abbildung von der leeren Menge, in jeder anderen Menge: $\varphi : \emptyset \longrightarrow A$, die als Tripel besser beschreibbar ist: $(\emptyset, A, \emptyset)$. Es ist die einzige Möglichkeit, weil $\emptyset \times A = \emptyset$.

Beispiele:

1. Seien $A = B = \{1, 2, 3\}$. Eine der folgenden Teilmengen von A^2 ist nicht eine Abbildung:

$$\begin{aligned} \Gamma_{f_1} &= \{(1,3),(2,3),(3,3)\} \\ \Gamma_{f_2} &= \{(1,1),(1,2),(1,3)\} \\ \Gamma_{f_3} &= \{(1,1),(2,2),(3,3)\} \end{aligned}$$

2. Für jede Menge M kann man die **identische Abbildung** der Menge M (oder Identität von M) definieren:
$$\mathrm{id}_M : M \longrightarrow M, \quad x \longmapsto x.$$

3. Ist folgende Teilmenge von $\mathbb{R}^2$ eine Abbildung?
$$\Gamma_f = \{(x, y) \in \mathbb{R}^2 \mid y^2 = x\}.$$

4. $f : \mathbb{N} \longrightarrow \mathbb{N}$ mit $f(n) = n - 1$ ist keine Abbildung.

5. $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ mit $f(n) = n - 1$ ist eine Abbildung.

6. $f : \mathbb{R} \longrightarrow \mathbb{R}$ mit $f(x) = \frac{1}{x}$ ist keine Abbildung

7. $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} \setminus \{0\}$ mit $f(x) = \frac{1}{x}$ ist eine Abbildung.

8. Für jedes $i \in \{1, 2\}$ können wir die **kanonische Projektion**
$$p_i : A_1 \times A_2 \longrightarrow A_i, \quad (a_1, a_2) \mapsto a_i.$$

9. Wenn $A' \subseteq A$ haben wir eine **kanonische Inklusion/Injektion**
$$i : A' \longrightarrow A, \quad a' \mapsto a'.$$

1.2.5 Operationen mit Mengen

Definition 1.18. Seien M_1, M_2 zwei Teilmengen einer Menge M.
Die **Vereinigung** (oder die Vereinigungsmenge) von M_1 und M_2 ist die Menge, die aus allen Elementen besteht, die in M_1 oder in M_2 (oder in beiden) enthalten sind:

$$M_1 \cup M_2 := \{x \in M \mid x \in M_1 \text{ oder } x \in M_2\}.$$

Der **Durchschnitt** (oder die Schnittmenge) von M_1 und M_2 ist die Menge, die aus allen Elementen besteht die sowohl in M_1 als auch in M_2 enthalten sind:

$$M_1 \cap M_2 := \{x \in M \mid x \in M_1 \text{ und } x \in M_2\}.$$

Die **Differenz** (oder Differenzmenge) von M_1 und M_2 ist die Menge, die aus allen Elementen von M_1 besteht die nicht in M_2 sind:

$$M_1 \setminus M_2 := \{x \in M \mid x \in M_1 \text{ und } x \notin M_2\}.$$

Wenn wir zusätzlich $M_2 \subseteq M_1$ voraussetzen, dann heißt die Differenz $M_1 \setminus M_2$ das **Komplement** von M_2 in M_1 und wird mit $\complement_{M_1} M_2$ bezeichnet. Wenn M_1 klar aus dem Kontext ist, dann schreiben wir nur $\complement M_2$.
Die Mengen M_1 und M_2 heißen **disjunkt** wenn $M_1 \cap M_2 = \emptyset$.
Für jede Menge M definieren wir die **Potenzmenge von M** als die Menge aller Teilmengen von M. Wir bezeichnen das als $\mathcal{P}(M)$ oder 2^M. (Können Sie raten warum man 2^M verwendet?) Also

$$2^M := \{N \ : \ N \subseteq M\}.$$

Insbesondere haben wir immer $\emptyset, M \in 2^M$.

Beispiele:

1. Wenn $M = \{1, 2, 3, 4\}$ und $N = \{2, 4, 6, 8\}$, dann gilt

$$M \cup N = \{1, 2, 3, 4, 6, 8\}, \qquad M \cap N = \{2, 4\}, \qquad M \setminus N = \{1, 3\}.$$

2. Für alle Mengen M gilt

$$M \cup M = M, \qquad M \cap M = M, \qquad M \setminus M = \emptyset.$$

3. Für welche Mengen hat $\mathcal{P}(M)$ ein einziges Element?

1.2.6 Familien von Mengen

Damit wir mit unendlich viele Mengen auf einmal operieren können, brauchen wir *Familien von Mengen* einzuführen. Diese bestehen aus einer *Indexmenge* I, dessen Elemente **Indizes** heißen, aus einer nicht-leerenMenge $\mathcal{M}$, dessen Elementen Mengen sind, und aus einer Abbildung $\mathcal{F} : I \longrightarrow \mathcal{M}$. Das heißt, dass wir jedem Index $i \in I$ genau eine Menge $M_i \in \mathcal{M}$ zuordnen. Wir sprechen dann von einer von I indizierten **Familie** von Mengen und schreiben dafür

$$\mathcal{F} = (M_i)_{i \in I}.$$

Wir können jetzt die Definition von Vereinigungsmenge, Schnittmenge und kartesisches Produkt auf Familien verallgemeinern.

Definition 1.19. Sei I eine Menge und $(M_i)_{i \in I}$ eine Familie von Mengen. Die **Vereinigung** und der **Durchschnitt** der Familie $(M_i)_{i \in I}$ sind die Mengen:

$$\bigcup_{i \in I} M_i \ := \ \{x \ : \ \exists\, i \in I \text{ sodass } x \in M_i\} \quad \text{beziehungsweise}$$

$$\bigcap_{i \in I} M_i \ := \ \{x \ : \ \forall\, i \in I \text{ gilt } \ x \in M_i\}.$$

Das **Kartesische Produkt** der Familie $(M_i)_{i \in I}$ ist die Menge

$$\prod_{i \in I} M_i := \{I \xrightarrow{f} \textstyle\bigcup_{i \in I} M_i \ : \ f(i) \in M_i \ \ \forall\, i \in I\}.$$

Wenn $I = \{1, \ldots, n\}$ mit $n > 0$, dann schreiben wir

$$\bigcup_{i=1}^{n} M_i := M_1 \cup \cdots \cup M_n := \bigcup_{i \in \{1,\ldots,n\}} M_i.$$

Wenn $I = \mathbb{N}$, dann schreiben wir

$$\bigcup_{i \geqslant 0} M_i := \bigcup_{i=0}^{\infty} M_i \quad := \quad M_0 \cup \ldots \cup M_n \cup \ldots := M_0 \cup M_1 \cup \ldots := \bigcup_{i \in \mathbb{N}} M_i.$$

Analoges gilt für den Durchschnitt und für das kartesische Produkt. Weiterhin, wenn $M_i = M$ für alle $i \in I$ und $I = \{1, \ldots, n\}$, dann schreiben wir

$$\prod_{i=1}^{n} M = M^n.$$

Wenn $I = \emptyset$, dann ist nach Definition die *leere Familie* die leere Abbildung von $\emptyset$ nach $\mathcal{M}$. In diesem Fall ist nach Definition 1.19 die Vereinigung der leeren Familie die leere Menge. Das kartesische Produkt der leeren Familie ist dann $\{\emptyset\}$, die Menge mit der leeren Abbildung als einziges Element. Für den Schnitt der leeren Familie ist es ein bisschen komplizierter, weil "$\forall\, i \in \emptyset$ gilt $x \in M_i$" ist für alle x wahr. Also dieser Schnitt muss "alle x" enthalten. Aber was heißt das? Wie wir die Sachen definiert haben, müssen wir die Menge $\mathcal{M}$ aus der Definition von Familie ins Spiel bringen. Dann kann man die identische Abbildung $\mathrm{id} : \mathcal{M} \longrightarrow \mathcal{M}$ als Familie mit Indexmenge $\mathcal{M}$ betrachten. Der Schnitt der leeren Familie ist dann die Vereinigung[13] von $(M)_{M \in \mathcal{M}}$.

Beispiele:

1. Das Komplement der Menge aller ganzen Zahlen in der Menge der reellen Zahlen ist

$$\begin{aligned} \mathbb{R} \setminus \mathbb{Z} &= \cdots \cup (-2,-1) \cup (-1,0) \cup (0,1) \cup (1,2) \ldots \\ &= \bigcup_{a \in \mathbb{Z}} (a, a+1), \end{aligned}$$

   wobei $(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$ das offene Intervall bezeichnet[14].

2. Es gelten auch

$$\bigcap_{i=1}^{\infty} (-\frac{1}{i}, \frac{1}{i}) = \{0\} \qquad \text{und} \qquad \bigcup_{i=1}^{\infty} (-i, i) = \mathbb{R}.$$

Definition 1.20. Für jede Familie von Mengen $(M_i)_{i \in I}$ heißen die Elemente des kartesischen Produktes $\prod_{i \in I} M_i$ Tupeln. Wenn $I = \{1, \ldots, n\}$ mit $n \in \mathbb{N}_{>0}$, dann nennen wir die Elemente der Menge $\prod_{i=1}^{n} M_i$ **n-Tupel**. Statt $f : \{1, \ldots, n\} \longrightarrow \bigcup_{i=1}^{n} M_i$, verwenden wir für n-Tupel die kompakte Schreibweise:

$$(x_1, \ldots, x_n), \qquad \text{wobei } x_i := f(i) \in M_i.$$

[13] In der axiomatischen Mengenlehre, wird die Existenz dieser Vereinigung als Axiom angenommen und das machen wir stillschweigend auch.

[14] Meistens in dieser Vorlesung wird (a,b) ein geordnetes Paar bezeichnen, cf. Definition 1.14.

Aus der obigen Definition folgt, dass

$$(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \quad \Longleftrightarrow \quad x_i = y_i \ \ \forall\, i = 1, \ldots, n^{15}.$$

Für den sorgfältigen Aufbau der Mengenlehre braucht man folgendes Axiom. Obwohl wir hier Mengen nicht axiomatisch eingeführt haben, erwähnen wir dieses Axiom weil wir es später explizit (oder implizit durch andere äquivalente Aussagen) anwenden werden. Insbesondere brauchen wir das um zu Beweisen, dass jeder Vektorraum eine Basis besitzt. Mehr zu diesem Axiom findet man in [Rau08, 2.7].

Das Auswahlaxiom. *Zu jeder Menge $\mathcal{P}$ von nicht-leeren Mengen gibt es eine Funktion f die jedem $X \in \mathcal{P}$ ein Element $f(X) \in X$ zuordnet.*

Bemerkung 1.21. Sei M eine Menge. Das Auswahlaxiom ist äquivalent zur Aussage:
Für jede Familie $(M_i)_{i \in I}$ von nicht-leeren Mengen ist das kartesische Produkt $\prod_{i \in I} M_i$ nicht leer.

Beispiele:

1. Wenn $A = \{0, 1\}$, dann haben wir

$$A^3 = \{(0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), (1,1,1)\}.$$

2. $\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$.

1.2.7 Eigenschaften von und Operationen mit, Abbildungen

Definition 1.22. Seien $A \xrightarrow{f} B \xrightarrow{g} C$ zwei Abbildungen. Die **Komposition**[16] von g mit f ist die Abbildung

$$g \circ f : A \longrightarrow C, \quad x \mapsto g(f(x)).$$

Achtung! Abbildungen kann man nicht immer verknüpfen!

Man kann zwei Abbildungen nur dann verknüpfen wenn der Wertebereich der zweiten Abbildung gleich dem Definitionsbereich der ersten Abbildung ist. Es kann also sein, dass $g \circ f$ existiert, aber $f \circ g$ nicht. Wenn $f : A \longrightarrow B$ und $g : B' \longrightarrow C$, aber Bild $f \subseteq B'$, dann würde $g(f(x))$ trotzdem Sinn machen und somit eine Abbildung von A nach C definieren. Wir bestehen aber darauf, dass die Komposition nur dann definiert ist wenn $B = B'$, was man in diesem Fall machen könnte, ist g auf Bild f einzuschränken und dann die zwei Abbildungen verknüpfen.

Bemerkung 1.23. (a) Die Verknüpfung von Abbildungen ist **assoziativ**. Das heißt, wenn man drei komponierbare Abbildungen $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$ hat, dann gilt

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

(b) Für jede Abbildung $f : A \longrightarrow B$ gilt $f \circ \mathrm{id}_A = \mathrm{id}_B \circ f = f$.

[15] $\forall\, i = 1, \ldots, n$ ist eine Schreibweise für $\forall\, i \in \{1, \ldots, n\}$.
[16] oder die Verknüpfung/Zusammensetzung/ Hintereinanderausführung/ Verkettung/ Hintereinanderschaltung

Beispiele:

1 **Vorsicht!** Abbildungen sind nicht immer miteinander Verknüpfbar. Das gilt nur wenn der Wertebereich der einen gleich mit dem Definitionsbereich der anderen sind. Zum Beispiel, die zwei Projektionen $p_1, p_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sind nicht miteinander Verknüpfbar, weil

$$p_1\big(p_2(x,y)\big) = p_1(y) - \text{ nicht definiert ist.}$$

2 Insbesondere, es kann oft sein, dass $g \circ f$ definiert ist, aber $f \circ g$ nicht. Zum beispiel, wenn $f = p_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ und $g : \mathbb{R} \longrightarrow \mathbb{R}$ mit $g(x) = x^2 \ \forall\, x \in \mathbb{R}$, dann gilt

$$g \circ p_1 : \mathbb{R}^2 \longrightarrow \mathbb{R} \quad \text{mit} \quad (g \circ p_1)(x,y) = x^2 \quad \forall\,(x,y) \in \mathbb{R}^2,$$

Aber $p_1 \circ g$ ist nicht definiert, weil p_1 von einer reellen Zahl (anstatt eines Paares) nicht definiert ist.

3 Auch wenn sowohl $f \circ g$ als auch $g \circ f$ definiert sind, kann es sein, dass $g \circ f \neq f \circ g$. Zum Beispiel, seien $f, g, h : \mathbb{R} \longrightarrow \mathbb{R}$ die drei Abbildungen gegeben durch

$$x \overset{f}{\longmapsto} x^2 \qquad x \overset{g}{\longmapsto} 2x \qquad x \overset{h}{\longmapsto} x - 1.$$

Dann gilt

$$x \overset{f\circ g}{\longmapsto} 4x^2 \qquad x \overset{f\circ g\circ h}{\longmapsto} 4x^2 - 8x + 4$$

$$x \overset{g\circ f}{\longmapsto} 2x^2 \qquad x \overset{h\circ g\circ f}{\longmapsto} 2x^2 - 1.$$

Definition 1.24. Eine Abbildung $f : A \longrightarrow B$ heißt

$$\begin{aligned}
\textbf{injektiv} &\iff \forall\, x, y \in A \text{ mit } x \neq y \Rightarrow f(x) \neq f(y).\\
\textbf{surjektiv} &\iff \forall\, b \in B, \quad \exists x \in A, \text{ sodass } f(x) = b.\\
\textbf{bijektiv} &\iff b \in B, \quad \exists! x \in A, \quad \text{sodass} \quad f(x) = b.\\
\textbf{invertierbar} &\iff \exists \overline{f} : B \longrightarrow A, \quad \text{sodass } f \circ \overline{f} = \mathrm{id}_B \quad \text{und} \quad \overline{f} \circ f = \mathrm{id}_A\,.
\end{aligned}$$

[5] 30.10.'23

Obwohl wir in Satz 1.27 zeigen werden, dass für Abbildungen *bijektiv sein* äquivalent zu *invertierbar sein* ist, macht es doch Sinn die beiden Eigenschaften eigenständig zu betrachten. Ein Grund dafür ist, dass in gewissen geometrischen Umständen, bijektiver Morphismus nicht mehr äquivalent zu invertierbarer Morphismus ist. Ein weiterer Grund ist, dass einerseits die Bijektivität intuitiver und in Beweisen einfacher anzugehen ist, anderseits ist Invertierbarkeit besser geeignet für abstraktere Situationen.

Bemerkung 1.25. Sei $f : A \longrightarrow B$ eine Abbildung. Es gilt

$$\begin{aligned}
f \text{ ist injektiv} &\iff \forall\, b \in B \text{ die Faser } f^{-1}(b) \text{ höchstens ein Element enthält.}\\
&\iff \forall\, x, y \in A \text{ mit } f(x) = f(y) \text{ folgt } x = y.\\
f \text{ ist surjektiv} &\iff \forall\, b \in B \text{ die Faser } f^{-1}(b) \text{ mindestens ein Element enthält.}\\
f \text{ ist bijektiv} &\iff \forall\, b \in B \text{ die Faser } f^{-1}(b) \text{ genau ein Element enthält.}\\
&\iff f \text{ ist injektiv und surjektiv.}
\end{aligned}$$

Beispiele:

1. $f : \mathbb{R} \longrightarrow \mathbb{R}^2, f(x) = (x, 0)$ (oder $(x-1, 2x+1)$ oder (x, x^2)) sind alle 3 injektiv. Man kann sich diese als Einbettungen der reellen Gerade in der reellen Ebene vorstellen.
2. Die Projektionen $p_1, p_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sind surjektiv.
3. Auch in konkreten Fällen, soll man Abbildungen mit dem Ausdruck der diese definiert **nicht** gleichstellen. Man betrachte $\cdot 3 : \mathbb{R} \longrightarrow \mathbb{R}, x \mapsto 3x$, die invertierbar (und bijektiv) mit Inverse gegeben durch $x \mapsto (1/3)x$ ist. Die Abbildung $\cdot 3 : \mathbb{Z} \longrightarrow \mathbb{Z}$ hingegen, die auch durch $x \mapsto 3x$ definiert ist, ist nicht invertierbar und nicht surjektiv.
4. Genauso wie im obigen Beispiel: $f : \mathbb{R} \longrightarrow \mathbb{R}$ definiert durch $f(x) = x^2$ ist nicht injektiv, aber $f|_{\mathbb{N}} : \mathbb{N} \longrightarrow \mathbb{R}$ ist injektiv.
5. $f : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R}$ definiert durch $f(x) = \frac{x}{x-1}$ ist auch injektiv.
6. Es gibt viele gute Beispiele in [Hou12].

Proposition 1.26. *Wenn eine Abbildung $f : A \longrightarrow B$ invertierbar ist, dann ist die Abbildung $\overline{f}$ aus der Definition 1.24 eindeutig bestimmt. Diese Abbildung ist dann die **Inverse** (oder die **Umkehrabbildung**) von f und wird auch mit f^{-1} bezeichnet.*

Beweis-Skizze: Wir können eigentlich eine stärkere Aussage beweisen:
Seien $\overline{f}, \widetilde{f} : B \longrightarrow A$ zwei Abbildungen, sodass $\overline{f} \circ f = \mathrm{id}_A$ und $f \circ \widetilde{f} = \mathrm{id}_B$ gelten. Dann gilt $f = \widetilde{f}$.
Aus $\overline{f} \circ f = \mathrm{id}_A$ folgt durch Verknüpfen mit $\widetilde{f}$ auf der rechten Seite, dass $(\overline{f} \circ f) \circ \widetilde{f} = \mathrm{id}_A \circ \widetilde{f}$. Aus Bemerkung 1.23 folgt also

$$\overline{f} = \overline{f} \circ \mathrm{id}_B = \overline{f} \circ (f \circ \widetilde{f}) = \widetilde{f}.$$

Q.E.D.

Eine Abbildung $\overline{f}$ wie in dem obigen Beweis heißt **Linksinverse** oder **Retraktion** von f. Eine Abbildung $\widetilde{f}$ wie hier oben heißt **Rechtsinverse** oder **Sektion** von f.

Satz 1.27. *Sei $f : A \longrightarrow B$ eine Abbildung zwischen zwei nicht-leeren Mengen. Dann gilt*

(i) *f ist injektiv $\iff$ f hat eine Retraktion.*

(ii) *f ist surjektiv $\iff$ f hat eine Sektion.*

(iii) *f ist bijektiv $\iff$ f ist invertierbar.*

Beweis-Skizze:

(i) $\boxed{\Rightarrow}$ Weil f injektiv ist, existiert für jedes $y \in f(A)$ ein einziges $x \in A$ mit $f(x) = y$. Wir bezeichnen das zu y zugeordnete x mit x_y. Sei $a \in A$ ein beliebiges Element. Wir definieren $\overline{f} : B \longrightarrow A$ durch

$$\overline{f}(y) = \begin{cases} x_y & \text{wenn } y \in f(A) \\ a & \text{wenn } y \notin f(A). \end{cases}$$

Dies ist eine Abbildung und erfüllt $\overline{f} \circ f = \mathrm{id}_A$.

$\boxed{\Leftarrow}$ Sei $\overline{f} : B \longrightarrow A$, sodass $\overline{f} \circ f = \mathrm{id}_A$. Seien $x, y \in A$ mit $f(x) = f(y)$. Dann haben wir

$$x = \mathrm{id}_A(x) = \overline{f}\,(f(x)) = \overline{f}\,(f(y)) = \mathrm{id}_A(y) = y.$$

Also f ist injektiv.

(ii) $\boxed{\Rightarrow}$ Wir wollen eine Abbildung $\widetilde{f} : B \longrightarrow A$, die $f \circ \widetilde{f} = \mathrm{id}_B$ erfüllt, definieren.
Sei $y \in B$ beliebig. Da f surjektiv ist, existiert mindestens ein $x \in A$, sodass $f(x) = y$. Wir wählen so ein x für jedes y und definieren $\widetilde{f}(y) := x$. Jedem $y \in B$ wurde genau ein $x \in A$ zugeordnet; wir haben also eine Abbildung von B in A definiert. Da $\widetilde{f}(y)$ aus der Menge $f^{-1}(x)$ gewählt wurde, haben wir $f \circ \widetilde{f} = \mathrm{id}_B$.
$\boxed{\Leftarrow}$ Sei $\widetilde{f} : B \longrightarrow A$, sodass $f \circ \widetilde{f} = \mathrm{id}_B$. Für jedes $y \in B$ haben wir $\widetilde{f}(y) \in f^{-1}(y)$, weil $f(\widetilde{f}(y)) = \mathrm{id}_B(y) = y$. Also f ist surjektiv.

(iii) $\boxed{\Rightarrow}$ Da f bijektiv ist, folgt aus Teil (i) und (ii) die Existenz von $\overline{f}$ und $\widetilde{f}$. Aus dem Beweis der Proposition 1.26 folgt das $\overline{f} = \widetilde{f} = f^{-1}$.
$\boxed{\Leftarrow}$ Wenn f invertierbar ist, dann existiert $f^{-1} : B \longrightarrow A$ mit

$$f^{-1} \circ f = \mathrm{id}_A \qquad \text{und} \qquad f \circ f^{-1} = \mathrm{id}_B\,.$$

Aus (i) und (ii) folgt, dass f injektiv und surjektiv ist.

Q.E.D.

[6] 1.11.'23

Bemerkung 1.28. Seien $\mathbb{A} \xrightarrow{f} B \xrightarrow{g} C$ zwei Abbildungen.

1. Wenn f und g invertierbar sind, dann ist $g \circ f$ invertierbar mit Inverse $f^{-1} \circ g^{-1}$.
2. Wenn f und g injektiv (bzw. surjektiv) sind, dann ist $g \circ f$ injektiv (bzw. surjektiv).

Die Umkehrung gilt allgemein nicht (d.h. $g \circ f$ - inj./surj./inv. $\not\Rightarrow$ f und g - inj./surj./inv.).

Aus der obigen Bemerkung und Satz 1.27 folgt:

Korollar 1.29. Seien $A \xrightarrow{f} B \xrightarrow{g} C$ zwei Abbildungen.

(i) Wenn f und g bijektiv sind, dann ist auch $g \circ f$ bijektiv.

(ii) Wenn $g \circ f$ injektiv ist, dann ist auch f injektiv

(iii) Wenn $g \circ f$ surjektiv ist, dann ist auch g surjektiv.

Hier ist eine weitere Charakterisierung[17] der Injektivität.

[17] Eine Abbildung mit dieser äquivalenten Eigenschaft heißt *Monomorphismus* und das kann allgemein für beliebige Kategorien und Morphismen definiert werden. Der Satz 1.30 zeigt also, dass in der Kategorie der Mengen *Monomorphismus sein* äquivalent zu *injektiv sein* ist.

Satz 1.30. *Eine Abbildung $f : A \longrightarrow B$ ist injektiv wenn und nur wenn für jede Menge A' und für alle Abbildungen $g, g' : A' \longrightarrow A$ gilt*

$$\text{aus } f \circ g = f \circ g' \text{ folgt } g = g'.$$

Beweis-Skizze: $\boxed{\Rightarrow}$ `Variante 1:` Sei f injektiv und seien $A', g,$ und g' beliebig mit der Eigenschaft, dass $f \circ g = f \circ g'$. Für jedes $x \in A'$ haben wir also $f(g(x)) = f(g'(x))$. Da f injektiv ist, folgt daraus, dass $g(x) = g(x')$. Wir haben also bewiesen, dass

$$g(x) = g'(x) \; \forall \, x \in A'.$$

Das heißt, dass $g = g'$.
`Variante 2:` Aus Satz 1.27 existiert eine Linksinverse $\overline{f}$ von f, also

$$f \circ g = f \circ g' \Rightarrow \overline{f} \circ (f \circ g) = \overline{f} \circ (f \circ g') \Rightarrow (\overline{f} \circ f) \circ g = (\overline{f} \circ f) \circ g' \Rightarrow \mathrm{id}_A \circ g = \mathrm{id}_A \circ g' \Rightarrow g = g'.$$

$\boxed{\Leftarrow}$ Die Voraussetzung ist, dass für jede Menge A' und für alle Abbildungen $g, g' : A' \longrightarrow A$ gilt

$$\text{“aus } f \circ g = f \circ g' \text{ folgt } g = g'.\text{”}$$

Wir wollen zeigen, dass f injektiv ist. Nehmen wir an, dass das Gegenteil wahr ist und zwar $\exists \, x \neq x' \in A$ mit $f(x) = f(x')$. Wir werden jetzt zeigen, dass die Negation der Voraussetzung gilt. Dafür wählen wir die Menge $A' = \{1, 2\}$ und definieren die Abbildungen g und g' durch

$$\Gamma_g = \{(1, x), (2, x')\} \quad \text{und} \quad \Gamma_{g'} = \{(1, x), (2, x)\}.$$

Dann haben wir offensichtlich $g \neq g'$, aber

$$\Gamma_{f \circ g} = \{(1, f(x)), (2, f(x')\} = \{(1, f(x), (2, f(x))\} = \Gamma_{f \circ g'}.$$

Wir haben also gezeigt, dass es A', g, g' gibt mit $f \circ g = f \circ g'$ und $g \neq g'$ - ein Widerspruch $\lightning$.

Q.E.D.

Surjektivität hat eine duale Beschreibung[18].

Satz 1.31. *Eine Abbildung $f : A \longrightarrow B$ ist surjektiv wenn und nur wenn für jede Menge B' und für alle Abbildungen $g, g' : B \longrightarrow B'$ gilt*

$$\text{aus } g \circ f = g' \circ f \text{ folgt } g = g'.$$

Beweis-Skizze: $\boxed{\Rightarrow}$ Aus Satz 1.27 folgt, dass f eine Sektion $s : B \longrightarrow A$ hat. Das heißt, dass

$$f \circ s = \mathrm{id}_B \, .$$

Wenn wir also beide Seiten von $g \circ f = g' \circ f$ mit s verknüpfen und die Assoziativität anwenden, bekommen wir $g = g'$.
$\boxed{\Leftarrow}$ Nehmen wir an, dass f nicht surjektiv ist. Dann existiert ein Element $b' \in B$, das nicht im

[18] Dieses Mal heißen solche Morphismen: *Epimorphismen.*

Bild von f liegt. Wir definieren dann $g, g' : B \longrightarrow \{0,1\}$ durch

$$g(b) = 0,\ \forall b \in B \qquad \text{und} \qquad g'(b) = \begin{cases} 0, & \text{wenn } b \neq b' \\ 1, & \text{sonst .} \end{cases}$$

Dann haben wir $g \circ f = g' \circ f$, aber $g \neq g'$ - ein Widerspruch ↯. Q.E.D.

1.2.8 Die Anzahl von Elementen in einer Menge

Wir werden hier erstmals eine naive Definition für endliche Mengen und deren Kardinalität verwenden. Diese ist intuitiv und freundlich, aber es basiert sehr stark auf der Bedeutung von *endlich* und *Anzahl* in der Umgangssprache. Deswegen werden wir später, in Teil 1.2.9, auch eine mathematisch genaue Definition sehen.

Definition 1.32. Wenn M eine Menge ist, die leer ist oder aus endlich vielen Elementen besteht, dann sagen wir, dass M eine **endliche Menge** ist. Wenn M endlich ist, dann nennt man die Anzahl ihrer Elemente **Kardinalität** (oder **Mächtigkeit**) von M. Man schreibt dafür $|M|$ oder $\#M$.
Wenn eine Menge M nicht endlich ist, dann sagen wir, dass diese eine **unendliche Menge** ist und schreiben dafür $\#M = \infty$.

Beispiele:

1. $|\{1,2,3,4\}| = 4$.
2. $\#\{1,2,\{3,4\}\} = 3$.
3. $\#\mathbb{N} = \infty$.

Man kann diese Definition etwas präziser gestalten, indem man sagt, dass eine Mengen M ist endlich von Kardinalität n, wenn es eine natürliche Zahl n und eine bijektive Abbildung zwischen M und $\{i \in \mathbb{N} \ : \ 1 \leqslant i \leqslant n\}$ gibt. Das Problem mit dieser Definition ist, dass die Eindeutigkeit stillschweigend vorausgesetzt wird und dass man die natürlichen Zahlen voraussetzen muss. Wir beseitigen diese Probleme im nächsten Teil.

1.2.9 Kardinalität

Die Definition *Eine Endliche Menge ist eine Menge die endlich viele Elementen hat*, obwohl sehr intuitiv, klingt "unmathematisch"[19]. Man kann und sollte jedoch weiterhin der Intuition von dem, was "endlich" und "Anzahl" bedeutet, vertrauen. Dennoch ist es wichtig, diese Begriffe gründlich zu definieren. Eine Definition ist "genau" wenn sie nur mit Hilfe von bereits mathematisch definierten Begriffen formuliert ist.

Definition 1.33. Zwei Mengen A und B heißen **gleichmächtig** genau dann, wenn es eine Bijektion von A in B gibt.

[19] Was ich hier meine ist, dass ein Begriff "endlich" scheint durch sich selbst ("endlich viele") definiert zu sein.

Das poetische an diesem Teil ist, dass die Definition von Endlichkeit ist "nicht Undendlichkeit". Das ist so, weil die Definition von unendliche Menge natürlicher[20] ist.

Definition 1.34 (Dedekind Unendlichkeit)**.** Eine Menge M ist **unendlich** genau dann, wenn es eine echte Teilmenge $M' \subsetneq M$ gibt die gleichmächtig mit M ist. Eine Menge ist **endlich** wenn diese nicht unendlich ist.

Man kann unendliche Mengen auch mit Hilfe von der Menge der natürlichen Zahlen definieren: Eine Menge M ist endlich, wenn es eine natürliche Zahl n und eine Bijektion $f : M \longrightarrow \mathbb{N}_{<n}$ gibt, wobei $\mathbb{N}_{<n} := \{a \in \mathbb{N} \ : \ a < n\} = \{0, 1, \ldots, n-1\}$. Da wir die natürlichen Zahlen hier nicht genau definiert haben, habe ich das nicht als Definition gegeben. Diese Aussage ist äquivalent zu Dedekinds Definition und wir werden das als Satz beweisen (cf. Satz 1.43).

Bemerkung 1.35. Eine Menge M ist unendlich wenn und nur wenn es eine injektive Abbildung $f : M \longrightarrow M$ gibt die nicht surjektiv ist. Diese findet man indem man den Wertebereich der Bijektion zwischen M und $M' \subsetneq M$ auf M erweitert.

Die Abbildung $f : \mathbb{N} \longrightarrow \mathbb{N}$, $f(n) = n + 1$ zeigt, dass die Menge der natürlichen Zahlen unendlich ist.

Definition 1.36. Eine Menge ist **abzählbar** wenn sie gleichmächtig zu der Menge der natürlichen Zahlen ist.

Satz 1.37. *Wenn eine Menge M eine unendliche Menge N als Teilmenge hat ($N \subseteq M$) dann ist die Menge M auch unendlich.*

Beweis-Skizze: Weil N unendlich ist, gibt es $f : N \longrightarrow N$ die injektiv aber nicht surjektiv ist. Wir definieren dann $\bar{f} : M \longrightarrow M$ durch

$$\bar{f}(x) = \begin{cases} x, & \text{wenn } x \in M \setminus N \\ f(x), & \text{wenn } x \in N. \end{cases}$$

Dann ist auch $\bar{f}$ injektiv aber nicht surjektiv, also ist M unendlich. Q.E.D.

Korollar 1.38. Jede Teilmenge einer endlichen Menge ist endlich.

Satz 1.39. *Es sei M eine Menge und $2^M = \{N \ : \ N \subseteq M\}$ die Potenzmenge davon. Es gibt keine bijektive Abbildung $f : M \longrightarrow 2^M$.*

Beweis-Skizze: (Siehe `Video` ab Minute `0:43:00`)
Wir nehmen an, es existiert eine bijektive Abbildung $f : M \longrightarrow 2^M$. Wir definieren dann

$$B := \{m \in M \ : \ m \notin f(m)\} \in 2^M.$$

Weil f bijektiv, also insbesondere surjektiv, ist, existiert $b \in M$ mit $f(b) = B$. Wir haben dann folgende Äquivalenzen:

$$b \notin B \iff b \notin f(b) \iff b \in B.$$

[20] Das heißt, sie wird durch die *Existenz* einer gewisser Art von Abbildung gegeben und Existenz ist schöner als Nichtexistenz.

Wir haben somit einen Widerspruch, es kann also keine Bijektion f geben. Q.E.D.

Bemerkung 1.40. Abzählbare Mengen sind die kleinsten unendlichen Mengen.

Satz 1.41. *Die Menge der reellen Zahlen ist nicht abzählbar.*

Beweis-Skizze: (Siehe `Video` ab Minute `0:01:00`) Q.E.D.

Satz 1.42. *Die Menge* $\mathbb{N}_{<n} := \{a \in \mathbb{N} \ : \ a < n\}$ *ist endlich.*

Beweis-Skizze: Wir beweisen das durch vollständige Induktion.

Der Induktionsanfang: $k = 0$. Wir haben $\mathbb{N}_{<0} = \emptyset$ und diese hat keine echte Teilmengen und ist also nach Definition 1.34 endlich.

Extra:
Wir schauen uns auch den Fall $k = 1$ an: $\mathbb{N}_{<1} = \{0\}$. Dann ist $\mathbb{N}_{<1} \times \mathbb{N}_{<1} = \{(0,0)\}$, also die einzige Abbildung $f : \mathbb{N}_{<1} \longrightarrow \mathbb{N}_{<1}$ ist $\mathrm{id}_{N_{<1}}$.

Der Induktionsschritt: $k \Rightarrow k+1$. Sei $k \geqslant 1$ und nehmen wir an, dass die Menge $N_{<k}$ endlich ist. Sei $f : \mathbb{N}_{<k+1} \longrightarrow \mathbb{N}_{<k+1}$ eine injektive Abbildung. Weil $\mathbb{N}_{<k} \subseteq \mathbb{N}_{<k+1}$, dürfen wir die Menge $f(\mathbb{N}_{<k})$ betrachten.

Fall 1: $f(\mathbb{N}_{<k}) \subseteq \mathbb{N}_{<k}$.
Dann ist aus induktiven Voraussetzung $f(\mathbb{N}_{<k}) = \mathbb{N}_{<k}$, also muss auch $f(k) = k$ gelten. Somit ist f surjektiv.

Fall 2: $f(N_{<k}) \not\subseteq N_{<k}$.
Dann existiert $a \in \mathbb{N}_{<k}$ mit* $f(a) = k$. Da f injektiv ist, muss $f(k) \neq k$ sein, also $f(k) = b \in \mathbb{N}_{<k}$. Wir definieren

$$\overline{f}(x) = \begin{cases} f(x) & \text{wenn } x \notin \{a, k\}, \\ b & \text{wenn } x = a, \\ k & \text{wenn } x = k. \end{cases}$$

Man überprüft direkt, dass $\overline{f}$ injektiv ist. Weiterhin gilt $\overline{f}(\mathbb{N}_{<k+1}) = f(\mathbb{N}_{<k+1})$ und $\overline{f}(\mathbb{N}_{<k}) \subseteq \mathbb{N}_{<k}$. Aus Fall 1 folgt, dass $\overline{f}$ surjektiv ist, und also auch f ist surjektiv. Q.E.D.

* Sonst gilt $f(a) < k$ für alle $a \in \mathbb{N}_{<k}$ und das widerspricht der Voraussetzung in Fall 2.

Satz 1.43. *Für jede endliche Menge M gibt es ein eindeutiges $n \in \mathbb{N}$, sodass M gleichmächtig mit $\mathbb{N}_{<n}$ ist.*

Beweis-Skizze: Sei M eine beliebige endliche Menge.

Existenz: Nehmen wir an, dass für alle $n \in \mathbb{N}$, M nicht gleichmächtig zu $\mathbb{N}_{<n}$ ist. Wenn $M = \emptyset$, dann ist $M = \mathbb{N}_{<0}$. Also $M \neq \emptyset$. Wir suchen einen Widerspruch. Zu diesem Ziel werden wir induktiv eine Familie von injektive Abbildungen $(f_n : \mathbb{N}_{<n} \longrightarrow M)_{n \geqslant 1}$ mit der Eigenschaft das $\mathrm{Bild}(f_n) \subsetneq \mathrm{Bild}(f_{n+1})$ konstruieren†.

$n = 1$ Da $M \neq \emptyset$ folgt, dass $\exists\, y_0 \in M$. Wir definieren dann $f_1 : \mathbb{N}_{<1} \longrightarrow M$ durch

$$f_1(0) = y_0.$$

$\boxed{n \Rightarrow n+1}$ Sei $f_n : \mathbb{N}_{<n} \longrightarrow M$ die injektive Abbildung die wir für den induktiven Schritt voraussetzen. Unsere Annahme ist, dass M mit keiner der Mengen $M_{<k}$ gleichmächtig ist. Insbesondere, darf f_n nicht surjektiv sein. Es gibt also $y_{n+1} \in M \setminus \text{Bild}(f_n)$. Wir definieren dann $f_{n+1} : \mathbb{N}_{<n+1} \longrightarrow M$ durch

$$f_{n+1}(a) := \begin{cases} f(a) & \text{wenn } a < n \\ y_{n+1} & \text{wenn } a = n. \end{cases}$$

Die Abbildung f_{n+1} ist wohl definiert, injektiv und $\text{Bild}(f_n) \subsetneq \text{Bild}(f_{n+1})$ weil $y_{n+1} \notin \text{Bild}(f_n)$. Wir haben also eine Teilmenge $Y = \{y_i \ : \ i \in \mathbb{N}\} \subseteq M$ definiert, die gleichmächtig zu $\mathbb{N}$ und somit unendlich ist. Aus Satz 1.37 folgt, dass M unendlich ist – ein Widerspruch ↯.
Es existiert also ein $n \in \mathbb{N}$, sodass M gleichmächtig mit $\mathbb{N}_{<n}$ ist.

Eindeutigkeit: Wenn es zwei verschiedene natürliche Zahlen m und n und zwei bijektive Abbildungen $f : M \longrightarrow \mathbb{N}_{<n}$ und $g : M \longrightarrow \mathbb{N}_{<m}$ gibt, dann können wir ohne die Allgemeinheit zu verlieren annehmen, dass $m < n$. Dann ist $\mathbb{N}_{<m} \subsetneq \mathbb{N}_{<n}$ und $g \circ f^{-1} : \mathbb{N}_{<n} \longrightarrow \mathbb{N}_{<m}$ eine bijektive Abbildung von $\mathbb{N}_{<n}$ in der echten Teilmenge $\mathbb{N}_{<m}$ - ein Widerspruch zur Endlichkeit von $\mathbb{N}_{<n}$ (Satz 1.42). Q.E.D.

† "Induktiv konstruieren" heißt zuerst f_1 definieren, dann annehmen, dass f_n definiert wurde und mit Hilfe von f_n eine Abbildung f_{n+1} definieren.

Definition 1.44. Sei M eine endliche Menge. Die **Mächtigkeit** (oder **Kardinalität**) von M ist die eindeutige natürliche Zahl für die M gleichmächtig mit $\mathbb{N}_{<n}$ ist.

Für unendliche Mengen definieren wir die Mächtigkeit einfach als ∞. Die Mächtigkeit von M wird mit $|M|$ oder $\#M$ bezeichnet.

Beispiele:

1. $|\{Hund, 1, M\}| = 3$ aber $|\{\{Hund, 1\}, M\}| = 2$.
2. $|\emptyset| = 0$, $|\{\emptyset\}| = 1$, $|\{\emptyset, \{\emptyset\}\}| = 2$, ...
3. $\#\{z \in \mathbb{Z} \mid x^2 - 5 \leqslant 0\} = 5$.
4. $\#\{z \in \mathbb{Q} \mid x^2 - 5 \leqslant 0\} = \infty$.
5. $|\{z \in \mathbb{R} \mid x^2 + 5 \leqslant 0\}| = 0$, das heißt es ist die leere Menge.
6. Wenn M endlich ist, dann gilt $\#2^M = 2^{\#M}$.
7. Wenn M und N endliche Mengen sind, dann gilt

$$\#\{f : M \longrightarrow N \ : \ f \text{ ist eine Abbildung}\} = \#N^{\#M}.$$

1.2.10 Die universelle Eigenschaft des Kartesischen Produktes

Universellen Eigenschaften sind ein Thema, das über die Ziele dieses Kapitels hinausgeht, daher können Sie diesen Teil gerne überspringen. Allerdings hat dieser alternative Ansatz zur Definition mathematischer Objekte seine Vorteile, daher kann es sich lohnen, einen Blick darauf zu werfen.

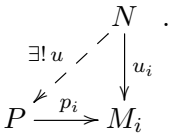
Grob gesagt, die Idee ist, dass bestimmte Konstruktionen so natürlich sind, dass sie vollständig durch eine bestimmte Eigenschaft charakterisiert sind. Mit anderen Worten, die Antwort auf einige Fragen ist kanonisch eindeutig. In solchen Fällen, anstatt ein Objekt durch eine explizite Konstruktion zu definieren, kann man es als das eindeutige Objekt mit der erforderlichen Eigenschaft definieren. Diese Eigenschaft wird oft verwendet und macht in der Regel Beweise natürlicher und schöner.

Universelle Eigenschaften sind meistens als Existenz und Eindeutigkeit gewisser Abbildungen ausgedrückt. Für das kartesische Produkt von Mengen ist die Idee, dass, wenn man für jeden Faktor eine Abbildung $N \longrightarrow M_i$ angibt, man daraus auf eindeutige Weise eine kompatible Abbildung $N \longrightarrow \prod_{i\in I} M_i$ definieren kann. Dies formulieren wir hier als einen Satz über das bereits konstruierte kartesische Produkt aus Definition 1.19.

Satz 1.45 (Universelle Eigenschaft des Kartesischen Produktes)**.** *Sei $(M_i)_{i\in I}$ eine Familie von Mengen, $P := \prod_{i\in I} M_i$ das Kartesische Produkt der Familie und sei $(p_i)_{i\in I}$ die Familie der kanonischen Projektionen*

$$p_j : P \longrightarrow M_j \qquad (m_i)_{i\in I} \overset{p_j}{\longmapsto} m_j, \quad \forall\, j \in I.$$

Für jede Menge N und für jede Familie $(u_i : N \longrightarrow M_i)_{i\in I}$ von Abbildungen existiert eine eindeutige Abbildung $u : N \longrightarrow P$, sodass $p_i \circ u = u_i, \forall\, i \in I$. Das heißt, das folgende Diagramm ist kommutativ für alle i:



Beweis-Skizze: **Existenz:** Man überprüft direkt, dass $u(n) := (u_i(n))_{i\in I}$ für alle $n \in N$ die erwünschte Eigenschaft hat.

Eindeutigkeit: Sei $v : N \longrightarrow P$ mit $p_i \circ v = u_i$ für alle $i \in I$. Wir wollen zeigen, dass $v = u$. Sei $n \in N$ beliebig und bezeichne das Tupel auf dem n abgebildet wird durch

$$v(n) = (v_{i,n})_{i\in I}.$$

Wir haben dann $\forall\, i \in I$, dass $v_{i,n} = p_i((v_{i,n})_{i\in I}) = p_i \circ v(n) = u_i(n)$. Also $u(n) = v(n) \quad \forall\, n \in N$ und somit gilt $u = v$. Q.E.D.

Bemerkung 1.46. Diese Eigenschaft bestimmt das kartesische Produkt im folgenden Sinne:

Sei Q eine Menge und $(q_i : Q \longrightarrow M_i)_{i\in I}$ eine Familie von Abbildungen, die die Eigenschaft aus Satz 1.45 haben. Das heißt, dass:

UE: *Für jede Menge N und für jede Familie $(u_i : N \longrightarrow M_i)_{i\in I}$ von Abbildungen existiert eine eindeutige Abbildung $u : N \longrightarrow Q$, sodass $q_i \circ u = u_i, \quad \forall\, i \in I$.*

Dann existiert eine eindeutige bijektive Abbildung $\varphi : Q \longrightarrow \prod_{i\in I} M_i$, die mit $(q_i)_{i\in I}$ und $(p_i)_{i\in I}$ im folgenden Sinne kompatibel ist:

$$p_i \circ \varphi = q_i \quad \forall\, i \in I.$$

In anderen Worten, Q ist auf eindeutiger Weise in sinnvoller 1-zu-1 Korrespondenz mit dem kartesischen Produkt.

Beweis-Skizze: Wenn **UE**(X) die universelle Eigenschaft für eine Menge X bezeichnet, dann:

$$\text{Aus } \mathbf{UE}(P) \text{ mit } N = Q \text{ und } u_i = q_i \text{ folgt } \exists!\, u : Q \longrightarrow P \text{ mit } p_i \circ u = q_i.$$

$$\text{Aus } \mathbf{UE}(Q) \text{ mit } N = P \text{ und } u_i = p_i \text{ folgt } \exists!\, v : P \longrightarrow Q \text{ mit } q_i \circ v = p_i.$$

Also die Eindeutigkeit und die Existenz gelten mit $\varphi = u$. Was fehlt ist die Bijektivität von u. Wir zeigen dafür, dass $u = v^{-1}$. Wir haben

$$q_i \circ (v \circ u) = (q_i \circ v) \circ u = p_i \circ u = q_i = q_i \circ \mathrm{id}_Q\,.$$

Aus der Eindeutigkeit in **UE**(Q) mit $N = Q$ und $u_i = q_i$ folgt, dass $v \circ u = \mathrm{id}_Q$. Analog folgt auch $u \circ v = \mathrm{id}_p$. Q.E.D.

1.3 Relationen

Definition 1.47. Eine binäre **Relation** auf einer Menge M ist eine Teilmenge $R \subseteq M \times M$.

Wir werden ausschließlich *binäre* Relationen betrachten und deswegen werden wir binär nicht mehr erwähnen. Eine Relation ist also eine Menge geordneter Paare. Wenn $(x, y) \in R$ schreiben wir

$$x \sim_R y \qquad \text{oder} \qquad x \sim y$$

und man sagt, dass x in Relation R zu y steht.

Definition 1.48. Eine Relation heißt

1. **reflexiv** $\iff$ $x \sim x \quad \forall\, x \in M.$
2. **symmetrisch** $\iff$ $x \sim y \Rightarrow y \sim x.$
3. **antisymmetrisch** $\iff$ $x \sim y$ und $y \sim x \Rightarrow x = y.$
4. **transitiv** $\iff$ $x \sim y$ und $y \sim z \Rightarrow x \sim z.$

Die einzige Relation, die die ersten drei Eigenschaften erfüllt ist die Gleichheit. Wir sind an Relationen interessiert, die 1., 2. und 4., oder die 1., 3. und 4. erfüllen. Diese heißen *Äquivalenzrelationen* – die Objekte identifizieren/gleichsetzen, beziehungsweise *Ordnungsrelationen* – die Objekte vergleichen und (manchmal) sagen welches größer/besser/schöner ist.

1.3.1 Äquivalenzrelationen

Definition 1.49. Eine **Äquivalenzrelation** auf einer Menge M ist eine Relation $\sim_R$ die *reflexiv*, *symmetrisch*, und *transitiv* ist.

Wenn $\sim_R$ eine Äquivalenzrelationist und $x \sim_R y$, dann sagen wir, dass x äquivalent zu y (unter R) ist.

Definition 1.50. Sei $\sim$ eine Äquivalenzrelation auf der Menge M. Eine **Äquivalenzklasse** für $\sim$ ist eine Teilmenge von M der Form:

$$[x]_\sim = \{m \in M \;:\; m \sim x\},$$

wobei $x \in M$. In diesem Fall heißt $[x]_\sim$ **die Äquivalenzklasse von** x.

[7] 6.11.'23

Wir bezeichnen Äquivalenzklassen auch mit $\widehat{x}$ oder $\widetilde{x}$ oder ähnliches.

Beispiel 1.51. Die Kongruenz modulo 3 ist die Relation "$\equiv \mod 3$" auf $\mathbb{Z}$ definiert durch

$$a \equiv b \bmod 3 \quad \iff \quad 3 \mid a - b.$$

Wir erinnern, dass $n \mid m$ für "n teilt m" steht und dass $n \in \mathbb{Z}$ teilt $m \in \mathbb{Z}$ wenn es eine ganze Zahl $k \in \mathbb{Z}$ existiert, sodass

$$m = n \cdot k.$$

Es gilt also $3 \mid 0 = a - a$, weil $0 = 3 \cdot 0$, und somit ist $\equiv \mod 3$ reflexiv. Symmetrie und Transitivität sind auch sehr einfach zu überprüfen.

Für die Kongruenz modulo 3 haben wir die Äquivalenzklassen:

$$\begin{aligned}
[0] &= \{3k \ : \ k \in \mathbb{Z}\} = \{\cdots -6, -3, 0, 3, 6, 9, \dots\}, \\
[1] &= \{3k+1 \ : \ k \in \mathbb{Z}\} = \{\cdots -5, -2, 1, 4, 7, 10, \dots\}, \\
[2] &= \{3k+2 \ : \ k \in \mathbb{Z}\} = \{\cdots -4, -1, 2, 5, 8, 11, \dots\}.
\end{aligned}$$

Es gibt keine weitere Äquivalenzklassen weil für alle $k \in \mathbb{Z}$ gilt:

$$[0] = [3k], \qquad [1] = [3k+1], \qquad [2] = [3k+2].$$

Insbesondere: $[0] = [3] = [99] = [123\,456\,789]$.

Beispiele:

1. Das "schärfste" Beispiel von Äquivalenzrelation ist die Gleichheit. Also

   $$R_= := \{(x, x) \mid x \in M\} \subseteq M \times M.$$

   Da wir Reflexivität brauchen, ist diese Relation in allen Äquivalenzrelationen enthalten. Die Menge $R_=$ heißt auch mit *Diagonale* von $M \times M$ und wird mit $\Delta_{M \times M}$ oder einfach Δ bezeichnet.

2. **Kongruenz Modulo** $m \in \mathbb{N}$ auf $\mathbb{Z}$. Diese wird mit $\equiv \mod m$ bezeichnet:

   $$x \equiv y \bmod m \quad \overset{\text{Def}}{\iff} \quad x - y \text{ ist durch } m \text{ teilbar.}$$

   Diese ist eine der wichtigsten Äquivalenzrelation der Algebra und Zahlentheorie.

3. Auf der Mengen der Karten in einem Stapel von 52 Spielkarten können wir folgende Äquivalenzrelation definieren: zwei Karten sind äquivalent, wenn diese dasselbe Zeichen haben. Dann gibt es vier Äquivalenzklassen: ♣, ♠, ♢, ♡.

4. Kongruenz und Ähnlichkeit für Dreiecke in der Euklidischen Ebene sind beide Äquivalenzrelationen.

5. Sei $Z = \mathbb{N}^2 = \{(a, b) \ : \ a, b \in \mathbb{N}\}$. Wir definieren die Äquivalenzrelation auf Z durch

   $$(a, b) \sim (c, d) \quad \iff \quad a + d = b + c.$$

   Die Äquivalenzklassen dieser Relation sind in Bijektion mit den ganzen Zahlen.

6. Sei $Q = \{(a, b) \mid a \in \mathbb{Z} \text{ und } b \in \mathbb{Z} \setminus \{0\}\}$. Eine Äquivalenzrelation auf Q ist
$$(a, b) \sim (c, d) \iff a \cdot d = b \cdot c.$$
Die Äquivalenzklassen dieser Relation sind in Bijektion mit den rationalen Zahlen.

7. Sei $R = \{(x_n)_{n\in\mathbb{N}} \in \mathbb{Q}^{\mathbb{N}} : (x_n)_{n\in\mathbb{N}} \text{ ist eine Cauchy Folge}\}$. Sie finden die Definition von Cauchy Folge hier unten[21]. Auf dieser Menge definieren wir die Äquivalenzrelation
$$(x_n) \sim (y_n) \iff \lim_{n\to\infty}(x_n - y_n) = 0.$$
Die Äquivalenzklassen dieser Relation sind in Bijektion mit den reellen Zahlen.

8. Auf der Menge $\mathbb{R}[x]$, der Polynome mit reellen Koeffizienten in einer Variable x, definieren wir die Äquivalenzrelation
$$f \sim g \iff (x^2 + 1) \,|\, (f - g).$$
Die Äquivalenzklassen dieser Relation sind in Bijektion mit den komplexen Zahlen.

9. Jede Abbildung $f : A \longrightarrow B$ definiert eine Äquivalenzrelation auf A:
$$a \sim_f b \iff f(a) = f(b).$$

10. Auf der Potenzmenge 2^M kann ist Gleichmächtigkeit eine Äquivalenzrelation:
$$A \sim_{gm} B \iff \exists f : A \longrightarrow B \text{ bijektiv.}$$
Ist Gleichmächtigkeit allgemein eine Äquivalenzrelation?

11. Die Relation $\subseteq$ auf 2^M ist nicht eine Äquivalenzrelation, weil diese nicht symmetrisch ist.

12. Die Relation $A \sim B \Leftrightarrow A \cap B = \emptyset$ auf 2^M ist nicht eine Äquivalenzrelation, weil sie nicht reflexiv und nicht transitiv ist.

13. Sei M die Menge aller lebendigen Menschen.
    (a) "$m_1 \sim_{AG} m_2 \Leftrightarrow m_1$ und m_2 denselben Arbeitgeber haben" ist symmetrisch, könnte transitiv sein (z.B. wenn niemand zwei Arbeitgeber hat), ist aber nicht reflexiv (Kinder, Arbeitslose, Rentner).
    (b) "$m_1 \sim_{SA} m_2 \Leftrightarrow m_1$ und m_2 dieselbe Staatsangehörigkeit haben" ist nicht transitiv (Doppelte Staatsangehörigkeit).

Wir haben in Definition 1.49 die Äquivalenzklasse von x als die Menge die *genau* die Elementen aus M enthält, die äquivalent zu x sind. Ich finde folgende Charakterisierung auch sehr suggestiv. Diese könnte die obige Definition ersetzen.

Satz 1.52. *Sei $\sim$ eine Äquivalenzrelation auf der Menge M. Eine Teilmenge $C \subseteq M$ ist genau dann eine Äquivalenzklasse für $\sim$, wenn folgende Axiome erfüllt sind:*

(ÄK 1) $C \neq \emptyset$.

(ÄK 2) *Wenn $a, b \in C$, dann $a \sim b$.*

(ÄK 3) *Für alle $m \in M$ gilt, wenn $\exists\, y \in C$ mit $y \sim m$, dann $m \in C$.*

[21] (x_n) ist eine Cauchy Folge wenn $\forall\, \varepsilon \in \mathbb{Q}_+,\ \exists N \in \mathbb{N},$ sodass $|x_n - x_m| < \varepsilon\ \forall\, m, n > N$.

Beweis-Skizze: $\boxed{\Rightarrow}$ Wenn C eine Äquivalenzklasse ist, dann existiert $x \in M$, sodass

$$C = [x]_\sim = \{m \in M \ : \ m \sim x\}.$$

Dann haben wir:

(ÄK 1): $x \in C$, also $C \neq \emptyset$.

(ÄK 2): Seien $a, b \in C = [x]$. Das heißt, dass $a \sim x$ und $b \sim x$, also wegen der Symmetrie haben wir $a \sim x$ und $x \sim b$. Aus der Transitivität folgt dann $a \sim b$.

(ÄK 3): Sei $m \in M$ und $y \in C = [x]$ mit $y \sim m$. Aus $y \in [x]$ folgt per Definition $x \sim y$. Aus der Transitivität folgt dann $x \sim m$, und somit $m \in [x] = C$.

$\boxed{\Leftarrow}$ Sei $C \subseteq M$ eine Menge die (ÄK 1), (ÄK 2) und (ÄK 3) erfüllt.

Aus (ÄK 1) folgt, dass es $x \in C$ existiert. Wir beweisen, dass $C = [x]$. Wir zeigen dafür die zwei Inklusionen.

Sei $c \in C$ beliebig. Weil $x, c \in C$, folgt aus (ÄK 2), dass $c \sim x$, also, dass $c \in [x]$. Somit haben wir $C \subseteq [x]$ gezeigt.

Sei $y \in [x]$ beliebig. Es gilt also für y, dass es $x \in C$ existiert mit $y \sim x$. Aus (ÄK 3) folgt $y \in C$. Somit haben wir auch $[x] \subseteq C$ bewiesen.

Q.E.D.

Definition 1.53. Ein **Repräsentant** der Äquivalenzklasse C ist ein Element $x \in C$.
Ein **Repräsentantensytem** für die Äquivalenzrelation $\sim$ ist eine Teilmenge $M' \subseteq M$ mit der Eigenschaft, dass $|M' \cap C| = 1$ für jede Äquivalenzklasse C.

Beispiele:

1. Für die Kongruenz modulo 3 auf $\mathbb{Z}$ ist für $i = 0, 1, 2$ jede Zahl der Form $3k + i$ ein Repräsentant der Äquivalenzklasse $[i]$. Es gibt also unendlich viele Repräsentantensysteme. Hier sind drei solche Beispiele:
$$\{0, 1, 2\}, \quad \{-1, 0, 1\}, \quad \{102, -17, 23\}.$$

2. Für die Äquivalenzrelation, die durch das Zeichen auf dem Stapel Spielkarten gegeben ist, hat jede Äquivalenzklasse 13 mögliche Repräsentanten:

   Äquivalenzklasse aller ♣ Karten $= [\mathbf{A}\clubsuit] = [\mathbf{2}\clubsuit] = \cdots = [\mathbf{10}\clubsuit] = [\mathbf{B}\clubsuit] = [\mathbf{D}\clubsuit] = [\mathbf{K}\clubsuit]$.

   Jedes Repräsentantensystem enthält vier Karten, eine von jedem Zeichen.

Satz 1.54. *Sei $\sim$ eine Äquivalenzrelation auf der Menge M.*

(i) *Wenn C_1 und C_2 zwei Äquivalenzklassen sind, dann gilt entweder $C_1 \cap C_2 = \emptyset$ oder $C_1 = C_2$.*

(ii) *Die Menge M ist die Vereinigung aller Äquivalenzklassen bezüglich $\sim$:*

$$M = \bigcup_{C = \text{Äq.Kl}} C.$$

Beweis-Skizze:

(i) Es reicht zu zeigen, dass $C_1 \cap C_2 \neq \emptyset \Rightarrow C_1 = C_2$. Aus der Symmetrie[a] der Aussage reicht es $C_1 \subseteq C_2$ zu zeigen. Wir haben:

$$C_1 \cap C_2 \neq \emptyset \quad \Leftrightarrow \quad \exists\, x \in C_1 \cap C_2 \quad \Leftrightarrow \quad \exists\, x \text{ mit } x \in C_1 \text{ und } x \in C_2.$$

Sei $y \in C_1$ beliebig. Aus (ÄK 2) folgt $x \sim y$. Da $x \in C_2$, folgt aus (ÄK 3), dass $y \in C_2$.

(ii) Äquivalenzklassen sind Teilmengen von M, also auch die Vereinigung aller Äquivalenzklassen ist in M enthalten. Für die andere Inklusion bemerken wir, dass für jedes $m \in M$ gilt wegen der Reflexivität der Äquivalenzrelation, dass $m \in [m]$.

Q.E.D.

[a] das heißt man kann C_1 und C_2 vertauschen.

Eine **Partition** (oder Unterteilung) einer Menge M ist eine Teilmenge $\mathcal{P} \subset 2^M$ mit der Eigenschaft, dass

$$A \cap B = \emptyset \;\; \forall\, A, B \in \mathcal{P} \qquad \text{und} \qquad \bigcup_{A \in \mathcal{P}} A = M.$$

Bemerkung 1.55. Satz 1.54 sagt, dass jede Äquivalenzrelation auf einer Menge eine Partition (in Äquivalenzklassen) definiert. Das heißt, dass jedes Element gehört genau einer Äquivalenzklasse.

Bemerkung 1.56. Die Umkehrung des Satzes 1.54 gilt auch:

Wenn $\mathcal{P}$ eine Partition von M ist, dann ist die Relation $\sim_{\mathcal{P}}$ definiert durch

$$x \sim_{\mathcal{P}} y \iff \exists\; A \in \mathcal{P} \;\text{ mit } x \in A \quad \text{und} \quad y \in A.$$

eine Äquivalenzrelation.

Beweis-Skizze: Reflexivität: Weil $M = \bigcup_{A \in \mathcal{P}} A$, existiert für jedes $m \in M$ ein $A \in \mathcal{P}$, sodass $m \in A$. Also $m \sim m$ für alle $m \in M$.

Symmetrie: Wenn $m \sim n$, dann existiert $A \in \mathcal{P}$ mit $m \in A$ und $n \in A$. Die Konjunktion " und " ist kommutativ, also gilt $n \in A$ und $m \in A$. Das bedeutet per Definition $n \sim m$.

Transitivität: Seien $m, n, p \in M$ mit $m \sim n$ und $n \sim p$. Es existieren also $A, B \in \mathcal{P}$, sodass

$$(m \in A \quad \text{und} \quad n \in A) \qquad \text{und} \qquad (n \in B \quad \text{und} \quad p \in B).$$

Aus der Assoziativität der Konjunktion folgt $n \in A \cap B$ und somit $A \cap B \neq \emptyset$. Weil $\mathcal{P}$ eine Partition ist, folgt dann, dass $A = B$. Somit haben wir $m, p \in A = B$ und per Definition $m \sim p$.

Q.E.D.

Die Faktormenge ist grundlegend für viele mathematische Begriffsbildungen. Man sollte sich genug Zeit nehmen, um folgende Definition gründlich zu verstehen.

Definition 1.57. Sei $\sim$ eine Äquivalenzrelation auf der Menge M. Die **Faktormenge** (oder Quotientenmenge, oder Menge der Äquivalenzklassen) von M durch $\sim$ ist die Menge

$$M/_{\sim} := \{C \;\; : \;\; C \;\text{ ist eine Äquivalenzklasse für } \sim \text{ in } M\}.$$

Die Quotientenabbildung (oder kanonische Projektion, oder kanonische Surjektion) ist die Abbildung die jedes Element in dessen Äquivalenzklasse abbildet:

$$p : M \longrightarrow M/_\sim \qquad x \mapsto [x].$$

Man kann diese Menge als auch $M/_\sim = \{[x] \ : \ x \in M\}$ beschreiben. Die Schreibweise in der Definition 1.57 soll betonen, dass die Elementen von $M/_\sim$ Mengen sind.

Bemerkung 1.58. Für zwei Äquivalenzklassen $[x], [y] \in M/_\sim$ haben wir

$$[x] = [y] \iff x \sim y.$$

Bemerkung 1.59. Man kann jede surjektive Abbildung als Quotientenabbildung für eine bestimmte Äquivalenzrelation sehen. Genauer gesagt: Wenn $q : M \longrightarrow N$ eine surjektive Abbildung ist, und $\sim_q$ die zugeordnete Äquivalenzrelation, dann gibt es eine bijektive Abbildung

$$g : N \longrightarrow M/_{\sim_q}$$

durch $g(y) = [x]$ definiert, wobei $x \in M$ die Eigenschaft $q(x) = y$ hat. So ein Element x gibt es immer, weil q surjektiv ist. Man muss noch überprüfen, dass wir tatsächlich eine Abbildung definiert haben. Das heißt, dass g von der Wahl von x unabhängig ist. Das ist eine Konsequenz des Satzes 1.60.

Die Faktormenge kann man auch durch eine universelle Eigenschaft definieren. wir formulieren diese als Satz hier. Dieser Satz wird oft angewendet um Abbildungen mit einer Quotientenmenge als Definitionsbereich zu definieren.

Satz 1.60. *Sei $\sim$ eine Äquivalenzrelation auf die Menge M. Sei $f : M \longrightarrow A$ eine beliebige Abbildung.*

(i) *Die Zuordnung $[m] \longmapsto f(m)$ definiert genau dann eine Abbildung $\widehat{f} : M/_\sim \longrightarrow A$, wenn*

$$a \sim b \Rightarrow f(a) = f(b).$$

(ii) *Wenn $\widehat{f}$ eine Abbildung ist, dann haben wir*

(a) $\operatorname{Bild} f = \operatorname{Bild} \widehat{f}$. *Insbesondere, $\widehat{f}$ ist genau dann surjektiv, wenn f ist surjektiv.*

(b) *$\widehat{f}$ ist genau dann injektiv, wenn $a \sim b \Leftrightarrow f(a) = f(b)$.*

Vor dem eigentlichen Beweis will ich versuchen die Idee des Argumentes darzustellen. Wir haben durch $[m] \mapsto f(m)$ genau dann eine Abbildung von $M/_\sim$ nach A definiert, wenn jedem Element (also Äquivalenzklasse) $[m] \in M/_\sim$ ein einziges Element aus A zugeordnet wird. Wir haben für jede $[m]$ mindestens eine Element: $f(m) \in A$. Wenn es eindeutig ist, dann heißt es, dass wenn wir einen andere Repräsentanten von $[m]$ wählen, zum Beispiel $m' \in M$ mit $[m] = [m']$, dann sollen wir dasselbe Element zuordnen. Also $f(m')$ sollte gleich mit $f(m)$ sein, wenn $[m] = [m']$.

Beweis-Skizze:

(i) $\boxed{\Rightarrow}$ Wir nehmen an, dass $\widehat{f}([m]) := f(m)$ eine Abbildung definiert. Seien $a, b \in M$ mit $a \sim b$. Dann haben wir $[a] = [b]$. Also $f(a) = \widehat{f}([a]) = \widehat{f}([b]) = f(b)$.

$\boxed{\Leftarrow}$ Um das obige Argument präzise zu machen, schauen wir uns den Graph der Zuordnung an:

$$\Gamma = \{([a], f(a)) \ : \ a \in M\} \subseteq (M/_\sim) \times A$$

Wir wollen überprüfen, dass es der Graph einer Abbildung im Sinne der Definition 1.17 ist.

Das heißt, dass es genau ein geordnetes Paar mit $[a]$ als erstes Element enthält. Für jedes $[a] \in M/_\sim$ haben wir ein Element in Γ, nämlich $([a], f(a))$. Es fehlt also nur, dass

$$[a] = [b] \Rightarrow ([a], f(a)) = ([b], f(b)).$$

Nach Bemerkung 1.58 ist äquivalent zu der Voraussetzung: $a \sim b \Rightarrow f(a) = f(b)$.

(ii) (a) Wir zeigen die Gleichheit der Mengen direkt, durch Äquivalenz:

$$b \in \text{Bild}\, f \iff \exists\, a \in M \text{ mit } f(a) = b \iff \exists\, [a] \in M/_\sim \text{ mit } \widehat{f}([a]) = b \iff b \in \text{Bild}\, \widehat{f}.$$

(b) $\boxed{\Rightarrow}$ Sei $\widehat{f}$ injektiv und seien $a, b \in M$. Wir haben

$$f(a) = f(b) \iff \widehat{f}([a]) = \widehat{f}([b]) \overset{\widehat{f}=\text{inj}}{\iff} [a] = [b] \iff a \sim b.$$

$\boxed{\Leftarrow}$ Seien $[a], [b] \in M/_\sim$ mit $\widehat{f}([a]) = \widehat{f}([b])$. Dann haben wir

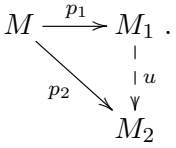
$$\widehat{f}([a]) = \widehat{f}([b]) \overset{\text{Def}}{\iff} f(a) = f(b) \overset{\text{Voraus.}}{\iff} a \sim b \overset{1.58}{\iff} [a] = [b].$$

Also $\widehat{f}$ ist injektiv.

Q.E.D.

Folgendes Korollar sagt, dass diese universelle Eigenschaft die Faktormenge eindeutig bestimmt.

Korollar 1.61. Wenn $p_1 : M \longrightarrow M_1$ und $p_2 : M \longrightarrow M_2$ zwei surjektive Abbildungen mit $\sim_{p_1} = \sim_{p_2}$ sind, dann gibt es eine *bijektive* Abbildung $u : M_1 \longrightarrow M_2$, sodass folgendes Diagramm kommutativ ist:



1.3.2 Ordnungsrelationen

Definition 1.62. Eine **Ordnungsrelation** auf der Menge M ist eine Relation $\preceq$ auf M die *reflexiv*, *antisymmetrisch*, und *transitiv* ist.

Bemerkung 1.63. In der Definition von Ordnungsrelation wird nicht verlangt, dass man alle Elemente miteinander vergleichbar sind. Die Teilbarkeit auf der Menge der natürlichen Zahlen ist ein gutes Beispiel dafür: Primzahlen sind unvergleichbar.

Folgende Bezeichnungen werden auftreten. Eine partiell **geordnete Menge** ist ein geordnetes Paar $(M, \preceq)$, wobei $\preceq$ eine Ordnungsrelation auf M ist. Manchmal wird "partiell" ausgelassen. Zwei Elemente $x, y \in M$ sind **vergleichbar** wenn $x \preceq y$ oder $y \preceq x$. Zwei Elemente sind **unvergleichbar** wenn diese nicht vergleichbar sind. Eine Ordnungsrelation heißt **total** wenn jede zwei Elemente vergleichbar sind. Eine **Wohlordnung** ist eine Ordnungsrelation für welche in jeder nicht leeren Teilmenge $M' \subseteq M$ ein minimales Element existiert, nämlich ein $x \in M'$, sodass $x \preceq y \;\; \forall\, y \in M'$. Wir sagen in diesem Fall, dass $(M, \preceq)$ **wohl geordnet** ist.

Beispiele:

1. Die "natürliche" Ordnung auf $\mathbb{N} = \{0, 1, 2, \dots\}$, definiert durch

$$a \leqslant b \iff \exists k \in \mathbb{N} \text{ sodass } b = a + k,$$

ist eine Ordnungsrelation. Eine ganz wichtige Eigenschaft dieser Relation ist, dass es eine Wohlordnung ist. Wir werden das hier beweisen.

Beweis-Skizze: Erstmals zeigen wir, dass es tatsächlich eine Ordnungsrelation ist.
Reflexivität $a = a + 0, \quad \forall\, a \in \mathbb{N}$ also $a \leqslant a$.
Antisymmetrie $a \leqslant b \Leftrightarrow b = a + k$ und $b \leqslant a \Leftrightarrow a = b + k'$. Also, wenn wir einsetzen, dann bekommen wir $a = a + k + k' \;\Rightarrow\; k = k' = 0 \;\Rightarrow\; a = b$.
Transitivität $a \leqslant b \leqslant c \Rightarrow b = a + k$ und $c = b + h \Rightarrow c = a + (k + h) \Rightarrow a \leqslant c$.

Bemerkung: Wir haben $0 \leqslant n \;\forall\, n \in \mathbb{N}$, weil $n = 0 + n$.

Wir zeigen jetzt, dass es eine Wohlordnung ist. Sei also $\emptyset \neq A \subseteq \mathbb{N}$ beliebig. Wir wollen zeigen, dass $\min A$ existiert (i.e. $\exists m \in A$ mit $m \leqslant a, \forall\, a \in A$)
Fall 1: $0 \in A$**.** Dann ist $\min A = 0$.
Fall 2: $0 \notin A$**.** Wir nehmen an, dass es $\min A$ nicht gibt, und suchen einen Widerspruch. Dieser wird $A = \emptyset$ sein.
Sei $B := \mathbb{N} \setminus A$. Wir beweisen durch vollständige Induktion, dass $\forall\, n \in \mathbb{N}$ gilt $\{0, \dots, n\} \subseteq B$. Daraus folgt dass $B = \mathbb{N}$, und somit unser Widerspruch: $A = \emptyset$.
Induktionsanfang $0 \in B$, weil wir im Fall $0 \notin A$ sind.
Induktionsschritt Die induktive Voraussetzung ist $\{0, \dots, n\} \subseteq B$. Wir wollen zeigen, dass $n + 1 \in B$. Aus der induktiven Voraussetzung folgt

$$n < a \quad \forall\, a \in A.$$

Also $n + 1 \leqslant a \quad \forall\, a \in A$. Es also eine unter Schranke für A. Wäre also $n + 1$ ein Element von A, dann wäre es auch das Minimum. Wir haben aber angenommen, dass $\min A$ nicht existiert, also $n + 1 \notin A$ und somit gilt $n + 1 \in B = \mathbb{N} \setminus A$.

Q.E.D.

2. $(\mathbb{N}, <)$ ist keine Ordnungsrelation, weil diese nicht reflexiv ist.

3. $(\mathbb{Z}, \leqslant)$ ist geordnet durch $a \leqslant b \Leftrightarrow b - a \in \mathbb{N}$. Das ist aber keine Wohlordnung, weil $\mathbb{Z}$ selbst kein minimales Element hat.

4. Eine andere Ordnungsrelation auf $\mathbb{Z}$ ist $a \preceq b \Leftrightarrow |a| < |b|$ oder $(|a| = |b|$ und $a \leqslant b)$.

5. Die Mengeninklusion sollte eine Ordnungsrelation sein. Wir können aber nicht sagen, dass es eine Ordnungsrelation ist, ohne eine Menge von Teilmengen zu erwähnen. Eine Relation braucht eine Menge, und die "Menge aller Mengen" existiert nicht. Also, für jede Menge M ist $(2^M, \subseteq)$ eine partiell geordnete Menge.

6. $(\mathbb{N}, |)$ wobei $|$ die Teilbarkeit der ganzen Zahlen ist auch eine partiell geordnete Menge (cf. Definition 1.65).

7. Die Teilbarkeit auf der Menge $\mathbb{Z}$ ist nicht antisymmetrisch: $a|-a$ und $-a|a$, aber $a \neq -a$ wenn $a \neq 0$.

Später in dieser Vorlesung (und nicht nur) werden wir folgendes Lemma brauchen. Mann kann zeigen, dass dieses Lemma zum Auswahlaxiom äquivalent ist und weiterhin zu mehrere Aussagen über geordnete Mengen (das Lemma von Tukey, Zermelo's Wohlordnungssatz). Die Beweise sind aber für eine "Anfängervorlesung" nicht ganz einfach. Eine Beweisskizze findet man in [Bri85, S.260]. Wir werden hier das Lemma ohne Beweis formulieren. Aber zu erst brauchen wir noch einige einfache Definitionen.

Sei $(M, \preceq)$ eine (partiell) geordnete Menge und $N \subseteq M$ eine Teilmenge. Ein **maximales Element** von N ist ein Element $n \in \mathbb{N}$ das nicht "kleiner" als ein anderes Element von N ist. Genauer formuliert heißt das:

$$n = \text{ maximales Element von } N \iff (n' \in N \text{ und } n \leqslant n') \Rightarrow n = n'.$$

Es ist wichtig zu bemerken, dass maximale Elemente nicht unbedingt eindeutig sind. Eine obere **Schranke** von N ist ein Element $x \in M$ mit

$$n \preceq x \quad \forall\, n \in N.$$

Schranken sind nicht unbedingt eindeutig und müssen nicht in N. Ein **Maximum** von N ist eine obere Schranke m von N, die zusätzlich $m \in N$ erfüllt. Wenn so ein Element existiert, kann man zeigen, dass es eindeutig ist (Übung). Man bezeichnet das mit $\max N$. Analog definiert man minimales Element, untere Schranke und Minimum.

Lemma 1.64 (Zorn)**.** *Sei $(M, \preceq)$ eine geordnete Menge. Wenn jede total geordnete Teilmenge $N \subseteq M$ eine obere Schranke hat, dann existiert ein maximales Element in M.*

1.4 Teilbarkeit

Definition 1.65. Für $a, b \in \mathbb{Z}$ sagen wir, dass a **teilt** b (oder b ist durch a teilbar) und schreiben $a|b$ (oder $b \vdots a$) genau dann wenn $\exists c \in \mathbb{Z}$, sodass $b = ac$. Wenn $a|b$, dann ist a ein **Teiler** von b und b ist ein **Vielfaches** von a.

Bemerkung 1.66. Teilbarkeit ist eine partielle Ordnungsrelation auf $\mathbb{N}$.

Beweis-Skizze: Reflexivität: $a = 1 \cdot a \quad \forall\, a \in \mathbb{N}$.
Antisymmetrie: $(a|b \text{ und } b|a) \Leftrightarrow (\exists c, c' \in \mathbb{Z}, \text{sodass } b = ac \text{ und } a = c'b)$. Wir bekommen dann $a = cc'a$, also $a(1 - cc') = 0$. $\overset{a\neq 0}{\Longrightarrow} 1 = cc' \overset{c,c'\in\mathbb{N}}{\Longrightarrow} c = c' = 1 \Rightarrow a = b$.
Transitivität: $(a|b \text{ und } b|c) \Leftrightarrow (\exists k, l \in \mathbb{Z}, \text{sodass } b = ak \text{ und } c = bl)$. Dann $c = a(kl) \overset{kl\in\mathbb{Z}}{\Longleftrightarrow} a|c$.
Q.E.D.

Bemerkung 1.67. Für $a, b \neq 0$ haben wir $a|b \Rightarrow a \leqslant b$. Aber, $a|0$ für alle $a \in \mathbb{N}$, also das größte Element in dieser Ordnung ist 0, das heißt $\max_| \mathbb{N} = 0$. Es gibt auch ein Minimum: $\min_| \mathbb{N} = 1$.

Bemerkung 1.68.
1. Teilbarkeit ist keine partielle Ordnungsrelation auf $\mathbb{Z}$, weil es nicht antisymmetrisch ist: $7|-7$ und $-7|7$ aber $7 \neq -7$.

2. Wenn Sie schon wissen was ein Ring ist, dann können Sie die Definition von Teilbarkeit in $\mathbb{Z}$ problemlos verallgemeinern. Ich will aber hier betonen, dass ein Teiler von Null nicht unbedingt ein Nullteiler ist. Alle Elemente des Ringes sind Teiler von Null, aber in $\mathbb{Z}$ gibt es keine Nullteiler.

Satz 1.69 (Division mit Rest)**.** *Wenn $a, b \in \mathbb{Z}$ mit $b > 0$, dann existieren eindeutige ganze Zahlen $q, r \in \mathbb{Z}$ mit*

$$a = qb + r \qquad \text{und} \qquad 0 \leqslant r < b.$$

Beweis-Skizze: Sei $S := \{a - sb \ : \ s \in \mathbb{Z} \ \text{ und } a - sb \geqslant 0\} \subseteq \mathbb{N}$. Wir werden die Wohlordnung von $\mathbb{N}$ anwenden um den Rest als das Minimum von S zu definieren. Dafür müssen wir zu erst zeigen, dass diese Menge nicht leer ist.

$\boxed{S \neq \emptyset}$ Wenn $a \geqslant 0$, dann $a = a - 0 \cdot b \in S$.

Wenn $a < 0$, dann, weil $b > 0 \Rightarrow 1 - b \leqslant 0$, haben wir $0 \leqslant a(1-b) = (a - ab) \in S$.

Weil $\mathbb{N}$ wohl geordnet ist $\Rightarrow$ hat S ein Minimum. Wir setzen

$$r := a - s_0 b := \min S \qquad \text{und} \qquad q := s_0.$$

$\boxed{r < b}$ Wenn $r \geqslant b$, dann ist $r - b \geqslant 0$. Also $r > r - b = a - s_0 b - b \geqslant 0$ - ↯ zu $r = \min S$.

$\boxed{\text{Eindeutigkeit}}$ Wenn es andere q' und r' mit $a = q'b + r'$ existieren, dann haben wir

$$(q' - q)b = r - r' \tag{1.5}$$

Wenn $q = q'$, dann haben wir offensichtlich $r = r'$.

Wenn $q \neq q'$, dann können wir ohne die Allgemeinheit zu Beschränken $q' > q$, also $q' - q \geqslant 1$, annehmen. Also, wenn wir beide Seiten mit b multiplizieren bekommen wir

$$(q' - q)b \geqslant b > |r - r'| \ \text{↯zu (1.5)}.$$

Die rechte Ungleichung gilt, weil, wenn $0 \leqslant r, r' < b$, dann $|r - r'| < b$. Q.E.D.

Bemerkung 1.70. In Satz 1.69 kann man die Voraussetzung $b > 0$ mit $b \neq 0$ ersetzten, wenn man die Schlussfolgerung zu "...$0 \leqslant r < |b|$" ändert.

Extra:
Die Teilbarkeit auf $\mathbb{N}$ ist keine Wohlordnung (es gibt unvergleichbare Elementen). Es hat aber die beste Eigenschaft die man von einer partiellen Ordnung erwarten kann: es gibt immer das Infimum und das Supremum zweier Elementen und diese Operationen sind zu einander distributiv: $\inf(a, \sup(b, c)) = \sup(\inf(a, b), \inf(a, c))$ und $\sup(a, \inf(b, c)) = \inf(\sup(a, b), \sup(a, c))$. Für diese Ordnung heißen diese Operationen ggT und kgV.

Definition 1.71. Seien $a, b \in \mathbb{Z}$.

1. Ein **größter gemeinsamer Teiler** von a und b ist eine Zahl $d \in \mathbb{N}$ mit den Eigenschaften:

(ggT 1) $d|a$ und $d|b$

(ggT 2) Wenn $d'|a$ und $d'|b$, dann $d'|d$.

2. Ein **kleinstes gemeinsames Vielfaches** von a und b ist eine Zahl $m \in \mathbb{N}$ mit den Eigenschaften:

(kgV 1) $a|m$ und $b|m$

(kgV 2) Wenn $a|m'$ und $b|m'$, dann $m|m'$.

[8] 8.11.'23

Bemerkung 1.72.

1. Der ggT und das kgV sind eindeutig, weil $d|d'$ und $d'|d \Rightarrow d = d'$.
2. Weiterhin: $\mathrm{ggT}(0,a) = a$, $\mathrm{ggT}(0,0) = 0 = \mathrm{kgV}(0,0)$, $\mathrm{kgV}(0,a) = 0$.
3. Wenn $a|b$, dann $\mathrm{ggT}(a,b) = |a|$ und $\mathrm{kgV}(a,b) = |b|$.

Lemma 1.73 (Lemma von Bézout)**.** *Wenn $d = \mathrm{ggT}(a,b)$, dann $\exists s,t \in \mathbb{Z}$, sodass $d = as + bt$.*

Beweis-Skizze: Wenn $a = b = 0$, dann ist es offensichtlich wahr. Seien also a, b, nicht beide Null. Sei $S = \{sa + tb \ : \ s,t \in \mathbb{Z} \text{ und } sa + tb > 0\}$. Weil $\mathbb{N}$ wohl geordnet ist und $S \neq \emptyset$, existiert ein eindeutiges $\min(S)$. Wir behaupten, dass $\min(S) = \mathrm{ggT}(a,b)$.
Sei $d_0 := as_0 + bt_0 := \min(S)$.
(ggT 1) Aus der Division mit Rest haben wir $a = (s_0a + t_0b)q + r$. Also entweder $r = 0$ oder $r \in S$. Da $r < d_0 \overset{d_0 = \min S}{\Longrightarrow} r = 0$ und somit $d_0|a$. Dass $d_0|b$, folgt völlig analog.
(ggT 2) Sei d ein gemeinsamer Teiler. Dann existieren $a', b' \in \mathbb{Z}$ mit $a = da', b = db'$. Also

$$d_0 = s_0d'a' + t_0d'b' = d'(s_0a' + t_0b') \Rightarrow d'|d_0.$$

Q.E.D.

Korollar 1.74. Für $a, b \in \mathbb{Z}$ gilt $\mathrm{ggT}(a,b) = 1 \Leftrightarrow \exists s,t \in \mathbb{Z}$, sodass $as + bt = 1$.

Beweis-Skizze: $\Rightarrow$ ist ein Sonderfall von Lemma 1.73.
$\Leftarrow$ Wenn für $d \in \mathbb{N}$, $d|a$ und $d|b$ gilt, dann $d|as + bt \Rightarrow d = 1$. Q.E.D.

Der Euklidische Algorithmus.
`Eingabe:` $a, b \in \mathbb{Z}$.
`Schritt 0:` Wenn $a = 0$, dann `Ausgabe` $= |b|$.
Wenn $b = 0$, dann `Ausgabe` $= |a|$.
`Schritt 1:` Definiere $i := 0$, $r_{-1} := a$, $r_0 := b$.
`Schritt 2:` Durch Anwenden des Divisionssatz 1.69 für r_{i-1} und r_i finde die eindeutigen q und r mit

$$r_{i-1} = qr_i + r, \qquad \text{und} \qquad 0 \leqslant r < r_i.$$

`Schritt 3:` Wenn $r \neq 0$ definiere $i := i + 1$, $r_i := r$ und wiederhole `Schritt 2` (für das neue i).
Sonst `Ausgabe`$= r_i$.

Beispiel 1.75. Seien $a = 161$ und $b = 28$. Wir haben

$$\begin{aligned} 161 &= 5 \cdot 28 + 21 \\ 28 &= 1 \cdot 21 + \boxed{7} \leftarrow \text{letzter Rest } \neq 0 \\ 21 &= 3 \cdot 7 + 0 \end{aligned}$$

Also $\mathrm{ggT}(161, 28) = 7$. Es gibt auch ein erweiterter Euklidischer Algorithmus, der uns auch die zwei ganze Zahlen s, t aus dem Lemma von Bézout (Lemma 1.73) gibt.

Satz 1.76. *Der euklidische Algorithmus endet und gibt als Antwort* $\mathrm{ggT}(a,b)$.

Beweis-Skizze: Es endet, weil $b = r_0 > r_1 > \cdots \geqslant 0$. Dass die Antwort der ggT ist, folgt aus Lemma 1.77 und Bemerkung 1.72. Q.E.D.

Lemma 1.77. *Seien $a, b \in \mathbb{Z}$ mit $b > 0$ und seien q, r die eindeutig-bestimmten ganze Zahlen aus Satz 1.69. Dann gilt* $\mathrm{ggT}(a, b) = \mathrm{ggT}(b, r)$.

Beweis-Skizze: Weil $a = bq + r$ haben wir ($d|r$ und $d|b \Rightarrow d|a$). Weil $r = a - qb$ haben wir ($d|a$ und $d|b \Rightarrow d|r$). Also die gemeinsame Teiler von a und b sind genau die gemeinsame Teiler von b und r. Q.E.D.

Lemma 1.78. *Für alle $a, b \in \mathbb{Z}$ gilt* $|ab| = \mathrm{ggT}(a, b) \cdot \mathrm{kgV}(a, b)$.

Beweis-Skizze: Wenn $a = 0$ oder $b = 0$ ist der Satz offensichtlich wahr. Es reicht also den Satz für $a, b > 0$ zu zeigen. Sei $d = \mathrm{ggT}(a, b) > 0$. Dann $d|ab$, also $\exists\, l \in \mathbb{N}$ mit $ab = dl$. Wir zeigen, dass $l = \mathrm{kgV}(a, b)$. Es existieren $k_a, k_b \in \mathbb{N}$ mit $dk_a = a$ und $dk_b = b$. Also $ld = ab = dk_a b = dk_b a \overset{d\neq 0}{\Longrightarrow} l = k_a b = k_b a$, also (kgV1). Sei $m \in \mathbb{N}$ mit $a|m$ und $b|m$. Also $\exists\, n_a, n_b \in \mathbb{N}$ mit $m = an_a = bn_b$. Aus Bézouts Lemma folgt $d = ar + bs$ mit $r, s \in \mathbb{Z}$. Also $md = mar + mbs = (bn_b)ar + (an_a)bs = ab(n_b r + n_a s) = dl(n_b r + n_a s) \overset{d\neq 0}{\Longrightarrow} l|m$. Q.E.D.

Lemma 1.79 (von Euklid)**.** *Wenn $d|ab$ und* $\mathrm{ggT}(a, d) = 1$*, dann $d|b$.*

Beweis-Skizze: Weil $\mathrm{ggT}(a, d) = 1 \overset{\text{Lemma } 1.73}{\Longrightarrow} \exists\, k, l \in \mathbb{Z}$ mit $ka + ld = 1$. Also $kab + lbd = b$. Aus $d|kab$ und $d|lbd$, folgt dann $d|b$. Q.E.D.

Definition 1.80. Eine **Primzahl** ist eine natürliche Zahl $p > 1$ die nur durch 1 und p teilbar ist.

Satz 1.81. *Die Zahl $p \in \mathbb{N}$ ist eine Primzahl genau dann, wenn "für alle $a, b \in \mathbb{N}$ mit $p|ab$ folgt $p|a$ oder $p|b$".*

Beweis-Skizze: $\boxed{\Rightarrow}$ Aus Lemma 1.79.
$\boxed{\Leftarrow}$ Wenn $d|p \Rightarrow p = dk$. Also $p|dk$. Dann ($p|d$ oder $p|k$) $\Rightarrow$($d = p$ oder $d = 1$). Q.E.D.

1.5 Modulare Arithmetik

Definition 1.82. Sei $n \in \mathbb{N}_{>0}$. Wir definieren die Relation **Kongruenz modulo n** auf $\mathbb{Z}$, die man mit $\equiv \bmod n$ bezeichnet, durch

$$a \equiv b \bmod n \quad \Longleftrightarrow \quad n|(a - b).$$

Bemerkung 1.83. Die Kongruenz modulo n ist eine Äquivalenzrelation: $n|(a-b)$ heißt per Definition $\exists\, q \in \mathbb{Z}$, sodass $a - b = cq$. Für die Reflexivität wählt man $q = 0$. Für die Symmetrie, $-q$. Wenn $a - b = nq_1$ und $b - c = nq_2$, dann $a - c = n(q_1 + q_2)$, also die Transitivität gilt auch.

Wir bezeichnen die Quotientenmenge mit $\mathbb{Z}/n\mathbb{Z}$ und die Äquivalenzklassen mit $[a]_n$ oder nur $[a]$ oder $\hat{a}$. Ein (kanonisches) Repräsentantensystem ist $\{0, 1, \ldots, n-1\}$. Also

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \ldots, [n-1]\}.$$

Wir definieren die **Addition** $(+)$ und die **Multiplikation** $(\cdot)$ **modulo** n als:

$$\begin{aligned} + &: \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z} \qquad [a]_n + [b]_n := [a+b]_n \\ \cdot &: \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z} \qquad [a]_n \cdot [b]_n := [a \cdot b]_n \end{aligned}$$

[9] 13.11.'23

Satz 1.84 (Modulare Operationen)**.** *Sei $n \in \mathbb{N}_{>0}$. Die Addition und die Multiplikation modulo n sind wohl-definierte Operationen auf $\mathbb{Z}/n\mathbb{Z}$ (d.h. unabhängig vom Repräsentanten). Weiterhin, diese erfüllen:*

1. $\forall\, [a], [b], [c] \in \mathbb{Z}/n\mathbb{Z}$ *gilt* $[a] + ([b] + [c]) = ([a] + [b]) + [c]$ *und* $[a] \cdot ([b] \cdot [c]) = ([a] \cdot [b]) \cdot [c]$.
2. $\forall\, [a] \in \mathbb{Z}/n\mathbb{Z}$ *gilt* $[0] + [a] = [a] + [0]$ *und* $[1] \cdot [a] = [a] \cdot [1]$.
3. $\forall\, [a] \in \mathbb{Z}/n\mathbb{Z}$ *gilt* $[a] + [-a] = [0]$.
4. $\forall\, [a], [b] \in \mathbb{Z}/n\mathbb{Z}$ *gilt* $[a] + [b] = [b] + [a]$ *und* $[a] \cdot [b] = [b] \cdot [a]$.

Beweis-Skizze: Um zu zeigen, dass diese Operationen wohldefiniert sind, muss man zeigen, dass $\forall\, a, b, c, d \in \mathbb{Z}$ mit der Eigenschaft, dass $a \equiv b \bmod n$ und $c \equiv d \bmod n$ gilt

$$a + c \equiv b + d \bmod n \qquad \text{und} \qquad ac \equiv bd \bmod n.$$

Das ist eine einfache Rechnung (Übung).
Punkt 1. und 4. folgen direkt aus der entsprechenden Eigenschaften der Operationen auf $\mathbb{Z}$.
2. Per Definition haben wir $[0] + [a] = [0 + a] = [a]$ und $[1] \cdot [a] = [1 \cdot a] = [a]$ für alle $[a] \in \mathbb{Z}/n\mathbb{Z}$.
3. $\forall\, [a] \in \mathbb{Z}$ haben wir $[a] + [-a] = [-a] + [a] = [a - a] = [0]$. Q.E.D.

Jedes Element, also jede Restklasse, hat ein "inverses Element" für die Addition. Das heißt, für jedes $[a]$ existiert ein "$-[a]$", sodass $[a] + (-[a]) = [0]$. Für die Multiplikation ist das nicht immer so. Das heißt, es gibt Restklassen die mit keiner anderen Restklasse multipliziert die $[1]$ geben. Wir können aber genau sagen wann das geht.

Satz 1.85. *Es sei $n \in \mathbb{N}$ und $[a] \in \mathbb{Z}/n\mathbb{Z}$ mit $[a] \neq [0]$. Es existiert ein Element $[a'] \in \mathbb{Z}/n\mathbb{Z}$ mit $[a] \cdot [a'] = [a'] \cdot [a] = [1]$ genau dann, wenn $\mathrm{ggT}(a, n) = 1$.*

Beweis-Skizze: $\mathrm{ggT}(a, n) = 1 \overset{Kor\ 1.74}{\iff} \exists\, s, t \in \mathbb{Z}$ mit $as + nt = 1 \Leftrightarrow [as] + [nt] = [1] \Leftrightarrow [a][s] + [0] = [1] \Leftrightarrow\ [a]^{-1} = [s]$. Q.E.D.

Satz 1.86 (kleiner Satz von Fermat)**.** *Wenn $p \in \mathbb{Z}$ eine Primzahl ist, dann gilt $a^p \equiv a \bmod p$ für alle $a \in \mathbb{Z}$.*

Diesen Satz werden wir hier nicht direkt beweisen. Es ist ein Korollar des Satzes von Lagrange (Satz 1.141)

1.6 Gruppen, Ringe, Körper

In diesem letzten Teil der *Grundlagen* werden die Grundbegriffe der abstrakten Algebra eingeführt. Es kommt (viel) mehr als das, was wir für die Vorlesungen Lineare Algebra I und II brauchen. Ich habe mich jedoch entschieden, es trotzdem hier zu lassen. Ein Grund dafür ist die relative Vollständigkeit und Kohärenz der Darstellung. Ein weiterer Grund ist das ganze Verfahren der Abstraktion; das ist ein wichtiges Beispiel dafür, "was" Mathematik ist. Hier werden wir nicht nur Objekte, die bereits bekannt sind, verallgemeinern. Wir führen nicht nur weitere "Zahlen" oder verallgemeinern Polynome in einer Variabel auf Polynome in zwei oder n Variablen. Stattdessen bringen wir alle bekannten Objekte in eine neue, noch abstraktere Welt. Dies geschieht durch die Auswahl von Eigenschaften der algebraischen Operationen ($+$ und $\times$), die wir als Axiome voraussetzen. Sobald wir das tun, entdecken wir eine Vielfalt von Strukturen, die diese Axiome erfüllen. [22] Ganz am Ende des Kapitels werden wir auf einem weiteren Schritt in Richtung Abstraktion deuten: Kategorien. Dort werden gemeinsame Eigenschaften von Mengen, Gruppen, Ringe, Körper, partiell Geordnete Teilmengen, Vektorräume, Moduln, topologische Räume, Mannigfaltigkeiten, usw. auf einer noch abstrakterer Ebene gebracht. Dort schaut man auf das Zusammenspiel zwischen Objekte und Morphismen - das heißt erlaubte "Pfeile" (Abbildungen) zwischen den Objekten. Wir werden aber nur eine Definition geben und "Nein, Kategorien kommen nicht in der Linearen Algebra I Klausur vor."

1.6.1 Innere Verknüpfungen

Definition 1.87. Eine **innere Verknüpfung** (oder **innere algebraische Operation**) auf einer Menge M ist eine Abbildung $* : M \times M \longrightarrow M$. Wir bezeichnen mit $a * b := *(a, b)$.

1. Die Verknüpfung $*$ heißt **assoziativ** wenn $a * (b * c) = (a * b) * c, \quad \forall\, a, b, c \in M$.
2. Die Verknüpfung $*$ heißt **kommutativ** wenn $a * b = b * a, \quad \forall\, a, b, \in M$.
3. Ein **neutrales Element** für $*$ ist ein Element $e \in M$ mit der Eigenschaft

$$e * m = m * e = m \quad \forall\, m \in M.$$

Bemerkung 1.88. Wenn es ein neutrales Element für $*$ gibt, dann ist dieses eindeutig.

Beweis-Skizze: Seien e, e' neutrale Elemente. Dann gilt $e' = e * e' = e$. Q.E.D.

Definition 1.89. Sei e ein neutrales Element der inneren Verknüpfung $*$ auf M.
Ein **linksinverses Element** für $m \in M$ bezüglich $*$ und e ist ein Element $m' \in M$ mit der Eigenschaft

$$m' * m = e.$$

Ein **rechtsinverses Element** für $m \in M$ bezüglich $*$ und e ist ein Element $m'' \in M$ mit der Eigenschaft

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

Ein Element das sowohl linksinvers, als auch rechtsinvers von m ist heißt einfach **inverses Element** von m.

[22] Diese Theorie der Gruppen, Ringe und Körper hat sich im 19. Jahrhundert entwickelt.

Wir haben schon links- und rechstinverse Elemente in Satz 1.27 gesehen. Wir sehen gleich in der nächsten Bemerkung, dass die Assoziativität zu guten Sachen führt.

Bemerkung 1.90. Sei $*$ assoziativ, mit neutralem Element $e \in M$. Wenn $m \in M$ sowohl ein linksinverses Element $m' \in M$ als auch ein rechtsinverses $m'' \in M$ bezüglich $*$ und e besitzt, dann sind diese gleich. Insbesondere, ist das inverse Element, wenn es existiert, eindeutig.

Beweis-Skizze: Seien m', m'' inverse Elementen von m bezüglich $*$ und e. Dann gilt

$$m' = m' * e = m' * (m * m'') = (m' * m) * m'' = e * m'' = m''.$$

Q.E.D.

Man kann auch nur links oder rechst neutrale Elemente definieren. Wir brauchen das nicht und machen das auch nicht.

Bezeichnung. Wir können also über *das* Inverse von m sprechen und wir werden es meistens durch m^{-1} bezeichnen. Wenn aber die Operation mit $+$ bezeichnet ist, dann bezeichnen wir auch das inverse Element von m mit $-m$.

Bemerkung 1.91. Sei M eine Menge und $*$ eine innere Verknüpfung auf M, die assoziativ ist und die ein neutrales Element e hat.

(i) Wenn $m \in M$ invertierbar ist, dann ist auch m^{-1} invertierbar und es gilt

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

(ii) Wenn $m, n \in M$ invertierbar sind, dann ist auch $m * n$ invertierbar und es gilt

$$(m * n)^{-1} = n^{-1} * m^{-1}.$$

Beispiele:

1. Addition und Multiplikation auf der $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$. Diese sind assoziativ, kommutativ, haben neutrales Element 0 bzw. 1. Inverse gibt es immer für die Addition. Für die Multiplikation gibt es Inverse in $\mathbb{Q}, \mathbb{R}$ beziehungsweise $\mathbb{C}$ genau dann, wenn $m \neq 0$. Multiplikative Inverse gibt es in $\mathbb{Z}$ nur für ± 1, und in $\mathbb{N}$ nur für die 1.

2. Für $X = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, bezeichnen wir mit $X^\times := X \setminus \{ 0 \}$. Die Multiplikation ist eine innere Verknüpfung auch auf $X^\times$, aber die Addition ist keine Operation für $\mathbb{Z}^\times, \mathbb{Q}^\times, R^\times, X^\times$, weil $1 + (-1) \notin \mathbb{Z}^\times$ usw. Für $\mathbb{N}_{>0}$ ist aber die Addition eine innere Verknüpfung.

3. Für $X = [-2, 2] \subset \mathbb{R}$ sind $+$ und $\cdot$ keine Operationen.

4. Die modulare Operationen auf $\mathbb{Z}/n\mathbb{Z}$ sind auch innere Verknüpfungen (cf. Teil 1.5).

5. Die Abbildung $\frown : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, definiert durch $a \frown b := a^b$ ist eine innere Verknüpfung auf $\mathbb{N}$. Diese ist aber nicht assoziativ:

$$2^{(3^4)} = 2^{81} = 2,417851\ldots \cdot 10^{24} \neq 4096 = (2^3)^4.$$

Diese hat ein rechst-neutrales Element: 1; es gibt aber kein links-neutrales Element.

6. Ein kleines und abstraktes Beispiel zeigt, dass ohne Assoziativität, das inverse Element nicht eindeutig sein könnte. Operationen auf kleinen endlichen Mengen kann man gut durch Verknüpfungstafeln darstellen. Zum Beispiel, wenn $M = \{ e, a, b \}$ mit $e \neq a \neq b \neq e$, dann können wir folgende Verknüpfungstafel definieren:

$*$	e	a	b
e	e	a	b
a	a	e	e
b	b	e	e

   Wir können daraus lesen, dass $e * x = x * e = x, \quad \forall\, x \in M$, und dass $x * y = e, \ \forall x, y \in \{ a, b \}$. Also e ist ein neutrales Element und a hat zwei verschiedene Inversen: a und b. Das heißt, dass die Verknüpfung $*$ nicht assoziativ sein kann.

   Übung: *Man finde ein Tripel* $(\alpha, \beta, \gamma) \in M^3$*, sodass* $(\alpha * \beta) * \gamma \neq a * (\beta * \gamma)$.

7. Wenn M eine Menge ist, dann ist $\circ$ eine innere Verknüpfung auf

$$\mathrm{Iso}(M) = \{ f : M \longrightarrow M \ : \ f \text{ ist bijektiv} \}.$$

   Diese ist assoziativ, hat das neutrale Element id_M und, weil bijektive Abbildungen invertierbar sind, hat jedes Element das inverse Element f^{-1}.

8. Aus der so-genannten "*uparrow* Bezeichnung von Knuth" haben wir eine Innere Verknüpfung $\uparrow\colon \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ mit $a \uparrow b := a^b = a \cdot (a \cdot \ldots (a \cdot a)))$ wobei a kommt b-Mal vor.

   Man kann das iterieren durch $\uparrow\uparrow\colon \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ indem man $a \uparrow\uparrow b := a \uparrow (a \uparrow \ldots (a \uparrow a))$, wobei a genau b-Mal vorkommt, definiert. Und so weiter: $a \uparrow^3 b := a \uparrow\uparrow\uparrow b := a \uparrow\uparrow (a \ldots (a \uparrow\uparrow a))$.

   Die größte Zahl die in einem mathematischen Beweis vorkam ist *Graham's number* g_{64}. Man fängt mit $g_1 := 3 \uparrow\uparrow\uparrow\uparrow 3 = 3 \uparrow^4 3$ an, und definiert $g_n := 3 \uparrow^{g_{n-1}} 3$. Schon g_1 ist unvorstellbar groß.

$$\underbrace{3^{3^{\cdot^{\cdot^{\cdot^{3}}}}} \Big\} 3^{3^{\cdot^{\cdot^{\cdot^{3}}}}} \Big\} \ldots \Big\} 3^{3^3} \Big\} 3}_{3\uparrow\uparrow\uparrow 3 = 3^{3^{\cdot^{\cdot^{\cdot^{3}}}}} \Big\} 3^{3^3} \} 3}$$

   Die Zahl $3 \uparrow\uparrow\uparrow 3$ ist also ein Potenzturm der Höhe $3^{3^3} = 7625597484987 \approx 7 \cdot 10^{12}$, und das ist $\ggg$ das Volumen des beobachtbares Universums in Planck-Volumen gemäßen! (1 PVE $\approx 4,222 \cdot 10^{-105} m^3$.)

   Dies sind Operationen, die weder assoziativ noch kommutativ sind und die kein neutrales Element besitzen (obwohl es ein rechts-neutrales Element gibt: $a \uparrow 1 = a \neq 1 \uparrow a$, wenn $a \neq 1$). Es existiert jedoch ein rechts-inverses Element: $a^0 = 1$. Wenn die Verknüpfung assoziativ wäre, würden sowohl ein rechts-neutrales als auch ein rechts-inverses Element existieren.

1.6.2 Grundlegende Definitionen der Gruppentheorie

Eine Gruppe wird oft als eine Menge zusammen mit einer Verknüpfung definiert. Dies vermittelt die richtige Intuition, jedoch kann das "zusammen mit" genauer formuliert werden:

Definition 1.92. Eine **Gruppe** ist ein geordnetes Paar $(G, *)$, wobei G eine Menge ist und eine innere Verknüpfung auf G ist, die folgende drei Axiome erfüllt:

Gr 1. $*$ ist **assoziativ**.

Gr 2. Es existiert ein **neutrales Element** $e \in G$.

Gr 3. Zu jedem $g \in G$ gibt es ein **inverses Element** g^{-1} bezüglich $*$ und e.

Eine Gruppe heißt **abelsch**[23] (oder kommutativ) wenn $*$ kommutativ ist.

Wann immer die Verknüpfung klar aus dem Kontext ist, schreiben wir einfach G für die Gruppe $(G, *)$. Zum Beispiel, wenn wir "die Gruppe $\mathbb{Z}$" schreiben, dann ist $(\mathbb{Z}, +)$ mit der gewöhnlichen Addition er ganzen Zahlen gemeint.

Aus den Bemerkungen 1.88 und 1.90 folgt, dass in jeder Gruppe das neutrale Element und das inverse Element immer eindeutig bestimmt sind. Aus **Gr 2.** folgt, dass $G \neq \emptyset$.

[10] 15.11.'23

Definition 1.93. Seien $(G_1, *)$ und $(G_2, \star)$ zwei Gruppen. Ein **Gruppenhomomorphismus** von G_1 nach G_2 ist eine Abbildung $\varphi : G_1 \longrightarrow G_2$ mit der Eigenschaft

$$\varphi(g * g') = \varphi(g) \star \varphi(g') \quad \forall\ g, g' \in G_1.$$

Ein Gruppenisomorphismus ist ein Gruppenhomomorphismus, das invertierbar als Gruppenhomomorphismus ist. Die genaue Formulierung ist die folgende.

Definition 1.94. Ein Gruppenhomomorphismus $\varphi : G_1 \longrightarrow G_2$ ist ein **Gruppenisomorphismus**, wenn es einen *Gruppenhomomorphismus* $\varphi^{-1} : G_2 \longrightarrow G_1$ gibt, sodass

$$\varphi \circ \varphi^{-1} = \mathrm{id}_{G_2} \qquad \text{und} \qquad \varphi^{-1} \circ \varphi = \mathrm{id}_{G_1}.$$

Wir sagen, dass zwei Gruppen G_1 und G_2 **isomorph** sind, wenn es einen Gruppenisomorphismus $\varphi : G_1 \longrightarrow G_2$ gibt. Wir schreiben in diesem Fall $G_1 \simeq G_2$.

In Lemma 1.103 werden wir sehen, dass ein bijektiver Gruppenhomomorphismus automatisch ein Isomorphismus ist. Die Formulierung aus Definition 1.94 hat den großen Vorteil, dass sie durch Ersetzen des Wortes "Gruppe" für viele andere mathematische Strukturen übernommen werden kann. In manchen Fällen sind bijektive Morphismen wieder äquivalent zu Isomorphismen, aber nicht immer. Zum Beispiel gibt es für topologische Räume bijektive Homomorphismen, die keine Isomorphismen sind.

Definition 1.95. Eine **Untergruppe** einer Gruppe $(G, *)$ ist eine Teilmenge $H \subseteq G$, die folgende Axiome erfüllt.

UG 1. Für alle $h_1, h_2 \in H$ gilt $h_1 * h_2 \in H$. (**Abgeschlossenheit**)

[23] nach dem norwegischen Mathematiker Niels Henrik Abel, 1802-1829.

UG 2. Das neutrale Element e von G liegt auch in H. (**Neutrales Element**)

UG 3. Für jedes $h \in H$ gilt $h^{-1} \in H$. (**Inverses Element**)

Wir schreiben dafür $(H, *) \leq (G, *)$, oft auch $H \leq G$. Wenn $H \leq G$, aber $H \neq G$, dann schreiben wir $H \lneqq G$, $H \lneq G$, oder $H < G$.

Wenn **UG,1.** erfüllt ist, dann ist $*|H \times H : H \times H \to H$ eine wohldefinierte, assoziative Verknüpfung auf H und heißt die **induzierte** Verknüpfung auf H. In diesem Fall sind [**UG,2.** und **UG,3.**] äquivalent zu [$(H, *|H \times H)$ ist eine Gruppe].

1.6.3 Wichtige Beispiele von Gruppen

In der folgenden Liste stehen $+$ und $\cdot$ für die übliche Addition, beziehungsweise Multiplikation.

1. Wir beginnen mit einem nicht-Beispiel. Die erste algebraische Struktur die man schon in der Schule lernt ist die Menge $\mathbb{N} = \{ 0, 1, 2, \dots \}$ der natürlichen Zahlen, zusammen mit der Addition. Diese ist assoziativ und hat ein neutrales Element: 0. Die positiven[24] natürlichen Zahlen sind aber nicht in $\mathbb{N}$ invertierbar:

   $$\text{Wenn } n > 0, \text{ dann } \nexists a \in \mathbb{N}, \text{ sodass } n + a = 0.$$

   Das heißt, dass $(\mathbb{N}, +)$ **keine Gruppe** ist.

   In der Schule lernt man gleich nach der Addition die Subtraktion. Auf der Menge der natürlichen Zahlen ist das keine innere Verknüpfung, weil $a - b$ nicht immer in $\mathbb{N}$ liegt. Auf der Menge der ganzen Zahlen ist es eine innere Verknüpfung, es ist jedoch nicht assoziativ und hat nur ein rechts-neutrales Element: die Null. Deshalb werden wir nicht über Subtraktion als algebraische Operation sprechen, sondern diese als Addition von Inversen betrachten. Das heißt, für $a, b \in \mathbb{Z}$ ist

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

2. Die schon bekannten Zahlenmengen sind *additive*[25] abelsche Gruppen:

   $$(\mathbb{Z}, +), \quad (\mathbb{Q}, +), \quad (\mathbb{R}, +), \quad (\mathbb{C}, +).$$

   Das neutrale Element ist die Null, und das inverse Element von a bezüglich der Addition ist $-a$.

3. Dieselben Mengen sind **nicht Gruppen** zusammen mit der Multiplikation:

   $$(\mathbb{Z}, \cdot), \quad (\mathbb{Q}, \cdot), \quad (\mathbb{R}, \cdot), \quad (\mathbb{C}, \cdot) - \textbf{nicht Gruppen!}.$$

   Die Multiplikation ist in allen vier Fällen eine innere Verknüpfung, die assoziativ ist, und die auch ein neutrales Element hat: die Eins. Dieses sind aber nicht Gruppen, weil 0 nicht invertierbar bezüglich der Multiplikation ist. Das heißt, $\nexists\, a \in \mathbb{C}$, sodass $0 \cdot a = 1$.

[24] das hießt größer als Null.
[25] das heißt nur, dass die Verknüpfung die gewöhliche Addition ist.

4. In manchen Fällen reicht es, die Null zu entfernen, um eine multiplikative Gruppe zu finden. Wir bezeichnen hier $\mathbb{Q}^\times := \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ und $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$. Die Multiplikation bleibt eine innere Verknüpfung auch auf diesen Mengen, weil wenn $a, b \neq 0$, dann auch $ab \neq 0$. Wir nehmen als bekannt an, dass die Gruppenaxiome für die folgenden drei Paare erfüllt sind.

$$(\mathbb{Q}^\times, \cdot), \quad (\mathbb{R}^\times, \cdot), \quad (\mathbb{C}^\times, \cdot).$$

5. Für $\mathbb{Z}$ reicht es nicht aus, nur die Null zu entfernen. Die einzigen ganzen Zahlen, die invertierbar bezüglich der Multiplikation sind, sind -1 und 1. Also

$$(\mathbb{Z} \setminus \{0\}, \cdot) \qquad \textbf{ist keine Gruppe.}$$

   Es ist wichtig zu betonen, dass wir immer über Invertierbarkeit *in der gegebenen Menge* sprechen. Das heißt, 22 ist nicht invertierbar in $\mathbb{Z}$, weil es keine ganze Zahl a gibt, sodass $22 \cdot a = 1$.

6. Die Modularen Operationen sind auch innere Verknüpfungen auf $\mathbb{Z}/n\mathbb{Z}$, wobei $n \in \mathbb{N}$. Es gilt:

$$(\mathbb{Z}/n\mathbb{Z}, +) \quad \text{ist eine Gruppe für alle } n \in \mathbb{N}.$$

   Genau wie im Fall der komplexen Zahlen müssen wir im Fall der Multiplikation nur die multiplikativ invertierbaren auswählen, um eine Gruppe zu erhalten. Die modulare Multiplikation ist assoziativ und hat $[1]_n$ als neutrales Element. Aus Bemerkung 1.91 folgt, dass die Einschränkung der Multiplikation eine innere Verknüpfung auf die Menge der multiplikativ invertierbaren Elemente in $\mathbb{Z}/n\mathbb{Z}$ ist:

$$U(\mathbb{Z}/n\mathbb{Z}) = \left\{ [a]_n \in \mathbb{Z}/n\mathbb{Z} \;:\; \exists [a']_n \in \mathbb{Z}/n\mathbb{Z} \text{ mit } [a]_n \cdot [a']_n = [1]_n \right\}.$$

   In Satz 1.85 haben wir gesehen, dass

$$U(\mathbb{Z}/n\mathbb{Z}) = \{ [a]_n \in \mathbb{Z}/n\mathbb{Z} \;:\; \mathrm{ggT}(a, n) = 1 \}.$$

   Insbesondere haben wir für jede Primzahl p, dass folgendes Paar eine Gruppe ist:

$$\left(\mathbb{Z}/p\mathbb{Z} \setminus \{[0]_p\}, \cdot\right)$$

7. Für jede Menge M ist die Menge aller bijektiven Selbstabbildungen

$$\mathrm{Sym}(M) = \{ f : M \longrightarrow M \;:\; f \text{ ist bijektiv} \}$$

   zusammen mit der Verknüpfung von Abbildungen eine Gruppe. Wir haben bereits in Bemerkung 1.23 gesehen, dass ganz allgemein die Verknüpfung von Abbildungen assoziativ ist. Wenn $f, g : M \longrightarrow M$, dann sind diese immer verknüpfbar zu $f \circ g : M \longrightarrow M$ und $g \circ f : M \longrightarrow M$. Wenn beide bijektiv sind, dann sind auch beide Verknüpfungen bijektiv (Satz 1.27 (iii) und Bemerkung 1.28). Das bedeutet, dass die Verknüpfung von Abbildungen eine assoziative innere Verknüpfung auf $\mathrm{Sym}(M)$ ist. Das neutrale Element ist id_M, und aus Satz 1.27 ist auch jedes Element in $\mathrm{Sym}(M)$ invertierbar.

   Diese Gruppe ist besonders wichtig wenn die Menge M endlich ist. In diesem Fall, kann man annehmen, dass $M = \{ 1, \ldots, n \}$ für ein $n \in \mathbb{N}_{>0}$ ist.

Definition 1.96. Die **symmetrische Gruppe** S_n ist die Gruppe $\mathrm{Sym}\big(\{1,\dots,n\}\big)$ aller Bijektionen von $\sigma : \{1,\dots,n\} \longrightarrow \{1,\dots,n\}$ zusammen mit der Verknüpfung von Abbildungen. Ein Element $\sigma \in S_n$ heißt **Permutation** und wird aufgeschrieben als

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

8. Für jede Menge M können wir auch die Menge aller reellwertigen Funktionen auf M definieren:

$$\mathcal{F}_{\mathbb{R}}(M) := \{\, f : M \longrightarrow \mathbb{R} \;:\; f \text{ ist eine Abbildung} \,\}.$$

Wenn M nicht $\mathbb{R}$ ist, dann kann man zwei Funktionen nicht mehr verknüpfen. Man kann aber eine Addition auf dieser Menge definieren. Für alle $f, g \in \mathcal{F}_{\mathbb{R}}(M)$ sei

$$f + g : M \longrightarrow \mathbb{R}, \qquad (f+g)(m) := f(m) + g(m) \;\; \forall\, m \in M.$$

Man sieht gleich, dass die Assoziativität der Addition auf $\mathbb{R}$ die Assoziativität der Addition auf $\mathcal{F}_{\mathbb{R}}(M)$ impliziert. Die konstante Funktion $0 : M \longrightarrow \mathbb{R}$, mit $0(m) = 0$ für alle $m \in M$ ist das neutrale Element, und das Inverse einer Funktion $f : M \longrightarrow \mathbb{R}$ bezüglich der Addition von Funktionen ist die Funktion

$$-f : M \longrightarrow \mathbb{R} \qquad (-f)(m) = -(f(m)) \;\; \forall\, m \in M.$$

Das ist also auch eine Gruppe.

9. Die Gruppe $\mathcal{Q}$ der Quaternionen. Diese ist eine nicht-kommutative Gruppe mit 8 Elementen:

$$\mathcal{Q} = \{\, \pm 1, \pm i, \pm j, \pm k \,\}.$$

Die Verknüpfung ist multiplikativ bezeichnet, sodass die Vorzeichen genau wie bei der Multiplikation der reellen Zahlen funktionieren. Insbesondere, ist 1 das neutrale Element von $\mathcal{Q}$, es gilt $(-1)^2 = 1$, und das Element (-1) kommutiert mit allen Elementen der Gruppe. Weiterhin gilt noch

$$i^2 = j^2 = k^2 = -1 \qquad \text{und} \qquad ijk = -1.$$

Aus diesen Relationen kann man dann beweisen, dass $ij = -ji$, $ik = -ki$ und $jk = -kj$. Welches der acht Elemente von $\mathcal{Q}$ wird gleich mit ij sein?

Diese sind einige wichtige Beispiele, die wir mit den jetzigen Kenntnissen beschreiben können. Viele geometrische Beispiele wie GL_n, SL_n, $O(n)$, $\mathrm{SO}(n)$, AGL_n, die für die Mathematik und Physik wesentlich sind, können wir noch nicht beschreiben.. Diese werden jedoch später in dieser Vorlesung eingeführt und studiert.

1.6.4 Erste Eigenschaften von Gruppen

Bemerkung 1.97. Sei $(G, *)$ eine Menge mit inneren Verknüpfung $*$. Wenn folgende Axiome gelten, dann ist $(G, *)$ eine Gruppe.

Gr 1. $*$ ist assoziativ.

Gr 2'. Es existiert ein Element $e \in G$, sodass $e * g = g$, $\forall\, g \in G$. (links-neutrales Element)

Gr 3'. Zu jedem $g \in G$ gibt es ein links-inverses Element g' bezüglich $*$ und bezüglich jedem links-neutrales Element e.

Beweis-Skizze: Wir zeigen zu erst, dass wenn **Gr 1.**, **Gr 2.**, und **Gr 3'.** gelten, dann ist ein links-inverses Element auch rechts-invers, also dass **Gr 3.** gilt.
Sei $g \in G$, und sei g' sodass $g' * g = e$. Sei g'' ein links-inverses Element von g', also $g'' * g' = e$. Wir wollen zeigen, dass $g * g' = e$, also das g' auch ein rechst-inverses Element für g ist. Wir haben

$$\begin{aligned} g * g' &= (e * g) * g' && (\textbf{Gr 2'.}) \\ &= ((g'' * g') * g) * g' && (\textbf{Gr 3'.}) \\ &= (g'' * (g' * g)) * g' && (\textbf{Gr 1.}) \\ &= (g'' * e) * g' && (\textbf{Gr 3'.}) \\ &= g'' * (e * g') && (\textbf{Gr 1.}) \\ &= g'' * g' && (\textbf{Gr 2'.}) \\ &= e && (\textbf{Gr 3'.}) \end{aligned}$$

Wir zeigen jetzt, dass **Gr 2.** gilt, indem wir zeigen, dass wenn **Gr 1.**, **Gr 2'.**, und **Gr 3.** gelten, dann ist ein links-neutrales Element auch rechts-neutral. Sei $g \in G$ und e ein links-neutrales Element. Wir wollen also zeigen, dass $g * e = g$. Wir haben

$$g * e \overset{\textbf{Gr 3}}{=} g * (g' * g) \overset{\textbf{Gr 1}}{=} (g * g') * g \overset{\textbf{Gr 3}}{=} e * g \overset{\textbf{Gr 2'}}{=} g.$$

Q.E.D.

Da offensichtlich **Gr 2.** $\Rightarrow$**Gr 2'.** und **Gr 3.**$\Rightarrow$**Gr 3'.**, haben wir gezeigt, dass

$$[\textbf{Gr 1.} \text{ und } \textbf{Gr 2.} \text{ und } \textbf{Gr 3.}] \iff [\textbf{Gr 1.} \text{ und } \textbf{Gr 2'.} \text{ und } \textbf{Gr 3'.}].$$

Das heißt aber nicht, dass **Gr 3'.**$\Leftrightarrow$**Gr 3.** oder **Gr 2'.**$\Leftrightarrow$**Gr 2.** (Siehe das Beispiel 5. hier oben, und Satz 1.27 + ein Beispiel von injektive, aber nicht surjektive Abbildung). Wir hätten also eine Gruppe auch durch die Axiome **Gr 1.**,**Gr 2'.**, und **Gr 3'.** definieren können. Manche Autoren machen das auch. Ich finde, dass Definition 1.92 klarer ausdrückt was eine Gruppe ist. Der Nachteil ist, dass man mehr überprüfen muss um zu zeigen, dass etwas eine Gruppe ist. Aber, um schneller zu beweisen, dass eine Menge mit einer Verknüpfung eine Gruppe bildet, kann man immer Bemerkung 1.97 anwenden.

Die **Verknüpfungstafel** einer endlichen Gruppe mit n Elementen ist eine $n \times n$ Tabelle deren Spalten und Zeilen von den Elementen der Gruppe indiziert sind, sodass in der g-Zeile an der h-Stelle das Gruppenelement $g * h$ vorkommt. Die Reihenfolge der Elemente von G, sowohl in der Indexierung der Zeilen, als auch in der Indexierung der Spalten, ist dieselbe. Üblicherweise kommt das neutrale Element als erstes vor. Zum Beispiel, wenn $G = \mathbb{Z}/4\mathbb{Z}$ (also die Verknüpfung ist die Addition modulo 4) dann haben wir

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Folgendes Lemma hat als Folgerung, dass in jeder Zeile und in jeder Spalte jedes Gruppenelement genau ein Mal vorkommt. Es gilt also eine Art "Sudoku-Regel". Das Inverse von g kann man als den Spalten-Index des neutrales Elementes in der g-Zeile ablesen. Die Kommutativität entspricht der Symmetrie

der Verknüpfungstafel bezüglich der Hauptdiagonale. Die Assoziativität kann man aus der Tafel nicht mehr gut lesen.

Lemma 1.98 (Kürzungsregeln)**.** *Sei* $(G, *)$ *eine Gruppe. Für alle* $g_1, g_2, g_3 \in G$ *gilt*

$$(g_1 * g_2 = g_1 * g_3 \implies g_2 = g_3) \qquad \text{und} \qquad (g_1 * g_3 = g_2 * g_3 \implies g_1 = g_2).$$

Beweis-Skizze: Wir verknüpfen linkst mit $g_1^{-1}*$ (bzw. rechts mit $* g_3^{-1}$) und wenden Assoziativität an. Q.E.D.

Bezeichnung. Wir haben die innere Verknüpfung mit $*$ bezeichnet, um zu betonen, dass es eine abstrakte Operation ist. Wir werden aber bald die bekannteren Symbole $+$ und $\cdot$ anwenden, mit der wichtigen Bemerkung, dass diese nicht unbedingt die Addition und die Multiplikation aus der Schule sind. Mit dieser Konvention, werden wir das neutrale Element der Verknüpfung $\cdot$ mit 1_G, oder einfach mit 1, bezeichnen. Das inverse Element bleibt g^{-1}. Für die Verknüpfung selbst werden wir

$$gh := g \cdot h$$

schreiben. Diese Verknüpfung muss nicht unbedingt kommutativ sein. Es kann also passieren, dass $gh \neq hg$.
Wenn wir die Verknüpfung auf G mit $+$ bezeichnen, ist die Konvention, dass diese Operation auch kommutativ ist. Wir haben unter dieser Notation für die Verknüpfung auch Soderbezeichnungen für neutrale und inverse Elemente: 0_G oder 0, und $-g$.

Die Assoziativität der Verknüpfung erlaubt uns die Verknüpfung endlich-vieler Elementen eindeutig zu definieren. Wenn die Operation $+$ oder $\cdot$ ist, dann schreiben wir

$$\begin{aligned} \textstyle\sum_{i=1}^n g_i &:= g_1 + g_2 + \cdots + g_n := ((g_1 + g_2) + \cdots + g_{n-1}) + g_n, \\ \textstyle\prod_{i=1}^n g_i &:= g_1 g_2 \dots g_n := ((g_1 \cdot g_2) \cdot \cdots \cdot g_{n-1}) \cdot g_n. \end{aligned}$$

Es ist wichtig, dass es *endlich* viele Elementen sind. Unendliche Summen und Produkte von Elementen sind in der abstrakten Algebra nicht definiert. Wenn $g_1 = \cdots = g_n = g$, dann schreiben wir

$$\begin{aligned} ng &:= \textstyle\sum_{i=1}^n g = g + \cdots + g, \\ g^n &:= \textstyle\prod_{i=1}^n g = g \cdot \cdots \cdot g. \end{aligned}$$

Wenn $n = 0$ dann ist die leere Summe per Definition gleich mit 0_G und das leere Produkt gleich mit 1_G. Für $n \in \mathbb{N}$ schreiben wir auch

$$\begin{aligned} (-n)g &:= \textstyle\sum_{i=1}^n -g = (-g) + \cdots + (-g), \\ g^{-n} &:= \textstyle\prod_{i=1}^n g^{-1} = g^{-1} \cdot \cdots \cdot g^{-1}. \end{aligned}$$

Bemerkung 1.99. Wenn $(G, \cdot)$ eine Gruppe ist, dann gilt für alle $a, b \in \mathbb{Z}$:

$$g^a \cdot g^b = g^{a+b}.$$

Insbesondere gilt $g^n \cdot g^{-n} = g^0 = e$, also $(g^n)^{-1} = g^{-n}$.

Beispiel 1.100. S_3 ist die kleinste nicht-kommutative Gruppe. Es hat als Elemente

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Wir haben $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4\sigma_5 = e$. Diese Gruppe ist nicht kommutativ, weil $\sigma_1\sigma_2 = \sigma_5 \neq \sigma_4 = \sigma_2\sigma_1$. Die Operationstafel von S_3 ist

$\circ$	e	σ_1	σ_2	σ_3	σ_4	σ_5
e	e	σ_1	σ_2	σ_3	σ_4	σ_5
σ_1	σ_1	e	σ_5	σ_4	σ_3	σ_2
σ_2	σ_2	σ_4	e	σ_5	σ_1	σ_3
σ_3	σ_3	σ_5	σ_4	e	σ_2	σ_1
σ_4	σ_4	σ_2	σ_3	σ_1	σ_5	e
σ_5	σ_5	σ_3	σ_1	σ_2	e	σ_4

Weiterhin, jede endliche Gruppe ist einer symmetrischen Gruppe S_n als Untergruppe "enthalten"[26] Die Korrespondenz basiert sich auf der Beobachtung, dass jedes Element g einer Gruppe G eine bijektive Abbildung definiert: $\cdot g : G \longrightarrow G$ durch $x \mapsto x \cdot g$. Diese Abbildung ist bijektiv, weil das Inverse die Inverse definiert: $\cdot g^{-1} = (\cdot g)^{-1}$. Für endliche Gruppen können wir dann einen injektiven Gruppenhomomorphismus (cf. Definition 1.93) $G \longrightarrow S_{|G|}$ durch

$$G \ni g \mapsto \cdot g \in S_{|G|}$$

definieren.

Satz 1.101. *Eine Teilmenge $H \subseteq G$ ist genau dann eine Untergruppe von G, wenn $H \neq \emptyset$ und*

$$\forall\ a, b \in H \ \Rightarrow\ ab^{-1} \in H. \tag{1.6}$$

Beweis-Skizze: $\boxed{\Rightarrow}$ Wir nehmen an, dass für H die Axiome aus Definition 1.95 gelten. Aus **UG 2.** gilt $e \in H$, also $H \neq \emptyset$. Seien jetzt $a, b \in H$ beliebig. Aus **UG 3.** folgt $b^{-1} \in H$ und aus **UG 1.**, dass $ab^{-1} \in H$.

$\boxed{\Leftarrow}$ Weil $H \neq \emptyset$ existiert $h \in H$. Wir können dann in (1.6) $a = b = h$ einsetzen und bekommen:

$$ab^{-1} = hh^{-1} = e \in H.$$

Also **UG 2.** gilt. Wir wählen jetzt ein beliebiges $h \in H$ und setzen $a = e \in H$ (weil wir **UG 2.** bewiesen haben, dürfen wir das machen) und $b = h$ in (1.6) ein. Es folgt

$$ab^{-1} = eh^{-1} = h^{-1} \in H.$$

Also Axiom **UG 3.** gilt auch. Wir wählen jetzt $h_1, h_2 \in H$ beliebig. Wir setzten $a = h_1$. Aus **UG 3.** gilt $h_2^{-1} \in H$, und wir setzen dann $b = h_2^{-1}$. Es folgt dann aus (1.6):

$$ab^{-1} = h_1(h_2^{-1})^{-1} = h_1h_2 \in H,$$

[26] wir brauchen Gruppenhomomorphismen, um das genau auszudrücken.

und somit gilt auch das Axiom **UG 1.**. Q.E.D.

Definition 1.102. Es sei $\varphi : G_1 \longrightarrow G_2$ ein Gruppenhomomorphismus. Der **Kern** von φ ist die Faser über e_2, wobei e_2 das neutrale Element von G_2 ist. Wir bezeichnen den Kern mit $\operatorname{Ker}\varphi$, und haben also

$$\operatorname{Ker}\varphi := \{\, g \in G_1 \;\; : \;\; \varphi(g) = e_2 \,\}.$$

Der Kern ist also per Definition eine Teilmenge von G_1. Für jede Abbildung hatten wir in Teil **??** das **Bild** definiert. Wir können insbesondere also für einen Gruppenhomomorphismus $\varphi : G_1 \longrightarrow G_2$ über

$$\operatorname{Bild}\varphi := \{\, g_2 \in G_2 \;\; : \;\; \exists\, g_1 \in G_1 \varphi(g_1) = g_2 \,\}$$

sprechen. Das ist a priori eine Teilmenge von G_2. Wir werden in Lemma 1.104 zeigen, dass beide diese Mengen sogar Untergruppen sind. Zu erst zeigen wir aber einige elementare Eigenschaften von Gruppenhomomorphismen.

Lemma 1.103. *Es seien $(G_1, *)$ und $(G_2, star)$ zwei Gruppen und $\varphi : G_1 \longrightarrow G_2$ ein Gruppenhomomorphismus. Bezeichne e_i das neutrale Element von G_i für $i = 1, 2$.*

(i) $\varphi(e_1) = e_2$.

(ii) $\varphi(g^{-1}) = \varphi(g)^{-1}$.

(iii) *φ ist injektiv* $\iff \operatorname{Ker}\varphi = \{\, e_1 \,\}$.

(iv) *φ ist surjektiv* $\iff \operatorname{Bild}\varphi = G_2$.

(v) *Wenn φ bijektiv, dann ist $\varphi^{-1} : G_2 \longrightarrow G_1$ auch ein Gruppenhomomorphismus.*

[11] 20.11.'23

Beweis-Skizze:

(i) Sei $g \in G_1$. Es gilt

$$\varphi(g) \star \varphi(e_1) = \varphi(g * e_1) = \varphi(g) = \varphi(g) \star e_2.$$

Nach der Kürzungsregel (Lemma 1.98 folgt $\varphi(e_1) = e_2$.

(ii) Sei $g \in G_1$ beliebig. Wir haben

$$\varphi(g) \star \varphi(g^{-1}) = \varphi(g * g^{-1}) = \varphi(e_1) = e_2.$$

Also, weil G_2 eine Gruppe ist, folgt aus der Eindeutigkeit des inverses Elementes, dass $\varphi(g)^{-1} = \varphi(g^{-1})$.

(iii) $\boxed{\Rightarrow}$ Wir haben aus Punkt (i), dass $e_1 \in \operatorname{Ker}\varphi$. Nehmen wir an, dass $g \in \operatorname{Ker}\varphi$ beliebig. Das bedeutet

$$\varphi(g) = e_2 = \varphi(e_1).$$

Aus der Injektivität von φ folgt dann $g = e_1$. Also $\operatorname{Ker}\varphi = \{\, e_1 \,\}$.

$\boxed{\Leftarrow}$ Es seien $g, h \in G_1$ mit $\varphi(g) = \varphi(h)$. Wenn wir beide Seiten der Gleichung rechts mit

$\varphi(h)^{-1}$ verknüpfen und Punkt (ii) anwenden, dann bekommen wir

$$\begin{aligned} \varphi(g) \star \varphi(h)^{-1} &= \varphi(h) \star \varphi(h)^{-1} \\ \varphi(g) \star \varphi\left(h^{-1}\right) &= e_2 \\ \varphi\left(g * h^{-1}\right) &= e_2. \end{aligned}$$

Also $g * h^{-1} \in \operatorname{Ker} \varphi$. Aus der Voraussetzung, ist $\operatorname{Ker} \varphi = \{e_1\}$, also

$$g * h^{-1} = e_1.$$

Wenn wir $*h$ auf beiden Seiten der Gleichheit anwenden, dann bekommen wir

$$g * h^{-1} * h = e_1 * h$$

also aus den Gruppenaxiomen, dass $g = h$. Somit ist φ injektiv.

(iv) Das ist einfach eine Umformulierung der Definition der Surjektivität (Definition **??**).

(v) Wir müssen zeigen, dass

$$\varphi^{-1}(g' \star h') = \varphi^{-1}(g') * \varphi^{-1}(h') \quad \forall\ g', h' \in G_2.$$

Es seien $g', h' \in G_2$ beliebig. Weil φ bijektiv ist, existieren eindeutige $g, h \in G_1$, sodass

$$\varphi(g) = g' \qquad \text{und} \qquad \varphi(h) = h'.$$

Es gilt also auch $\varphi^{-1}(g') = g$ und $\varphi^{-1}(h') = h$. Wir haben dann

$$\begin{aligned} \varphi^{-1}(g' \star h') &= \varphi^{-1}(\varphi(g) \star \varphi(h)) && \text{(Definition von } g' \text{ und } h'.) \\ &= \varphi^{-1}(\varphi(g * h)) && \text{(weil } \varphi \text{ ein Homomorphismus ist)} \\ &= (\varphi^{-1} \circ \varphi)(g * h)) && \text{(Verknüpfung von Abbildungen)} \\ &= \mathrm{id}_{G_1}(g * h)) && \text{(weil } \varphi^{-1} \text{ die Inverse ist)} \\ &= g * h && \text{(Identische Abbildung)} \\ &= \varphi^{-1}(g') * \varphi^{-1}(h'). && \text{(Definition von } g' \text{ und } h'.) \end{aligned}$$

Q.E.D.

Lemma 1.104. *Es seien $(G_1, *)$ und $(G_2, \star)$ zwei Gruppen und $\varphi : G_1 \longrightarrow G_2$ ein Gruppenhomomorphismus. Es gilt:*

(i) $\operatorname{Ker} \varphi$ *ist eine Untergruppe von* G_1.

(ii) $\operatorname{Bild} \varphi$ *ist eine Untergruppe von* G_2.

Beweis-Skizze:

(i) Seien $g, h \in \operatorname{Ker}\varphi$. Laut Satz 1.101 müssen wir zeigen, dass $g * h^{-1} \in \operatorname{Ker}\varphi$. Wir haben

$$\varphi(g * h^{-1}) = \varphi(g) \star \varphi(h^{-1}) = \varphi(g) \star (\varphi(h))^{-1} = e_2 \star (e_2)^{-1} = e_2.$$

Also $g * h^{-1} \in \operatorname{Ker}\varphi$.

(ii) Seien $g', h' \in \operatorname{Bild}\varphi$. Das heißt, es existieren $g, h \in G_1$, sodass

$$\varphi(g) = g' \qquad \text{und} \qquad \varphi(h) = h'.$$

Wir haben nach Bemerkung 1.103 (iii), dass $(h')^{-1} = (\varphi(h))^{-1} = \varphi(h^{-1})$. Es gilt also

$$g' \star (h')^{-1} = \varphi(g) \star \varphi(h^{-1}) = \varphi(g * h^{-1}) \in \operatorname{Bild}\varphi.$$

Q.E.D.

Satz 1.105. *Jede Untergruppe von $(\mathbb{Z}, +)$ hat die Form $n\mathbb{Z} := \{\, nz \;:\; z \in \mathbb{Z} \,\}$, für ein gewisses $n \in \mathbb{N}$.*

Beweis-Skizze: Sei $H < \mathbb{Z}$. Wenn $H = \{\, 0 \,\}$, dann $H = 0\mathbb{Z}$. Wenn $H \neq \{\, 0 \,\}$, dann $\exists\, h \in H$, mit $h \neq 0$. Da auch $-h \in H$, folgt, dass die Menge $H_+ := \{\, h \in H \;:\; h > 0 \,\}$ nicht leer ist. Weil $\mathbb{N}$ wohl geordnet ist, folgt dass es ein Minimum $n := \min H_+$ hat. Wir zeigen jetzt, dass $H = n\mathbb{Z}$.
$\boxed{\supseteq}$ Gilt offensichtlich aus (UG1) und (UG3).
$\boxed{\subseteq}$ Sei $h \in H$. Wir nehmen an, dass $h > 0$ (sonst ersetzten wir es mit $-h$). Aus Satz 1.69 folgt $\exists\, q, r \in \mathbb{Z}$ mit $0 \leqslant r < n$ so dass $h = qn + r$. Es gilt also

$$r = h - qn.$$

Aus (UG1), dass $qn \in H$, aus (UG3), dass $-qn \in H$ und wieder aus (UG1), dass $h - qn \in H$. Also wenn $r \neq 0$, haben wir $r \in H_+$ mit $r < n$ – ein Widerspruch ↯ zu $n = \min H_+$. Q.E.D.

Satz 1.106. *Wenn $\varphi : \mathbb{Z} \longrightarrow \mathbb{Z}$ ein Gruppenhomomorphismus[27] ist, dann existiert ein $a \in \mathbb{Z}$, sodass*

$$\varphi(z) = a \cdot z \quad \forall\; z \in \mathbb{Z}.$$

Beweis-Skizze: Wir zeigen, dass dieses $a = \varphi(1)$ ist. Wir haben für $n \in \mathbb{N}_{>0}$

$$\varphi(n) = \varphi(1 + \cdots + 1) = \varphi(1) + \cdots + \varphi(1) = n \cdot \varphi(1) = a \cdot n,$$

wobei $1 + \cdots + 1$ und $\varphi(1) + \cdots + \varphi(1)$ genau n Summanden haben. Für 0 gilt auch $\varphi(0) = 0 = a \cdot 0$. Wenn $z < 0$, dann gilt $z = -n$ mit $n \in \mathbb{N}$ und wir haben $\varphi(-1) = -\varphi(1) = -a$. Es gilt also

$$\varphi(z) = \varphi(-n) = \varphi\big((-1) + \cdots + (-1)\big) = \varphi(-1) + \cdots + \varphi(-1) = n \cdot \varphi(-1) = n \cdot (-a) = a \cdot (-n) = a \cdot z.$$

Q.E.D.

[27] Wenn man $\mathbb{Z}$ als Gruppe erwähnt, ohne eine Verknüpfung anzugeben, dann versteht man immer die Addition der ganzen Zahlen als Gruppenoperation.

1.6.5 Das Direkte Produkt von Gruppen

Wenn $(G_1, *)$ und $(G_2, \star)$ zwei Gruppen sind, dann kann man eine Gruppenstruktur auf dem kartesischen Produkt $G_1 \times G_2$ der zwei Mengen definieren:

$$(g_1, g_2) \diamond (h_1, h_2) := (g_1 * h_1, g_2 \star h_2).$$

Das ist eine innere Verknüpfung auf $G_1 \times G_2 = \{ (g_1, g_2) \ : \ g_i \in G_i, i = 1, 2 \}$. Die Assoziativität von $\diamond$ folgt direkt aus der Assoziativität von $*$ und $\star$. Das neutrale Element in $G_1 \times G_2$ ist (e_1, e_2), wobei für $i = 1, 2$ ist e_i das neutrale Element von G_i. Das inverse Element von (g_1, g_2) ist

$$(g_1, g_2)^{-1} = (g_1^{-1}, g_2^{-1}).$$

Die obig definierte Gruppe $(G_1 \times G_2, \diamond)$ heißt das **direkte Produkt** der Gruppen G_1 und G_2.

Beispiele:

1. $(\mathbb{R}^2, +)$ wobei $(x_1, x_2) + (x_1', x_2') := (x_1 + x_1', x_2 + x_2)'$. Genau so auch $(\mathbb{Q}^2, +)$, $(\mathbb{Z}^2, +)$, $(\mathbb{C}^2, +)$, $(\mathbb{Q}^{\times 2}, \cdot)$, $(\mathbb{R}^{\times 2}, \cdot)$, $(\mathbb{C}^{\times 2}, \cdot)$.
2. Man kann auch mehr machen: $(\mathbb{R}^3, +)$, $(\mathbb{R}^4, +), \ldots,$ $(\mathbb{R}^n, +)$ usw. Es spielt keine Rolle, dass es die reellen Zahlen sind, das geht auch für beliebige Gruppen.
3. Allgemeiner: wenn $(G, *)$ eine Gruppe ist und $n \in \mathbb{N}_{>0}$, dann kann man auf $G^n := \underbrace{G \times \cdots \times G}_{n-\text{Mal}}$ eine Gruppenstruktur definieren, indem man
$$(g_1, \ldots, g_n) * (h_1, \ldots, h_n) := (g_1 * h_1, \ldots, g_n * h_n)$$
setzt. Assoziativität ist klar. Das neutrale Element ist $(e, \ldots, e)$ und das inverse ist $(g_1, \ldots, g_n)^{-1} = (g_1^{-1}, \ldots, g_n^{-1})$. Konkret:
4. Noch allgemeiner: Wenn $(G_1, *_1), \ldots, (G_n, *_n)$ Gruppen sind, dann kann man auf $\prod_{i=1}^n G_i$ eine Gruppenstruktur definieren, indem man
$$(g_1, \ldots, g_n) * (h_1, \ldots, h_n) := (g_1 *_1 h_1, \ldots, g_n *_n h_n)$$
5. $(\mathbb{R}_{>0}, \cdot)$ und $(\mathbb{R}/\mathbb{Z}, +) \simeq S^1$ sind auch Gruppen, und $\mathbb{R}_{>0} \times S^1 \simeq \mathbb{C}^\times$. Dieser Isomorphismus entspricht der Polardarstellung der komplexen Zahlen.

1.6.6 Die Ordnung eines Elementes

Sei $(G, \cdot)$ eine Gruppe mit neutralem Element e und sei $g \in G$ ein beliebiges Element.

Bemerkung 1.107. Wenn die Menge $\{ g^i \ : \ i \in \mathbb{N}_{>0} \}$ endlich ist, dann existiert ein $d \in \mathbb{N}_{>0}$, sodass

$$g^d = e.$$

Weil die Menge $\{ g^i \ : \ i \in \mathbb{N}_{>0} \}$ endlich ist, können nicht alle Potenzen von g paarweise unterschiedlich sein. Es existieren also $n, m \in \mathbb{N}_{>0}$ mit $n \neq m$ und $g^n = g^m$. Wir dürfen annehmen, dass $n > m$. Wenn wir also beide Seiten von $g^n = g^m$ mit $g^{-m} = (g^m)^{-1}$ multiplizieren, dann bekommen wir

$$g^{n-m} = g^{n-n} = e.$$

Definition 1.108. Für jedes Element $g \in G$ ist die **Ordnung des Elementes** g

$$\operatorname{ord} g = \begin{cases} \min \{ k \in \mathbb{N}_{>0} \ : \ g^k = e \} & , \text{ wenn } \{ k \in \mathbb{N}_{>0} \ : \ g^k = e \} \neq \emptyset, \\ \infty & , \text{ sonst.} \end{cases}$$

Proposition 1.109. *Sei $(G, \cdot)$ eine Gruppe und $g \in G$ mit $\operatorname{ord} g < \infty$. Die Abbildung $f_g : \mathbb{Z} \longrightarrow G$ definiert durch*

$$f_g(n) = g^n$$

ist ein Gruppenhomomoprhismus mit $\operatorname{Ker} f_g = (\operatorname{ord} g)\mathbb{Z}$.

Beweis-Skizze: Aus Lemma 1.104 ist $\operatorname{Ker} f_g$ eine Untergruppe von $\mathbb{Z}$. Aus Satz 1.105 existiert also ein $d \in \mathbb{N}_{>0}$, sodass $\operatorname{Ker} f_g = d\mathbb{Z}$. In dem Beweis des Satzes haben auch gesehen, dass $d = \min \{ n \in \mathbb{Z}_{>0} \ : \ f_g(d) = 0 \}$. Das ist per Definition die Ordnung von g. Q.E.D.

Korollar 1.110. Sei $(G, \cdot)$ eine Gruppe, $g \in G$ und $m \in \mathbb{Z}$. Wenn $g^m = e$, dann gilt $\operatorname{ord} g \,|\, m$.

1.6.7 Die Symmetrische Gruppe

Sei $n \geqslant 1$. Wir haben schon in Definition 1.96 auf Seite 59 die symmetrische Gruppe S_n als die Menge aller Bijektionen $\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ zusammen mit der Verknüpfung von Abbildungen als Gruppenoperation definiert. Ein Element $\sigma \in S_n$ heißt Permutation und wird aufgeschrieben als

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Ein einfacher induktiver Beweis, oder direktes Zählen, geben uns die Kardinalität von S_n:

$$\# S_n = n! = 1 \cdot \cdots \cdot n.$$

Insbesondere, haben S_1 und S_2 jeweils 1, beziehungsweise 2 Elementen. Das heißt, dass beide abelsche Gruppen sind. Wir haben gesehen (oder werden gesehen haben), dass Gruppen mit p Elementen, wenn p eine Primzahl ist, zyklisch, und somit kommutativ, sind. Außerdem, gibt es bis auf Isomorphismus nur zwei Gruppen mit 4 Elementen: $\mathbb{Z}/4\mathbb{Z}$ und $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Das heißt, dass wenn $\#G < 5$, dann ist G eine abelsche Gruppe. Die Gruppe S_3 hat $3! = 1 \cdot 2 \cdot 3 = 6$ Elemente, und ist nicht kommutativ. Es hat die Elemente:

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Die Gruppe S_3 ist nicht kommutativ, weil

$$\sigma_1 \sigma_2 = \sigma_5 \neq \sigma_4 = \sigma_2 \sigma_1.$$

Das heißt, dass S_3 die kleinste[28] nicht-kommutative Gruppe ist. Für jedes $n \geqslant 4$ können wir auch Elemente $\sigma_1, \ldots, \sigma_5 \in S_n$ finden, die dieselbe Wirkung auf $1, 2, 3$ wie die $\sigma_k \in S_3$ haben, und alle $i \geqslant 4$ auf sich selbst abgebildet werden. Das beweist die folgende Bemerkung.

[28] Eigentlich haben wir nur gezeigt, dass es keine nicht-kommutative Gruppe mit weniger Elementen gibt. Es könnte aber theoretisch auch andere, nicht zu S_3 isomorphe Gruppen mit 6 Elementen geben, die auch nicht kommutativ sind. Wir werden aber sehen, dass die einzige andere Gruppe mit 6 Elementen $\mathbb{Z}/6\mathbb{Z}$ ist.

Bemerkung 1.111. Die Gruppe S_n ist genau dann kommutativ, wenn $n \geqslant 2$.

Sei $k \in \mathbb{N}_{>1}$. Ein **k-Zyklus** (oder zyklische Permutation der Länge k) ist eine Permutation $\sigma \in S_n$ mit der Eigenschaft, dass es paarweise unterschiedliche $i_1, \ldots, i_k \in \{1, \ldots, n\}$ gibt, sodass

$$\sigma(i_j) = i_{j+1} \quad \forall\, j = 1, \ldots, k \text{ (wobei die Indizes modulo } k \text{ verstanden werden}^{29}).$$

und $\sigma(\ell) = \ell$ für alle $\ell \in \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$. Wir schreiben in diesem Fall

$$\sigma =: (i_1\ i_2\ \ldots\ i_k).$$

In S_3 haben wir also außer der Identität drei 2-Zyklen: (12), (13), und (23) und zwei 3-Zyklen: (123) und (132). Man muss aufpassen, dass die Notation der Zyklen nicht eindeutig ist:

$$(123) = (231) = (312).$$

Aus der Zyklus-Notation allein ist auch nicht klar ob der obige 3-Zyklus in S_3, S_4 oder S_{101} lebt. Das sollte man vorher klar machen.

Zwei Zyklen $\gamma_1 = (i_1 \ldots i_r)$ und $\gamma_2 = (j_1 \ldots j_s)$ sind **disjunkt** wenn

$$\{\, i_1, \ldots, i_r \,\} \cap \{\, j_1, \ldots, j_s \,\} = \emptyset.$$

Für jede Permutation $\sigma \in S_n$ und jedes $i \in \{\, 1, \ldots, n \,\}$ definieren wir die **Bahn von i unter σ** als die Menge

$$\mathrm{Bahn}_\sigma(i) = \left\{\, \sigma^k(i) \;\; : \;\; k \in \mathbb{N} \,\right\}.$$

Wir nennen eine Bahn *nichttrivial* wenn $\#\,\mathrm{Bahn} > 1$.

Bemerkung 1.112. Für $i, j \in \{\, 1, \ldots, n \,\}$ gilt $\mathrm{Bahn}_\sigma(i) = \mathrm{Bahn}_\sigma(j)$ oder $\mathrm{Bahn}_\sigma(i) \cap \mathrm{Bahn}_\sigma(j) = \emptyset$. Daraus folgt, dass jede Permutation eine Äquivalenzrelation auf $\{\, 1, \ldots, n \,\}$ definiert:

$$i \sim_\sigma j \iff \mathrm{Bahn}_\sigma(i) = \mathrm{Bahn}_\sigma(j).$$

Die Äquivalenzklassen sind dann genau die Bahnen.

Beweis-Skizze: Es reicht zu zeigen, dass

$$\mathrm{Bahn}_\sigma(i) \cap \mathrm{Bahn}_\sigma(j) \neq \emptyset \quad \Rightarrow \quad \mathrm{Bahn}_\sigma(i) = \mathrm{Bahn}_\sigma(j).$$

Sei $k \in \mathrm{Bahn}_\sigma(i) \cap \mathrm{Bahn}_\sigma(j)$. Das heißt, es existieren $r, s \in \mathbb{N}_{>0}$, sodass

$$\sigma^r(i) = k = \sigma^s(j).$$

Daraus folgt, dass $\sigma^{r+(\mathrm{ord}\,\sigma - s)}(i) = \sigma^{\mathrm{ord}\,\sigma}(j) = \mathrm{id}(j) = j$. Wir bezeichnen mit $m := r + \mathrm{ord}\,\sigma - s$ und zeigen, dass $\mathrm{Bahn}_\sigma(j) \subseteq \mathrm{Bahn}_\sigma(i)$.

Sei also $a \in \mathrm{Bahn}_\sigma(j)$ beliebig. Es existiert dann ein $\ell \in \mathbb{N}_{>0}$, sodass $\sigma^\ell(j) = a$. Es folgt

$$a = \sigma^\ell(j) = \sigma^\ell(\sigma^m(i)) = \sigma^{\ell+m}(i) \in \mathrm{Bahn}_\sigma(i).$$

[29] Das heißt, dass wir $k+1$ und 1 identifizieren. Das spart uns die Fallunterscheidung: $\sigma(i_j) = i_{j+1}$ für $j = 1, \ldots, k-1$ und $\sigma(i_k) = i_1$.

Also $\text{Bahn}_\sigma(j) \subseteq \text{Bahn}_\sigma(i)$. Analog beweist man die andere Inklusion, also $\text{Bahn}_\sigma(i) = \text{Bahn}_\sigma(j)$.
Q.E.D.

Für ein k-Zyklus $\gamma = (i_1 \ \dots \ i_k)$, mit $k \geqslant 2$, dann hat γ genau eine nichttriviale Bahn. Es gilt

$$\text{Bahn}_\gamma(i) = \begin{cases} \{ i_1, \dots, i_k \} & , \text{ wenn } i \in \{ i_1, \dots, i_k \} \\ \{ i \} & , \text{ wenn } i \notin \{ i_1, \dots, i_k \}. \end{cases}$$

Wir bezeichnen die eindeutige Bahn eines Zyklus γ mit

$$B(\gamma) = \{ i \ : \ \gamma(i) \neq i \}.$$

Die eindeutige Bahn bestimmt aber nicht den Zyklus. Disjunkte Zyklen kommutieren. Das folgt aus der allgemeineren Bemerkung

Bemerkung 1.113. Wenn $\sigma = \gamma_1 \cdots \gamma_r$ mit $B(\gamma_k) \cap B(\gamma_\ell) = \emptyset$ für $k \neq \ell$, dann gilt

$$\forall\, i \ \ \exists!\, k \quad \gamma_k(i) = \sigma(i) \qquad \text{und} \qquad \gamma_\ell(i) = i \quad \text{if } \ell \neq k.$$

Insbesondere, wenn γ und α disjunkte Zyklen sind, dann gilt

$$\gamma \cdot \alpha = \alpha \cdot \gamma.$$

Beweis-Skizze: Wenn $\sigma(i) \neq i$, aber es existiert nicht ein eindeutiges k wie oben, dann gibt es zwei Möglichkeiten:

Fall 1: Entweder $\gamma_k(i) = i$ für alle k. Aber dann gilt $(\gamma_1 \cdots \gamma_r)(i) = i \neq \sigma(i)$ - ↯.

Fall 2: Es existieren $k \neq \ell$ mit $\gamma_k(i) \neq i \neq \gamma_\ell(i)$. Aber das ist ein Widerspruch weil die Zyklen disjunkt sind.
Q.E.D.

Satz 1.114. *Es sei $n \in \mathbb{N}_{>0}$. Für jede Permutation $\sigma \in S_n$ existieren eindeutige disjunkte Zyklen $\gamma_1, \dots, \gamma_r$, sodass*

$$\sigma = \gamma_1 \cdots \gamma_r.$$

Beweis-Skizze: Existenz.

Variante 1: Für jede Bahn die mehr als 1 Element hat, definieren wir einen Zyklus. Genauer gesagt, seien $\text{Bahn}_\sigma(i_1), \dots, \text{Bahn}_\sigma(i_r)$ alle unterschiedliche Bahnen mit $b_k := \#\,\text{Bahn}_\sigma(i_k) > 1$. Wir definieren dann

$$\gamma_k := (i_k \ \sigma(i_k) \ \dots \ \sigma^{b_k - 1}(i_k)).$$

Nach Bemerkung 1.112 sind die Zyklen $\gamma_1, \dots, \gamma_r$ disjunkt. Für jedes $i \in \{ 1, \dots, n \}$ haben wir:

Fall 1: Wenn $\#\,\text{Bahn}_\sigma(i) = 1$, dann gilt $\sigma(i) = i$ und auch $\gamma_k(i) = i$ für alle $k = 1, \dots, r$. Also

$$\sigma(i) = (\gamma_1 \cdots \gamma_r)(i).$$

Fall 2: Wenn $\#Bahn_\sigma(i) > 1$, dann existiert nach Bemerkung 1.112 ein einziges $k \in \{ 1, \dots, r \}$, sodass $\text{Bahn}_\sigma(i) = \text{Bahn}_\sigma(i_k)$. Das heißt, es existiert auch ein $j \in \{ 0, \dots, b_k - 1 \}$, sodass $i = \sigma^j(i_k)$. Daraus folgt,

$$\sigma(i) = \sigma(\sigma^j(i_k) = \sigma^{j+1}(i_k) = \gamma_k(i),$$

und $\gamma_\ell(i) = i$ für alle $\ell \neq k$. Es gilt also

$$\sigma(i) = (\gamma_1 \cdots \gamma_r)(i).$$

Variante 2: Man kann die Existenz auch durch Induktion beweisen. Dafür werden wir erstens für jede Permutation σ die Menge der von σ bewegten Elementen definieren:

$$B(\sigma) = \bigcup_{\#\,\mathrm{Bahn}_\sigma(i)>1} \mathrm{Bahn}_\sigma(i).$$

Das heißt, für ein $i \in \{\, 1, \ldots, n \,\}$ gilt

$$\sigma(i) \notin i \iff i \notin B(\sigma).$$

Wir werden dann n fixieren und dann die Aussage für alle $\sigma \in S_n$ durch Induktion nach $\#B(\sigma)$ beweisen. Die Aussage ist:

$$\mathcal{A}(k): \quad \#B(\sigma) = k \Rightarrow \ \sigma \text{ ist Produkt von disjunkten Zyklen.}$$

$\boxed{k = 0}$ Wenn $k = 0$, dann gilt $\sigma(i) = i$ für alle i, also $\sigma = \mathrm{id}$. Die Identität ist das leere Produkt disjunkter Zyklen.

$\boxed{\mathcal{A}(j) \ \ \forall j \leqslant k \Rightarrow \mathcal{A}(k+1)}$ Wir verwenden also die starke Induktion (siehe Bemerkung **??**). Wir suchen dann das kleinste $i \in \{\, 1, \ldots, n \,\}$ mit $\sigma(i) \neq i$. Es gilt dann $\mathrm{Bahn}_\sigma(i) = \{\, i, \sigma(i), \ldots, \sigma^{b-1}(i) \,\}$, mit $b > 1$ und $\#\,\mathrm{Bahn}_\sigma(i) = b$. Wir definieren dann den b-Zyklus $\gamma \in \S_n$ durch

$$\gamma = (i \ \sigma(i) \ \ldots \ \sigma^{b-1}(i)).$$

Wir definieren auch $\tau \in S_n$ durch

$$\tau(j) = \begin{cases} \sigma(j) & , \text{ wenn } j \notin \mathrm{Bahn}_\sigma(i) \\ j & , \text{ wenn } j \in \mathrm{Bahn}_\sigma(i). \end{cases}$$

Es gilt dann $\#B(\tau) = \#B(\sigma) - b \geqslant k$ und $\mathbb{B}(\tau) \cap B(\gamma) = \emptyset$. Es gilt auch

$$\sigma = \gamma \cdot \tau,$$

also aus der induktiven Voraussetzung existieren $\gamma_2, \ldots, \gamma_r$ disjunkte Zyklen mit $\tau = \gamma_2 \cdots \gamma_r$. Es gilt auch $B(\gamma_k) \subseteq B(\tau)$, also $\gamma \cap \gamma_k = \emptyset$ für alle $k = 2, \ldots, r$. Somit haben wir

$$\sigma = \gamma \cdot \gamma_2 \cdots \gamma_r$$

ist ein Produkt disjunkter Zyklen.

Eindeutigkeit.

Jeder Zyklus hat genau eine nichttriviale Bahn. Das Produkt von r disjunkten Zyklen hat genau die r entsprechenden nichttriviale Bahnen. Also wenn $\gamma_1 \cdots \gamma_r = \sigma = \alpha_1 \cdots \alpha_s$, dann muss $r = s$ und $\text{Bahn}_{\gamma_k} = \text{Bahn}_{\alpha_k}$ gelten. Für jedes $i \in \{1, \ldots, n\}$ mit $\sigma(i) \neq i$ ein einziges k existiert mit $\sigma(i) = \gamma_k(i) = \alpha_k(i)$ und für alle anderen $\ell \neq k$ gilt $\gamma_\ell(i) = \alpha_\ell(i) = i$. Es folgt daraus, dass $\gamma_k = \alpha_k$ für alle $k = 1, \ldots, r$ und somit die Eindeutigkeit. Q.E.D.

Lemma 1.115. *Für jeden k-Zyklus $\gamma = (i_1 \ \ldots \ i_k) \in S_n$ gilt*

$$\gamma = (i_1 \ i_2) \cdots (i_{k-1} \ i_k).$$

Beweis-Skizze: Übung Q.E.D.

Lemma 1.116. *Für jede Transposition $(i \ j) \in S_n$ können wir annehmen, dass $i < j$ und es gilt*

$$\begin{aligned} (i \ j) &= (i \ i+1) \cdots (j-1 j) \\ &= (1 \ j) \cdot (1 \ j-1) \cdots (1 \ i) \end{aligned}$$

Korollar 1.117. Jede Permutation $\sigma \in S_n$ kann sowohl als Produkt der Transpositionen $(1\ 2), \ldots, (n-1\ n)$, als auch als Produkt der Transpositionen $(1\ 2), \ldots, (1, n)$ geschrieben werden.

Die Darstellungen aus dem obigen Korollar sind nicht unbedingt eindeutig. Zum Beispiel

$$\begin{aligned} (1\ 2\ 3) &= (1\ 2)(2\ 3) = (2\ 3)(1\ 2)(2\ 3)(1\ 2) \\ &= (1\ 3)(1\ 2) = (1\ 2)(1\ 3)(1\ 2)(1\ 3) \end{aligned}$$

Die Signatur einer Permutation

Auch wenn die Anzahl von Transpositionen nicht eindeutig ist, wir werden gleich sehen, dass die Parität der Anzahl von Transpositionen konstant ist. Deswegen werden wir sagen, dass eine Permutation **gerade** ist, wenn die Anzahl der Faktoren in einer Zerlegung als Produkt von Permutationen gerade ist. Wenn diese Anzahl ungerade ist, dann sagen wir dass die Permutation **ungerade** ist. Man muss aber zu erst beweisen, dass diese Definitionen Sinn[31] haben. Dafür führen wir folgender Begriff ein.

Definition 1.118. Das **Vorzeichen** (oder das **Signum** oder die **Signatur** oder die **Parität**) einer Permutation $\sigma \in S_n$ ist

$$\text{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Das ist intrinsisch definiert, also man muss sich keine Sorgen machen, dass es Sinn macht.
Man kann das Signum auch durch über folgender Begriff beschreiben,

Definition 1.119. Sei $\sigma \in S_n$. Ein **Fehlstand** (oder eine **Inversion**) ist ein geordnetes Paar $(i, j) \in \{1, \ldots, n\}^2$ mit $i < j$ und $\sigma(i) > \sigma(j)$. Wir bezeichnen die Menge aller Fehlstände von σ mit

$$\text{inv}(\sigma) = \{(i, j) \mid 1 \leq i < j \leq n \text{ und } \sigma(i) > \sigma(j)\}.$$

[31] Das heißt, dass diese Parität wirklich invariant für jede Permutation ist.

Übung. Zeigen[32] Sie, dass für $\sigma \in S_n$ gilt $\operatorname{sgn}(\sigma) = (-1)^{|\operatorname{inv}(\sigma)|}$.

Bemerkung 1.120. Für die Identität $\mathrm{id}_n \in S_n$ gilt $\operatorname{sgn}(\mathrm{id}_n) = 1$.

Satz 1.121. *Die Abbildung* $\operatorname{sgn} : S_n \longrightarrow \{-1, 1\}$ *ist ein Gruppenhomomorphismus, wobei die Gruppenoperation auf* $\{-1, 1\}$ *die Multiplikation ist.*

Beweis-Skizze: Wir müssen also zeigen, dass für alle $\sigma, \pi \in S_n$ gilt $\operatorname{sgn}(\sigma \cdot \pi) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\pi)$. Der Trick ist der folgende:

$$\begin{aligned} \operatorname{sgn}(\sigma \cdot \pi) &= \prod_{i<j} \frac{\sigma\big(\pi(j)\big) - \sigma\big(\pi(i)\big)}{j - i} \\ &= \prod_{i<j} \frac{\sigma\big(\pi(j)\big) - \sigma\big(\pi(i)\big)}{\pi(j) - \pi(i)} \cdot \prod_{i<j} \frac{\pi(j) - \pi(i)}{j - i} \end{aligned}$$

Das zweite Produkt ist per Definition $\operatorname{sgn}(\pi)$. Wir müssen nur noch bemerken, dass das erste Produkt $\operatorname{sgn}(\sigma)$ ist. Dafür behaupten wir, dass jeder Bruch ein Bruch aus der Definition von $\operatorname{sgn}(\sigma)$ ist. Wenn wir $k := \pi(j)$ und $\ell := \pi(i)$ definieren, dann ist das einzige das stören könnte, dass $k < \ell$ vorkommen könnte. Aber das kann gleich wieder gut gemacht werden, weil

$$\frac{\sigma(k) - \sigma(\ell)}{k - \ell} = \frac{\sigma(\ell) - \sigma(k)}{\ell - k}.$$

Q.E.D.

Korollar 1.122. Wenn $\sigma \in S_n$, dann gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Bemerkung 1.123. Für eine Transposition $\tau = (i, j) \in S_n$ gilt $\operatorname{sgn}(\tau) = -1$.

Beweis-Skizze: Wir verwenden die obige Beschreibung des Signums als $(-1)^{|\operatorname{inv}(\sigma)|}$. Das heißt, das $\operatorname{sgn}(1\ 2) = -1$, weil es eine einzige Inversion hat.

Das kann man auch direkt beweisen. Bezeichne $\tau = (1\ 2)$. Es gilt dann

$$\begin{aligned} \operatorname{sgn}(\tau) &= \textstyle\prod_{1 \leqslant i < j \leqslant n} \frac{\tau(j) - \tau(i)}{j - i} \\ &= \left(\textstyle\prod_{\substack{i=1 \\ j=2}} \frac{\tau(j) - \tau(i)}{j - i}\right) \cdot \left(\textstyle\prod_{\substack{i=1 \\ 3 \leqslant j \leqslant n}} \frac{\tau(j) - \tau(i)}{j - i}\right) \cdot \left(\textstyle\prod_{\substack{i=2 \\ 3 \leqslant j \leqslant n}} \frac{\tau(j) - \tau(i)}{j - i}\right) \cdot \left(\textstyle\prod_{3 \leqslant i < j \leqslant n} \frac{\tau(j) - \tau(i)}{j - i}\right) \\ &= \left(\tfrac{1-2}{2-1}\right) \cdot \left(\textstyle\prod_{3 \leqslant j \leqslant n} \frac{j-2}{j-1}\right) \cdot \left(\textstyle\prod_{3 \leqslant j \leqslant n} \frac{j-1}{j-2}\right) \cdot \left(\textstyle\prod_{3 \leqslant i < j \leqslant n} \frac{\cancel{j-i}}{\cancel{j-i}}\right) \\ &= (-1) \cdot \left(\textstyle\prod_{3 \leqslant j \leqslant n} \frac{\cancel{j-2}}{\cancel{j-1}} \cdot \frac{\cancel{j-1}}{\cancel{j-2}}\right) \cdot \left(\textstyle\prod_{3 \leqslant i,j \leqslant n} 1\right) \end{aligned}$$

Eine Beliebige Transposition $(i\ j)$ ist gleich[a] mit

$$(i\ j) = (1\ i) \cdot (2\ j) \cdot\ (1\ 2) \cdot (2\ j) \cdot (1\ i).$$

[32] Sie finden einen Eleganten Beweis dafür in [Fis09, S.288]

Aus Satz 1.121 und Korollar 7.31 folgt, weil $(k\ \ell)^{-1} = (k\ \ell)$

$$\begin{aligned}\operatorname{sgn}(i\ j) &= \operatorname{sgn}(1\ i)\cdot\operatorname{sgn}(2\ j)\cdot\ \operatorname{sgn}(1\ 2)\cdot\operatorname{sgn}(2\ j)^{-1}\cdot\operatorname{sgn}(1\ i)^{-1}\\ &= \operatorname{sgn}(1\ 2)\cdot\operatorname{sgn}(1\ i)\cdot\operatorname{sgn}(2\ j)\cdot\operatorname{sgn}(2\ j)^{-1}\cdot\operatorname{sgn}(1\ i)^{-1}\\ &= -1.\end{aligned}$$

Q.E.D.

[a] wenn $i = 1$ oder $j = 2$, dann verwenden wir die Konvention, dass $(a\ a) = \mathrm{id}$.

Wir kommen endlich zur Invarianz der Parität der Anzahl von Transpositionen in der Zerlegung einer Permutation.

Korollar 1.124. (a) Wenn $\sigma \in S_n$ das Produkt der Transpositionen $\tau_1, \ldots, \tau_r \in S_n$ ist, dann gilt $\operatorname{sgn}(\sigma) = (-1)^r$.

(b) Wenn für $\sigma \in S_n$ gilt $\sigma = \tau_1 \cdot \cdots \cdot \tau_r = \tau'_1 \cdot \cdots \cdot \tau'_s$, mit τ_i und τ'_j Transpositionen für alle $i = 1, \ldots, r$ und alle $j = 1, \ldots, s$, dann gilt

$$r \equiv s \bmod 2.$$

(c) Die Menge $A_n = \{\, \sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1 \,\} = \ker \operatorname{sgn}$ ist eine Untergruppe von S_n.

Die Untergruppe A_n heißt die **alternierende Untergruppe** von S_n.

Bemerkung 1.125. Für jede Transposition $\tau = (i, j) \in S_n$ haben wir die disjunkte Vereinigung

$$S_n = A_n \sqcup \tau A_n.$$

Weil $\operatorname{sgn}\tau = -1$, gilt auch $\langle\tau\rangle \cap A_n = \{\, \mathrm{id} \,\}$. Wenn $\langle\tau\rangle = \{\, \mathrm{id}, \tau \,\}$ die Untergruppe die von τ erzeugt wird bezeichnet, dann gilt

$$S_n = A_n \cdot \langle\tau\rangle = \{\, \alpha \cdot \beta \ : \ \alpha \in A_n, \ \ \beta \in \langle\tau\rangle \,\}.$$

Um das zu sehen, müssen wir wenn $\sigma \in A_n$, dann kann man $\alpha = \sigma$ und $\beta = \mathrm{id}$ wählen. Wenn $\sigma \notin A_n$, dann gilt $\sigma\tau \in \mathbb{A}_n$ und wir können dann $\alpha = \sigma\tau$ und $\beta = \tau$ wählen. Anders gesagt, ist S_n das semidirekte Produkt[33] der Untergruppen A_n und $\operatorname{Span}_{\mathbb{K}}\{\tau\}$. Das wird als $S_n = A_n \rtimes \operatorname{Span}_{\mathbb{K}}\{\tau\}$ geschrieben.

Wir haben die **Alternierende Untergruppe** $A_n = \{\, \sigma \in S_n \ : \ \operatorname{sign}(\sigma) = 1 \,\}$ hat Index $[S_n : A_n] = 2$. Sei $\{\, \pm 1 \,\}$ die Gruppe mit zwei Elementen unter multiplikativen Bezeichnung. Die Abbildung $\operatorname{sign} : S_n \longrightarrow \{\, \pm 1 \,\}$ ist ein Gruppenhomomorphismus[34]. Die Alternierende Gruppe ist dann der Kern des Homomorphismus sign.

Der Satz von Cayley

Satz 1.126 (von Cayley[35])**.** *Für jede endliche Gruppe G existiert ein $n \in \mathbb{N}$, sodass G isomorph zu einer Untergruppe von S_n ist.*

[33] Der Begriff von semiderektes Produkt wurde nicht eingeführt. Machen Sie sich keine weitere Gedanken darüber.
[34] Das braucht einen kleinen Beweis
[35] Englischer Mathematiker (1821-1895)

Beweis-Skizze: Jedes Element g einer Gruppe G definiert eine bijektive Abbildung definiert: $m_g : G \longrightarrow G$ durch

$$x \mapsto m_g(x) := g \cdot x.$$

Diese Abbildung ist bijektiv, weil das Inverse Element von g die Inverse Abbildung definiert: $m_{g^{-1}} = (m_g)^{-1}$. Wir haben also eine Abbildung $\Phi : G \longrightarrow S_{\#G}$ gegen durch

$$g \mapsto m_g.$$

Aus der Assoziativität der Gruppenoperation von G haben wir für alle $g, h \in \mathbb{G}$:

$$m_{g \cdot h}(x) = (g \cdot h) \cdot x = g \cdot (h \cdot x) = m_g(m_h(x)) = (m_g \circ m_h)(x) \quad \forall\, x \in G.$$

Also Φ ist ein Gruppenhomomorphismus. Weiterhin, es gilt

$$m_g(x) = \mathrm{id}(x) \quad \forall\, x \in G \Rightarrow m_g(e) = g \cdot e = e \Rightarrow g = e.$$

Also $\operatorname{Ker} \Phi = \{\, e \,\}$ und somit ist Φ injektiv. Also $G \simeq \operatorname{Bild} \Phi \leq S_{\#G}$. Q.E.D.

Beispiel 1.127.

Wenn $G = \mathbb{Z}/3\mathbb{Z}$ dann haben wir $\Phi : G \longrightarrow S_3$ definiert durch durch

$$1 \;\mapsto\; \{\, 1,2,3 \,\} \xrightarrow{+0} \{\, 1,2,3 \,\} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \mathrm{id},$$

$$2 \;\mapsto\; \{\, 1,2,3 \,\} \xrightarrow{+1} \{\, 1,2,3 \,\} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1\ 2\ 3),$$

$$3 \;\mapsto\; \{\, 1,2,3 \,\} \xrightarrow{+2} \{\, 1,2,3 \,\} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1\ 3\ 2).$$

Also die Untergruppe von S_3 ist $\left\{\, \mathrm{id}, \gamma, \gamma^2 \,\right\}$, wobei γ der 3-Zyklus (1 2 3) ist.

1.6.8 Erzeuger von Gruppen

Wir beweisen erstmals ein Lemma, das uns garantiert, dass "die von einer Teilmenge erzeugte Untergruppe", tatsächlich eine Untergruppe sein wird.

Lemma 1.128. *Sei $(G, \cdot)$ eine Gruppe. Wenn $(H_i)_{i \in I}$ eine Familie von Untergruppen*[36] *ist, dann gilt*

$$\bigcap_{i \in I} H_i \leq G.$$

[36] Das heißt, dass $H_i \leq G$ für alle $i \in I$.

Beweis-Skizze: Wir werden Satz 1.101 zwei Mal anwenden.

Seien $a, b \in \bigcap_{i\in I} H_i$ beliebig. Das heißt,

$$a, b \in H_i \quad \forall\, i \in I.$$

Weil $H_i \leq G$ für alle $i \in I$, folgt aus Satz 1.101, dass $ab^{-1} \in H_i$ für alle $i \in I$. Also, aus der Definition des Durchschnittes einer Familie (siehe Definition **??**) haben wir

$$ab^{-1} \in \bigcap_{i\in I} H_i.$$

Q.E.D.

Die folgende Definition ist nicht spezifisch für die Gruppentheorie. In vielen anderen algebraischen oder geometrischen Strukturen[37], kann man über die von einer Menge erzeugte Unterstruktur sprechen. In der linearen Algebra lernt man schon am Anfang des Studiums über den Untervektorraum, der von einer Menge von Vektoren erzeugt ist. Die Idee wird auch dort dieselbe sein: Wenn S eine beliebige Teilmenge ist, dann ist die davon erzeugte Unterstruktur, die **kleinste** Unterstruktur die diese Menge enthält. Für Gruppen wird also "die von S erzeugte Untergruppe" die kleinste Untergruppe die S enthält sein. "Kleinste" bezieht sich hier auf der Mengeninklusion. Das heißt, es wird die Untergruppe H_0 sein, mit $S \subseteq H_0$ und wenn $S \subseteq H$ für eine Untergruppe H gilt, dann gilt auch $H_0 \subseteq H$. Das kann man genauer in der folgenden Form ausdrücken.

Definition 1.129. Sei $(G, \cdot)$ eine Gruppe und sei $S \subseteq G$ eine beliebige Teilmenge der unterliegenden Menge G. Die **von S erzeugte Untergruppe** von G ist die Untergruppe

$$\langle S\rangle = \bigcap_{\substack{H\leq G\\ S\subseteq H}} H.$$

Es gilt insbesondere, dass $\langle\emptyset\rangle = \bigcap_{H\leq G} = \{\, e \,\}$, die Gruppe mit einem Element.

Die Definition von $\langle S\rangle$ sagt nichts über die Form der Elemente der Untergruppe. Wir werden aber gleich beweisen, dass die von S erzeugte Untergruppe genau die Elementen, die durch endlich viele Verknüpfungen von Elementen aus S und deren Inversen erhalten werden können, enthält.

Satz 1.130. *Sei $(G, \cdot)$ und $S \subseteq G$ eine Teilmenge. Man bezeichne mit $S^{-1} = \left\{\, s^{-1} \;:\; s \in S \,\right\}$. Es gilt*

$$\langle S\rangle = \left\{\, \textstyle\prod_{i=1}^{r} s_i \;:\; r \in \mathbb{N} \text{ und } s_i \in S \cup S^{-1} \;\; \forall\, i \,\right\}.$$

Wir erinnern hier, dass wenn $r = 0$, dann ist $\prod_{i=1}^{0} s_i$ das leere Produkt, also gleich mit dem neutralen Element von G.

Beweis-Skizze: Wir bezeichnen mit $\mathcal{P} = \left\{\, \prod_{i=1}^{r} s_i \;:\; r \in \mathbb{N} \text{ und } s_i \in S \cup S^{-1} \;\; \forall\, i \,\right\}$. Wir wollen also zeigen, dass

$$\mathcal{P} = \bigcap_{\substack{H\leq G\\ S\subseteq H}} H.$$

[37] Zum Beispiel in Ringen,in Moduln, in Algebren, in Vektorräumen, in affinen Räumen, in projektive Räumen

Wenn wir $r = 1$ und $s_1 = s \in S$ setzen, dann bekommen wir $s \in \mathcal{P}$ für alle $s \in S$. Also $S \subseteq \mathcal{P}$. Wenn wir beweisen, dass $\mathcal{P} \leq G$, dann folgt $\langle S \rangle \subseteq \mathcal{P}$. Wenn wir auch beweisen, dass $\mathcal{P} \subseteq H$ für alle $H \leq G$ mit $S \subseteq H$, dann bekommen wir auch $\mathcal{P} \subseteq \bigcap_{\substack{H \leq G \\ S \subseteq H}} H = \langle S \rangle$. Es reicht also folgende Aussagen zu beweisen:

1. $\mathcal{P} \leq G$.
2. $\mathcal{P} \subseteq H$ für alle $H \leq G$ mit $S \subseteq H$.

1. Wir verwenden wieder den Satz 1.101. Es seien $a, b \in \mathcal{P}$. Das heißt es existieren $r, \ell \in \mathbb{N}$ und $s_1, \ldots, s_r, t_1, \ldots, t_\ell \in S \cup S^{-1}$ mit

$$a = s_1 \cdots s_r \qquad \text{und} \qquad b = t_1 \cdots t_\ell.$$

   Es gilt dann, $b^{-1} = t_\ell^{-1} \cdots t_1^{-1}$ und, weil $t_i \in S \cup S^{-1}$, auch $t_i^{-1} \in S \cup S^{-1}$. Also

$$ab^{-1} = s_1 \cdots s_r \cdot t_\ell^{-1} \cdots t_1^{-1} \in \mathcal{P}.$$

2. Sei $g = s_1 \cdots s_r \in \mathcal{P}$ beliebig und sei $H \leq G$ mit $S \subseteq H$. Weil $S \subseteq H \leq G$, gilt aus **UG 3.** $s^{-1} \in H$ für alle $s \in S$. Also $s_1, \ldots, s_r \in H$, und aus **UG 1.** folgt $g \in H$.

Q.E.D.

Die Elemente einer Menge S mit $\langle S \rangle = G$ heißen **Erzeuger** von G. Wir sagen auch, dass S ein **Erzeugendensystem** von G ist. Ein Erzeugendensystem S von G ist minimal, wenn alle echte Teilmengen von S keine Erzeugendensysteme von G sind. Also wenn

$$\langle S \rangle = G \qquad \text{und} \qquad \langle S' \rangle \neq G \ \ \forall\ S' \subsetneq S.$$

Wir sagen, dass die Gruppe G **endlich erzeugt** ist, wenn es eine endliche Menge $S \subseteq G$ gibt, die ein Erzeugendensystem von G ist. Eine Gruppe G heißt **zyklische Gruppe** wenn es von einem einzigen Element erzeugt werden kann.

Bemerkung 1.131. Jede Gruppe hat mindestens ein Erzeugendensystem: $\langle G \rangle = G$. Interessanter sind aber minimale Erzeugendensysteme. Es gibt meistens mehrere davon, und es kann auch passieren, dass minimale Erzeugendensysteme verschiedene Kardinalitäten haben. Zum Beispiel:

$$\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle = \langle 2, 3 \rangle = \langle 11, 25 \rangle .$$

Alle vier sind minimale Erzeugendensysteme. In den ersten zwei Fällen ist das sehr einfach: jede ganze Zahl ist eine endliche Summe von 1 und -1. Es gilt allgemeiner, dass wenn wir schon $G = \langle S \rangle$ wissen, dann gilt

$$G = \langle T \rangle \iff S \subseteq \langle T \rangle .$$

Die direkte Implikation ist trivial, und die Umkehrung folgt aus Satz 1.130. In unserem Fall gilt also

$$\mathbb{Z} = \langle a, b \rangle \iff 1 \in \langle a, b \rangle = \{\, \lambda \cdot a + \mu \cdot b \ \ : \ \ \lambda, \mu \in \mathbb{Z} \,\} .$$

Also, aus Korollar 1.74 gilt $\mathbb{Z} = \langle a, b \rangle \Leftrightarrow \operatorname{ggT}(a, b) = 1$. Wie würde man das für $\mathbb{Z} = \langle a_1, \ldots, a_n \rangle$ verallgemeinern? Wann ist $\{\, a_1, \ldots, a_n \,\}$ ein minimales Erzeugendensystem von $\mathbb{Z}$?

Bemerkung 1.132. Für jedes Element $g \in G$ gilt $\operatorname{ord} g = \# \langle g \rangle$.

In Korollar 1.117 haben wir also bewiesen, dass

$$S_n = \langle (1\ 2), \ldots, (n-1\ n) \rangle = \langle (1\ 2), \ldots, (1\ n) \rangle .$$

Es ist eine gute Übung zu überprüfen, dass beide der obigen Erzeugendensysteme minimal sind. Dabei könnte folgende Bemerkung helfen.

Bemerkung 1.133. Ein Erzeugendensystem $S \subseteq G$ ist genau dann minimal, wenn

$$s \notin \langle S \setminus s \rangle \quad \forall\ s \in S.$$

Also wenn $S = \{ s_1, \ldots, s_n \}$, dann ist S genau ein minimales Erzeugendensystem, wenn[38]

$$G = \langle s_1, \ldots, s_n \rangle \qquad \text{und} \qquad G \neq \langle s_1, \ldots, \widehat{s_i}, \ldots, s_n \rangle \quad \forall\ i = 1, \ldots, n.$$

1.6.9 Normalteiler und die Faktorgruppe

Wir werden in diesem Teil die Kongruenz modulo einer ganzen Zahl verallgemeinern. Wir haben gesehen, dass jede Untergruppe von $\mathbb{Z}$ die Form $n\mathbb{Z}$ hat. Wir haben mit $\mathbb{Z}/n\mathbb{Z}$ die Menge der Restklassen modulo n bezeichnet, weil es ein Sonderfall folgender Relation ist.

Sei G eine Gruppe. Für jede Untergruppe $H \leq G$ definieren wir die **rechte Kongruenz Modulo** H als die Relation

$$a \sim_H b \iff ab^{-1} \in H.$$

Wenn $G = (\mathbb{Z}, +)$ und $H = n\mathbb{Z}$ dann ist "ab^{-1}" das Element $a - b$, und $a - b \in n\mathbb{Z}$ ist äquivalent zu $n \mid a - b$, also

$$a \sim_{n\mathbb{Z}} b \Leftrightarrow a \equiv b \bmod n.$$

Lemma 1.134. *Für alle $H \leq G$ ist die Relation $\sim_H$ ist eine Äquivalenzrelation. Die Äquivalenzklassen haben die Form $Hg := \{ hg \ : \ h \in H \}$, mit $g \in G$.*

Beweis-Skizze: Aus **UG 2.** folgt $gg^{-1} = e \in H$ für alle $g \in G$. Also $g \sim_H g$ für alle $g \in G$ und somit ist $\sim_H$ reflexiv.

Wenn $g_1 \sim_H g_2$, dann gilt per Definition $g_1 g_2^{-1} \in H$. Aus **UG 3.** folgt $(g_1 g_2^{-1})^{-1} = g_2 g_1^{-1} \in H$. Also $g_2 \sim_H g_1$ und somit ist $\sim_H$ symmetrisch.

Für die Transitivität, seien $g_1, g_2, g_3 \in G$ mit $g_1 \sim_H g_2$ und $g_2 \sim_H g_3$. Es gilt also

$$g_1 g_2^{-1} \in H \quad \text{und} \quad g_2 g_3^{-1} \in H.$$

Wir haben dann $g_1 g_3^{-1} = (g_1 e) g_3^{-1} = (g_1 (g_2^{-1} g_2)) g_3^{-1} = (g_1 g_2^{-1})(g_2 g_3^{-1}) \in H$.

[38] Die Bezeichnung $\{ s_1, \ldots, \widehat{s_i}, \ldots, s_n \}$ bedeutet $S \setminus \{ s_i \}$.

Sei $[g]_H$ eine Äquivalenzklasse für $\sim_H$. Das heißt,

$$\begin{aligned}
[g]_H &= \{\, x \in G \;:\; x \sim_H g \,\} \\
&= \{\, x \in G \;:\; xg^{-1} \in H \,\} \\
&= \{\, x \in G \;:\; \exists h \in H \text{ mit } xg^{-1} = h \,\} \\
&= \{\, x \in G \;:\; \exists h \in H \text{ mit } x = hg \,\} \\
&= \{\, hg \;:\; h \in H \,\}.
\end{aligned}$$

Q.E.D.

Für jede Untergruppe $H \leq G$, eine **Rechtsnebenklasse** von H ist eine Teilmenge von G der Form

$$Hg = \{\, hg \;:\; h \in H \,\}.$$

Für eine Untergruppe $H \leq G$ kann man auch eine linke Kongruenz Modulo H definieren

$$a \;{}_H\!\sim b \Leftrightarrow a^{-1}b \in H.$$

Analog zu dem Beweis von Lemma 1.134, zeigt man, dass auch ${}_H\!\sim$ eine Äquivalenzrelation auf G ist. Die Äquivalenzklasse in diesem Fall heißen **Linksnebenklassen** und sind Mengen der Form

$$gH = \{\, gh \;:\; h \in H \,\}$$

wobei $g \in G$. Wenn die Gruppe G kommutativ, dann gilt $gH = Hg$ für alle $g \in G$ und $H \leq H$. Für nicht-kommutative Gruppen gilt diese Gleichheit nicht immer.

Beispiel 1.135. In der symmetrischen Gruppe S_3 seien $H_2 = \langle(1\ 2)\rangle$ und $H_3 = \langle(1\ 2\ 3)\rangle$. Für H_2 haben wir folgende Nebenklassen:

$$\begin{array}{llllll}
H_2\,\mathrm{id} &= \{\,\mathrm{id}, (1\ 2)\,\} & H_2(1\ 3) &= \{\,(1\ 3), (1\ 3\ 2)\,\} & H_2(2\ 3) &= \{\,(2\ 3), (1\ 2\ 3)\,\} \\
\mathrm{id}\,H_2 &= \{\,\mathrm{id}, (1\ 2)\,\} & (1\ 3)H_2 &= \{\,(1\ 3), (1\ 2\ 3)\,\} & (2\ 3)H_2 &= \{\,(2\ 3), (1\ 3\ 2)\,\}
\end{array}$$

Es gilt also $H_2(1\ 3) \neq (1\ 3)H_2$ und $H_2(2\ 3) \neq (2\ 3)H_2$. Für H_3 haben wir nur zwei Nebenklassen: H_3 und $G \setminus H_3$.

$$\begin{aligned}
\mathrm{id}\,H_3 &= (1\ 2\ 3)H_3 = (1\ 3\ 2)H_3 &&= \{\,\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\,\}, \\
H_3\,\mathrm{id} &= H_3(1\ 2\ 3) = H_3(1\ 3\ 2) &&= \{\,\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\,\},
\end{aligned}$$

$$\begin{aligned}
(1\ 2)H_3 &= (1\ 3)H_3 = (2\ 3)H_3 &&= \{\,(1\ 2), (1\ 3), (2\ 3)\,\}, \\
H_3(1\ 2) &= H_3(1\ 3) = H_3(2\ 3) &&= \{\,(1\ 2), (1\ 3), (2\ 3)\,\}.
\end{aligned}$$

Genau wie wir Restklassen modulo n addieren können, wollen wir auch modulo einer Untergruppe rechnen. Das heißt, wir würden gerne eine innere Verknüpfung auf der Menge der rechts Nebenklassen, mit Hilfe der algebraischen Operation auf G, definieren. Das heißt, wir wollen, dass

$$(Hg) \cdot (Hg') := H(gg')$$

eine Wohldefinierte innere Verknüpfung ist. Das heißt, dass es unabhängig von der Wahl der Repräsentanten der Äquivalenzklassen ist. Wir werden jetzt sehen, dass das genau dann möglich ist, wenn die $\sim_H$ und ${}_H\!\sim$ dieselbe Äquivalenzrelation ist.

Lemma 1.136. *Für eine Untergruppe $H \leq G$ sind folgende Aussagen äquivalent.*

(i) $gH = Hg$ *für alle* $g \in G$.

(ii) *Für alle* $a, b, c, d \in G$ *gilt: wenn* $a \sim_H b$ *und* $c \sim_H d$, *dann gilt* $ac \sim_H bd$.

Beweis-Skizze: (i)⇒ (ii) Wir haben $ab^{-1}, cd^{-1} \in H$ und wollen zeigen, dass $ac(bd)^{-1} \in H$.

Aus $ab^{-1}, cd^{-1} \in H$ folgt die Existenz von $h_1, h_2 \in H$, sodass

$$a = h_1 b \qquad \text{und} \qquad c = h_2 d.$$

Wir suchen ein $h \in H$, sodass $ac = hbd$. Weil $bh_2 \in bH$ und aus (i) gilt $bH = Hb$, folgt, dass ein $h_3 \in H$ existiert, sodass $bh_2 = h_3 b$. Es gilt also

$$ac = (h_1 b)(h_2 d) = h_1(bh_2)d = h_1 h_3 (bd).$$

Weil $H \leq G$, gilt auch $h_1 h_3 \in H$, also $(ac)(bd)^{-1} \in H$. Also $ac \sim_H bd$.

(ii)⇒(i) Sei $g \in G$ beliebig. Wir zeigen $gH = Hg$ indem wir beide Inklusionen zeigen.

"$\subseteq$" Sei $x \in gH$ beliebig. Es gibt also ein $h \in H$ mit $x = gh$. Wir haben $g \sim_H g$ und $h \sim_H e$, also aus (ii) folgt $gh \sim_h g$. Das heißt, es existiert ein $h' \in H$, sodass

$$(gh)g^{-1} = h'$$

Also $x = gh = h'g \in Hg$.

"$\supseteq$" Sei $y \in Hg$ beliebig. Also $y = hg$ für ein bestimmtes $h \in H$. Weil $g^{-1} \sim_H g^{-1}$ und $h \sim_H e$, folgt aus (ii), dass $g^{-1}h \sim_H g^{-1}$. Das heißt, es existiert ein $h' \in H$, sodass $(g^{-1}h)(g^{-1})^{-1} = h'$. Wenn wir beide Seiten links mit g multiplizieren, bekommen wir

$$hg = gh' \in gH.$$

Also $y \in gH$, und auch die zweite Inklusion gilt. Q.E.D.

Das Lemma sagt uns also, dass die Gruppenoperation eine Operation auf der Faktormenge $G/\sim_H$ genau dann induziert, wenn $gH = Hg$ für alle $g \in G$. Wir führen deswegen folgenden Begriff ein.

Definition 1.137. Eine Untergruppe $H \leq G$ heißt **normale Untergruppe** (oder **Normalteiler**) von G genau dann, wenn $gH = Hg$ für alle $g \in G$. Wir schreiben dafür $H \triangleleft G$.

Vorsicht! Die Gleichheit $gH = Hg$ bedeutet **nicht**, dass $gh = hg \ \ \forall\, g \in G, h \in H.$ In Beispiel 1.135 haben wir gesehen, dass $H = \langle (1\ 2\ 3) \rangle$ ein Normalteiler ist. Es gilt also $(1\ 2)H = H(1\ 2)$, aber

$$(1\ 2)(1\ 2\ 3) = (2\ 3) \neq (1\ 3) = (1\ 2\ 3)(1\ 2).$$

Bemerkung 1.138. Für eine Untergruppe $H \leq G$ gilt

$$H \triangleleft G \iff gHg^{-1} = H, \ \forall\, g \in G.$$

Die Bedingung auf der rechten Seite heißt auch, dass H invariant unter **Konjugation** mit g für jedes $g \in G$ ist.

Satz 1.139. *Sei G eine Gruppe, $H \triangleleft G$ ein Normalteiler und bezeichne $G/H = \{ gH \ : \ g \in G \}$ die Menge der Äquivalenzklassen von $\sim_H$. Die Menge G/H zusammen mit der inneren Verknüpfung*

$$\begin{array}{ccc} G/H \times G/H & \longrightarrow & G/H \\ (Ha, Hb) & \longmapsto & Hab \end{array}$$

ist eine Gruppe.

Beweis-Skizze: Aus Lemma 1.136 ist die Abbildung wohl definiert, also tatsächlich eine innere Verknüpfung. Die Assoziativität folgt direkt aus der Assoziativität der Verknüpfung auf G. Direkt aus der Definition der Verknüpfung folgt auch, dass das neutrale Element $H = He$ ist und, dass das Inverse von Hg, die Nebenklasse Hg^{-1} ist. Q.E.D.

Wenn $H \triangleleft G$, dann heißt die Gruppe G/H aus Satz 1.139 die **Faktorgruppe** von G modulo H.

Endliche Gruppen

Eine Gruppe heißt **endliche Gruppe** wenn die Menge G eine endliche Menge ist.

Definition 1.140. Sei G eine **endliche** Gruppe. Die **Ordnung der Gruppe** G ist die Mächtigkeit $|G|$ der endlichen Menge G. Für jedes Element $g \in G$ ist die **Ordnung des Elementes** g ist $\operatorname{ord}(g) := |g| := |\langle g \rangle|$. Wenn $H \leq G$, dann ist der **Index der Untergruppe** H in G die Mächtigkeit der Faktormenge G/H und wird mit $[G : H]$ bezeichnet.

Beispiele:

1. $|\mathbb{Z}/6\mathbb{Z}| = 6$, $|2\mathbb{Z}/6\mathbb{Z}| = 3$, $[\mathbb{Z}/6\mathbb{Z} : 2\mathbb{Z}/6\mathbb{Z}] = 2$.
2. Die symmetrische Gruppe S_n hat Ordnung $n!$. Das kann man per Induktion nach n beweisen. Die Alternierende Untergruppe $A_n = \{\sigma \in S_n \ : \ \operatorname{sign}(\sigma) = 1\}$ hat Index $[S_n : A_n] = 2$.

Satz 1.141 (Lagrange)**.** *Sei G eine endliche Gruppe und $H \leq G$. Wir haben*

$$|G| = |H|[G : H].$$

Insbesondere $|H|$ teilt $|G|$ für jede Untergruppe von G.

Beweis-Skizze: Die Rechtsnebenklassen sind Äquivalenzklassen, also ist G eine disjunkte Vereinigung der Ha. Insbesondere $|G| = \sum_{a \in R} |Ha|$, wobei R ist eine Repräsentantensystem für $\sim_H$. Es reicht zu zeigen, dass $|H| = |Ha|$, $\forall\ a \in G$. Das heißt (cf. Definition 1.33), dass es eine Bijektion zwischen den beiden Mengen gibt. Wir haben für jedes $a \in G$ die bijektive Abbildung $\cdot a : H \longrightarrow Ha$. Die Umkehrabbildung ist $\cdot a^{-1}$. Q.E.D.

Weil die Ordnung eines Elementes gleich mit der Ordnung der davon erzeugten Untergruppe ist, haben wir folgendes Korollar des Satzes von Lagrange.

Korollar 1.142. Wenn G eine endliche Gruppe ist und $g \in G$, dann ist $\operatorname{ord} G$ ein Teiler von $|G|$.

Als Folgerung dieses Korollars haben wir den kleinen Satz von Fermat (Satz 1.86). Die Beziehung kommt aus der Tatsache, dass wenn p eine Primzahl ist, dann ist $(\mathbb{Z}/p\mathbb{Z}, \cdot)$ eine endliche Gruppe von Ordnung $p-1$. Also für jedes $a \not{equiv} 0 \bmod p$ gilt

$$\operatorname{ord}([a]) \mid p-1.$$

Das heißt, dass ein $k \in \mathbb{N}$ existiert, sodass $p-1 = k \cdot \operatorname{ord}([a])$. Also

$$[a]^{p-1} = [a]^{k \cdot \operatorname{ord}([a])} = \left([a]^{\operatorname{ord}([a])}\right)^k = [1]^k = [1].$$

Wenn $a \equiv 0 \bmod p$, dann gilt auch $a^p \equiv 0 \bmod p$. Also, wenn p eine Primzahl ist, dann gilt $a^p \equiv a \bmod p$ für alle $a \in \mathbb{Z}$.

1.6.10 Unitäre Ringe

Diese Strukturen werden auch Ringe mit Eins genannt. Es gibt eine Theorie der nicht-unitären Ringen auch. Hier werden aber alle Ringe eine Eins haben und alle Ringhomomorphismen werden die 1 auf der 1 abbilden.

Definition 1.143. Ein **Ring** ist ein Tripel $(R, +, \cdot)$ wobei: R ist eine Menge und $+$ und $\cdot$ sind innere Verknüpfungen, die Addition beziehungsweise Multiplikation genannt werden, sodass folgende Axiome erfüllt sind

(R 1) $(R, +)$ ist eine abelsche Gruppe.

(R 2) Die Multiplikation ist assoziativ und hat neutrales Element.

(R 3) (*Distributivitätsgesetze*) Für alle $a, b, c \in R$ gilt

$$(a+b) \cdot c = a \cdot c + b \cdot c \qquad \text{und} \qquad a \cdot (b+c) = a \cdot b + a \cdot c.$$

Ein **kommutativer Ring** ist ein Ring in dem die Multiplikation kommutativ ist.

Bezeichnung. Das neutrale Element der Addition wird mit 0_R oder nur mit 0 bezeichnet. Das neutrale Element der Multiplikation wird mit 1_R oder nur mit 1 bezeichnet.

Außer dem Fall in dem die "nicht-Kommutativität" ausdrücklich angegeben wird, wird für uns ein Ring **immer kommutativ** sein.

Bemerkung 1.144. Für jeder (nicht unbedingt kommutativer) Ring R gelten

1. $0 \cdot a = 0 \quad \forall\, a \in R.$
2. $-a = (-1) \cdot a \quad \forall\, a \in R.$
3. $(-a) \cdot b = -(a \cdot b) \quad \forall\, a, b \in R.$

1.6.11 Körper

Definition 1.145. Ein **Körper** ist ein Tripel $(\mathbb{K}, +, \cdot)$ wobei $\mathbb{K}$ ist eine Menge und $+$ und $\cdot$ sind Verknüpfungen die Addition beziehungsweise Multiplikation genannt sind, sodass folgende Axiome erfüllt sind

($\mathbb{K}\,1$) $(\mathbb{K}, +)$ ist eine abelsche Gruppe, dessen neutrale Element 0 heißt.

($\mathbb{K}\,2$) $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

($\mathbb{K}\,3$) (*Distributivitätsgesetze*) Für alle $a, b, c \in \mathbb{K}$ gilt

$$(a+b)\cdot c = a\cdot c + b\cdot c \qquad \text{und} \qquad a\cdot(b+c) = a\cdot b + a\cdot c.$$

Manche Autoren verlangen die Kommutativität der Multiplikation nicht und nennen Objekte die die Definition 1.145 erfüllen ein Feld[39]. Wir bleiben aber bei der Definition von Körper die eine kommutative Multiplikation voraussetzt und nennen **Schiefkörper** ein Tripel das alle Axiome eines Körpers außer der Kommutativität der Multiplikation erfüllt.

Bemerkung 1.146. Ein Körper ist also ein kommutativer Ring, in dem jedes nicht-triviales Element invertierbar bezüglich der Multiplikation ist. Ein Element a ist **trivial** wenn $a = 0$ und **nicht-trivial** wenn $a \neq 0$.

[12] 22.11.'23

1.6.12 Beispiele

1. Nullring: $R = \{0\}$. In diesem Fall ist $0 = 1$. Und $0 = 1$ gilt nur in diesem Fall: Wenn $0 = 1$ und $a \in R$, dann haben wir $a = 1 \cdot a = 0 \cdot a = 0$.
2. Der Nullring ist kein Körper, weil dann $(\mathbb{K} \setminus \{0\})$ keine Gruppe ist. (Gruppen müssen nicht-leere Mengen sein, weil ein neutrales Element existieren muss.)
3. $\mathbb{Z}$ ist ein Ring. $\mathbb{Q}$, $\mathbb{R}$ sind Körper.
4. $\mathbb{C} = \{a + ib \;:\; a, b \in \mathbb{R}\}$ mit "normales Rechnen" unter der Bedingung, dass $i^2 = -1$:
$$\begin{aligned}(a+ib)+(c+id) &= (a+c)+i(b+d)\\ (a+ib)\cdot(c+id) &= (ac-bd)+i(ad+bc)\end{aligned}$$
5. $\mathbb{Z}[i] = \{a + ib \;:\; a, b \in \mathbb{Z} \text{ und } i^2 = -1\}$ mit der Addition und Multiplikation der komplexen Zahlen ist ein Ring: der Ring der Gaußschen Zahlen.
6. $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \;:\; a, b \in \mathbb{Q} \text{ und } (\sqrt{2})^2 = 2\}$.

[39] Das entspricht auch der englischen, französischen und italienischen Terminologie: *field, champ, campo.*

7. Ein **Polynom** in einer Variable x mit Koeffizienten in dem Ring R ist ein formaler Ausdruck

$$f = a_0 + a_1 x^1 + \cdots + a_n x^n$$

mit $n \in \mathbb{N}$ und $a_i \in R \; \forall \, i = 1 \ldots n$. Die Menge aller solchen Elementen wird mit $R[x]$ bezeichnet. Der **Koeffizient** von x^n ist das Element $a_n \in R$. Um klar zu machen wann zwei Polynome gleich sind und auch um die Addition von Polynome einfacher zu definieren, ist es praktisch Polynome als $\mathbb{N}$-Tupel[40] die ab einer gewissen Stelle nur Null-Einträge haben. Die Schreibweise kann dann auch kompakter sein:

$$f = \sum_{i \in \mathbb{N}} a_i x^i.$$

In Konkreten Fällen lassen wir alle Koeffizienten die Null sind aus:

$$f = 1 + 3x^3 + x^5 = 1 \cdot x^0 + 0 \cdot x + 0 \cdot x^2 + 3 \cdot x^3 + 0 \cdot x^4 + 1 \cdot x^5 + 0 \cdot x^6 + \ldots$$

Das **Nullpolynom** ist das Polynom dessen Koeffizienten alle Null sind. Das entspricht dann der Null-Abbildung $0 : \mathbb{N} \longrightarrow R$, das heißt dem $\mathbb{N}$-Tupel mit allen Einträgen gleich mit Null. Wir bezeichnen es (auch) mit 0. Zwei Polynome $f = \sum a_i x^i$ und $g = \sum b_i x^i$ sind gleich, wenn $a_i = b_i$ für alle $i \in \mathbb{N}$. Der **Grad** eines nichtnulles Polynoms $f = \sum a_i x^i$ wird mit $\deg f$ bezeichnet und ist

$$\deg f := \max\{k \in \mathbb{N} \;\; : \;\; a_k \neq 0\},$$

also das maximale k für das der Koeffizient von x^k nicht Null ist; dieser Koeffizient wird **Leitkoeffizient** genannt. Der Grad des Nullpolynoms ist nicht bestimmt, wir verwenden aber die Konvention, dass das Nullpolynom jedes mögliche Grad haben kann/darf. Also das Nullpolynom erfüllt

$$0 \in \{f \;\; : \;\; \deg f = n\}, \quad \forall \, n \in \mathbb{N}.$$

Für jedes Polynom f definieren wir den Koeffizient a_i von x^i als Null wenn $i > \operatorname{Grad} f$.

Die **Summe** von $f = \sum a_i x^i$ und $g = \sum b_i x^i$ ist das Polynom

$$f + g := \sum_{i \in \mathbb{N}} (a_i + b_i) x^i.$$

Das **Produkt** von $f = \sum a_i x^i$ und $g = \sum b_j g^j$ ist das Polynom

$$fg := \sum a_i b_j x^{i+j} = \sum p_k x^k.$$

Der Koeffizient von x^k in fg ist also $p_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i+j=k} a_i b_j$.

Satz 1.147. *Sei R ein Ring (also kommutativ und mit 1). Die Menge aller Polynome $R[x]$ zusammen mit der Addition und dem Produkt der Polynome ist ein Ring. Der Ring R ist ein Unterring (Definition 1.157) von $R[x]$ wenn man dessen Elementen mit Polynome von Grad Null identifiziert.*

[40] Also als Abbildungen $f : \mathbb{N} \longrightarrow R$ mit $\exists \; n$ sodass $f(i) = 0, \; \forall \, i > n$.

8. Eine 2×2 **Matrix** mit reellen Koeffizienten ist eine Tabelle A der Form

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \text{wobei } a, b, c, d \in \mathbb{R}.$$

Wir bezeichnen mit $\mathrm{Mat}_{2\times 2}(\mathbb{R})$ die Menge aller solchen Matrizen. Wenn $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$, dann definieren wir

$$\begin{aligned} A + B &:= \begin{pmatrix} a+p & b+q \\ c+r & d+s \end{pmatrix} \\ A \cdot B &:= \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix} \end{aligned}$$

Die Addition von Matrizen ist kommutativ, aber die Multiplikation ist nicht (siehe Beispiel 1.149).

Satz 1.148. *Das Tripel* $(\mathrm{Mat}_{2\times 2}(\mathbb{R}), +, \cdot)$ *ist ein nicht-kommutativer Ring, mit*

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad -A = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix} \qquad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

1.6.13 Mehr über Ringe

Ein **Nullteiler** in einem Ring R ist ein Element $a \in R$ mit der Eigenschaft, $\exists b \neq 0$, sodass $ab = 0$. Ein Element $a \in R$ ist ein nicht-Nullteiler wenn es nicht ein Nullteiler ist. Ein nicht-Nullteiler ist also ein Element $a \in R$ für dem

$$ab = 0 \Rightarrow b = 0$$

gilt.

Beispiele: In $\mathbb{Z}/6\mathbb{Z}$ sind $0, 2, 3, 4$ Nullteiler und $1, 5$ nicht-Nullteiler. Allgemein gilt $m \in \mathbb{Z}/n\mathbb{Z}$ ist ein Nullteiler genau dann, wenn $\mathrm{ggT}(m, n) \neq 1$.

Beispiel 1.149. $a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ mit

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Bemerkung 1.150. Ein Nullteiler $a \in R$ ist nie invertierbar, weil $ab = 0 \overset{\exists a^{-1}}{\Longrightarrow} a^{-1} \cdot (ab) = a^{-1} \cdot 0 \Leftrightarrow b = 0$. Es gibt aber nicht-invertierbare Elemente die nicht-Nullteiler sind: e.g. $2 \in \mathbb{Z}$.

Ein Ring ist ein **Integritätsbereich** genau dann, wenn es keine nicht-triviale Nullteiler hat.

Beispiel 1.151.

1. $\mathbb{Z}$, $\mathbb{K}[x]$ sind Integritätsbereiche.
2. $\mathbb{Z}/n\mathbb{Z}$ ist ein Integritätsbereich genau dann, wenn n prim ist.

Bemerkung 1.152. Für eine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$, haben wir immer

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix}.$$

Also A ist invertierbar $\iff ad - bc \neq 0$.

Eine **Einheit** ist ein invertierbares Element eines Ringes. Wir bezeichnen mit

$$U(R) = \{u \in R \ : \ u \text{ ist eine Einheit}\}$$

die Menge aller Einheiten. Wir haben: $(U(R), \cdot)$ ist eine Gruppe.

Beispiele:

1. $U(\mathbb{Z}) = \{\pm 1\}$
2. $U(\mathbb{K}[x]) = \mathbb{K}^\times$.
3. $U(\mathrm{Mat}_{2\times 2}(\mathbb{R}) = \mathrm{GL}_2(\mathbb{R}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \ : \ ad - bc \neq 0\}$.

Bemerkung 1.153. Wenn R_1 und R_2 Ringe sind, dann kann man immer auf dem Kartesischen Produkt $R_1 \times R_2$ eine Ringstruktur mit der Komponenten-weise Addition und Multiplikation definieren. Das Problem ist, dass wenn $\mathbb{K}_1$ und $\mathbb{K}_2$ Körper sind, dann geben die Komponenten-weise Verknüpfungen **nie** eine Körper Struktur auf $\mathbb{K}_1 \times \mathbb{K}_2$. Das ist so, weil in einem Körper $0 \neq 1$ und $(0,1) \cdot (1,0) = (0,0)$ - es gibt also nicht-invertierbare Elementen.

Ringhomomorphismen, Unterringe und Ideale

Definition 1.154. Seien R und S zwei Ringe. Ein **Ringhomomorphismus** von R nach S ist eine Abbildung $\varphi : R \longrightarrow S$ die folgende Axiome erfüllt

(RHom 1) $\varphi(a + b) = \varphi(a) + \varphi(b)$.

(RHom 2) $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$.

(RHom 3) $\varphi(1) = 1$.

Um die Definition eines Ringisomorphismus zu bekommen, muss man nur "Gruppen" mit "Ring" in Definition 1.94 ersetzen. Es ist auch hier einfach zu zeigen, dass alle bijektive Ringhomomorphismen auch Isomorphismen sind.
Ein Ringhomomorphismus von R nach R heißt **Endomorphismus** von R und ein bijektives Endomorphismus von R heißt **Automorphismus** von R.
Der **Kern** eines Ringhomomorphismus φ ist $\mathrm{Ker}\,\varphi = \{r \in R \ : \ \varphi(r) = 0\}$.

Bemerkung 1.155. Die Abbildung $0 : R \longrightarrow S$ mit $0(a) := 0$ ist ein Ringhomomorphismus nur wenn $1 = 0$ in S gilt, also wenn S der Null-Ring ist.

Bemerkung 1.156. Der Ring $\mathbb{Z}$ hat folgende Eigenschaft: für jeden Ring R gibt es genau einen Ringhomomorphismus $f : \mathbb{Z} \longrightarrow R$. In einer Kategorie heißt so ein Objekt Anfangsobjekt. Für die Kategorie der Mengen ist das die leere Menge. Für die Kategorie der Körper gibt es so etwas nicht.

Für Ringe ist es nicht mehr wahr, dass Epimorphismen[41] dasselbe wie surjektive Homomorphismen sind. Zum Beispiel, $\mathbb{Z} \longrightarrow \mathbb{Q}$ ist eine Epimorphismus, aber es ist nicht surjektiv.

Beispiele:

1. $f : \mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, f(x) = [x]$
2. $\varphi : \mathbb{Z} \longrightarrow R$ - für jeder Ring R gibt es genau ein solches Ringhomomorphismus und es ist von $\varphi(1)$ bestimmt. Das erlaubt uns jedes $n \in \mathbb{Z}$ als Element in jedem Ring zu sehen.
3. $ev_\alpha : R[x] \longrightarrow R, \quad ev_\alpha(f) = f(\alpha)$. Wobei, wenn $f = a_0 + a_1 x + \cdots + a_n x^n$, dann ist $f(\alpha) = a_0 + a_1\alpha + \cdots + a_n\alpha^n$ für jedes $\alpha \in \mathbb{R}$.

[13] 27.11.'23

Definition 1.157. Ein **Unterring** (oder Teilring) eines Ringes $(R, +, \cdot)$ ist eine Teilmenge $S \subseteq R$ mit $1 \in S$ und die abgeschlossen bezüglich der Subtraktion und der Multiplikation ist.

Extra:
Das heißt, dass ein Unterring ist eine Teilmenge die selber zusammen mit den Einschränkungen der Addition und Multiplikation ein Ring ist. Dieser Begriff ist aber nicht so flexibel wie der Begriff von (normale) Untergruppe. Insbesondere, kann man auf der Faktormenge keine natürliche Ringstruktur definieren und der Kern eines Ringhomomorphismus ist fast nie ein Unterring:

Wenn $S \subseteq R$ ein Unterring ist, dann ist insbesondere $(S, +) \triangleleft (R, +)$ und also $(R/S, +)$ eine Gruppe. Die Faktormenge R/S ist definiert also durch $a \sim_S b \Leftrightarrow a - b \in S$. Ist die Multiplikation wohl definiert? Nicht unbedingt: $(r+s)r' = rr' + r's$, aber es kann sein, dass $r's \notin S$. (Z.B. $\mathbb{Z} \subset \mathbb{Q}$, $[\frac{2}{3}] \cdot [\frac{1}{2}] \neq [\frac{2}{3}] \cdot [\frac{1}{2} + 1]$. Deswegen, führt man für Ringe folgende Definition ein:)

Definition 1.158. Sei R ein Ring (also kommutativ mit 1). Ein **Ideal**[42] von R ist eine Teilmenge $I \subseteq R$ die folgende Axiome erfüllt:

(I 1) I ist eine Untergruppe von $(R, +)$.

(I 2) Wenn $a \in I$ und $r \in R$, dann gilt $ra \in I$.

Bemerkung 1.159. Eine Teilmenge I ist ein Ideal genau dann, wenn $a - b \in I$ und $ra \in I$, $\forall\, a, b \in I$ und $r \in R$. Oder genau dann, wenn:

$$r_1 a_1 + \cdots + r_n a_n \in I \quad \forall\, a_i \in I, r_i \in R.$$

[41] Diese sind Ringhomomorphismen die die Bedingung aus Satz 1.31 erfüllen.
[42] Der Begriff kommt von *Ideale Zahl*. Es wurde von Richard Dedekind (1831-1916) und Leopold Kronecker (1823-1891) eingeführt, die Ernst-Eduard Kummers (1810-1893) Ideen von Ideale Zahlen erweiterten.

Bemerkung 1.160. Wenn $I \subseteq R$ ein Ideal ist, dann ist R/I ein Ring mit

$$\begin{aligned} \sim_I b &\Leftrightarrow a - b \in I, \\ [a] + [b] &= [a + b], \\ [a][b] &= [ab]. \end{aligned}$$

Der Ring R/I ist der **Faktorring** von R modulo I.

Bemerkung 1.161. Wenn $\varphi : R \longrightarrow S$ ein Ringhomomorphismus ist, dann ist der Kern $\operatorname{Ker}\varphi := \{a \in R \ : \ \varphi(a) = 0\}$ ein Ideal. Es ist nicht ein Teilring, weil $\varphi(1) \neq 0$ wenn S nicht der Nullring ist.

Beispiele:

1. $(0) = \{0\}$
2. R ist immer ein Ideal von R. Der Faktorring ist dann isomorph zum Nullring.
3. $\mathbb{Z}$ hat keine Unterringe, aber die Menge aller Ideale ist $\{n\mathbb{Z} \ : \ n \in \mathbb{N}\}$. Wir haben $a\mathbb{Z} \subseteq b\mathbb{Z} \Leftrightarrow b|a$ und $n \in a\mathbb{Z} \Leftrightarrow a|n$.

Für eine Familie von Ideale $\{I_j\}_{j\in J}$ mit $I_j \subseteq R$ ist auch $\cap_{j\in J} I_j \subseteq R$ ein Ideal. Die Vereinigung aber nicht ($2\mathbb{Z} \cup 3\mathbb{Z} \neq n\mathbb{Z}$). Wir können aber die Menge $I_1 + I_2 = \{a_1 + a_2 \ : \ a_i \in I\}$. Das ist das kleinste Ideal das $I_1 \cup I_2$ enthält.
Für ein Element $a \in R$ definieren wir das von a erzeugte **Hauptideal** $(a) = \{ra \ : \ r \in R\}$. Es ist das kleinste Ideal das a enthält. Ein Ideal ist ein Hauptideal genau dann, wenn es von einem einzigen Element von R erzeugt werden *kann*. Genauso können wir für eine endliche Teilmengen $\{a_1, \ldots, a_n\} \subseteq R$ das von $a_1, \ldots, a_n$ erzeugte Ideal definieren:

$$(a_1, \ldots, a_r) := (a_1) + \cdots + (a_r) = \{\textstyle\sum_{i=1}^{n} r_i a_i \ : \ r_i \in R\}.$$

Es ist das (bezüglich der Inklusion) kleinste Ideal von R das alle a_i enthält.

Beispiele:

1. $(1) = R$. Wenn $a \in U(R)$ und $a \in I$, dann $I = R = (1)$.
2. $\mathbb{Z} \supset (4, 6) = (4) + (6) = (2)$. $(4) \cap (6) = (12)$.
3. $(2, x) \in \mathbb{Z}[x]$ ist nicht ein Hauptideal.
4. $(x, y) = (x - y, x + y) \subset \mathbb{K}[x, y]$ ist nicht ein Hauptideal.

Da der Divisionssatz auch in $\mathbb{K}[x]$ gilt (wenn $\mathbb{K}$ ein Körper ist!), mit dem selben Beweis wie dem von Satz 1.105, haben wir dass alle Ideale von $\mathbb{K}[x]$ Hauptideale sind.

Definition 1.162. Ein Ideal $P \subsetneq R$ ist ein **Primideal** wenn $ab \in P \Rightarrow a \in P$ oder $b \in P$.
Ein Ideal $\mathfrak{m} \subsetneq R$ ist ein **Maximalideal** wenn aus $\mathfrak{m} \subseteq I \subseteq R$ folgt $I = \mathfrak{m}$ oder $I = R$.

Satz 1.163. *Sei R ein Ring. Es gilt*

(i) *Jedes Maximalideal ist auch ein Primideal*

(ii) P *ist ein Primideal* $\Longleftrightarrow R/P$ *ein Integritätsbereich ist.*

(iii) $\mathfrak{m}$ *ist ein Maximalideal* $\Longleftrightarrow R/\mathfrak{m}$ *ein Körper ist.*

Beweis-Skizze: (i) Sei $\mathfrak{m} \subsetneq R$ ein Maximal ideal und $a, b \in R$ mit $ab \in \mathfrak{m}$. Wenn $a \notin \mathfrak{m}$, dann wollen wir zeigen, dass $b \in \mathfrak{m}$.
Aus $a \notin \mathfrak{m}$ folgt $\mathfrak{m} \subsetneq \mathfrak{m} + (a)$, also $\mathfrak{m} + (a) = R$. Es gibt also $m \in \mathfrak{m}$ und $r \in R$, sodass $m + ra = 1$. Also $b = bm + rab \in \mathfrak{m}$.

(ii) $ab \in P \Leftrightarrow [a][b] = [0] \in R/P$.

(iii) $\mathfrak{m}$ maximal $\Leftrightarrow \forall\, a \in R \setminus \mathfrak{m}$ gilt $\mathfrak{m} + (a) = (1) \Leftrightarrow \forall\, [a] \neq [0] \in R/\mathfrak{m}, \exists [r] \in R/\mathfrak{m}$ mit $[a][r] + [m] = [a][r] = 1 \Leftrightarrow R/\mathfrak{m}$ ist ein Körper. Q.E.D.

1.7 Kategorien

Eine **Kategorie** $\mathcal{C}$ besteht aus Objekten und Morphismen. Die Objekten bilden eine *Klasse* $Obj(\mathcal{C})$. Das heißt, eine Zusammenfassung die einem höheren Ordens als Mengen gehört. Man soll sich hier nicht viele Gedanken darüber machen was genau das ist. Wichtig ist nur, dass die Klasse aller Mengen nicht zu einem Widerspruch führt. Wir schreiben auch für Klassen $A, B \in Obj(\mathcal{C})$ und das heißt, dass A und B Objekte der Kategorie $\mathcal{C}$ sind. Für jedes Paar von Objekten $A, B \in Obj(\mathcal{C})$ gibt es auch eine Menge[43] von Morphismen (oder einfach Pfeile) von A nach B. Diese wir mit $\mathrm{Hom}_{\mathcal{C}}(A, B)$ bezeichnet[44]. Weiterhin, gibt es eine Verknüpfung von Morphismen die assoziativ ist und es gibt für alle Objekte ein identischer Morphismus. Genauer gesagt:

- $\forall\, A, B, C \in Obj(\mathcal{C})$ haben wir eine Abbildung $\circ : \mathrm{Hom}_{\mathcal{C}}(A, B) \times \mathrm{Hom}_{\mathcal{C}}(B, C) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(A, C)$ und wir bezeichnen das Bild von (f, g) darunter mit $g \circ f$.
- $\forall\, A_1, \ldots, A_4 \in Obj(\mathcal{C})$ und alle $f_i \in \mathrm{Hom}_{\mathcal{C}}(A_i, A_{i+1})$ für $i = 1, 2, 3$ gilt
$$(f_3 \circ f_2) \circ f_1 = f_3 \circ (f_2 \circ f_1).$$
- $\forall\, A \in Obj(\mathcal{C})$ existiert ein Morphismus $\mathrm{id}_A \in \mathrm{Hom}_{\mathcal{C}}(A, A)$ sodass
$$id_A \circ f = f \circ \mathrm{id}_A = f, \quad \forall\, f \in \mathrm{Hom}_{\mathcal{C}}(A, A).$$

Was wir hier gemacht haben ist eine gemeinsame Abstraktion von Mengen, Gruppen, Ringe und so weiter.

[43] eigentlich auch Klasse, aber nicht für uns hier.
[44] allgemeiner wäre $\mathrm{Mor}_{\mathcal{C}}(A, B)$, aber Hom reicht für uns hier.

Beispiele: Wir geben hier unten eine Liste von Beispielen durch Objekte und die Art von Morphismen an.

1. Objekte = Mengen, Morphismen = alle Abbildungen.
2. Objekte = partiell geordnete Mengen, Morphismen = Abbildungen f mit $a \leq b \Rightarrow f(a) \leq f(b)$.
3. Objekte = Gruppen, Morphismen = Gruppenhomomorphismen.
4. Objekte = Abelsche Gruppen, Morphismen = Gruppenhomomorphismen.
5. Objekte = Unitäre Ringe, Morphismen = Ringhomomorphismen.
6. Objekte = Kommutative Ringe, Morphismen = Ringhomomorphismen.
7. Objekte = Körper, Morphismen = Ringhomomorphismen.
8. Objekte = $\mathbb{K}$-Vektorräume, Morphismen = lineare Abbildungen. (Das ist das Hauptthema dieser Vorlesung).

Überblick

Alles in dieser Liste ist abgekürzt und unvollständig aufgeschrieben!

Mengen: $x \in M$ ist eine Aussage + $M \in M$ ist immer falsch.
Teilmengen: $N \subseteq M \Leftrightarrow (x \in N \Rightarrow x \in M)$.
Gleichheit: $M = N \Leftrightarrow (N \subseteq M \text{ und } M \subseteq N)$.
Produktmengen: $M \times N = \{(m,n) \ : \ m \in M, n \in N\}$,
wobei $(m,n) = (m',n') \Leftrightarrow m = m' \text{ und } n = n'$.
Abbildungen: $f : M \longrightarrow N \quad \forall x \in M, \ \exists! f(x) \in N$.
- injektiv: $f(a) = f(b) \Rightarrow a = b$.
- surjektiv: $\forall n \in N, \exists \ m \in M$, sodass $f(m) = n$.
- bijektiv: $\forall n \in N, \exists! \ m \in M$, sodass $f(m) = n$.
- invertierbar: $\exists f^{-1} : N \longrightarrow M$, sodass $f \circ f^{-1} = \mathrm{id}_N \quad \text{und } f^{-1} \circ f = \mathrm{id}_M$.

$$\begin{aligned} \text{injektiv} &\Leftrightarrow \exists \text{ links-Inverse} \\ \text{surjektiv} &\Leftrightarrow \exists \text{ rechts-Inverse} \\ \text{injektiv} &\Leftrightarrow \exists \text{ Inverse} \end{aligned}$$

Äquivalenzrelation:
- $x \sim x$ (Reflexivität),
- $x \sim y \Rightarrow y \sim x$ (Symmetrie),
- $(x \sim y \text{ und } y \sim z) \Rightarrow x \sim z$ (Transitivität)

Äquivalenzklassen: $[a] = \{b \ : \ b \sim a\}$
Faktormenge: $M/\sim \ = \{[a] \ : \ a \in M\}$ - die Menge aller Äquivalenzklassen.
Die Universelle Eigenschaft (Satz 1.60) sagt wie man eine Abbildung auf $M/\sim$ durch eine Abbildung auf M definieren muss.
Ordnungsrelation:
- $x \sim x$ (Reflexivität),
- $(x \sim y \text{ und } y \sim x) \Rightarrow x = y$ (Antisymmetrie),
- $(x \sim y \text{ und } y \sim z) \Rightarrow x \sim z$ (Transitivität)

Wohlordnung: Jede Teilmenge hat ein Minimum. $\mathbb{N}$ ist wohl geordnet.
Gruppen: (Menge, Verknüpfung) mit:
- Assoziativität,
- neutrales Element,
- inverses Element.

nicht unbedingt kommutativ.
Untergruppen: Teilmenge mit der Einschränkung der Verknüpfung ist wieder eine Gruppe.
Ringe: (Menge, $+$, $\cdot$) mit:
- $(R,+)$ = abelsche Gruppe,
- $\cdot$ ist assoziativ und hat 1,
- Distributivität: $a(b+c) = ab + ac$, $(a+b)c = ac + bc$.

Körper Kommutativer Ring mit $\forall \, a \neq 0, \exists a^{-1}$.
Homomorphismen: Abbildungen die die zusätzliche Struktur berücksichtigen.

Kapitel 2

Matrizen und Lineare Gleichungssysteme

"Classification of mathematical problems as linear and nonlinear is like classification of the Universe as bananas and non-bananas."

2.1 Definition und Bezeichnungen

2.1.1 Matrizen

Seien $m, n \in \mathbb{N}_{>0}$ und sei R ein Ring.

Definition 2.1. Eine $(m \times n)$–**Matrix** mit Einträge (oder Koeffizienten) in R ist eine Abbildung

$$A : \{1, \dots, n\} \times \{1, \dots, m\} \longrightarrow R.$$

Wir bezeichnen mit $a_{ij} := A((i, j))$ und stellen A als Tabelle dar

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Dabei heißt der erste Index i der **Zeilenindex** und der zweite Index j heißt **Spaltenindex** von A. Wenn m und n klar aus der Kontext sind, schreiben wir einfach $A = (a_{ij})$. Die Menge aller $(m \times n)$-Matrizen mit Einträge in R wird mit $\mathrm{Mat}_{m\times n}(R)$ bezeichnet.

- Der **Typ** der Matrix ist das geordnete Paar (m, n). Wir sagen auch, dass der Typ $m \times n$ ist.
- Eine Matrix ist **quadratisch** wenn $m = n$.
- Eine **Teilmatrix** einer $(m\times n)$-Matrix A ist eine Einschränkung von $A : \{1, \dots, m\}\times\{1, \dots, n\} \longrightarrow R$ zu einer Teilmenge des Definitionsbereichs der Form $I \times J$ mit $I \subseteq \{1, \dots, m\}$ und $J \subseteq \{1, \dots, n\}$. Wir bezeichnen so eine Teilmatrix von A mit $A|_{I\times J}$.
- Eine **Zeilenmatrix** ist eine $1 \times n$ Matrix.
- Eine **Spaltenmatrix** ist eine $m \times 1$ Matrix.

- Die **Zeilen**, beziehungsweise die **Spalten**, einer $m \times n$ Matrix sind die m $(1 \times n)-$ Teilmatrizen, beziehungsweise die n $(m \times 1)$-Teilmatrizen, von $A = (a_{ij})$.

$$Z_{\mathbf{i}} = \begin{pmatrix} a_{\mathbf{i}1} & \dots & a_{\mathbf{i}n} \end{pmatrix} \quad , \mathbf{i} = 1, \dots, m.$$

$$S_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad , j = 1, \dots, n.$$

  Wir sagen, dass Z_i die *Zeile i von A*, oder die *i-te Zeile von A* ist. Wenn nötig, schreiben wir noch das Etikett "A" irgendwie dazu: ${}^{A}Z_i$, ${}_{A}Z_i$, Z_i^A, $Z_i(A)$, o.ä. Analog für Spalten.

- Eine **Blockdarstellung** einer Matrix A ist eine Schreibweise die gewisse Teilmatrizen betont. Diese Teilmatrizen müssen *zusammenhängend* sein, dass heißt, dass die Zeilen- und Spalten-Indizes konsekutiv sind. Ein $r \times s$-Block von A ist also eine Teilmatrix der Form $B_{IJ} := A|_{I \times J}$ mit $I = \{i, i+1, i+2, \dots, i+r\} \subseteq \{1, \dots, m\}$ und $J = \{j, j+1, j+2, \dots, s\} \subseteq \{1, \dots, n\}$. In einer Blockdarstellung müssen Blöcke disjunkt sein (d.h. $(I_1 \times J_1) \cap (I_2 \cap J_2) = \emptyset$). Zum Beispiel:

$$\left(\begin{array}{c|ccc} a & b & c & d \\ \hline e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right), \quad \text{mit } A = (a)\,,\ B = \begin{pmatrix} b & c & d \\ f & g & h \end{pmatrix},\ C = \begin{pmatrix} e \\ i \\ m \end{pmatrix},\ D = \begin{pmatrix} j & j & l \\ n & o & p \end{pmatrix}.$$

  Am meisten werden wir die Blockaufteilung in Zeilen oder Spalten verwenden

$$A = \left(\begin{array}{c} Z_1 \\ \hline Z_2 \\ \hline \vdots \\ \hline Z_m \end{array}\right) = \left(\begin{array}{c|c|c|c} S_1 & S_2 & \dots & S_n \end{array}\right)$$

  und die **Erweiterung** einer $(m \times n)-$Matrix A durch eine m-Spalte $\mathbf{b}$:

$$(A|\mathbf{b}) := \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right).$$

- Die $(m \times n)$-**Nullmatrix** ist die Matrix A mit $a_{ij} = 0 \ \ \forall\, i, j$.

- Die **Transponierte** Matrix einer $(m \times n)-$Matrix $A = (a_{ij})$ ist die $(n \times m)-$Matrix $A^T := (a_{ji})$.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Rechnen mit Matrizen

Die **Matrizenaddition** ist die Verknüpfung $+ : \mathrm{Mat}_{m\times n}(R) \times \mathrm{Mat}_{m\times n}(R) \longrightarrow \mathrm{Mat}_{m\times n}(R)$ gegeben durch

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

Bemerkung 2.2. $(\mathrm{Mat}_{m\times n}(R), +)$ ist eine abelsche Gruppe.
Beweis: Weil die Addition in R assoziativ und kommutativ ist, hat auch die Matrizenaddition diese Eigenschaften. Das neutrale Element ist die $(m \times n)$ Nullmatrix $0 = 0_{m\times n}$, und jede Matrix $A = (a_{ij})$ hat ein inverses Element: $-A = (-a_{ij})$.

Die **Skalarmultiplikation** ist die (äußere) Verknüpfung $\cdot : R \times \mathrm{Mat}_{m\times n}(R) \longrightarrow \mathrm{Mat}_{m\times n}(R)$ durch

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

Das Produkt $A \cdot B$ zweier Matrizen ist **nur dann definiert** wenn die Anzahl von Spalten in A gleich mit der Anzahl von Zeilen in B ist. Seien also $m, n, l \in \mathbb{N}_{>0}$.
Die **Matrixmultiplikation** ist die Verknüpfung $\cdot : \mathrm{Mat}_{m\times n}(R) \times \mathrm{Mat}_{n\times p}(R) \longrightarrow \mathrm{Mat}_{m\times p}(R)$ definiert durch

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} a_{1k}b_{k1} & \sum_{k=1}^{n} a_{1k}b_{k2} & \dots & \sum_{k=1}^{n} a_{1k}b_{kn} \\ \sum_{k=1}^{n} a_{2k}b_{k1} & \sum_{k=1}^{n} a_{2k}b_{k2} & \dots & \sum_{k=1}^{n} a_{2k}b_{kn} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{k=1}^{n} a_{mk}b_{k1} & \sum_{k=1}^{n} a_{mk}b_{k2} & \dots & \sum_{k=1}^{n} a_{mk}b_{kn} \end{pmatrix}$$

Man kann also Zeilenmatrizen mit Spaltenmatrizen multiplizieren (und nicht umgekehrt) genau dann, wenn sie die selbe Anzahl von Einträge haben. Also

$$Z \cdot S = \begin{pmatrix} z_1 & \dots & z_n \end{pmatrix} \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} z_1 s_1 + \dots + z_n s_n \end{pmatrix}$$

Wenn ${}^{A}Z_i$ die Zeilen von A sind und S_j^B die Spalten von B, dann kann man AB beschreiben als

$$AB = \begin{pmatrix} {}^{A}Z_1 S_1^B & \dots & {}^{A}Z_1 S_p^B \\ \vdots & & \vdots \\ {}^{A}Z_m S_1^B & \dots & {}^{A}Z_m S_p^B \end{pmatrix}$$

Bemerkung 2.3. Direktes Rechnen und Anwenden der Assoziativität und Distributivität von $+$ und $\cdot$ in R zeigt, dass auch die Matrixmultiplikation assoziativ ist:

Wenn $A \in \mathrm{Mat}_{m\times n}(R)$, $B \in \mathrm{Mat}_{n\times p}(R)$, $C \in \mathrm{Mat}_{p\times q}(R)$, dann gilt

$$A(BC) = (AB)C = \left(\sum_{\substack{k=1\ldots n \\ l=1\ldots p}} a_{\mathbf{i}k} b_{kl} c_{l\mathbf{j}} \right)_{\substack{\mathbf{i}=1\ldots m \\ \mathbf{j}=1\ldots q}}.$$

Allgemein ($m \neq n \neq p$) kann man über Kommutativität nicht sprechen, weil AB definiert ist, aber BA nicht.

Beispiel 2.4. Hier ist ein lustiges Beispiel für die Multiplikation zweier Matrizen. Das ist natürlich keine Regel, aber nur ein Zufall. Können Sie andere/alle ähnliche Beispiele finden?

$$\begin{pmatrix} 3 & 4 \\ 8 & 7 \end{pmatrix} \cdot \begin{pmatrix} 7 & 2 \\ 4 & 9 \end{pmatrix} = \begin{pmatrix} 37 & 42 \\ 84 & 79 \end{pmatrix}$$

Quadratische Matrizen

Im Fall $m = n$ bezeichnen wir mit $\mathrm{Mat}_n(R) := \mathrm{Mat}_{n\times n}(R)$. Die Matrixmultiplikation ist in diesem Fall ($m = n = p$) eine innere Verknüpfung auf $\mathrm{Mat}_n(R)$ mit neutralem Element die ($n\times n$)-**Einheitsmatrix** (oder Identitätsmatrix):

$$I_n := \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Solange $1 \neq 0$ in R ist die Matrixmultiplikation auf $\mathrm{Mat}_n(R)$ nicht kommutativ (Beispiel 1.149).

Bemerkung 2.5. Für jedes $n \in \mathbb{N}_{>0}$ und jeden Ring R ist $(\mathrm{Mat}_n(R), +, \cdot)$ ein nicht-kommutativer Ring.

Beweis: Das einzige was noch zu überprüfen wäre ist die Distributivität, und diese kann man durch direktes Ausrechnen sehen:

$$[(a_{ij}) + (b_{ij})]\,(c_{ij}) = \left(\sum_{k=1}^{n} (a_{ik} + b_{ik}) c_{kj} \right) = \left(\sum_{k=1}^{n} (a_{ik}c_{kj}) + \sum_{k=1}^{n} (b_{ik}c_{kj}) \right) = (a_{ij})(c_{ij}) + (b_{ij})(c_{ij}).$$

Genauso zeigt man auch, dass $A \cdot (B + C) = A \cdot B + A \cdot C$.

Definition 2.6. Die **allgemeine lineare Gruppe** ist die multiplikative Gruppe $U(\mathrm{Mat}_n(R))$ der invertierbaren Elementen in $\mathrm{Mat}_n(R)$. Wir bezeichnen die Menge der Elementen dieser Gruppe mit

$$\mathrm{GL}_n(R) := \{A \in \mathrm{Mat}_n(R) \ : \ \exists A^{-1} \in \mathrm{Mat}_n(R) \text{ so dass } A^{-1}A = AA^{-1} = I_n\}.$$

Bemerkung 2.7. Weil die Matrixmultiplikation nicht kommutativ ist, ist es nicht offensichtlich, dass eine links-Inverse auch eine rechts-Inverse ist. Weil die Matrixmultiplikation assoziativ ist, haben wir nach Bemerkung 1.97, dass eine links-Inverse auch eine rechts-Inverse ist, und umgekehrt. Es gilt also

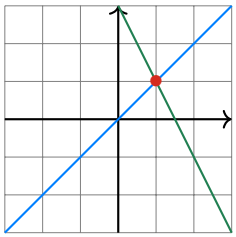
$$\begin{aligned} \mathrm{GL}_n(R) &= \{A \in \mathrm{Mat}_n(R) \ : \ \exists\, A^{-1} \in \mathrm{Mat}_n(R) \text{ so dass } A^{-1}A = I_n\}, \\ &= \{A \in \mathrm{Mat}_n(R) \ : \ \exists\, A^{-1} \in \mathrm{Mat}_n(R) \text{ so dass } AA^{-1} = I_n\}. \end{aligned}$$

[14] 29.11.'23

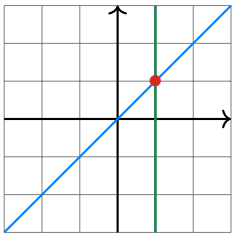
2.2 $\mathbb{K}$-Lineare Gleichungssysteme

Ab jetzt, außer dem Fall in dem es ausdrücklich anders angegeben ist, werden wir **nur Matrizen mit Koeffizienten in einem Körper** $\mathbb{K}$ betrachten.

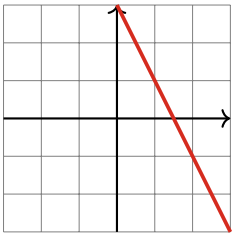
Beispiele:



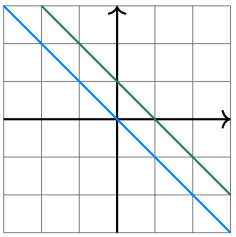
$$\begin{cases} x - y = 0 \\ 2x + y = 3 \end{cases},$$



$$\begin{cases} x - y = 0 \\ 3x \quad\ = 3 \end{cases},$$



$$\begin{cases} -4x - 2y = -6 \\ \frac{5}{3}x + \frac{5}{6}y = \frac{5}{2} \end{cases},$$



$$\begin{cases} x + y = 0 \\ x + y = 1 \end{cases}$$

Es sollte intuitiv klar sein, was eine Gleichung ist. Wir wollen hier diese Intuition nicht abschaffen. Die kurze Einführung hier unten ist nur ein flüchtiger Blick in der Richtung einer mathematisch gründlichen Definition einer Gleichung.

Eine Gleichung mit Unbekannten $x_1, \ldots, x_n$ ist ein Prädikat in dem das "=" Symbol genau ein Mal vorkommt, und für das man die Unbekannten x_i mit Elementen einer Menge ersetzten kann, so dass man eine Aussage bekommt (also ein Satz über dem entschieden werden kann ob es wahr oder falsch ist). Das Wesentliche an einer Gleichung ist das es eine Lösungsmenge hat; diese besteht aus den n-Tupeln die man einsetzten kann um eine wahre Aussage zu bekommen.

Eine **lineare Gleichung** mit Unbekannten $x_1, \ldots, x_n$ und mit Koeffizienten in einem Körper $\mathbb{K}$ ist eine Gleichung in der auf beiden Seiten des "=" Zeichens Polynome von Grad $\leqslant 1$ in $\mathbb{K}[x_1, \ldots, x_n]$ vorkommen. Durch addieren auf beiden Seiten mit demselben Polynom bleibt wegen der Kürzungsregel die Lösungsmenge unverändert. Wir sagen das so eine Operation die Form der Gleichung ändert. Jede lineare Gleichung kann also auf der Form

$$a_1x_1 + \cdots + a_nx_n = b, \qquad \text{mit } a_i, b \in \mathbb{K} \ \ \forall\, i = 1, \ldots, n,$$

gebracht werden. Eine **Lösung** der Gleichung ist ein Element $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ mit der Eigenschaft, dass $a_1\alpha_1 + \cdots + a_n\alpha_n = b$ gilt. Wir nennen eine lineare Gleichung mit Koeffizienten in $\mathbb{K}$ auch $\mathbb{K}$-lineare Gleichung.

Definition 2.8. Ein **lineares Gleichungssystem** (LGS) mit m Gleichungen in n Unbekannten $x_1, \ldots, x_n$ und Koeffizienten im Körper $\mathbb{K}$ ist eine Sammlung von m $\mathbb{K}$-linearen Gleichungen in $x_1, \ldots, x_n$. Eine **Lösung** des Gleichungssystems ist ein Element von $\mathbb{K}^n$ das (gleichzeitig) eine Lösung für alle m Gleichungen ist.

Jedes lineare Gleichungssystem kann auf folgender Form gebracht werden

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases} \tag{2.1}$$

Man definiert für jedes LGS dieser Form die Matrizen $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ und $\mathbf{b} \in \mathrm{Mat}_{m\times 1}(\mathbb{K})$, und sei $\mathbf{x} = (x_1 \dots x_n)^T \in \mathrm{Mat}_{n\times 1}(\mathbb{K}[x_1, \dots, x_n])$. Man kann das LGS aus (2.1) als

$$A \cdot \mathbf{x} = \mathbf{b} \tag{2.2}$$

kodieren. Wir bezeichnen so ein lineares Gleichungssystem LGS$(A, \mathbf{b})$. Die **Koeffizientenmatrix** des Gleichungssystems ist die Matrix A, und die **erweiterte Koeffizientenmatrix** ist die Matrix $(A|\mathbf{b})$. Ein LGS ist **homogen** wenn $\mathbf{b} = 0$. Wenn $\mathbf{b} \neq 0$ dann ist das LGS **inhomogen**.

Die Multiplikation von Matrizen mit n-Tupeln wird durch Identifizieren der Elementen $(\alpha_1, \dots, \alpha_n)$ der Menge $\mathbb{K}^n$ mit $(n \times 1)-$Matrizen $\begin{pmatrix}\alpha_1 \dots \alpha_n\end{pmatrix}^T$ in $\mathrm{Mat}_{n\times 1}(\mathbb{K})$ durchgeführt.

Definition 2.9. Die **Lösungsmenge** eines LGS$(A, \mathbf{b})$ ist die Menge

$$\mathcal{L}(A, \mathbf{b}) := \{\boldsymbol{\alpha} \in \mathbb{K}^n \ : \ A\boldsymbol{\alpha} = \mathbf{b}\}.$$

Ein LGS$(A, \mathbf{b})$ ist **lösbar** wenn $\mathcal{L}(A, \mathbf{b}) \neq \emptyset$; sonst ist es unlösbar.

Satz 2.10. *Seien* $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ *und* $\mathbf{b} \in \mathrm{Mat}_{n\times 1}(\mathbb{K})$. *Sei* $\boldsymbol{\alpha} \in \mathcal{L}(A, \mathbf{b})$ *eine Lösung des* LGS$(A, \mathbf{b})$. *Dann gilt*

$$\mathcal{L}(A, \mathbf{b}) = \boldsymbol{\alpha} + \mathcal{L}(A, 0) = \{\boldsymbol{\alpha} + \boldsymbol{\gamma} \ : \ \boldsymbol{\gamma} \in \mathcal{L}(A, 0)\}.$$

Beweis-Skizze: $\boxed{\subseteq}$ Sei $\boldsymbol{\beta} = (\beta_1 \dots \beta_n)^T \in \mathcal{L}(A, \mathbf{b})$; also $A\boldsymbol{\beta} = \mathbf{b}$. Dann gilt

$$A\boldsymbol{\beta} - A\boldsymbol{\alpha} = A(\boldsymbol{\beta} - \boldsymbol{\alpha}) = \mathbf{b} - \mathbf{b} = 0.$$

Also $\boldsymbol{\gamma} := \boldsymbol{\beta} - \boldsymbol{\alpha} \in \mathcal{L}(A, 0)$ und $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{\gamma} \in (\boldsymbol{\alpha} + \mathcal{L}(A, 0))$.
$\boxed{\supseteq}$ Sei $\boldsymbol{\gamma} \in \mathcal{L}(A, 0)$, also $\boldsymbol{\alpha} + \boldsymbol{\gamma} \in (\boldsymbol{\alpha} + \mathcal{L}(A, 0))$. Dann gilt

$$A(\boldsymbol{\alpha} + \boldsymbol{\gamma}) = A\boldsymbol{\alpha} + A\boldsymbol{\gamma} = \mathbf{b} + 0 = \mathbf{b},$$

also $\boldsymbol{\alpha} + \boldsymbol{\gamma} \in \mathcal{L}(A, \mathbf{b})$. Q.E.D.

Satz 2.11. *Für die Lösungsmenge* $\mathcal{L}(A, 0)$ *eines homogenes* LGS$(A, 0)$ *gilt*

(i) $0 \in \mathcal{L}(A, 0)$

(ii) $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{L}(A, 0) \Rightarrow \boldsymbol{\alpha} + \boldsymbol{\beta} \in \mathcal{L}(A, 0)$

(iii) $\boldsymbol{\alpha} \in \mathcal{L}(A, 0)$ *und* $\lambda \in \mathbb{K} \Rightarrow \lambda \cdot \boldsymbol{\alpha} \in \mathcal{L}(A, 0)$.

Beweis-Skizze: Übung Q.E.D.

Lemma 2.12. *Sei* $U \in \mathrm{GL}_m(\mathbb{K})$ *eine invertierbare Matrix, und sei* $\mathrm{LGS}(A, \mathbf{b})$ *ein lineares Gleichungssystem mit* $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ *und* $\mathbf{b} \in \mathrm{Mat}_{m\times 1}$. *Dann gilt*

$$\mathcal{L}(A, \mathbf{b}) = \mathcal{L}(U \cdot A, U \cdot \mathbf{b}).$$

Beweis-Skizze:

$$\boldsymbol{\alpha} \in \mathcal{L}(A, \mathbf{b}) \quad \Leftrightarrow \quad A \cdot \boldsymbol{\alpha} = \mathbf{b} \quad \underset{U^{-1}\cdot}{\overset{U\cdot}{\Longrightarrow}} \quad U \cdot A \cdot \boldsymbol{\alpha} = U \cdot \mathbf{b} \quad \Leftrightarrow \quad \boldsymbol{\alpha} \in \mathcal{L}(U \cdot A, U \cdot \mathbf{b}).$$

Q.E.D.

2.2.1 Elementaroperationen

Das **Kronecker-Delta** oder Kronecker Symbol δ_{ij} ist für jedes Paar (i, j) definiert als

$$\delta_{ij} := \begin{cases} 0 & \text{wenn } i \neq j \\ 1 & \text{wenn } i = j. \end{cases}$$

Insbesondere wird für uns $\delta_{ij} \in \mathbb{K}$ das additive oder das multiplikative neutrale Element in $\mathbb{K}$ sein.

Definition 2.13. Seien $m \in \mathbb{N}_{>0}$ und seien $k, l \in \{1, \ldots, n\}$. Die $(n \times n)-$**Standardmatrix** $E_{k,l} = (e_{ij})$ ist die quadratische $(m \times m)-$Matrix mit

$$e_{ij} = \begin{cases} 0 & \text{wenn } (i, j) \neq (k, l) \\ 1 & \text{wenn } (i, j) = (k, l). \end{cases}$$

Man kann das kompakter aufschreiben als $E_{kl} = (\delta_{ik}\delta_{lj})_{1\leqslant i,j\leqslant m}$. Wichtig ist, dass E_{kl} ein einziger nicht-trivialer Eintrag hat: eine 1 an der Stelle (k, l). Wir haben

$$I_m = \sum_{i=1}^{m} E_{ii}.$$

Wir packen die "m" in dieser Notation nicht dazu, da wir meistens mit einem von Anfang an fixierten m arbeiten werden.

Bemerkung 2.14.

1. Für die Standardmatrizen gilt
$$E_{ij}E_{kl} = \begin{cases} 0, & \text{wenn } j \neq k \\ E_{il}, & \text{wenn } j = k \end{cases}$$
2. Für jede Matrix A mit Zeilen $Z_1, \ldots, Z_m$, seien $Y_1, \ldots, Y_m$ die Zeilen der Matrix $E_{kl} \cdot A$. Wir haben
$$Y_i = \begin{cases} 0 & \text{wenn } i \neq k \\ Z_l & \text{wenn } i = k \end{cases}.$$
Also $E_{kl} \cdot A$ hat eine einzige nichttriviale Zeile: die Zeile k, und diese ist gleich mit der Zeile l von A.

Definition 2.15. Die **elementare Zeilenumformungen** sind Abbildungen $\Upsilon_\bullet : \mathrm{Mat}_{m\times n}(\mathbb{K}) \longrightarrow \mathrm{Mat}_{m\times n}(\mathbb{K})$ eine der folgenden Wirkungen haben:

$\boxed{\Upsilon_{k\leftrightarrow l}}$ Vertauscht die Zeilen k und l, mit $k, l \in \{1, \ldots, m\}$:

$$A = \begin{pmatrix} \vdots \\ Z_{\mathbf{k}} \\ \vdots \\ Z_l \\ \vdots \end{pmatrix} \overset{\Upsilon_{k\leftrightarrow l}}{\longmapsto} \begin{pmatrix} \vdots \\ Z_l \\ \vdots \\ Z_{\mathbf{k}} \\ \vdots \end{pmatrix}$$

Dadurch verstehen wir, dass die Zeilen Z_k und Z_l vertauscht werden, und alles andere unverändert bleibt.

$\boxed{\Upsilon_{k\to\lambda\cdot k}}$ Multipliziert die Zeile k mit dem nicht-trivialen Skalar $\lambda \in \mathbb{K}^\times$:

$$A = \begin{pmatrix} \vdots \\ Z_k \\ \vdots \end{pmatrix} \longmapsto \begin{pmatrix} \vdots \\ \lambda \cdot Z_k \\ \vdots \end{pmatrix}$$

Dadurch verstehen wir, dass jeder Eintrag in der Zeile Z_k mit λ multipliziert wird, und alle andere Einträge unverändert bleiben.

$\boxed{\Upsilon_{k\to k+\lambda\cdot l}}$ Addiert λ mal die Zeile l zu der Zeile k, mit $\lambda \in \mathbb{K}$:

$$A = \begin{pmatrix} \vdots \\ Z_{\mathbf{k}} \\ \vdots \\ Z_l \\ \vdots \end{pmatrix} \longmapsto \begin{pmatrix} \vdots \\ Z_{\mathbf{k}} + \lambda \cdot Z_l \\ \vdots \\ Z_l \\ \vdots \end{pmatrix}$$

Dadurch verstehen wir, dass jeder Eintrag a_{kj} in der Zeile Z_k durch $a_{kj} + \lambda a_{lj}$ ersetzt wird, und alle andere Einträge unverändert bleiben.

Bemerkung 2.16. Die Zeilenumformung $\Upsilon_{k\to\lambda\cdot k}$ ist ein Sonderfall der Zeilenumformung $\Upsilon_{k\to k+\lambda\cdot l}$, nämlich mit $k = l$ und $\lambda \neq -1$.

Satz 2.17. *Jede Elementare Zeilenumformung ist durch links-multiplizieren mit einer der folgenden drei Typen von invertierbaren Matrix erhalten:*

$$\begin{aligned} U_{k\leftrightarrow l} &= I_n - E_{kk} - E_{ll} + E_{k,l} + E_{l,k} \\ U_{k\to\lambda k} &= I_n + (\lambda - 1)E_{kk} \\ U_{k\to k+\lambda l} &= I_n + \lambda E_{kl} \end{aligned}$$

Insbesondere sind elementare Zeilenumformungen bijektive Abbildungen.

Beweis-Skizze: Aus Bemerkung 2.14 sieht man, dass die $\Upsilon_\bullet$ Abbildungen genau links-multiplizieren mit der entsprechenden $U_\bullet$ Matrix gegeben sind. Die Invertierbarkeit folgt aus

$$\begin{aligned} U_{k\leftrightarrow l} \cdot U_{k\leftrightarrow l} &= I_n \\ U_{i\to\lambda i} \cdot U_{i\to\frac{1}{\lambda}i} &= I_n \\ U_{i\to i+\lambda j} \cdot U_{i\to i-\lambda j} &= I_n \text{ wenn } i \neq j. \end{aligned}$$

Q.E.D.

Definition 2.18. Eine **Elementarmatrix** ist eine Matrix der Form $U_{k\leftrightarrow l}$, $U_{k\to\lambda k}$, oder $U_{k\to k+\lambda l}$ aus Satz 7.2.2. Wenn die Form nicht festgelegt ist, dann schreiben wir $U_\bullet$ für eine beliebige Elementarmatrix.

Definition 2.19. Eine **Zeilenumformung** ist die Verknüpfung endlich-vieler elementaren Zeilenumformungen. Wir sagen, dass eine Matrix $A' \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ eine Zeilenumformung einer Matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ ist, genau dann, wenn es eine Zeilenumformung $\Upsilon : \mathrm{Mat}_{m\times n}(\mathbb{K}) \longrightarrow \mathrm{Mat}_{m\times n}(\mathbb{K})$ gibt so dass $\Upsilon(A) = A'$.

Eine Zeilenumformung ist also eine bijektive Abbildung $\Upsilon : \mathrm{Mat}_{m\times n}(\mathbb{K}) \longrightarrow \mathrm{Mat}_{m\times n}(\mathbb{K})$, die durch links-multiplizieren mit einer Matrix U die in der Untergruppe $\mathrm{Span}_{\mathbb{K}}\{U_\bullet = \text{ Elementarmatrix}\} \leq \mathrm{GL}_n(\mathbb{K})$. Wir werden in Korollar 2.38 sehen, dass eigentlich die Elementarmatrizen ganz $\mathrm{GL}_n(\mathbb{K})$ erzeugen. Wir nennen eine solche Matrix eine **Zeilenumformmatrix** (ZU-Matrix) oder (ZUM).

Bemerkung 2.20. Für jede $(n \times n)-$Elementarmatrix haben wir

$$U_\bullet = \Upsilon_\bullet(I_n).$$

Bemerkung 2.21. Für jede $(n \times n)-$Standardmatrix E_{kl} haben wir, dass die $((n+1) \times (n+1))-$Standardmatrix

$$E_{k+1,l+1} = \left(\begin{array}{c|ccc} 0 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & E_{kl} & \\ 0 & & & \end{array}\right) \quad \text{und} \quad I_{n+1} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & I_n & \\ 0 & & & \end{array}\right)$$

Aus dieser einfacher Tatsache folgt, dass für eine $(n \times n)-$Elementarmatrix $U_\bullet$, die Matrix

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & U_\bullet & \\ 0 & & & \end{array}\right) = U'_\bullet$$

eine $((n+1) \times (n+1))-$Elementarmatrix ist.

Korollar 2.22. Seien $A, A' \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ und $\mathbf{b}, \mathbf{b}' \in \mathrm{Mat}_{m\times 1}(\mathbb{K})$. Setze $C := (A|\mathbf{b})$ und sei $C' = (A'|\mathbf{b}')$ eine Zeilenumformung von C. Dann gilt

$$\mathcal{L}(A, \mathbf{b}) = \mathcal{L}(A', \mathbf{b}')$$

Beweis-Skizze: Es folgt aus Satz 7.2.2 und Lemma 2.12. Q.E.D.

[15] 4.12.'23

2.2.2 Gaußscher Algorithmus

Wir werden nun ein Verfahren definieren, das es uns ermöglicht, die Lösungsmenge eines linearen Gleichungssystems schnell zu bestimmen. Als Nebenprodukt werden wir auch eine schnelle Methode finden, um die Invertierbarkeit zu überprüfen und die Inverse zu berechnen. Dies basiert auf Zeilenumformungen.

Definition 2.23. Eine Matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ ist in **Zeilenstufenform** (ZSF), wenn es $r \in \{1,\dots,m\}$ und $1 \leqslant st_1 < st_2 < \cdots < st_r \leqslant n$ existieren, sodass

(ZSF 1) $a_{ij} = 0$ für $i > r$ und $j = 1 \dots n$,

(ZSF 2) $a_{ik} = 0$ für $i = 1 \dots r$ und $k = 1 \dots (st_i - 1)$,

(ZSF 3) $a_{ist_i} \neq 0$ für $i = 1 \dots r$.

Eine Matrix A ist in **reduzierte Zeilenstufenform** (RZSF), wenn sie in ZSF ist und es gelten

(RZSF 1) $a_{ist_i} = 1$ für $i = 1 \dots r$,

(RZSF 2) $a_{kst_i} = 0$ für alle $i = 2 \dots r$ und $k = 1 \dots (i-1)$.

Wenn A in ZSF ist, dann sagen wir: *die Matrix A hat r Stufen*, und nennen die Einträge a_{ist_i} **Stufen**.

Beispiele: Seien A, B, C folgende Matrizen mit Koeffizienten in $\mathbb{R}$:

$$A = \begin{pmatrix} 2 & 5 & 6 & 10 \\ 4 & 10 & 15 & 26 \\ 6 & 15 & 21 & 36 \end{pmatrix}$$ ist nicht in ZSF.

$$B = \begin{pmatrix} 2 & 5 & 6 & 10 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$ ist in ZSF mit $r_B = 2$, $st_1 = 1$, $st_2 = 3$, aber nicht in RZSF.

$$C = \begin{pmatrix} 1 & 5/2 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$ ist in RZSF mit $r_C = 2$, $st_1 = 1$, $st_2 = 3$.

Die Matrizen B und C können durch folgende elementare Zeilenumformungen aus A erhalten werden:

$$\begin{aligned} B &= U_{\mathbf{3}\to\mathbf{3}+(-1)\cdot\mathbf{2}} \cdot U_{\mathbf{3}\to\mathbf{3}+(-3)\cdot\mathbf{1}} \cdot U_{\mathbf{2}\to\mathbf{2}+(-2)\cdot\mathbf{1}} \cdot A \\ C &= U_{\mathbf{1}\to\mathbf{1}+(-3)\cdot\mathbf{2}} \cdot U_{\mathbf{2}\to(\frac{1}{3})\cdot\mathbf{2}} \cdot U_{\mathbf{1}\to(\frac{1}{2})\cdot\mathbf{1}} \cdot B. \end{aligned}$$

Bemerkung 2.24. Wenn $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ in ZSF oder RZSF ist, und $m' \leqslant m$ und $n' \leqslant n$, dann ist die Teilmatrix $A|_{\{1\dots m'\}\times\{1\dots n'\}}$ auch in ZSF, beziehungsweise RZSF.

Bemerkung 2.25. Seien $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ und $B \in \mathrm{Mat}_{m\times p}(\mathbb{K})$ zwei Matrizen mit der Gleichen Anzahl von Zeilen: m. Sei $U \in \mathrm{Mat}_{q\times m}(\mathbb{K})$ eine Matrix mit m Spalten. Aus der Definition der Matrixmultiplikation folgt direkt, dass

$$U \cdot \left(\begin{array}{c|c} A & B \end{array}\right) = \left(\begin{array}{c|c} UA & UB \end{array}\right).$$

Satz 2.26 (Gausscher Algorithmus)**.** *Für jede Matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ gibt es eine Zeilenumformung Υ so dass $\Upsilon(A)$ in RZSF ist.*

Beweis-Skizze: Wir beweisen zu erst durch vollständige Induktion nach m, dass A in ZSF durch Zeilenumformung gebracht werden kann.

IA: $m = 1$

Dann ist A schon in ZSF: wenn $A = 0$, mit $r_0 = 0$,
wenn $A \neq 0$, mit $r_A = 1$ und $st_1 = \min\{j \ : \ a_{1j} \neq 0\}$.

IS: $m - 1 \Rightarrow m$

Wenn $A = 0$ dann ist A schon in ZSF mit $r = 0$.
Wenn $A \neq 0$, dann $\exists\ (i, j)$ mit $a_{ij} \neq 0$. Wir setzen $st_1 := \min\{j \ : \ \exists\ i \text{ mit } a_{ij} \neq 0\}$ und wählen ein $i_1 \in \{1, \ldots, m\}$ mit $a_{i_1 st_1} \neq 0$. Wir tauschen die Zeilen 1 und i_1:

$$A' := \Upsilon_{i_1 \leftrightarrow 1}(A) = U_{i_1 \leftrightarrow 1} \cdot A.$$

Wir definieren für jedes $i = 2, \ldots, m$ den Skalar

$$\lambda_i = \frac{a'_{ist_1}}{a'_{1st_1}} \in \mathbb{K}.$$

(**Hier war es essenziell, dass die Koeffizienten in einem Körper sind!**)
Durch Zeilenumformungen der Form $\Upsilon_{i \to i - \lambda_i \cdot 1}$ für jedes $i = 2, \ldots, m$ bekommen wir

$$A'' := \left(\prod_{i=2}^{m} U_{i \to i - \lambda_i \cdot 1}\right) \cdot A' = \left(\begin{array}{cccc|c} 0 & \ldots & 0 & a'_{1st_1} & * \ \ldots \ * \\ \hline 0 & \ldots & 0 & 0 & \\ \vdots & & \vdots & \vdots & A_{m-1} \\ 0 & \ldots & 0 & 0 & \end{array}\right).$$

Aus der induktiven Voraussetzung existiert eine ZU-Matrix $U' \in \mathrm{Mat}_{(m-1)\times n}(\mathbb{K})$, so dass die Matrix $B_{m-1} := U' \cdot A_{m-1}$ in ZSF ist. Nach Bemerkung 2.21 ist auch die Matrix

$$U := \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{array}\right)$$

eine ZU-Matrix. Wir bekommen die ZSF von A als $B := U \cdot A''$. Wir bezeichnen die "Stufen" mit $w_i := b_{ist_i}$.

Um auf der RZSF zu kommen, multiplizieren wir B zu erst mit der ZU-Matrix

$$\prod_{i=1}^{r_B} U_{(\frac{1}{w_i})\cdot i}$$

und dann für jedes $i = 2 \dots r_B$ mit der ZU-Matrix

$$\prod_{k=1}^{i-1} U_{k \to k - b_{k st_i} \cdot i}.$$

Q.E.D.

Satz 2.27. *Die durch dem Gaußschen Algorithmus erhaltene RZSF Matrix ist eindeutig bestimmt.*

Beweis-Skizze: Wir beweisen den Satz durch Induktion nach der Anzahl von Spalten n.

IA: $n = 1$

Wenn $A = 0$ dann ist die RZSF auch A.
Wenn $A \neq 0$, dann ist die RZSF $\begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}^T$.

IS: $n - 1 \Rightarrow n$ Sei $A = (A'|\mathbf{a})$ und seien $B = (B'|\mathbf{b})$ und $C = (C'|\mathbf{c})$ Matrizen in RZSF die durch Zeilenumformung aus A erhalten wurden. Dann gilt aus Korollar 2.22

$$\mathcal{L}(A, 0) = \mathcal{L}(B, 0) = \mathcal{L}(C, 0).$$

Aus der induktiven Voraussetzung folgt $B' = C'$, weil beide in RZSF sind, aus A' durch ZU erhalten wurden, und $n - 1$ Spalten haben. Wir müssen also nur noch, dass $\mathbf{b} = \mathbf{c}$ zeigen.
Behauptung: Für jede Matrix M die in RZSF mit r und $1 \leqslant st_1 < \dots < st_r \leqslant n$ ist, haben wir

$$(\forall\ \boldsymbol{\alpha} \in \mathcal{L}(M, 0) \ \text{ gilt } \alpha_n = 0) \iff st_r = n.$$

Beweis der Behauptung:
$\Rightarrow$ Wenn $st_r < n$, sei $0 \neq t \in \mathbb{K}$ und $\boldsymbol{\alpha} \in \mathbb{K}^n$ mit

$$\alpha_j = \begin{cases} -m_{in} \cdot t & \text{wenn } j = st_i \\ t & \text{wenn } j = n \\ 0 & \text{sonst} \end{cases}$$

Dann ist $\boldsymbol{\alpha} \in \mathcal{L}(M, 0)$ mit $\alpha_n = t \neq 0$.
$\Leftarrow$ Wenn $st_r = n$, dann ist die $r-$te Gleichung $x_n = 0$. ■

Wir nehmen an, dass $\mathbf{b} \neq \mathbf{c}$. Dann $\exists\ i$ mit $b_{in} \neq c_{in}$. Die Differenz der i-ten Gleichungen in LGS$(B, 0)$ und LGS$(C, 0)$ ist die Gleichung $(b_{in} - c_{in})x_n = 0$. Da jede Lösung auch diese Gleichung erfüllen muss, gilt also

$$\text{wenn } \mathbf{b} \neq \mathbf{c}, \ \text{ dann } \boldsymbol{\alpha} \in \mathcal{L}(B, 0) = \mathcal{L}(C, 0) \Rightarrow \alpha_n = 0.$$

Aus der Behauptung folgt dann $st_r = n$ sowohl für B als auch für C, und also

$$\mathbf{b} = \mathbf{c} = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \end{pmatrix}.$$

– ein Widerspruch ↯ zu $\mathbf{b} \neq \mathbf{c}$. Q.E.D.

Wir können jetzt von **der** reduzierten Zeilenstufenform einer Matrix sprechen:

Definition 2.28. Sei $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ eine Matrix. Die **reduzierte Zeilenstufenform** von A ist eindeutig-bestimmte RZSF Matrix aus Satz 2.26 und Satz 2.27. Wir bezeichnen diese mit $\mathrm{RZ}(A)$.

Definition 2.29. Der **Zeilenstufenrang** einer Matrix A ist die Anzahl der Stufen in $\mathrm{RZ}(A)$. Wir bezeichnen es mit ${}^{\mathrm{ZS}}\mathrm{Rang}\, A$.

Wir werden später sehen, dass es mehrere natürliche Möglichkeiten gibt diese Zahl zu beschreiben, und werden den Zeilenstufenrang später einfach *Rang* der Matrix nennen.

[17] 6.12.'23

2.2.3 RZSF und die Lösungsmenge eines LGS

Satz 2.30. *Sei* $A\mathbf{x} = \mathbf{b}$ *ein LGS mit* $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ *und* $\mathbf{b} \in \mathrm{Mat}_{m\times 1}(\mathbb{K})$. *Sei* $(A'|\mathbf{b}') := \mathrm{RZ}(A|\mathbf{b})$, *mit* $\mathbf{b}' = \begin{pmatrix} b'_1 & b'_2 & \dots & b'_r & b'_{r_{A'}+1} & 0 & \dots & 0 \end{pmatrix}^T$, *wobei* $r_{A'}$ *der Zeilenstufenrang von* A' *ist. Wir haben*

$$\mathcal{L}(A, \mathbf{b}) \neq \emptyset \iff b'_{r_{A'}+1} = 0.$$

In diesem Fall gibt es eine bijektive Abbildung $\Phi : \mathbb{K}^{n-r} \longrightarrow \mathcal{L}(A, \mathbf{b})$.

Bemerkung 2.31. Die Bedingung $b'_{r_{A'}+1} = 0$ in Satz 2.30 ist äquivalent zu ${}^{\mathrm{ZS}}\mathrm{Rang}(A'|\mathbf{b}') = {}^{\mathrm{ZS}}\mathrm{Rang}\, A'$. Anders gesagt, $\mathrm{RZ}(A|\mathbf{b})$ hat keine Stufe in der letzten Spalte. Die Äquivalenz aus dem Satz ist also:

Ein LGS hat mindestens eine Lösung, genau dann wenn der Rang der Koeffizientenmatrix gleich mit dem Ran

Beweis-Skizze: Wir haben nach Korollar 2.22, dass $\mathcal{L}(A, \mathbf{b}) = \mathcal{L}(A', \mathbf{b}')$. Die $(r+1)$-te Gleichung im $\mathrm{LGS}(A', \mathbf{b}')$ ist

$$0x_1 + \dots + 0x_n = b_{r+1}.$$

Also, wenn $b_{r+1} \neq 0$, dann ist $\mathcal{L}(A', \mathbf{b}') = \emptyset$.
Wenn $b_{r+1} = 0$, dann haben wir $\boldsymbol{\alpha} \in \mathcal{L}(A, b) = \mathcal{L}(A', b')$ wobei

$$\alpha_k = \begin{cases} b'_i & \text{falls } k = st_i \\ 0 & \text{sonst} \end{cases}.$$

Sei $r := r_{A'}$. Um die Bijektion zu definieren, nehmen wir an, dass $st_i = i \ \forall i = 1 \dots r$. [a] Dann, definieren wir $\varphi : \mathbb{K}^{n-r} \longrightarrow \mathcal{L}(A', 0)$ durch

$$\varphi(t_1, \dots, t_{n-r}) = (-\sum_{j=r+1}^{n} a'_{1j} t_{j-r}, \dots, -\sum_{j=r+1}^{n} a'_{rj} t_{j-r}, t_1, t_2, \dots, t_{n-r}).$$

Wir haben $\varphi(\mathbf{t}) \in \mathcal{L}(A', 0)$ für alle $\mathbf{t} \in \mathbb{K}^{n-r}$, weil jede nicht-triviale Gleichung in $\mathrm{LGS}(A, 0)$ die

Form

$$x_i + a_{i,r+1}x_{r+1} + \cdots + a_{in}x_n \quad = \quad 0 \tag{2.3}$$

hat für ein $i \in \{1, \ldots, r\}$. Also φ ist wohldefiniert. Offensichtlich ist φ injektiv (man muss sich nur die Komponenten $(r+1)$ bis n dafür anschauen). Weiterhin, aus (2.3) ist jede Lösung von $\mathrm{LGS}(A', 0)$ der Form $\varphi(\mathbf{t})$ mit $\mathbf{t} \in \mathbb{K}^{n-r}$. Also $\varphi : \mathbb{K}^{n-r} \longrightarrow \mathcal{L}(A', 0)$ ist bijektiv. Aus Satz 2.10 folgt jetzt, dass die Abbildung $\Phi : \mathbb{K}^{n-r} \longrightarrow \mathcal{L}(A, \mathbf{b}) = \mathcal{L}(A', \mathbf{b}')$ gegeben durch

$$\Phi(\mathbf{t}) := \boldsymbol{\alpha} + \varphi(\mathbf{t})$$

die gesuchte Bijektion ist. Q.E.D.

[a] Wenn das nicht so ist, dann können wir die Variablen "neu benennen" so dass es stimmt. (Dieses neue-beschriftete LGS ist völlig analog zum alten; technisch haben wir eine Permutation der Index Menge angewendet, und das ist immer reversibel).

Korollar 2.32. Jedes homogenes LGS $(A|0)$ mit $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ hat eine nicht-triviale Lösung wenn $m < n$.

Bemerkung 2.33. Der Satz 2.30 sagt also, dass

$$\mathrm{LGS}(A, \mathbf{b}) \text{ hat eine Lösung} \quad \Longleftrightarrow \ {}^{\mathrm{ZS}}\mathrm{Rang}(A|\mathbf{b}) = {}^{\mathrm{ZS}}\mathrm{Rang}(A).$$

Korollar 2.34. Ein lineares Gleichungssystem $A \cdot \mathbf{x}^T = \mathbf{b}$ mit n Variablen hat eine eindeutige Lösung, genau dann, wenn ${}^{\mathrm{ZS}}\mathrm{Rang}(A|\mathbf{b}) = {}^{\mathrm{ZS}}\mathrm{Rang}(A) = n$.

[17] 11.12.'23

2.2.4 RZSF und invertierbare Matrizen

Satz 2.35. *Eine quadratische Matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ ist invertierbar genau dann, wenn ${}^{\mathrm{ZS}}\mathrm{Rang}\, A = n$.*

Beweis-Skizze: Die einzige $(n \times n)-$Matrix in RZSF mit n Stufen ist die Einheitsmatrix I_n. Wenn ${}^{\mathrm{ZS}}\mathrm{Rang}\, A = n$, dann existiert eine ZU-Matrix U mit

$$U \cdot A = \mathrm{RZ}(A) = I_n,$$

also A ist invertierbar.
Wenn A invertierbar ist, dann ist auch $\mathrm{RZ}(A)$ invertierbar, weil es das Produkt zweier invertierbaren Matrizen ist. Wäre der ${}^{\mathrm{ZS}}\mathrm{Rang}\, A < n$, dann wäre die letzte Zeile von $\mathrm{RZ}(A)$ die Nullzeile $\begin{pmatrix}0 & \ldots & 0\end{pmatrix}$. Dann wäre auch die letzte Zeile in $I_n = \mathrm{RZ}(A) \cdot (\mathrm{RZ}(A))^{-1}$ die Nullzeile – ein Widerspruch ↯. Q.E.D.

Satz 2.36. *Für eine quadratische invertierbare Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ gilt*

$$\mathrm{RZ}\left(A \;\middle|\; I_n\right) = \left(I_n \;\middle|\; A^{-1}\right).$$

Beweis-Skizze: Weil A invertierbar ist, folgt aus Satz 2.35, dass $^{\text{ZS}}\text{Rang}\, A = n$. Es existiert eine also ZUM U mit $U \cdot A = I_n$. Also, weil I_n die einzige $n \times n$ Matrix die in RZSF n Stufen hat ist, haben wir $U = A^{-1}$. Aus Bemerkung 2.25 folgt also

$$U \cdot \left(A \;\middle|\; I_n\right) = \left(U \cdot A \;\middle|\; U \cdot I_n\right) = \left(I_n \;\middle|\; A^{-1}\right) = \text{RZ}\left(A \;\middle|\; I_n\right).$$

Alternative: Man kann das auch anders zeigen. Weil A invertierbar ist, hat es nach Satz 2.35 $^{\text{ZS}}\text{Rang}\, A = n$, und somit hat $(A \,|\, I_n)$ folgende reduzierte Zeilenstufenform:

$$\text{RZ}\left(A \;\middle|\; I_n\right) = \left(I_n \;\middle|\; X\right).$$

Es gilt

$$\left(A \;\middle|\; I_n\right) \cdot \begin{pmatrix} -I_n \\ \hline A \end{pmatrix} = 0,$$

wobei 0 die $(n \times n)$-Nullmatrix ist. Das heißt, dass die Spalten der Matrix $\left(-I_n \;\middle|\; A\right)^T$ Lösungen des homogenes LGS $\left(A \;\middle|\; I_n\right) \cdot \mathbf{x} = 0$ sind. Also, weil $\left(I_n \;\middle|\; X\right)$ durch Zeilenumformung erhalten wurde, hat das assoziierte homogene LGS dieselben Lösungen. Das heißt,

$$\left(I_n \;\middle|\; X\right) \cdot \begin{pmatrix} -I_n \\ \hline A \end{pmatrix} = I_n \cdot (-I_n) + A \cdot X = 0.$$

Q.E.D.

Korollar 2.37. Für eine quadratische $(n \times n)-$Matrix A, das Gaußsche Verfahren angewendet an $\left(A \;\middle|\; I_n\right)$ überprüft die Invertierbarkeit von A und, falls A invertierbar ist, liefert es auch die Inverse von A.

Korollar 2.38. Die $(n \times n)-$Elementarmatrizen erzeugen die allgemeine lineare Gruppe $\text{GL}_n(\mathbb{K})$.

Beweis-Skizze: Die Elementarmatrizen erzeugen alle ZU-Matrizen. Es reicht also zu zeigen, dass jede invertierbare Matrix eine ZU-Matrix ist.
Sei $A \in \text{GL}_n(\mathbb{K})$. Dann ist $A^{-1} \in \text{GL}_n(\mathbb{K})$. Aus dem Beweis von Satz 2.35 folgt, dass es eine ZU-Matrix U gibt mit $U \cdot A^{-1} = \text{RZ}(A^{-1}) = I_n$. Also $A = U$ ist eine Zeilenumformmatrix. Q.E.D.

Beispiele: 1. Sei $A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$. Wir wenden den Gaußschen Algorithmus an um die (eventuelle) Inverse zu finden. ("eventuelle" weil wir nicht a priori wissen ob A invertierbar ist).

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \xrightarrow{2 \to 2 - \frac{3}{2}1} \left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{-3}{2} & 1 \end{array}\right) \xrightarrow{1 \to 1 - 6 \cdot 2} \left(\begin{array}{cc|cc} 2 & 0 & 10 & -6 \\ 0 & \frac{1}{2} & \frac{-3}{2} & 1 \end{array}\right)$$

$$\left(\begin{array}{cc|cc} 2 & 0 & 10 & -6 \\ 0 & \frac{1}{2} & \frac{-3}{2} & 1 \end{array}\right) \xrightarrow{1 \to \frac{1}{2}1} \left(\begin{array}{cc|cc} 1 & 0 & 5 & -3 \\ 0 & \frac{1}{2} & -\frac{3}{2} & 1 \end{array}\right) \xrightarrow{2 \to 2 2} \left(\begin{array}{cc|cc} 1 & 0 & 5 & -3 \\ 0 & 1 & -3 & 2 \end{array}\right)$$

Probe:

$$\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \cdot \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 2\cdot 5 - 3\cdot 3 & -2\cdot 3 + 3\cdot 2 \\ 3\cdot 5 - 5\cdot 3 & -3\cdot 3 + 5\cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

2. Das gleiche für $A = \begin{pmatrix} 1 & -2 & 9 \\ 2 & 0 & 3 \\ 0 & -1 & 4 \end{pmatrix}$.

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 9 & 1 & 0 & 0 \\ 2 & 0 & 3 & 0 & 1 & 0 \\ 0 & -1 & 4 & 0 & 0 & 1 \end{array}\right) \xrightarrow{2\to 2-2\cdot 1} \left(\begin{array}{ccc|ccc} 1 & -2 & 9 & 1 & 0 & 0 \\ 0 & 4 & -15 & -2 & 1 & 0 \\ 0 & -1 & 4 & 0 & 0 & 1 \end{array}\right) \xrightarrow{2\leftrightarrow 3} \left(\begin{array}{ccc|ccc} 1 & -2 & 9 & 1 & 0 & 0 \\ 0 & -1 & 4 & 0 & 0 & 1 \\ 0 & 4 & -15 & -2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow{3\to 3+4\cdot 2} \left(\begin{array}{ccc|ccc} 1 & -2 & 9 & 1 & 0 & 0 \\ 0 & -1 & 4 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 4 \end{array}\right) \xrightarrow{2\to(-1)\cdot 2} \left(\begin{array}{ccc|ccc} 1 & -2 & 9 & 1 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & -1 \\ 0 & 0 & 1 & -2 & 1 & 4 \end{array}\right) \xrightarrow{1\to 1+2\cdot 2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & -1 \\ 0 & 0 & 1 & -2 & 1 & 4 \end{array}\right)$$

$$\xrightarrow[2\to 2+4\cdot 3]{1\to 1-3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & -6 \\ 0 & 1 & 0 & -8 & 4 & 15 \\ 0 & 0 & 1 & -2 & 1 & 4 \end{array}\right)$$

Probe:

$$\begin{pmatrix} 1 & -2 & 9 \\ 2 & 0 & 3 \\ 0 & -1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -6 \\ -8 & 4 & 15 \\ -2 & 1 & 4 \end{pmatrix} =$$

$$\begin{pmatrix} 1\cdot 3 + (-2)\cdot(-8) + 9\cdot(-2) & 1\cdot(-1) + (-2)\cdot 4 + 9\cdot 1 & 1\cdot(-6) + (-2)\cdot 15 + 9\cdot 4 \\ 2\cdot 3 + 0\cdot(-8) + 3\cdot(-2) & 2\cdot(-1) + 0\cdot 4 + 3\cdot 1 & 2\cdot(-6) + 0\cdot 15 + 3\cdot 4 \\ 0\cdot 3 + (-1)\cdot(-8) + 4\cdot(-2) & 0\cdot(-1) + (-1)\cdot 4 + 4\cdot 1 & 0\cdot(-6) + (-1)\cdot 15 + 4\cdot 4 \end{pmatrix} =$$

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

3. Finde die Lösungsmenge des LGS mit 3 Unbekannten, 3 Gleichungen und Koeffizienten in $\mathbb{R}$:

$$\begin{cases} x + y + z = 2 \\ x - y + 2z = 1 \\ x - 3y + 3z = 0 \end{cases}.$$

Wir haben als Koeffizientenmatrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & -3 & 3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$, und als erweiterte Koeffizientenmatrix $(A|\mathbf{b}) = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & -3 & 3 & 0 \end{pmatrix} \in \mathrm{Mat}_{3\times 4}(\mathbb{R})$. Um die Lösungsmenge zu bestimmen, berechnen wir die RZ der

erweiterten Koeffizientenmatrix:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & -3 & 3 & 0 \end{array}\right) \xrightarrow[3\to 3-1]{2\to 2-1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -2 & 1 & -1 \\ 0 & -4 & 2 & -2 \end{array}\right) \xrightarrow{3\to 3-2\cdot 2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{2\to -\frac{1}{2}\cdot 2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{1\to 1-2} \left(\begin{array}{ccc|c} 1 & 0 & \frac{3}{2} & \frac{3}{2} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Die erweiterte Koeffizientenmatrix hat also 2 Stufen, mit Stufenindizes $\mathscr{s}t_1 = 1, \mathscr{s}t_2 = 2$. Weil $2 < 4$ haben wir also keine Stufe in der letzten Spalte, also nach Satz 2.30 ist $\mathcal{L}(A|\mathbf{b}) \neq \emptyset$. Nach demselben Satz, bekommen wir die Beschreibung der Lösungsmenge als

$$\mathcal{L}(A|\mathbf{b}) = \{(\frac{3}{2} - \frac{3}{2}t\ ,\ \ \frac{1}{2} + \frac{1}{2}t\ ,\ \ t)\ \ :\ \ t \in \mathbb{R}\} = \{(\frac{3-3t}{2}\ ,\ \ \frac{1+t}{2}\ ,\ \ t)\ \ :\ \ t \in \mathbb{R}\}$$

Probe:

$$\frac{3-3t}{2} + \frac{1+t}{2} + t = 2, \qquad \frac{3-3t}{2} - \frac{1+t}{2} + 2t = 1, \qquad \frac{3-3t}{2} - 3\frac{1+t}{2} + 3t = 0 \qquad \forall\ t \in \mathbb{R}.$$

[18] 13.12.'23

Kapitel 3

$\mathbb{K}$-Vektorräume

L'algèbre n'est qu'une géométrie écrite, la géométrie n'est qu'une algèbre figurée.
Sophie Germain

Chronologie

∼1630 René Descartes, Pierre de Fermat: analytische Geometrie

1804 Bernard Bolzano[1]: Operationen mit Punkten

1827 August Ferdinand Möbius: Barizentrische Koordinaten

1828 C.V. Mourney[2]: Existenz einer "größeren" Algebra + geometrische Darstellung komplexer Zahlen

∼1840 Giusto Bellavitis: Bipunkte (Vektoren); Jean-Robert Argand: Quaternionen; William Rowan Hamilton: Biquaternionen

∼1850 Hermann Günther Grassmann: Abstrakte Opeartionen, lineare Unabhängigkeit, Dimension, Skalarpordukt

1857 Arthur Cayley: Matrizen

1867 Laguerre: lineare Kombinationen in $\mathbb{R}^2$, $\mathbb{R}^4$, $\mathbb{R}^8$.

1888 Giuseppe Peano: moderne Definition von Vektorraum und lineare Abbildung

∼1900 Henri Lebesque: Funktionräume

∼1920 Stefan Banach: Banach Räume; David Hilbert: Hilbert Räume

[1] Bernardus Placidus Johann Nepomuk Bolzano
[2] Hat unter diesen Namen in Paris veröffentlicht, die wahre Identität ist aber unbekannt

3.1 Definition und Beispiele

In diesem Kapitel wird $\mathbb{K}$ ein beliebiger Körper bezeichnen. Die Elemente dieses Körpers $\mathbb{K}$ werden im Kontext der Vektorräume als **Skalare** bezeichnet. Dies soll betonen, dass die Konzepte und Aussagen, mit denen wir arbeiten, nicht nur für reelle und komplexe Vektorräume sinnvoll sind. Die gesamte Theorie basiert lediglich auf den algebraischen Eigenschaften der reellen und komplexen Zahlen, nämlich darauf, dass $(\mathbb{K}, +)$ und $(\mathbb{K}^\times, \cdot)$ abelsche Gruppen sind, sowie auf der Distributivität der Multiplikation bezüglich der Addition. Tatsächlich könnte man sogar die Kommutativität der Multiplikation in $\mathbb{K}$ vernachlässigen. Um die Sprache einfach zu halten (d.h. um links- und rechts-Vektorräume zu vermeiden), setzen wir die Kommutativität von $(\mathbb{K}^\times, \cdot)$ immer voraus.

Definition 3.1. Ein **$\mathbb{K}$-Vektorraum** ist ein Tripel $(V, +, \cdot)$ wobei:

V ist eine Menge,

$+ : V \times V \longrightarrow V$ ist eine innere Verknüpfung die **Vektoraddition** genannt wird,

$\cdot : \mathbb{K} \times V \longrightarrow V$ ist eine äußere Verknüpfung die **skalare Multiplikation** genannt wird,

sodass folgende Axiome gelten

(VR 1) $(V, +)$ ist eine abelsche Gruppe

(VR 2) Die skalare Multiplikation erfüllt

- (VR 2.1) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$ für alle $\lambda, \mu \in \mathbb{K}$ und $v \in V$.
- (VR 2.2) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ für alle $\lambda \in \mathbb{K}$ und $v, w \in V$.
- (VR 2.3) $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$ für alle $\lambda, \mu \in \mathbb{K}$ und $v \in V$.
- (VR 2.4) $1 \cdot v = v$ für alle $v \in V$, wobei $1 \in \mathbb{K}$.

Die Elemente der unterliegenden Menge V des $\mathbb{K}$-Vektorraumes werden als **Vektoren** bezeichnet.

Aus praktischen Gründen identifizieren wir einen $\mathbb{K}-$Vektorraum, also ein Tripel $(V, +, \cdot)$ mit der unterliegenden Menge der Vektoren. Daher sagen wir einfach: "V ist ein $\mathbb{K}$-Vektorraum" und nennen in den meisten Fällen nicht explizit die Addition und die skalare Multiplikation. In der obigen Definition haben wir die "Tripel"-Formulierung gewählt, um die Wichtigkeit der Struktur zu betonen. Insbesondere kann man auf derselben Menge verschiedene Vektorraumstrukturen definieren.

Jeder Vektorraum enthält mindestens ein Element: das neutrale Element der Gruppe $(V, +)$. Wir werden dieses Element mit **0** oder 0_V bezeichnen. Um den Nullvektor von dem Skalar Null zu unterscheiden, schreiben wir, wenn nötig, $0_\mathbb{K}$ für das neutrale Element der Addition in $\mathbb{K}$.
Für einen Vektor $v \in V$ bezeichnen wir mit $-v$ das inverse Element von v bezüglich der Vektoraddition.

Lemma 3.2. *Für jeden $\mathbb{K}$-Vektorraum V gelten*

(i) $\lambda \cdot 0_V = 0_V$ *für alle* $\lambda \in \mathbb{K}$.

(ii) $0_\mathbb{K} \cdot v = 0_V$ *für alle* $v \in V$.

(iii) $(-\lambda) \cdot v = \lambda \cdot (-v) = -(\lambda \cdot v)$ *für alle* $\lambda \in \mathbb{K}$ *und* $v \in V$.

(iv) *Aus* $\lambda \cdot v = 0_V$ *für* $\lambda \in \mathbb{K}$ *und* $v \in V$ *folgt* $\lambda = 0_\mathbb{K}$ *oder* $v = 0_V$.

Beweis-Skizze:

(i) Wir haben $\lambda \cdot O_V = \lambda \cdot (O_V + O_V) = \lambda \cdot O_V + \lambda \cdot O_V$ und die Aussage folgt aus der Kürzungsregel in der abelschen Gruppe $(V, +)$.

(ii) Es gilt $0_{\mathbb{K}} \cdot v = (0_{\mathbb{K}} + 0_{\mathbb{K}}) \cdot v = 0_{\mathbb{K}} \cdot v + 0_{\mathbb{K}} \cdot v$ und die Aussage folgt aus der Kürzungsregel in der abelschen Gruppe $(V, +)$.

(iii) Es gilt $\lambda \cdot v + (-\lambda) \cdot v = (\lambda - \lambda) \cdot v = 0_V$. Analog für $\lambda \cdot (-v)$.

(iv) Sei $\lambda \cdot v = 0_V$. Wenn $\lambda = 0_{\mathbb{K}}$ sind wir fertig. Wenn $\lambda \neq 0_{\mathbb{K}}$, dann existiert $\lambda^{-1} \in \mathbb{K}$. Also

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

Q.E.D.

Bemerkung 3.3. Die Distributivitätsgesetze (VR.2.1) und (VR.2.2) gelten auch für endliche Familien $v, v_i, w_i \in V$, $\lambda, \lambda_i, \mu_i \in \mathbb{K}$ für $i = 1, \ldots, n$:

$$\begin{aligned} \lambda \cdot \textstyle\sum_{i=1}^n v_i &= \textstyle\sum_{i=1}^n \lambda \cdot v_i, \\ (\textstyle\sum_{i=1}^n \lambda_i) \cdot v &= \textstyle\sum_{i=1}^n \lambda_i \cdot v, \\ \textstyle\sum_{i=1}^n \lambda_i v_i + \textstyle\sum_{i=1}^n \mu_i v_i &= \textstyle\sum_{i=1}^n (\lambda_i + \mu_i) \cdot v_i. \end{aligned}$$

Beispiele:

1. Der kleinste $\mathbb{K}$-Vektorraum ist $V = \{0\}$, mit $0 + 0 = 0$ und $\lambda \cdot 0 = 0$ für alle $\lambda \in \mathbb{K}$. Dieser Vektorraum heißt der **Nullraum** und wir schrieben einfach $V = 0$ dafür.

2. Für jedes $n \in \mathbb{N}$ ist der n-Dimensionale **$\mathbb{K}$-Standardraum** $\mathbb{K}^n$ mit der komponentenweise Addition $+\mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}^n$, gegeben für alle $\mathbf{v}, \mathbf{w} \in \mathbb{K}^n$ durch

   $$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) = (v_1 + w_1, \ldots, v_n + w_n),$$

   und der komponentenweise skalaren Multiplikation $\cdot : \mathbb{K} \times \mathbb{K}^n \longrightarrow \mathbb{K}^n$, gegeben durch

   $$\lambda \cdot (v_1, \ldots, v_n) = (\lambda v_1, \ldots, \lambda v_n) \qquad \forall\, \mathbf{v} \in V \text{ und } \lambda \in \mathbb{K}.$$

   Wir haben schon gesehen, dass $(\mathbb{K}^n, +)$ eine abelsche Gruppe ist. Die andere Axiome sind sehr einfach zu überprüfen.

3. (a) Der Polynomring $\mathbb{K}[x]$ in einer Variable x mit Koeffizienten in $\mathbb{K}$. Die Addition ist die Addition der Polynome und die skalare Multiplikation ist die Multiplikation mit Polynomen von Grad Null.

   (b) $\mathbb{K}[x]_{\leqslant d} = \{f \in \mathbb{K}[x] \;:\; \deg f \leqslant d\}$

   (c) $\mathbb{K}[x_1, \ldots, x_n] = \{\sum_{\boldsymbol{\alpha} \in A} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \;:\; A \subset \mathbb{N}^n, \text{ mit } A \text{ endlich, und } c_{\boldsymbol{\alpha}} \in \mathbb{K}\}$, wobei $\mathbf{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdot \cdots \cdot x_n^{\alpha_n}$.

(d) $\mathbb{K}[\mathbf{x}]_{\leqslant d} = \{f \in \mathbb{K}[\mathbf{x}] \ : \ \deg f \leqslant d\}$, wobei

$$\begin{aligned} \deg \mathbf{x^d} &= \deg(x_1^{d_1} x_2^{d_2} \dots x_n^{d_n}) = d_1 + \dots + d_n \\ \deg f &= \deg \textstyle\sum_{c_\mathbf{d} \in \mathbb{K}} c_\mathbf{d} \mathbf{x^d} = \max\{\deg \mathbf{x^d} \ : \ c_\mathbf{d} \neq 0\}. \end{aligned}$$

(e) $\mathbb{K}[\mathbf{x}]_d = \{f \in \mathbb{K}[\mathbf{x}] \ : \ \deg f = d\}$:

$$\mathbb{K}[x,y,z]_3 = \{a_1x^3+a_2x^2y+a_3x^2z+a_4xy^2+a_5xyz+a_6xz^2+a_7y^3+a_8y^2z+a_9yz^2+a_{10}z^3 \ : \ a_i \in \mathbb{K}\}.$$

In diesem Beispiel sieht man warum es wichtig war das Nullpolynom vom beliebigen Grad zu betrachten.

4. $\mathrm{Mat}_{m\times n}(\mathbb{K})$ ist ein $\mathbb{K}$-VR.

5. Wenn $\mathbb{K}$ ein Teilkörper des Körpers $\mathbb{L}$ ist, dann ist $\mathbb{L}$ ein $\mathbb{K}$-Vektorraum. Insbesondere ist jeder Körper $\mathbb{K}$ ein $\mathbb{K}$-Vektorraum. Auch $\mathbb{Q}[\sqrt{2}]$, $\mathbb{R}$, $\mathbb{C}$ sind $\mathbb{Q}$-Vektorräume. Der Körper mit vier Elementen $\mathbb{F}_4$ ist ein $\mathbb{Z}/2\mathbb{Z}$-VR und $\mathbb{C}$ ist ein $\mathbb{R}$-VR.

6. $\mathbb{Q}$ ist aber kein $\mathbb{R}$-VR

7. Die abelsche Gruppe $\mathrm{GL}_n(\mathbb{K})$ mit der üblichen Gruppenstruktur und Skalarmultiplikation der Matrizen ist kein $\mathbb{K}$-Vektorraum.

[19] 18.12.'23

8. Die Gruppe $(\mathbb{R}^n, +)$ ist kein $\mathbb{C}$-VR mit der Skalarmultiplikation $\lambda \cdot v := Re(\lambda) \cdot v$, weil $(i^2)v = -v$ aber $i(iv) = 0$.

9. Die Gruppe $(\mathbb{R}^n, +)$ ist kein $\mathbb{C}$-VR mit $\lambda \cdot v := |\lambda| \cdot v$, weil VR2.3 gilt nicht.

10. Die Gruppe $(\mathbb{Z}, +)$ ist kein $\mathbb{K}$-VR, unabhängig von $\mathbb{K}$:
Wenn $\mathrm{char}\ \mathbb{K} = p > 0$, das impliziert $\sum_{i=1}^{p} 1_\mathbb{K} = 0$, dann haben wir in $\mathbb{Z}$:

$$0_\mathbb{Z} \neq p = 1_\mathbb{Z} + \dots + 1_\mathbb{Z} = 1_\mathbb{K} \cdot 1_\mathbb{Z} + \dots + 1_\mathbb{K} \cdot 1_\mathbb{Z} = (1_\mathbb{K} + \dots + 1_\mathbb{K}) \cdot 1_\mathbb{Z} = 0_\mathbb{K} \cdot 1_\mathbb{Z} = 0_\mathbb{Z}.$$

Wenn $\mathrm{char}\ \mathbb{K} = 0$, dann haben wir einen injektiven Körperhomomorphismus $\mathbb{Q} \longrightarrow \mathbb{K}$. Das bedeutet wir dürfen annehmen, dass der Körper $\mathbb{K}$ den Körper $\mathbb{Q}$ enthält. Wenn es eine Multiplikation mit Skalare gäbe, dann würde eine ganze Zahl $z \in \mathbb{Z}$ existieren, sodass

$$\tfrac{1}{2} \cdot 1_\mathbb{Z} = z.$$

Daraus würde folgen, dass $1_\mathbb{Z} = 1_\mathbb{K} \cdot 1_\mathbb{Z} = (\frac{1}{2} + \frac{1}{2}) \cdot 1_\mathbb{Z} = z + z$. Es gibt aber keine ganze Zahl z mit der Eigenschaft $z + z = 1$.

11. Sei X eine beliebige Menge. Eine **$\mathbb{K}$-wertige Funktion** auf X ist einfach Abbildung $f : X \longrightarrow \mathbb{K}$, ohne weitere Vorgaben. Wir bezeichnen die Menge aller $\mathbb{K}$-wertigen Funktionen als $\mathrm{Abb}(X, \mathbb{K})$. Auf dieser Menge kann man mit Hilfe der zwei Körper-Operationen aus $(\mathbb{K}, +, \cdot)$ eine $\mathbb{K}$-VR Struktur definieren durch:

$$\begin{aligned} f + g : X &\longrightarrow \mathbb{K}, & x &\mapsto f(x) + g(x) \\ \lambda \cdot f : X &\longrightarrow \mathbb{K}, & x &\mapsto \lambda \cdot f(x) \end{aligned}$$

Man kann direkt überprüfen, dass die Assoziativität, Kommutativität, und Distributivität für die roten $+$ und $\cdot$ aus den entsprechenden Eigenschaften der blauen Verknüpfungen folgen. Weiterhin haben wir $\mathbf{0} : X \longrightarrow \mathbb{K}$ mit $x \mapsto 0$, $\forall x \in X$ als Neutrales Element für $+$, und $-f : X \longrightarrow \mathbb{K}$ mit $x \mapsto -f(x)$ als inverses *Element* bezüglich $+$ (nicht als inverse Abbildung bezüglich der Verknüpfung! Die Verknüpfung würde so wie so nur für $X = \mathbb{K}$ Sinn haben). Das heißt, dass $(\mathrm{Abb}(X, \mathbb{K}), +, \cdot)$ ein $\mathbb{K}$-Vektorraum ist.

12. Wenn V_1, V_2 zwei $\mathbb{K}$-VR sind, dann ist auch $V_1 \times V_2$ ein $\mathbb{K}$-VR mit den Komponenten-weise Operationen:

$$\begin{aligned} (v_1, v_2) + (w_1, w_2) &= (v_1 + v_2, w_1 + w_2) \\ \lambda \cdot (v_1, v_2) &= (\lambda \cdot v_1, \lambda \cdot v_2) \\ 0_{V_1 \times V_2} &= (0_{V_1}, 0_{V_2}) \\ -(v_1, v_2) &= (-v_1, -v_2) \end{aligned}$$

für alle $v_1, w_1 \in V_1$, $v_2, w_2 \in V_2$ und $\lambda \in \mathbb{K}$.

13. Wenn $(V_i)_{i \in I}$ eine Familie von $\mathbb{K}$-VR, ist $V = \prod_{i \in I} V_i$ ein $\mathbb{K}$-VR mit

$$(v_i)_{i \in I} + (w_i)_{i \in I} = (v_i + w_i)_{i \in I}, \tag{3.1}$$
$$\lambda \cdot (v_i)_{i \in I} = (\lambda \cdot v_i)_{i \in I}. \tag{3.2}$$

Dieser Vektorraum heißt das **direkte Produkt** der Familie von $\mathbb{K}$-Vektorräumen $(V_i)_{i \in I}$.

3.2 $\mathbb{K}$-lineare Abbildungen

Definition 3.4. Seien V, W zwei $\mathbb{K}$-Vektorräume. Eine **$\mathbb{K}$-lineare Abbildung** (oder ein **Homomorphismus von $\mathbb{K}$-Vektorräume**) von V nach W ist eine Abbildung $f : V \longrightarrow W$ mit den Eigenschaften

(LA 1) $f(v_1 + v_2) = f(v_1) + f(v_2)$ für alle $v_1, v_2 \in V$.

(LA 2) $f(\lambda v) = \lambda f(v)$ für alle $\lambda \in \mathbb{K}$ und $v \in V$.

Die Menge aller $\mathbb{K}$-linearen Abbildungen von V nach W bezeichnen wir als $\mathrm{Hom}_{\mathbb{K}}(V, W)$.

Definition 3.5. Seien V und W zwei $\mathbb{K}$-Vektorräume. Eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ ist ein **$\mathbb{K}$-linearer Isomorphismus**, wenn es eine $\mathbb{K}$-lineare Abbildung $f' : W \longrightarrow V$ gibt, sodass

$$f \circ f' = \mathrm{id}_W \quad \text{und} \quad f' \circ f = \mathrm{id}_V .$$

Bemerkung 3.6. Eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ ist ein Isomorphismus, wenn und nur wenn es bijektiv ist.

Beweis. $\boxed{\Rightarrow}$ Ist klar: f' ist die Inverse Abbildung.
$\boxed{\Leftarrow}$ Wenn f bijektiv ist, dann existiert die[3] Inverse $f^{-1} : W \longrightarrow V$. Wir müssen aber noch zeigen, dass die Inverse auch $\mathbb{K}$-linear ist. Weil f bijektiv ist, existiert für jedes $w \in W$ ein eindeutiges $v \in V$, sodass

[3] wenn es eine Inverse Abbildung gibt, dann ist diese eindeutig.

$f(v) = w$, und somit auch $f^{-1}(w) = v$. Seien $w_1, w_2 \in W$ beliebig, und $v_1, v_2 \in V$ die eindeutigen entsprechenden Vektoren mit $f(v_i) = w_i$. Dann gilt:

$$f^{-1}(w_1 + w_2) = f^{-1}(f(v_1) + f(v_2)) = f^{-1}(f(v_1 + v_2)) = v_1 + v_2 = f^{-1}(w_1) + f^{-1}(w_2).$$

Analog, für einen beliebigen $w \in W$, mit $f^{-1}(w) = v$, und $\lambda \in \mathbb{K}$ haben wir

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

□

Eine $\mathbb{K}$-lineare Abbildung von V nach V heißt **$\mathbb{K}$-linearer Endomorphismus** von V. Ein Endomorphismus der auch ein Isomorphismus ist, heißt **$\mathbb{K}$-linearer Automorphismus** von V. Wir bezeichnen die Menge aller Endomorphismen von V mit $\mathrm{End}_{\mathbb{K}}(V)$ und die Menge aller Automorphismen von V mit $\mathrm{Aut}_{\mathbb{K}}(V)$.

Bemerkung 3.7. Das Axiom (LA1) ist genau das Axiom das Gruppenhomomorphismen erfüllen müssen. Also, eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ ist insbesondere ein Gruppenhomomorphismus $f : (V, +) \longrightarrow (W, +)$. Das heißt, es gelten

(i) $f(0_V) = 0_W$.

(ii) $f(-v) = -f(v)$.

(iii) f ist injektiv genau dann, wenn $\mathrm{Ker}(f) := \{v \in V \ : \ f(v) = 0_W\} = \{0_V\}$.

Beispiele:

1. Die Identität $\mathrm{id}_V : V \longrightarrow V$ ist eine $\mathbb{K}$-lineare Abbildung.
2. Die Inklusion $i_W : W \hookrightarrow V$ eines $\mathbb{K}$-UVR $W \subseteq V$ ist eine $\mathbb{K}$-lineare Abbildung.
3. Die Nullabbildung $0 : V \longrightarrow W$, für die $0(v) := 0_W$ für alle $v \in V$.
4. Wenn $V \xrightarrow{f} W \xrightarrow{g} U$ zwei $\mathbb{K}$-lineare Abbildungen sind, dann ist auch $g \circ f : V \longrightarrow U$ eine $\mathbb{K}$-lineare Abbildung. Das ist sehr einfach zu sehen: Für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ gilt:

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

5. Für jeden $\mathbb{K}$-Vektorraum V und jeden nicht-nullen Skalar $\lambda \in \mathbb{K}^\times$ ist die **Homothetie** $h_\lambda : V \longrightarrow V$, gegeben durch
$$h_\lambda(v) := \lambda v.$$
Wir bezeichnen diese Abbildung auch durch $\cdot\lambda$.
6. Die Drehung der reellen Ebene $\mathbb{R}^2$ um den Ursprung $(0, 0)$ mit Winkel θ ist eine lineare Abbildung $D_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Es gilt
$$D_\theta(\mathbf{0}) = \mathbf{0}, \quad D_\theta(1, 0) = (\cos\theta, \sin\theta), \quad D_\theta(0, 1) = (-\sin\theta, \cos\theta).$$
Für $v = (x, y) \in \mathbb{R}^2$ gilt $v = xe_1 + ye_2$, also
$$D_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

7. Sei $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$. Die Abbildung $f_A : \mathbb{K}^n \xrightarrow{A} \mathbb{K}^m$ gegeben durch

$$(x_1, \ldots, x_n)^T \longmapsto A \cdot (x_1, \ldots, x_n)^T$$

ist eine $\mathbb{K}$-lineare Abbildung. Es ist sehr einfach das zu überprüfen, dass f_A eine $\mathbb{K}$-lineare Abbildung ist:

$$\begin{aligned}
f_A(\mathbf{x}+\mathbf{y}) &= \begin{pmatrix} a_{11}(x_1+y_1)+\cdots+a_{1n}(x_n+y_n) \\ \vdots \\ a_{m1}(x_1+y_1)+\cdots+a_{mn}(x_n+y_n) \end{pmatrix}^T \\
&= \begin{pmatrix} a_{11}x_1+\cdots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n \end{pmatrix}^T + \begin{pmatrix} a_{11}y_1+\cdots+a_{1n}y_n \\ \vdots \\ a_{m1}y_1+\cdots+a_{mn}y_n \end{pmatrix}^T \\
&= f_A(\mathbf{x}) + f_A(\mathbf{y}). \\
f_A(\lambda\cdot\mathbf{x}) &= \begin{pmatrix} a_{11}(\lambda x_1)+\cdots+a_{1n}(\lambda x_n) \\ \vdots \\ a_{m1}(\lambda x_1)+\cdots+a_{mn}(\lambda x_n) \end{pmatrix}^T \\
&= \begin{pmatrix} \lambda(a_{11}x_1+\cdots+a_{1n}x_n) \\ \vdots \\ \lambda(a_{m1}x_1+\cdots+a_{mn}x_n) \end{pmatrix}^T \\
&= \lambda\cdot f_A(\mathbf{x}).
\end{aligned}$$

[20] 20.12.'23

3.3 $\mathbb{K}$-Untervektorräume und Erzeugendensysteme

Die Untervektorräume von V sind genau die Teilmengen von V die mit den Einschränkungen der Vektoraddition und der Multiplikation mit Skalaren, selber $\mathbb{K}$-Vektorräume sind. Das bedeutet insbesondere, dass die Einschränkungen wohl definierte Abbildungen sind, also dass die Teilmengen abgeschlossen unter der Addition und der skalaren Multiplikation sind. Man kann das genauer formulieren.

Definition 3.8. Sei V ein $\mathbb{K}$-Vektorraum. Ein **$\mathbb{K}$-Untervektorraum**[4] ist eine Teilmenge $U \subseteq V$ mit den Eigenschaften:

(UVR 1) $U \neq \emptyset$,

(UVR 2) Wenn $v, w \in U$, dann $v + w \in U$,

(UVR 3) Wenn $\lambda \in \mathbb{K}$ und $v \in U$, dann $\lambda \cdot v \in U$.

Bemerkung 3.9. Eine Teilmenge $U \subseteq V$ ist genau dann ein $\mathbb{K}$-Untervektorraum, wenn $U \neq \emptyset$ und

$$\lambda_1 u_1 + \lambda_2 u_2 \in U \quad \forall\, u_1, u_2 \in U \quad \text{und } \forall\, \lambda_1, \lambda_2 \in \mathbb{K}.$$

[4] auch $\mathbb{K}$-linearer Unterraum genannt; Abkürzungen $\mathbb{K}$-UVR oder UVR.

Beispiele:

1. Die Teilmengen $\{0\}$ und V sind $\mathbb{K}$-UVR von V für alle V. Der erste heißt der **Nullraum** und wird auch **trivialer Unterraum** genannt. Ein Unterraum $U \subseteq_{\mathbb{K}} V$ mit $U \neq V$ heißt **echter Unterraum**. Man kann also sagen, dass $V \subseteq_{\mathbb{K}} V$ der "unechte Vektorraum" ist.

2. Wenn $m < n$, dann kann man $\mathbb{K}^m$ als $\mathbb{K}$-UVR von $\mathbb{K}^n$ auffassen:

$$(v_1, \ldots, v_m) \mapsto (v_1, \ldots, v_m, 0, \ldots, 0).$$

   Es ist aber nicht die einzige Möglichkeit:

   2..1. Achsen und Diagonalen in $\mathbb{R}^2$ und $\mathbb{R}^3$,

   2..2. Koordinaten Ebenen in $\mathbb{R}^3$

   2..3. $\forall\, v \in \mathbb{R}^2$ haben wir $\mathbb{R} \cdot v := \{\lambda \cdot v \;:\; \lambda \in \mathbb{R}\} \subset \mathbb{R}^2$.

   2..4. $\mathbb{R}^2 \simeq \{(t_1, t_2, t_1 - t_2)\} \subset \mathbb{R}^3$.

   2..5. $\{(x, y) \in \mathbb{R}^2 \;:\; ax + by = c\} \subset \mathbb{R}^2$ ist ein $\mathbb{K}$-UVR $\iff c = 0$.

3. $\{f : \mathbb{R} \longrightarrow \mathbb{R} \;:\; f \text{ ist stetig}\} \subset \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ als $\mathbb{R}$-VR. Auch differenzierbare/polynomielle Funktionen sind ein $\mathbb{K}$-UVR.

4. Symmetrische Matrizen in Matrizen

5. $\mathbb{R} \subset \mathbb{C}$, $\mathbb{Q} \subset \mathbb{C}$.

6. Die Lösungsmenge eines homogenes LGS mit n-Unbekannten ist ein $\mathbb{K}$-UVR von $\mathbb{K}^n$.

Lemma 3.10. *Sei $(U_i)_{i \in I}$ eine nicht-leere*[5] *Familie von $\mathbb{K}$-Untervektorräumen von V. Die Schnittmenge $U = \bigcap_{i \in I} U_i$ ist ein $\mathbb{K}$-Untervektorraum von V.*

Beweis-Skizze: (UVR 1) Wir haben für alle $i \in I$, dass $\mathbf{0} \in U_i$. Also, weil $I \neq \emptyset$, $\mathbf{0} \in U$ und somit $U \neq \emptyset$.

(UVR 2) Wenn $v, w \in U$, dann $v, w \in U_i \;\; \forall\, i \in I$, dann $v + w \in U_i \;\; \forall\, i \in I$, also $v + w \in U$.

(UVR 3) Wenn $\lambda \in \mathbb{K}$ und $v \in U$, dann $v \in U_i \;\; \forall\, i \in I$, dann $\lambda \cdot v \in U_i \;\; \forall\, i \in I$, also $\lambda \cdot v \in U$.

Q.E.D.

Definition 3.11. Sei $S \subseteq V$ eine Teilmenge des $\mathbb{K}$-Vektorraumes V. Der von S in V **erzeugten linearen Unterraum** (oder die **lineare Hülle** von S in V) ist der $\mathbb{K}$-Untervektorraum:

$$\mathrm{Span}_{\mathbb{K}} S := \bigcap_{S \subseteq U \subseteq_{\mathbb{K}} V} U$$

Weil $S \subseteq V$ ist die Familie $(U \;:\; S \subseteq U \subseteq V, U = \mathbb{K} - \mathrm{UVR})$ nicht leer, also nach Lemma 3.10 ist $\mathrm{Span}_{\mathbb{K}} S$ tatsächlich ein $\mathbb{K}$-UVR. Wir sagen, dass S den $\mathbb{K}$-UVR $\mathrm{Span}_{\mathbb{K}} S$ **erzeugt**, oder dass S ein **Erzeugendensystem** von $\mathrm{Span}_{\mathbb{K}} S$ ist.

Manche Autoren bezeichnen den von S erzeugten linearen Unterraum als $\mathrm{Span}\, S$.

[5] Das heißt, dass $I \neq \emptyset$. Manchmal, wenn die Indexmenge I nicht explizit angegeben ist, ist das nicht offensichtlich und kann einen Beweis brauchen.

Bemerkung 3.12. Für jeden $\mathbb{K}$-UVR $U \subseteq V$ gilt

$$\text{wenn } S \subseteq U, \text{ dann } \operatorname{Span}_{\mathbb{K}} S \subseteq U.$$

Anders gesagt, ist $\operatorname{Span}_{\mathbb{K}} S$ der kleinste $\mathbb{K}$-UVR der S enthält. Genauer formuliert man das als Minimum bezüglich der Inklusion[6]:

$$\operatorname{Span}_{\mathbb{K}} S = \min\{ U \quad : \quad S \subseteq U \subseteq_{\mathbb{K}} V\}.$$

Also

$$W = \operatorname{Span}_{\mathbb{K}} S \iff \begin{cases} W \subseteq_{\mathbb{K}} V & \text{und} \\ S \subseteq W & \text{und} \\ (U \subseteq_{\mathbb{K}} V \text{ und } S \subseteq U) \Rightarrow W \subseteq U. \end{cases} \tag{3.3}$$

Bemerkung 3.13. Wenn U ein $\mathbb{K}$-UVR ist und $S, T \subset V$ beliebige Teilmengen, dann gilt

(i) $U = \operatorname{Span}_{\mathbb{K}} U$.

(ii) $S \subseteq T \Rightarrow \operatorname{Span}_{\mathbb{K}} S \subseteq \operatorname{Span}_{\mathbb{K}} T$.

(iii) $\operatorname{Span}_{\mathbb{K}} \emptyset = \operatorname{Span}_{\mathbb{K}}\{0_V\} = 0$.

Definition 3.14. Seien U_1, U_2 zwei lineare Unterräume des Vektorraumes V. Die **Summe** von U_1 und U_2 ist der $\mathbb{K}$-Untervektorraum

$$U_1 + U_2 := \operatorname{Span}_{\mathbb{K}} U_1 \cup U_2.$$

Also $U_1 + U_2$ ist der kleinste UVR der sowohl U_1 als auch U_2 enthält.

Satz 3.15. *Wenn U_1, U_2 zwei $\mathbb{K}$-Untervektorräume des Vektorraumes V sind, dann gilt*

$$U_1 + U_2 = \{v_1 + v_2 \quad : \quad v_1 \in U_1, v_2 \in U_2\}. \tag{3.4}$$

Beweis-Skizze: Sei W die Menge auf der rechten Seite der Gleichung (3.4). Wir verwenden die Äquivalenz (3.3) aus Bemerkung 3.12. Wir müssen also zu erst zeigen, dass W ein $\mathbb{K}$-UVR ist:
(UVR 1) $0 = 0 + 0 \in W$, also $W \neq \emptyset$;
(UVR 2) $(v_1 + v_2) + (w_1 + w_2) = (v_1 + w_1) + (v_2 + w_2) \in W$;
(UVR 3) $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 \in W$.

Wir müssen dann überprüfen, dass $U_1 \cup U_2 \subseteq W$: wenn $u \in U_1$, dann gilt $u = u + 0_V \in W$; wenn $u \in U_2$, dann $u = 0_V + u \in W$.

Für die dritte Bedingung aus (3.3) sei $U \subseteq_{\mathbb{K}} V$ mit $U_1 \cup U_2 \subseteq U$ und sei $v \in W$ beliebig. Dann existieren $v_1 \in U_1$ und $v_2 \in U_2$ mit $v = v_1 + v_2$. Es gilt auch $v_1, v_2 \in U_1 \cup U_2 \subseteq U$, also aus (UVR 2) für U folgt $v = v_1 + v_2 \in U$. Q.E.D.

Definition 3.16. Sei $S \subseteq V$ eine Menge von Vektoren. Eine **Linearkombination** von Elementen aus S ist ein Vektor v für den es $n \in \mathbb{N}$, $v_1, \ldots, v_n \in S$ und Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ existieren, sodass

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Eine Linearkombination $\lambda_1 v_1 + \cdots + \lambda_n v_n$ ist **nicht-trivial** wenn $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$.

[6] Die Inklusion ist eine partielle Ordnungsrelation.

Diese Definition funktioniert auch für $S = \emptyset$, weil die leere Summe ist auch definiert: es ist Null. Das passt gut zusammen mit $\mathrm{Span}_{\mathbb{K}} \emptyset = 0$, wie wir gleich in Satz 3.18 sehen werden.

Bemerkung 3.17. Die Anzahl n, die Vektoren $v_1, \ldots, v_n \in S$ und die Skalare $\lambda_1, \ldots, \lambda_n$ sind nicht unbedingt eindeutig. Zum Beispiel für $S = \{(0,1), (1,1), (1,0)\} \subset \mathbb{R}^2$ und $v = (2,2)$ haben wir

$$(2,2) = 2 \cdot (1,1) = 2 \cdot (1,0) + 2 \cdot (0,1) = (1,0) + (1,1) + (0,1) = \frac{3}{2}(1,0) + \frac{1}{2}(1,1) + \frac{3}{2}(0,1).$$

Satz 3.18. *Sei $S \subseteq V$ eine Teilmenge von V. Es gilt*

$$\mathrm{Span}_{\mathbb{K}} S = \{v \in V \quad : \quad v \text{ ist eine Linearkombination von Vektoren in } S\}.$$

Beweis-Skizze: Sei L die Menge aller Linearkombinationen von Vektoren in S. Wir zeigen, dass L ein $\mathbb{K}$-UVR ist:

$$\begin{aligned} 0_V &\in L, \\ \textstyle\sum_{i=1}^m \lambda_i v_i + \sum_j^m \mu_j w_j &= \lambda_1 v_1 + \cdots + \lambda_n v_n + \mu_1 w_1 + \cdots + \mu_m w_m, \\ \lambda \cdot (\textstyle\sum_{i=1}^n \lambda_i v_i) &= \textstyle\sum_{i=1}^n ((\lambda \lambda_i) v_i). \end{aligned}$$

Der Rest ist Analog zum Beweis von Satz 3.15. Q.E.D.

Korollar 3.19. Für jede nicht-leere Teilmenge $S \subseteq V$ gilt

$$v \in \mathrm{Span}_{\mathbb{K}} S \iff \exists\, v_1, \ldots, v_n \in S \text{ und } \lambda_1, \ldots, \lambda_n \in \mathbb{K}, \text{ sodass } v = \textstyle\sum_{i=1}^n \lambda_i v_i.$$

Lemma 3.20. *Seien $T \subseteq S$ zwei Mengen von Vektoren. Es gilt*

$$\mathrm{Span}_{\mathbb{K}} T = \mathrm{Span}_{\mathbb{K}} S \iff v \in \mathrm{Span}_{\mathbb{K}} T \quad \forall\, v \in S.$$

Beweis-Skizze: $\boxed{\Rightarrow}$ Diese Implikation gilt, weil $v \in \mathrm{Span}_{\mathbb{K}} S \quad \forall\, v \in S$.

$\boxed{\Leftarrow}$ Die Inklusion $\mathrm{Span}_{\mathbb{K}} T \subseteq \mathrm{Span}_{\mathbb{K}} S$ gilt weil $T \subseteq S$. Wir brauchen also nur noch "$\supseteq$" zu zeigen. Sei $v \in \mathrm{Span}_{\mathbb{K}} S$. Aus Korollar 3.19 folgt $\exists\, v_1, \ldots, v_n \in S$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, sodass

$$v = \textstyle\sum_{i=1}^n \lambda_i v_i. \tag{3.5}$$

Aus der Voraussetzung gilt $v_i \in \mathrm{Span}_{\mathbb{K}} T \quad \forall\, i = 1, \ldots, n$. Wir wenden Korollar 3.19 für jeden Vektor v_i an und bekommen $\forall\, i = 1, \ldots, n, \quad \exists w_{i1}, \ldots, w_{in_i} \in T$ und $\mu_{i1}, \ldots, \mu_{in_i} \in \mathbb{K}$ sodass

$$v_i = \textstyle\sum_{j=1}^{n_i} \mu_{ij} w_{ij}. \tag{3.6}$$

Wenn wir (3.6) in (3.5) einsetzten, bekommen wir $v \in \mathrm{Span}_{\mathbb{K}} T$.

Q.E.D.

Definition 3.21. Sei V ein $\mathbb{K}$-Vektorraum. Ein **Erzeugendensystem** von V ist eine Teilmenge $S \subset V$ mit der Eigenschaft, dass

$$\mathrm{Span}_{\mathbb{K}} S = V.$$

Ein $\mathbb{K}$-VR heißt **endlich erzeugt** wenn es eine endliche Menge S gibt mit $\mathrm{Span}_{\mathbb{K}} S = V$.

Beispiele:

1. $e_i = (\delta_{1i}, \dots, \delta_{ni}) \in \mathbb{K}^n$ für $\mathbb{K}^n$.
2. Standardmatrizen für $\mathrm{Mat}_{m\times n}(\mathbb{K})$.
3. Monome für $\mathbb{K}[\mathbf{x}]$.
4. Sei $n \in \mathbb{N}_{>0}$ und für $i = 1, \dots, n$ seien $\mathbf{v}_i \in \mathbb{K}^m$. Sei $\mathbf{b} \in \mathbb{K}^m$ und sei $A = \left(\mathbf{v}_1^T | \quad \dots | \quad \mathbf{v}_n^T\right)$. Das LGS $A \cdot \mathbf{x} = \mathbf{b}^T$ ist genau dann lösbar, wenn $\mathbf{b} \in \mathrm{Span}_{\mathbb{K}} \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Insbesondere gilt

$$\mathrm{Span}_{\mathbb{K}} \{\mathbf{v}_1, \dots, \mathbf{v}_n\} = \{\mathbf{b} \in \mathbb{K}^m \quad : \quad \mathcal{L}(A|\mathbf{b}^T) \neq \emptyset\}.$$

Wir sagen, dass eine Menge $S \subseteq V$ ein **minimales Erzeugendensystem** von V ist wenn folgende zwei Bedingungen erfüllt sind:

(a) $\mathrm{Span}_{\mathbb{K}} S = V$

(b) $\mathrm{Span}_{\mathbb{K}} S \setminus v \subsetneq \mathrm{Span}_{\mathbb{K}} S, \quad \forall\, v \in S.$

Die Bedingung (b) ist äquivalent zu

(b') $\mathrm{Span}_{\mathbb{K}} T \neq \mathrm{Span}_{\mathbb{K}} S, \quad \forall\, T \subsetneq S.$

3.4 Lineare Unabhängigkeit

Definition 3.22. Sei $S = \{v_1, \dots, v_n\}$ eine endliche Menge von Vektoren eines $\mathbb{K}$-Vektorraumes V. Die Menge S heißt **linear unabhängig** (über $\mathbb{K}$) genau dann, wenn

$$\text{aus } \lambda_1 v_1 + \dots + \lambda_n v_n = 0, \text{ mit } \lambda_i \in \mathbb{K} \text{ notwendig } \lambda_1 = \dots = \lambda_n = 0 \text{ folgt.}$$

Eine unendliche Menge $S \subseteq V$ ist linear unabhängig, wenn jede endliche Teilmenge davon linear unabhängig ist. Eine Menge $S \subseteq V$ heißt **linear abhängig** (über $\mathbb{K}$) genau dann, wenn es nicht linear unabhängig ist.

Wenn $S = \{v_1, \dots, v_n\}$ eine linear unabhängige (bzw. abhängige) Menge ist, sagen wir auch einfach, dass die Vektoren linear unabhängig (bzw. abhängig) sind. Statt Menge, wird manchmal S auch *System von Vektoren* genannt. In diesem Fall sagen wir *linear unabhängiges System* von Vektoren, usw. Die leere Menge ist linear unabhängig.

[21] 8.1.'24

Bemerkung 3.23. Eine Menge $S = \{v_1, \dots, v_n\}$ ist also linear abhängig wenn es $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ gibt, sodass $\sum_{i=1}^n \lambda_i v_i = 0$.

Bemerkung 3.24. Wenn $S = \{v_1, \dots, v_n\}$ linear unabhängig ist, dann gelten

(i) $v_i \neq 0, \quad \forall\, i.$

(ii) $v_i \neq v_j, \quad \forall i \neq j.$

(iii) Jede Teilmenge $T \subseteq S$ ist auch linear unabhängig.

Satz 3.25. *Sei S eine Teilmenge eines $\mathbb{K}$-Vektorraumes V. Es gilt*

$$S \text{ ist linear unabhängig} \iff v \notin \mathrm{Span}_{\mathbb{K}} S \setminus v, \ \forall\, v \in S.$$

Beweis-Skizze: Wenn $S = \emptyset$, dann ist S linear unabhängig und die rechte Seite wahr. Wir nehmen also an, dass $S \neq \emptyset$.

$\boxed{\Rightarrow}$ Nehmen wir an, dass es ein $v \in S$ existiert, sodass $v \in \text{Span}_{\mathbb{K}} S \setminus v$. Aus Satz 3.18 folgt, dass es Vektoren $v_1, \ldots, v_n \in S \setminus v$ und Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gibt, sodass

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Weil $-1 \neq 0$ und $v_i \in S \setminus v$, bekommen wir eine nicht-triviale lineare Kombination von Vektoren aus S:

$$-1 \cdot v + \lambda_1 v_1 + \cdots + \lambda_n v_n = 0,$$

↯ ein Widerspruch ↯ zur linearen Unabhängigkeit von S.

$\boxed{\Leftarrow}$ Nehmen wir an, dass S linear abhängig ist. Dann existieren $v_1, \ldots, v_n \in S$ unterschiedliche Vektoren und $0 \neq (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$, sodass $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. O.B.d.A. dürfen wir annehmen, dass $\lambda_1 \neq 0$. Weil $\mathbb{K}$ ein Körper ist, existiert $\lambda_1^{-1} \in \mathbb{K}$. Wir bekommen dann

$$v_1 = (-\lambda_1^{-1}\lambda_2)v_2 + \cdots + (-\lambda_1^{-1}\lambda_n)v_n \in \text{Span}_{\mathbb{K}} \{v_2, \ldots, v_n\} \subseteq \text{Span}_{\mathbb{K}} S \setminus v_1$$

ein Widerspruch ↯. Q.E.D.

Beispiele:

1. Die Vektoren $(1,1), (1,0) \in \mathbb{R}^2$ sind linear unabhängig.
2. Die Vektoren $(1,2,3), (4,5,6), (7,8,9) \in \mathbb{R}^3$ sind linear abhängig weil
$$1 \cdot (1,2,3) - 2 \cdot (4,5,6) + 1 \cdot (7,8,9) = (0,0,0).$$
3. Die Vektoren $e_i = (\delta_{1i}, \ldots, \delta_{ni}) \in \mathbb{K}^n$ sind linear unabhängig.
4. Seien $v_1, \ldots, v_m \in \mathbb{K}^n$. Wir verstehen diese hier als Spaltenvektoren, also als $n \times 1$ Matrizen. Sei $A = \begin{pmatrix} v_1 & \ldots & v_m \end{pmatrix} \in \text{Mat}_{n\times m}(\mathbb{K})$, die Matrix mit den Vektoren $v_1, \ldots, v_m$ als Spaten. Die Vektoren $v_1, \ldots, v_m$ sind genau dann linear unabhängig, wenn $\mathcal{L}(A, 0) = \{\mathbf{0}\}$. Insbesondere, wenn $m = n$ sind $v_1, \ldots, v_n$ linear unabhängig wenn und nur wenn die Matrix A invertierbar ist.

Wir sagen, dass eine Menge $S \subseteq V$ eine **maximale linear unabhängige Menge** ist wenn folgende zwei Bedingungen erfüllt sind:

(c) S ist linear unabhängig.

(d) $S \cup \{v\}$ ist linear abhängig $\forall\ v \in V \setminus S$.

Die Bedingung (d) äquivalent zu

(d') T ist linear abhängig $\forall\ S \subsetneq T \subseteq V$.

3.5 Basen

Definition 3.26. Eine Menge $S \subseteq V$ ist eine **Basis**[7] von V wenn S folgende zwei Axiome erfüllt:

(B 1) S ist linear unabhängig.

(B 2) S ist ein Erzeugendensystem für V.

Beispiele:

1. Für jedes $i = 1, \dots, n$, sei $e_i = (\delta_{1i}, \dots, \delta_{in})$. Also
$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1).$$
Die Menge $\{e_1, \dots, e_n\}$ ist eine Basis und heißt die **Standardbasis** von $\mathbb{K}^n$.
2. Die Gruppe $(\mathbb{C}, +)$ der komplexen Zahlen mit der Addition hat sowohl eine $\mathbb{C}$-Vektorraum Struktur als auch eine $\mathbb{R}$-Vektorraum Struktur. Die Menge $\{1, i\}$ ist eine $\mathbb{R}$-Basis vom $\mathbb{C}$ das heißt eine Baisi des $\mathbb{R}$-Vektorraumes $\mathbb{C}$. Es ist aber keine $\mathbb{C}$-Basis von $\mathbb{C}$, weil es nicht linear unabhängig über $\mathbb{C}$ ist: $(i) \cdot 1 + (-1) \cdot i = 0$.
3. Die unendliche Menge $\{x^n\}_{n \in \mathbb{N}}$ ist eine Basis von $\mathbb{K}[x]$.
4. Der Nullraum $V = 0$ hat auch eine Basis: die leere Menge. Wir haben schon gesehen, dass $\text{Span}_{\mathbb{K}} \emptyset = 0$. Hier war die Definition von $\text{Span}_{\mathbb{K}}$ als Schnitt von $\mathbb{K}$-UVR praktischer als die Beschreibung aus Satz 3.18. Es bleibt also nur das Axiom (B 2) zu überprüfen. Es gibt aber keine $\mathbb{K}$-lineare Kombinationen mit Vektoren aus $\emptyset$, also da ist nichts zu überprüfen. Die Menge $\{0\}$ ist ein Erzeugendensystem des Nullraumes, es ist aber nicht linear unabhängig, weil $1 \cdot 0 = 0$.

Bemerkung 3.27. Wenn $\{v_1, \dots, v_n\}$ eine Basis von V ist, dann gilt
$$\forall\, v \in V, \exists!\, \lambda_1, \dots, \lambda_n \in \mathbb{K}, \text{ sodass } v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Beweis-Skizze: Existenz Folgt weil $\text{Span}_{\mathbb{K}} \{v_1, \dots, v_n\} = V$.
Eindeutigkeit Wenn $v = \lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n$, dann folgt
$$(\lambda_1 - \mu_1) v_1 + \dots + (\lambda_n - \mu_n) v_n = \mathbf{0}.$$
Aus der linearen Unabhängigkeit folgt $\lambda_1 - \mu_1 = \dots = \lambda_n - \mu_n = 0$. Q.E.D.

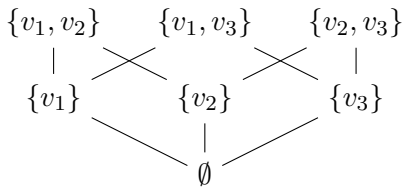
Unser nächstes Ziel ist zu zeigen, dass jeder Vektorraum mindestens eine Basis hat. Die Strategie dafür besteht darin, mit einem Erzeugendensystem zu beginnen (zum Beispiel ganz V) und in diesem Erzeugendensystem eine maximale linear unabhängige Menge zu finden. Hier ist ein konkretes Beispiel, das sich sehr einfach auf den endlich erzeugten Fall verallgemeinern lässt.

[7] Wenn der Körper der Skalare nicht klar aus dem Kontext bestimmt ist, dann sagt man auch **$\mathbb{K}$-Basis**.

Beispiel 3.28. Seien $v_1 = (1,0), v_2 = (0,1), v_3 = (1,1) \in \mathbb{R}^2$, und $S = \{v_1, v_2, v_3\}$. Wir haben

$$(x,y) = (x-1)\cdot v_1 + (y-1)\cdot v_2 + 1 \cdot v_3, \quad \forall\ (x,y) \in \mathbb{R}^2.$$

Also $\operatorname{Span}_{\mathbb{K}} S = \mathbb{R}^2$. Die einzige linear abhängige Teilmenge von S ist S selber, weil $v_1 + v_2 - v_3 = 0$, und die Matrizen $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ und $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ invertierbar sind. Das Hasse Diagramm der linear unabhängigen Teilmengen von $\{v_1, v_2, v_3\}$ ist also



Die maximale linear unabhängige Teilmengen von S sind also $T_3 = \{v_1, v_2\}$, $T_2\{v_1, v_3\}$ und $T_1 = \{v_2, v_3\}$. Um zu überprüfen, dass jede der drei Mengen eine Basis von $\mathbb{R}^2$ ist, muss man zeigen, dass $\operatorname{Span}_{\mathbb{K}} T_i = \operatorname{Span}_{\mathbb{K}} S = \mathbb{R}^2$. Aus Lemma 3.20 reicht es zu zeigen, dass $v_i \in \operatorname{Span}_{\mathbb{K}} T_i$. Das ist wahr, weil $v_1 + v_2 - v_3 = 0$.

Diese Strategie funktioniert auch ganz allgemein: es gibt immer ein Erzeugendensystem; das enthält mindestens eine linear unabhängige Menge; jede maximale linear unabhängige Teilmenge eines Erzeugendensystems ist eine Basis.

Satz 3.29. *Sei S ein Erzeugendensystem für V. Wenn $T \subseteq S$ eine maximale linear unabhängige Teilmenge von S ist, dann ist T eine Basis von V. Insbesondere, ist jede maximale linear unabhängige Menge eine Basis.*

Beweis-Skizze: Weil T linear unabhängig ist, müssen wir nur noch zeigen, dass $\operatorname{Span}_{\mathbb{K}} T = V$. Weil $V = \operatorname{Span}_{\mathbb{K}} S$, reicht es nach Lemma 3.20 zu zeigen, dass $v \in \operatorname{Span}_{\mathbb{K}} T$ für alle $v \in S$. Das ist trivial wenn $v \in T$.
Wenn $v \in S \setminus T$, dann, weil T maximal ist, muss $T \cup \{v\}$ linear abhängig sein. Es gibt also eine Linearkombination $\lambda_v v + \lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ mit $(\lambda_v, \lambda_1, \ldots, \lambda_n) \neq \mathbf{0}$, und $v_i \in T$. Wenn $\lambda_v = 0$, dann hätten wir eine nicht-triviale Linearkombination von Vektoren aus T die Null gibt, und das wäre ein Widerspruch ↯ zur linearen Unabhängigkeit von T. Also $\lambda_v \neq 0$, und, weil $\mathbb{K}$ ein Körper ist, haben wir

$$v = (-\lambda_v^{-1}\lambda_2)v_1 + \cdots + (-\lambda_v^{-1}\lambda_n)v_n \in \operatorname{Span}_{\mathbb{K}} \{v_1, \ldots, v_n\} \subseteq \operatorname{Span}_{\mathbb{K}} T. \qquad \text{Q.E.D.}$$

Satz 3.30. *Jeder Vektorraum hat eine Basis.*

Beweis-Skizze: Für den Nullraum haben wir in Beispiel 3.5.4. gesehen, dass die leere Menge eine Basis ist. Wir können also annehmen, dass V nicht trivial ist. Nach Satz 3.29 reicht es zu zeigen, dass es immer eine maximale linear unabhängige Menge gibt. Dafür wenden wir Zorns Lemma an:

Zorn's Lemma: Sei $(M, \preceq)$ eine geordnete Menge. Wenn jede total geordnete Teilmenge $N \subseteq M$ eine obere Schranke hat, dann existiert ein maximales Element in M.

In unserem Fall ist die partielle Ordnung die Mengeninklusion "$\subseteq$" und die Menge:

$$\mathcal{M} := \{S \subset V \ : \ S \text{ ist linear unabhängig}\}$$

Wir müssen also nur überprüfen, dass für jede total geordnete Teilmenge $\mathcal{N} \subseteq \mathcal{M}$ es eine obere Schranke $O \in \mathcal{M}$ gibt. Also eine linear unabhängige Menge O, mit $T \subseteq O$ für alle $T \in \mathcal{N}$.
Sei $\mathcal{N} \subseteq \mathcal{M}$ eine total geordnete Teilmenge von $\mathcal{M}$. Wir definieren

$$O := \bigcup_{T \in \mathcal{N}} T.$$

Sei $\{v_1, \dots, v_n\} \subseteq O$ eine endliche Teilmenge. Es gibt also $T_1, \dots, T_n \in \mathcal{N}$, sodass $v_i \in T_i$. Weil $\mathcal{N}$ total geordnet ist, können wir o.B.d.A annehmen, dass $T_1 \subseteq T_2 \subseteq \cdots \subseteq T_n$. Also $\{v_1, \dots, v_n\} \subseteq T_n$, und weil $T_n \in \mathcal{M}$ ist auch $\{v_1, \dots, v_n\}$ linear unabhängig. Offensichtlich gilt $T \subseteq O$ für alle $T \in \mathcal{N}$, also ist O die gesuchte obere Schranke für $\mathcal{N}$. Q.E.D.

Wir haben bis jetzt gezeigt, dass maximale linear unabhängige Mengen Basen sind. Es gilt auch, dass minimale Erzeugendensysteme Basen sind:

Satz 3.31. *Sei S ein minimales Erzeugendensystem eines $\mathbb{K}$-Vektorraumes V. Dann ist S eine Basis von V.*

Beweis-Skizze: Wir müssen zeigen, dass S linear unabhängig ist. Wir werden dafür Satz 3.25 anwenden. Sei $v \in S$ beliebig. Wir wollen zeigen, dass $v \notin \operatorname{Span}_{\mathbb{K}} S \setminus v$. Wenn $v \in \operatorname{Span}_{\mathbb{K}} S \setminus v$, dann aus Lemma 3.20 folgt $\operatorname{Span}_{\mathbb{K}} S \setminus v = \operatorname{Span}_{\mathbb{K}} S$ ↯ ein Widerspruch ↯ zur Minimalität des Erzeugendensystems S. Q.E.D.

[22] 10.1.'24

Korollar 3.32. Jedes Erzeugendensystem eines $\mathbb{K}$-Vektorraumes enthält eine Basis.

Korollar 3.33. Sei $S = \{v_1, \dots, v_n\} \subseteq V$. Folgende Aussagen sind äquivalent

(i) S ist eine maximale linear unabhängige Menge.

(ii) S ist ein minimales Erzeugendensystem von V.

(iii) S ist eine Basis von V.

Um zu zeigen, dass unendlich-erzeugte Vektorräume eine Basis besitzen, haben wir Zorn's Lemma angewendet. Dieses Lemma ist zum Auswahlaxiom äquivalent. Für gewisse unendlich-erzeugte Vektorräume braucht man dieses Axiom nicht um die Existenz einer Basis zu beweisen. (Z.B. $\mathbb{Q}[x]$ hat die unendliche $\mathbb{Q}$-Basis $\{1, x, x^2, \dots\}$.) Für andere unendlich-erzeugte Räume wissen wir aber nur, dass Basen existieren. Das ist der Fall für den $\mathbb{Q}$-Vektorraum $\mathbb{R}$: Es gibt Modelle der Mengenlehre in denen das Auswahlaxiom nicht gilt. In einem solchen Modell kann man zeigen, dass $\mathbb{R}$ keine Basis als $\mathbb{Q}$-Vektorraum besitzt. Wenn wir eine $\mathbb{Q}$-Basis von $\mathbb{R}$ explizit aufschreiben könnten, dann würde die Existenz der Basis nicht mehr von Auswahlaxiom abhängig sein.

Wir werden immer annehmen, dass das Auswahlaxiom gilt. Der Punkt hier ist, dass in gewisse unendlich-dimensionale Fälle wissen wir nur, dass es eine Basis gibt, aber wir können diese nicht aufschreiben.

Für endlich-erzeugte Vektorräume ist die Existenz einer Basis einfacher zu beweisen. Wir fangen mit einer einfachen Bemerkung an, die aber oft angewendet wird. In dieser Bemerkung ist die Voraussetzung, dass $\mathbb{K}$ ein Körper ist, entscheidend.

Bemerkung 3.34. Wenn $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ mit $\lambda_j \neq 0$, dann gilt $v_j \in \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, \widehat{v_j}, \ldots, v_n)$, und somit, nach Lemma 3.20, gilt

$$\mathrm{Span}_{\mathbb{K}}(v_1, \ldots, \widehat{v_j}, \ldots, v_n) = \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_n).$$

Die Idee des Beweises haben wir schon gesehen: $\lambda_j \neq 0 \overset{\mathbb{K}=\text{Körper}}{\Longrightarrow} \exists\ \lambda_j^{-1} \in \mathbb{K}$ sodass $\lambda_j \cdot \lambda_j^{-1} = 1$. Wir multiplizieren die lineare Kombination damit und bekommen

$$v_j = (-\lambda_j^{-1}\lambda_1)v_1 + \cdots + (-\lambda_j^{-1}\lambda_{j-1})v_{j-1} + (-\lambda_j^{-1}\lambda_{j+1})v_{j+1} + \cdots + (-\lambda_j^{-1}\lambda_n)v_n.$$

Satz 3.35. *Wenn $S = \{v_1, \ldots, v_n\} \subseteq V$ ein endlicher Erzeugendensystem ist, dann enthält S eine Basis.*

Beweis-Skizze: Wenn S linear abhängig ist, dann wenden wir Bemerkung 3.34 höchstens n Mal an um eine Basis zu bekommen. Wenn wir die Bemerkung tatsächlich n Mal anwenden müssen, dann bleibt nur die leere Menge übrig, und somit $V = 0$. Q.E.D.

3.6 Ergänzen eines linear unabhängiges Systems

Satz 3.36 (Ergänzungssatz)**.** *Sei V ein $\mathbb{K}$-Vektorraum. Sei $T = \{v_1, \ldots, v_m\} \subseteq V$ eine linear unabhängige Menge und sei $S = \{w_1, \ldots, w_n\} \subseteq V$ ein Erzeugendensystem von V. Dann gilt*

(i) $m \leqslant n$

(ii) Nach eventuellem Umbenennen der Vektoren in S gilt $\mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_m, w_{m+1}, \ldots, w_n\} = V$.

Beweis-Skizze: Wir beweisen den Satz durch Induktion nach m.
$\boxed{m = 1}$ Wenn $n < m = 1$, dann ist $S = \emptyset$. Dann wäre $V = 0$, und dieser Vektorraum enthält keine nicht-leere linear unabhängige Mengen. Also $n \geq 1 = m$.
Weil $\mathrm{Span}_{\mathbb{K}} S = V$, existieren nach Korollar 3.19 Skalare $\lambda_1, \ldots, \lambda_n$, sodass

$$v_1 = \lambda_1 w_1 + \cdots + \lambda_n w_n.$$

Weil T linear unabhängig ist, ist $v_1 \neq 0$, also existiert ein $i \in \{1, \ldots, n\}$, sodass $\lambda_i \neq 0$. Wir permutieren die Elementen in S, sodass $\lambda_1 \neq 0$. Weil $\mathbb{K}$ ein Körper ist, existiert λ_1^{-1}. Es gilt dann

$$w_1 = \lambda_1^{-1} v_1 - (\lambda_1^{-1}\lambda_2)w_2 - \cdots - (\lambda_1^{-1}\lambda_n)w_n.$$

Aus Lemma 3.20 haben wir

$$V = \mathrm{Span}_{\mathbb{K}} S = \mathrm{Span}_{\mathbb{K}}\{v_1, w_1, \ldots, w_n\} = \mathrm{Span}_{\mathbb{K}}\{v_1, w_2, \ldots, w_n\}.$$

$\boxed{m - 1 \Rightarrow m}$ Wir nehmen an, dass $\mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_{m-1}, w_m, \ldots, w_n\} = V$ und $m - 1 \leqslant n$. Nach Korollar 3.19 existieren Skalare $\mu_1, \ldots, \mu_n$, sodass

$$v_m = \mu_1 v_1 + \cdots + \mu_{m-1} v_{m-1} + \mu_m w_m + \cdots + \mu_n w_n.$$

Wenn $m - 1 = n$ oder wenn $\mu_m = \cdots = \mu_n = 0$, dann haben wir eine nicht-triviale Linearkombination von Vektoren aus T die Null ist. Das kann nicht sein, weil T linear unabhängig ist. Es gilt

also $m-1 < n$ (also (i)) und $\exists\, j \in \{m, \ldots, n\}$ mit $\mu_j \neq 0$. Nach einer eventuellen Permutation der Vektoren $w_m, \ldots, w_n$, können wir annehmen, dass $j = m$. Weil $\mathbb{K}$ ein Körper ist, existiert μ_m^{-1}, also

$$w_m = (-\mu_m^{-1}\mu_1)v_1 + \cdots + (\mu_m^{-1}\mu_{m-1})v_{m-1} + \mu_m^{-1}v_m + (-\mu_m^{-1}\mu_{m+1})w_{m+1} + \cdots + (-\mu_m^{-1}\mu_n)w_n.$$

Aus Lemma 3.20 haben wir

$$\begin{aligned} V &= \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_{m-1}, w_m, \ldots, w_n\} \\ &= \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_{m-1}, v_m, w_m, \ldots, w_n\} \\ &= \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_m, w_{m+1}, \ldots, w_n\}. \end{aligned}$$

Q.E.D.

Sowohl für den Beweis des folgenden Korollars, als auch für den Beweis von Satz 3.39 werden wir folgendes Lemma brauchen. Dieses Lemma gibt uns eine Methode um zu überprüfen, dass eine Menge endlich ist.

Lemma 3.37. *Sei A eine Menge. Wenn ein $n \in \mathbb{N}$ existiert, sodass $|A'| \leqslant n$ für alle endliche Teilmengen $A' \subseteq A$, dann ist A selbst eine endliche Menge.*

Beweis-Skizze: Sobald es ein n wie in der Aussage gibt, können wir ein minimales n_0 mit dieser Eigenschaft finden:

$$n_0 := \min\{n \in \mathbb{N} \;:\; |A'| \leqslant n \;\forall A' \subseteq A \text{ mit } A' \text{ endlich}\}.$$

Das heißt, es existiert $A' \subseteq A$ mit $|A'| = n_0$. Wenn A unendlich wäre, dann müsste $A \setminus A' \neq \emptyset$ gelten. Es existiert also ein $a \in A$ mit $a \notin A'$. Dann ist $A' \cup \{a\}$ eine endliche Teilmenge von A mit mehr als n_0 Elementen – ↯ ein Widerspruch.

Q.E.D.

Korollar 3.38. Sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum. Jedes linear unabhängiges System in V kann zu einer endlichen Basis ergänzt werden.

Beweis-Skizze: Weil V endlich erzeugt ist, existiert ein $n \in \mathbb{N}$ und ein Erzeugendensystem $S = \{w_1, \ldots, w_n\}$. Sei T eine linear unabhängige Menge. Nach Satz 3.36.(i) hat jede endliche Teilmenge davon höchstens n Elemente. Also, nach Lemma 3.37, ist T selbst endlich. Wir bezeichnen die Elementen von T mit $v_1, \ldots, v_m$. Nach Satz 3.36 kann man T zu einem Erzeugendensystem $\{v_1, \ldots, v_m, w_{m+1}, \ldots, w_m\}$ ergänzen. Wenn dieses nicht linear unabhängig ist, dann existieren $\lambda_1, \ldots, \lambda_n$, nicht alle Null, sodass

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + \lambda_{m+1} w_{m+1} + \cdots + \lambda_n w_n = 0.$$

Es muss aber ein $j > m$ geben, sodass $\lambda_j \neq 0$, sonst wären $v_1, \ldots, v_m$ nicht linear unabhängig. OBdA können wir $j = n$ annehmen, also $w_n \in \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_m, w_{m+1}, \ldots, w_{n-1})$. Man kann das weiterführen, und nur w-s los werden, bis eine Basis übrig bleibt. Q.E.D.

Ich will hier nochmals betonen, wie wichtig es ist, dass die Skalare aus einem Körper sind. Die Definitionen für lineare Unabhängigkeit, Erzeugendensystem und Basis kann man dann formulieren, wenn die Skalare aus einem beliebigen kommutativen Ring mit Eins R kommen. Zum Beispiel wenn $R = \mathbb{Z}$. Das macht man auch: solche Strukturen heißen **R-Moduln**. Die meisten Sätze über Erzeugendensysteme und linear unabhängige Mengen gelten aber nicht. Zum Beispiel, wenn wir $\mathbb{Z}$ als $\mathbb{Z}$-Modul betrachten, dann ist $S = \{2, 3\}$ ein Erzeugendensystem:

$$\forall\, z \in \mathbb{Z},\ \exists 2z, -z \in Z \quad \text{sodass} \quad z = (2z) \cdot 2 + (-z) \cdot 3.$$

Das ist aber keine Basis, weil es linear abhängig ist:

$$3 \cdot 2 + (-2) \cdot 3 = 0.$$

Es ist aber einfach zu sehen, dass sowohl $\mathrm{Span}_{\mathbb{Z}} 2 \neq \mathbb{Z}$ als auch $\mathrm{Span}_{\mathbb{Z}} 3 \neq \mathbb{Z}$. Also enthält dieses Erzeugendensystem keine Basis. Weiterhin, jede Teilmenge $\{z\} \subseteq \mathbb{Z}$ mit $z \neq \pm 1, 0$ ist linear unabhängig, kann aber nicht zu einer Basis ergänzt werden, weil jede Menge mit zwei Elementen linear abhängig ist.

Das $\mathbb{Z}$-Modul $\mathbb{Z}$ hat aber zwei mögliche Basen: $\{1\}$ und $\{-1\}$. Es gibt aber endlich erzeugte $\mathbb{Z}$-Moduln die gar keine Basis besitzen. Zum Beispiel jedes $\mathbb{Z}/n\mathbb{Z}$ mit $n \in \mathbb{N}_{>1}$.

Man sollte also die Existenz eines multiplikatives Inverses für nicht-Null Skalare gut schätzen.

3.7 Dimension

Korollar 3.39. Sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum.

(i) Es gibt eine endliche Basis von V.

(ii) Wenn B und C Basen von V sind, dann gilt $|B| = |C|$.

Beweis-Skizze: (i) Folgt aus Korollar 3.38 weil V endlich erzeugt ist.
(ii) Sei B die endliche Basis die aus (i) existiert, und sei C eine beliebige Basis. Es reicht zu zeigen, dass $|B| = |C|$.

Wir nehmen erstmals an, dass auch C endlich ist. Wir werden Teil (i) aus Satz 3.36 zwei Mal anwenden: ein Mal für $T = C$ und $S = B$, und ein Mal für $T = B$ und $S = C$. Es folgt also

$$|C| \leqslant |B| \qquad \text{und} \qquad |B| \leqslant |C|,$$

Also $|B| = |C|$.

Satz 3.36 gilt nur für endliche Mengen. Weil C a priori unendlich sein könnte, müssen wir noch zeigen, dass C endlich ist. Wir werden Lemma 3.37. Sei also C' eine beliebige endliche Teilmenge von C. Weil C linear unabhängig ist, dann ist per Definition auch C' linear unabhängig. Weil B eine Basis ist, und somit auch ein Erzeugendensystem, folgt aus Satz 3.36 Teil (i), für $T = C'$ und $S = B$, folgt $|C'| \leq |B|$. Weil das für alle endliche Teilmengen von C gilt, folgt, dass C selbst endlich sein muss.

Q.E.D.

Definition 3.40. Sei V ein $\mathbb{K}$-Vektorraum. Die **Dimension** von V ist

$$\dim_{\mathbb{K}} V := \begin{cases} n, & \text{falls } V \text{ eine endliche Basis mit } n \text{ Elementen besitzt,} \\ \infty, & \text{falls } V \text{ keine endliche Basis besitzt.} \end{cases}$$

Ein $\mathbb{K}$-Vektorraum ist **endlichdimensional** wenn $\dim_{\mathbb{K}} V \in \mathbb{N}$. Ein $\mathbb{K}$-Vektorraum ist **unendlichdimensional** wenn $\dim_{\mathbb{K}} V = \infty$.

[23] 15.1.'24

Satz 3.36 hat auch folgendes Korollar.

Korollar 3.41. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $W \subseteq V$ ein $\mathbb{K}$-UVR. Dann gilt

(i) W ist endlichdimensional mit $\dim_{\mathbb{K}} W \leqslant \dim_{\mathbb{K}} V$.

(ii) Wenn $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} V$, dann $W = V$.

Beweis-Skizze: (i) Sei B eine Basis von V. Weil V endlich erzeugt ist, ist auch $|B| < \infty$. Für jede endliche linear unabhängige Teilmenge $T \subseteq W$ folgt aus Satz 3.36, dass $|T| \leqslant |B|$. Also muss jede linear unabhängige Menge in W endlich sein. Also W ist endlichdimensional, und wieder aus Satz 3.36 folgt $\dim_{\mathbb{K}} W \leq \dim_{\mathbb{K}} V$.
(ii) Sei B' eine Basis von W. Wenn $W = \operatorname{Span}_{\mathbb{K}} B' \neq V$, dann existiert $v \in V$ mit $v \notin \operatorname{Span}_{\mathbb{K}} B'$. Aus Satz 3.25 folgt, dass $B' \cup \{v\}$ linear unabhängig in V ist, also $\dim_{\mathbb{K}} V > |B| = \dim_{\mathbb{K}} W$.
Q.E.D.

Beispiele:

1. $\dim_{\mathbb{K}} \mathbb{K}^n = n$, weil $\{e_i = (\delta_{1i}, \ldots, \delta_{ni}) \ : \ i = 1, \ldots, n\}$ eine Basis ist. Diese wird die **Standardbasis** des n-dimensionalen $\mathbb{K}$-Standardraumes genannt.

2. Wenn $v_1, \ldots, v_r$ linear unabhängige Vektoren eines $\mathbb{K}$-Vektorraumes V sind, dann gilt
$$\dim_{\mathbb{K}} \operatorname{Span}_{\mathbb{K}} \{v_1, \ldots, v_r\} = r.$$

3. Wenn $v_1 = (a, c), v_2 = (b, d) \in \mathbb{R}^2$, dann ist $\{v_1, v_2\}$ eine Basis genau dann, wenn $ad - bc \neq 0$. Das gilt, weil $\dim_{\mathbb{R}} \mathbb{R}^2 = 2$, also $\{v_1, v_2\}$ ist eine Basis wenn und nur wenn es eine linear unabhängige Menge ist. Wir haben schon gesehen, dass das äquivalent zu $ad - bc \neq 0$ ist.

4. Der Körper der komplexen Zahlen ist ein 2-dimensionaler $\mathbb{R}$-Vektorraum. Eine $\mathbb{R}$-Basis ist $\{1, i\}$. Eine andere Basis ist $\{1 + i, 1 - i\}$. Sind $\{1, 3, 2i\}$ oder $\{1 + i, -1 - i\}$ auch $\mathbb{R}$-Basen von $\mathbb{C}$?

5. **Übung.** Sei $\mathbb{Q}[\sqrt[3]{2}] := \{a + b\sqrt[3]{2} + c\sqrt[3]{4} \ : \ a, b, c \in Q\}$. Zeigen Sie, dass $\mathbb{Q}[\sqrt[3]{2}]$ ein Körper der $\mathbb{Q}$ enthält ist. Insbesondere ist $\mathbb{Q}[\sqrt[3]{2}]$ ein $\mathbb{Q}$-Vektorraum. Zeigen Sie, dass $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$ eine $\mathbb{Q}$-Basis von $\mathbb{Q}[\sqrt[3]{2}]$ ist. (Hinweis, das Polynom $x^3 - 2$ ist in $\mathbb{Q}$ unzerlegbar).

6. Die Menge $\{\text{Monome in } x_1, \ldots, x_n\}$ ist eine $\mathbb{K}$-Basis von $\mathbb{K}[x_1, \ldots, x_n]$. Ein Monom ist ein Produkt der Form $x_1^{d_1} \cdot \ldots \cdot x_n^{d_n}$ mit $d_i \in \mathbb{N}$. Wenn $d_i = 0$, gilt $x_i^{d_i} = 1$ und, wenn nicht alle d_i Null sind, können wir es weglassen. Zum Beispiel: $x_1^3 x_3^2 x_4^{103242}$ ist ein Monom mit $(d_1, d_2, d_3, d_4) = (3, 2, 0, 103242)$. Wenn alle $d_i = 0$, dann ist das Monom einfach 1.

7. Die Menge $\{\binom{x+i}{i} \ : \ i \in \mathbb{N}\}$ ist eine $\mathbb{Q}$-Basis von $\mathbb{Q}[x]$, wobei

$$\binom{x+i}{i} := \frac{(x+i)(x+i-1)\dots(x+1)}{1 \cdot 2 \cdot \dots \cdot i}.$$

Satz 3.42. *Seien V und W zwei endlich dimensionale $\mathbb{K}$-Vektorräume. Wenn $B_V = \{v_1, \dots, v_n\}$ eine Basis von V ist, und $B_W = \{w_1, \dots, w_m\}$ eine Basis von W ist, dann ist*

$$B := \{(v_1, \mathbf{0}_W), \dots, (v_n, \mathbf{0}_W), (\mathbf{0}_V, w_1), \dots, (\mathbf{0}_V, w_m)\}$$

eine Basis des $\mathbb{K}$-Vektorraumes $V \times W$. Insbesondere, ist $V \times W$ auch ein endlichdimensionaler Vektorraum und es gilt

$$\dim_{\mathbb{K}} V \times W = \dim_{\mathbb{K}} V + \dim_{\mathbb{K}} W.$$

Beweis-Skizze: Seien $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m \in \mathbb{K}$, sodass

$$\lambda_1(v_1, \mathbf{0}_W) + \dots + \lambda_n(v_n, \mathbf{0}_W) + \mu_1(\mathbf{0}_V, w_1) + \dots + \mu_m(\mathbf{0}_V, w_m) = \mathbf{0}.$$

Aus der Definition des Kartesisches Produktes folgen

$$\begin{aligned} \lambda_1 v_1 + \dots + \lambda_n v_n &= \mathbf{0}_V \\ \mu_1 w_1 + \dots + \mu_m w_m &= \mathbf{0}_W. \end{aligned}$$

Weil B_V und B_W Basen sind, folgt $\lambda_1 = \dots = \lambda_n = \mu_1 = \dots = \mu_m = 0$, also B ist linear unabhängig.

Sei $(v, w) \in V \times W$ beliebig. Also $v \in V$ und $w \in W$. Weil B_V und B_W Basen (also Erzeugendensysteme) sind, existieren $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m \in \mathbb{K}$, sodass

$$\begin{aligned} v &= \lambda_1 v_1 + \dots + \lambda_n v_n \\ w &= \mu_1 w_1 + \dots + \mu_m w_m. \end{aligned}$$

Es folgt $(v, w) = \lambda_1(v_1, \mathbf{0}_W) + \dots + \lambda_n(v_n, \mathbf{0}_W) + \mu_1(\mathbf{0}_V, w_1) + \dots + \mu_m(\mathbf{0}_V, w_m)$, also ist B auch ein Erzeugendensystem von $V \times W$. Q.E.D.

Kapitel 4

Neue Vektorräume aus alte Vektorräume

4.1 Unabhängigkeit und Span unter lineare Abbildungen

Wir erinnern zu erst, dass, wenn V und W zwei $\mathbb{K}$-Vektorräume sind, dann ist eine Abbildung $f : V \longrightarrow W$ genau dann $\mathbb{K}$-linear, wenn

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) \qquad \forall\, \lambda_1, \lambda_2 \in \mathbb{K} \text{ und } \forall\, v_1, v_2 \in V.$$

Die Menge aller $\mathbb{K}$-linearen Abbildungen von V nach W bezeichnen wir als $\mathrm{Hom}_{\mathbb{K}}(V, W)$. Nach Definition 3.5 und Bemerkung 3.6 ist f genau dann ein Isomorphismus von $\mathbb{K}$-Vektorräumen, wenn f $\mathbb{K}$-linear und bijektiv ist.

Satz 4.1. *Sei $f : V \longrightarrow W$ eine $\mathbb{K}$-lineare Abbildung.*

(i) *Wenn $U \subseteq_{\mathbb{K}} V$, dann $f(U) \subseteq_{\mathbb{K}} W$. Insbesondere* $\mathrm{Bild}\, f = f(V)$ *ist ein $\mathbb{K}$-Unterraum von W.*

(ii) *Wenn $U \subseteq_{\mathbb{K}} W$, dann $f^{-1}(U) \subseteq_{\mathbb{K}} V$. Insbesondere ist* $\mathrm{Ker}\, f = f^{-1}(0)$ *ein $\mathbb{K}$-Unterraum von V.*

(iii) *Für jede Teilmenge $S \subseteq V$ gilt* $f(\mathrm{Span}_{\mathbb{K}} S) = \mathrm{Span}_{\mathbb{K}} f(S)$.

(iv) *Wenn eine Teilmenge $S \subseteq V$ linear abhängig ist, dann ist auch $f(S) \subseteq W$ linear abhängig. Insbesondere,* $\dim_{\mathbb{K}} f(V) \leqslant \dim_{\mathbb{K}} V$.

(v) *Wenn f ein Isomorphismus ist, dann gilt* $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$.

Beweis-Skizze:

(i) $0 \in U$, also $f(0) = 0 \in f(U) \neq \emptyset$. Seien $\lambda_1, \lambda_2 \in \mathbb{K}$ und $w_1, w_2 \in f(U)$ beliebig. Dann $\exists u_1, u_2 \in U$ mit $f(u_i) = w_i$. Also

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 f(u_1) + \lambda_2 f(u_2) = f(\lambda_1 u_1 + \lambda_2 u_2) \in f(U),$$

weil U ein Untervektorraum ist, also $\lambda_1 u_1 + \lambda_2 u_2 \in U$.

(ii) Analog zu (i)

(iii) Wir haben

$$\begin{aligned} w \in f(\mathrm{Span}_{\mathbb{K}} S) &\iff w = f(\lambda_1 v_1 + \cdots + \lambda_n v_n) \text{ mit } \lambda_i \in \mathbb{K} \text{ und } v_i \in S \\ &\iff w = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) \\ &\iff w \in \mathrm{Span}_{\mathbb{K}} f(S). \end{aligned}$$

(iv) Wir zeigen beweisen, dass die Kontraposition wahr ist:

$$f(S) \text{ ist linear unabhängig } \Rightarrow S \text{ ist linear unabhängig.}$$

Seien $v_1, \ldots, v_n \in S$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ mit $\sum_{i=1}^n \lambda_i v_i = 0$. Dann gilt auch

$$\textstyle\sum_{i=1}^n \lambda_i f(v_i) = f\left(\sum_{i=1}^n \lambda_i v_i\right) = f(0) = 0.$$

Also, weil $f(v_1), \ldots, f(v_n) \in f(S)$ und $f(S)$ linear unabhängig ist, folgt $\lambda_1 = \cdots = \lambda_n = 0$.

Insbesondere, wenn B eine Basis von $f(V)$ ist, dann, weil für alle $w \in B$ existiert $v \in V$ mit $f(v) = w$, ist $B = f(A)$ mit $A \subseteq V$ und $\#A = \#B$. Wir haben gerade bewiesen, dass A linear unabhängig ist weil B linear unabhängig ist. Es folgt also

$$\dim_{\mathbb{K}} f(V) = \#B = \#A \leqslant \dim_{\mathbb{K}} V.$$

(v) **Übung**

Q.E.D.

4.2 Dimension von Bild und Kern

Der folgende Satz ist einer der wichtigsten[1] der linearen Algebra 1 Vorlesung.

Satz 4.2 (Dimensionssatz)**.** *Seien V und W zwei $\mathbb{K}$-Vektorräume, mit V endlichdimensional. Wenn $f : V \longrightarrow W$ ein Homomorphismus von $\mathbb{K}$-Vektorräume ist, dann gilt*

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \mathrm{Ker}(f) + \dim_{\mathbb{K}} \mathrm{Bild}(f).$$

Beweis-Skizze: Aus Satz 4.1 ist $\mathrm{Ker}(f)$ ein $\mathbb{K}$-Unterraum von V. Nach Korollar 3.41 ist $\mathrm{Ker}(f)$ auch endlichdimensional. Sei $\{v_1, \ldots, v_r\}$ eine Basis von $\mathrm{Ker}(f)$. Aus Satz 3.36 kann man diese zu einer Basis von V ergänzen: $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$. Wir haben also $\dim_{\mathbb{K}} \mathrm{Ker}(f) = r \leqslant n = \dim_{\mathbb{K}} V$. Es reicht zu zeigen, dass $\{f(v_{r+1}), \ldots, f(v_n)\}$ eine Basis von $\mathrm{Bild}(f)$ ist.

$\boxed{\{f(v_{r+1}), \ldots, f(v_n)\} \text{ ist l.u.}}$ Seien $\lambda_{r+1}, \ldots, \lambda_n \in \mathbb{K}$, sodass

$$\lambda_{r+1} f(v_{r+1}) + \cdots + \lambda_n f(v_n) = 0.$$

[1] Unter "wichtig" verstehe ich, dass es sehr viele Folgerungen und Anwendungen hat.

Weil f $\mathbb{K}$-linear ist, haben wir $f(\lambda_{r+1}v_{r+1} + \cdots + \lambda_n v_n) = \lambda_{r+1}f(v_{r+1}) + \cdots + \lambda_n f(v_n) = 0$. Das heißt

$$\lambda_{r+1}v_{r+1} + \cdots + \lambda_n v_n \in \mathrm{Ker}(f) = \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\}.$$

Aus Korollar 3.19 existieren dann $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$, sodass $\lambda_{r+1}v_{r+1} + \cdots + \lambda_n v_n = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Also

$$(-\lambda_1)v_1 + \cdots + (-\lambda_r)v_r + \lambda_{r+1}v_{r+1} + \cdots + \lambda_n v_n = 0.$$

Weil $\{v_1, \ldots, v_n\}$ eine Basis ist, folgt $\lambda_i = 0, \quad \forall i = 1, \ldots, n$. Also $\{f(v_{r+1}), \ldots, f(v_n)\}$ ist linear unabhängig.

$\boxed{\mathrm{Span}_{\mathbb{K}}\{f(v_{r+1}), \ldots, f(v_n)\} = \mathrm{Bild}(f)}$ Sei $w \in \mathrm{Bild}(f)$ beliebig. Das heißt, es existiert ein $v \in V$, sodass $f(v) = w$. Weil $V = \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_n\}$, existieren $\lambda_i \in \mathbb{K}$, sodass $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Wir haben also

$$w = f(v) = f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_r f(v_r) + \lambda_{r+1} f(v_{r+1}) + \cdots + \lambda_n f(v_n).$$

Aus $f(v_i) = 0$ für alle $i = 1, \ldots, r$ folgt dann $w = \lambda_{r+1}f(v_{r+1}) + \cdots + \lambda_n v_n$. Q.E.D.

[24] 17.1.'24

Korollar 4.3 (Grassmann)**.** Seien U_1, U_2 zwei $\mathbb{K}$-Untervektorräume eines endlich dimensionales $\mathbb{K}$-Vektorraumes V. Dann gilt

$$\dim_{\mathbb{K}}(U_1 + U_2) = \dim_{\mathbb{K}} U_1 + \dim_{\mathbb{K}} U_2 - \dim_{\mathbb{K}}(U_1 \cap U_2).$$

Beweis-Skizze: Sei $f : U_1 \times U_2 \longrightarrow V$ gegeben durch $f(v_1, v_2) := v_1 - v_2$. Für alle $v_i, u_i \in U_i$ und $\lambda \in \mathbb{K}$ haben wir

$$\begin{aligned} f((v_1, v_2) + (u_1, u_2)) &= f(v_1 + u_1, v_2 + u_2) = v_1 - v_2 + u_1 - u_2 = f(v_1, v_2) + f(u_1, u_2) \\ f(\lambda(v_1, v_2)) &= f(\lambda v_1, \lambda v_2) = \lambda v_1 - \lambda v_2 = \lambda f(v_1, v_2). \end{aligned}$$

Also f ist eine $\mathbb{K}$-lineare Abbildung. Aus Satz 3.15 folgt gleich, dass $\mathrm{Bild}(f) = U_1 + U_2$. Also

$$\dim_{\mathbb{K}} \mathrm{Bild}(f) = \dim_K(U_1 + U_2).$$

Aus Satz 3.42 haben wir

$$\dim_{\mathbb{K}} U_1 \times U_2 = \dim_{\mathbb{K}} U_1 + \dim_{\mathbb{K}} U_2.$$

Wir zeigen jetzt, dass $\mathrm{Ker}(f) = \{(v, v) \quad : \quad v \in U_1 \cap U_2\}$.

Die Inklusion "$\supseteq$" ist klar: $f(v, v) = v - v = 0$.
Die Inklusion "$\subseteq$" gilt, weil $f(v_1, v_2) = 0 \Leftrightarrow v_1 - v_2 = 0 \Leftrightarrow v_1 = v_2$. Also $v_1 = v_2 \in U_2$ und $v_2 = v_1 \in U_1$.

Die Abbildung $g : U_1 \cap U_2 \longrightarrow \{(v, v) \quad : \quad v \in U_1 \cap U_2\}$ gegeben durch $g(v) := (v, v)$ ist ein Isomorphismus von $\mathbb{K}$-Vektorräume, also aus Satz 4.1 Teil (v) haben wir

$$\dim_{\mathbb{K}} \mathrm{Ker}(f) = \dim_{\mathbb{K}}(U_1 \cap U_2).$$

Die Formel folgt jetzt direkt aus Satz 4.2. Q.E.D.

Lemma 4.4. *Wenn V und W zwei $\mathbb{K}$-Vektorräume mit $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W < \infty$ sind und wenn $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, dann sind folgende Aussagen äquivalent:*

(i) *f ist ein Isomorphismus.*

(ii) *f ist injektiv.*

(iii) *f ist surjektiv.*

Beweis-Skizze: Wir haben aus Satz 4.2, dass

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \mathrm{Ker}(f) + \dim_{\mathbb{K}} \mathrm{Bild}(f). \tag{4.1}$$

(i) $\Rightarrow$(ii) f ist ein Isomorphismus, also bijektiv und somit injektiv.

(ii)$\Rightarrow$(iii) f ist injektiv $\iff \mathrm{Ker}(f) = 0 \iff \dim_{\mathbb{K}} \mathrm{Ker}(f) = 0$. Also aus (4.1) folgt

$$\dim_{\mathbb{K}} \mathrm{Bild}(f) = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W.$$

Weil $\mathrm{Bild}(f) \subseteq_{\mathbb{K}} W$, folgt aus Korollar 3.41, dass $\mathrm{Bild}(f) = W$, also dass f surjektiv ist.

(iii)$\Rightarrow$(i) f ist surjektiv $\iff \mathrm{Bild}(f) = W \iff \dim_{\mathbb{K}} \mathrm{Bild}(f) = \dim_{\mathbb{K}} W$. Also aus (4.1) folgt

$$\dim_{\mathbb{K}} \mathrm{Ker}(f) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} \mathrm{Bild}(f) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} W = 0 \iff \mathrm{Ker}_{\mathbb{K}}(f) = 0.$$

Also f injektiv und, weil es auch surjektiv ist, ist es bijektiv. Also f ist ein Isomorphismus.

Q.E.D.

4.3 Direkte Summen

Wir werden in diesem Abschnitt mit dem ganz allgemeinen und abstrakten Fall beginnen. Danach werden wir uns dem einfacheren und speziellen Fall der direkten Summen von Untervektorräumen zuwenden. Ich habe diese Reihenfolge aus folgendem Grund gewählt: Die direkte Summe und das direkte Produkt einer Familie von $\mathbb{K}$-Vektorräumen unterscheiden sich nur, wenn die Familie unendlich ist. Also, um zu zeigen, dass es sich tatsächlich um etwas Neues handelt, beginnen wir damit.

Sei $(V_i)_{i\in I}$ eine Familie von $\mathbb{K}$-Vektorräumen. Wir haben gesehen (3.1), dass das direkte Produkt

$$\prod_{i\in I} V_i = \{(v_i)_{i\in I} \ : \ \text{mit } v_i \in V_i \ \forall \ i \in I\}$$

ein $\mathbb{K}$-Vektorraum mit der Komponenten-weise Addition und der Skalaren Multiplikation $\lambda\cdot(v_i) = (\lambda v_i)$. Aus den Definitionen folgt, dass die Teilmenge

$$\{(v_i)_{i\in I} \ : \ |\{i \in I \ : \ v_i \neq 0\}| < \infty\}$$

ein $\mathbb{K}$-UVR von $\prod_{i\in I} V_i$ ist. Dieser $\mathbb{K}$-Vektorraum ist die direkte Summe.

Definition 4.5. Die **direkte Summe** der Familie $(V_i)_{i\in I}$ ist der $\mathbb{K}$-Vektorraum

$$\bigoplus_{i\in I} V_i := \{(v_i) \in \prod_{i\in I} V_i \ : \ \text{nur für endliche viele Indizes } i \text{ gilt } v_i \neq 0\}.$$

Aus den Definitionen ist es klar, sowohl, dass wenn $|I| = \infty$, dann

$$\prod_{i\in I} V_i \neq \bigoplus_{i\in I} V_i.$$

als auch, dass wenn $I = \{1, \dots, n\}$, dann

$$\prod_{i=1}^{n} V_i = V_1 \times \cdots \times V_n = V_1 \oplus \cdots \oplus V_n = \bigoplus_{i=1}^{n} V_i.$$

Man könnte eigentlich beide Konzepte durch universelle Eigenschaften definieren. Dadurch kann man auch den Unterschied (und eine Dualität) zwischen direkte Produkte und direkte Summen bemerken[2]. Für das direkte Produkt, werden folgende surjektive $\mathbb{K}$-lineare Abbildungen, die *kanonische Projektionen* genannt werden, die universelle Eigenschaft bestimmen:

$$\forall\ k \in I, p_k : \prod_{i\in I} V_i \longrightarrow V_k, \quad p_k((v_i)) := v_k.$$

Für die direkte Summe, werden folgende injektive $\mathbb{K}$-lineare Abbildungen, die *kanonische Inklusionen* genannt werden, die universelle Eigenschaft bestimmen:

$$\forall\ k \in I, i_k : V_k \hookrightarrow \bigoplus_{i\in I} V_i \quad i_k(v) := (\delta_{ik} v)_{i\in I},$$

wobei δ_{ik} das Kronecker Delta ist.

Bemerkung 4.6. Jeder nicht trivialer Vektor in $\bigoplus_{k\in I} V_k$ ist auf eindeutiger Weise als (endliche) Summe von nicht trivialen Vektoren der Form $i_k(v_k)$ ausdrückbar. Genauer gesagt, wenn $v \in \bigoplus_{k\in I} V_k$, dann existieren eindeutig bestimmte $k_1, \dots, k_r \in I$ und $v_{k_1} \in V_{k_1} \setminus \{0\}, \dots, v_{k_r} \in V_{k_r} \setminus \{0\}$, sodass

$$v = i_{k_1}(v_1) + \cdots + i_{k_r}(v_{k_r}).$$

Beweis-Skizze: Für die Existenz, hat man als $k_1, \dots, k_r \in I$ genau die Indizes der Stellen an denen die Einträge in $v = (v_k)$ nicht Null sind. Es gibt mindestens eine solche Stelle, weil $v \neq 0$, und per Definition sind es auch endlich-viele: sagen wir r. Wenn $I = \mathbb{Z}$ wäre, dann könnten wir das auch darstellen:

$$v = (\dots, 0, v_{k_1}, 0, \dots, 0, v_{k_2}, 0, \dots\dots, 0, v_{k_r}, 0, \dots)$$

Wir hätten dann, wieder per Definition

$$\begin{aligned} v \;=\; & (\dots, 0, v_{k_1}, 0, \dots, 0, 0, 0, \dots\dots, 0, 0, 0, \dots) + \\ & +(\dots, 0, 0, 0, \dots, 0, v_{k_2}, 0, \dots\dots, 0, 0, 0, \dots) + \\ & \vdots \\ & +(\dots, 0, 0, 0, \dots, 0, 0, 0, \dots\dots, 0, v_{k_r}, 0, \dots). \end{aligned}$$

[2]Diese universellen Eigenschaften ermöglichen es, "direkte Produkte" und "direkte Summen" in beliebigen Kategorien zu definieren. In so einem abstrakten Kontext werden die Objekte, die die universelle Eigenschaft der direkten Summe von Vektorräumen erfüllen, als *co-Produkte* bezeichnet. Diese sind nicht immer jedoch direkte Summen.

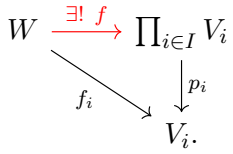
Das gilt auch für beliebige Indexmengen I, also

$$v = i_{k_1}(v_{k_1}) + i_{k_2}(v_{k_2}) + \cdots + i_{k_r}(v_{k_r}).$$

Das diese Schreibweise eindeutig ist folgt aus der Injektivität der Abbildungen i_k und aus der Voraussetzung, dass alle $v_{k_j} \neq 0$. Q.E.D.

Satz 4.7. *Sei $(V_i)_{i\in I}$ eine Familie von $\mathbb{K}$-Vektorräumen. Das direkte Produkt $\prod_{i\in I} V_i$ ist bis auf eindeutigen Isomorphismus der einzige $\mathbb{K}$-Vektorraum mit folgender universeller Eigenschaft:*

Für jeden anderen $\mathbb{K}$-Vektorraum W, der zusammen mit einer Familie von $\mathbb{K}$-linearen Abbildungen $(f_i : W \longrightarrow V_i)_{i\in I}$ gegeben ist, existiert genau eine $\mathbb{K}$-lineare Abbildung $f : W \longrightarrow \prod_{i\in I} V_i$, sodass $p_i \circ f = f_i$ für alle $i \in I$; dass heißt, sodass folgende Diagramme kommutativ für alle $i \in I$ sind:

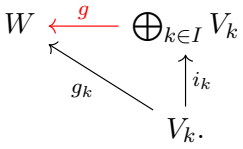


Beweis-Skizze: Der Beweis ist kürzer als die Aussage. Um zu zeigen, dass $\prod V_i$ die Eigenschaft hat, setzt man $f(w) := (f_i(w))_{i\in I}$. Für die Eindeutigkeit: wenn P auch diese Eigenschaft hätte, dann sind die eindeutig bestimmten Abbildungen von P nach $\prod V_i$ und von $\prod V_i$ nach P invers zu einander. Q.E.D.

Die direkte Summe hat eine ähnliche, aber duale Eigenschaft.

Satz 4.8. *Sei $(V_i)_{i\in I}$ eine Familie von $\mathbb{K}$-Vektorräumen. Die direkte Summe $\bigoplus_{i\in I} V_i$ ist bis auf eindeutigen Isomorphismus der einzige $\mathbb{K}$-Vektorraum mit folgender universeller Eigenschaft:*

Für jeder andere $\mathbb{K}$-Vektorraum W, der zusammen mit einer Familie von $\mathbb{K}$-linearen Abbildungen $(g_i : V_i \longrightarrow W)_{i\in I}$ gegeben ist, existiert genau eine $\mathbb{K}$-lineare Abbildung $g : \bigoplus_{i\in I} V_i \longrightarrow W$, sodass $g \circ i_k = g_k$ für alle $k \in I$; dass heißt, sodass folgende Diagramme kommutativ für alle $k \in I$ sind



Beweis-Skizze: Um g zu definieren, braucht man Bemerkung 4.6. Sei also $v \in \bigoplus_{k\in I} V_k$, und sei

$$v = i_{k_1}(v_{k_1}) + i_{k_2}(v_{k_2}) + \cdots + i_{k_r}(v_{k_r})$$

die eindeutige nicht-triviale Zerlegung von v. Dann definiert man $g : \bigoplus_{k\in I} V_k \longrightarrow W$ durch

$$g(v) := g_{k_1}(v_{k_1}) + g_{k_2}(v_{k_2}) + \cdots + g_{k_r}(v_{k_r}).$$

Die $\mathbb{K}$-Linearität folgt aus der $\mathbb{K}$-Linearität der g_k und der i_k. Die Kommutativität der Diagramme

dann folgt per Definition. Die Eindeutigkeit bis auf eindeutiger Isomorphismus funktioniert analog zum Satz 4.7. Q.E.D.

Beispiele:

1. Sei für jedes $d \in \mathbb{N}$, sei $V_d = \mathbb{K}$, der 1-dimensionale $\mathbb{K}$-Standardraum. Wir betrachten den Polynomring $\mathbb{K}[x]$ als $\mathbb{K}$-Vektorraum und haben dann

$$\mathbb{K}[x] \simeq \bigoplus_{d\in\mathbb{N}} \mathbb{K}.$$

   Der Isomorphismus ist durch $P = a_0 + a_1 x + \cdots + a_n x^n \longmapsto (a_i)_{i\in\mathbb{N}}$, wobei $a_i = 0$ für $i > n$.

   Mit Hilfe der direkten Summen kann man eigentlich Polynomringe in mehrere Variablen mit Koeffizienten in einem Körper genauer definieren. Dafür braucht man aber auch Dualräume und das Tensorprodukt.

   Das kartesische (oder direkte) Produkt dieser Vektorräume ist $\mathbb{K}$-isomorph zu dem $\mathbb{K}$-Vektorraum der **formalen Potenzreihen**

$$K[[x]] = \{a_0 + a_1 x + \cdots + a_r x^r + \ldots \ : \ a_i \in \mathbb{K}\}.$$

2. Wenn die Indexmenge I endlich ist, dann ist $\bigoplus_{i\in I} V_i = \prod_{i\in I} V_i$. Wir haben insbesondere

$$\begin{aligned} \bigoplus_{i=1}^{n} \mathbb{K} &= \mathbb{K}^n \\ \bigoplus_{i=1}^{r} \mathbb{K}^{m_i} &= \mathbb{K}^{m_1} \oplus \cdots \oplus \mathbb{K}^{m_r} \simeq \mathbb{K}^{m_1+\cdots+m_r}. \end{aligned}$$

[25] 22.1.'24

Was man sich aus der obigen Definition merken muss, ist, dass die direkte Summe von endlich vielen oder unendlich vielen Vektorräume immer definiert ist. Jetzt werden wir uns direkte Summen von Untervektorräumen eines gegebenen Vektorraumes anschauen. Einige Autoren machen einen Unterschied zwischen *inneren* und *äußeren* direkten Summen. Wir verwenden diese Bezeichnung jedoch nicht.

Für zwei $\mathbb{K}$-Untervektorräume $U_1, U_2 \subseteq V$ eines $\mathbb{K}$-Vektorraumes V haben wir in Definition 3.14 die Summe $U_1 + U_2$ als $\mathrm{Span}_{\mathbb{K}}\{U_1 \cup U_2\}$ definiert, und dann in Satz 3.15 bewiesen, dass $U_1 + U_2 = \{u_1 + u_2 \ : \ u_i \in U_i\}$. Genauso definiert man für (endlich-viele) $\mathbb{K}$-UVR $U_1, \ldots, U_r \subseteq V$

$$U_1 + \cdots + U_r = \mathrm{Span}_{\mathbb{K}}(U_1 \cup \cdots \cup U_r).$$

Es folgt genau wie im Beweis von Satz 3.15, dass

$$U_1 + \cdots + U_r = \{u_1 + \cdots + u_r \ : \ u_i \in U_i\}. \tag{4.2}$$

Satz 4.9. *Seien $U_1, \ldots, U_r$ Unterräume eines $\mathbb{K}$-Vektorraumes V. Folgende Aussagen sind äquivalent.*

(i) $U_1 + \cdots + U_r = V$ *und* $U_i \cap (\sum_{j\neq i} U_j) = \{0_V\}$ *für alle* $i = 1, \ldots, n$

(ii) *Für jeder* $v \in V$*, existieren* eindeutige $u_i \in U_i$ *so dass*

$$v = u_1 + \cdots + u_r.$$

(iii) *Für jeder* $v \in V$ *existieren* $u_i \in U_i$*, sodass* $v = u_1 + \cdots + u_r$ *und wenn* $w_1 + \cdots + w_r = 0_V$ *mit* $w_i \in U_i$*, dann folgt* $w_i = 0_V$*,* $\forall\ i = 1, \ldots, r$*.*

(iv) *Die Abbildung* $f : U_1 \oplus \cdots \oplus U_r \longrightarrow V$ *gegeben durch*

$$f(u_1, \ldots, u_r) = u_1 + \cdots + u_r$$

ist ein $\mathbb{K}$*-Vektorraum Isomorphismus.*

Beweis-Skizze: (i)⇒(ii) Aus $V = U_1 + \cdots + U_r$ und (4.2) folgt die Existenz. Wenn

$$v = u_1 + \cdots + u_r = u_1' + \cdots + u_r' \qquad \text{mit } u_i, u_i' \in U_i\ \forall\ i,$$

dann folgt für jeden Index $j = 1, \ldots, r$, dass

$$U_j \ni u_j - u_j' = \sum_{i \neq j} (u_i' - u_i) \in U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_r.$$

Also $u_j - u_j' \in U_i \cap (\sum_{j \neq i} U_j) = \{0_V\}$, und $u_j = u_j'\ \forall\ j = 1, \ldots, r$.

(ii)⇒(iii) Das ist trivial.

(iii)⇒(iv) Die $\mathbb{K}$-Linearität ist offensichtlich. Wir brauchen also nur die Bijektivität zu zeigen.

Injektivität: Wenn $f(\mathbf{u}) = f(\mathbf{u}')$, dann $(u_1 - u_1') + \cdots + (u_r - u_r') = 0_V$, also aus (iii) folgt $\mathbf{u} = \mathbf{u}'$.

Surjektivität: Folgt direkt aus $\forall\ v \in V$ existieren $u_i \in U_i$ für jeder $i = 1, \ldots, r$ mit $v = u_1 + \cdots + u_r$.

(iv)⇒(i) $U_1 + \cdots + U_r = V$ folgt aus der Surjektivität. Wenn $u \in U_i \cap (\sum_{j \neq i} U_j)$, dann haben wir $u = \sum_{j \neq i} u_j$ mit $u_j \in U_j$. Wir haben also

$$f(0, \ldots, \underset{i}{u}, \ldots, 0) = f(u_1, \ldots, u_{i-1}, \underset{i}{0}, u_{i+1}, \ldots, u_r).$$

Aus der Injektivität von f folgt $u = 0$. Q.E.D.

Wenn die $\mathbb{K}$-Untervektorräume $U_1, \ldots, U_r \subseteq V$ die äquivalente Bedingungen aus Satz 4.9 erfüllen, dann sagen wir, dass V die **direkte Summe** von $U_1, \ldots, U_r$ ist, und wir schreiben einfach

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_r.$$

In der Literatur (z.B. [**?**, Abschnitt 1.6,S.45]) findet man auch folgende Ausdrucksweise:

" Die Summe [...] von linearen Unterräumen $U_1, \ldots, U_r \subset V$ *wird auch mit* $U_1 + \cdots + U_r$ *bezeichnet, und man schreibt* $U_1 \oplus \cdots \oplus U_r$ ***falls diese Summe direkt ist.*** *"*

Das kann zu der Verwirrung führen, dass zum Beispiel $U \oplus U$ nicht erlaubt wäre, weil die Summe $U + U$ nicht "direkt ist". Aber $\mathbb{K} \oplus \mathbb{K}$ ist erlaubt und ist isomorph zu $\mathbb{K}^2$. Wir haben eine einzige ("echte",

allgemeine, oder für manche "äußere") direkte Summe definiert. Und das obige Zitat würde man so interpretieren: Wir schreiben $U_1 + \cdots + U_r = U \oplus \cdots \oplus U_r$ falls diese zwei Vektorräume isomorph sind; wir sagen einfach, dass *die Summe* $U_1 + \cdots + U_r$ *ist direkt*, anstatt *die Summe* $U_1 + \cdots + U_r$ *ist isomorph zu der direkten Summe* $U_1 \oplus \cdots \oplus U_r$.

Korollar 4.10. Wenn $V = U_1 \oplus \cdots \oplus U_r$ und $\forall\ i \in \{1, \ldots, n\}$ die Menge B_i eine Basis von U_i ist, dann ist

$$B = B_1 \cup \cdots \cup B_i$$

eine Basis von V. Insbesondere, wenn V endlichdimensional ist, dann haben wir

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} U_1 + \cdots + \dim_{\mathbb{K}} U_n.$$

Bemerkung 4.11. Wenn $U_1, U_2 \subseteq V$ zwei $\mathbb{K}$-UVR von V sind, dann haben wir

$$\begin{aligned} V = U_1 \oplus U_2 &\iff U_1 \cap U_2 = \{0\} \quad \text{und} \quad U_1 + U_2 = V \\ &\iff U_1 \cap U_2 = \{0\} \quad \text{und} \quad \dim_{\mathbb{K}} U_1 + \dim_{\mathbb{K}} U_2 = \dim_{\mathbb{K}} V \\ &\iff U_1 + U_2 = V \quad \text{und} \quad \dim_{\mathbb{K}} U_1 + \dim_{\mathbb{K}} U_2 = \dim_{\mathbb{K}} V \end{aligned}$$

Das ganze folgt ganz einfach aus der kurzen exakten Folge

$$0 \longrightarrow U_1 \cap U_2 \xrightarrow{g} U_1 \oplus U_2 \xrightarrow{f} U_1 + U_2 \longrightarrow 0 \tag{4.3}$$

wobei $g(u) := (u, -u)$ und $f(u_1, u_2) = u_1 + u_2$.

Definition 4.12. Sei $U \subseteq V$ ein $\mathbb{K}$-Unterraum von V. Ein **Komplementärraum** von U in V ist ein $\mathbb{K}$-Unterrwaum $W \subseteq V$ mit der Eigenschaft, dass $U \oplus W = V$.

Das heißt ein Komplementärraum W von U erfüllt $U + W = V$ und $U \cap W = \{0\}$. Es ist wichtig zu bemerken, dass ein $\mathbb{K}$-Unterraum von V mehrere Komplementärräume hat. Zum Beispiel $W = \mathrm{Span}_{\mathbb{K}} \{(x, y)\} \subseteq \mathbb{R}^2$ ist ein Komplementärraum von $U = \mathrm{Span}_{\mathbb{K}} \{(1, 0)\} \subseteq \mathbb{R}^2$ genau dann, wenn $y \neq 0$.

4.4 Quotientenräume

Sei $U \subseteq V$ ein $\mathbb{K}$-UVR des $\mathbb{K}$-Vektorraumes V. Wir haben insbesondere, dass $(U, +) \leq (V, +)$, das heißt $(U, +)$ ist eine Untergruppe von $(V, +)$. Weil $(V, +)$ abelsch ist, dann ist automatisch $(U, +)$ ein Normalteiler von $(V, +)$. Nach Teil 1.6.9 aus Kapitel 1 folgt also, dass die Relation $\sim_U$ auf V gegeben durch

$$v \sim_U v' \iff v - v' \in U$$

eine Äquivalenzrelation ist. Eine alternative Bezeichnung der Äquivalenzrelation ist $\equiv \mod U$. Das heißt wir schreiben

$$v \equiv v' \mod U \iff v - v' \in U.$$

Weiterhin, aus Satz 1.139 hat die Quotientenmenge V/U eine Gruppenstruktur mit der Addition $+ : V/U \times V/U \longrightarrow V/U$ gegeben durch

$$[v] + [v'] := [v + v'].$$

Wir erinnern noch, dass die Äquivalenzklasse eines Vektors $v \in V$ auch als $v + U$ bezeichnet werden kann. Diese Schreibweise soll betonen, dass

$$[v] = v + U = \{v + u \ : \ u \in U\}.$$

Definition 4.13. Sei V eine $\mathbb{K}$-Vektorraum. Eine Teilmenge $A \subseteq V$ ist ein **affiner Unterraum** $A = \emptyset$ oder wenn es ein Vektor v existiert und ein $\mathbb{K}$-Unterraum $U \subseteq_{\mathbb{K}} V$, sodass $A = v + U$.

Es ist einfach zu überprüfen, dass wenn A ein affiner Unterraum ist, dann ist der Unterraum U, dass ist es eindeutig ist. Deswegen kann man die **Dimension** des affines Raumes als $\dim A := \dim_{\mathbb{K}} U$ definieren.

Beispiele:

1. Jede Teilmenge mit ein Element ist ein 0-dimensionaler affiner Raum.
2. Es gibt ein einziger 2-dimensionaler affiner Raum in $\mathbb{R}^2$: $\mathbb{R}^2$ selber. Allgemein, es gibt ein einziger affiner Unterraum maximaler Dimension, weil es ein einziger linearer Unterraum maximaler Dimension gibt.
3. Für alle $a, b, c \in \mathbb{R}$ mit $(a, b) \neq (0, 0)$ ist die Menge $\{(x, y) \in \mathbb{R}^2 \quad : \quad ax + by = c\}$ ein affiner Unterraum von $\mathbb{R}^2$. Alle diese affine Unterräume sind die Geraden in $\mathbb{R}^2$.
4. Wir haben in Satz 2.10 gesehen, dass, weil $\mathcal{L}(A|\mathbf{0}) \subseteq \mathbb{K}^n$ ein $\mathbb{K}$-UVR ist, alle $\mathcal{L}(A|\mathbf{b})$ affine Unterräume sind. Wir werden sehen, dass alle affine Unterräume von $\mathbb{K}^n$ diese Form haben.
5. Wenn $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, dann ist für jeder $w \in W$ die Faser $f^{-1}(w)$ ein affiner Unterraum: Entweder ist $f^{-1}(w) = \emptyset$, oder es gibt $v \in f^{-1}(w)$, dann ist
$$f^{-1}(w) = v + \mathrm{Ker}(f).$$
6. Der Durchschnitt zweier affine Unterräume ist wieder ein affiner Raum - Übung.

Wir werden jetzt sehen, dass auf der Faktormenge V/U auf natürlicher Weise eine $\mathbb{K}$-Vektorraum Struktur definiert werden kann. Die Skalarmultiplikation $\cdot : \mathbb{K} \times V/U \longrightarrow V/U$ ist definiert durch

$$\lambda \cdot [v] := [\lambda \cdot v], \quad \forall \lambda \in \mathbb{K} \text{ und } [v] \in V/U.$$

Wir müssen noch überprüfen, dass $\cdot$ wohldefiniert ist. Das heißt, dass es unabhängig von der Wahl des Repräsentanten der Äquivalenzklasse ist. Seien also $v \equiv v' \bmod U$. Das heißt $[v] = [v']$. Wir wollen zeigen, dass $\lambda \cdot [v] = \lambda \cdot [v'] \ \ \forall \ \lambda \in \mathbb{K}$. Wir haben

$$\lambda \cdot [v] = \lambda \cdot [v'] \iff [\lambda v] = [\lambda v'] \iff \lambda v - \lambda v' \in U \iff \lambda(v - v') \in U.$$

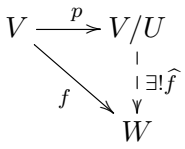
Die letzte Aussage ist wahr, weil aus $[v] = [v']$ folgt $v - v' \in U$, und weil U ein $\mathbb{K}$-UVR ist, folgt auch $\lambda(v - v') \in U, \ \forall \ \lambda \in \mathbb{K}$. Die Skalarmultiplikation auf V/U ist also wohldefiniert.

Satz 4.14. *Sei V ein $\mathbb{K}$-Vektorraum und $U \subseteq_{\mathbb{K}} V$ ein Untervektorraum.*

[26] 24.1.'24

1. *Die Quotientenmenge V/U zusammen mit den obigen Verknüpfungen ist ein $\mathbb{K}$-Vektorraum.*
2. *Die kanonische Projektion $p : V \longrightarrow V/U$ ist eine $\mathbb{K}$-lineare Abbildung mit* $\mathrm{Ker}(p) = U$.

3. [Universelle Eigenschaft] *Für jede andere* $\mathbb{K}$*-lineare Abbildung* $f : V \longrightarrow W$ *mit* $U \subseteq \operatorname{Ker} f$*, gibt es eine eindeutige* $\mathbb{K}$*-lineare Abbildung* $\widehat{f} : V/U \longrightarrow W$*, sodass folgendes Diagramm kommutiert:*



(d.h. $\widehat{f} \circ p = f$).

Des Weiteren:

(i) $\widehat{f}$ *ist surjektiv* $\Leftrightarrow$ f *ist surjektiv.*

(ii) $\widehat{f}$ *ist injektiv* $\Leftrightarrow \operatorname{Ker} f = U$.

Beweis-Skizze: Dass V/U ein $\mathbb{K}$-Vektorraum ist, dass p $\mathbb{K}$-linear ist, und dass $\operatorname{Ker}(p) = U$ folgt auf einfacher und direkter Weise aus den Definitionen. Der Rest folgt aus Satz 1.60. Q.E.D.

Definition 4.15. Sei V ein $\mathbb{K}$-Vektorraum und $U \subseteq V$ ein $\mathbb{K}$-UVR. Der **Quotientenvektorraum** von V nach U (oder auch **Quotientenraum** oder **Faktorraum**) ist der $\mathbb{K}$-Vektorraum V/U aus Satz 4.14

Korollar 4.16. Für jedes Homomorphismus von $\mathbb{K}$-Vektorräume $\varphi : V \longrightarrow W$ gilt folgenden Isomorphismus:

$$\operatorname{Bild} \varphi \simeq V/\operatorname{Ker} \varphi.$$

Beweis-Skizze: In Satz 4.14 nimmt man $U = \operatorname{Ker} \varphi$, $W = \operatorname{Bild} \varphi$ und als f die Einschränkung des Wertebereichs von φ, also $f : V \longrightarrow \operatorname{Bild} \varphi \subseteq W$, mit $f(v) = \varphi(v)$. Q.E.D.

Korollar 4.17. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $U \subseteq_{\mathbb{K}} V$ ein Unterraum. Dann ist der $\mathbb{K}$-Vektorraum V/U auch endlichdimensional mit

$$\dim_K(V/U) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} U.$$

Beweis-Skizze: Nach Satz 4.14 haben wir $\operatorname{Ker}(p) = U$, wobei $p : V \longrightarrow V/U$ die kanonische Projektion ist. Weil p surjektiv ist (also $\operatorname{Bild}(p) = U/V$), folgt die Aussage aus dem Dimensionssatz 4.2 für p.

Q.E.D.

Korollar 4.18. Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, B eine Basis eines $\mathbb{K}$-Unterraumes $U \subseteq_{\mathbb{K}} V$ und $p : V \longrightarrow V/U$ die kanonische Projektion. Sei $C \subseteq V$ eine Teilmenge mit $\#C = \#p(C)$. Es gilt

$$p(C) \text{ ist eine Basis von } V/U \iff C \cap B = \emptyset \text{ und } B \cup C \text{ ist eine Basis von } V.$$

Beweis-Skizze: Zuerst bezeichnen wir die Vektoren der zwei Mengen mit $B = \{v_1, \ldots, v_s\}$ und $C = \{w_1, \ldots, w_r\}$.

$\boxed{\Rightarrow}$ Wenn $B \cap C \neq \emptyset$, dann existiert ein j mit $w_j \in B$. Also $p(w_j) = [0]$ - ↯.

Um zu zeigen, dass $B \cup C$ eine Basis ist, reicht es, nach Korollar 4.17, zu zeigen, dass es linear unabhängig ist oder, dass es ein Erzeugendensystem ist. Wir zeigen hier, dass es ein Erzeugendensystem ist.

Sei $v \in V$. Weil $p(C)$ eine Basis von V/U ist, existieren $\mu_1, \dots, \mu_r \in \mathbb{K}$, sodass

$$[v] = \mu_1[w_1] + \cdots + \mu_r[w_r].$$

Weil $\mu_1[w_1]+\cdots+\mu_r[w_r] = [\mu_1 w_1+\cdots+\mu_r w_r]$, heißt das, dass $v-\mu_1 w_1-\cdots-\mu_r w_r \in U = \operatorname{Span}_{\mathbb{K}} B$. Daraus folgt, dass $v \in \operatorname{Span}_{\mathbb{K}}(B \cup C)$.

$\boxed{\Leftarrow}$ Aus Korollar 4.17 und $\#p(C) = \#C$ wissen wir schon, dass $\#p(C) = \dim_{\mathbb{K}} V/U$. Es reicht also wieder nur eine der zwei Axiome für Basen zu zeigen. Wir zeigen hier, dass $p(C) = \{[w_1], \dots, [w_r]\}$ linear unabhängig ist. Seien $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ so dass

$$\lambda_1[w_1] + \cdots + \lambda_r[w_r] = [0].$$

Es folgt also, dass es existieren $\mu_1, \dots, \mu_s \in \mathbb{K}$ so dass

$$\lambda_1 w_1 + \cdots + \lambda_r w_r = \mu_1 v_1 + \cdots + \mu_s v_s.$$

Weil $B \cap C = \emptyset$ und weil $B \cup C$ eine Basis folgt, dass alle Skalare Null sind. Q.E.D.

Wir werden hier unten einen zentralen Begriff der homologischen Algebra einführen. Für uns wird es nur eine Methode sein, Zusammenhänge zwischen mehrere Vektorräume und lineare Abbildungen kompakt darzustellen.

Definition 4.19. Sei $(V_i)_{i\in\mathbb{Z}}$ eine Familie von $\mathbb{K}$-Vektorräume. Für jedes Index $i \in \mathbb{Z}$ sei $d_i : V_i \longrightarrow V_{i-1}$ ein Homomorphismus von $\mathbb{K}$-Vektorräume. Wir haben also eine Kette von $\mathbb{K}$-linearen Abbildungen:

$$\cdots \xrightarrow{d_{i+2}} V_{i+1} \xrightarrow{d_{i+1}} V_i \xrightarrow{d_i} V_{i-1} \xrightarrow{d_{i-1}} \cdots$$

Wenn $d_i \circ d_{i+1} = 0$ für alle $i \in \mathbb{Z}$, dann sagen wir, dass eine solche Folge von Vektorräume, zusammen mit den Abbildungen d_i, ein **Kettenkomplex** ist. Die Bedingung $d_i \circ d_{i+1} = 0$ ist per Definition äquivalent zu $\operatorname{Bild} d_{i+1} \subseteq \operatorname{Ker} d_i$. Wir sagen, dass ein Kettenkomplex wie oben **exakt an der Stelle** i (oder in V_i) ist, wenn $\operatorname{Bild} d_{i+1} = \operatorname{Ker} d_i$. Eine **kurze exakte Folge** (oder kurze exakte Sequenz) ist ein Kettenkomplex der Form

$$\cdots \quad 0 \longrightarrow V_2 \xrightarrow{d_2} V_1 \xrightarrow{d_1} V_0 \longrightarrow 0 \quad \cdots \tag{4.4}$$

das exakt an jeder Stelle ist. Die "$\cdots$" heißen, dass $V_i = 0$ wenn $i < 0$ oder wenn $i > 2$; und werden dann immer weggelassen.

Bemerkung 4.20. In einer kurzen exakten Folge wie in (4.4) gilt immer:

$\cdots$ Die $\mathbb{K}$-lineare Abbildungen $d_3 : 0 \longrightarrow V_2$ und $d_0 : V_0 \longrightarrow 0$ können nur Nullabbildungen sein. Bei d_0 ist das klar, bei d_3 folgt es aus der $\mathbb{K}$-Linearität.

2. Exaktität an der Stelle 2 heißt also: $0 = \operatorname{Bild} 0 = \operatorname{Ker} d_2$, also d_2 ist injektiv.

0. Exaktität an der Stelle 0 heißt also: $\operatorname{Bild} d_1 = \operatorname{Ker} d_0 = V_0$, also d_1 ist surjektiv.

1. Exaktität an der Stelle 1 heißt also, nach Korollar 4.16, dass

$$V_0 \;=\; \operatorname{Bild} d_1 \;\simeq\; V_1/\operatorname{Ker} d_1 \simeq\; V_1/V_2.$$

Korollar 4.16 kann also als "Folgende kurze Sequenz ist exakt: $0 \longrightarrow \operatorname{Ker}\varphi \longrightarrow V \xrightarrow{\varphi} \operatorname{Bild}\varphi \longrightarrow 0$. " umformuliert werden.

Kettenkomplexe können auch mit homogene lineare Gleichungssysteme in Zusammenhang gebracht werden. Ein homogenes LGS mit n_0 Gleichungen und n_1 Unbekannten ist durch eine Matrix $A \in \mathrm{Mat}_{n_0,n_1}$ gegeben. Diese Matrix gibt dann eine Abbildung

$$\mathbb{K}^{n_1} \xrightarrow{A\cdot} \mathbb{K}^{n_0} \ .$$

Das System zu lösen, also die Lösungsmenge zu bestimmen, ist dann äquivalent mit, einen $\mathbb{K}$-Vektorraum U und eine $\mathbb{K}$-lineare Abbildung $i : U \longrightarrow \mathbb{K}^n$ zu finden, sodass folgender Kettenkomplex exakt in $\mathbb{K}^{n_1}$ ist.

$$U \xrightarrow{i} \mathbb{K}^{n_1} \xrightarrow{A\cdot} \mathbb{K}^{n_0} \ .$$

Bemerkung 4.21. Für zwei $\mathbb{K}$-Unterräume $U_1, U_2 \subseteq V$ gelten:

(i) $(U_1 \oplus U_2)/(U_1 \cap U_2) \simeq U_1 + U_2$

(ii) $(U_1 + U_2)/U_1 \simeq U_2/(U_1 \cap U_2)$

(iii) $U_1 \oplus U_2 = V \iff$ die kanonische Abbildung $U_2 \longrightarrow V/U_1$ ist ein Isomorphismus

Beweis-Skizze: Das ist eine gute Gelegenheit kurze Exakte folgen anzuwenden.

(i) Folgt aus der Folge (4.3)

(ii) Folgt aus folgender kurzen exakten Folge:

$$0 \longrightarrow U_1 \xrightarrow{g} U_1 + U_2 \xrightarrow{f} U_2/(U_1 \cap U_2) \longrightarrow 0$$

wobei $g(u_1) := u_1 + 0$, und das ist offensichtlich injektiv, und $f(u_1 + u_2) := [u_2]$. Vor der Surjektivität von f (die offensichtlich ist) müssen wir überprüfen, dass f wohl definiert[a] ist. Wenn $u_1 + u_2 = u_1' + u_2'$, dann haben wir $U_2 \ni u_2 - u_2' = u_1' - u_1 \in U_1$ und somit in $U_1 \cap U_2$.

(iii) Das folgt einfach aus Satz 4.9 und aus Punkt (ii).

Q.E.D.

[a] Der Grund ist: die Definition von f ist von der Darstellung eines Vektors in $U_1 + U_2$ als $u_1 + u_2$ abhängig; diese kann nicht eindeutig sein! Zum Beispiel, in $\mathbb{R}^3$, wenn $U_1 = \mathrm{Span}_{\mathbb{R}}\{(1,0,0),(0,1,1)\}$ und $U_2 = \mathrm{Span}_{\mathbb{R}}\{(0,1,1),(0,0,1)\}$ haben wir $(1,1,1) + (0,0,1) = (1,0,0) + (0,1,2)$. Also um die Wohldefiniertheit zu überprüfen, sollte man zeigen, dass $(0,0,1) \equiv (0,1,2) \bmod U_1 \cap U_2$. Und das stimmt auch, weil $(0,1,2) - (0,0,1) = (0,1,1) \in U_1 \cap U_2$.

[27] 29.1.'24

4.5 Der $\mathbb{K}$-Vektorraum der Homomorphismen

Wir haben in Abschnitt 1.2.7 gesehen, dass Abbildungen verknüpft werden können. Das funktioniert aber nur wenn der Wertebereich der einen Abbildung gleich mit dem Definitionsbereich der anderen

ist. Dann haben wir in Beispiele 3.2.4. gesehen, dass die Verknüpfung zweier linearen Abbildungen, wieder linear ist. In diesem Teil werden wir neue Operationen für $\mathbb{K}$-lineare Abbildungen einführen. Dieses Mal kommen diese nicht mehr von Operationen für allgemeine Abbildungen, sondern nutzen die Vektorraumstruktur des Wertebereichs aus. Diese werden analog zu den Operationen auf $\mathrm{Abb}(X, \mathbb{K})$ definiert. Alles funktioniert aber nur für fixierte Definitons- und Wertebereiche!

Sei X eine Menge und seien V und W zwei $\mathbb{K}$-Vektorräume. Wir definieren die Mengen:

$$\begin{aligned} \mathrm{Abb}(X, W) &:= \{f : X \longrightarrow W \mid f = \text{Abbildung}\} \quad \text{und} \\ \mathrm{Hom}_{\mathbb{K}}(V, W) &:= \{\varphi : V \longrightarrow W \mid \varphi = \mathbb{K}\text{-lineare Abbildung}\}. \end{aligned}$$

Auf beiden Mengen definieren wir eine $\mathbb{K}$-Vektorraumstruktur[3], völlig analog wie im Beispiel 3.1.11. für $\mathrm{Abb}(X, \mathbb{K})$. Das heißt, wir verwenden die Vektoraddition $+$ in W und die Multiplikation mit Skalaren $\cdot$ auf W verwenden. Seien f, g, entweder beide in $\mathrm{Hom}_{\mathbb{K}}(V, W)$ oder beide in $\mathrm{Abb}(X, W)$, aber nicht gemischt. Wir setzen für jedes $x \in X$, beziehungsweise $x \in V$:

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \quad \text{und} \\ (\lambda \cdot f)(x) &:= \lambda \cdot f(x), \quad \forall\ \lambda \in \mathbb{K}. \end{aligned}$$

Wenn $f, g \in \mathrm{Abb}(X, W)$, dann haben wir offensichtlich Abbildungen definiert: $f + g, \lambda \cdot f \in \mathrm{Abb}(X, W)$, für alle $\lambda \in \mathbb{K}$. Wenn $\varphi, \gamma \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, dann haben wir erstmals $\varphi + \gamma, \lambda \cdot \varphi \in \mathrm{Abb}(V, W)$, wir wollen aber, dass $\varphi + \gamma, \lambda \cdot \varphi \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, also dass diese auch $\mathbb{K}$-linear sind. Das ist einfach zu überprüfen: Für $v, w \in V$ und $\mu \in \mathbb{K}$ beliebig gelten:

$$\begin{aligned} (\varphi + \gamma)(v + w) &= \varphi(v + w) + \gamma(v + w) \\ &= \varphi(v) + \varphi(w) + \gamma(v) + \gamma(w) \\ &= \varphi(v) + \gamma(v) + \varphi(w) + \gamma(w) \\ &= \varphi(v + w) + \gamma(v + w). \end{aligned}$$

$$\begin{aligned} (\varphi + \gamma)(\mu \cdot v) &= \varphi(\mu \cdot v) + \gamma(\mu \cdot v) \\ &= \mu \cdot \varphi(v) + \mu \cdot \gamma(v) \\ &= \mu \cdot (\varphi(v) + \gamma(v)) \\ &= \mu \cdot (\varphi + \gamma)(v). \end{aligned}$$

Völlig analog zeigt man, dass für alle $\lambda \in \mathbb{K}$ die Abbildung $\lambda \cdot \varphi$ $\mathbb{K}$-linear ist. Also $\forall\ \varphi, \gamma \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ und $\forall\ \lambda \in \mathbb{K}$ gilt

$$\varphi + \gamma \in \mathrm{Hom}_{\mathbb{K}}(V, w), \quad \text{und} \quad \lambda \cdot \varphi \in \mathrm{Hom}_{\mathbb{K}}(V, w).$$

Satz 4.22. *Seien V, W zwei $\mathbb{K}$-Vektorräume. Das Tripel $(\mathrm{Hom}_{\mathbb{K}}(V, W), +, \cdot)$ ist ein $\mathbb{K}$-Vektorraum.*

Beweis-Skizze: $(\mathrm{Hom}_{\mathbb{K}}(V, W), +)$ ist eine abelsche Gruppe:

- **Das neutrale Element** ist die Nullabbildung $\mathbf{0} : V \longrightarrow W$, mit $\mathbf{0}(v) = 0_W\ \forall\ v \in V$.
- **Das inverse Element** von $\varphi \in \mathrm{Hom}_K(V, W)$ ist $-\varphi : V \longrightarrow W$ gegeben durch

$$(-\varphi)(v) := -\varphi(v),\ \forall\ v \in V.$$

[3] Das heißt, das wir eine Vektoraddition und eine Multiplikation mit Skalare definieren.

- **Die Assoziativität** und **die Kommutativität** folgen direkt aus den entsprechenden Eigenschaften der Vektoraddition $+$ auf W.

Die Axiome (VR 2.$\star$) für $\star = 1 \dots 4$ folgen direkt aus den Axiomen für den $\mathbb{K}$-Vektorraum W.
Q.E.D.

Wir erinnern hier, dass wenn $X' \subseteq X$ eine Teilmenge ist und $f : X \longrightarrow Y$ eine Abbildung ist, dann bezeichnet $f|_{X'} : X' \longrightarrow N$ die Einschränkung der Abbildung f auf dem Definitionsbereich X'.

Bemerkung 4.23. In der Definition der Addition und Skalarmultiplikation auf $\mathrm{Hom}_{\mathbb{K}}(V, W)$ haben wir die $\mathbb{K}$-Vektorraumstruktur von V nicht verwendet. Wir bekommen also vom Beweis von Satz 4.22, dass auch $(\mathrm{Abb}(X, W), +, \cdot)$ ein $\mathbb{K}$-Vektorraum ist.

Satz 4.24. *Seien V und W zwei $\mathbb{K}$-Vektorräume und sei $B = \{v_1, \dots, v_n\}$ eine Basis von V. Dann ist die Abbildung*

$$\Phi : \mathrm{Hom}_{\mathbb{K}}(V, W) \longrightarrow \mathrm{Abb}(B, W) \qquad \text{mit } \Phi(f) := f|_B$$

ein $\mathbb{K}$-Vektorraum Isomorphismus. Anders gesagt, ist jede $\mathbb{K}$-lineare Abbildung durch die Bilder der Vektoren einer Basis eindeutig bestimmt.

Die Zusatzaufgabe 5 auf Blatt 13 (Januar 2024) zeigt, dass auch die andere Implikation in Satz 4.24 gilt: das heißt, B ist eine Basis genau dann, wenn Φ ein Isomorphismus ist. Wir werden aber nur die Formulierung hier oben später anwenden.

Beweis-Skizze: Die $\mathbb{K}$-Linearität folgt aus der Definition:

$$(f + g)|_B = f|_B + g|_B \quad \text{und } (\lambda f)|_B = \lambda(f|_B).$$

Wir müssen noch zeigen, dass Φ bijektiv ist, oder, äquivalent, dass Φ eine Inverse besitzt. Wir wählen den zweiten Weg. Für jede Abbildung $\alpha \in \mathrm{Abb}(B, W)$ definieren wir $f_\alpha : V \longrightarrow W$ durch

$$f_\alpha(v) := \lambda_1 \alpha(v_1) + \dots + \lambda_n \alpha(v_n), \quad \forall\, v \in V,$$

wobei $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ die eindeutige Darstellung von v als lineare Kombination der Elementen der Basis B ist (cf. Bemerkung 3.27). Wegen der Eindeutigkeit der λ_i ist f_α tatsächlich eine Abbildung von V nach W. Wir müssen noch überprüfen, dass $f_\alpha \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, also dass f_α $\mathbb{K}$-linear ist. Sei ein $v \in V$ wie oben, beliebig, und sei auch $w \in V$ ein beliebiger Vektor mit eindeutiger Darstellung durch B als $w = \mu_1 v_1 + \dots + \mu_n v_n$. Dann ist $v + w = (\lambda_1 + \mu_1)v_1 + \dots + (\lambda_n + \mu_n)v_n$, also

$$f_\alpha(v + w) = f_\alpha(v) + f_\alpha(w).$$

Genauso haben wir für jedes $\lambda \in \mathbb{K}$, dass $\lambda v = (\lambda\lambda_1)v_1 + \dots + (\lambda\lambda_n)v_n$, also

$$f_\alpha(\lambda v) = \lambda f_\alpha(v).$$

Also $f_\alpha \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Da offensichtlich $f_\alpha|_B = \alpha$ gilt, haben wir die inverse Abbildung von Φ definiert, und somit gezeigt, dass Φ ein Isomorphismus von $\mathbb{K}$-Vektorräume ist. Q.E.D.

Satz 4.24 zusammen mit Bemerkung 3.27 wird es uns erlauben, Matrizen $\mathbb{K}$-linearer Abbildungen zuzuordnen, wann immer V und W beide endlich-dimensional sind. Als nächstes werden wir jedoch untersuchen, wie sich Injektivität und Surjektivität unter $\Phi : \mathrm{Hom}_{\mathbb{K}}(V, W) \longrightarrow \mathrm{Abb}(B, W)$ verhalten.

Lemma 4.25. *Sei $f : V \longrightarrow W$ eine $\mathbb{K}$-lineare Abbildung. Folgende Aussagen sind äquivalent.*

(i) *f ist injektiv.*

(ii) *Für jede linear unabhängige Teilmenge $\{v_1, \ldots, v_r\} \subseteq V$ ist $\{f(v_1), \ldots, f(v_n)\}$ linear unabhängig.*

(iii) *Es existiert eine Basis B von V, sodass $f(B) \subseteq W$ linear unabhängig ist.*

Beweis-Skizze: (i)⇒(ii) Nehmen wir an, dass $\lambda_1 f(v_1) + \cdots + \lambda_r f(v_r) = 0$. Es folgt aus der $\mathbb{K}$-Linearität von f, dass $f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0$. Weil f injektiv ist, haben wir also $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$, und somit, weil $v_1, \ldots, v_r$ linear unabhängig sind, haben wir $\lambda_1 = \cdots = \lambda_r = 0$.
(ii)⇒(iii) Es ist trivial per Definition 3.22 der linearen Unabhängigkeit.
(iii)⇒(i) Sei $v \in \operatorname{Ker} f$. Weil B eine Basis ist, existieren $v_1, \ldots, v_n \in B$ und eindeutig bestimmte $\lambda_i \in \mathbb{K}$ sodass $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Es gilt dann

$$0 = f(v) = f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n),$$

also, aus der linearen Unabhängigkeit von $f(B)$ folgt $\lambda_i = 0$, $\forall\, i$, und somit $v = 0$. Q.E.D.

Bemerkung 4.26. Aus Satz 4.1 (iii) folgt es gleich, dass für Surjektivität eine ähnliche Beschreibung gilt. Das heißt, für eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ sind folgende Aussagen äquivalent.

(i) f ist surjektiv.

(ii) Für jedes Erzeugenden System $S \subseteq V$ ist auch $f(S)$ ein Erzeugenden System von W.

(iii) Es existiert eine Basis B von V, sodass $\operatorname{Span}_{\mathbb{K}} f(B) = W$.

Lemma 4.25 und Bemerkung 4.26 geben uns:

Korollar 4.27. Für eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ sind folgende Aussagen äquivalent.

(i) f ist ein Isomorphismus.

(ii) Für jede Basis B von V ist auch $f(B)$ eine Basis von W.

(iii) Es existiert eine Basis B von V, sodass $f(B)$ eine Basis von W ist.

4.5.1 Endomorphismen sind $\mathbb{K}$-Algebren

Wenn V, W, U drei $\mathbb{K}$-Vektorräume sind, und wenn $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$ und $g \in \operatorname{Hom}_K(W, U)$, dann ist $g \circ f \in \operatorname{Hom}_K(V, U)$. In dem Sonderfall $V = W = U$ ist diese Verknüpfung eine innere Verknüpfung auf $\operatorname{End}_{\mathbb{K}}(V) := \operatorname{Hom}_{\mathbb{K}}(V, V)$

$$\circ : \operatorname{End}_{\mathbb{K}}(V) \times \operatorname{End}_{\mathbb{K}}(V) \longrightarrow \operatorname{End}_{\mathbb{K}}(V).$$

Das Tripel $(\operatorname{End}_{\mathbb{K}}(V), +, \circ)$ ist ein nicht-kommutativer Ring. Die additive abelsche Gruppe ist dieselbe wie im Fall der $\mathbb{K}$-Vektorraumstruktur. Die Assoziativität der Komposition gilt immer, und das

multiplikative Neutralelement ist die identische Abbildung id_V. Zu überprüfen bleiben also nur die Distributivitätsgesetze aus dem Ring-Axiom (R3):

$$\begin{aligned}((f+g)\circ h)(v) &= (f+g)(h(v)) = f(h(v)) + g(h(v)) = (f\circ h)(v) + (g\circ g)(v), \quad \forall\ v \in V,\\ (h\circ(f+g))(v) &= h((f+g)(v)) = f(h(v)) + g(h(v)) = (f\circ h)(v) + (g\circ g)(v), \quad \forall\ v \in V.\end{aligned}$$

Eigentlich kann man (und soll man) die Ring- und $\mathbb{K}$-Vektorraumstrukturen auf $\mathrm{End}_{\mathbb{K}}(V)$ gleichzeitig betrachten, indem man $\mathrm{End}_{\mathbb{K}}(V)$ als $\mathbb{K}$-Algebra zusammenfasst. Ein (nicht unbedingt kommutativer) Ring R ist eine **$\mathbb{K}$-Algebra** wenn es ein injektiver Ringhomomorphismus $u : \mathbb{K} \hookrightarrow R$ gibt, so dass $u(\mathbb{K}) \subseteq Z(R) = \{a \in R \ : \ ab = ba\ \forall\ b \in R\}$. In unserem Fall ist die Abbildung $u : \mathbb{K} \hookrightarrow \mathrm{End}_{\mathbb{K}}(V)$ gegeben durch $\lambda \mapsto \lambda \cdot \mathrm{id}_V$.

Kapitel 5

Die Matrix eines Homomorphismus bezüglich fixierten Basen

Das Ziel ist, lineare Abbildungen zwischen endlich-dimensionale $\mathbb{K}$-Vektorräumen durch Matrizen zu beschreiben. Das bedeutet, eine Korrespondenz zwischen $\mathrm{Hom}_{\mathbb{K}}(V,W)$ und $\mathrm{Mat}_{m,n}(\mathbb{K})$ herzustellen, und dann zu verstehen wie viel Information[1] unter dieser Korrespondenz erhalten bleibt.

Wir sehen zunächst kurz, wie diese Korrespondenz grob funktioniert, und dann kümmern wir uns um die Eigenschaften und Einzelheiten. Wir fixieren:

$$V \text{ mit geordneter Basis } B = \{b_1, \dots, b_n\}, \qquad W \text{ mit geordneter Basis } C = \{c_1, \dots, c_m\}.$$

Um eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ zu beschreiben, reicht es, die Bilder der Basis-Elemente zu kennen: $f(b_1), \dots, f(b_n)$. Diese sind Vektoren in W und können auf eindeutige Weise als lineare Kombination der Basis-Elemente in C dargestellt werden. Das heißt, für jeden Homomorphismus $f : V \longrightarrow W$ existieren eindeutig bestimmte $\alpha_{ji} \in \mathbb{K}$ für $j = 1, \dots, n$ und $i = 1, \dots, m$, sodass

$$\begin{aligned} f(b_1) &= \alpha_{1,1}\, c_1 + \alpha_{2,1}\, c_2 + \cdots + \alpha_{m,1}\, c_m \\ f(b_2) &= \alpha_{1,2}\, c_1 + \alpha_{2,2}\, c_2 + \cdots + \alpha_{m,2}\, c_m \\ &\vdots \\ f(b_n) &= \alpha_{1,n}\, c_1 + \alpha_{2,n}\, c_2 + \cdots + \alpha_{m,n}\, c_m. \end{aligned}$$

Die $\alpha_{i,j}$ sind von der Wahl der Basen B und C abhängig. Wir sammeln diese in einer Matrix, die wir mit $M_C^B(f) \in \mathrm{Mat}_{m,n}(\mathbb{K})$ bezeichnen, und wir nennen diese die **zugeordnete Matrix** (oder Darstellungsmatrix) von f bezüglich der *geordneten*[2] Basen B und C. Um das ganze mit Abbildungen der Form $\mathbf{x} \mapsto A \cdot \mathbf{x}$ kompatibel zu machen, werden wir die Koordinaten von $f(b_i)$ bezüglich der Basis C in der Spalte i eintragen. Das ist sicher nur eine Konvention, aber diese ist ziemlich weit verbreitet. Das heißt,

$$M_C^B(f) = \begin{pmatrix} \alpha_{1,1} & \alpha_{12,} & \dots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \dots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \dots & \alpha_{m,n} \end{pmatrix} \in \mathrm{Mat}_{\dim W, \dim V}(\mathbb{K}).$$

[1] Zum Beispiel: sind Operationen damit verträglich? Kann man Injektivität und Surjektivität von der Matrix ablesen?
[2] Siehe Definition 5.3 weiter unten.

5.1 Einschränkung auf $\mathbb{K}^n$

Der folgende Satz besagt, dass bis auf Isomorphie $\mathbb{K}^n$ die einzigen endlich-dimensionalen $\mathbb{K}$-Vektorräume sind.

Satz 5.1. *Alle $\mathbb{K}$-Vektorräume von Dimension $n \in \mathbb{N}$ sind untereinander isomorph. Insbesondere, wenn $\dim_{\mathbb{K}} V = n \in \mathbb{N}$, dann $V \simeq \mathbb{K}^n$.*

Beweis-Skizze: Nach Satz 3.30 und der Definition 3.40 existieren Basen $B = \{b_1, \ldots, b_n\}$ von V und $\{c_1, \ldots, c_n\}$ von W. Gemäß Satz 4.24 und Korollar 4.27 gibt die Zuordnung $b_i \mapsto c_i$ ein Isomorphismus zwischen V und W. Q.E.D.

Wir haben in Beispiel 3.2.7. gesehen, dass eine Matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$ eine Abbildung $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ definiert, durch

$$f_A(\mathbf{x}) := A \cdot \mathbf{x}.$$

Um die Transponierten Notation zu vermeiden, werden wir, wann immer es günstig ist, die Vektoren in $\mathbb{K}^n$ als Spaltenmatrizen betrachten, damit $A \cdot \mathbf{x}$ definiert ist.

Satz 5.2. *Die Zuordnung $A \mapsto f_A$ definiert Isomorphismus $f_\bullet : \mathrm{Mat}_{m,n}(\mathbb{K}) \longrightarrow \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$. Die Inverse davon ist die Abbildung $M : \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m) \longrightarrow \mathrm{Mat}_{m,n}(\mathbb{K})$, definiert durch:*

$$M(f) := \left(\begin{array}{c|c|c} & & \\ f(e_1) & \ldots & f(e_n) \\ & & \end{array} \right).$$

Beweis-Skizze: Die $\mathbb{K}$-Linearität von $f_\bullet$ folgt direkt aus den Eigenschaften der Operationen für Matrizen:

$$f_{\lambda A + \mu B} = \lambda f_A + \mu f_B, \text{ weil } (\lambda A + \mu B) \cdot \mathbf{x} = \lambda(A \cdot x) + \mu(B \cdot \mathbf{x}), \quad \text{für alle } \mathbf{x} \in \mathbb{K}^n.$$

Das die zwei invers zueinander sind, gilt weil einerseits, wenn $A = (a_{ij})$, dann gilt

$$f_A(e_i) = A \cdot e_i = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{n,i} \end{pmatrix} = \text{die } i\text{-te Spalte von } A,$$

anderseits, für jedes $\mathbf{x} = (x_1 \ldots x_n)^T \in \mathbb{K}^n$ haben wir

$$f(\mathbf{x}) = f(x_1 e_1 + \cdots + x_n e_n) = x_1 f(e_1) + \cdots + x_n f(e_n) = \left(\begin{array}{c|c|c} & & \\ f(e_1) & \ldots & f(e_n) \\ & & \end{array} \right) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = f_{(M(f))}(\mathbf{x}).$$

Q.E.D.

Die gesamte Idee hinter der zugeordneten Matrix einer $\mathbb{K}$-linearen Abbildung bezüglich gegebener Basen ist es, Satz 5.2, Satz 5.1 und Korollar 4.27 zusammenzubringen. Dafür müssen wir eine Basis und eine Reihenfolge für die Elemente der Basis festlegen, was äquivalent zur Wahl eines Isomorphismus nach $\mathbb{K}^n$ ist.

Definition 5.3. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum von Dimension n. Eine **geordnete Basis** von V ist ein n-Tupel[3] $v_1, \ldots, v_n \in V^n$ mit der Eigenschaft, dass $\{v_1, \ldots, v_n\}$ eine Basis ist.

Eine geordnete Basis schreibt man also als eine Liste: $v_1, \ldots, v_n$, obwohl es rein formell ein n-Tupel aus V^n ist.

Bemerkung 5.4. Wir haben aus Satz 5.1, dass jeder endlichdimensionaler $\mathbb{K}$-Vektorraum V isomorph zu $\mathbb{K}^{\dim_{\mathbb{K}} V}$ ist. Korollar 4.27 gibt uns sogar die Beschreibung aller Isomorphismen. Es gilt nämlich, dass wenn V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n \in \mathbb{N}$ ist, dann gibt jede geordnete Basis $C = c_1, \ldots, c_n$ von $\mathbb{K}^n$ eine Bijektion zwischen den Mengen

Aus Satz 5.1 folgt, dass jeder endlichdimensionale $\mathbb{K}$-Vektorraum V isomorph zu $\mathbb{K}^{\dim_{\mathbb{K}} V}$ ist. Korollar 4.27 gibt uns sogar eine Beschreibung aller Isomorphismen. Genauer gesagt, wenn V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n \in \mathbb{N}$ ist, dann haben wir für jede fixierte geordnete Basis $C = c_1, \ldots, c_n$ von $\mathbb{K}^n$ eine Bijektion zwischen den Mengen:

$$\begin{aligned} \{\text{Geordnete Basen } B \text{ von } V\} &\xrightarrow{\sim} \mathrm{Iso}_{\mathbb{K}}(V, \mathbb{K}^n) \\ B &\longmapsto \varphi_C^B : V \longrightarrow \mathbb{K}^n. \end{aligned}$$

Wenn $B = b_1, \ldots, b_n$, dann bezeichnet $\varphi_C^B : V \longrightarrow \mathbb{K}^n$ den durch $\varphi_C^B(b_i) = c_i$ eindeutig bestimmten Isomorphismus.

Wir bezeichnen mit $\mathcal{E} = e_1, \ldots, e_n$ die kanonische geordnete Basis von $\mathbb{K}^n$. Seien V und W zwei endlich dimensionale $\mathbb{K}$-Vektorräume mit geordneten Basen $B = b_1, \ldots, b_n$, beziehungsweise $C = c_1, \ldots, c_m$. Wir betrachten die Isomorphismen:

$$\begin{aligned} \varphi_B^{\mathcal{E}} : \ \mathbb{K}^n &\xrightarrow{\sim} V & \qquad \varphi_{\mathcal{E}}^C : W &\xrightarrow{\sim} \mathbb{K}^m \\ e_i &\longmapsto b_i & \qquad c_j &\longmapsto e_j \end{aligned}$$

Für festgelegte B und C wie oben haben wir zu jeder $\mathbb{K}$-linearen Abbildung $f : V \to W$ eine eindeutige entsprechende $\mathbb{K}$-lineare Abbildung zwischen den Standardräumen: $\varphi_{\mathcal{E}}^C \circ f \circ \varphi_B^{\mathcal{E}} : \mathbb{K}^n \to \mathbb{K}^m$. Wir haben also das kommutative Diagramm:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle \varphi_B^{\mathcal{E}}}\uparrow\wr & & \wr\downarrow{\scriptstyle \varphi_{\mathcal{E}}^C} \\ \mathbb{K}^n & \xrightarrow{\varphi_{\mathcal{E}}^C \circ f \circ \varphi_B^{\mathcal{E}}} & \mathbb{K}^m. \end{array} \tag{5.1}$$

Die senkrechten Pfeile sind Isomorphismen und entsprechen der Wahl der geordnete Basen.

Bemerkung 5.5. Unter den obigen Voraussetzungen ist die zugeordnete Matrix einer $\mathbb{K}$-linearen Abbildung f bezüglich der geordneten Basen B und C die Matrix

$$M_C^B(f) = M(\varphi_{\mathcal{E}}^C \circ f \circ \varphi_B^{\mathcal{E}}) \in \mathrm{Mat}_{m,n}(\mathbb{K}),$$

[3] Wenn $V = \mathbb{K}^n$, dann haben wir ein n-Tupel von n-Tupeln, was etwas verwirrend sein kann. Es ist günstiger in diesem Fall eine geordnete Basis als eine $n \times n$-Matrix zu sehen, in der die Spalten die Vektoren sind. Wenn es aber nicht Sinn macht eine geordnete Basis als Matrix anzugeben, werden wir die Klammern für das n-Tupel, das eine geordnete Basis ist, völlig weg lassen: also $v_1, \ldots, v_n$ bezeichnet eine geordnete Basis und $(v_1, \ldots, v_n)$ wird einen Element aus X^n für irgendwelches X bezeichnen.

Wobei $M : \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m) \longrightarrow \mathrm{Mat}_{m,n}(\mathbb{K})$ die Abbildung aus Satz 5.2 ist. Das heißt nach Satz 5.2, dass die rote Abbildung durch die links-Multiplikation mit der Darstellungsmatrix gegeben ist. Also

$$(\varphi_{\mathcal{E}}^{C} \circ f \circ \varphi_{B}^{\mathcal{E}})(\mathbf{x}) = \big(M_C^B(f)\big) \cdot \mathbf{x}, \qquad \text{für alle } \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n. \tag{5.2}$$

Die Bemerkung betont, wie die Untersuchung der $\mathbb{K}$-linearen Abbildungen auf den Fall von Standardräumen eingeschränkt werden kann. Sie warnt auch davor, direkten "$f = f_{M_C^B(f)}$" einzusetzen! Es gibt zwar eine bijektive Korrespondenz wie in (5.2), jedoch gilt eine echte Gleichheit nur, wenn $V = \mathbb{K}^n$, $W = \mathbb{K}^m$, $B = \mathcal{E}$ und $C = \mathcal{E}$.

Korollar 5.6. Seien V und W endlich dimensionale $\mathbb{K}$-Vektorräume mit geordneten Basen B beziehungsweise C. Die Abbildung

$$M_C^B : \mathrm{Hom}_{\mathbb{K}}(V, W) \xrightarrow{\sim} \mathrm{Mat}_{\dim W, \dim V}(\mathbb{K}),$$

die jedem $\mathbb{K}$-Homomorphismus die Darstellungsmatrix bezüglich der angegebenen geordneten Basen B und C zuordnet, ist ein Isomorphismus von $\mathbb{K}$-Vektorräumen.

Beweis-Skizze: Das folgt aus Satz 5.2 und der Bemerkung, dass links und rechst Verknüpfen mit den Isomorphismen $\varphi_B^{\mathcal{E}} : \mathbb{K}^n \to V$ und $\varphi_{\mathcal{E}}^{C} : W \longrightarrow \mathbb{K}^m$ in (5.3) einen Isomorphismus definiert:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbb{K}}(V, W) & \xrightarrow{\sim} & \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m) \\ f & \longmapsto & \varphi_{\mathcal{E}}^{C} \circ f \circ \varphi_B^{\mathcal{E}} \end{array}.$$

Die Abbildung ist offensichtlich linear, mit Inverser $g \mapsto \varphi_C^{\mathcal{E}} \circ g \circ \varphi_{\mathcal{E}}^{B}$. Q.E.D.

[28] 31.1.'24

Korollar 5.7. Für endlich-dimensionale $\mathbb{K}$-Vektorräume V und W haben wir:

$$\dim_{\mathbb{K}} \mathrm{Hom}_{\mathbb{K}}(V, W) = \dim_{\mathbb{K}} V \cdot \dim_{\mathbb{K}} W.$$

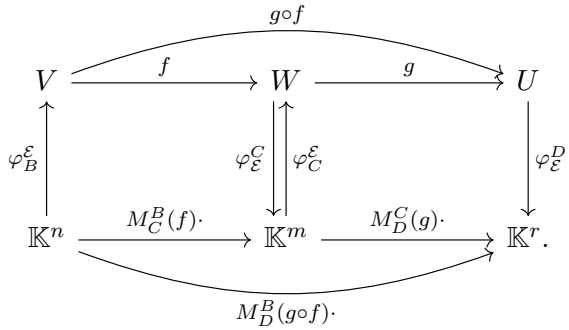
Die Zuordnung $M_C^B : \mathrm{Hom}_{\mathbb{K}}(V, W) \longrightarrow \mathrm{Mat}_{m,n}(\mathbb{K})$ ist ein $\mathbb{K}$-Isomorphismus ist, also bijektiv, verträglich mit der Addition und mit der Multiplikation mit Skalaren. Was noch fehlt, ist die Verträglichkeit mit der Verknüpfung von Abbildungen. Dies gilt ebenfalls:

Satz 5.8. *Seien V, W, U drei $\mathbb{K}$-Vektorräume endlicher Dimension n, m, beziehungsweise r. Seien B, C, D geordnete Basen von V, W, beziehungsweise U. Wenn f und g $\mathbb{K}$-lineare Abbildungen mit $V \xrightarrow{f} W \xrightarrow{g} U$ sind, dann gilt*

$$M_D^B(g \circ f) = M_D^C(g) \cdot M_C^B(f).$$

Beweis-Skizze: Um auf Matrixmultiplikation zu kommen, wenden wir Diagramm (5.3) zwei Mal

an:



(5.3)

Wir haben also aus Bemerkung 5.5, dass die links-Multiplikation mit der $r \times n$ Matrix $M_D^B(g \circ f)$ gleich mit der Verknüpfung der links-Multiplikation mit $M_C^B(f)$ und der links-Multiplikation mit $M_D^C(g)$ ist. Das ist genau das, was zu zeigen war. Q.E.D.

Beispiele:

1. Für die Nullabbildung $0 : V \longrightarrow W$, mit $0(v) = 0_W, \forall\, v \in V$, haben wir $M_C^B(0) = 0_{m\times n}$ für jede Wahl der geordneten Basen.
2. Die Identität $\mathrm{id}_V : V \longrightarrow V$ hat $M_B^B(\mathrm{id}_V) = I_n$ für jede Wahl einer Basis B.
3. Die Abbildung $\mathrm{ev}_1 : \mathbb{K}[x]_{\leqslant 2} \longrightarrow \mathbb{K}$ gegeben durch die Evaluation des Polynoms in 1, also $ev_1(P(x)) := P(1)$ ist $\mathbb{K}$-linear. Die Matrix bezüglich der Basen $B = 1, x, x^2$ und $\mathcal{E} = 1$ ist $\begin{pmatrix}1 & 1 & 1\end{pmatrix}$. Für die Abbildung ev_0 bekommen wir $\begin{pmatrix}1, 0, 0\end{pmatrix}$. Allgemein bekommen wir

$$M_{\mathcal{E}}^B(\mathrm{ev}_a) = \begin{pmatrix}1 & a & a^2\end{pmatrix}.$$

5.2 Quadratische Matrizen und Endomorphismen

Das Korollar 5.6 ist besonders interessant in dem Fall, wenn $V = W$ und $B = C$. Dann ist es nicht mehr notwendig, B doppelt zu schreiben. Die **zugeordnete Matrix** des Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ bezüglich der Basis B bezeichnen wir mit

$$M^B(f) := M_B^B(f) \in \mathrm{Mat}_{\dim V}(\mathbb{K}).$$

In diesem Fall haben sowohl $\mathrm{End}_{\mathbb{K}}(V)$ als auch $\mathrm{Mat}_n(\mathbb{K})$ eine zusätzliche innere Verknüpfung (die Komposition von Abbildungen, beziehungsweise die Matrixmultiplikation). Zusammen mit der üblichen Addition bilden diese eine Ringstruktur. Satz 5.8 hat dann folgende Konsequenz.

Satz 5.9. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und B eine geordnete Basis davon. Die Abbildung*

$$M^B : \mathrm{End}_{\mathbb{K}}(V) \longrightarrow \mathrm{Mat}_n(\mathbb{K})$$

ist ein Ringisomorphismus. Insbesondere ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ invertierbar genau dann, wenn die Matrix $M^B(f)$ invertierbar ist.

Beweis-Skizze: Wir müssen Bijektivität und die Axiome aus Definition 9.2 überprüfen.
Die Bijektivität und (RHom1), das ist die Verträglichkeit mit der Addition, gelten nach Korollar 5.6.
Das Axiom (RHom2) ist dann $M^B(g \circ f) = M^B(g) \cdot M^B(f)$, und das ist ein Sonderfall von Satz 5.8.
Das Axiom (RHom3) sagt, dass die 1 auf 1 abgebildet werden muss. In unserem Fall ist die 1 in $\mathrm{End}_{\mathbb{K}}(V)$ die Identität $\mathrm{id}_V : V \longrightarrow V$ und die 1 in $\mathrm{Mat}_n(\mathbb{K})$ ist die Identitätsmatrix $I_n = (\delta_{ij})_{i,j=1\ldots n}$. Es gilt offensichtlich dass

$$M^B(\mathrm{id}_V) = I_n.$$

Der letzte Teil ist allgemein eine Folgerung von Ringisomorphismen:

$$f \circ g = \mathrm{id}_V \iff M^B(f) \cdot M^B(g) = I_n.$$

Q.E.D.

Wir haben also bewiesen, dass für einen n-dimensionalen $\mathbb{K}$-Vektorraum V, mit $n < \infty$, die Isomorphie

$$\mathrm{End}_{\mathbb{K}}(V) \quad \simeq \quad \mathrm{Mat}_n(\mathbb{K})$$

nicht nur als $\mathbb{K}$-Vektorräume gilt, sondern auch als (nicht kommutative) Ringe. Dies lässt sich in einem Begriff zusammenfassen: die beiden sind als $\mathbb{K}$-Algebren isomorph. Allerdings werden wir das hier nicht anwenden. Der "insbesondere"-Teil im Satz 5.9 liefert uns auch einen Gruppenisomorphismus

$$\mathrm{Aut}_{\mathbb{K}}(V) \simeq \mathrm{GL}_n(\mathbb{K}).$$

Dies ist jedoch kein Isomorphismus von Ringen oder $\mathbb{K}$-Algebren mehr, da die beiden Strukturen nicht abgeschlossen bezüglich der Addition sind. Zum Beispiel $\mathrm{id}_V, -\mathrm{id}_V \in \mathrm{Aut}_{\mathbb{K}}(V)$, aber deren Summe ist die Nullabbildung und somit nicht invertierbar.

5.3 Basiswechsel

Die Matrixbeschreibung eines $\mathbb{K}$-Homomorphismus hängt offensichtlich von der Wahl einer Basis ab. Eine Änderung (oder eine Neuwahl) der Basis nennen wir **Basiswechsel**. Der kann wiederum durch eine Matrix beschrieben werden. In der Fachliteratur wird das oft unabhängig von der Darstellungsmatrix einer $\mathbb{K}$-linearen Abbildung definiert, und erst später wird es in Zusammenhang mit Darstellungsmatrizen gebracht. Hier führen wir jedoch kein neues Verfahren ein, um Matrizen aus zwei geordneten Basen zu erzeugen. Die **Basiswechselmatrix** (auch als **Transformationsmatrix** oder **Übergangsmatrix** bezeichnet) definieren wir als einen Sonderfall der Darstellungsmatrix. Satz 5.13 ergibt sich dann als eine einfache Folgerung von Satz 5.8. Die Aussage dieses Satzes lässt sich somit leichter merken. Hier ist zunächst die Definition.

Definition 5.10. Seien $B = v_1, \ldots, v_n$ und $B' = v'_1, \ldots, v'_n$ zwei geordnete Basen eines endlichdimensionalen $\mathbb{K}$-Vektorraumes V. Die **Basiswechselmatrix** von B' nach B ist die Matrix

$$M_B^{B'} := M_B^{B'}(\mathrm{id}_V).$$

Das heißt, *die* ***alte Basis*** B' *wir durch die* ***neue Basis*** B *ausgedrückt.* Explizit geschrieben, gilt nach der Definition von Darstellungsmatrix: wenn $v_i' = \alpha_{1i}v_1 + \cdots + \alpha_{ni}v_n$, dann ist

$$M_B^{B'} = (\alpha_{ij})_{i,j=1,\ldots,n}.$$

Beispiel 5.11. Seien $B = (1,-1),(2,0)$ und $B' = (1,2),(1,1)$ zwei geordnete Basen von $V = \mathbb{R}^2$. Dann haben wir

$$\begin{aligned}(1,2) &= -2\cdot(1,-1) + \frac{3}{2}\cdot(2,0),\\ (1,1) &= -1\cdot(1,-1) + 1\cdot(2,0),\end{aligned}$$

also

$$M_B^{B'} = \begin{pmatrix} -2 & -1 \\ \frac{3}{2} & 1 \end{pmatrix}.$$

Wenn wir das umgekehrt machen, dann haben wir

$$\begin{aligned}(1,-1) &= -2\cdot(1,2) + 3\cdot(1,1),\\ (2,0) &= -2\cdot(1,2) + 4\cdot(1,1),\end{aligned}$$

also

$$M_{B'}^{B} = \begin{pmatrix} -2 & -2 \\ 3 & 4 \end{pmatrix}.$$

Wir merken hier gleich, dass

$$\begin{pmatrix} -2 & -1 \\ \frac{3}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} -2 & -2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Das ist kein Zufall (cf. Bemerkung 5.12). Wir machen erstmal mit dem Beispiel weiter und beobachten was mit den Koordinaten bezüglich der zwei geordneten Basen passiert. Sei $v = (0,1) \in \mathbb{R}^2$. Die Koordinaten von v bezüglich B sind $(-1, \frac{1}{2})_B$ und die Koordinaten von v bezüglich B' sind $(1,-1)_{B'}$ weil

$$\begin{aligned}(0,1) &= -1\cdot(1,-1) + \frac{1}{2}\cdot(2,0),\\ (0,1) &= 1\cdot(1,2) + (-1)\cdot(1,1).\end{aligned}$$

Um die Koordinaten bezüglich B' von den Koordinaten bezüglich B zu bekommen, müssen wir mit der Basiswechselmatrix von B nach B' multiplizieren. In diesem Fall

$$M_{B'}^{B} \cdot \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix}_B = \begin{pmatrix} -2 & -2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix}_B = \begin{pmatrix} 1 \\ -1 \end{pmatrix}_{B'}.$$

Man kann sich diese Formel merken, oder man kann sie immer von Diagramm (5.3) herleiten.

Bemerkung 5.12. Jede Basiswechselmatrix $M_{B'}^{B}$ ist invertierbar mit Inverse $M_B^{B'}$. Also

$$\left(M_{B'}^{B}\right)^{-1} = M_B^{B'}.$$

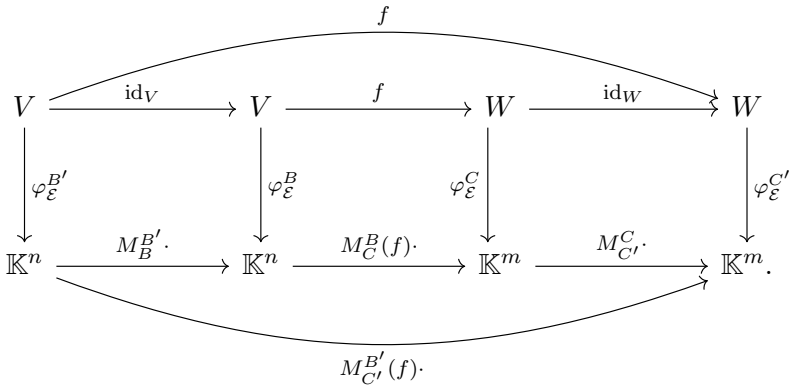
Das gilt weil nach Satz 5.8 ist

$$M_{B'}^{B} \cdot M_B^{B'} = M_{B'}^{B}(\mathrm{id}_V) \cdot M_B^{B'}(\mathrm{id}_V) = M_B^{B}(\mathrm{id}_V) = I_n.$$

Satz 5.13. *Seien B und B' geordnete Basen von V, und C und C' geordnete Basen von W. Für jedes $f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$ gilt*

$$M_{C'}^{B'}(f) = M_{C'}^{C} \cdot M_{C}^{B}(f) \cdot M_{B}^{B'}.$$

Beweis-Skizze: Sei $n = \dim_{\mathbb{K}} V$ und $m = \dim_{\mathbb{K}} W$. Wir wenden Satz 5.8 für folgende Verknüpfung:



Q.E.D.

Korollar 5.14. Wenn $V = W$, $B = C$, $B' = C'$, und $T = M_B^{B'}$ dann haben wir für jede $f \in \mathrm{End}_{\mathbb{K}}(V)$

$$M^{B'}(f) = T^{-1} \cdot M^{B}(f) \cdot T.$$

Wir werden uns mit dieser Relation noch viel beschäftigen, aber erst in Kapitel 10.

5.4 Zeilen- und Spaltenrang einer Matrix

Satz 5.1 und Korollar 5.6 zusammen sagen uns , dass jede $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$ als eine Abbildung $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, mit $f_A(\mathbf{x}) = A \cdot \mathbf{x}$ interpretiert werden kann. Dies gibt uns einen weiteren Grund, Matrizen zu untersuchen und besser zu verstehen.

Definition 5.15. Sei $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ mit Zeilen und Spalten:

$$A = \left(\begin{array}{c} z_1 \\ \hline z_2 \\ \hline \vdots \\ \hline z_m \end{array}\right) = \left(\begin{array}{c|c|c|c} & & & \\ s_1 & s_2 & \dots & s_n \\ & & & \end{array}\right).$$

Der **Zeilenraum** von A, beziehungsweise der **Spaltenraum** von A, ist

$$\mathrm{ZRaum}(\mathbf{A}) := \mathrm{Span}_{\mathbb{K}}\{z_1,\dots,z_m\} \subseteq \mathbb{K}^n \quad \text{beziehungsweise} \quad \mathrm{SRaum}(\mathbf{A}) := \mathrm{Span}_{\mathbb{K}}\{s_1,\dots,s_n\} \subseteq \mathbb{K}^m.$$

Der **Zeilenrang** von A, beziehungsweise der **Spaltenrang** von A, ist

$$\mathrm{Rang_Z}(\mathbf{A}) := \dim_{\mathbb{K}} \mathrm{Span}_{\mathbb{K}}\{z_1,\dots,z_m\} \quad \text{beziehungsweise} \quad \mathrm{Rang_S}(\mathbf{A}) := \dim_{\mathbb{K}} \mathrm{Span}_{\mathbb{K}}\{s_1,\dots,s_n\}$$

Insbesondere gilt $\mathrm{Rang_Z}(A) \leqslant \min\{m,n\}$ und $\mathrm{Rang_S}(A) \leqslant \min\{m,n\}$.

Beispiele:

1. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$. Dann gilt

$$\mathrm{Rang}_{\mathrm{Z}}(A) = \begin{cases} 0 & \text{wenn } A = 0 \\ 1 & \text{wenn } ad - bc = 0 \\ 2 & \text{wenn } ad - bc \neq 0 \end{cases}$$

2. $\mathrm{Rang}_{\mathrm{Z}}(I_n) = \mathrm{Rang}_{\mathrm{S}}(I_n) = n$.

3. Wenn A in Zeilenstufenform ist, dann ist $\mathrm{Rang}_{\mathrm{Z}}(A) = {}^{\mathrm{ZS}}\mathrm{Rang}(A)$, nämlich die Anzahl der Stufen.

Satz 5.16. *Sei $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ eine Matrix und Υ eine Zeilenumformung. Dann gilt*

$$\mathrm{ZRaum}(A) = \mathrm{ZRaum}(\Upsilon(A)).$$

Insbesondere gilt für alle Matrizen

$${}^{\mathrm{ZS}}\mathrm{Rang}(A) = \mathrm{Rang}_{\mathrm{Z}}(A).$$

Beweis-Skizze: Es reicht die erste Aussage für elementare Zeilenumformungen zu zeigen. Wir bezeichnen die Zeilen von A beziehungsweise von $\Upsilon(A)$ mit

$$A = \left(\begin{array}{c} z_1 \\ \hline z_2 \\ \hline \vdots \\ \hline z_m \end{array}\right), \qquad \Upsilon(A) = \left(\begin{array}{c} y_1 \\ \hline y_2 \\ \hline \vdots \\ \hline y_m \end{array}\right).$$

$\boxed{\Upsilon = \Upsilon_{k\leftrightarrow l}}$ Dann gilt $\{z_1, \ldots, z_m\} = \{y_1, \ldots, y_m\}$.

$\boxed{\Upsilon = \Upsilon_{k\to\lambda\cdot k}}$ (mit $\lambda \neq 0$). Dann gilt $\{y_1, \ldots, y_m\} = \{z_1, \ldots, z_{k-1}, \lambda z_k, z_{k+1}, \ldots, z_m\}$. Also

$$\begin{aligned} v \in \mathrm{Span}_{\mathbb{K}}\{y_1, \ldots, y_m\} &\iff v = \textstyle\sum_{i=1}^m \mu_i y_i = \mu_1 z_1 + \cdots + (\mu_k\lambda) z_k + \cdots + z_m \\ &\iff v \in \mathrm{Span}_{\mathbb{K}}\{z_1, \ldots, z_m\}, \end{aligned}$$

wobei für die letzte Äquivalenz $\lambda \neq 0$ wichtig ist.

$\boxed{\Upsilon = \Upsilon_{k\to k+\lambda\cdot l}}$ Dann gilt $\{y_1, \ldots, y_m\} = \{z_1, \ldots, z_{k-1}, z_k + \lambda z_l, z_{k+1}, \ldots, z_m\}$. Also

$$\begin{aligned} v \in \mathrm{Span}_{\mathbb{K}}\{y_1, \ldots, y_m\} \iff v &= \mu_1 y_1 + \cdots + \mu_m y_m \\ &= \mu_1 z_1 + \cdots + \mu_k(z_k + \lambda z_l) + \cdots + z_m \\ &= \mu_1 z_1 + \cdots + \mu_k z_k + \cdots + (\lambda + \mu_l) z_l + \cdots + z_m \\ \iff v &\in \mathrm{Span}_{\mathbb{K}}\{z_1, \ldots, z_m\}. \end{aligned}$$

Q.E.D.

Analog zur Zeilenstufenform kann man eine Spaltenstufenform definieren und damit auch einen Spaltenstufenrang. Mit einem analogen Beweis zeigt man dann, dass der Spaltenstufenrang immer gleich zum Spaltenrang ist. In Teil 5.5 werden wir zeigen, dass alle diese Versionen von Rang immer dasselbe Ergebnis liefern. Es ist wichtig zu bemerken, dass, obwohl $\text{Rang}_{\text{Z}}(A) = \text{Rang}_{\text{S}}(A)$ immer gilt, $\text{ZRaum}(A) \neq \text{SRaum}(A)$ sein kann; zum Beispiel wenn $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

Bemerkung 5.17. Seien $v_1, \ldots, v_m \in \mathbb{K}^n$, und sei A die $m \times n$ Matrix mit Zeilen v_i. Dann gilt

(i) $\{v_1, \ldots, v_m\}$ ist genau dann linear unabhängig, wenn $\text{Rang}_{\text{Z}}(A) = m$.

(ii) $\{v_1, \ldots, v_m\}$ ist genau dann ein Erzeugendensystem, wenn $\text{Rang}_{\text{Z}}(A) = n$.

(iii) $\{v_1, \ldots, v_m\}$ ist genau dann eine Basis, wenn $m = n$ und A invertierbar ist.

5.5 Rang einer Matrix

Definition 5.18. Der **Rang** einer $\mathbb{K}$-linearen Abbildung f ist die Dimension des Bildes von f:

$$\text{Rang}(f) := \dim_{\mathbb{K}} \text{Bild}(f).$$

Wenn $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ durch links-Multiplikation mit A definiert ist, dann gilt

$$\text{Bild}\, f_A = \text{Span}_{\mathbb{K}}\{f_A(e_1), \ldots, f_A(e_n)\} = \text{Span}\{S_1(A), \ldots, S_n(A)\},$$

wobei $S_i(A)$ die i. Spalte von A ist. Wir haben also

$$\text{Rang}(f_A) = \text{Rang}_{\text{S}}(A).$$

Satz 5.19. *Sei $A \in \text{Mat}_{m,n}(\mathbb{K})$ eine Matrix. Dann gilt $\text{Rang}_{\text{Z}}(A) = \text{Rang}_{\text{S}}(A)$.*

Beweis-Skizze: Wir betrachten $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$. Dann gilt $\text{Bild}(f_A) = \text{SRaum}(A)$. Wir haben aus Satz 4.2 dass

$$\dim_{\mathbb{K}} \mathbb{K}^n - \dim_{\mathbb{K}} \text{Ker}(f_A) = \dim_{\mathbb{K}} \text{Bild}(f_A) = \text{Rang}(f_A) = \text{Rang}_{\text{S}}(A).$$

Es gilt auch $\text{Ker}(f_A) = \mathcal{L}(A \,|\, \mathbf{0})$. Aus dem Gaußschen Algorithmus folgt

$$\dim_K \mathcal{L}(A|\mathbf{0}) = n - {}^{\text{ZS}}\text{Rang}(A).$$

Aus Satz 5.16 haben wir ${}^{\text{ZS}}\text{Rang}(A) = \text{Rang}_{\text{Z}}(A)$, also

$$\text{Rang}_{\text{S}}(A) = n - (n - \text{Rang}_{\text{Z}}(A)) = \text{Rang}_{\text{Z}}(A).$$

Q.E.D.

Definition 5.20. Der **Rang** einer Matrix $A \in \text{Mat}_{m,n}(\mathbb{K})$ ist der Rang der Abbildung $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$.

Bemerkung 5.21. Für $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$ gilt dann

$$\mathrm{Rang}(A) = {}^{\mathrm{ZS}}\mathrm{Rang}(A) = \mathrm{Rang}_{\mathrm{Z}}(A) = \mathrm{Rang}_{\mathrm{S}}(A) = \mathrm{Rang}(f_A).$$

Wir haben festgestellt, dass man durch Anwendung des Gaußschen Algorithmus auf die Matrix mit den Zeilenvektoren $v_1, \ldots, v_r \in \mathbb{K}^n$ eine Basis für den von diesen Vektoren erzeugten Unterraum $\mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\}$ finden kann. Es war einfach zu zeigen, dass dieser Ansatz funktioniert und ermöglicht, eine "einfachere" Basis zu finden. Mit Hilfe von Satz 5.19 können wir sogar eine Basis $B \subseteq \{v_1, \ldots, v_n\}$ für $\mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\}$ finden, indem wir die Vektoren $v_1, \ldots, v_r$ als Spalten einer Matrix A betrachten und den Gaußschen Algorithmus anwenden (siehe Kapitel 6 Algorithmus 1. indem die leere Menge ergänzt wird).

Korollar 5.22. Seien $v_1, \ldots, v_r \in \mathbb{K}^n$ Vektoren, sei $U = \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\}$ der davon erzeugte $\mathbb{K}$-Unterraum und sei

$$A = \left(\begin{array}{c|c|c|c} & & & \\ v_1 & v_2 & \ldots & v_r \\ & & & \end{array}\right) \in \mathrm{Mat}_{n,r}(\mathbb{K})$$

die Matrix mit den Vektoren $v_1, \ldots, v_r$ als Spalten. Sei $Y = \Upsilon(A \,|\, I_n)$ die RZSF der mit der Einheitsmatrix I_n ergänzten Matrix

$$\left(\begin{array}{c|c} & \\ A & I_n \\ & \end{array}\right) \in \mathrm{Mat}_{n,r+n}(\mathbb{K}).$$

k Diese Matrix hat Rang n weil der Spaltenraum ganz $\mathbb{K}^n$ ist; seien $\mathit{st}_1, \ldots, \mathit{st}_n$ die Stufenindizes von Y. Es gelten

(i) Die Menge $B = \{v_i \;:\; \exists\, j \text{ mit } i = \mathit{st}_j \leqslant r\}$ ist eine Basis von U.

(ii) Die Menge $B \cup \{e_i \;:\; \exists\, j \text{ mit } \mathit{st}_j - r = i\}$ ist eine Ergänzung der Basis B von U zu einer Basis von V.

Beweis-Skizze: Es reicht Punkt (i) zu zeigen, und wir können uns nur auf den ersten r Spalten konzentrieren, also auf der RZSF von A. Sei St die Menge der Stufenindizes in der RZSF von A. Weil Zeilenumformungen die Lösungsmenge von $\mathcal{L}(A|\mathbf{0})$ nicht ändern, haben wir, dass aus $\sum_{i \in St} \lambda_i v_i = \mathbf{0}$ folgt dass $(\lambda'_1, \ldots, \lambda'_r) \in \mathcal{L}(A|\mathbf{0})$, wobei

$$\lambda'_i = \begin{cases} \lambda_i & \text{wenn } i \in St \\ 0 & \text{sonst} \end{cases}$$

Aus der RZSF folgt also $\lambda_i = 0$ für alle $i \in St$. Das heißt, dass $B = \{v_i \;:\; i \in St\}$ linear unabhängig ist. Aus Satz 5.19 haben wir

$$|St| = {}^{\mathrm{ZS}}\mathrm{Rang}(A) = \mathrm{Rang}_{\mathrm{S}}(A) = \dim_K \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\},$$

wir haben also, dass B eine Basis ist. Q.E.D.

Kapitel 6

Verfahren um alles mögliche in $\mathbb{K}^n$ zu berechnen

Wir werden uns in diesem Teil nur um den Standardraum über einem Körper $\mathbb{K}$ kümmern. Nach Satz 5.1 ist alles was wir hier machen in jedem endlich-dimensionalen $\mathbb{K}$-Vektorraum übertragbar. Insbesondere, wollen wir einige Methoden sehen die uns erlauben konkret Basen zu finden. Wir wollen folgende Fragen beantworten.

1. Gegeben $T = \{v_1, \ldots, v_r\} \subseteq \mathbb{K}^n$ und erzeugenden System $S = \{w_1, \ldots, w_n\}$ von $\mathbb{K}^n$, finde eine Basis von $\mathrm{Span}_{\mathbb{K}} T$ und man ergänze diese mit Vektoren aus $\{w_1, \ldots, w_n\}$ zu einer Basis von $\mathbb{K}^n$.
2. Gegeben $U = \mathrm{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\}$, finde eine "schöne"[1] Basis von U.
3. Gegeben $U \subseteq_{\mathbb{K}} V$, man finde eine Basis von V/U.
4. Gegeben B =Basis von $\mathbb{K}^n$, C = Basis von $\mathbb{K}^m$, und $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, Bestimme $M_C^B(f)$.
5. Gegeben $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$, man finde Basen von $\mathrm{Ker}\, f_A$ und von $\mathrm{Bild}\, f_A$, wobei $f_A : \mathbb{K}^n \xrightarrow{A\cdot} \mathbb{K}^m$.
6. Wie bestimmt man den Rang einer linearen Abbildung?
7. Gegeben $U, W \subseteq \mathbb{K}^n$, man finde eine Basis von $U \cap W$.

Der Schlüssel ist immer das Gaußsche Eliminationsverfahren. Es bleibt nur zu verstehen an welcher Matrix man das anwenden muss, und wie man das Ergebnis liest.

[1] "schön" = viele Nullen.

Algorithmus 1. Basisergänzung

Eingabe: $T = \{v_1, \ldots, v_r\}$ und $S = \{w_{r+1}, \ldots, w_m\}$ Teilmengen von $\mathbb{K}^n$.

Schritt 1: Setze diese Vektoren als Spalten einer Matrix

$$A = \left(\begin{array}{c|c|c|c|c|c} v_1 & \ldots & v_r & w_{r+1} & \ldots & w_m \end{array}\right).$$

Schritt 2: Berechne eine ZSF von A.

Schritt 3: Die Stufen sind die Indizes $i_1 < \cdots < i_s \leqslant r < i_{s+1} < \cdots < i_n$.

Ausgabe: $v_{i_1}, \ldots, v_{i_s}, w_{i_{s+1}}, \ldots, w_{i_n}$ ist eine Basis von $\mathbb{K}^n$.

Kommentar:

- Einfacher gesagt: die Stufen zeigen die Positionen der Vektoren die man nehmen soll.
- Der Algorithmus hängt von der Reihenfolge der Angabe. Das heißt, die erst eingetragene Vektoren werden bevorzugt.
- Wenn $T = \emptyset$, dann findet man eine Basis die im erzeugenden System S enthalten ist.
- Wenn $S = \emptyset$, dann findet man eine Basis von $\operatorname{Span}_{\mathbb{K}} T$.

Algorithmus 2. Die beste Basis

Eingabe: $v_1, \ldots, v_r \in \mathbb{K}^n$

Schritt 1:Setze die Vektoren als Zeilen einer Matrix

$$A = \left(\begin{array}{c} v_1 \\ \hline v_2 \\ \hline \vdots \\ \hline v_r \end{array}\right).$$

Schritt 2: Berechne die $\mathrm{RZ}(A)$.

Ausgabe: Die nicht-Null Zeilen von $\mathrm{RZ}(A)$ sind die schöne Basis.

Kommentar:

- Das hat Vorteile wenn man mit dieser Basis weiter rechnen will.
- Man bekommt aber keine Teilmenge von $v_1, \ldots, v_r$.
- Wenn $\operatorname{Span}_{\mathbb{K}} v_1, \ldots, v_r = \mathbb{K}^n$, dann bekommt man die kanonische Basis.

Algorithmus 3. Basis von Quotientenraum

Eingabe: $U = \operatorname{Span}_{\mathbb{K}} v_1, \ldots, v_r$.

Schritt 1: Finde eine Basis B_1 von U und ergänze diese zu einer Basis B von $\mathbb{K}^n$:

$$u_1, \ldots, u_s, b_{s+1}, \ldots, b_n,$$

mit $u_i \in U$ und $w_i \notin U$. Die Prozedur dafür ist oben in Algorithmus 1. gegeben.

Ausgabe: $\widehat{b}_{s+1}, \ldots, \widehat{b}_n$ ist eine Basis von $\mathbb{K}^n/U$.

Algorithmus 4. Darstellungsmatrix

Eingabe: $b_1, \ldots, b_n$ Basis von $\mathbb{K}^n$, $c_1 \ldots, c_m$ Basis von $\mathbb{K}^m$.

Schritt 1: Setze die Vektoren c_i, und $f(b_j)$ als Spalten einer Matrix

$$\big(C \mid f(B)\big) = \left(\begin{array}{c|c|c|c|c|c} & & & & & \\ c_1 & \ldots & c_m & f(b_1) & \ldots & f(b_n) \\ & & & & & \end{array}\right).$$

Schritt 2: Berechne $\mathrm{RZ}(C \mid f(B)) = (I_m \mid M)$.

Ausgabe: Ausgabe $M_C^B(f) = M$.

Kommentar:

- Was wir machen ist das $LGS(C|f(b_i))$ für jedes $i = 1, \ldots, n$ gleichzeitig zu lösen.
- Das funktioniert bestimmt, weil C eine Basis ist, also das erste Block in $\mathrm{RZ}(C \mid f(B))$ ist immer I_m.
- Achtung! Man bekommt nicht $f(\mathbf{x}) = M_C^B(f) \cdot \mathbf{x}$. Es funktioniert so nur wenn $\mathbf{x}$ die Koordinaten bezüglich B sind, und was man bekommt sind die Koordinaten bezüglich C.

Algorithmus 5. Basen von Bild und Ker

Eingabe: $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$.

Berechne: die $\mathrm{RZ}(A)$ mit Stufen $i_1 < \cdots < i_r$.

Ausgabe Dimensionen: $\dim_{\mathbb{K}} \operatorname{Ker} f_A = n - r \qquad \dim_{\mathbb{K}} \operatorname{Bild} f_A = r.$

Ausgabe Ker: Für die Basis des Kerns, bestimme die Lösungsmenge von $LGS(A|0)$. (siehe VL-14, Minute 15:00-25:00) für ein Beispiel.)

Ausgabe Bild: Die Spalten von A mit Indizes $i_1, \ldots, i_r$.

Kommentar:

- Für die Basis von Bild f es ist wichtig die Spalten von A und **nicht** die von $\mathrm{RZ}(A)$ zu nehmen.

Algorithmus 6. Rang

Eingabe: Eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$

Schritt 1: Finde eine Basis B von V.

Schritt 2: Finde eine Basis C von W.

Schritt 3: Bestimme $M_C^B(f)$.

Schritt 4: Berechne die $\mathrm{RZ}(M_C^B(f))$

Ausgabe: Rang $f = \sharp$ Stufen.

Algorithmus 7. Schnitt

Eingabe: $U = \mathrm{Span}_{\mathbb{K}}\{u_1, \ldots, u_r\} \subseteq \mathbb{K}^n$ und $W = \mathrm{Span}_{\mathbb{K}}\{w_1, \ldots, w_s\} \subseteq \mathbb{K}^n$.

Schritt 1: Setze die u_i und die $-w_j$ als Spalten einer Matrix:

$$A = \big(U \mid -W\big) = \left(\begin{array}{c|c|c|c|c|c} u_1 & \ldots & u_r & w_1 & \ldots & w_s \end{array}\right).$$

Schritt 2: Finde eine Basis des zugeordneten homogenen $LGS(A \mid 0)$: $\ell_1, \ldots, \ell_t \in \mathbb{K}^{r+s}$

Ausgabe: Die Spalten von

$$\left(\begin{array}{c|c|c|c|c|c} u_1 & \ldots & u_r & 0 & \ldots & 0 \end{array}\right) \cdot \left(\begin{array}{c|c|c} \ell_1 & \ldots & \ell_t \end{array}\right).$$

Kommentar

- $u \in U \cap W$ wenn es $\lambda_1, \ldots, \lambda_r$ und $\mu_1, \ldots, \mu_s$ existieren mit

  $$u = \lambda_1 u_1 + \cdots + \lambda_r u_r = \mu_1 w_1 + \cdots + \mu_s w_s.$$

  Das heißt, wir suchen alle Lösungen von $LGS\big(U| -W| \, \mathbf{0}\big)$, und brauchen dann nur eine der zwei linearen Kombinationen.

- Wenn man aus theoretischen Gründen die schlussfolgern kann, dass $\dim_{\mathbb{K}} U \cap W = 1$, dann kann es sich lohnen das lineare Gleichungssystem "per Hand" zu lösen. Es reicht eine einzige nicht triviale Lösung in diesem Fall zu finden.

Kapitel 7

Determinanten

Jeder quadratischen $n \times n$ Matrix mit Einträgen aus einem Ring wird ein Element des Ringes, den wir **Determinante der Matrix** nennen, zugeordnet. Diese Zuordnung kann man durch verschiedene Formeln geben. Alle Formeln werden kompliziert wenn n groß wird. Direkt zu zeigen, dass zwei solche Formeln die selbe Funktion definieren ist auch schwer. Glücklicher Weise hat die Determinante sehr gute Eigenschaften, und man kann zeigen, dass drei dieser Eigenschaften die Determinante eindeutig bestimmen. So kann man auch zeigen, dass die verschiedene Formeln tatsächlich die selbe Funktion definieren. Wir werden mit einer der Formeln anfangen, und die "gute" Eigenschaften für diese Formel finden.

7.1 Naive Einführung

Sei R ein Ring und $n \in \mathbb{N}_{>0}$. Determinanten sind Abbildungen $\det : \mathrm{Mat}_n(R) \longrightarrow R$ die multilinear, alternierend und normiert sind. Wir werden sehen (Satz 7.42 und Satz **??**), dass es immer genau eine Abbildung mit diesen Eigenschaften gibt. In diesem Teil werden wir aber erstmals erklären was genau multilinear heißt und wozu können Determinanten dienen.

7.1.1 Multilineare Abbildungen

Wir fangen mit bilineare Abbildungen an. Sei jetzt $\mathbb{K}$ ein Körper und seien V, U, W drei $\mathbb{K}$-Vektorräume. Eine **bilineare Abbildung** von $V \times U$ nach W ist eine Abbildung $\psi(\cdot,\cdot) : V \times U \longrightarrow W$ mit der Eigenschaft:

$$\begin{array}{llll} \forall v \in V & \text{ist die Abbildung} & \psi(v,\cdot) : U \longrightarrow W & \mathbb{K}\text{-linear, und} \\ \forall u \in U & \text{ist die Abbildung} & \psi(\cdot,u) : V \longrightarrow W & \mathbb{K}\text{-linear} \end{array}$$

Das kann auch konkreter formuliert werden als: für alle $v_i \in V, u_i \in U$, und $\lambda_i, \mu_i \in \mathbb{K}$ gilt

$$\begin{array}{rcl} \psi(\lambda_1 v_1 + \lambda_2 v_2, \mu_1 u_1 + \mu_2 u_2) & = & \lambda_1\mu_1 \cdot \psi(v_1,u_1) + \lambda_1\mu_2 \cdot \psi(v_1,u_2) \\ & = & \lambda_2\mu_1 \cdot \psi(v_2,u_1) + \lambda_2\mu_2 \cdot \psi(v_2,u_2) \end{array}$$

Wir werden uns in Kapitel 11 mit bilineare Abbildungen mehr beschäftigen. Wir geben hier nur einige Beispiele um die Intuition hinter multilineare Abbildungen zu entwickeln.

Beispiele:

1. Jede quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ definiert eine bilineare Abbildung $\psi_A : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}$ durch
$$\psi_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \cdot A \cdot \mathbf{y}, \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{K}^n,$$
wobei wir $\mathbf{x}$ und $\mathbf{y}$ als Spalten zusammenfassen. Um ganz konkret zu werden: wenn $n = 3$, $A = I_3$ und $\mathbb{K}{=}\mathbb{R}$ dann haben wir $\psi_{I_3} : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ gegeben durch
$$\psi_{I_3}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1y_1 + x_2y_2 + x_3y_3.$$
Wir werden diese bilineare Abbildung später als Skalarprodukt[1] kennen.

2. Seien V, W zwei $\mathbb{K}$-Vektorräume. Dann hat $\mathrm{Hom}_{\mathbb{K}}(V, W)$ auch eine $\mathbb{K}$-Vektorraum Struktur. Wir definieren $\langle \cdot, \cdot \rangle : V \times \mathrm{Hom}_{\mathbb{K}}(V, W) \longrightarrow W$ durch
$$\langle v, f \rangle := f(v), \quad \forall v \in V, f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$$
Aus der Definition von $\mathbb{K}$-lineare Abbildung, und aus der Definition der Vektorraum Operationen auf $\mathrm{Hom}_{\mathbb{K}}(V, W)$ folgt es direkt, dass diese eine bilineare Abbildung ist.

3. Auf $\mathbb{K}^2 \times \mathbb{K}^2$ definieren wir $d : \mathbb{K}^2 \times \mathbb{K}^2 \longrightarrow \mathbb{K}$ durch
$$d\big((a, b), (c, d)\big) = ad - bc, \quad \forall\, (a, b), (c, d) \in \mathbb{K}^2.$$
Man kann direkt überprüfen, dass d bilinear ist. Oder man kann auch bemerken, dass $d = \psi_M$ für $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, und die bilinearität folgt aus Beispiel 1. hier oben. Dieses Beispiel kann man auch so interpretieren: Man identifiziert $\mathbb{K}^2 \times \mathbb{K}^2$ mit $\mathrm{Mat}_2(\mathbb{K})$ durch $\big((a, b), (c, d)\big) \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, und bekommt also eine Abbildung $d : \mathrm{Mat}_2(\mathbb{K}) \longrightarrow \mathbb{K}$ die *bilinear in den Zeilen* ist. Weiterhin, wenn $(a, b) = (c, d)$, also wenn $a = c$ und $b = d$, dann hat man
$$d\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ad - bc = ab - ba = 0.$$
Diese Eigenschaft, dass wenn zwei der Argumenten gleich sind, dann ordnet die bilineare (multilineare) Abbildung 0 zu, heißt, dass die Abbildung *alternierend* ist. Letztens, ist es einfach zu sehen, dass $d(I_2) = 1$, und das heißt, dass die Abbildung normiert ist. Was wir hier sehen ist also nichts anderes als die Determinante für $2 \times\ 2$ Matrizen.

7.1.2 Anwendungen von 2×2 Determinanten

Für 2×2 Matrizen ist die Determinante durch
$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

[1] Nicht zu verwechseln mit der Multiplikation mit Skalaren!

definiert. Wir schauen uns hier konkret an, was uns dieser Wert über die Matrix und über zugeordnete Strukturen sagen kann.

Invertierbarkeit. Eine Übung aus der linearen Algebra 1 (und auch aus der Zentralübung 1 im SoSe21) zeigt, dass

$$A \in \mathrm{Mat}_2(\mathbb{K}) \text{ ist invertierbar } \iff \det A \neq 0.$$

Eine einfache Rechnung zeigt auch, dass in diesem Fall

$$A^{-1} = \frac{1}{\det A} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Das zeigt auch, dass die Determinante die Invertierbarkeit der Matrizen mit Koeffizienten in einem Ring R bestimmt:

$$A \in \mathrm{Mat}_2(R) \text{ ist invertierbar } \iff \det A \text{invertierbar}^{2} \text{ in } R \text{ ist.}$$

Lösungen linearer Gleichungssysteme. Wenn $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(R)$ eine invertierbare Matrix ist, und $\beta_1, \beta_2 \in R$, dann hat das LGS $A \cdot \mathbf{x} = \mathbf{b}$ die Lösungen

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \frac{1}{\det A} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} d\beta_1 - b\beta_2 \\ -c\beta_1 + a\beta_2 \end{pmatrix}.$$

Wir merken jetzt, dass

$$\det \begin{pmatrix} a & \beta_1 \\ c & \beta_2 \end{pmatrix} = -c\beta_1 + a\beta_2 \qquad \text{und} \qquad -\det \begin{pmatrix} b & \beta_1 \\ d & \beta_2 \end{pmatrix} = d\beta_1 - b\beta_2.$$

Die Lösungen des LGS $A\mathbf{x} = \mathbf{b}$ können also durch die Determinanten der 2×2 Teilmatrizen der erweiterten Koeffizienten Matrix ausgedrückt werden. Wenn also

$$(A|\mathbf{b}) = \begin{pmatrix} a & b & \beta_1 \\ c & d & \beta_2 \end{pmatrix} \qquad B_1 = \begin{pmatrix} b & \beta_1 \\ d & \beta_2 \end{pmatrix} \qquad B_2 = \begin{pmatrix} a & \beta_1 \\ c & \beta_2 \end{pmatrix}$$

Dann hat die eindeutige Lösung die Koordinaten

$$x_1 = \frac{-\det B_1}{\det A} \qquad x_2 = \frac{\det B2}{\det A}.$$

Flächeninhalt und orientiertes Volumen. Um die geometrische Bedeutung der Determinante zu betonen, werden wir am Anfang Matrizen mit Einträgen aus $\mathbb{R}$ betrachten. Für jede $n \times n$ Matrix A, sei $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ der zugeordnete Endomorphismus

$$f_A(x_1, \ldots, x_n) = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Wenn $n = 1$, dann ist $A = (a)$ und die Abbildung $f_A : \mathbb{R} \longrightarrow \mathbb{R}$ ist die Multiplikation mit a. Das heißt, dass die kanonische Basis $\{1\}$ auf $\{a\}$ abgebildet wird. Insbesondere, die "Länge" des Vektors 1 wird

mit $|a|$ multipliziert. Wenn $a > 0$, dann "zeigt" $f(a)$ in derselben Richtung, und wenn $a < 0$, zeigt es in der Gegenrichtung.
Sei $n = 2$ und

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Die kanonische Basis $\{e_1, e_2\} = \{(1,0),(0,1)\}$ von $\mathbb{R}^2$ wird also auf $\{(a,c),(b,d)\}$ abgebildet. Wir werden jetzt sehen, wie sich der Flächeninhalt des Quadrates mit Ecken $\{(0,0),(1,0),(0,1),(1,1)\}$ (den wir zu 1 setzten) durch die Abbildung f_A ändert. Eine einfache Rechnung zeigt, dass der Flächeninhalt des Parallelogrammes mit Ecken $\{(0,0),(a,c),(b,d),(a+b,c+d)\}$ gleich mit $|ad-bc|$ ist. Wir werden die Determinante von A als

$$\begin{aligned} \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= ad - bc \\ &= a \cdot \det(d) - c \cdot \det(b) \end{aligned}$$

definieren. Das heißt, dass $\det(A) = 0 \Leftrightarrow (a,c) = t \cdot (b,d) \Leftrightarrow$ der Flächeninhalt Null ist. Weiterhin, der Vorzeichen kontrolliert ob die im "Uhrzeigersinn" Reihenfolge von $f(e_1)$ und $f(e_2)$ sich ändert oder nicht. Das heißt ob sich die *Orientierung* des Parallelogrammes ändert. Das gilt auch allgemein für $\mathbb{R}^n$, indem man Länge und Flächeninhalt mit Volumen ersetzt.

[1] 12.4.'21

7.2 Definition, erste Eigenschaften, Existenz, Eindeutigkeit über $\mathbb{K}$

Sei R ein kommutativer Ring mit 1 und $n \in \mathbb{N}_{>0}$. Wir werden Determinanten axiomatisch definieren. Das garantiert nicht die Existenz, sondern diese muss bewiesen werden. Das werden wir machen indem wir "eine" Determinante induktiv definieren (durch die Laplace Entwicklung), und dann zeigen, dass diese die Axiome erfüllt (Satz **??**). Der Beweis der Eindeutigkeit ist etwas einfacher wenn $R = \mathbb{K}$ ein Körper ist.(Satz **??**). Die Eindeutigkeit über Ringe wird durch die Leibniz Formel bewiesen (Satz 7.42). Der Weg zu diesen Beweisen wird mehrere Eigenschaften der Determinanten aufzeigen.

Definition 7.1. Eine Abbildung $d : \mathrm{Mat}_n(R) \longrightarrow R$ heißt **Determinante** wenn folgende Axiome gelten:

(D 1) d ist *linear in jeder Zeile*, das heißt für alle $i = 1, \ldots, n$ und $\lambda, \mu \in R$ und $Z_i, Z_i' \in R^n$ gilt

$$d\begin{pmatrix} \vdots \\ \hline \lambda Z_i + \mu Z_i' \\ \hline \vdots \end{pmatrix} = \lambda \cdot d\begin{pmatrix} \vdots \\ \hline Z_i \\ \hline \vdots \end{pmatrix} + \mu \cdot d\begin{pmatrix} \vdots \\ \hline Z_i' \\ \hline \vdots \end{pmatrix}$$

(D 2) d ist *alternierend*, das heißt, wenn A zwei gleiche Zeilen hat, dann ist $d(A) = 0$.

(D 3) d ist *normiert*, das heißt $d(I_n) = 1$.

Das Ziel ist jetzt folgenden Satz zu beweisen.

Satz 7.2. *Für jeden Ring R undq $n \in \mathbb{N}_{>0}$ existiert genau eine Determinante* $\det : \mathrm{Mat}_n(R) \longrightarrow R$.

Der Beweis dieses Satzes wird die nächsten zwei Sektionen füllen.

7.2.1 Existenz

Wir brauchen zu erst folgende Bezeichnung. Für jede Matrix $A = (a_{ij}) \in \mathrm{Mat}_n(R)$ und für alle $i, j \in \{1, \dots, n\}$ bezeichnen wir mit A_{ij} die Untermatrix von A die durch das löschen der i-ten Zeile und j-ten Spalte erhalten wird. Zum Beispiel,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad A_{21} = \begin{pmatrix} \blacksquare & 2 & 3 \\ \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & 8 & 9 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}$$

Wir beweisen die Existenz induktiv. Das heißt, dass wir induktiv Abbildungen $\det : \mathrm{Mat}_n(R) \longrightarrow R$ definieren, und dann zeigen, dass diese (D 1), (D 2), und (D 3) erfüllen.

Definition 7.3. Sei $\det : \mathrm{Mat}_n(R) \longrightarrow R$ die induktiv definierte Abbildung gegeben durch

$$\det A = \begin{cases} a_{11} & \text{wenn } n = 1 \\ \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det A_{i1} & \text{wenn } n \geq 2. \end{cases}$$

Diese Formel durch welcher wir det definiert haben heißt die **Laplace Entwicklung**[3] nach der ersten Spalte. Wenn wir die Summe aus der induktiven Definition 7.3 ausklappen, dann bekommen wir

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + \cdots + (-1)^{n-1} a_{n1} \det A_{n1}. \tag{7.1}$$

Beispiele:

a. Wenn $n = 2$ bekommt man die 2×2 Determinante die wir schon kennen:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \cdot \det\big((a_{22})\big) - a_{21} \cdot \det\big((a_1 2)\big) = a_{11}a_{22} - a_{21}a_{12}.$$

b. Im obigen Beispiel $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ haben wir

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} &= 1 \cdot \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 4 \cdot \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} + 7 \cdot \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \\ &= 1 \cdot (5 \cdot 9 - 8 \cdot 6) - 4 \cdot (2 \cdot 9 - 8 \cdot 3) + 7 \cdot (2 \cdot 6 - 5 \cdot 3) \\ &= 1 \cdot (-3) - 4 \cdot (-6) + 7 \cdot (-3) \\ &= 0 \end{aligned}$$

Proposition 7.4. *Das Axiom (D 1) gilt für* det. *Das heißt, dass die Abbildung* det *aus Definition 7.3 ist* linear in der Zeilen der Matrix. *Damit wird folgendes für alle $\lambda, \mu \in R$ gemeint:*

$$\det \begin{pmatrix} \vdots \\ \hline \lambda \cdot Z + \mu \cdot Y \\ \hline \vdots \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} \vdots \\ \hline Z \\ \hline \vdots \end{pmatrix} + \mu \cdot \det \begin{pmatrix} \vdots \\ \hline Y \\ \hline \vdots \end{pmatrix}.$$

[3] Wir werden sehen, dass das eine solche Entwicklung für jede Spalte und für jede Zeile möglich ist. Für unsere Zwecke reicht im Moment die Laplace Entwicklung nach der ersten Spalte.

Beweis-Skizze: Übung 1, Hausaufgabe 1. Q.E.D.

Unser nächstes Ziel ist (D 2) zu beweisen. Wir brauchen dafür folgende drei Bemerkungen.

Bemerkung 7.5. Wenn zwei aufeinanderfolgende Zeilen von A gleich sind, dann ist $\det A = 0$.

Beweis-Skizze: Wir zeigen das durch Induktion nach n. Für $n = 2$ ist das klar:

$$\det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = ab - ba = 0.$$

Für den Induktionsschritt, seien $Z_j = Z_{j+1}$ die konsekutive gleiche Zeilen. Wir wenden (7.1) an, und, weil A_{i1} zwei gleiche Zeilen haben für $i \notin \{j, j+1\}$, bekommen wir in diesen Fällen aus der Induktiven Voraussetzung, dass $\det A_{i,1} = 0$. Also

$$\det A = (-1)^{j-1}(a_{j1} \det A_{j1} - a_{j+1,1} \det A_{j+1,1}).$$

Aus $Z_j = Z_{j+1}$ bekommen wir, dass $A_{j1} = A_{j+1,1}$ und $a_{j1} = a_{j+1,1}$, also $\det A = 0$. Q.E.D.

Bemerkung 7.6. Wenn die Matrix A' aus der Matrix A durch Addieren des Vielfaches Zeile an einer benachbarten Zeile erhalten wurde, dann bleibt die Determinante unverändert.

Beweis-Skizze:

$$\det \begin{pmatrix} \vdots \\ Z \\ Y + \lambda Z \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ Z \\ Y \\ \vdots \end{pmatrix} + \lambda \cdot \det \begin{pmatrix} \vdots \\ Z \\ Z \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ Z \\ Y \\ \vdots \end{pmatrix}$$

Q.E.D.

Bemerkung 7.7. Wenn zwei aufeinanderfolgende Zeilen der Matrix vertauscht werden, dann wird die Determinante mit -1 multipliziert.

Beweis-Skizze:

$$\det \begin{pmatrix} \vdots \\ Z \\ Y \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ Z \\ Y - Z \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ Z + (Y - Z) \\ Y - Z \\ \vdots \end{pmatrix} =$$

$$= \det \begin{pmatrix} \vdots \\ Y \\ Y - Z \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ Y \\ (Y - Z) - Y \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ Y \\ Z \\ \vdots \end{pmatrix}$$

Q.E.D.

Proposition 7.8. *Das Axiom (D 2) gilt für die Abbildung* det *aus Definition 7.3.*

Beweis-Skizze: Wenn A zwei gleiche Zeilen $Z_i = Z_j$ hat, mit $i < j$, dann tauschen wir konsekutive Zeilen $i - j - 1$ Mal um Z_i in der Position $j - 1$ zu bekommen. Wir nennen diese neue Matrix A'. Aus Bemerkungen 7.7 und 7.5 gilt

$$\det A = (-1)^{j-i-1} \det A' = 0.$$

Q.E.D.

Bemerkung 7.6 wurde nicht direkt im Beweis von Proposition 7.8 angewendet. Diese wurde diese Bemerkung haben wir für den Beweis von Bemerkung 7.7 gebraucht.
Jetzt fehlt nur noch Axiom (D 3), das ist aber sehr einfach:

Bemerkung 7.9. Die Determinante der $n \times n$-Einheitsmatrix $I_n = (\delta_{ij})$ ist 1.

Beweis-Skizze: Offensichtlich gilt $\det I_1 = 1$. Per Induktion haben wir

$$\det I_n = 1 \cdot I_{n-1} - 0 \cdot \det \bullet + \cdots + (-1)^{n-1} \cdot \det \bullet = 1 \cdot 1 = 1.$$

Wir schreiben $\bullet$ für die Matrix die dort vorkommt, weil es keine Rolle spielt: dessen Determinante wird mit 0 multipliziert. Q.E.D.

Wir haben also folgenden Satz bewiesen.

Satz 7.10. *Für jeden Ring R und jede positive natürliche Zahl $n \in \mathbb{N}_{>0}$ existiert eine Abbildung* $\det : \mathrm{Mat}_n(R) \longrightarrow R$ *die multilinear in den Zeilen, alternierend und normiert ist.*

7.2.2 Eindeutigkeit über $\mathbb{K}$

Wir brauchen die Voraussetzung, dass die Einträge aus einem Körper sind nur ganz am Ende, wenn wir die Existenz und Eindeutigkeit der reduzierten Zeilen-Stufen-Form brauchen. Bis dann, gilt alles was wir hier machen für Matrizen mit Einträgen in einem beliebigen kommutativen Ring R.

Sei also $d : \mathrm{Mat}_n(R) \longrightarrow R$ *eine* Abbildung die die Axiome aus Definition 7.1 erfüllt.

Bemerkung 7.11. Wenn eine Zeile von A eine Nullzeile ist, dann haben wir $d(A) = 0$.

Beweis-Skizze: Wir können die Nullzeile mit 0 multiplizieren, und die Matrix bleibt unverändert. Es gilt also aus (D 1), dass $\det A = 0 \cdot \det A = 0$. Q.E.D.

Bemerkung 7.12. In den Beweisen von Bemerkungen 7.6, 7.6 und 7.7 sieht man, dass (D 1) und (D 2) dafür reichen. Weiterhin, weil (D 2), im Vergleich zu Bemerkung 7.5, auch für nicht aufeinanderfolgende Zeilen gilt, bekommen wir, dass die Bemerkungen 7.6, 7.6 und 7.7 auch für nicht konsekutive Zeilen gelten.

Wir erinnern hier, die Definition der elementaren Zeilenumformungen für Matrizen über Ringe:

Definition. Die **elementare Zeilenumformungen** sind Abbildungen $\Upsilon_\bullet : \mathrm{Mat}_{m\times n}(\mathbb{K}) \longrightarrow \mathrm{Mat}_{m\times n}(\mathbb{K})$ eine der folgenden Wirkungen haben:

$\boxed{\Upsilon_{k\leftrightarrow\ell}}$ Vertauscht die Zeilen k und ℓ, mit $k, \ell \in \{1, \dots, m\}$:

$$A = \begin{pmatrix} \vdots \\ Z_{\mathbf{k}} \\ \vdots \\ Z_\ell \\ \vdots \end{pmatrix} \overset{\Upsilon_{k\leftrightarrow\ell}}{\longmapsto} \begin{pmatrix} \vdots \\ Z_\ell \\ \vdots \\ Z_{\mathbf{k}} \\ \vdots \end{pmatrix}.$$

Dadurch verstehen wir, dass die Zeilen Z_k und Z_ℓ vertauscht werden, und alles andere unverändert bleibt.

$\boxed{\Upsilon_{k\to\lambda\cdot k}}$ Multipliziert die Zeile k mit dem invertierbaren Skalar[4] $\lambda \in R^\times$:

$$A = \begin{pmatrix} \vdots \\ Z_k \\ \vdots \end{pmatrix} \longmapsto \begin{pmatrix} \vdots \\ \lambda \cdot Z_k \\ \vdots \end{pmatrix}.$$

Dadurch verstehen wir, dass jeder Eintrag in der Zeile Z_k mit λ multipliziert wird, und alle andere Einträge unverändert bleiben.

$\boxed{\Upsilon_{k\to k+\lambda\cdot l}}$ Addiert λ mal die Zeile ℓ zu der Zeile k, mit $\lambda \in R$:

$$A = \begin{pmatrix} \vdots \\ Z_{\mathbf{k}} \\ \vdots \\ Z_\ell \\ \vdots \end{pmatrix} \longmapsto \begin{pmatrix} \vdots \\ Z_{\mathbf{k}} + \lambda \cdot Z_\ell \\ \vdots \\ Z_\ell \\ \vdots \end{pmatrix}.$$

Dadurch verstehen wir, dass jeder Eintrag a_{kj} in der Zeile Z_k durch $a_{kj} + \lambda a_{\ell j}$ ersetzt wird, und alle andere Einträge unverändert bleiben.

Satz. *Jede Elementare Zeilenumformung ist durch links-multiplizieren mit einer der folgenden drei Typen von invertierbaren Matrix erhalten:*

$$\begin{aligned} U_{\mathbf{k}\leftrightarrow\mathbf{l}} &= I_n - E_{kk} - E_{ll} + E_{k,l} + E_{l,k} \\ U_{\mathbf{k}\to(\lambda)\cdot\mathbf{k}} &= I_n + (\lambda - 1)E_{kk} \\ U_{\mathbf{k}\to\mathbf{k}+(\lambda)\cdot\mathbf{l}} &= I_n + \lambda E_{kl} \end{aligned}$$

Insbesondere sind elementare Zeilenumformungen bijektive Abbildungen.

Bemerkung 7.13. Eine Elementarmatrix $U_\bullet$ bekommt man von der Einheitsmatrix I_n durch die entsprechende Zeilenumformung $\Upsilon_\bullet$. Aus Axiom (D 3) haben wir $d(I_n) = 1$. Aus den Axiomen (D 1) und (D 2), zusammen mit der Bemerkung 7.12, bekommen wir, dass

$$\det U_{\mathbf{k}\leftrightarrow\mathbf{l}} = -1 \tag{7.2}$$

$$\det U_{\mathbf{k}\to(\lambda)\cdot\mathbf{k}} = \lambda \tag{7.3}$$

$$\det U_{\mathbf{k}\to\mathbf{k}+(\lambda)\cdot\mathbf{l}} = 1. \tag{7.4}$$

[4] Wenn $R = \mathbb{K}$ ein Körper ist, heißt das, dass $\lambda \neq 0$.

Proposition 7.14. *Sei $U_\bullet$ eine Elementarmatrix und A eine beliebige $n \times n$ Matrix. Dann gilt*

$$\det(U_\bullet \cdot A) = (\det U_\bullet)(\det A).$$

Insbesondere, haben wir

$$\begin{aligned} \det(U_{k\leftrightarrow l} \cdot A) &= -\det A && (7.5)\\ \det(U_{k\to(\lambda)\cdot k} \cdot A) &= \lambda \det A && (7.6)\\ \det(U_{k\to k+(\lambda)\cdot l} \cdot A) &= \det A && (7.7) \end{aligned}$$

Beweis-Skizze: Die Proposition folgt direkt aus den Bemerkungen 7.12 und 7.13. Q.E.D.

Bis zu diesem Punkt, gelten alle Beweise für Matrizen mit Einträgen in einem beliebigen Ring. Der Satz 7.15 gilt auch über R, aber der Beweis den wir jetzt geben braucht die Voraussetzung $R = \mathbb{K}$ ist ein Körper.

Jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ kann durch endlich-viele elementare Zeilenumformungen auf einer eindeutigen reduzierten Zeilenstufenform gebracht werden. Das heißt, es existieren Elementarmatrizen $U_1, \dots, U_r$ sodass

$$\mathrm{RZ}\, A = U_1 \cdots U_r \cdot A.$$

Weiterhin haben wir entweder $\mathrm{RZ}\, A = I_n$ oder $\mathrm{RZ}\, A$ hat eine Nullzeile, also

$$\det \mathrm{RZ}\, A = \begin{cases} 1 & \text{wenn } A \text{ invertierbar ist} \\ 0 & \text{wenn } A \text{ nicht invertierbar ist} \end{cases} \qquad (7.8)$$

Es folgt also aus Proposition 7.14, weil alle $U_\bullet$ invertierbar sind, dass $(\det U_\bullet) \cdot (\det U_\bullet^{-1}) = \det I_n = 1$, also $\det U_\bullet$ ist invertierbar in $\mathbb{K}$. Weiterhin, aus derselben Proposition 7.14 haben wir

$$\det(A) = \det(U_1)^{-1} \cdots \det(U_r)^{-1} \cdot \det(\mathrm{RZ}\, A). \qquad (7.9)$$

Wir haben endlich unseren Satz für Matrizen über Körper.

Satz 7.15. *Für jeden Körper $\mathbb{K}$existiert genau eine Abbildung $d : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{K}$ die multilinear in den Zeilen, alternierend und normiert ist.*

Beweis-Skizze: Bemerkung 7.13 zeigt, dass d für die Elementarmatrizen von den Axiomen (D 1–3) eindeutig bestimmt sind. Für die $\mathrm{RZ}\, A$ ist $d(\mathrm{RZ}\, A)$ auch eindeutig bestimmt, und hängt nur von der Invertierbarkeit von A ab. Proposition 7.14 zusammen mit der Eindeutigkeit der reduzierten Zeilenstufenform für Matrizen über $\mathbb{K}$ folgt, folgt die Aussage des Satzes. Q.E.D.

[2] 13.4.21

7.2.3 Eigenschaften

Bevor wir die Eindeutigkeit der Determinante für allgemeine Ringe beweisen, sammeln wir hier einige nützliche Eigenschaften von Determinanten über Körper. Manche gelten allgemeiner in $\mathrm{Mat}_n(R)$, aber den Beweis im allgemeinen Fall müssen wir verschieben. Wir fangen mit einer Eigenschaft die *nicht* gilt: Linearität.

Bemerkung 7.16. Für $n > 1$ ist die Determinante $\det : \mathrm{Mat}_n(R) \longrightarrow \mathbb{R}$ **nicht linear**[5]: für alle $\lambda \in R$ und alle $A \in \mathrm{Mat}_n(R)$ gilt

$$\det(\lambda \cdot A) = \lambda^n \cdot \det A. \tag{7.10}$$

Also wenn $\lambda^n \neq \lambda$ scheitert die Kompatibilität mit der Multiplikation mit Skalaren. Das gibt uns ein einfaches Gegenbeispiel auch für die Verträglichkeit mit der Addition wenn $1 + 1 \neq 0$:

$$\det I_2 + \det(-I_2) = 1 + 1 \neq 0 = \det(I_2 - I_2) = \det(\mathbf{0}).$$

Es ist wichtig zu betonen, dass die Verträglichkeit mit der Addition in den meisten Fällen[6] scheitert. Hier ist noch ein einfaches Gegenbeispiel: $A = \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}$ und $B = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$. Es gilt

$$\begin{aligned} \det(A+B) = \det \begin{pmatrix} 1+2 & 3+4 \\ 5+6 & 7+8 \end{pmatrix} &= \det \begin{pmatrix} 1 & 3 \\ 5+6 & 7+8 \end{pmatrix} + \det \begin{pmatrix} 2 & 4 \\ 5+6 & 7+8 \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix} + \det \begin{pmatrix} 1 & 3 \\ 6 & 8 \end{pmatrix} + \det \begin{pmatrix} 2 & 4 \\ 5 & 7 \end{pmatrix} + \det \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \\ &= \det A + \det B - 14. \end{aligned}$$

Jetzt zeigen wir, dass die Invertierbarkeit einer Matrix an der Determinante ablesbar ist. Das gilt auch für Ringe, aber wir werden es erst in Sektion 7.4 formulieren und beweisen.

Satz 7.17. *Für eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind folgende Aussagen äquivalent:*

(a) *A ist invertierbar.*

(b) $\mathrm{Rang}\, A = n$.

(c) $\det A \neq 0$.

Beweis-Skizze: Die Äquivalenz von (a) und (b) ist uns schon bekannt[a]. Die Äquivalenz von (a) und (c) ist eine direkte Folgerung von (7.8) und (7.9). Q.E.D.

[a] Zur Erinnerung: diese folgt aus der Definition der zugeordneten $\mathbb{K}$-linearen Abbildung $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$, der Definition von Rang und aus dem Dimensionssatz für $\mathbb{K}$-lineare Abbildungen.

Folgender Satz gibt uns eine der wichtigsten Eigenschaften der Determinante. Es verallgemeinert Proposition 7.14 zu dem Produkt beliebiger Matrizen, und gilt auch über Ringe.d

Satz 7.18. *Für $A, B \in \mathrm{Mat}_n(\mathbb{K})$ gilt* $\det(A \cdot B) = \det A \cdot \det B$.

[5] Wir haben das formell nicht definiert. Aber Matrizen über R kann man genau so gut mit Skalaren aus R multiplizieren, also R-linear heißt genau was man erwartet: $d(\lambda A + \mu B) = \lambda d(A) + \mu d(B)$.

[6] "in den meisten Fällen" ist nicht eine genaue Aussage. Was ich damit meine, ist, dass es keine gute/natürliche Bedingung die uns $\det(A + B) \neq \det A + \det B$.

Beweis-Skizze: Aus Proposition 7.14 gilt die Aussage wenn A eine Elementarmatrix ist. Wenn A invertierbar ist, dann existieren $U_1, \dots, U_r$ Elementarmatrizen, sodass $A = U_1 \cdots U_r$. Aus 7.14 folgt $\det A = (\det U_1) \cdots (\det U_r)$, und auch

$$\det(AB) = \det(U_1 \cdots U_r B) = (\det U_1) \cdots (\det U_r)(\det B) = (\det A)(\det B).$$

Wenn A nicht invertierbar ist, dann ist auch AB nicht invertierbar (sonst hätten wir $A \cdot (B \cdot (AB)^{-1}) = I_n$). Aus Satz 7.17 haben wir dann $\det(AB) = 0 = 0 \cdot \det B = (\det A)(\det B)$. Q.E.D.

Korollar 7.19. Wenn $A \in \mathrm{GL}_n(\mathbb{K})$, dann gilt

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

Beweis-Skizze: $(\det A)(\det A^{-1}) = \det(A \cdot A^{-1}) = \det I_n = 1$. Q.E.D.

Satz 7.20. *Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ und $A^\top$ die Transponierte von A. Es gilt*

$$\det A = \det A^\top.$$

Beweis-Skizze: Die Aussage ist wahr für Elementarmatrizen wegen Proposition 7.14. Weil $(AB)^\top = B^\top A^\top$ folgt die Aussage für invertierbare Matrizen aus der Zerlegung als Produkt von Elementarmatrizen und aus Satz 7.18. Wenn die Matrix nicht invertierbar ist, dann ist die Transponierte auch nicht invertierbar, also nach Satz 7.17 sind beide Determinanten Null. Q.E.D.

Korollar 7.21. Man kann "Zeilen" mit "Spalten" überall in diesem Kapitel ersetzen.

7.3 Die Symmetrische Gruppe

Damit wir ähnliche Eigenschaften für Matrizen mit Koeffizienten in Ringe beweisen, brauchen wir eine andere Formel für die Determinante: die Formel von Leibniz. Aber dazu brauchen wir erstmals die symmetrische Gruppe zu verstehen.

Sei $n \in \mathbb{N}_{>0}$. Die **symmetrische Gruppe** S_n ist die Gruppe aller bijektiven Selbstabbildungen der Menge $\{1, \dots, n\}$, mit der Verknüpfung von Abbildungen als Gruppenoperation. Also

$$(S_n, \cdot) := (\{\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\} \mid f \text{ bijektiv}\}, \circ).$$

Die Verknüpfung von Abbildungen ist immer assoziativ. Die identische Abbildung id_n ist bijektiv, es liegt also in S_n, und es ist das neutrale Element der Gruppe. Eine Abbildung ist invertierbar genau dann wenn es bijektiv ist. Also ist S_n tatsächlich eine Gruppe. Wir werden gleich sehen, dass diese Gruppe nicht kommutativ ist wenn $n \geq 3$. Die Elementen von S_n heißen **Permutationen**. Wir schreiben eine Permutation $\sigma \in S_n$ als eine $2 \times n$ Tafel mit $1 \dots n$ in der ersten Zeile und unter jedem i das Bild von i via σ:

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

Weil σ bijektiv ist, kommt in der zweiten jede Zahl von 1 bis n genau ein Mal vor. Die Reihenfolge ist aber permutiert.

Beispiele:

1. Sei $n = 4$. Für $\sigma, \pi \in S_4$ um $\sigma \cdot \pi$ zu bestimmen brauchen wir $(\sigma \circ \pi)(i) = \sigma\big(\pi(i)\big)$ für jeden $i \in \{1, 2, 3, 4\}$ zu berechnen:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \qquad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad \sigma \cdot \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

$$\shortparallel$$

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix} \quad \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad \pi \cdot \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

Man merkt also, dass S_4 nicht Abelsch ist. Übung: Was ist $(\sigma \cdot \pi)^{-1}$ und was ist der Zusammenhang mit σ^{-1} und π^{-1}.

2. Wenn $n = 1$, dann gibt es eine einzige Abbildung: die Identität $\mathrm{id} : \{1\} \longrightarrow \{1\}$. Also S_1 ist die Gruppe mit einem einzigen Element. Diese Gruppe ist kommutativ.

3. Wenn $n = 2$, dann gibt es 2^2 Abbildungen von $\{1, 2\} \longrightarrow \{1, 2\}$, aber nur 2 davon sind bijektiv: id und $\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. $S_2 \simeq \mathbb{Z}/2\mathbb{Z}$ ist die[7] Gruppe mit zwei Elementen. Diese ist auch kommutativ.

4. Wenn $n = 3$, dann gibt es drei Möglichkeiten 1 abzubilden: auf 1, auf 2 oder auf 3. Sobald das fixiert wurde, dann haben wir in jedem der Fälle zwei Varianten die 2 abzubilden: $\sigma(2) \in \{1, 2, 3\} \setminus \{\sigma(1)\}$. Sonst wäre σ nicht injektiv, und somit auch nicht bijektiv. Sobald $\sigma(1)$ und $\sigma(2)$ bekannt ist, dann bleibt wegen der Bijektivität nur eine Mögliche Wahl für $\sigma(3)$. Also $|S_3| = 3 \cdot 2 \cdot 1 = 3! = 6$.

Proposition 7.22. *Die Gruppe S_n hat Ordnung $n! = 1 \cdot 2 \cdots n$.*

Beweis-Skizze: Für Beweis durch vollständige Induktion habe wir in Beispiel **2.** hier oben den Induktionsanfang gezeigt. Der Induktionsschritt ist völlig analog zu dem Beweis in Beispiel **4.**. Q.E.D.

Weil für $n = 0$ oder $n = 1$ die Gruppe S_n die triviale Gruppe ist, werden wir für den Rest dieses Abschnittes immer $n \geq 2$ annehmen.

Eine Permutation $\sigma \in S_n$ heißt **Zyklus**, wenn es $a \in \{1, \ldots, n\}$ und $r \geq 2$ existieren, sodass

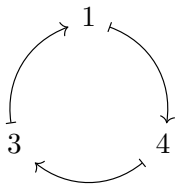
$$\begin{array}{rcl} |\{a, \sigma(a), \ldots, \sigma^{r-1}(a)\}| & = & r \ \text{ und} \\ i \notin \{a, \sigma(a), \ldots, \sigma^{r-1}(a)\} & \Leftrightarrow & \sigma(i) = i. \end{array}$$

Wir sagen in diesem Fall, dass die **Länge** des Zyklus r ist, oder, dass σ ein r-Zyklus ist. Der einzige 1-Zyklus ist die Identität. Ein Zyklus von Länge zwei heißt **Transposition**. Um einen Zyklus zu bestimmen, reicht es die Elementen $a, \sigma(a), \ldots, \sigma^{r-1}(a)$ in Reihenfolge zu kennen. Wir schreiben deswegen ein r-Zyklus σ einfach als $(a, \sigma(a), \ldots, \sigma^{r-1}(a))$. Zum Beispiel

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 3 & 5 \end{pmatrix} =: (1\ 4\ 3)$$

[7] Wir sagen die Gruppe mit zwei Elementen, weil alle Gruppen mit zwei Elementen isomorph zueinander sind.

ist ein 3-Zyklus in S_5. Das Element a in dem obigen Beispiel ist 1. Das ist aber nicht eindeutig bestimmt. Mann kann es auch 4 oder 3 wählen, also, wenn wir die Fixpunkte[8] weglassen, dann bekommen wir folgende Darstellung:



$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 3 & 5 \end{pmatrix} = (1\ 4\ 3) = (4\ 3\ 1) = (3\ 1\ 4).$$

Allgemein, kann man a mit jedem $\sigma^i(a)$ ersetzten. Wichtig ist, dass die Reihenfolge im Zyklus erhalten wird. Also $(1\ 4\ 3) \neq (3\ 1\ 4)$. Noch eine Bemerkung dazu: aus der Zyklus Bezeichnung ist es nicht klar was n ist. Also $(1\ 3\ 4)$ könnte Element von S_n für alle $n \geq 4$ sein.

Übung. Zeigen Sie, dass wenn $\sigma = (a, \sigma(a), \ldots, \sigma^{r-1}(a))$ ein $r-$Zyklus ist, mit $2 \leq r \leq n$, dann gilt

(a) $\sigma^r = \mathrm{id}_n$

(b) $\sigma^i \neq \mathrm{id}_n$ für $0 < i < r$.

Bemerkung 7.23. Für eine Transposition τ gilt $\tau = \tau^{-1}$.

Zwei Zyklen $(i_1, \ldots, i_r)$ und $(j_1, \ldots, j_s) \in S_n$ sind **disjunkt** wenn $\{i_1, \ldots, i_r\} \cap \{j_1, \ldots, j_s\} = \emptyset$.

Bemerkung 7.24. Zwei Zyklen $\gamma = (i_1, \ldots, i_r)$ und $\delta = (j_1, \ldots, j_s) \in S_n$ kommutieren wenn diese disjunkt sind.

Beweis-Skizze: Seien γ und δ disjunkt. Dann gilt

$$(\gamma \cdot \delta)(a) = \begin{cases} a & \text{wenn } a \notin \{i_1, \ldots, i_r, j_1, \ldots, j_s\}, \\ i_{k+1} & \text{wenn } a = i_k, \\ j_{\ell+1} & \text{wenn } a = j_\ell, \end{cases}$$

wobei die Indizes von j in $\mathbb{Z}/s\mathbb{Z}$ sind und die von i in $\mathbb{Z}/r\mathbb{Z}$. Also $(\gamma \cdot \delta)(a) = (\delta \cdot \gamma)(a)$ für alle a. Q.E.D.

Übung. Man finde zwei verschiedene und nicht disjunkte Zyklen die kommutieren.

Proposition 7.25. *Jede Permutation $\sigma \in S_n$ kann als Produkt disjunkter Zyklen geschrieben werden. Diese Schreibweise ist, bis auf der Reihenfolge der Zyklen, eindeutig.*

Beweis-Skizze: Sei $\sigma \in S_n$. Für jeden $1 \leq i \leq n$ sei $A_i := \{\sigma^m(i) \ : \ m \in \mathbb{N}\}$. Für alle $i \neq j$ sind A_i und A_j entweder gleich oder disjunkt. Man wählt dann eine Teilmenge $\{j_1, \ldots, j_r\}$ sodass $|A_{j_k}| > 1$ und jede Menge A_i mit $|A_i| > 1$ zu einer Menge $A_{j_\bullet}$ gleich ist. Dann ist σ das Produkt folgender disjunkter Zyklen, wobei das leere Produkt gleich dem Neutralelement $\mathrm{id} \in S_n$ ist.

$$\sigma = (j_1\ \sigma(j_1)\ \ldots) \cdot \ldots \cdot (j_r\ \sigma(j_r)\ \ldots).$$

[8] Das heißt i mit $\sigma(i) = i$.

Q.E.D.

Satz 7.26. *Sei $n \geq 2$. Für jede Permutation $\sigma \in S_n$ gibt es $k \in \{0, \ldots, n-1\}$ und $\tau_1, \ldots, \tau_k$ Transpositionen, sodass $\sigma = \tau_1 \ldots \tau_k$.*

Beweis-Skizze: Aus Proposition 7.25 reicht es die Aussage für Zyklen zu beweisen. Nach eventuelle Umbenennung, haben wir

$$\begin{aligned}(1\ 2\ 3\ \ldots\ k) &= (1\ k)\cdot(1\ k-1)\ \ldots\ (1\ 3)\cdot(1\ 2) \quad \text{oder}\\ (1\ 2\ 3\ \ldots\ k) &= (1\ 2)\cdot(2\ 3)\ \ldots\ (k-2\ k-1)\cdot(k-1\ k).\end{aligned}$$

Q.E.D.

[3] 19.4.'21

Wie man im Beweis sehen kann, ist eine solche Zerlegung nicht eindeutig. Nicht einmal die Anzahl der Transpositionen ist eindeutig:

$$(1\ 3) = (1\ 2)\cdot(2\ 3)\cdot(1\ 2).$$

Jedoch, wir werden gleich sehen, dass die Parität der Anzahl von Transpositionen konstant bleibt. Deswegen werden wir sagen, dass eine Permutation **gerade** ist, wenn die Anzahl der Faktoren in einer Zerlegung als Produkt von Permutationen gerade ist. Wenn diese Anzahl ungerade ist, dann sagen wir dass die Permutation **ungerade** ist. Man muss aber zu erst beweisen, dass diese Definitionen Sinn[9] haben. Dafür führen wir folgender Begriff ein.

Definition 7.27. Das **Vorzeichen** (oder das **Signum** oder die **Signatur** oder die **Parität**) einer Permutation $\sigma \in S_n$ ist

$$\operatorname{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Das ist intrinsisch definiert, also man muss sich keine Sorgen machen, dass es Sinn macht.

Man kann das Signum auch durch über folgender Begriff beschreiben,

Definition 7.28. Sei $\sigma \in S_n$. Ein **Fehlstand** (oder eine **Inversion**) ist ein geordnetes Paar $(i, j) \in \{1, \ldots, n\}^2$ mit $i < j$ und $\sigma(i) > \sigma(j)$. Wir bezeichnen die Menge aller Fehlstände von σ mit

$$\operatorname{inv}(\sigma) = \{(i, j) \mid 1 \leq i < j \leq n \text{ und } \sigma(i) > \sigma(j)\}.$$

Übung. Zeigen[10] Sie, dass für $\sigma \in S_n$ gilt $\operatorname{sgn}(\sigma) = (-1)^{|\operatorname{inv}(\sigma)|}$.

Bemerkung 7.29. Für die Identität $\mathrm{id}_n \in S_n$ gilt $\operatorname{sgn}(\mathrm{id}_n) = 1$.

Satz 7.30. *Die Abbildung $\operatorname{sgn} : S_n \longrightarrow \{-1, 1\}$ ist ein Gruppenhomomorphismus, wobei die Gruppenoperation auf $\{-1, 1\}$ die Multiplikation ist.*

[9] Das heißt, dass diese Parität wirklich invariant für jede Permutation ist.
[10] Sie finden einen Eleganten Beweis dafür in [Fis09, S.288]

Beweis-Skizze: Wir müssen also zeigen, dass für alle $\sigma, \pi \in S_n$ gilt $\operatorname{sgn}(\sigma \cdot \pi) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\pi)$. Der Trick ist der folgende:

$$\begin{aligned} \operatorname{sgn}(\sigma \cdot \pi) &= \prod_{i<j} \frac{\sigma\big(\pi(j)\big) - \sigma\big(\pi(i)\big)}{j - i} \\ &= \prod_{i<j} \frac{\sigma\big(\pi(j)\big) - \sigma\big(\pi(i)\big)}{\pi(j) - \pi(i)} \cdot \prod_{i<j} \frac{\pi(j) - \pi(i)}{j - i} \end{aligned}$$

Das zweite Produkt ist per Definition $\operatorname{sgn}(\pi)$. Wir müssen nur noch bemerken, dass das erste Produkt $\operatorname{sgn}(\sigma)$ ist. Dafür behaupten wir, dass jeder Bruch ein Bruch aus der Definition von $\operatorname{sgn}(\sigma)$ ist. Wenn wir $k := \pi(j)$ und $\ell := \pi(i)$ definieren, dann ist das einzige das stören könnte, dass $k < \ell$ vorkommen könnte. Aber das kann gleich wieder gut gemacht werden, weil

$$\frac{\sigma(k) - \sigma(\ell)}{k - \ell} = \frac{\sigma(\ell) - \sigma(k)}{\ell - k}.$$

Q.E.D.

Korollar 7.31. Wenn $\sigma \in S_n$, dann gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Bemerkung 7.32. Für eine Transposition $\tau = (i, j) \in S_n$ gilt $\operatorname{sgn}(\tau) = -1$.

Beweis-Skizze: Wir verwenden die obige Beschreibung des Signums als $(-1)^{|\operatorname{inv}(\sigma)|}$. Das heißt, das $\operatorname{sgn}(1\ 2) = -1$, weil es eine einzige Inversion hat. Eine Beliebige Transposition $(i\ j)$ ist aber gleich[a] mit

$$(i\ j) = (1\ i) \cdot (2\ j) \cdot\ (1\ 2) \cdot (2\ j) \cdot (1\ i).$$

Aus Satz 1.121 und Korollar 7.31 folgt was wir wollen. Q.E.D.

[a] wenn $i = 1$ oder $j = 2$, dann verwenden wir die Konvention, dass $(a\ a) = \mathrm{id}$.

Wir kommen endlich zur Invarianz der Parität der Anzahl von Transpositionen in der Zerlegung einer Permutation.

Korollar 7.33. (a) Wenn $\sigma \in S_n$ das Produkt der Transpositionen $\tau_1, \ldots, \tau_r \in S_n$ ist, dann gilt $\operatorname{sgn}(\sigma) = (-1)^r$.

(b) Die Menge $A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$ ist eine Untergruppe von S_n.

Die Untergruppe A_n heißt die **alternierende Untergruppe** von S_n.

Übung. Die Untergruppe A_n ist ein Normalteiler von S_n.

Bemerkung 7.34. Für jede Transposition $\tau = (i, j) \in S_n$ haben wir die disjunkte Vereinigung

$$S_n = A_n \sqcup \tau A_n.$$

Anders gesagt, ist S_n das semidirekte Produkt[11] der Untergruppen A_n und $\operatorname{Span}_{\mathbb{K}}\{\tau\}$. Das wird als $S_n = A_n \rtimes \operatorname{Span}_{\mathbb{K}}\{\tau\}$ geschrieben.

[11] Der Begriff von semiderektes Produkt wurde nicht eingeführt. Machen Sie sich keine weitere Gedanken darüber.

7.3.1 Permutationsmatrizen

Ein Permutationsmatrix ist eine Matrix $P \in \mathrm{Mat}_n(R)$ mit der Eigenschaft, dass links-Multiplizieren mit P die Zeilen permutiert[12].

Definition 7.35. Sei $\sigma \in S_n$. Die zu σ zugeordnete **Permutationsmatrix** mit Koeffizienten im Ring R ist die Matrix $P_\sigma \in \mathrm{Mat}_n(R)$ gegeben durch

$$P_\sigma := \left(\begin{array}{c} e_{\sigma(1)} \\ \hline \vdots \\ \hline e_{\sigma(n)} \end{array}\right)$$

Wenn $A \in \mathrm{Mat}_{n,r}(R)$, dann bezeichnen wir mit $Z_i(A) \in \mathrm{Mat}_{1n}(R)$ die i-te Zeile von A. Für $P_\sigma \in \mathrm{Mat}_n(R)$ und für jede Matrix $A \in \mathrm{Mat}_{n,r}(R)$ mit $r \in \mathbb{N}_{>0}$ beliebig gilt

$$Z_i(P_\sigma \cdot A) = Z_{\sigma(i)}(A).$$

Wenn P eine Permutationsmatrix ist, dann bezeichnet σ_P die entsprechende Permutation also $P = P_{\sigma_P}$.

Beispiele:

1. Sei $n = 3$ und $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. Also σ ist der 3-Zykel (1 3 2). Dann haben wir

$$P_\sigma \cdot \mathbf{x} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}.$$

   Man merkt hier das die Wirkung der Linksmultiplikation mit P_σ dieselbe ist mit der Wirkung von σ auf der Menge $\{1, 2, 3\}$.

2. Die Zeilenumformmatrizen $U_{\mathbf{k} \leftrightarrow \mathbf{l}} = P_{(k\ l)}$.

Bemerkung 7.36. Jede Spalte von P_σ enthält genau eine 1 und sonst 0, und jede Zeile von P_σ enthält genau eine 1 und sonst 0. Umgekehrt, jede solche Matrix ist eine Permutationsmatrix.

Dieser Satz und dieser Beweis haben bis am 28.4.2021 einen Fehler enthalten. Nämlich stand hier, dass $\sigma \mapsto P_\sigma$ ein Gruppenhomomorphismus war, das heißt, dass die Reihenfolge der Gruppenoperationen nicht vertauscht wurde. Das gilt aber nur für $\sigma \mapsto P_{\sigma^{-1}}$.

Satz 7.37. *Die Abbildung $P_\bullet : S_n \to \mathrm{GL}_n(R)$ gegeben durch $\sigma \mapsto P_\sigma$ ist ein Gruppen-Antihomomorphismus. Das bedeutet, dass die Reihenfolge der Gruppenoperation wird umgedreht:*

$$P_{\sigma \cdot \pi} = P_\pi \cdot P_\sigma.$$

[12] Es gibt zwei natürliche Möglichkeiten einer Permutationsmatrix P eine Permutation σ zuzuordnen:

1. Die Permutation die der Wirkung von $V1(\sigma)$ auf den **Zeilenindizes** entspricht.
2. Die Permutation die der Wirkung von $V2(\sigma)$ auf den **Zeilen** entspricht.

Variante 1. heißt, dass $V1(\sigma) \cdot \mathbf{x} = (x_{\sigma(1)} \dots x_{\sigma(n)})^\top$. Variante 2. heißt, dass x_i wird an der stelle $\sigma(i)$ in $V2(\sigma) \cdot \mathbf{x}$ vorkommen. Der Zusammenhang der beiden ist einfach: $V1(\sigma) = V2(\sigma^{-1})$. Wir werden die Variante 1 wählen, die zur Zeit verbreiteter zu sein scheint. (In [Fis09] und jetzt, in 2021 auf Wikipedia (de, en) wird die Variante 1 präsentiert. In [Bos08, Art91] und auf Wikipedia in 2019 ist es die Variante 2.)

Beweis-Skizze: Seien $\sigma, \pi \in S_n$. Wir müssen zeigen, dass $P_{(\sigma\cdot\pi)} = P_\pi \cdot P_\sigma$. Es gilt

$$P_\pi \cdot P_\sigma. = P_\pi \cdot \begin{pmatrix} \text{---}\, e_{\sigma(1)} \,\text{---} \\ \vdots \\ \text{---}\, e_{\sigma(n)} \,\text{---} \end{pmatrix} = \begin{pmatrix} \text{---}\, e_{\sigma(\pi(1))} \,\text{---} \\ \vdots \\ \text{---}\, e_{\sigma(\pi(n))} \,\text{---} \end{pmatrix} = P_{\sigma\cdot\pi}$$

Der wichtigste Schritt ist in der zweite. I Dort muss man folgen welches e_j aus $P_\sigma$cdurch links-Multiplizieren mit P_π auf der i-ten Stelle gebracht wird. Es ist nämlich das e_j das in der Zeile mit Index $\pi(i)$ steht. In P_σ steht in der Zeile j immer $e_{\sigma(j)}$. Also in der Zeile i von des Produktes kommt genau $e_{\sigma(\pi(i))}$ Q.E.D.

Dieser Satz und dieser Beweis haben bis am 28.4.2021 einen Fehler enthalten. Nämlich stand hier, dass $\sigma \mapsto P_\sigma$ ein Gruppenhomomorphismus war, das heißt, dass die Reihenfolge der Gruppenoperationen nicht vertauscht wurde. Das gilt aber nur für $\sigma \mapsto P_{\sigma^{-1}}$.

Wir erinnern, dass eine $n \times n$ Standardmatrix eine Matrix der Form $E_{ij} = (\delta_{ki} \cdot \delta_{jl})_{k,l=1\ldots n}$ ist. Die folgende Bemerkungen sind Umformungen der derselben einfachen Tatsache.

Bemerkung 7.38. (a) $P_\sigma = E_{1,\sigma(1)} + \cdots + \mathbb{E}_{n,\sigma(n)}$.

(b) Die i-te Zeile von P_σ ist der standard Zeilenvektor $e_{\sigma(i)}$.

(c) Die j-te Spalte von P_σ ist der standard Spaltenvektor $e_{\sigma^{-1}(j)}$.

(d) Die Einträge p_{ij} von P_σ erfüllen

$$p_{ij} = \begin{cases} 0 & \text{wenn } j \neq \sigma(i) \\ 1 & \text{wenn } j = \sigma(i) \end{cases}.$$

7.4 Die Eindeutigkeit der Determinante über einen Ring

Das Ziel in diesem Teil ist den Satz 7.42 zu bewiesen. Sei also R ein Ring und $d : \mathrm{Mat}_n(R) \longrightarrow R$ eine Determinante, das heißt eine Abbildung die die Axiome (D 1-3) aus Definition 7.1 erfüllt. Wir werden "eine Determinante d" statt "die Determinante det" benutzen, bis die Eindeutigkeit bewiesen ist.

Bemerkung 7.39. Elementarmatrizen $U_\bullet$ kann man in $\mathrm{Mat}_n(R)$ genauso wie in $\mathrm{Mat}_n(\mathbb{K})$ definieren:

$$U_{\mathbf{k}\leftrightarrow\mathbf{l}} = \begin{array}{c} \\ \mathbf{k} \\ \\ \mathbf{l} \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccc} & & & \mathbf{k} & & \mathbf{l} & \\ \end{array} \\ \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}$$

$$U_{\mathbf{k}\to(\lambda)\cdot\mathbf{k}} = \begin{array}{c} \\ \mathbf{k} \end{array} \overset{\overset{\mathbf{k}}{\downarrow}}{\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}}$$

$$U_{\mathbf{k}\to\mathbf{k}+(\lambda)\cdot\mathbf{l}} = \begin{array}{c} \\ \mathbf{k} \\ \mathbf{l} \end{array} \overset{\overset{\mathbf{k}}{\downarrow}\quad\overset{\mathbf{l}}{\downarrow}}{\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \lambda & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}}$$

wobei $k, l \in \{1, \dots, n\}$ und $\lambda \in R$. Es ist wichtig zu bemerken, dass $U_{\mathbf{k}\leftrightarrow\mathbf{l}}$ und $U_{\mathbf{k}\to\mathbf{k}+(\lambda)\cdot\mathbf{l}}$ invertierbar bleiben, mit

$$U_{\mathbf{k}\leftrightarrow\mathbf{l}}^{-1} = U_{\mathbf{k}\leftrightarrow\mathbf{l}} \quad \text{und} \quad U_{\mathbf{k}\to\mathbf{k}+(\lambda)\cdot\mathbf{l}}^{-1} = U_{\mathbf{k}\to\mathbf{k}+(-\lambda)\cdot\mathbf{l}}.$$

Die Matrix $U_{\mathbf{k}\to(\lambda)\cdot\mathbf{k}}$ ist invertierbar, genau dann, wenn λ invertierbar in R ist. In diesem Fall haben wir

$$U_{\mathbf{k}\to(\lambda)\cdot\mathbf{k}}^{-1} = U_{\mathbf{k}\to(\lambda^{-1})\cdot\mathbf{k}}.$$

Sei $(i\ j) \in S_n$ die Transposition die i und j vertauscht und alle anderen k stehen läßt. Es gilt also $P_{(i\ j)} = U_{\mathbf{i}\leftrightarrow\mathbf{j}}$. Aus Satz 7.26 ist jede Permutation $\sigma \in S_n$ Produkt von Transpositionen: $\sigma = \prod_{k=1}^{r}(i_k\ j_k)$. Es gilt dann nach Satz 7.37, dass

$$P_\sigma = \prod_{k=1}^{r} U_{i_k\leftrightarrow j_k}.$$

Also, aus Bemerkung 7.13, Proposition 7.14, und Satz 7.30, folgt:

Bemerkung 7.40. Für jede Determinante d und jede Permutationsmatrix $P_\sigma \in \mathrm{Mat}_n(R)$ gilt

$$d(P_\sigma) = \mathrm{sgn}(\sigma).$$

Aus Proposition 7.14 folgt, dass für jede Matrix $A \in \mathrm{Mat}_n(R)$ und jede Permutation $\sigma \in S_n$ gilt

$$d(P_\sigma \cdot A) = \mathrm{sgn}(\sigma) \cdot d(A) \tag{7.11}$$

Im Beweis von Satz 7.42 werden folgende Bemerkung brauchen.

Bemerkung 7.41. Sei d eine Determinante und $Q \in \mathrm{Mat}_n(R)$ eine Matrix mit genau n Einträge gleich 1 und alle anderen Einträge 0. Dann gilt

$$d(Q) \neq 0 \iff Q \text{ ist eine Permutationsmatrix.}$$

Beweis-Skizze: $\boxed{\Leftarrow}$ Wenn Q eine Permutationsmatrix P_σ ist, dann haben wir aus Bemerkung 7.40 dass $d(Q) = \operatorname{sgn}(\sigma) \neq 0$.

$\boxed{\Rightarrow}$ Wir zeigen, dass wenn Q nicht eine Permutationsmatrix ist, dann ist $d(Q) = 0$. Wenn Q eine Nullzeile hat, dann $d(Q) = 0$ folgt aus Bemerkung 7.11. Wenn Q keine Nullzeile hat, dann, weil Q genau n Einträge gleich 1 hat, gilt

$$\text{Jede Zeile von } Q \text{ hat genau eine 1.} \tag{7.12}$$

Da $Q \neq P_\sigma$ für alle $\sigma \in S_n$, gibt es eine Spalte mit mindestens zwei Einträge gleich 1, und aus (7.12) sind die entsprechenden Zeilen gleich. Also aus (D2) folgt $d(Q) = 0$. Q.E.D.

Bei der Eindeutigkeit der Determinante über $\mathbb{K}$ (Satz 7.15) haben wir die reduzierte Zeilenstufenform angewendet, und diese existiert nicht immer über einem Ring. Also obwohl die Aussage gleich ist, muss der Beweis angepasst werden.

Satz 7.42 (Eindeutigkeit der Determinante)**.** *Wenn $A = (a_{ij}) \in \operatorname{Mat}_n(R)$ ist, und $d : \operatorname{Mat}_n(R) \longrightarrow R$ eine Determinate ist, dann gilt*

$$d(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \tag{7.13}$$

Insbesondere, wenn eine Determinante existiert, ist diese eindeutig.

Beweis-Skizze: Sei $A = (a_{ij}) \in \operatorname{Mat}_n(R)$. Wir wenden jetzt das Axiom (D1) n^n Mal an. Zu erst, haben wir

$$d(A) = d\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ \hline Z_2(A) & & & \\ \hline \vdots & & & \\ \hline Z_n(A) & & & \end{pmatrix} + d\begin{pmatrix} 0 & a_{12} & \dots & 0 \\ \hline Z_2(A) & & & \\ \hline \vdots & & & \\ \hline Z_n(A) & & & \end{pmatrix} + \cdots + d\begin{pmatrix} 0 & 0 & \dots & a_{1n} \\ \hline Z_2(A) & & & \\ \hline \vdots & & & \\ \hline Z_n(A) & & & \end{pmatrix}$$

Für jede der n Matrizen auf der rechten Seite machen wir weiter mit der zweiten Zeile, und so weiter, bis wir $d(A)$ als Summe von n^n Determinanten von Matrizen Q_k schreiben, wobei jede Zeile einer solcher Matrix $n-1$ Einträge gleich 0 hat, und einen Eintrag gleich mit einem Eintrag a_{ij} aus A. Aus (D1) könne wir diese Skalare vorne bringen. Aus Bemerkung 7.41 sind nur die Determinanten nicht Null, die Permutationsmatrizen entsprechen; und jede Permutationsmatrix kommt genau ein Mal vor. Wir haben also aus Bemerkung 7.40, dass

$$\begin{aligned} d(A) &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} d(P_\sigma) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \end{aligned}$$

Q.E.D.

Der Ausdruck der Determinante in (7.13) ist als **Leibnizformel** bekannt.

[4] 20.4.'21

Wir können also ab jetzt "offiziell" von **der** Determinante über einer Matrix in $\operatorname{Mat}_n(R)$ sprechen.

7.5 Und wieder Eigenschaften von Determinanten, dieses Mal aber über Ringe

Satz 7.43. *Für die Determinante* $\det : \mathrm{Mat}_n(R) \longrightarrow R$ *gilt für alle* $A, B \in \mathrm{Mat}_n(R)$*:*

(a) $\det A = \det A^\top$.

(b) $\det(AB) = (\det A)(\det B)$

Beweis-Skizze: (a) Wir haben $A^\top = (a_{ij}^\top)$ mit $a_{ij}^\top = a_{ji}$ für alle $i, j \in \{1, \ldots, n\}$. Es gilt also

$$\begin{aligned}
\det A^\top &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)}^\top \cdots a_{n\sigma(n)}^\top \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \\
&= \det A.
\end{aligned}$$

Die dritte Gleichheit folgt weil R kommutativ und σ bijektiv ist. Die vierte weil $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ und weil $\sigma \mapsto \sigma^{-1}$ eine bijektive Selbstabbildung von S_n.

(b) Die Einträge von AB sind $(AB)_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$. Es gilt also

$$\begin{aligned}
Z_i(AB) &= \begin{pmatrix} \sum_{k=1}^n a_{ik} b_{k1} & \ldots & \sum_{k=1}^n a_{ik} b_{kn} \end{pmatrix} \\
&= \sum_{k=1}^n a_{ik} \begin{pmatrix} b_{k1} & \ldots & b_{kn} \end{pmatrix} \\
&= \sum_{k=1}^n a_{ik} Z_k(B).
\end{aligned}$$

Es gilt also

$$\begin{aligned}\det(AB) &= \det\begin{pmatrix}\sum_{k=1}^n a_{1k}Z_k(B)\\ Z_2(AB)\\ \vdots\\ Z_n(AB)\end{pmatrix} = \sum_{k=1}^n a_{1k}\cdot\det\begin{pmatrix}Z_k(B)\\ Z_2(AB)\\ \vdots\\ Z_n(AB)\end{pmatrix}\\ &= \sum_{1\le k_1,\ldots,k_n\le n} a_{1k_1}\cdots a_{nk_n}\cdot\det\begin{pmatrix}Z_{k_1}(B)\\ Z_{k_2}(B)\\ \vdots\\ Z_{k_n}(B)\end{pmatrix}\end{aligned}$$

Die Summe hat ein Summand für jedes n-Tupel $(k_1,\ldots,k_n)\in\{1,\ldots,n\}^n$. Wenn es aber $i\neq j$ mit $k_i=k_j$ gibt, dann ist nach (D2)

$$\det\begin{pmatrix}Z_{k_1}(B)\\ Z_{k_2}(B)\\ \vdots\\ Z_{k_n}(B)\end{pmatrix}=0.$$

Wir brauchen also nur die Summanden mit $\{k_1,\ldots,k_n\}=\{1,\ldots,n\}$ zu betrachten. Jeder Summand entspricht dann der Permutation $\sigma:\mathbb{N}\longrightarrow\mathbb{N}$ gegeben durch $\sigma(i)=k_i$. Wir bekommen also, dass

$$\begin{aligned}\det(AB) &= \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)}\cdot\det\begin{pmatrix}Z_{\sigma(1)}(B)\\ Z_{\sigma(2)}(B)\\ \vdots\\ Z_{\sigma(n)}(B)\end{pmatrix}\\ &= \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)}\cdot\det(P_\sigma\cdot B)\\ &= \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)}\cdot\operatorname{sgn}(\sigma)\cdot\det B\\ &= (\det A)(\det B),\end{aligned}$$

wobei die dritte Gleichheit aus Bemerkung 7.40 folgt. Q.E.D.

Wir erinnern, dass wir für eine Matrix $A\in\mathrm{Mat}_n(R)$ und $i,j\in\{1,\ldots,n\}$ wir durch A_{ij} die $(n-1)\times(n-1)$ Untermatrix von A bezeichnen, die durch Streichen der i-ten Zeile und der j-ten Spalte entsteht.

Wir haben in Definition 7.3 die Determinante rekursiv eingeführt. Jetzt zeigen wir, dass man die rekursive Formel für jede Spalte oder Zeile gilt.

Satz 7.44 (Laplacescher Entwicklungssatz)**.** *Für alle $A=(a_{ij})\in\mathrm{Mat}_n(R)$ und alle $k\in\{1,\ldots,n\}$*

gilt

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} \cdot a_{ik} \cdot \det A_{ik} \tag{7.14}$$

$$= \sum_{k=1}^{n} (-1)^{i+k} \cdot a_{kj} \cdot \det A_{kj}. \tag{7.15}$$

Beweis-Skizze: Satz 7.10 zeigt, dass (7.14) für $k = 1$ gilt. Aus Satz 7.43 (a) gilt auch (7.15) für $k = 1$. Für beliebige k müssen wir die k-Zeile an der ersten Stelle bringen, und die anderen in derselben Reihenfolge lassen. Das wird durch Multiplikation mit

$$U_{1\leftrightarrow 2} \cdot U_{2\leftrightarrow 3} \cdot \cdots \cdot U_{k-1\leftrightarrow k}.$$

Aus Bemerkung 7.40 und Satz 7.43 (b) bekommen wir also die $(-1)^{k-1}$ als Vorfaktor.

Q.E.D.

7.6 Die Cramer'sche Regel(n), Minoren und der Rang

Das erste Ziel ist die Inverse einer Matrix mit Hilfe von Determinanten zu beschreiben. Dafür werden wir zu erst Minoren für beliebige Matrizen Definieren.

Definition 7.45. Sei $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(R)$ eine (nicht unbedingt quadratische) Matrix. Sei $t \in \mathbb{N}$ mit $1 \leq t \leq \min\{m, n\}$. Für die Zeilen-Indizes $1 \leq i_1 < \cdots < i_t \leq m$ und die Spalten-Indizes $1 \leq j_1 < \cdots < j_t \leq n$ bezeichnen wir die Untermatrix von A die nur diese Indizes nimmt durch

$$A_{[i_1 \ldots i_t | j_1 \ldots j_t]}.$$

Der **Minor** von A mit Zeilen-Indizes $1 \leq i_1 < \cdots < i_t \leq m$ und Spalten-Indizes $1 \leq j_1 < \cdots < j_t \leq n$ ist die Determinante der entsprechenden Untermatrix. Wir bezeichnen das durch

$$m_{[i_1 \ldots i_t | j_1 \ldots j_t]} = \det A_{[i_1 \ldots i_t | j_1 \ldots j_t]}.$$

Die Zahl t heißt die Ordnung des Minor. Man sagt, auch dass $m_{[i_1 \ldots i_t | j_1 \ldots j_t]}$ ein $t \times t$-Minor oder ein t-Minor von A ist.

Beispiel 7.46.

1. Wenn $t = 1$ dann sind die 1-Minoren einfach die Einträge der Matrix.
2. Die 2-Minoren von $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ sind $m_{[1,2|1,2]} = -1, \quad m_{[1,2|1,3]} = -6, \quad m_{[1,2|2,3]} = -3.$
3. Eine 4×4-Matrix hat aber $\binom{4}{2} \cdot \binom{4}{2} = 36$ Minoren. Zum Beispiel

$$m_{[2,3|1,3]} = a_{2,1}a_{3,3} - a_{2,3}a_{3,1}.$$

Eine der wichtigsten Anwendungen von Minoren ist folgende Beschreibung des Ranges einer Matrix dadurch.

Satz 7.47. *Sei $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$ eine Matrix mit Einträgen in einem Körper. Dann gilt*

$$\mathrm{Rang}\, A = \max\{t \in \mathbb{N} \ : \ \exists\ t - \textit{Minor} \neq 0\}.$$

Beweis-Skizze: Wenn $m_{[\mathbf{i}|\mathbf{j}]} \neq 0$ ein nicht-trivialer t-Minor ist, dann hat die $t \times n$-Teilmatrix von A mit Zeilenindizes in $\mathbf{i}$ maximalen Rang, uns somit auch die Matrix A hat Rang mindestens t. Das zeigt $\operatorname{Rang} A \geq \max\{t \in \mathbb{N} \ : \ \exists\, t - \text{Minor} \neq 0\}$.
Sei jetzt $t > \operatorname{Rang} A$, und sei A' eine $t \times t$-Teilmatrix von A. Dann ist eine Zeilenumformung von A' eine $t \times t$-Teilmatrix von $\mathrm{RZ}(A)$. Diese letzte Matrix hat aber nur $\operatorname{Rang} A$ nicht-triviale Zeilen, und somit muss eine Zeilenumformung von A' eine Nullzeile haben. Das heißt $\det A' = 0$. Das zeigt die andere Ungleichung. Q.E.D.

‘

In diesem Teil brauchen wir nur die $(n-1) \times (n-1)-$Minoren von quadratischen Matrizen. In diesem Fall ist es günstiger die Indizes die wir weglassen aufzulisten, anstatt die die wir brauchen. Wir schreiben also wie bis her $A_{ij} = A_{[1\ldots\widehat{i}\ldots n|1\ldots\widehat{j}\ldots n]}$, und in Konsequenz schreiben wir auch

$$m_{ij} := \det A_{ij}.$$

Wir sammeln alle diese $(n-1)-$Minoren, transponiert und mit alternierenden Vorzeichen, in einer neuen Matrix:

Definition 7.48. Sei $A \in \operatorname{Mat}_n(R)$. Die **komplementäre Matrix** zu A (oder die **klassische Adjugierte** von A) ist die Matrix $\widetilde{A} = (\widetilde{a}_{ij}) \in \operatorname{Mat}_n(R)$ gegeben durch

$$\widetilde{a}_{ij} := (-1)^{i+j} m_{ji} = (-1)^{i+j} \det A_{ji}.$$

! Beachten Sie die vertauschte Reihenfolge von i und j!

Beispiel 7.49. Wenn $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \operatorname{Mat}_2(R)$, dann ist $\widetilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Man kann in diesem Fall beobachten, dass $A \cdot \widetilde{A} = \det(A) \cdot I_2$. Das gilt nicht nur für $n = 2$, sondern ganz allgemein.

Proposition 7.50. *Für jede Matrix $A \in \operatorname{Mat}_n(R)$ gilt*

$$A \cdot \widetilde{A} = \widetilde{A} \cdot A = (\det A) \cdot I_n.$$

Beweis-Skizze: Für $i, j \in \{1, \ldots, n\}$ definieren wir die Matrix

$$B^{(ij)} = \mathbf{i} \begin{pmatrix} a1\ 1 & \ldots & a_{1\ j-1} & a_{1\ j} & a_{1\ j+1} & \ldots & a_{1\ n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ ai-1\ 1 & \ldots & a_{i-1\ j-1} & a_{i-1\ j} & a_{i-1\ j+1} & \ldots & a_{i-1\ n} \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ ai+1\ 1 & \ldots & a_{i+1\ j-1} & a_{i+1\ j} & a_{i+1\ j+1} & \ldots & a_{i+1\ n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ an\ 1 & \ldots & a_{n\ j-1} & a_{n\ j} & a_{n\ j+1} & \ldots & a_{n\ n} \end{pmatrix} = \begin{pmatrix} Z_1(A) \\ \vdots \\ Z_{i-1}(A) \\ \hline e_j \\ \hline Z_{i+1}(A) \\ \vdots \\ Z_n(A) \end{pmatrix}$$

(Spalte $\mathbf{j}$ markiert.)

wobei $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$, mit der Eins an der Stelle j. Es gilt

$$A^{(ij)} = U_{\mathbf{1} \to \mathbf{1} + (-a_{1\ j}) \cdot \mathbf{i}} \cdot \cdots \cdot U_{\mathbf{n} \to \mathbf{n} + (-a_{n\ j}) \cdot \mathbf{i}} A^{(ij)},$$

also aus Bemerkung 7.13 und Proposition 7.14, dass

$$\det A^{(ij)} = \det B^{(ij)}.$$

Sei jetzt $A \cdot \widetilde{A} = (c_{ij})$. Wir haben

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} a_{ik} \cdot \widetilde{a}_{kj} = \sum_{k=1}^{n} a_{ik} \cdot \det A^{(jk)} = \sum_{k=1}^{n} a_{ik} \cdot \det B^{(jk)} \\ &= \sum_{k=1}^{n} a_{ik} \cdot \det \left(\begin{array}{c} Z_1(A) \\ \vdots \\ Z_{j-1}(A) \\ \hline \mathbf{e}_k \\ \hline Z_{j+1}(A) \\ \vdots \\ Z_n(A) \end{array}\right) = \det \left(\begin{array}{c} Z_1(A) \\ \vdots \\ Z_{j-1}(A) \\ \hline \sum_{k=1}^{n} a_{ik} \cdot \mathbf{e}_k \\ \hline Z_{i+1}(A) \\ \vdots \\ Z_n(A) \end{array}\right) = \det \left(\begin{array}{c} Z_1(A) \\ \vdots \\ Z_{j-1}(A) \\ \hline Z_i(A) \\ \hline Z_{j+1}(A) \\ \vdots \\ Z_n(A) \end{array}\right) \\ &= \delta_{ij} \det A. \end{aligned}$$

Für $\widetilde{A} \cdot A$ kann man genauso berechnen, oder, aus $(A^\top \cdot \widetilde{A^\top}) = \det(A^\top) \cdot I_n$, weil $\widetilde{A^\top} = (\widetilde{A})^\top$, haben wir $(\widetilde{A} \cdot A)^\top = \det A^\top \cdot I_n$, und wir transponieren nochmals. Q.E.D.

Wir haben folgendes Korollar.

Korollar 7.51 (Cramer 1)**.** Eine Matrix $A \in \mathrm{Mat}_n(R)$ ist invertierbar genau dann, wenn $\det A \in R$ invertierbar ist. In diesem Fall, ist die Inverse

$$A^{-1} = (\det A)^{-1} \cdot \widetilde{A}.$$

Insbesondere, gibt die Einschränkung der Determinante ein Gruppenhomomorphismus

$$\det : \mathrm{GL}_n(R) \longrightarrow R^\times.$$

Als Anwendung können wir die Lösung eines LGS mit eindeutiger Lösung durch Determinanten beschreiben.

Korollar 7.52. Sei $A \in \mathrm{GL}_n(\mathbb{K})$ und $\mathbf{b} \in \mathbb{K}^n$. Also $A \cdot \mathbf{x} = \mathbf{b}$ hat eine eindeutige Lösung: $\mathbf{x} = A^{-1}\mathbf{b}x$. Es gilt

$$x_i = \frac{\det B_i}{\det A},$$

wobei, wenn $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathrm{Mat}_{n,1}(\mathbb{K})$ die Spalten von A sind, dann ist

$$B_i = \left(\begin{array}{c|c|c|c|c|c|c} \mathbf{a}_1 & \ldots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \ldots & \mathbf{a}_n \end{array}\right)$$

Beweis-Skizze: Aus Korollar 7.51 haben wir

$$\det A \cdot x_i = \sum_{j=1}^{n} (-1)^{i+j} \det A_{ij} b_j = \det B_i.$$

Q.E.D.

7.7 Determinanten von Endomorphismen

In diesem Teil arbeiten wir wieder über einem Körper $\mathbb{K}$. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und sei $f : V \longrightarrow V$ eine $\mathbb{K}$-lineare Abbildung. Für jede geordnete Basis $B = \{v_1, \ldots, v_n\}$ von V ist $M_B(f) = (b_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ die zugeordnete Matrix von f bezüglich der Basis B, also für alle $j \in \{1, \ldots, n\}$ gilt

$$f(v_j) = \sum_{i=1}^{n} b_{ij} v_i.$$

Wenn $B' = \{v'_1, \ldots, v'_n\}$ eine andere geordnete Basis von V ist, dann bezeichnet $T := M_{B'}^{B} = (t_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ die Basiswechselmatrix von B nach B'. Es gilt also

$$T \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix}.$$

Lemma 7.53. *Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und seien B und B' zwei geordnete Basen von V. Es gilt*

$$\det M_B(f) = \det M_{B'}(f).$$

Beweis-Skizze: Wir bezeichnen mit T die Basiswechselmatrix. Aus der "Lineare Algebra 1" Vorlesung (Satz 4.27 im Skript) haben wir

$$M_{B'}(f) = T^{-1} \cdot M_B(f) \cdot T.$$

Es folgt also aus Satz 7.43 Teil (b), dass

$$\begin{aligned} \det M_{B'}(f) &= \det \left(T^{-1} \cdot M_B(f) \cdot T\right) = \det \left(T^{-1}\right) \cdot \det \left(M_B(f)\right) \cdot \det (T) = \\ &= \det (T)^{-1} \cdot \det \left(M_B(f)\right) \cdot \det (T) = \det M_B(f). \end{aligned}$$

Q.E.D.

Definition 7.54. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Die **Determinante Abbildung** $\det : \mathrm{End}_{\mathbb{K}} \longrightarrow \mathbb{K}$ ist definiert durch

$$\det f := \det M_f^B.$$

für eine beliebige geordnete Basis B von B.

Aus Lemma 7.53 ist die obige Definition unabhängig von der Wahl der geordneten Basis B.

Kapitel 8

Der Dualraum

Sei $\mathbb{K}$ ein Körper. Wir untersuchen in diesem Kapitel, in dem Fall der $\mathbb{K}$-Vektorräumen, das Zusammenspiel zwischen einem Raum (oder allgemeiner einem geometrischen Objekt) und die kompatiblen Funktionen die darauf definiert sind. Insbesondere im endlich-dimensionalen Fall ist der Zusammenhang sehr eng. Wir fangen mit der Definition von "kompatiblen Funktion" in unserem Fall: Linearform.

Warum ist das interessant? Man kann sich vorstellen, dass Dualräume die unterliegenden Ausdrücke der linearen Gleichungen verallgemeinern: Genau wie ein beliebiger Vektor v die Verallgemeinerung eines Vektors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ ist, ist eine Linearform die Verallgemeinerung eines Ausdruckes der Form $a_1x_1 + \cdots + a_nx_n$ mit $a_i \in \mathbb{R}$. Im Lösen eines LGS (mit dem Gaußschen Algorithmus z.B.) werden lineare Gleichungen addiert und mit Skalare multipliziert. Genau diese Operationen entsprechen der Vektorraum-Struktur des Dualraumes. Eines unserer Ziele ist eine Methode zu entwickeln um einen Untervektorraum als Lösungsmenge eines linearen Gleichungssystems zu beschreiben.

8.1 Der Vektorraum der Linearformen

Definition 8.1. Sei V ein $\mathbb{K}$-Vektorraum. Eine **Linearform** auf V (oder ein **lineares Funktional** auf V, oder eine **1-Form** auf V) ist eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow \mathbb{K}$. Das heißt $f \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$.

Definition 8.2. Sei V ein $\mathbb{K}$-Vektorraum. Der(algebraische) **Dualraum** von V ist der $\mathbb{K}$-Vektorraum

$$V^* := \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K}).$$

Beispiele:

1. Seien $a_1, \ldots, a_n \in \mathbb{R}$. Die Abbildung $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ gegeben durch
$$f(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n$$
ist eine Linearform auf $\mathbb{R}^n$.
2. Die Abbildung $f := \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit $f(x, y) = x^2 + y^2$ ist nicht eine Linearform.
3. Sei V der $\mathbb{R}$-Vektorraum der stetigen Funktionen $f : [a, b] \longrightarrow \mathbb{R}$ (mit $a < b \in \mathbb{R}$). Die Funktion $\varphi : V \longrightarrow \mathbb{R}$ gegeben durch
$$\varphi(f) := \int_a^b f(t)dt, \quad \forall\ f \in V,$$

eine Linearform auf V.

4. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, mit Basis $B = \{v_1, \ldots, v_n\}$. Für jeden Index $i = 1, \ldots, n$ definieren wir die Linearform $v_i^* \in V^*$ durch

$$v_i^*(v_j) := \delta_{ij}.$$

Das reicht um eine Linearform eindeutig zu bestimmen. Diese Linearformen werden **Koordinatenfunktionen** bezüglich der Basis B genannt. Konkreter: Auf $\mathbb{R}^n$ sind die Koordinatenfunktionen bezüglich der kanonischen Basis $\{e_1, \ldots, e_n\}$ die Abbildungen

$$e_i^* : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \text{mit } e_i^*(x_1, \ldots, x_n) = x_i, \quad \forall i = 1, \ldots, n.$$

Satz 8.3. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Die Menge $B^* = \{v_1^*, \ldots, v_n^*\}$ der Koordinatenfunktionen bezüglich der Basis B auf V ist eine Basis des Dualraumes V^*. Insbesondere ist V^* auch endlichdimensional mit*

$$\dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V.$$

Beweis-Skizze: lineare Unabhängigkeit Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, sodass

$$\lambda_1 v_1^* + \cdots + \lambda_n v_n^* = \mathbf{0}_{V^*}. \tag{8.1}$$

Für jeden $j = 1, \ldots, n$ berechnen wir

$$(\lambda_1 v_1^* + \cdots + \lambda_n v_n^*)(v_j) = \lambda_j v_j^*(v_j) = \lambda_j.$$

Aus (8.1) folgt also $\lambda_j = \mathbf{0}_{V^*}(v_j) = 0$.
Erzeugendensystem Sei $f \in V^*$, und seien $\lambda_i := f(v_i) \in \mathbb{K}$, $\forall\ i = 1, \ldots, n$. Wir wenden zeigen, dass

$$f = \lambda_1 v_1^* + \cdots + \lambda_n v_n^*. \tag{8.2}$$

Es reicht zu zeigen, dass beide Seiten in (8.2) denselben Wert auf Elementen aus B nehmen. Es ist offensichtlich klar aus den Definitionen von λ_i und v_i^*, dass

$$f(v_j) = (\lambda_1 v_1^* + \cdots + \lambda_n v_n^*)(v_j), \quad \forall\ j = 1, \ldots, n.$$

Q.E.D.

Definition 8.4. Sei $B = \{v_1, \ldots, v_n\}$ eine Basis des $\mathbb{K}$-Vektorraumes V. Die **Dualbasis** von B ist die Menge $B^* = \{v_1^*, \ldots, v_n^*\}$ der Koordinatenfunktionen bezüglich der Basis B.

Beispiel 8.5. Es ist wichtig zu betonen, dass jedes v_i^* von den anderen v_j, also von ganz B abhängig ist. Um die Notation nicht unnötig zu belasten, schreiben wir aber keinen weiteren Index dazu (wie "$v_{B,i}^*$"). Hier ist aber ein konkretes Beispiel als Warnung. Seien $e_1, e_2 \in \mathbb{R}^2$ die Vektoren der kanonischen Basis und sei $v = (1, 1)$. Wenn wir $B = \{e_1, e_2\}$ dualisieren, dann gilt:

$$e_1^*(e_1) = 1 \quad e_1^*(e_2) = 0 \quad \Rightarrow \quad e_1^*(v) = e_1^*(e_1 + e_2) = 1 + 0 = 1.$$

Wenn wir aber $B' = \{e_1, v\}$ dualisieren, dann haben wir per Definition

$$e_1^*(v) = 0.$$

Das ist kein Widerspruch, es sind einfach zwei verschiedene Linearformen die gleich bezeichnet werden.

Man kann Satz 8.3 wie folgt stärken:
Sei V ein (nicht unbedingt endlichdimensionaler) $\mathbb{K}$-Vektorraum und sei B eine Basis von V. Dann ist die Menge B^* der Koordinatenfunktionen bezüglich B eine linear unabhängige Menge in V^*. Weiterhin, B^* ist eine Basis von V^* genau dann, wenn V endlichdimensional ist.

Das schwierige an dieser Verallgemeinerung ist zu zeigen, dass B^* kein Erzeugendensystem ist, falls V unendlichdimensional ist. Die Idee ist, dass das Funktional $f : V \longrightarrow \mathbb{K}$, das durch $f(v_i) = 1,\ \forall\ v_i \in B$ bestimmt ist, kann nicht als Linearkombination[1] von Elementen aus B^* geschrieben werden.
Konkreter kann man dieses Verhalten durch in folgendes Beispiel sehen: Sei $V = \{(a_n)_{n\in\mathbb{N}} \ : \ a_n \in \mathbb{R}\ \forall\ n \in \mathbb{N}$ und $\exists n_0 \in \mathbb{N}$ sodass $a_n = 0\ \forall\ n > n_0\}$ der $\mathbb{R}$-Vektorraum aller endlichen folgen. Insbesondere, ist $V = \bigoplus_{i\in\mathbb{N}} \mathbb{R}$, und es hat V eine abzählbare $\mathbb{R}$-Basis: $\{e_i\}_{i\in\mathbb{N}}$. Der Dualraum ist $V^* = \{f : V \longrightarrow R \ : \ f((a_n)) = \sum_{n\in\mathbb{N}} b_n \cdot a_n$ mit $b_n \in \mathbb{N}$ beliebig$\}$. Der Dualraum ist also zum direkten Produkt $\prod_{i\in\mathbb{N}} \mathbb{R}$ isomorph. Man kann beweisen, dass V^* keine abzählbare Basis besitzen kann, also, dass $V \ncong V^*$.

Beispiele:

1. Wir identifizieren $\mathbb{K}^n \simeq \mathrm{Mat}_{n,1}(\mathbb{K})$ mit Spaltenmatrizen. Wir schreiben dann die Vektoren als Spalten auf, und haben die kanonische Basis

$$B = \left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad , e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

   Der Dualraum entspricht $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K})$, und das ist isomorph zu $\mathrm{Mat}_{1,n}(\mathbb{K})$. Dann ist die Dualbasis

$$B^* = \left\{ e_1^* = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}, \dots, e_n^* = \begin{pmatrix} 0 & 0 & \dots & 1 \end{pmatrix} \right\}.$$

2. Mit der obigen Vereinbarung, dass Vektoren von $\mathbb{K}^n$ Spalten sind und Vektoren von $(\mathbb{K}^n)^*$ Zeilen, hat man: Wenn

$$B = \left(\begin{array}{c|c|c|c} & & & \\ v_1 & v_2 & \dots & v_n \\ & & & \end{array} \right) \in \mathrm{Mat}_n(\mathbb{K})$$

   eine Matrix deren Spaltenmenge eine Basis von $\mathbb{K}^n$ ist, dann entsprechen die Zeilen der Inverse von B zu der Dualbasis. Die Interpretation von $\mathrm{Mat}_{1,n}(\mathbb{K})$ als Linearformen auf $\mathrm{Mat}_{n,1}(\mathbb{K})$ ist durch das Produkt von Matrizen.

[1] Für uns sind alle Linearkombinationen endlich.

8.2 Der Dualraum des Dualraumes

Wir werden jetzt zeigen, dass wenn V endlichdimensional ist, dann gilt $(V^*)^* \simeq V$. Wir schreiben einfach V^{**} für $(V^*)^*$. Wir definieren zu erst eine Abbildung $\mathrm{ev} : V \longrightarrow V^{**}$. Dafür müssen wir für jeden Vektor $v \in V$ eine Linearform $\mathrm{ev}_v : V^* \longrightarrow \mathbb{K}$ definieren. Das heißt, wir müssen jedem $f \in V^*$ ein Element aus $\mathbb{K}$ vorschreiben. Die kanonische Wahl ist:

$$\mathrm{ev}_v(f) := f(v) \in \mathbb{K}, \quad \forall\, f \in V^* \text{ und } \forall\, v \in V.$$

Wir haben eine Abbildung $\mathrm{ev}_v : V^* \longrightarrow \mathbb{K}$ definiert. Wir müssen noch zeigen, dass ev_v eine Linearform auf V^* für alle $v \in V$ ist. Wir haben:

$$\begin{aligned} \mathrm{ev}_v(f+g) &= (f+g)(v) = f(v) + g(v) = \mathrm{ev}_v(f) + \mathrm{ev}_v(g). \\ \mathrm{ev}_v(\lambda f) &= (\lambda f)(v) = \lambda(f(v)) = \lambda \cdot \mathrm{ev}_v(f). \end{aligned}$$

Wir haben also eine Abbildung $\mathrm{ev} : V \longrightarrow V^{**}$ definiert. Weiterhin gilt:

$$\begin{aligned} \mathrm{ev}_{v_1+v_2}(f) &= f(v_1+v_2) = f(v_1) + f(v_2) = \mathrm{ev}_{v_1}(f) + \mathrm{ev}_{v_2}(f), \quad \forall\, v_1, v_2 \in V, f \in V^*, \\ \mathrm{ev}_{\lambda v}(f) &= f(\lambda v) = \lambda f(v) = \lambda \cdot \mathrm{ev}_v(f), \quad \forall\, \lambda \in \mathbb{K}, v \in V, f \in V^*, \end{aligned}$$

also $\mathrm{ev} : V \longrightarrow V^{**}$ ist $\mathbb{K}$-linear.

Satz 8.6. *Für jeden endlichdimensionalen $\mathbb{K}$-Vektorraum V ist die oben definierte Abbildung* $\mathrm{ev} : V \longrightarrow V^{**}$ *einen Isomorphismus von $\mathbb{K}$-Vektorräume.*

Beweis-Skizze: Wir haben oben gezeigt, dass $\mathrm{ev} \in \mathrm{Hom}_{\mathbb{K}}(V, V^{**})$. Wir brauchen noch, dass ev bijektiv ist. Weil $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V^{**}$ (Satz 8.3), reicht es nur die Injektivität zu zeigen.
Wir zeigen dafür, dass $\ker \mathrm{ev} = \{0\}$. Sei $v \in \ker \mathrm{ev}$. Das heißt, $\mathrm{ev}_v(f) = 0$ für alle $f \in V^*$. Wir nehmen an, dass $v \neq 0$, und werden einen Widerspruch finden. Wir haben $v \neq 0 \iff \{v\}$ ist linear unabhängig. Es existiert also eine Basis B von V mit $v \in B$. Dann ist aber v^* ein Element der dualen Basis $B^* \subseteq V^*$, mit $v^*(v) = 1 \neq 0$ - ein Widerspruch ↯ zu $v \in \ker \mathrm{ev}$. Q.E.D.

8.3 Duale Homomorphismen

Sei $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, und seien V^* und W^* die Dualräume von V, beziehungsweise W. Wenn $\alpha \in W^*$, dass heißt $\alpha \in \mathrm{Hom}_{\mathbb{K}}(W, \mathbb{K})$, dann ist $\alpha \circ f \in V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. Der Grund ist, dass die Verknüpfung zweier $\mathbb{K}$-linearen Abbildungen wieder $\mathbb{K}$-linear ist. Die Abbildung

$$f^* : W^* \longrightarrow V^* \quad \text{gegeben durch } f^*(\alpha) := \alpha \circ f, \quad \forall\, \alpha \in W^*, \tag{8.3}$$

ist also wohldefiniert. Wir haben

$$\begin{aligned} f^*(\alpha+\beta) &= (\alpha+\beta) \circ f = \alpha \circ f + \beta \circ f = f^*(\alpha) + f^*(\beta), \quad \forall\, \alpha, \beta \in W^*, \\ f^*(\lambda\alpha) &= (\lambda\alpha) \circ f = \lambda(\alpha \circ f) = \lambda f^*(\alpha), \quad \forall\, \lambda \in \mathbb{K} \text{ und } \alpha \in W^*. \end{aligned}$$

Also $f^* \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$. Wir können das wie folgt zusammenfassen.

Definition 8.7. Seien V, W zwei $\mathbb{K}$-Vektorräume und sei $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Die **duale $\mathbb{K}$-lineare Abbildung** von f (oder das **duale Homomorphismus** von f) ist die Abbildung f^*: $W^* \longrightarrow V^*$ gegeben durch (8.3).

Lemma 8.8. *Seien V, W, U drei $\mathbb{K}$-Vektorräume und seien $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ und $g \in \mathrm{Hom}_{\mathbb{K}}(W, U)$ zwei verknüpfbare Homomorphismen von Vektorräumen. Es gelten*

(i) $(g \circ f)^* = f^* \circ g^*$.

(ii) $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$.

Beweis-Skizze: Übung. Q.E.D.

Satz 8.9. *Seien V, W zwei endlich dimensionale $\mathbb{K}$-Vektorräume, und sei $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Sei $A = \{v_1, \ldots, v_n\}$ eine Basis von V und $B = \{w_1, \ldots, w_m\}$ eine Basis von W. Dann gilt*

$$M_{A^*}^{B^*}(f^*) = \left(M_B^A(f)\right)^\top.$$

Beweis-Skizze: Sei $M_{A^*}^{B^*}(f^*) = (b_{ij})$ und $M_B^A(f) = a_{ij}$. Wir haben für jeden $i \in \{1, \ldots, n\}$ und $j \in \{1, \ldots, m\}$. Aus der Definition von zugeordneter Matrix und von dualer Basis haben wir

$$f^*(w_j^*)(v_i) = w_j^*(f(v_i)) = w_j^*\left(\sum_{k=1}^{m} a_{ki} w_k\right) = \sum_{k=1}^{m} a_{ki} w_j^*(w_k) = \sum_{k=1}^{m} a_{ki}\delta_{kj} = a_{ji}.$$

Wenn wenn wir aber zu erst $f^*(w_j^*)$ mit Hilfe der zugeordneten Matrix ausdrücken, dann bekommen wir

$$f^*(w_j^*)(v_i) = \left(\sum_{k=1}^{n} b_{kj} v_k^*\right)(v_i) = \sum_{k=1}^{m} b_{kj} v_k^*(v_i) = \sum_{k=1}^{n} b_{kj}\delta_{ki} = b_{ij}.$$

Q.E.D.

Korollar 8.10. Seien V und W endlich dimensionale $\mathbb{K}$-Vektorräume und $f : V \longrightarrow W$ eine $\mathbb{K}$-lineare Abbildung. Dann gilt: $\mathrm{Rang}\, f = \mathrm{Rang}\, f^*$.

Beweis-Skizze: Es folgt direkt aus Satz 8.9 und aus $\mathrm{Rang}\, A = \mathrm{Rang}\, A^\top$. Diese letzte Gleichheit gilt, weil der Spaltenraum von A gleich mit dem Zeilenraum von $A^\top$ ist. Q.E.D.

Satz 8.11. *Sei $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ ein $\mathbb{K}$-Vektorraumhomomorphismus zwischen endlich dimensionale $\mathbb{K}$-Vektorräume, und sei $f^* \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$ das duale Homomorphismus von f. Es gelten*

(a) *Wenn f surjektiv ist, dann ist f^* injektiv.*

(b) *Wenn f injektiv ist, dann ist f^* surjektiv.*

Beweis-Skizze: Variante 1: (a) Wir haben

$$\dim_{\mathbb{K}} W^* \overset{①}{=} \dim_{\mathbb{K}} W \overset{②}{=} \text{Rang } f \overset{③}{=} \text{Rang } f^* \overset{④}{=} \dim_{\mathbb{K}} \text{Bild } f^* \overset{⑤}{=} \dim_{\mathbb{K}} W^* - \dim_{\mathbb{K}} \text{Ker } f^*.$$

Wobei die Begründungen sind die folgenden:

① Satz 8.3, ② f = surjektiv, ③ Korollar 8.10, ④ Definition ⑤ Dimensionssatz.

Es gilt also $\dim_{\mathbb{K}} \text{Ker } f^* = 0$.
(b) Mit einer ähnlichen Strategie haben wir

$$\dim_{\mathbb{K}} \text{Bild } f^* \overset{ⓐ}{=} \text{Rang } f^* \overset{ⓑ}{=} \text{Rang } f \overset{ⓒ}{=} \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} \text{Ker } f \overset{ⓓ}{=} \dim_{\mathbb{K}} V \overset{ⓔ}{=} \dim_{\mathbb{K}} V^*.$$

Mit den Begründungen:

ⓐ Definition ⓑ Korollar 8.10, ⓒ Dimensionssatz, ⓓ f = injektiv, ⓔ Satz 8.3.

Also Bild $f^* = V^*$.
Variante 2:
(a) Seien $\alpha, \beta \in W^*$, sodass $f^*(\alpha) = f^*(\beta)$. Das heißt, $\alpha \circ f = \beta \circ f$. Weil f surjektiv ist folgt , dass es eine Abbildung $f'' : W \longrightarrow V$ gibt, sodass $f \circ f'' = \text{id}_W$. Es folgt also

$$(\alpha \circ f) \circ f'' = (\beta \circ f) \circ f'' \iff \alpha \circ (f \circ f'') = \beta \circ (f \circ f'') \iff \alpha \circ \text{id}_W = \beta \circ \text{id}_W \iff \alpha = \beta.$$

(b) Weil f injektiv ist folgt, dass V isomorph zu Bild(f) ist. Also wenn $\{v_1, \ldots, v_n\}$ eine Basis von V ist, dann ist $\{f(v_1), \ldots, f(v_n)\}$ eine Basis des $\mathbb{K}$-UVR Bild$(f) \subseteq W$. Insbesondere ist $\{f(v_1), \ldots, f(v_n)\}$ linear unabhängig in W, und wir ergänzen es zu einer Basis von W:

$$B = \{f(v_1), \ldots, f(v_n), w_{n+1}, \ldots, w_m\}.$$

Sei jetzt $\alpha \in V^*$. Wir definieren $\beta : W \longrightarrow \mathbb{K}$ indem wir die Werte auf der Basis B angeben:

$$\begin{cases} \beta(f(v_i)) := \alpha(v_i) & i = 1, \ldots, n \\ \beta(w_j) := 0 & j = (n+1), \ldots, m. \end{cases}$$

Wir haben dann $f^*(\beta) = f \circ \beta = \alpha$. Q.E.D.

8.4 Der Annullator eines Unterraumes

Definition 8.12. Sei V ein $\mathbb{K}$-Vektorraum. Der **Annullator** des $\mathbb{K}$-Unterrvektorraumes $U \subseteq V$ ist der $\mathbb{K}$-UVR von V^*:

$$U^0 := \{f \in V^* \ : \ f|_U = 0\} \subseteq V^*.$$

Es ist sehr einfach zu sehen, dass U^0 tatsächlich ein $\mathbb{K}$-UVR ist wenn U ein $\mathbb{K}$-UVR ist.
Wir bezeichnen[2] mit $Gr(V) := \{U \subseteq_{\mathbb{K}-UVR} V\}$ die Menge aller $\mathbb{K}$-Untervektorräume von V.

[2] Die Bezeichnung Gr kommt von Hermann Graßmann (1809-1877).

Satz 8.13. *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, und die Abbildung*

$$\Psi : Gr(V) \longrightarrow Gr(V^*), \quad \text{gegeben durch } U \longmapsto U^0.$$

Es gelten:

(a) *Wenn $U_1 \subseteq U_2$, dann $U_1^0 \supseteq U_2^0$.*

(b) $\dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^0 = \dim_{\mathbb{K}} V$.

(c) *Ψ ist bijektiv.*

Beweis-Skizze: (a) Folgt direkt aus der Definition.
(b) Sei U ein $\mathbb{K}$-UVR und $i_U : U \longrightarrow V$ die kanonische Inklusion. Nach Lemma 8.11 ist $i_U^* : V^* \longrightarrow U^*$ ein surjektiver Homomorphismus von $\mathbb{K}$-VR, also $\text{Bild}(i_U^*) = U^*$. Weiterhin $\alpha \in \text{Ker}(i_U^*) \iff \alpha \circ i = 0 \iff \alpha \in U^0$, also $\text{Ker}(i_U^*) = U^0$. Es folgt aus dem Dimensionssatz und aus Satz 8.3, dass

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} \text{Ker}(i_U^*) + \dim_{\mathbb{K}} \text{Bild}(i_U^*) = \dim_{\mathbb{K}} U^0 + \dim_{\mathbb{K}} U^* = \dim_{\mathbb{K}} U^0 + \dim_{\mathbb{K}} U.$$

(c) Sei $\Psi' : Gr(V^*) \longrightarrow Gr(V^{**})$ die Abbildung mit $\Psi'(W) = W^0$. Durch den kanonischen Isomorphismus $\text{ev} : V \longrightarrow V^{**}$ aus Satz 8.6 identifizieren wir sowohl V mit V^{**} also auch $Gr(V^{**})$ mit $Gr(V)$. Dadurch bekommen wir $U \subseteq (U^0)^0$ wie folgt. Für jeden $u \in U$ und $f \in U^0$ gilt $\text{ev}_u(f) = f(u) = 0$, also $\text{ev}_u \in (U^0)^0$. Aus Teil (b) gilt aber $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} (U^0)^0$, also

$$U = (U^0)^0, \quad \forall\, U \subseteq_{\mathbb{K}-\text{UVR}} V.$$

Das heißt, dass $\Psi' \circ \Psi = \text{id}_{Gr(V)}$. Analog, zeigt man, dass auch $\Psi \circ \Psi' = \text{id}_{Gr(V^*)}$.

Q.E.D.

Bemerkung 8.14. Seien U, U_1, U_2 $\mathbb{K}$-UVR eines endlichdimensionalen $\mathbb{K}$-Vektorraum V. Es gelten

1. $\{0_V\}^0 = V^*$.
2. $V^0 = \{0_{V^*}\}$.
3. $(U_1 \cap U_2)^0 = U_1^0 + U_2^0$.
4. $(U_1 + U_2)^0 = U_1^0 \cap U_2^0$.
5. Wenn wir V und V^{**} durch den Isomorphismus aus Satz 8.6 identifizieren, dann gilt $U = (U^0)^0$.

Bemerkung 8.15. Der Zusammenspiel zwischen U und U^0 liefert eine Verbindung zwischen die Beschreibung eines $\mathbb{K}$-UVR durch eine Basis (oder Erzeugendensystem) und die Beschreibung als Lösungsmenge eines LGS. Wenn $B = \{v_1, \ldots, v_r\}$ eine Basis von U ist und $\{f_1, \ldots, f_k\}$ eine Basis von U^0 eine Basis des Annullators ist, dann haben wir, wenn $\dim_{\mathbb{K}} V = n$,

$$U = \text{Span}_{\mathbb{K}}\{v_1, \ldots, v_r\} = \{v \in V \;:\; f_{r+1}(v) = \cdots = f_n(v) = 0\}.$$

8.5 Ein Algorithmus für $\mathbb{K}^n$

Wenn $V = \mathbb{K}^n$ verwenden wir folgende Konvention: die Elementen von $\mathbb{K}^n$ werden als $(x_1, \ldots, x_n)$ bezeichnet. Die Elementen von $(\mathbb{K}^n)^*$ werden durch $[a_1, \ldots, a_n]$ bezeichnet, wobei $[a_1, \ldots, a_n] = a_1 e_1^* + \cdots + a_n e_n^*$. Es gilt also

$$[a_1, \ldots, a_n](x_1, \ldots, x_n) = \begin{pmatrix} a_1 & \ldots & a_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_i x_i.$$

Für einen $\mathbb{K}$-UVR der durch Erzeuger gegeben ist, haben wir folgender Algorithmus um eine Darstellung als Lösungsmenge eines Gleichungssystems zu berechnen.

Bemerkung 8.16. Denselben Algorithmus kann man anwenden um aus der Beschreibung durch Gleichungen eine Basis eines $\mathbb{K}$-UVR von $\mathbb{K}^n$ zu berechnen. Der Grund dafür ist Punkt (iv) aus Bemerkung 8.14.

`Algorithmus zum Berechnen einer Basis des Annullators`

`Eingabe:` $ES_U = \{v_1, \ldots, v_r\}$ ein Erzeugendensystem von $U \subseteq \mathbb{K}^n$.

`Schritt 1:` Finde eine Basis $B_U \subseteq ES_U$ von U und ergänze diese mit B' zu einer Basis von $\mathbb{K}^n$.

`Schritt 2:` Wenn $B_U = \{v_1, \ldots, v_r\}$ und $B' = \{v_{r+1}, \ldots, v_n\}$, setze $B = \left(\begin{array}{c|c|c} v_1^T & \ldots & v_n^T \end{array} \right)$ und berechne B^{-1}.

`Ausgabe:` Wenn $z_1, \ldots, z_n$ die Zeilen von B^{-1} sind, dann ist die Ausgabe: $z_{r+1}, \ldots, z_n$.

Beweis-Skizze: Wir zeigen hier, warum dieses Verfahren eine Basis des Annullators ausgibt. Weil z_i die Zeilen von B^{-1} sind, gilt $z_i \cdot v_j = \delta_{ij}$. Also für alle $i = r+1, \ldots, n$ und alle $j = 1, \ldots, r$ gilt $z_i \cdot v_j = 0$, und somit $z_{r+1}, \ldots, z_n \in U^0$. Aus Satz 8.13 (b) ist die Dimension von $U^0 = n - r$. Es reicht also zu zeigen, dass unsere Kandidaten linear unabhängig sind. Das gilt weil die z_i die Zeilen einer invertierbaren Matrix sind. Q.E.D.

Beispiel 8.17. `Eingabe:` Sei $U = \mathrm{Span}_{\mathbb{K}}\{\{\}(1,2,3,4),(5,6,7,8),(9,10,11,12),(13,14,15,16)\} \subseteq \mathbb{R}^4$. `Schritt 1:` Wir wissen a priori nicht ob die vier Vektoren eine Basis von U bilden. Wir können aber gleichzeitig eine Basis finden und diese zu einer Basis von $\mathbb{R}^4$ ergänzen. Wir berechnen dafür (mit dem Gaußschen Verfahren) eine ZSF der Matrix:

$$\left(\begin{array}{cccc|cccc} 1 & 5 & 9 & 13 & 1 & 0 & 0 & 0 \\ 2 & 6 & 10 & 14 & 0 & 1 & 0 & 0 \\ 3 & 7 & 11 & 15 & 0 & 0 & 1 & 0 \\ 4 & 8 & 12 & 16 & 0 & 0 & 0 & 1 \end{array} \right) \left(\begin{array}{cccc|cccc} \boxed{1} & 5 & 9 & 13 & 1 & 0 & 0 & 0 \\ 0 & \boxed{1} & 2 & 3 & 1/2 & -1/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & -2 & 1 \end{array} \right)$$

Das heißt, dass die ersten zwei Spalten von A eine Basis von U sind, und das die 5-te und 6-te Spalte von A diese zu einer Basis von $\mathbb{R}^4$ ergänzen. Wir haben also $B_U = \{(1,2,3,4),(5,6,7,8)\}$ und

$B' = \{(1,0,0,0),(0,1,0,0)\}$.
`Schritt 2:` Wir setzen

$$B = \begin{pmatrix} 1 & 5 & 1 & 0 \\ 2 & 6 & 0 & 1 \\ 3 & 7 & 0 & 0 \\ 4 & 8 & 0 & 0 \end{pmatrix}$$

und berechnen B^{-1} indem wir die RZSF der erweiterten Matrix $\Big(B \mid I_4\Big)$ berechnen:

$$\big(B \mid I_4\big) = \left(\begin{array}{cccc|cccc} 1 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 5 & 6 & 0 & 1 & 0 & 1 & 0 & 0 \\ 9 & 10 & 0 & 0 & 0 & 0 & 1 & 0 \\ 13 & 14 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \quad \longmapsto \quad \big(I_4 \mid B^{-1}\big) = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & -2 & 7/4 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & -3/4 \\ 0 & 0 & 1 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{-3} & \mathbf{2} \\ 0 & 0 & 0 & 1 & \mathbf{0} & \mathbf{1} & \mathbf{-2} & \mathbf{1} \end{array}\right)$$

`Ausgabe:` $\{[1,0,-3,2],[0,1,-2,1]\}$

Probe:

$$\begin{aligned} [1,0,-3,2](1,2,3,4) &= 1 - 9 + 8 = 0 \\ [1,0,-3,2](5,6,7,8) &= 5 - 21 + 16 = 0 \\ [0,1,-2,1](1,2,3,4) &= 2 - 6 + 4 = 0 \\ [0,1,-2,1](5,6,7,8) &= 6 - 14 + 8 = 0 \end{aligned}$$

Also, das heißt, dass $w_1 = [1,0,-3,2]$ und $w_2 = [0,1,-2,1]$ in U^0 liegen. Weil w_1, w_2 Zeilen einer invertierbaren Matrix sind (B^{-1}), sind diese linear unabhängig. Weil $\dim_{\mathbb{R}} U^0 = \dim_{\mathbb{R}} \mathbb{R}^4 - \dim_{\mathbb{R}} U = 4 - 2$, sind diese auch eine Basis, also $U^0 = \operatorname{Span}_{\mathbb{K}} \{\{\}\, [1,0,-3,2],[0,1,-2,1]\}$. Wir haben also

$$U = \operatorname{Span}_{\mathbb{K}} \{\{\}\, (1,2,3,4),(5,6,7,8)\} = \{(x,y,z,w) \in \mathbb{R}^4 \;\; : \;\; x - 3z + 2w = 0 \;\; \text{und } y - 2z + w = 0\}.$$

Aus Bemerkung 8.16 haben wir auch

$$\begin{aligned} U^0 &= \operatorname{Span}_{\mathbb{K}} \{\{\}\, [1,0,-3,2],[0,1,-2,1]\} \\ &= \{[x,y,z,w] \in (\mathbb{R}^4)^* \;\; : \;\; x + 2y + 3z + 4w = 0 \;\; \text{und } 5x + 6y + 7z + 8w = 0\}. \end{aligned}$$

Kapitel 9

Polynome und $\mathbb{K}$-Algebren

Dieses Kapitel war ursprünglich ein kleiner Teil in Kapitel 10 Abschnitt 10.3. Es ist weiter noch kein allgemeiner Überblick der Theorie der Polynome und der $\mathbb{K}$-Algebren. Es enthält nur gezieltes Material das für Charakteristische Polynome angewendet sein soll.

Ich wiederhole hier kurz die Definition von Ring und von Ringhomomorphismus.

Definition 9.1. Ein **Ring** ist ein Tripel $(R, +, \cdot)$ wobei: R ist eine Menge, und $+$ und $\cdot$ sind innere Verknüpfungen, die Addition beziehungsweise Multiplikation genannt werden, sodass folgende Axiome erfüllt sind

(R 1) $(R, +)$ ist eine abelsche Gruppe.

(R 2) Die Multiplikation ist assoziativ und hat ein neutrales Element.

(R 3) (*Distributivitätsgesetze*) Für alle $a, b, c \in R$ gilt

$$(a + b) \cdot c = a \cdot c + b \cdot c \qquad \text{und} \qquad a \cdot (b + c) = a \cdot b + a \cdot c.$$

Ein **kommutativer Ring** ist ein Ring in dem die Multiplikation kommutativ ist, also

$$a \cdot b = b \cdot a, \quad \forall\ a, b \in R.$$

Bezeichnung. Das neutrale Element der Addition wird mit 0_R oder nur mit 0 bezeichnet. Das neutrale Element der Multiplikation wird mit 1_R oder nur mit 1 bezeichnet.

Definition 9.2. Seien R und S zwei Ringe. Ein **Ringhomomorphismus** von R nach S ist eine Abbildung $\varphi : R \longrightarrow S$ die folgende Axiome erfüllt

(RHom 1) $\varphi(a + b) = \varphi(a) + \varphi(b)$.

(RHom 2) $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$.

(RHom 3) $\varphi(1_R) = 1_S$.

Ein Ringisomorphismus ist ein Ringhomomorphismus der invertierbar ist, und dessen Inverse auch ein Ringisomorphismus ist. Es ist auch einfach zu zeigen, dass alle bijektive Ringhomomorphismen automatisch[1] Isomorphismen sind.

Der **Kern** eines Ringhomomorphismus φ ist $\operatorname{Ker}\varphi = \{r \in R \ : \ \varphi(r) = 0\}$, und es gilt auch für Ringhomomorphismen, dass φ injektiv ist $\Leftrightarrow \operatorname{Ker}\varphi = 0$.

Bemerkung 9.3. Wenn $\mathbb{K}$ ein Körper ist, R ein Ring und $\varphi : \mathbb{K} \longrightarrow R$ ein Ringhomomorphismus, dann ist f injektiv.

Beweis-Skizze: Es reicht zu zeigen, dass wenn $a \neq 0$, dann gilt $a \notin \operatorname{Ker}\varphi$. Sei also $0 \neq a \in \mathbb{K}$. Weil $\mathbb{K}$ ein Körper ist, existiert $a^{-1} \in \mathbb{K}$ mit $a \cdot a^{-1} = 1$. Dann gilt

$$\varphi(a) \cdot \varphi(a^{-1}) = \varphi(a \cdot a^{-1}) = \varphi(1_{\mathbb{K}}) = 1_R,$$

Also $\varphi(a) \neq 0$. Q.E.D.

9.1 Der Polynomring über $\mathbb{K}$

Wir werden den Polynomring in der Variable x mit Koeffizienten in $\mathbb{K}$ mit $\mathbb{K}[x]$ bezeichnen. Eine Möglichkeit[2] den Polynomring $\mathbb{K}[x]$ formal zu definieren ist die folgende. Man nimmt die direkte Summe (als $\mathbb{K}$-Vektorräume) von $\mathbb{N}$ Kopien des Körpers $\mathbb{K}$. So bekommt man schon eine abelsche Gruppe. Dann definiert man eine innere Multiplikation und überprüft, dass die Ring-Axiome erfüllt sind. Sei also

$$\mathbb{K}[x] := \bigoplus_{i\in\mathbb{N}} \mathbb{K} = \{(c_i)_{i\in\mathbb{N}} \ : \ c_i \in \mathbb{K} \text{ mit } c_i \neq 0 \text{ nur für endlich viele } n\}.$$

Polynome sind somit unendliche Tupeln über $\mathbb{N}$ indiziert, die aber nur endlich viele nicht-triviale[3] Einträge haben. So bekommt $\mathbb{K}[x]$ automatisch eine $\mathbb{K}$-Vektorraum Struktur, insbesondere ist $(\mathbb{K}[x], +)$ eine abelsche Gruppe mit der komponentenweise Addition:

$$(c_i)_{i\in\mathbb{N}} + (d_i)_{i\in\mathbb{N}} = (c_i + d_i)_{i\in\mathbb{N}}, \qquad \forall\ (c_i)_{i\in\mathbb{N}}, (d_i)_{i\in\mathbb{N}}\mathbb{K}[x].$$

Um einen Ring zu definieren brauchen wir eine (innere) Multiplikation auf $\mathbb{K}[x]$ anzugeben. Das machen wir für alle $(c_i)_{i\in\mathbb{N}}, (d_i)_{i\in\mathbb{N}} \in \mathbb{K}[x]$ wie folgt:

$$(c_i)_{i\in\mathbb{N}} \cdot (d_i)_{i\in\mathbb{N}} := (b_i)_{i\in\mathbb{N}} \in \mathbb{K}[x], \qquad \text{wobei } b_i := \sum_{k+\ell=i} c_k \cdot d_\ell.$$

Es ist sehr einfach direkt zu überprüfen, dass diese Multiplikation assoziativ ist, dass $(1, 0, 0, \dots)$ das neutrale Element dafür ist, und dass die Distributivität bezüglich der Addition ist.

[1] Das gilt aber nicht für alle mathematische Strukturen: Es gibt bijektive Morphismen von topologischen Räumen die kein Isomorphismus sind.

[2] Man kann das allgemeiner als die so genannte *symmetrische Algebra* machen, aber in einer Variable reicht diese Definition hier.

[3] das heißt nicht-Null.

Wir haben aber mehr als einen Ring definiert. Es gibt auch eine Multiplikation mit Skalare:

$$\lambda \cdot (c_i)_{i\in\mathbb{N}} := (\lambda \cdot c_i)_{i\in\mathbb{N}}, \qquad \forall\ \lambda \in \mathbb{K} \ \text{ und } \forall\ (c_i)_{i\in\mathbb{N}} \in \mathbb{K}[x].$$

Weiterhin, für jeden $i \in \mathbb{N}$ ist $e_i = (0, \dots, 0, 1, 0, \dots\dots) \in \mathbb{K}[x]$ ein Element der Standardbasis, wobei die 1 an der $i$ten Stelle vorkommt. Dann haben wir also

$$e_i \cdot e_j = e_{i+j}, \qquad \forall\ i, j \in \mathbb{N}.$$

Damit wir die klassische Bezeichnung für Polynome zurückgewinnen, setzen wir $x := e_1$, $x^0 := 1 = e_0$, und dann gilt $x^i = e_i$ für alle $i \in \mathbb{N}$.
Wenn $p = (c_i)_{i\in\mathbb{N}} \in \mathbb{K}[x]$ und wenn $n \in \mathbb{N}$ die Eigenschaft "$c_j = 0$ für alle $j > n$" hat, dann hat p eine eindeutige Schreibweise in der Form

$$p = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0.$$

Wir nehmen nicht immer an, dass $c_n \neq 0$. Wir bezeichnen Polynome mit einem Kleinbuchstaben $p, q, \dots \in \mathbb{K}[x]$, oder, um die Variable x hervorzuheben, mit $p(x), q(x) \dots \in \mathbb{K}[x]$. Diese letzte Schreibweise soll aber nicht mit dem Wert von p an einer Zahl verwechselt werden (cf. Definition 9.14).

Beispiel 9.4.

$$\begin{aligned}(x^4 + x^3 - \frac{1}{3}x + 1) + (-x^4 - 2x^2 + \frac{4}{3}x - 1) &= x^3 - 2x^2 + x \\ x \cdot (x-1)^2 = x \cdot (x^2 - 2x + 1) &= x^3 - 2x^2 + x\end{aligned}$$

Definition 9.5. Sei $n \in \mathbb{N}$ und $p = c_n x^n + \cdots + c_1 x + c_0 \in \mathbb{K}[x] \setminus \{0\}$. Der **Grad des Polynoms** p ist

$$\deg p := \max\{j \in \mathbb{N} \ : \ c_j \neq 0\}.$$

Der **Grad des Nullpolynoms** ist nicht bestimmt. Wir verwenden die Konvention, dass das Nullpolynom jedes mögliche Grad haben darf.

Ein Polynom von Grad n heißt **normiert**, wenn der Koeffizient von x^n gleich 1 ist. Zum Beispiel $p = 2x^2 + 3x + 4$ ist nicht normiert, aber $\frac{1}{2} \cdot p = x^2 + \frac{3}{2}x + 2$ ist normiert.

Lemma 9.6. *Für $p, q \in \mathbb{K}[x] \setminus \{0\}$ gilt*

(a) $\deg(p + q) \leqslant \max\{\deg p, \deg q\}$.

(b) $\deg(p \cdot q) = \deg(p) + \deg(q)$.

Beweis-Skizze: Übung Q.E.D.

Satz 9.7. *Seien $p, q \in \mathbb{K}[x]$ mit $q \neq 0$. Dann existieren eindeutig bestimmte Polynome $s, r \in \mathbb{K}[x]$, sodass*

$$p = q \cdot s + r \quad \text{und} \quad \deg r < \deg q.$$

Beweis-Skizze: [Bosch, Teil 5.2, Satz 2] Q.E.D.

Definition 9.8. Seien $p, q \in \mathbb{K}[x]$ zwei Polynome. Wir sagen, dass q **teilt** p, wenn es ein Polynom $s \in \mathbb{K}[x]$ existiert sodass $p = q \cdot s$. Wir schreiben also

$$q \,|\, p \quad \Longleftrightarrow \quad \exists\, s \in \mathbb{K}[x] \;\; \text{sodass} p = q \cdot s.$$

Wenn q, p teilt, dann ist das eindeutig bestimmte Polynom r aus Satz 9.7 das Nullpolynom.

Beispiel 9.9. Seien $p = 2x^5 + 3x^4 + x^3 - 2x^2 - 4x + 1$ und $q = x^2 + 2x - 1$ zwei Polynome in $\mathbb{R}[x]$. Hier ist ein Beispiel für das Ausrechnen der Polynome s und r aus Satz 9.7.

$$\begin{array}{rrrrrl}
2x^5 + 3x^4 & + x^3 & - 2x^2 & - 4x & + 1 = \big(x^2 + 2x - 1\big)\big(2x^3 - x^2 + 5x - 13\big) & + 27x - 12 \\
\underline{- 2x^5 - 4x^4 + 2x^3} & & & & & \\
- x^4 + 3x^3 & - 2x^2 & & & & \\
\underline{x^4 + 2x^3} & \underline{- x^2} & & & & \\
5x^3 & - 3x^2 & - 4x & & & \\
\underline{- 5x^3} & \underline{- 10x^2} & \underline{+ 5x} & & & \\
& - 13x^2 & + x & + 1 & & \\
& \underline{13x^2} & \underline{+ 26x} & \underline{- 13} & & \\
& & 27x & - 12 & &
\end{array}$$

Also $s = 2x^3 - x^2 + 5x - 13$ und $r = 27x - 12$.

9.2 $\mathbb{K}$-Algebren und Einsetzen in Polynome

Für jedes Polynom $p(x) \in \mathbb{K}[x]$ kann man die Unbekannte x mit einer konkreter Zahl (Element aus $\mathbb{K}$) ersetzen und eine eindeutig bestimmte Zahl bekommen. Man kann genauso x mit einer quadratischen Matrix, oder mit einem Endomorphismus ersetzen und eine eindeutige Matrix, bzw. Endomorphismus bekommen. Allgemein, kann man x mit "etwas" ersetzten wofür die Multiplikation mit Skalare, die Addition und Selbstmultiplikation wohl definiert sind. Für eine einheitliche und genaue Behandlung dieses Verfahrens, definieren wir als nächstes $\mathbb{K}$-Algebren. Man könnte dieses Begriff noch allgemeiner für einen Ring R, statt für einen Körper $\mathbb{K}$, definieren, und mann kann auch "weniger Struktur" der $\mathbb{K}$-Algebra verlangen [siehe Bosch Kapitel 5.1]. Für unsere Zwecke reicht aber folgende Definition.

Definition 9.10. Sei $\mathbb{K}$ ein Körper. Eine **$\mathbb{K}$-Algebra** ist ein Paar $(A, \varphi : \mathbb{K} \longrightarrow A)$ wobei A ein (nicht unbedingt kommutativer) Ring mit Eins ist, und $\varphi : \mathbb{K} \longrightarrow A$ ein Ringhomomorphismus mit der Eigenschaft

$$\varphi(\lambda) \cdot a = a \cdot \varphi(\lambda), \quad \forall\, \lambda \in \mathbb{K} \;\; \text{und } a \in A$$

ist[4].

Wir werden uns meistens mit einem der folgenden Beispiele beschäftigen.

[4] Es gibt sehr gute Gründe auf der Assoziativität der Multiplikation in A und der Existenz einer Eins in A verzichten. In einem solchen Kontext würde der Begriff den wir hier definiert haben *assoziative und unitäre* $\mathbb{K}$-Algebra heißen. Nicht assoziative Algebren (zum Beispiel Lie Algebren) spielen eine wichtige Rolle in der Quantenmechanik und in der Theorie der Elementarteilchen.

Beispiele:

1. $\mathbb{K}$ ist selber eine $\mathbb{K}$-Algebra mit $\varphi = \mathrm{id}_{\mathbb{K}}$.
2. $\mathrm{Mat}_n(\mathbb{K})$ ist eine $\mathbb{K}$-Algebra mit $\varphi : \mathbb{K} \longrightarrow \mathrm{Mat}_n(\mathbb{K})$ definiert als
$$\varphi(\lambda) := \lambda \cdot I_n.$$
3. Wenn V ein $\mathbb{K}$-Vektorraum ist, dann ist $\mathrm{End}_{\mathbb{K}}(V)$ ist eine $\mathbb{K}$-Algebra mit $\varphi : \mathbb{K} \longrightarrow \mathrm{End}_{\mathbb{K}}(V)$ definiert als
$$\varphi(\lambda) := \lambda \cdot \mathrm{id}_V\,.$$
4. Der Polynomring $\mathbb{K}[x]$ ist eine $\mathbb{K}$-Algebra mit $\varphi : \mathbb{K} \longrightarrow \mathbb{K}[x]$ definiert als
$$\varphi(\lambda) = \lambda + 0 \cdot x + 0 \cdot x^2 + \ldots$$

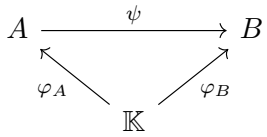
Bemerkung 9.11. (a) Jede $\mathbb{K}$-Algebra $\varphi : \mathbb{K} \longrightarrow A$ bekommt eine $\mathbb{K}$-Vektorraum Struktur mit der Addition auf A und der Skalarmultiplikation $\cdot : \mathbb{K} \times A \longrightarrow A$ gegeben durch
$$\lambda \cdot a := \varphi(\lambda) \cdot a,$$
wobei das Zeichen "$\cdot$" in $\varphi(\lambda) \cdot a$ die Multiplikation auf A bezeichnet. Also eine $\mathbb{K}$-Algebra ist ein $\mathbb{K}$-Vektorraum der auch eine Ringstruktur hat, und diese Ringstruktur ist mit der $\mathbb{K}$-Vektorraumstruktur kompatibel.

(b) Weil $\mathbb{K}$ ein Körper ist, ist jeder Ringhomomorphismus $\mathbb{K} \longrightarrow A$ injektiv. Wir identifizieren deswegen die Elementen aus $\lambda \in \mathbb{K}$ mit $\varphi(\lambda) = \lambda \cdot \varphi(1_{\mathbb{K}}) = \lambda \cdot 1_A$. Während für $\mathbb{K}[x]$ die Schreibweise für $\lambda \in \mathbb{K}$ und $\lambda \in \mathbb{K}[x]$ als Polynom von Grad 0 dieselbe ist, verstehen wir durch "$\lambda \in \mathrm{End}_K(V)$" den Endomorphismus $h_\lambda = \lambda \cdot \mathrm{id}_V$ und durch "$\lambda \in \mathrm{Mat}_n(\mathbb{K})$" die Matrix $\lambda \cdot I_n = \mathrm{diag}(\lambda, \ldots, \lambda)$.

Wir sagen manchmal einfach, dass $\varphi : \mathbb{K} \longrightarrow A$ eine $\mathbb{K}$-Algebra ist. Falls der Ringhomomorphismus selbstverständlich ist, schreiben wir nur "$\mathbb{K} \hookrightarrow A$ ist eine $\mathbb{K}$-Algebra" oder, am einfachsten, "A ist eine $\mathbb{K}$-Algebra".

Wie immer, wenn wir neue algebraische Strukturen (Objekte) definieren, sollten wir auch gleich sagen was Homomorphismen zwischen solche Strukturen sein sollen. Da eine $\mathbb{K}$-Algebra gleichzeitig ein Ring und ein $\mathbb{K}$-Vektorraum ist, sollte ein Homomorphismus von $\mathbb{K}$-Algebren eine linearer Ringhomomorphismus sein. Das kann man eleganter durch kommutative Diagramme von Ringhomomorphismen definieren.

Definition 9.12. Seien A, B zwei $\mathbb{K}$-Algebren. Ein **Homomorphismus von $\mathbb{K}$-Algebren** (oder ein $\mathbb{K}$-Algebrahomomorphismus) ist ein Ringhomomorphismus $\psi : A \longrightarrow B$ mit der Eigenschaft, dass das Diagramm



kommutativ ist. Das heißt, dass $\big(\psi \circ \varphi_A\big)(\lambda) = \psi(\varphi_A(\lambda)) = \varphi_B(\lambda)$ für alle $\lambda \in \mathbb{K}$.

Wenn die strukturelle Ringhomomorphismen selbstverständlich sind, dann lässt sich ein $\mathbb{K}$-Algebra Homomorphismus durch die Eigenschaft

$$\psi(\lambda \cdot a) = \lambda \cdot \psi(a), \quad \forall\, \lambda \in \mathbb{K} \quad \text{und} \quad a \in A,$$

charakterisieren. Somit kann man sehen, dass ein $\mathbb{K}$-Algebrahomomorphismus insbesondere $\mathbb{K}$-linear ist.
Für jeden Ring A und jedes Element $a \in A$, bezeichnen wir mit

$$\begin{aligned} a^i &:= \underbrace{a \cdot a \cdots a}_{\text{i-Mal}}, \quad \forall i > 1 \\ a^0 &:= 1_A. \end{aligned}$$

Also, für ein $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist $f^3 = f \circ f \circ f$ und $f^0 = \mathrm{id}_V$. Genauso, für jede quadratische Matrix A^0 die Identitätsmatrix I_n.

Satz 9.13. *Seien $\mathbb{K}$ ein Körper, $\mathbb{K}[x]$ der Polynomring in einer Variable x und A eine beliebige $\mathbb{K}$-Algebra. Zu jedem Element $a \in A$ gibt es einen eindeutig bestimmten Homomorphismus von $\mathbb{K}$-Algebren $\mathrm{ev}_a : \mathbb{K}[x] \longrightarrow A$ mit $\mathrm{ev}_a(x) = a$.*

Beweis-Skizze: Eindeutigkeit Wenn $\mathrm{ev}_a : \mathbb{K}[x] \longrightarrow A$ ein $\mathbb{K}$-Algebrahomomorphismus der geforderten Art ist, so gilt für jedes Polynom $p = c_n x^n + \cdots + c_1 x + c_0 \in \mathbb{K}[x]$, dass

$$\mathrm{ev}_a(p) = \mathrm{ev}_a\left(\sum_{i=0}^{n} c_i x^i\right) = \sum_{i=0}^{n} c_i \,\mathrm{ev}_a(x^i) = \sum_{i=0}^{n} c_i (\mathrm{ev}_a(x))^i = \sum_{i=0}^{n} c_i a^i,$$

wobei die 2te und 3te Gleichheit aus der $\mathbb{K}$-Algebrahomomorphismus Bedingung folgen, und die letzte aus der Forderung, dass $\mathrm{ev}_a(x) = a$. So haben wir gezeigt, dass das Bild von x das Bild aller Polynome eindeutig bestimmt.
Existenz Man setzt für jedes Polynom $p = \sum_{i=0}^{n} c_i x^i \in \mathbb{K}[x]$

$$ev_a(p) := \sum_{i=0}^{n} c_i a^i.$$

Man sieht unmittelbar, dass die so-definierte Abbildung $ev_a : \mathbb{K}[x] \longrightarrow A$ ein Homomorphismus von $\mathbb{K}$-Algebren ist. Q.E.D.

Definition 9.14. Sei A eine $\mathbb{K}$-Algebra und sei $a \in A$. Der **Einsetzungshomomorphismus** (oder die **Evaluationsabbildung**) der a an Stelle von x einsetzt, ist der eindeutig bestimmte Homomorphismus von $\mathbb{K}$-Algebren

$$ev_a : \mathbb{K}[x] \longrightarrow A$$

aus Satz 9.13. Wir bezeichnen für jedes $a \in A$ und jedes $p(x) \in \mathbb{K}[x]$ mit

$$p(a) := \mathrm{ev}_a(p).$$

Bemerkung 9.15. Es gelten also für $\lambda \in \mathbb{K}$, $a \in A$ und $p, q \in \mathbb{K}[x]$:

(i) $(p+q)(a) = p(a) + q(a)$.

(ii) $(p \cdot q)(a) = p(a) \cdot q(a)$.

(iii) $(\lambda \cdot p)(a) = \lambda \cdot p(a)$.

Für konstante Polynome c gilt $c(\lambda) = c \cdot 1_A$.

9.3 Nullstellen von Polynome

Definition 9.16. Sei $p \in \mathbb{K}[x]$. Die **zu p gehörige Polynomabbildung** ist die Abbildung $\phi_p : \mathbb{K} \longrightarrow \mathbb{K}$ gegeben durch

$$\phi_p(\lambda) := p(\lambda).$$

Eine Abbildung $\phi : \mathbb{K} \longrightarrow \mathbb{K}$ heißt Polynomabbildung wenn es ein $p \in \mathbb{K}[x]$ existiert, sodass $\phi = \phi_p$.

Bemerkung 9.17. Wenn $\mathbb{K}$ ein endlicher Körper ist, dann ist die Zuordnung $p \mapsto \phi_p$ nicht injektiv. Zum Beispiel, wenn $\mathbb{K} = \mathbb{Z}_3$, dann geben sowohl das Nullpolynom als auch das Polynom $p(x) = x^3 + 2x = x(x-1)(x-2)$ die Nullabbildung. Wir werden später sehen, dass diese Zuordnung genau dann injektiv ist, wenn $\mathbb{K}$ unendlich ist. Das ist einer der Gründe dafür, dass wir Polynome nicht als konkrete Funktionen, sondern als formale Ausdrücke definiert haben.

Definition 9.18. Sei $p \in \mathbb{K}[x]$ ein Polynom. Eine **Nullstelle** von p ist ein Element $\lambda \in \mathbb{K}$, sodass $p(\lambda) = 0$.

Lemma 9.19. *Wenn $\lambda \in \mathbb{K}$ eine Nullstelle des Polynoms $p \in \mathbb{K}[x]$ ist, dann gilt $(x - \lambda) \,|\, p$. Es gibt also ein Polynom $s \in \mathbb{K}[x]$, sodass*

$$p = (x - \lambda) \cdot s.$$

Beweis-Skizze: Wir führen die Division mit Rest von p durch $x - \lambda$. Es gibt also eindeutig bestimmte Polynome $s, r \in \mathbb{K}[x]$ mit

$$p = (x - \lambda) \cdot s + r \qquad \text{mit } \deg r < \deg(x - \lambda) = 1.$$

Also r ist ein konstantes Polynom. Wir setzten jetzt λ auf beiden Seiten ein und bekommen

$$0 = p(\lambda) = (\lambda - \lambda) \cdot s(\lambda) + r(\lambda) = r.$$

Q.E.D.

Man spricht in dem obigen Kontext von der Aufspaltung des Linearfaktors $(x - \lambda)$.

Definition 9.20. Sei $p \in \mathbb{K}[x] \setminus \{0\}$ ein nicht-nulles Polynom und sei $\lambda \in \mathbb{K}$ eine Nullstelle von p. Die **Vielfachheit der Nullstelle** (oder **Multiplizität der Nullstelle**) λ von p ist die positive natürliche Zahl

$$\mu_a(p, \lambda) := \max\{i \in \mathbb{N} \;:\; (x - \lambda)^i \,|\, p\}.$$

Satz 9.21. *Sei $p \in \mathbb{K}[x]$ ein nicht-nulles Polynom mit $m \geq 0$ verschiedene Nullstellen $\lambda_1, \ldots, \lambda_m$. Dann gibt es ein Polynom $q \in \mathbb{K}[x]$ ohne Nullstellen, sodass*

$$p = (x - \lambda_1)^{l_1} \cdot \cdots \cdot (x - \lambda_m)^{l_m} \cdot q, \qquad \text{wobei } l_i = \mu_a(p, \lambda_i) \;\; \forall\, i = 1, \ldots, m.$$

Insbesondere, p hat höchstens $\deg p$ mit Vielfachheit gezählte Nullstellen.

Beweis-Skizze: Wir führen Induktion über $n = \deg p$. Wenn $n = 0$, dann ist $p = a_0 \in \mathbb{K} \setminus \{0\}$ und man muss nichts zeigen.
$\boxed{n \Rightarrow n+1}$ Sei p mit $\deg p = n+1$. Wenn p keine Nullstellen hat, dann setzen wir $q := p$. Sonst, sei λ eine Nullstelle von p. Aus Lemma 9.19 folgt, dass es ein Polynom $s \in \mathbb{K}[x]$ gibt, sodass $p = (x - \lambda) \cdot s$. Aus Lemma 9.6 folgt, dass $\deg s = n$. Aus der induktiven Voraussetzung existiert $q \in \mathbb{K}[x]$ ohne Nullstellen, sodass $s = (x - \lambda_1)^{l_1} \cdot \dots \cdot (x - \lambda_m)^{l_m} \cdot q$. Also

$$p = (x - \lambda) \cdot (x - \lambda_1)^{l_1} \cdot \dots \cdot (x - \lambda_m)^{l_m} \cdot q.$$

Wir müssen nur noch die Vielfachheiten überprüfen. Wenn $\lambda \notin \{\lambda_1, \dots, \lambda_m\}$ dann sind $\lambda_1, \dots, \lambda_m, \lambda$ die Nullstellen von p mit Vielfachheiten $l_1, \dots, l_m, 1$. Sonst, existiert $i \in \{1, \dots, m\}$ mit $\lambda = \lambda_i$; dann sind $\lambda_1, \dots, \lambda_m$ die Nullstellen, und $l_1, \dots, l_{i-1}, l_i + 1, l_{i+1}, \dots, l_m$ die Vielfachheiten. Q.E.D.

Korollar 9.22. Wenn der Körper $\mathbb{K}$ unendlich viele Elemente hat, dann ist die lineare Abbildung $F : \mathbb{K}[x] \longrightarrow \mathrm{Abb}(\mathbb{K}, \mathbb{K})$, gegeben durch $F(p) := \phi_p$, injektiv.

Beweis-Skizze: Die Abbildung F ist offensichtlich $\mathbb{K}$-linear. Also injektiv ist äquivalent zu $\mathrm{Ker}\, F = \{0\}$. Das gilt, weil wenn ϕ_p die Nullabbildung ist, dann hat p unendlich viele Nullstellen - Widerspruch zu Satz 9.21. Q.E.D.

Definition 9.23. Ein Körper $\mathbb{K}$ heißt **algebraisch abgeschlossen** wenn jedes Polynom vom Grad ≥ 1 mindestens eine Nullstelle hat.

Beispiele:

1. Die Körper $\mathbb{Q}$ und $\mathbb{R}$ sind nicht algebraisch abgeschlossen: das Polynom $x^2 + 1$ hat keine Nullstelle in $\mathbb{Q}$ oder $\mathbb{R}$.
2. Endliche Körper sind nicht algebraisch abgeschlossen (Übung).
3. Der Körper der komplexen Zahlen ist algebraisch abgeschlossen, cf. Satz 9.25.

Satz 9.24. *Seien $\mathbb{K}$ ein algebraisch abgeschlossener Körper und $p \in \mathbb{K}[x]$ ein Polynom mit $\deg p \geq 1$ und mit verschiedenen Nullstellen $\lambda_1, \dots, \lambda_m$. Dann existiert ein Skalar $a \in \mathbb{K} \setminus \{0\}$, sodass*

$$p = a \cdot (x - \lambda_1)^{l_1} \cdot \dots \cdot (x - \lambda_m)^{l_m} \text{ wobei } l_i = \mu_a(p, \lambda_i) \quad \forall\ i = 1, \dots, m.$$

Beweis-Skizze: Aus Satz 9.21 existiert ein Polynom $q \in \mathbb{K}[x]$, das keine Nullstellen hat, sodass $p = q \cdot \prod_{i=1}^{m} (x - \lambda_i)^{l_i}$. Weil $\mathbb{K}$ algebraisch abgeschlossen ist und q keine Nullstellen hat, folgt dass $q \neq 0$ und $\deg q = 0$. Also $a := q$. Q.E.D.

Wir sagen in der obigen Situation, dass *das Polynom p zerfällt vollständig in Linearfaktoren.*

Satz 9.25 (Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Der Beweis dieses Satzes erfordert Methoden außerhalb der linearen Algebra. Diese Methoden sind auch unabhängig von der Theorie die wir hier weiter entwickeln, sodass kein Zirkelschluss zu befürchten ist. Üblicherweise sieht man einen Beweis dieses Satzes in einer Algebra und Zahlentheorie Vorlesung, oder einer Funktionentheorie Vorlesung.

9.4 Wie findet man Nullstellen von Polynome?

Menschen beschäftigen sich damit seit mehr als 5000 Jahre. Eine allgemeine Methode (durch Radikale) gibt es aber nur für Polynome von Grad höchstens vier (siehe den Satz von Abel-Ruffini und Galois Theorie). Es gibt aber Algorithmen die in dem allgemeinen Fall numerische Annäherungen finden. Wir beschränken uns aber auf das "Eins-Zwei-Drei-Verfahren", d.h. wir überprüfen, ob $\pm 1, \pm 2, \pm 3, \ldots$ vielleicht Nullstellen sind. Wir werden auch die Lösungsformel für Polynome von Grad 2 anwenden.
Sei $p = ax^2 + bx + c \in \mathbb{C}[x]$ mit $a \neq 0$. Die Nullstellen von p sind

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \in \mathbb{C}$$

Falls $a, b, c \in \mathbb{R}$, existieren Nullstellen in $\mathbb{R}$ genau dann, wenn die Diskriminante $\Delta = b^2 - 4ac$ nicht-negativ ist. Eine doppelte Nullstelle existiert genau dann, wenn $\Delta = 0$.

Folgende einfache Bemerkung kann ein Trick in der Suche nach Nullstellen sein.

Bemerkung 9.26. Sei $p = x^n + a_{n-1}x^{n-1} \cdots + a_1 x + a_0 \in \mathbb{K}[x]$ mit n Nullstellen $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$. Dann gilt

$$\begin{aligned} -a_{n-1} &= \lambda_1 + \cdots + \lambda_n \\ a_0 &= (-1)^n \lambda_1 \cdots \lambda_n. \end{aligned}$$

Insbesondere, wenn alle a_i und λ_j ganze Zahlen sind, dann ist jede Nullstelle ein Teiler von a_0.

Mit dieser empirischen Methode zum Bestimmen einer Nullstelle, können wir dann ein Verfahren für die Zerlegung in Linearfaktoren eines Polynoms $p \in \mathbb{K}[x]$ zusammenstellen:

`Schritt` 1. Wenn p keine Nullstelle hat, dann kann man keine Linearfaktoren abspalten. Also Abbrechen.

`Schritt` 2. Sonst, bestimme eine Nullstelle λ von p (das ist das schwierige Problem).

`Schritt` 3. Finde durch Division mit Rest das Polynom s, sodass $p = (x - \lambda) \cdot s$.

`Schritt` 4. Wenn $\deg s \geq 2$, dann wiederhole Schritt 1 mit $p = s$.

Dieses Verfahren ist eher ungenau, da man nicht immer sagen kann ob Polynome eine Nullstelle haben.

Beispiel 9.27. Sei $p = x^4 - 5x^3 + 7x^2 + x - 12$. Durch Ausprobieren finden wir $p(-1) = 0$. Wir teilen also p durch $x + 1$ um s zu finden:

$$\begin{array}{rl} x^4 - 5x^3 + 7x^2 + x - 12 = \big(x+1\big)\big(x^3 - 6x^2 + 13x - 12\big) & \\ \underline{-x^4 - x^3} & \\ -6x^3 + 7x^2 & \\ \underline{6x^3 + 6x^2} & \\ 13x^2 + x & \\ \underline{-13x^2 - 13x} & \\ -12x - 12 & \\ \underline{12x + 12} & \\ 0 & \end{array}$$

Also $s = x^3 - 6x^2 + 13x - 12$. Wir sehen dann, wieder durch Ausprobieren, dass $s(3) = 0$. Wir teilen s durch $x - 3$:

$$\begin{array}{rl} x^3 - 6x^2 + 13x - 12 = & \big(x - 3\big)\big(x^2 - 3x + 4\big) \\ \underline{- x^3 + 3x^2} & \\ - 3x^2 + 13x & \\ \underline{3x^2 \quad - 9x} & \\ 4x - 12 & \\ \underline{- 4x + 12} & \\ 0 & \end{array}$$

und finden $s' = x^2 - 3x + 4$, dessen Lösungen

$$\lambda_{1,2} = \frac{3 \pm \sqrt{9 - 16}}{2} = \frac{3 \pm i\sqrt{7}}{2}.$$

Wir haben also die Zerlegung von p in Linearfaktoren:

$$p = (x + 1)(x - 3)(x - \frac{3 + i\sqrt{7}}{2})(x - \frac{3 - i\sqrt{7}}{2}).$$

9.4.1 Ein Trick

Es gibt eigentlich einen Trick den man probieren kann um ein Polynom in Faktoren zu zerlegen. Nämlich, man kann die irreduzible Faktoren finden die mit einer Vielfachheit von mehr als eins vorkommen.

Sei $f \in \mathbb{K}[x]$, mit[5] $\operatorname{char} \mathbb{K} = 0$. Polynome kann man wie gewöhnlich ableiten. Man muss sich jetzt keine Gedanken machen was infinitesimale Prozesse über beliebige Körper heißen. Die Ableitungen funktionieren formell einfach durch lineares erweitern der Regel $\frac{\partial}{\partial x} x^n = n \cdot x^{n-1}$. Zum Beispiel, wenn $f = x^3 + 4x^2 - 2x + 1$, dann gilt

$$f' = 3x^2 + 8x - 2.$$

Die übliche Regel für die Ableitung des Produktes gilt auch:

$$(f \cdot g)' = f' \cdot g + f \cdot g'.$$

Nehmen wir an, dass $f = p^d \cdot q$ für irgendwelche $p, q \in \mathbb{K}[x]$ von Grad > 0, dann gilt

$$f' = (p^d \cdot q)' = (p^d)' \cdot q + p^d \cdot q' = d \cdot p^{d-1} \cdot p' \cdot q + p^d \cdot q' = p^{d-1} \cdot (d \cdot p' \cdot q + p \cdot q').$$

Also, wenn in der Zerlegung von f in Faktoren in höheren Potenzen vorkommen, dann teilt die nächst kleinere Potenz davon auch die Ableitung von f'. Also, um f in Faktoren zu zerlegen, kann man unter den gemeinsamen Teiler von f und f' anfangen. Genauer gesagt, der größte gemeinsame Teiler wird ein Faktor von f sein. Um den ggT zu berechnen kann man, wie bei ganzen Zahlen, den erweiterten euklidischen Algorithmus. Das funktioniert wegen Satz 9.7.

[5] Wir nehmen das an, damit wir Situationen in denen die Ableitung von nicht-trivialen Polynomen verschwindet vermeiden. Zum Beispiel, die Ableitung von $x^2 + 1 \in \mathbb{F}_2[x]$ ist Null.

Teil II

Lineare Algebra 2

Kapitel 10

Die Klassifikation von Endomorphismen

Wir wollen die Struktur von Endomorphismen eines endlichdimensionalen Vektorraumes besser Verstehen. Für das ganze Kapitel haben wir:

- $\mathbb{K}$ ist ein Körper
- alle $\mathbb{K}$-Vektorräume sind endlichdimensional
- $n, m \in \mathbb{N}_{>0}$ sind positive ganze Zahlen und meistens wird $n = \dim_{\mathbb{K}} V$ bezeichnen.

Wir werden dafür die Zuordnungen von Matrizen benutzen. Eine solche Zuordnung entspricht der Wahl einer Basis. Die Frage die wir uns stellen ist:

Welche ist die schönste/beste/schlauste *Matrix die man einer linearen Abbildung zurodnen kann?*

Das heißt, in diesem Kapitel werden Matrizen als *einer linearen Abbildung zugeordnet* analysiert. Außer Abschnitt 10.1.1 werden es ausschließlich quadratische Matrizen die einem Endomorphismus zugeordnet sind.

10.1 Ähnlichkeit von Matrizen

Die Idee ist eine Äquivalenzrelation auf der Menge aller $(m \times n)-$Matrizen mit Einträge in einem Körper zu definieren, die in Zusammenhang mit dem Basiswechsel steht. Wir werden gleich im Abschnitt 10.1.1 sehen, dass im Fall $m \neq n$ diese Äquivalenzrelation keine zusätzliche Information bringt. Für quadratische Matrizen wird das wesentlich komplizierter sein. Dafür werden wir Begriffe wie Eigenwerte und Eigenvektoren einführen.

10.1.1 Matrizen in $\mathrm{Mat}_{m,n}(\mathbb{K})$ mit $m \neq n$.

Der Hintergrund ist folgender. Wenn $f : V \longrightarrow W$ eine lineare Abbildung zwischen zwei endlich dimensionale Vektorräume ist, dann suchen wir geordnete Basen B und C von V, beziehungsweise von W, sodass die zugeordnete Matrix $M_C^B(f)$ möglichst einfach ist. Wenn zwei andere geordnete Basen B' und C' gewählt werden, dann ist der Zusammenhang zwischen den zugeordneten Matrizen durch

$$M_{C'}^{B'}(f) = \left(M_C^{C'}\right)^{-1} \cdot M_C^B(f) \cdot M_B^{B'}$$

gegeben, wobei $M_B^{B'} = M_B^{B'}(\mathrm{id}_V)$ und $M_C^{C'} = M_C^{C'}(\mathrm{id}_W)$. Basiswechsel Matrizen sind invertierbar, und die Umkehrung gilt auch: jede invertierbare Matrix ist eine Basiswechsel Matrix.

Definition 10.1. Zwei Matrizen $A, B \in \mathrm{Mat}_{m,n}(\mathbb{K})$ sind **äquivalent** wenn es Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(\mathbb{K})$ gibt, sodass

$$A = S^{-1} \cdot B \cdot T.$$

Übung. Die obige Relation ist eine Äquivalenzrelation.

Aus der Lineare Algebra 1 Vorlesung bekommen wir den folgenden Satz.

Satz 10.2. *Eine Matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$ von Rang r ist zu der Matrix*

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

äquivalent. Insbesondere, sind zwei Matrizen $A, B \in \mathrm{Mat}_{m,n}(\mathbb{K})$ äquivalent genau dann, wenn $\mathrm{Rang}\, A = \mathrm{Rang}\, B$.

Diese Äquivalenzrelation bringt uns also nicht mehr Information als der Rang. Es gibt also genau $\min\{m,n\} + 1$ Äquivalenzklassen, und es gibt kanonische Repräsentanten: Die Nullmatrix, oder für $r = 1, \ldots, \min\{m,n\}$ die Matrizen mit $a_{ii} = 1$ für $i = 1, \ldots, r$ und alle andere Einträge Null.

[2]25.4.'22

10.1.2 Quadratische Matrizen

Die Suche nach einer möglichst einfachen Matrixdarstellung eines Endomorphismus $f : V \longrightarrow V$ ist nicht so einfach. Der Grund ist, dass wir die zugeordnete Matrix $M^B(f)$ bezüglich einer einzigen geordneten Basis B von V definieren. Somit dürfen (und wollen) wir nicht unterschiedliche Basen für den Definitionsbereich und den Wertebereich wählen. Wenn wir eine andere geordnete Basis B' von V wählen und wenn $T = M_B^{B'}$ die Basiswechselmatrix bezeichnet, dann ändert sich die zugeordnete Matrix wie folgt:

$$M^{B'}(f) = T^{-1} \cdot M^B(f) \cdot T.$$

Das ist der Grund für folgende wichtige Definition.

Definition 10.3. Zwei Matrizen $A, B \in \mathrm{Mat}_n(\mathbb{K})$ sind **ähnlich** wenn es eine Matrix $T \in \mathrm{GL}_n(\mathbb{K})$ gibt, sodass

$$A = T^{-1} \cdot B \cdot T.$$

Wir bezeichnen das mit $A \sim B$.

Übung. Ähnlichkeit ist eine Äquivalenzrelation auf $\mathrm{Mat}_n(\mathbb{K})$.

Man kann gleich sehen, dass $\lambda \cdot I_n \not\sim \mu \cdot I_n$ wenn $\lambda \neq \mu$. Wenn $\mathbb{K}$ unendlich ist, gibt es also unendlich viele Ähnlichkeitsklassen[1]. In dem nächsten Teil führen wir Begriffe ein, die essentiell in der Beschreibung der Ähnlichkeitsklassen sind. Das Ziel ist Kriterien zu finden, die zwischen $A \sim B$ und $A \not\sim B$ entscheiden.

[1] Das heißt Äquivalenzklassen für die Ähnlichkeit der Matrizen.

10.2 Erste Eigenschaften von Eigenwerte, -vektoren und -räume

Der *einfachste*[2] Endomorphismus eines $\mathbb{K}$-Vektorraumes V ist die Identität id_V. Man kann behaupten, dass die nächst-einfachste Art von Endomorphismus durch die Multiplikation mit einem Skalar $\lambda \in \mathbb{K}$ gegeben ist. Geometrisch heißt es, dass der Vektorraum V mit Faktor λ gestreckt wird. Für jeden Skalar $\lambda \in \mathbb{K}$ werden wir den entsprechenden Endomorphismus mit $h_\lambda : V \longrightarrow V$ bezeichnen, also

$$h_\lambda(v) = \lambda \cdot v, \quad \forall\, v \in V.$$

Da $\mathrm{End}_{\mathbb{K}}(V)$ selber eine $\mathbb{K}$-Vektorraumstruktur hat, können wir einfach $h_\lambda = \lambda \cdot \mathrm{id}_V$ schreiben. Diese Art von Endomorphismus ist relativ selten, wobei man die Seltsamkeit in diesem Fall ganz konkret ausgedrückt werden kann (Übung). Der nächste Schritt ist zu verstehen, ob solche Dehnungen für kleinere Teile von V stattfinden. Dafür definieren wir folgende Begriffe.

Definition 10.4. Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Ein Skalar $\lambda \in \mathbb{K}$ ist ein **Eigenwert** von f, wenn es einen nicht-Null Vektor $v \in V \setminus \{0\}$ gibt, sodass

$$f(v) = \lambda \cdot v.$$

In diesem Fall heißt der Vektor v **Eigenvektor** von f (zum Eigenwert λ). Für jeden Skalar $\lambda \in \mathbb{K}$ ist der **λ-Eigenraum** von f die Teilmenge

$$\mathrm{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda \cdot v\}.$$

Wir erinnern, dass jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ den Endomorphismus $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ durch

$$f_A(\mathbf{x}) := A \cdot \mathbf{x} \quad \forall\, \mathbf{x} \in \mathbb{K}^n$$

definiert.

Definition 10.5. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ eine quadratische Matrix. Ein Eigenwert (beziehungsweise Eigenvektor oder -raum) von A ist ein Eigenwert (beziehungsweise Eigenvektor oder -raum) von f_A. Wir bezeichnen mit $\mathrm{Eig}(A, \lambda) := \mathrm{Eig}(f_A, \lambda)$.

Die Menge aller Eigenwerte eines Endomorphismus (oder einer Matrix) wird auch **Spektrum** genannt. Diese Bezeichnung wird aber meistens für lineare Operatoren auf unendlich-dimensionale Vektorräume verwendet.

Bemerkung 10.6. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$.

(a) Es gilt $\mathrm{Eig}(f, 0) = \mathrm{Ker}(f)$. Also der Skalar $0 \in \mathbb{K}$ ist genau dann ein Eigenwert von f, wenn f **nicht** injektiv ist.

(b) Die Eigenräume von f sind $\mathbb{K}$-Untervektorräume von V, weil

$$\mathrm{Eig}(f, \lambda) = \mathrm{Ker}(f - h_\lambda),$$

wobei $h_\lambda = \lambda \cdot \mathrm{id}_V$.

[3]27.4.'22

(c) Ein Skalar $\lambda \in \mathbb{K}$ ist ein Eigenwert von f genau dann, wenn $\mathrm{Eig}(f, \lambda) \neq \{0\}$.

[2] "Einfach" ist ein relativer Begriff. Ich finde die Identität "am einfachsten" weil es für alle Mengen (ohne weitere Strukturen/Verknüpfungen darauf) immer die Identische Abbildung gibt. Eine andere "einfache" lineare Abbildung, die auch immer existiert, ist die Nullabbildung. Beide sind Sonderfälle von Homothetien: $\mathrm{id} = h_1$ und $0 = h_0$.

(d) Jeder Vektor $v \in \mathrm{Eig}(f, \lambda) \setminus \{0\}$ ist ein Eigenvektor von f zum Eigenwert λ.

(e) Ein Eigenraum $E = \mathrm{Eig}(f, \lambda)$ ist f-invariant, das heißt

$$f(E) \subseteq E.$$

(f) Für $\lambda, \mu \in \mathbb{K}$ mit $\lambda \neq \mu$ gilt

$$\mathrm{Eig}(f, \lambda) \cap \mathrm{Eig}(f, \mu) = \{0\}.$$

(g) Sei $A \in \mathrm{Mat}_n(\mathbb{K})$. Für $\lambda \in \mathbb{K}$ ist

$$\mathrm{Eig}(A, \lambda) = \{\mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = \lambda\mathbf{x}\} = \mathcal{L}(A - \lambda I_n \mid \mathbf{0}),$$

Wobei $\mathrm{LGS}(A - \lambda I_n \mid \mathbf{0})$ das homogene lineare Gleichungssystem $(A - \lambda I_n) \cdot \mathbf{x} = \mathbf{0}$ und $\mathcal{L}(A - \lambda I_n \mid \mathbf{0})$ dessen Lösungsmenge bezeichnen.

Beispiele:

1. Wenn $f = \mathrm{id}_V$, dann ist der einzige Eigenwert $\lambda = 1$ und

   $$\mathrm{Eig}(\mathrm{id}, 1) = V.$$

   Also jeder Vektor $0 \neq v \in V$ ist ein Eigenvektor zum Eigenwert 1.

2. Wenn $f = h_\lambda$ für $\lambda \in \mathbb{K}$, dann ist der einzige Eigenwert λ und

   $$\mathrm{Eig}(h_\lambda, \lambda) = V.$$

   Also jeder Vektor $0 \neq v \in V$ ist ein Eigenvektor von h_λ zum Eigenwert λ.

3. Sei $V = \mathbb{R}^2$ und $f = D_\theta$ die Drehung um den Nullpunkt $\mathbf{0} \in \mathbb{R}^2$ mit Winkel θ. Für die Standardbasis B von $\mathbb{R}^2$ haben wir dann

   $$M^B(D_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

   Ein Vektor v ist Eigenvektor von D_θ genau dann, wenn v,$D_\theta(v)$ und $\mathbf{0}$ kollinear sind. Daher D_θ hat Eigenvektoren genau dann, wenn $\theta \in \{k \cdot \pi \mid k \in \mathbb{Z}\}$. Wenn $\theta = 2k\pi$ für $k \in \mathbb{Z}$, dann ist $D_\theta = \mathrm{id}_{\mathbb{R}^2}$. Wenn $\theta = (2k+1)\pi$ für $k \in \mathbb{Z}$, dann ist $D_\theta = -\mathrm{id}_{\mathbb{R}^2} = h_{-1}$.

4. Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ und sei $A \in \mathrm{Mat}_n(\mathbb{K})$ die Matrix

   $$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

   wobei alle Einträge die nicht auf der Diagonale liegen Null sind. Die Eigenwerte von A sind $\lambda_1, \ldots, \lambda_n$, weil $A \cdot e_i = \lambda \cdot e_i$. Falls die λ_i paarweise verschieden sind, dann haben wir

   $$\mathrm{Eig}(A, \lambda_i) = \mathrm{Span}_{\mathbb{K}}\{e_i\} \quad \text{und} \quad \mathbb{K}^n = \mathrm{Eig}(A, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(A, \lambda_n).$$

Wir werden in Satz 10.32 sehen, dass allgemeiner gilt. Der Beweis in dem allgemeinen Fall ist nicht sehr kurz, weil $V = U_1 \oplus \cdots \oplus U_m$ äquivalent zu

$$V = U_1 + \cdots + U_m \quad \text{und } U_i \cap \left(U_1 + \cdots + \widehat{U_i} + \cdots + U_m\right) = 0$$

ist. Um die zweite Bedingung zu zeigen werden wir Lemma 10.31 brauchen. In diesem Beispiel haben wir aber beide dieser Bedingungen offensichtlich erfüllt, weil jeder Unterraum von einem Vektor der Standardbasis erzeugt wird.

Satz 10.7. *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Seien $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $\lambda \in \mathbb{K}$. Folgende Aussagen sind äquivalent:*

(a) *λ ist ein Eigenwert von f.*

(b) *$\lambda \mathrm{id}_V - f$ ist nicht bijektiv.*

(c) *$\lambda \mathrm{id}_V - f$ ist nicht injektiv.*

(d) *$\lambda \mathrm{id}_V - f$ ist nicht surjektiv.*

(e) $\det(\lambda \mathrm{id}_V - f) = 0$.

Beweis-Skizze: (b)⇔(c)⇔(d) Das folgt aus dem Dimensionssatz (cf. Lemma 4.4) weil V endlich dimensional ist.

(a)⇒(c) Sei $v \in V$ ein Eigenvektor zum Eigenwert λ. Also $v \neq 0$ und wir haben

$$(\lambda \mathrm{id}_V - f)(v) = \lambda \mathrm{id}_V(v) - f(v) = \lambda v - \lambda v = 0 = (\lambda \mathrm{id}_V - f)(0).$$

Also $\lambda \mathrm{id}_V - f$ ist nicht injektiv.

(c)⇒(a) Weil $f - \lambda \mathrm{id}_V$ eine lineare Abbildung ist, existiert $0 \neq v \in \mathrm{Ker}(f - \lambda \mathrm{id}_V)$. Jeder solcher Vektor ist dann ein Eigenvektor zum Eigenwert λ, weil

$$0 = (\lambda \mathrm{id}_V - f)(v) = \lambda v - f(v).$$

(b)⇔(e) Ein Endomorphismus ist bijektiv genau dann, wenn jede zugeordnete Matrix invertierbar ist. Die Äquivalenz folgt dann aus Korollar 7.51. Q.E.D.

10.3 Das Charakteristische Polynom

Für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ können wird die Matrix $x \cdot I_n - A \in \mathrm{Mat}_n(\mathbb{K}[x])$ definieren. Die Determinante dieser Matrix ist also ein Polynom $\det(x \cdot I_n - A) \in \mathbb{K}[x]$. Aus Satz 10.7 bekommen wir das $\lambda \in \mathbb{K}$ ein Eigenwert von A ist genau dann, wenn λ eine Nullstelle des Polynoms $\det(x \cdot I_n - A)$ ist. In diesem Abschnitt werden wir uns mit diesem Polynom beschäftigen.

Definition 10.8. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ eine quadratische Matrix und sei $x \cdot I_n - A \in \mathrm{Mat}_n(\mathbb{K}[x])$. Das **charakteristische Polynom der Matrix** A ist

$$\chi_A(x) := \det(x \cdot I_n - A) \in \mathbb{K}[x].$$

Die **Spur** der Matrix $A = (a_{ij})$ ist die Zahl

$$\mathrm{Spur}(A) := a_{11} + \cdots + a_{nn} \in \mathbb{K}.$$

Bemerkung 10.9. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ aus der Leibniz-Formel bekommt man

$$\chi_A(x) = x^n - \mathrm{Spur}(A) \cdot x^{n-1} + \cdots + (-1)^n \cdot \det A.$$

Insbesondere ist χ_A ein normiertes Polynom[3] von Grad n.

Lemma 10.10. *Für zwei ähnliche Matrizen $A, B \in \mathrm{Mat}_n(\mathbb{K})$ gilt $\chi_A(x) = \chi_B(x)$.*

Beweis-Skizze: Seien $A, B \in \mathrm{Mat}_n(\mathbb{K})$ mit $A \sim B$. Es existiert also $T \in \mathrm{GL}_n(\mathbb{K})$ mit $A = T^{-1}BT$. Es gilt dann

$$\chi_A(x) = \det(xI_n - A) = \det\left(T^{-1}(xI_n)T - T^{-1}BT\right) = \det(T^{-1}) \cdot \det(xI_n - B) \cdot \det(T) = \chi_B(x).$$

Q.E.D.

Bemerkung 10.11. Die Umkehrung im obigen Lemma gilt nicht. Das heißt, es gibt Matrizen die nicht ähnlich sind, aber dasselbe charakteristische Polynom haben. Zum Beispiel I_2 und $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ sind nicht ähnlich (weil I_2 zu keiner andere Matix außer sich selbst ähnlich ist), haben aber

$$\chi_{I_2}(x) = \chi_A(x) = (x-1)^2.$$

Für einen Endomorphismus sind alle zugeordnete Matrizen zueinander ähnlich. Deswegen, dürfen wir nach Lemma 10.10 folgende Definition geben.

Definition 10.12. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$. Das **charakteristische Polynom des Endomorphismus** f ist das Polynom

$$\chi_f(x) = \chi_{M_f^B}(x),$$

wobei B eine (beliebige) geordnete Basis von V ist.

[4]2.5.'22

Satz 10.13. (a) *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Die Eigenwerte eines Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ sind die Nullstellen des charakteristischen Polynoms χ_f.*

(b) *Die Eigenwerte einer Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind die Nullstellen des charakteristischen Polynoms χ_A.*

Beweis-Skizze: Der Beweis folgt direkt aus Satz 10.7. Q.E.D.

Korollar 10.14. (i) Jeder Endomorphismus von V hat höchstens $\dim_{\mathbb{K}} V$ Eigenwerte.

(ii) Eine $n \times n$ Matrix hat höchstens n Eigenwerte.

[3] Ich erinnere, dass normiert heißt mit Leitkoeffizient 1. Das ist auch der Grund warum ich χ_A als $\det(xI_n - A)$, statt $\det(A - xI_n)$, definiert habe.

(iii) Ähnliche Matrizen haben dieselben Eigenwerte.

Wir werden die übrigen Aussagen in diesem Abschnitt nur für quadratische Matrizen formulieren. Eine gute Übung ist diese für Endomorphismen auszudrücken.

In Definition 10.8 haben wir eigentlich eine Abbildung definiert:

$$\chi : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \{p \in \mathbb{K}[x] \ : \ p \text{ ist normiert und hat Grad } n\}.$$

In Lemma 10.10 haben wir gezeigt, dass diese Abbildung auch eine Abbildung von der Faktormenge modulo Matrizenähnlichkeit definiert (siehe die Universelle Eigenschaft der Faktormenge 1.60):

$$\chi : \mathrm{Mat}_n(\mathbb{K})/_\sim \longrightarrow \{p \in \mathbb{K}[x] \ : \ p \text{ ist normiert und hat Grad } n\}.$$

In Bemerkung 10.11 haben wir gezeigt, dass, sogar als Abbildung von der Faktormenge, χ nicht injektiv ist. Der nächste Satz sagt uns, dass χ (in beiden Fällen) surjektiv ist.

Satz 10.15. *Jedes normierte Polynom $p \in \mathbb{K}[x]$ ist das charakteristische Polynom einer Matrix.*

Beweis-Skizze: Sei $p(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 \in \mathbb{K}[x]$ ein normiertes Polynom. Gesucht ist eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ mit $\chi_A(x) = p(x)$. Sei $A_p = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ die Matrix mit

$$\begin{array}{lcll} a_{i+1,i} & = & 1 & \forall\ i = 1, \ldots, n-1 \\ a_{in} & = & -c_{i-1} & \forall\ i = 1, \ldots, n \\ a_{ij} & = & 0 & \text{sonst.} \end{array}$$

Also A hat die Form

$$A_p = \begin{pmatrix} 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & \ldots & 0 & -c_1 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & 0 & \ddots & 0 & -c_{n-2} \\ 0 & 0 & \ldots & 1 & -c_{n-1} \end{pmatrix}. \tag{10.1}$$

Wir beweisen durch Induktion über $n = \deg p$, dass $\chi_{A_p}(x) = p(x)$.

$\boxed{n = 1}$ Dann sind $p = x + c_0$ und $A_p = (-c_0)$, also $\chi_{A_p}(x) = \det(x \cdot 1 - (-c_0)) = x + c_0 = p(x)$.

$\boxed{n - 1 \Rightarrow n}$ Wir haben

$$x \cdot I_n - A_p = \begin{pmatrix} x & 0 & \ldots & 0 & c_0 \\ -1 & x & \ldots & 0 & c_1 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & 0 & \ddots & x & c_{n-2} \\ 0 & 0 & \ldots & -1 & x + c_{n-1} \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{K}[x]).$$

Sei $q = x^{n-1} + c_{n-1}x^{n-2} + \cdots + c_2 x + c_1$ und A_q die entsprechende Matrix aus (10.1). Die Laplacesche

Entwicklung nach der ersten Zeile der Determinante von $x \cdot I_n - A_p$ gibt

$$\begin{aligned}
\chi_{A_p}(x) &= x \cdot \det(x \cdot I_{n-1} - A_q) + (-1)^{n+1} c_0 \cdot \det \begin{pmatrix} -1 & x & \dots & 0 & 0 \\ 0 & -1 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & -1 & x \\ 0 & 0 & \dots & 0 & -1 \end{pmatrix} \\
&= x\left(x^{n-1} + c_{n-1}x^{n-1} + \cdots + c_2 x + c_1\right) + (-1)^{n+1} \cdot (-1)^{n-1} \cdot c_0 \\
&= x^n + c_{n-1}x^{n-1} + \cdots + c_2 x^2 + c_1 x + c_0 = p(x).
\end{aligned}$$

Q.E.D.

Für jedes Polynom $p = c_n x^n + \cdots + c_1 x + c_0$ von Grad n (also mit $c_n \neq 0$), hat das Polynom $\frac{1}{c_n} \cdot p$ dieselben Nullstellen wie p. Wir haben also als Folgerung von Satz 10.15 die nächste Aussage.

Korollar 10.16. Ein Körper $\mathbb{K}$ ist algebraisch abgeschlossen genau dann, wenn jede quadratische Matrix mit Einträgen aus $\mathbb{K}$ einen Eigenwert besitzt.

Zum ersten Mal spielen zusätzliche Eigenschaften, die nicht für alle Körper gelten, eine Rolle.

Beispiel 10.17. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Wir können in beiden Fällen die Drehung $f_{D_\theta} : \mathbb{K}^2 \longrightarrow \mathbb{K}^2$ betrachten, wobei

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

Das Charakteristische Polynom ist $\chi_{D_\theta}(x) = x^2 - 2\cos\theta \cdot x + 1$. Das hat die komplexen Nullstellen

$$\lambda_{1,2} = \cos\theta \pm i \sin\theta.$$

Einerseits sieht man, dass, wenn $\mathbb{K} = \mathbb{R}$, muss $\sin\theta = 0$ gelten um eine reelle Nullstelle zu bekommen. Die Matrix $D_\theta \in \mathrm{Mat}_2(\mathbb{R})$ hat also Eigenwerte genau dann, wenn $\theta \in \{k \cdot \pi \mid k \in \mathbb{Z}\}$. Anderseits, für alle $\theta \in \mathbb{R}$ hat die Matrix $D_\theta \in \mathrm{Mat}_2(\mathbb{C})$ Eigenwerte. Für jedes $\theta \in \mathbb{R} \setminus \{k \cdot \pi \ : \ k \in \mathbb{Z}\}$ hat D_θ sogar zwei verschiedene Eigenwerte.

10.4 Vielfachheit von Eigenwerten

Die Begriffe aus Abschnitt 9.3, insbesondere Definition 9.20, werden hier vorausgesetzt.

Definition 10.18. Es seien V ein endlich dimensionaler $\mathbb{K}$-Vektorraum, $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus und $\lambda \in \mathbb{K}$ ein Eigenwert von f.

(a) Die **algebraische Vielfachheit** des Eigenwertes λ von f ist die Vielfachheit von λ als Nullstelle des Polynoms χ_f. Wir bezeichnen diese mit

$$\mu_a(f, \lambda) := \mu_a(\chi_f, \lambda) = \max\{a \in \mathbb{N} \ : \ (x - \lambda)^a \mid \chi_f\}.$$

(b) Die **geometrische Vielfachheit** des Eigenwertes λ von f ist die Dimension des $\mathbb{K}$-Eigenraumes $\mathrm{Eig}(f, \lambda)$. Wir bezeichnen diese mit

$$\mu_g(f, \lambda) := \dim_{\mathbb{K}} \mathrm{Eig}(f, \lambda).$$

Definition 10.19. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ eine quadratische Matrix und λ ein Eigenwert von A. Die algebraische (bzw. geometrische) Vielfachheit des Eigenwertes λ von A ist die algebraische (bzw. geometrische) Vielfachheit des Eigenwertes λ von f_A. Wir bezeichnen diese mit $\mu_a(A,\lambda)$ (bzw. $\mu_g(A,\lambda)$). Also

$$\begin{aligned} \mu_a(A,\lambda) &= \mu_a(\chi_A,\lambda) \\ \mu_g(A,\lambda) &= \dim_{\mathbb{K}} \mathrm{Eig}(A,\lambda). \end{aligned}$$

Bemerkung 10.20. Die geometrische Vielfachheit eines Eigenwertes λ gibt uns also die maximale Anzahl von linear unabhängigen Eigenvektoren zum Eigenwert λ.

Proposition 10.21. *Seien $A, B \in \mathrm{Mat}_n(\mathbb{K})$ zwei ähnliche Matrizen und sei $\lambda \in \mathbb{K}$ ein Eigenwert von A, also*[4] *auch von B. Es gelten*

(a) $\mu_a(A,\lambda) = \mu_a(B,\lambda)$.

(b) $\mu_g(A,\lambda) = \mu_g(B,\lambda)$.

Wenn die Matrix $T \in \mathrm{GL}_n(\mathbb{K})$ die Ähnlichkeit von A und B gibt, das heißt $A = T^{-1}BT$, dann gilt

(c) $f_T(\mathrm{Eig}(A,\lambda)) = \mathrm{Eig}(B,\lambda)$.

Insbesondere haben Eigenräume ähnlicher Matrizen zum selben Eigenwert nur dieselbe Dimension, diese müssen aber nicht gleich als $\mathbb{K}$-Unterraum sein.

Beweis-Skizze: (a) Folgt aus $\chi_A = \chi_B$ (Lemma 10.10).

(b) Weil $A \sim B$, existiert $T \in \mathrm{GL}_n(\mathbb{K})$, sodass $A = T^{-1}BT$. Oder äquivalent:

$$B = TAT^{-1}.$$

Behauptung: $f_T(\mathrm{Eig}(A,\lambda)) \subseteq \mathrm{Eig}(B,\lambda)$.

Sei $\mathbf{y} \in f_T(\mathrm{Eig}(A,\lambda))$, also $\mathbf{y} = T \cdot \mathbf{x}$, mit $\mathbf{x} \in \mathrm{Eig}(A,\lambda)$. Wir haben also $A \cdot \mathbf{x} = \lambda \cdot \mathbf{x}$. Es folgt

$$B \cdot \mathbf{y} = B \cdot f_T(\mathbf{x}) = (TAT^{-1}) \cdot (T \cdot \mathbf{x}) = T(A \cdot \mathbf{x}) = T(\lambda \cdot \mathbf{x}) = \lambda \cdot f_T(\mathbf{x}) = \lambda \cdot \mathbf{y}.$$

Also $\mathbf{y} \in \mathrm{Eig}(B,\lambda)$ und, weil $\mathbf{y}$ beliebig gewählt wurde, haben wir die Behauptung gezeigt.

Es folgt aus der Behauptung, dass

$$\dim_{\mathbb{K}} f_T(\mathrm{Eig}(A,\lambda)) \leqslant \dim_{\mathbb{K}} \mathrm{Eig}(B,\lambda). \tag{10.2}$$

Weil $T \in \mathrm{GL}_n(\mathbb{K})$, ist f_T ein Automorphismus von $\mathbb{K}^n$. Also $\dim_K f_T(U) = \dim_{\mathbb{K}} U$ für alle $U \subseteq_{UVR} \mathbb{K}^n$. Wir haben also aus (10.2)

$$\dim_{\mathbb{K}} \mathrm{Eig}(A,\lambda) \leqslant \dim_{\mathbb{K}} \mathrm{Eig}(B,\lambda).$$

[4] Aus Korollar 10.14 (iii).

Analog zeigt man $f_{T^{-1}}(\mathrm{Eig}(B,\lambda) \subseteq \mathrm{Eig}(A,\lambda)$, und somit auch die andere Ungleichung. Also $\mu_g(A,\lambda) = \mu_g(B,\lambda)$ folgt.

(c) Folgt aus $f_T(\mathrm{Eig}(B,\lambda)) \subseteq \mathrm{Eig}(A,\lambda)$ und $\mu_g(A,\lambda) = \mu_g(B,\lambda)$. Q.E.D.

Korollar 10.22. Für jeden Eigenwert $\lambda \in \mathbb{K}$ von $f \in \mathrm{End}_{\mathbb{K}}(V)$ und für jede geordnete Basis $\mathcal{B}$ von V gilt

$$\mu_a(f,\lambda) = \mu_a(M_f^{\mathcal{B}},\lambda) \qquad \text{und} \qquad \mu_g(f,\lambda) = \mu_g(M_f^{\mathcal{B}},\lambda).$$

Satz 10.23. *Es seien V ein endlich dimensionaler $\mathbb{K}$-Vektorraum, $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $\lambda \in \mathbb{K}$ ein Eigenwert von f. Es gilt*

$$\mu_g(f,\lambda) \leqslant \mu_a(f,\lambda).$$

Beweis-Skizze: Sei $r = \mu_g(f,\lambda)$, und seien $v_1,\dots,v_r \in \mathrm{Eig}(f,\lambda)$ linear unabhängige Vektoren. Wir ergänzen $v_1,\dots,v_r$ zu einer geordneten Basis $\mathcal{B}$ von V. Dann gilt

$$M_f^{\mathcal{B}} = \left(\begin{array}{c|c} I_r & * \\ \hline \mathbf{0} & A \end{array}\right)$$

Es folgt aus der wiederholter Laplace Entwicklung der Determinante nach der ersten r Spalten (siehe auch Hausaufgabe 3, Zusatzaufgabe 3), dass

$$\chi_f(x) = (x-\lambda)^r \cdot \chi_A(x).$$

Also $\mu_g(f,\lambda) = r \leqslant \mu_a(\chi_f,\lambda) = \mu_a(f,\lambda)$. Q.E.D.

Korollar 10.24. Sei $\lambda \in \mathbb{K}$ ein Eigenwert der Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$. Es gilt

$$\mu_g(A,\lambda) \leqslant \mu_a(A,\lambda).$$

Folgendes Beispiel zeigt, dass allgemein nur die Ungleichung zwischen den Vielfachheiten gilt.

Beispiel 10.25. Sei $A = \begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$. Wir haben $\chi_A = (x-3)^2$, also ist $3 \in \mathbb{R}$ der einzige Eigenwert, und es hat $\mu_a(A,3) = 2$. Der Eigenraum zu 3 findet man als Lösungsmenge des homogenes LGS $\mathcal{L}(3I_2 - A \mid \mathbf{0})$ also von

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Das gibt uns $\mathrm{Eig}(A,3) = \{(t,0) \in \mathbb{R}^2 \mid t \in \mathbb{R}\}$, also $\mu_g(A,3) = 1$.

Bemerkung 10.26. Wenn $\lambda \in \mathbb{K}$ ein Eigenwert von f mit $\mu_a(f,\lambda) = 1$, dann gilt $\mu_g(f,\lambda) = \mu_a(f,\lambda)$.

[5]4.5.'22

10.5 Diagonalisierbarkeit

Definition 10.27. Eine Matrix $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ ist eine **Diagonalmatrix** wenn $a_{ij} = 0$ für alle $1 \leqslant i \neq j \leqslant n$. Es reicht also die Einträge auf der Diagonale aufzulisten und wir schreiben dann

$$A = \mathrm{diag}(a_{11}, \ldots, a_{nn}).$$

Definition 10.28. (a) Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist **diagonalisierbar** wenn es eine geordnete Basis $\mathcal{B}$ gibt, sodass $M^{\mathcal{B}}(f)$ eine Diagonalmatrix ist.

(b) Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ ist **diagonalisierbar** wenn es einer Diagonalmatrix ähnlich ist (cf. Definition 10.3).

Lemma 10.29. *Für $f \in \mathrm{End}_{\mathbb{K}}(V)$ sind folgende Aussagen äquivalent.*

(a) *Es gibt eine Basis von V die nur aus Eigenvektoren besteht.*

(b) *f ist diagonalisierbar.*

Beweis-Skizze: (a)⇒(b) Seien $\mathcal{B} = v_1, \ldots, v_n$ die Basis und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, sodass für jeden $i = 1, \ldots, n$, der Vektor $v_i \in V$ ein Eigenvektor zum Eigenwert λ_i ist. Die λ_i müssen nicht unbedingt paarweise verschieden sein. Es gilt dann $M^{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.
(b)⇒(a) Wenn f diagonalisierbar ist, dann existiert per Definition eine geordnete Basis $\mathcal{B} = v_1, \ldots, v_n$ sodass $M^{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Es gilt für jedes $i = 1, \ldots, n$, dass $f(v_i) = \lambda_i v_i$. Somit sind alle Basisvektoren auch Eigenvektoren. Q.E.D.

Beispiel 10.30. Sei $A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$. Dann haben wir $A \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, also ist $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ein Eigenvektor von A zum Eigenwert 2. Wir haben auch $A \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$, also ist $v_2 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$ ein Eigenvektor von A zum Eigenwert 1. Die Menge $B' = \{v_1, v_2\}$ ist linear unabhängig, also, weil $\dim_{\mathbb{R}} \mathbb{R}^2 = 2$, ist es eine Basis aus Eigenvektoren. Also A ist diagonalisierbar. Um die Diagonalform zu berechnen, betrachten wir A als $M^{\mathcal{B}}(f_A)$ mit $\mathcal{B}$ die Standardbasis von $\mathbb{R}^2$. Die Basiswechselmatrix von der Standardbasis $\mathcal{B}$ zur Basis $\mathcal{B}'$ ist dann $M^{\mathcal{B}}_{\mathcal{B}'} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$. Weiterhin haben wir $(M^{\mathcal{B}}_{\mathcal{B}'})^{-1} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$. Also

$$M^{\mathcal{B}'}(f_A) = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Folgender Satz steht hinter der vollständigen Beschreibung diagonalisierbarer Endomorphismen.

Lemma 10.31. *Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$. Wenn $v_1, \ldots, v_r$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ sind, dann sind $v_1, \ldots, v_r$ linear unabhängig.*

Beweis-Skizze: Wir beweisen das durch vollständige Induktion über r.
$r = 1$ Die Aussage ist wahr weil der Eigenvektor $v_1 \neq 0$.
$r - 1 \Rightarrow r$ Wir nehmen an, dass $v_1, \ldots, v_{r-1}$ linear unabhängig sind. Seien $\mu_1, \ldots, \mu_r \in \mathbb{K}$, sodass

$$\mu_1 v_1 + \cdots + \mu_r v_r = 0. \tag{10.3}$$

Es folgt auf einer Seite, dass

$$0 = f(0) = f(\mu_1 v_1 + \cdots + \mu_r v_r) = \mu_1 \lambda_1 v_1 + \cdots + \mu_r \lambda_r v_r. \tag{10.4}$$

Anderseits, durch Multiplikation mit λ_r in (10.3) bekommen wir

$$0 = \lambda_r 0 = \mu_1 \lambda_r v_1 + \cdots + \mu_r \lambda_r v_r. \tag{10.5}$$

Wenn wir von (10.4) die Gleichung (10.5) abziehen, dann bekommen wir

$$0 = \mu_1(\lambda_1 - \lambda_r)v_1 + \cdots + \mu_{r-1}(\lambda_{r-1} - \lambda_r)v_{r-1}.$$

Aus der induktiven Voraussetzung folgt also

$$\mu_1(\lambda_1 - \lambda_r) = \cdots = \mu_{r-1}(\lambda_{r-1} - \lambda_r) = 0.$$

Da die λ_i paarweise verschieden sind, folgt $\mu_1 = \cdots = \mu_{r-1} = 0$. Aus der Gleichung (10.3) bleibt also nur

$$\mu_r v_r = 0$$

übrig, und, weil $v_r \neq 0$, folgt auch $\mu_r = 0$. Q.E.D.

In Lemma 10.29 haben wir bewiesen, dass ein Endomorphismus f genau dann diagonalisierbar ist, wenn es eine Basis aus Eigenvektoren besitzt. Wir erweitern jetzt die Beschreibung der Diagonalisierbarkeit.

Satz 10.32. *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus von V. Folgende Aussagen sind äquivalent.*

(a) *Der Endomorphismus f ist diagonalisierbar.*

(b) *Es gibt eine Basis von V die nur aus Eigenvektoren von f besteht.*

(c) (i) *Das charakteristische Polynom χ_f zerfällt vollständig in Linearfaktoren und*

(ii) *für jeden Eigenwert λ von f gilt $\mu_g(f, \lambda) = \mu_a(f, \lambda)$.*

(d) *Wenn $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ die verschiedenen Eigenwerte von f sind, dann gilt*

$$V = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_m).$$

Beweis-Skizze: Seien $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ die verschiedenen Eigenwerte von f.

(a)⇔(b) Das ist Lemma 10.29.

(b)⇒(c) Sei $\mathcal{B}$ die Basis aus Eigenvektoren, die wir so ordnen, dass

$$M^{\mathcal{B}}(f) = \begin{pmatrix} \boxed{\lambda_1 I_{m_1}} & & & \\ & \boxed{\lambda_2 I_{m_2}} & & \\ & & \ddots & \\ & & & \boxed{\lambda_r I_{m_r}} \end{pmatrix}, \tag{10.6}$$

Punkt (i) gilt weil $\chi_f = (x-\lambda_1)^{m_1} \cdot \dots \cdot (x-\lambda_r)^{m_r}$. Es folgt auch, dass $\mu_a(f,\lambda_i) = m_i$. Aus (10.6) folgt, dass es mindestens m_i linear unabhängige Vektoren zu dem Eigenwert λ_i gibt. Das heißt $\mu_g(f,\lambda_i) \geqslant \mu_a(f,\lambda_i)$ und aus Satz 10.23 folgt die Gleichheit und somit (ii).

(c)⇒(d) Wir definieren den $\mathbb{K}$-Unterraum $W = \mathrm{Eig}(f,\lambda_1)+\dots+\mathrm{Eig}(f,\lambda_r) \subseteq V$. Aus Lemma 10.31 zusammen mit der Beschreibung der direkten Summen aus Satz 4.9 folgt, dass

$$W = \mathrm{Eig}(f,\lambda_1) \oplus \dots \oplus \mathrm{Eig}(f,\lambda_r).$$

Wir haben also

$$\dim_{\mathbb{K}} W = \sum_{i=1}^{r} \dim_{\mathbb{K}} \mathrm{Eig}(f,\lambda_i) = \sum_{i=1}^{r} \mu_g(f,\lambda_i) = \sum_{i=1}^{r} \mu_a(f,\lambda_i) = \dim_{\mathbb{K}} V,$$

wobei die dritte Gleichheit aus (c)(ii) folgt und die vierte Gleichheit aus (c)(i). Also $W = V$.

(d)⇒(b) Wir wählen für jeden Eigenraum $\mathrm{Eig}(f,\lambda_i)$ eine geordnete Basis $\mathcal{B}_i$. Da V die direkte Summe der Eigenräume ist, folgt es, dass $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ eine geordnete Basis von V ist, bezüglich welcher (10.6) gilt. Q.E.D.

Bemerkung 10.33. Man kann den obigen Satz auch eine quadratische Matrix A formulieren, indem man f und V mit A beziehungsweise $\mathbb{K}^n$ ersetzt.

Bemerkung 10.34. (a) Zwei Diagonalmatrizen $A = \mathrm{diag}(a_1,\dots,a_n)$ und $B = \mathrm{diag}(b_1,\dots,b_n)$ aus $\mathrm{Mat}_n(\mathbb{K})$ sind ähnlich, genau dann, wenn es eine Permutation $\sigma \in S_n$ gibt, sodass $a_i = b_{\sigma(i)}$. Das gilt, weil die Einträge auf der Diagonale genau die Eigenwerte sind, und jeder Eigenwert λ_i kommt $\mu_a(A,\lambda_i)$ oft vor. Ähnliche Matrizen haben aber gleiche Eigenwerte mit gleicher Vielfachheit.

(b) Wenn $\mathbb{K}$ algebraisch abgeschlossen ist, dann gilt die Bedingung (i) aus Satz 10.32 Punkt (c) immer.

(c) Eine "allgemeine" Matrix in $\mathrm{Mat}_n(\mathbb{C})$ hat unterschiedliche Eigenwerte, und ist somit diagonalisierbar.

[6]9.5.'22

10.5.1 Das Verfahren zur Diagonalisierung

Sei $f \in \mathrm{End}_K(V)$ gegeben. Folgendes Verfahren bestimmt ob f diagonalisierbar ist, und falls ja, findet eine Diagonalform.

`Schritt` 1. Man stelle fest, ob $\chi_f(x)$ in Linearfaktoren zerfällt.

`Schritt` 2.
a. Falls "Nein", dann ist f nicht diagonalisierbar.
b. Falls "Ja", dann zerlege $\chi_f(x)$ in Linearfaktoren und somit berechne $\mu_a(f, \lambda_i)$.

`Schritt` 3. Man berechne die Eigenräume und somit $\mu_g(f, \lambda_i)$.

a. Falls es einen Eigenwert λ_i mit $\mu_g(f, \lambda_i) < \mu_a(f, \lambda_i)$ gibt, dann ist f nicht diagonalisierbar.
b. Sonst, man bilde die Basis aus Eigenvektoren als Vereinigung der Basen der Eigenräume.

Für eine Matrix A, man wendet das obige Verfahren für $f_A \in \mathrm{End}_{\mathbb{K}}(\mathbb{K}^n)$ an. Im letzten Schritt kann man dann genauer sein:

`Schritt` 3'. Sei $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ die geordnete Basis aus Eigenvektoren. Dann wählen wir

$$T = \left(\begin{array}{c|c|c|c} & & & \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_n \\ & & & \end{array}\right) \in \mathrm{GL}_n(\mathbb{K})$$

und bekommen

$$T^{-1}AT = \begin{pmatrix} \boxed{\lambda_1 I_{m_1}} & & & \\ & \boxed{\lambda_2 I_{m_2}} & & \\ & & \ddots & \\ & & & \boxed{\lambda_r I_{m_r}} \end{pmatrix}, \qquad \text{mit } m_i = \mu_a(A, \lambda_i).$$

Beispiel 10.35. Sei $f \in \mathrm{End}_R(\mathbb{R}^3)$, gegeben durch

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 + x_3 \\ -3x_1 - 2x_2 + 3x_3 \\ -2x_1 - 2x_2 + 3x_3 \end{pmatrix}.$$

Bezüglich der Standardbasis $\mathcal{B} = \{e_1, e_2, e_3\}$ von $\mathbb{R}^3$ haben wir also

$$M^{\mathcal{B}}(f) = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}.$$

Also

$$\begin{aligned}\chi_f(x) &= \det\begin{pmatrix} x & 1 & -1 \\ 3 & x+2 & -3 \\ 2 & 2 & x-3 \end{pmatrix} \\ &= x(x+2)(x-3) + 1\cdot(-3)\cdot 2 + (-1)\cdot 3\cdot 2 - \\ &\quad -(-1)\cdot 2\cdot(x+2) - 1\cdot 3\cdot(x-3) - (-3)\cdot 2\cdot x \\ &= x^3 - x^2 - x + 1 \\ &= (x-1)(x^2-1) \\ &= (x-1)^2(x+1).\end{aligned}$$

Es gibt also zwei Eigenwerte 1, mit $\mu_a(f,1) = 2$, und -1, mit $\mu_a(f,-1) = 1$. Es gilt

$$\mathrm{Eig}(f,1) = \mathcal{L}\left(\left(\begin{array}{ccc} 1 & 1 & -1 \\ 3 & 3 & -3 \\ 2 & 2 & -2 \end{array}\right) \middle| \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = \mathrm{Span}_{\mathbb{K}}\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}.$$

Also $\mu_g(f,1) = 2 = \mu_a(f,1)$.
Da $\mu_a(f,-1) = 1$ und jeder Eigenwert geometrische Vielfachheit mindestens 1 hat, gilt auch $\mu_g(f,-1) = \mu_a(f,1) = 1$, also f ist diagonalisierbar. Wir suchen noch eine Basis aus Eigenvektoren. Wir haben schon eine Basis von $\mathrm{Eig}(f,1)$, wir suchen also nur noch eine Basis für $\mathrm{Eig}(f,-1)$. Wir bekommen (z.B. mit dem Gaußschen Algorithmus), dass

$$\mathrm{Eig}(f,-1) = \mathcal{L}\left(\left(\begin{array}{ccc} -1 & 1 & -1 \\ 3 & 1 & -3 \\ 2 & 2 & -4 \end{array}\right) \middle| \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = \{(t,3t,2t) \mid t \in \mathbb{R}\} = \mathrm{Span}_{\mathbb{K}}\left\{\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}\right\}.$$

Wir haben also die Basis $\mathcal{B} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ und setzen

$$T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}.$$

Wieder mit dem gaußschen Algorithmus bekommen wir

$$T^{-1} = \frac{1}{2}\cdot\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Kontrolle:

$$\frac{1}{2}\cdot\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}\cdot\begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}\cdot\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

10.6 Trigonalisierbare Endomorphismen

Definition 10.36. Eine Matrix $A = (a_{ij})_{i,j\in\{1,\dots,n\}} \in \mathrm{Mat}_n(\mathbb{K})$ heißt **obere Dreiecksmatrix** falls

$$a_{ij} = 0 \qquad \text{für alle } 1 \leqslant j < i \leqslant n$$

gilt. Das heißt, dass alle Einträge unterhalb der Diagonale Null sind.

Definition 10.37. (a) Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist **trigonalisierbar** falls es eine geordnete Basis $\mathcal{B}$ von V gibt, sodass $M^{\mathcal{B}}(f)$ eine obere Dreiecksmatrix ist.

(b) Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ ist **trigonalisierbar** wenn sie einer oberen Dreiecksmatrix ähnlich ist.

Bemerkung 10.38. Die Einträge auf der Diagonale einer oberen Dreiecksmatrix in $\mathrm{Mat}_n(\mathbb{K})$ sind genau die Eigenwerte der Matrix.

Satz 10.39. *Für eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$, beziehungsweise für einen Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$, sind die folgenden Aussagen äquivalent:*

(i) *A, beziehungsweise f, ist trigonalisierbar.*

(ii) *Das charakteristische Polynom von A, beziehungsweise von f, zerfällt vollständig in Linearfaktoren.*

Beweis-Skizze: Wir beweisen den Satz nur für Matrizen. Für Endomorphismen folgt dann die Aussage indem man eine Matrix zuordnet.

(i)⇒ (ii) Sei $T \in \mathrm{GL}_n(\mathbb{K})$ sodass

$$T^{-1}AT = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n-1} & b_{1n} \\ 0 & b_{22} & \dots & b_{2n-1} & b_{2n} \\ 0 & 0 & \ddots & b_{3n-1} & b_{3n} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & b_{nn} \end{pmatrix}$$

Aus Lemma 10.10 haben wir $\chi_A(x) = \chi_{T^{-1}AT}(x)$. Durch sukzessive Anwendung der Laplace Entwicklung folgt also

$$\chi_A(x) = \chi_{T^{-1}AT}(x) = (x - b_{11})(x - b_{22}) \dots (x - b_{nn}).$$

(ii)⇒ (i) Wir beweisen diese Richtung durch vollständige Induktion über n.
Im Fall "$n = 1$" ist jede Matrix eine obere Dreiecksmatrix, also da ist nichts zu beweisen.
Wir zeigen jetzt "$n - 1 \Rightarrow n$". Die induktive Voraussetzung ist, dass wenn das charakteristische Polynom einer Matrix in $\mathrm{Mat}_{n-1}(\mathbb{K})$ vollständig in Linearfaktoren zerfällt, dann ist die $(n-1) \times (n-1)$ Matrix trigonalisierbar. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ mit $\chi_A = (x-\lambda_1) \dots (x-\lambda_n)$ wobei $\lambda_i \in \mathbb{K}$ für alle i. Also, nach Satz 10.13 ist λ_1 ein Eigenwert von A. Es existiert also ein Eigenvektor $v_1 \in \mathbb{K}^n \setminus \{0\}$ zum Eigenwert λ_1. Weil $v_1 \neq 0$, können wir $\{v_1\}$ zu einer Basis $\mathcal{B} = \{v_1, \dots, v_n\}$ von $\mathbb{K}^n$ ergänzen.

Wir bilden die Matrix $T = \left(v_1 \mid v_2 \mid \ldots \mid v_n\right) \in \mathrm{GL}_n(\mathbb{K})$. Dann, weil $M^{\mathcal{B}}(f_A) = T^{-1}AT$ haben wir

$$T^{-1}AT = \left(\begin{array}{c|ccc} \lambda_1 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & \mathrm{A}' & \\ 0 & & & \end{array}\right).$$

Und es gilt $\chi_A = (x-\lambda)\chi_{A'}$, und somit zerfällt auch $\chi_{A'}$ vollständig in Linearfaktoren. Nach der Induktionsvoraussetzung existiert also $S' \in \mathrm{GL}_{n-1}(\mathbb{K})$ mit

$$(S')^{-1}A'S' = \begin{pmatrix} b'_{22} & b'_{23} & \ldots & b'_{2n-2} & b'_{2n-1} \\ 0 & b'_{33} & \ldots & b'_{3n-2} & b'_{3n-1} \\ 0 & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & b'_{n-1n-1} \end{pmatrix}$$

Wir setzen

$$S = \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \mathrm{S}' & \\ 0 & & & \end{array}\right)$$

Weil $\det S = 1 \cdot \det(S') \neq 0$ folgt $S \in \mathrm{GL}_n(\mathbb{K})$. Per Definition haben wir dann

$$S^{-1}(T^{-1}AT)S = (TS)^{-1}A(TS) = \left(\begin{array}{c|cccc} \lambda_1 & * & * & \ldots & * & * \\ \hline 0 & b'_{22} & b'_{23} & \ldots & b'_{2n-2} & b'_{2n-1} \\ 0 & 0 & b'_{33} & \ldots & b'_{3n-2} & b'_{3n-1} \\ 0 & 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & b'_{n-1n-1} \end{array}\right)$$

Q.E.D.

[7]11.5.'22

Bemerkung 10.40. Wenn $\mathbb{K}$ algebraisch abgeschlossen ist, dann ist die Bedingung (ii) in Satz 10.39 immer erfüllt. Insbesondere, alle Matrizen in $\mathrm{Mat}_n(\mathbb{C})$ sind trigonalisierbar. Das ist aber nicht völlig zufrieden stellend, weil eine Matrix zu vielen oberen Dreiecksmatrizen ähnlich sein kann, und es unter denen keine kanonische Wahl gibt.

Definition 10.41. Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$. Ein $\mathbb{K}$-Untervektorraum $U \subseteq V$ ist **f-invariant**, wenn

$$f(U) \subseteq U.$$

Beispiele:

1. Alle $\mathbb{K}$-Unterräume sind $h_\lambda = \lambda \,\mathrm{id}_V$ invariant, für jeden $\lambda \in \mathbb{K}$.
2. Die Drehung $d_\theta \in \mathrm{End}_{\mathbb{R}}\,\mathbb{R}^2$ mit Winkel θ hat für $\theta \notin \pi \cdot \mathbb{Z}$ keine invariante echte[5] Unterräume

Bemerkung 10.42. Wenn ein $\mathbb{K}$-Vektorraum V als Direkte Summe invarianter $\mathbb{K}$-Unterräume geschrieben werden kann: $V = U_1 \oplus \cdots \oplus U_r$, mit $f(U_i) \subseteq U_i$, dann gibt es eine Basis $\mathcal{B}$ von V, sodass $M_f^{\mathcal{B}}$ aus Blöcke der Größe $\dim_K U_i$ auf der Diagonale besteht, und sonst Null.

Definition 10.43. Sei V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n$. Eine **Fahne** ist eine Kette von $n+1$ $\mathbb{K}$-Untervektorräume mit der Eigenschaft:

$$\{0\} = W_0 \subsetneq W_1 \subsetneq \cdots \subsetneq W_n = V.$$

Insbesondere, gilt $\dim_{\mathbb{K}} W_i = i$ für alle $i = 0, \ldots, n$. Eine Fahne heißt ***f*-invariant** wenn $f(V_i) \subseteq V_i$ für alle $i = 0, \ldots, n$.

Korollar 10.44. Ein Endomorphismus $f \in \mathrm{End}_K(V)$ ist trigonalisierbar, genau dann, wenn es eine f-invariante Fahne in V gibt.

10.7 Der Satz von Cayley-Hamilton

Wir haben schon gesehen, dass sowohl $\mathrm{Mat}_n(\mathbb{K})$ also auch $\mathrm{End}_{\mathbb{K}}(V)$ eine $\mathbb{K}$-Algebra Struktur haben, und dass für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$, beziehungsweise für jeden Endomorphismus f, es ein eindeutiger bestimmter Homomorphismus von $\mathbb{K}$-Algebren $\mathrm{ev}_A : \mathbb{K}[x] \longrightarrow \mathrm{Mat}_n(\mathbb{K})$, beziehungsweise $\mathrm{ev}_f : \mathbb{K}[x] \longrightarrow \mathrm{End}_{\mathbb{K}}(V)$, gibt, mit $\mathrm{ev}_A(x) = A$, beziehungsweise $\mathrm{ev}_f(x) = f$. Wir bezeichnen mit

$$\mathbb{K}[A] := \mathrm{Bild}(ev_A) \qquad \text{beziehungsweise} \qquad \mathbb{K}[f] := \mathrm{Bild}(ev_f).$$

Als Mengen haben wir also $\mathbb{K}[A] = \{c_d A^d + \cdots + c_1 A + c_0 \;:\; d \in \mathbb{N} \text{ und } c_i \in \mathbb{K}\}$ und analog für $\mathbb{K}[f]$.

Bemerkung 10.45. Die Bilder $\mathbb{K}[A]$, beziehungsweise $\mathbb{K}[f]$ der Evaluationsabbildungen sind kommutative Teilringe von $\mathrm{Mat}_n(\mathbb{K})$, beziehungsweise $\mathrm{End}_{\mathbb{K}}(V)$. Das heißt:

1. $I_n = 1(A) \in \mathbb{K}[A]$ beziehungsweise $\mathrm{id}_V = 1(f) \in \mathbb{K}[f]$.
2. $(\mathbb{K}[A], +) \leq (\mathrm{Mat}_n(\mathbb{K}), +)$, beziehungsweise $(\mathbb{K}[f], +) \leq (\mathrm{End}_{\mathbb{K}}(V))$ [6].
3. $B, C \in \mathbb{K}[A] \Rightarrow B \cdot C \in \mathbb{K}[A]$, beziehungsweise $g, h \in \mathbb{K}[f] \Rightarrow g \circ h \in \mathbb{K}[f]$.
4. $\forall B, C \in \mathbb{K}[A]$ gilt $BC = CB$, beziehungsweise $\forall g, h \in \mathbb{K}[f]$ gilt $g \circ h = h \circ g$.

Damit sind insbesondere $\mathbb{K}[A]$ und $\mathbb{K}[f]$ selber kommutative Ringe, und wir können über $\mathrm{Mat}_m(\mathbb{K}[A])$ und $\mathrm{Mat}_m(\mathbb{K}[f])$ sprechen. Also gibt es auch die Determinante von Matrizen deren Einträge Matrizen, beziehungsweise Endomorphismen sind. Man muss nur aufpassen, dass für $B \in \mathrm{Mat}_m(\mathbb{K}[A])$ gilt

$$\det(B) \in \mathbb{K}[A] \subseteq \mathrm{Mat}_n(\mathbb{K}),$$

also die Determinante ist selber eine Matrix. Analog ist $\det(B) \in \mathbb{K}[f] \subseteq \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Der nächste Satz, zeigt dass eine Matrix, beziehungsweise ein Endomorphismus, ihre charakteristische Gleichung erfüllt.

[5] Das heißt nicht der Nullraum oder der gesamte Raum.
[6] wobei $\leq$ als "ist Untergruppe von" zu lesen ist

Satz 10.46 (Cayley[7]-Hamilton[8]). *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, und $n \in \mathbb{N}_{>0}$.*

(a) *Für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ gilt für das charakteristische Polynom χ_A von A, dass*

$$\chi_A(A) = 0.$$

(b) *Für jeden Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ gilt für das charakteristische Polynom χ_f von f, dass*

$$\chi_f(f) = 0.$$

Beweis-Skizze: Wir beweisen nur Teil (a). Teil (b) folgt dann aus der Identität $\chi_f = \chi_{M^{\mathcal{B}}(f)}$ für jede geordnete Basis $\mathcal{B}$ von V.

(a) Sei $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$. Wir definieren nun eine $n \times n$ Matrix dessen Einträge selber Matrizen aus $\mathbb{K}[A]$ sind:

$$C := (c_{ij}) = \begin{pmatrix} A - a_{11}I_n & -a_{12}I_n & \dots & -a_{1n}I_n \\ -a_{21}I_n & A - a_{22}I_n & \dots & -a_{2n}I_n \\ \vdots & & \ddots & \vdots \\ -a_{n1}I_n & -a_{n2}I_n & \dots & A - a_{1n}I_n \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{K}[A]).$$

Es gilt also für $i, j \in \{1, \dots, n\}$, dass

$$c_{ij} = \begin{cases} A - a_{ij}I_n & , \text{ wenn } i = j, \\ -a_{ij}I_n & , \text{ wenn } i \neq j. \end{cases}$$

Wir haben also jeden Eintrag der Matrix $xI_n - A \in \mathrm{Mat}_n(\mathbb{K}[x])$ mit dessen Wert unter ev_A ersetzt. Da ev_A ein Homomorphismus von $\mathbb{K}$-Algebren ist, gilt

$$\det C = \mathrm{ev}_A\left(\det(xI_n - A)\right) = \mathrm{ev}_A\left(\chi_A(x)\right) = \chi_A(A).$$

Also unser Ziel ist zu zeigen, dass $\det C = 0$, wobei hier 0 für die Nullmatrix in $\mathrm{Mat}_n(\mathbb{K})$ steht. Für $i \in \{1, \dots, n\}$ bezeichnen wir mit $e_i \in \mathbb{K}^n$ den Spaltenvektor mit 1 an der i-ten Stelle und 0 sonst. Wir werden folgende Gleichheiten später im Beweis brauchen:

$$\sum_{i=1}^{n} c_{ij} \cdot e_i = A \cdot e_j - \sum_{i=1}^{n} a_{ij} e_i = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} - \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = 0, \qquad \forall\ j = 1, \dots, n. \tag{10.7}$$

Behauptung: Für $k = 1, \dots, n$ gilt $\det(C) \cdot e_k = 0$.
Zu erst bemerken wir, dass aus der Behauptung folgt, dass jede Spalte der Matrix $\det C$ die Nullspalte ist, also $\det C = 0$. Wir brauchen also nur noch die Behauptung zu zeigen.

[7] Arthur Cayley, 1821-1895, englischer Mathematiker
[8] William Rowan Hamilton, 1805-1865, irischer Mathematiker und Physiker

Beweis der Behauptung. Sei $k \in \{1, \ldots, n\}$. Sei $\widetilde{C} = (\widetilde{c}_{ij}) \in \mathrm{Mat}_n(\mathbb{K}[A])$ die komplementäre Matrix von C (cf. Definition 7.48). Wir erinnern, dass nach Proposition 7.50 gilt $C \cdot \widetilde{C} = \widetilde{C} \cdot C = \det(C) \cdot I_n(\mathbb{K}[A])$, wobei $I_n(\mathbb{K}[A])$ die Einheitsmatrix in $\mathrm{Mat}_n(\mathbb{K}[A])$ bezeichnet. Das heißt, dass wenn wir die Einträge von $C \cdot \widetilde{C}$ mit c'_{ij} bezeichnen, es gilt für alle $i, j \in \{1, \ldots, n\}$, dass

$$c'_{ij} = \begin{cases} \det(C) & , \text{ wenn } i = j, \\ 0 & , \text{ wenn } i \neq j. \end{cases}$$

Wir bekommen also

$$\begin{aligned} \det(C) \cdot e_k = c'_{kk} e_k &= \textstyle\sum_{i=1}^n c'_{ik} e_i \\ &= \textstyle\sum_{i=1}^n \left(\sum_{j=1}^n c_{ij} \widetilde{c}_{jk} \right) e_i \\ &= \textstyle\sum_{i=1}^n \left(\sum_{j=1}^n \widetilde{c}_{jk} c_{ij} \right) e_i \\ &= \textstyle\sum_{i=1}^n \sum_{j=1}^n \widetilde{c}_{jk} \left(c_{ij} e_i \right) \\ &= \textstyle\sum_{j=1}^n \widetilde{c}_{jk} \left(\sum_{i=1}^n c_{ij} \cdot e_i \right) \\ &= \textstyle\sum_{j=1}^n \widetilde{c}_{jk} \cdot 0 \\ &= 0. \end{aligned}$$

Wobei die erste Gleichheit per Definition von c'_{ik} gilt; die zweite per Definition des Matrizenproduktes; die dritte aus der Kommutativität von $\mathbb{K}[A]$; die vierte und fünfte folgen aus Umschreiben der Summe (grundsätzlich Distributivität, Assoziativität, und Kommutativität der Addition); die sechste folgt aus (10.7). Q.E.D.

10.7.1 Das Minimalpolynom

Der Satz von Cayley-Hamilton liefert für jede Matrix A und für jeden Endomorphismus f jeweils ein Polynom $p \in \mathbb{K}[x]$ mit der Eigenschaft $p(A) = 0 \in \mathbb{K}[A] \subseteq \mathrm{Mat}_n(\mathbb{K})$, beziehungsweise $p(f) = 0 \in \mathbb{K}[f] \subseteq \mathrm{End}_{\mathbb{K}}(V)$. Wir suchen jetzt ein/das kleinstmögliche Polynom mit dieser Eigenschaft. "Klein" wird in diesem Fall bezüglich der der Teilbarkeit[9] sein. Die Eindeutigkeit wird "bis auf Multiplikation mit einer Konstante" gelten. Genauer gesagt gilt:

Satz 10.47. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Es existiert genau ein normiertes Polynom $\mathrm{mPol}_f \in \mathbb{K}[x]$ mit den Eigenschaften:*

(i) $\mathrm{mPol}_f(f) = 0$.

(ii) *Wenn $q \in \mathbb{K}[x]$ und $q(f) = 0$, dann $\mathrm{mPol}_f \,|q$.*

[9] Teilbarkeit ist eine partielle Ordnung auf $\mathbb{K}[x]$ modulo Multiplikation mit einer Konstanten.

Beweis-Skizze: Aus Satz 10.46 ist die Menge $\{p \in \mathbb{K}[x] \setminus \{0\} \ : \ p(f) = 0\}$ nicht leer. Man kann also

$$d_0 = \min\{\deg p \ : \ p \in \mathbb{K}[x] \setminus \{0\} \quad \text{und} \quad p(f) = 0\}$$

definieren, und $p \in \mathbb{K}[x]$ mit $p(f) = 0$ und $\deg(p) = d_0$ wählen. Sei $p = b_{d_0}x^{d_0} + \cdots + b_1 x + b_0$, mit $b_{d_0} \neq 0$. Wir definieren das normierte Polynom

$$\mathrm{mPol}_f = \frac{1}{b_{d_0}} \cdot p.$$

Per Definition gilt (i). Für (ii), sei $q \in \mathbb{K}[x]$ mit $q(f) = 0$. Aus der Division mit Rest in $\mathbb{K}[x]$ existieren eindeutige $s, r \in \mathbb{K}[x]$ mit $q = \mathrm{mPol}_f \cdot s + r$ und $\deg(r) < \deg(\mathrm{mPol}_f) = d_0$. Weil ev_f ein $\mathbb{K}$-Algebrahomomorphismus ist, haben wir

$$0 = q(f) = \mathrm{mPol}_f(f) \cdot s(f) + r(f) = 0 \cdot s(f) + r(f) = r(f).$$

Weil $\deg(r) < d_0$, muss also $r = 0 \in \mathbb{K}[x]$ sein, und das heißt $\mathrm{mPol}_f \,|q$. Also (ii) ist auch bewiesen, und somit die Existenz.
Zur Eindeutigkeit, sei $p' \in \mathbb{K}[x]$ ein normiertes Polynom das (i) und (ii) erfüllt. Es folgt also, dass $\mathrm{mPol}_f \,|p'$ und $p'|\,\mathrm{mPol}_f$. Das heißt, dass $\deg(\mathrm{mPol}_f) = \deg(p')$, und somit muss es eine Konstante $c \in \mathbb{K}$ existieren, sodass $p' = c \cdot \mathrm{mPol}_f$. Weil beide Polynome normiert sind, muss $c = 1$ sein, und somit $p' = \mathrm{mPol}_f$. Q.E.D.

Eine analoge Aussage gilt mit einem analogen Beweis für quadratische Matrizen:

Satz 10.48. *Seien $n \in \mathbb{N}_{>0}$ und $A \in \mathrm{Mat}_n(\mathbb{K})$ eine Matrix. Es existiert genau ein normiertes Polynom $\mathrm{mPol}_A \in \mathbb{K}[x]$ mit den Eigenschaften:*

(i) $\mathrm{mPol}_A(A) = 0$.

(ii) *Wenn $q \in \mathbb{K}[x]$ und $q(A) = 0$, dann $\mathrm{mPol}_f \,|q$.*

Definition 10.49. (a) Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$. Das **Minimalpolynom** von f ist das Polynom mPol_f aus Satz 10.47.

(b) Sei $A \in \mathrm{Mat}_n(\mathbb{K})$. Das **Minimalpolynom** von A ist das Polynom mPol_A aus Satz 10.48.

Korollar 10.50. Sei $n \in \mathbb{N}_{>0}$ und V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n$. Seien $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$. Es gelten

(a) $\mathrm{mPol}_f \,|\chi_f$ und $\mathrm{mPol}_A \,|\chi_A$.

(b) $\deg(\mathrm{mPol}_f) \leqslant n$ und $\deg(\mathrm{mPol}_A) \leqslant n$.

Bemerkung 10.51. Im Beweis von Satz 10.47 haben wir $\chi_f(f) = 0$ angewendet. Es hätte aber gereicht, dass es irgendein Polynom $p \in \mathbb{K}[x]$ existiert mit $p(f) = 0$. Das kann man elementar zeigen (Übung). Der Satz von Cayley-Hamilton ist aber essentiell für die Ungleichung $\deg(\mathrm{mPol}_f) \leqslant n$.

[8]16.5.'22

Satz 10.52. *Seien $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$. Die Nullstellen von mPol_f, beziehungsweise von mPol_A sind genau die Eigenwerte von f, beziehungsweise von A.*

Beweis-Skizze: Wir zeigen den Satz nur für A. Für f funktioniert der Beweis analog. Wir müssen also zeigen, dass, für $\lambda \in \mathbb{K}$ gilt $\mathrm{mPol}_A(\lambda) = 0 \iff \chi_A(\lambda) = 0$.
$\boxed{\Rightarrow}$ Sei $\lambda \in \mathbb{K}$ mit $\mathrm{mPol}_A(\lambda) = 0$. Wir haben, dass $\mathrm{mPol}_A \,|\chi_A$, also es existiert $s \in \mathbb{K}[x]$, sodass $\chi_A = \mathrm{mPol}_A \cdot s$. Also $\chi_A(\lambda) = \mathrm{mPol}_A(\lambda) \cdot s(\lambda) = 0 \cdot s(\lambda) = 0$.
$\boxed{\Leftarrow}$ Sei $\lambda \in \mathbb{K}$ mit $\chi_A(\lambda) = 0$. Aus Satz 10.13 ist λ ein Eigenwert von A, also es existiert $\mathbf{x} \in \mathbb{K}^n \setminus \{0\}$, sodass $A \cdot \mathbf{x} = \lambda \cdot \mathbf{x}$. Sei $\mathrm{mPol}_A = x^m + c_{m-1}x^{m-1} + \cdots + c_0$. Es gilt

$$\begin{aligned} 0 = 0 \cdot \mathbf{x} &= \mathrm{mPol}_A(A) \cdot \mathbf{x} = (A^m + c_{m-1}A^{m-1} + \cdots + c_1 A + c_0 I_n) \cdot \mathbf{x} \\ &= (\lambda^m + c_{m-1}\lambda^{m-1} + \cdots + c_1\lambda + c_0) \cdot \mathbf{x} = \mathrm{mPol}_A(\lambda) \cdot \mathbf{x}. \end{aligned}$$

Weil $\mathbf{x} \neq 0$, muss $\mathrm{mPol}_A(\lambda) = 0$ gelten. Q.E.D.

[9]18.5.'22
Übersicht VL

10.8 Das Lemma von Fitting

Für jeden $f \in \mathrm{End}_{\mathbb{K}}(V)$ gilt $v \in \mathrm{Ker}\, f^i \Rightarrow f^{i+1}(v) = f(f^i(v)) = f(0) = 0$. Wir haben also eine steigende Kette

$$\{0\} = \mathrm{Ker}\, f^0 \subseteq \mathrm{Ker}\, f \subseteq \mathrm{Ker}\, f^2 \subseteq \cdots \subseteq \mathrm{Ker}\, f^n \subseteq \ldots. \tag{10.8}$$

Bemerkung 10.53. Sobald zwei konsekutive Kerne gleich sind, dann wird diese Kette stationär. Das heißt

$$\mathrm{Ker}\, f^i = \mathrm{Ker}\, f^{i+1} \Rightarrow \mathrm{Ker}\, f^i = \mathrm{Ker}\, f^{i+k} \quad \forall\, k \in \mathbb{N}. \tag{10.9}$$

Beweis-Skizze: Wir zeigen das durch Induktion nach k.

Wenn die Aussage für $k-1$ gilt, also wenn $\mathrm{Ker}\, f^i = \mathrm{Ker}\, f^{i+k}$, dann wollen wir zeigen, dass $\mathrm{Ker}\, f^{i+k} = \mathrm{Ker}\, f^i$. Sei $v \in \mathrm{Ker}\, f^{i+k}$ beliebig. Es gilt $0 = f^{i+k}(v) = f^{i+k-1}(f(v))$ und somit $f(v) \in \mathrm{Ker}\, f^{i+k-1} = \mathrm{Ker}\, f^i$. Das heißt $f^{i+1}(v) = f^i(f(v)) = 0$, also $v \in \mathrm{Ker}\, f^{i+1} = \mathrm{Ker}\, f^i$, und die Aussage ist bewiesen. Q.E.D.

Für die Bilder der Potenzen von f haben wir eine fallende Inklusionskette:

$$V = \mathrm{Bild}\, f^0 \supseteq \mathrm{Bild}\, f^1 \supseteq \mathrm{Bild}\, f^2 \supseteq \cdots \supseteq \mathrm{Bild}\, f^n \supseteq \ldots. \tag{10.10}$$

Man kann auch in diesem Fall zeigen, dass, wenn $\mathrm{Bild}\, f^i = \mathrm{Bild}\, f^{i+1}$, dann $\mathrm{Bild}\, f^i = \mathrm{Bild}\, f^{i+k}\ \forall\, k \in \mathbb{N}$:

$$w \in \mathrm{Bild}\, f^i \quad \Rightarrow \quad w = f^i(v_0) = f^{i+1}(v_1) = f(f^i(v_1)) = f(f^{i+1}(v_2)) \ldots \text{usw.}$$

Wenn $\dim_{\mathbb{K}} V = \infty$, dann können alle Inklusionen der Ketten (10.8) und (10.10) streng[10] sein. Aber wenn V endlichdimensional ist, werden beide Ketten immer stationär, und das passiert sobald es eine Gleichheit gibt.

Bemerkung 10.54. Wir haben bis jetzt nur festgestellt, dass es minimale $i_1 \in \mathbb{N}$ und $i_2 \in \mathbb{N}$ gibt, sodass $\mathrm{Ker}\, f^{i_1} = \mathrm{Ker}\, f^{i_1+k}$ und $\mathrm{Bild}\, f^{i_2} = \mathrm{Bild}\, f^{i_2+k}$ für alle $k \in \mathbb{N}$. Aus dem Dimensionssatz wissen, dass

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \mathrm{Ker}\, f^i + \dim_{\mathbb{K}} \mathrm{Bild}\, f^i.$$

Es folgt aus (10.8) und (10.10), dass die minimale Potenzen gleich sind $i_1 = i_2 =: i_o$. Das heißt, die Ketten (10.8) und (10.10) werden an derselben Stelle stationär.

[10] d.h. $\subsetneq$

Definition 10.55. Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$. Sei $i_0 \in \mathbb{N}$ mit der Eigenschaft, dass $\mathrm{Ker}\, f^{i_0} = \mathrm{Ker}\, f^{i_0+1}$ und $\mathrm{Bild}\, f^{i_0} = \mathrm{Bild}\, f^{i_0+1}$. Wir bezeichnen den $\mathbb{K}$-Unterraum, der gleich mit $\mathrm{Ker}\, f^{i_0+k}$ für alle $k \in \mathbb{N}$ ist, mit

$$\mathrm{Ker}\, f^{\infty} := \mathrm{Ker}\, f^{i_0}.$$

Wir bezeichnen den $\mathbb{K}$-Unterraum, der gleich mit $\mathrm{Bild}\, f^{i_0+k}$ für alle $k \in \mathbb{N}$ ist, mit

$$\mathrm{Bild}\, f^{\infty} := \mathrm{Bild}\, f^{i_0}.$$

Lemma 10.56 (von Fitting[11])**.** *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus von V. Sei*

$$i_0 := \min\{i \ \in \mathbb{N} \mid \ \mathrm{Ker}\, f^i = \mathrm{Ker}\, f^{i+1}\}.$$

Es gelten

(a) $\mathrm{Ker}\, f^{\infty}$ *und* $\mathrm{Bild}\, f^{\infty}$ *sind f-invariante $\mathbb{K}$-Untervektorräume von V.*

(b) $f|_{\mathrm{Ker}\, f^{\infty}} : \mathrm{Ker}\, f^{\infty} \longrightarrow \mathrm{Ker}\, f^{\infty}$ *ist nilpotent, mit* $\mathrm{mPol}_{f|_{\mathrm{Ker}\, f^{\infty}}}(x) = x^{i_0}$.

(c) $f|_{\mathrm{Bild}\, f^{\infty}} : \mathrm{Bild}\, f^{\infty} \longrightarrow \mathrm{Bild}\, f^{\infty}$ *ist ein Isomorphismus.*

(d) $V = \mathrm{Ker}\, f^{\infty} \oplus \mathrm{Bild}\, f^{\infty}$.

(e) *Es gibt eine geordnete Basis $\mathcal{B}$ von V mit*

$$M^{\mathcal{B}}(f) = \left(\begin{array}{c|c} N & \mathbf{0} \\ \hline \mathbf{0} & B \end{array}\right)$$

wobei N nilpotent (mit $N^{i_0} = 0$) ist, und B invertierbar ist.

(f) $i_0 \leqslant \dim_{\mathbb{K}} \mathrm{Ker}\, f^{\infty} = \mu_a(f, 0)$.

Beweis-Skizze:

(a) Sei $v \in \mathrm{Ker}\, f^{\infty}$. Insbesondere, $v \in \mathrm{Ker}\, f^{i_0} = \mathrm{Ker}\, f^{i_0+1}$. Es gilt also $f^{i_0}(f(v)) = f^{i_0+1}(v) = 0$. Sei $v \in \mathrm{Bild}\, f^{\infty} = \mathrm{Bild}\, f^{i_0} = \mathrm{Bild}\, f^{i_0+1}$. Also $\exists w \in V$ mit $v = f^{i_0}(w)$, und wir bekommen $f(v) = f(f^{i_0}(w)) \in \mathrm{Bild}\, f^{i_0+1} = \mathrm{Bild}\, f^{\infty}$.

(b) Es gilt offensichtlich $f^{i_0}|_{\mathrm{Ker}\, f^{\infty}} = 0$, also das Minimalpolynom ist ein Teiler von x^{i_0}. Es gilt aber auch $f^{i_0-1}|_{\mathrm{Ker}\, f^{\infty}} \neq 0$, weil $\mathrm{Ker}\, f^{i_0-1} \subsetneq \mathrm{Ker}\, f^{i_0}$, und die Schlussfolgerung folgt.

(c) Die Abbildung $f|_{\mathrm{Bild}\, f^{i_0}} : \mathrm{Bild}\, f^i \longrightarrow \mathrm{Bild}\, f^{i+1}$ ist surjektiv für alle i, weil für alle $w \in \mathrm{Bild}\, f^{i+1}$, existiert $v \in V$, sodass $w = f^{i+1}(v) = f(f^i(v))$. Also, weil $\dim_{\mathbb{K}} V < \infty$, haben wir für $i \geqslant i_0$ den erwünschten Isomorphismus.

(d) Aus dem Dimensionssatz haben wir schon, dass $\dim_{\mathbb{K}} \mathrm{Ker}\, f^{i_0} + \dim_{\mathbb{K}} \mathrm{Bild}\, f^{i_0} = \dim_{\mathbb{K}} V$. Es reicht also $\mathrm{Ker}\, f^{i_0} \cap \mathrm{Bild}\, f^{i_0} = \{0\}$ zu zeigen. Sei $v \in \mathrm{Ker}\, f^{i_0} \cap \mathrm{Bild}\, f^{i_0}$. Also $f^{i_0}(v) = 0$ und $\exists\ w \in V$, sodass $v = f^{i_0}(w)$. Also $f^{2i_0}(w) = f^{i_0}(v) = 0$. Das heißt, dass $w \in \mathrm{Ker}\, f^{2i_0} = \mathrm{Ker}\, f^{i_0}$, also $v = f^{i_0}(w) = 0$.

[11] Hans Fitting, deutscher Mathematiker, 1906-1938.

(e) Es folgt direkt aus (a-d), wobei N dem direkten Summand $\operatorname{Ker} f^\infty$, und B dem direkten Summand $\operatorname{Bild} f^\infty$ entspricht.

(f) Wenn $\operatorname{Ker} f = \{0\}$, dann sind alle f^i injektiv, also $i_0 = 0$. Sonst haben wir

$$0 < \dim_{\mathbb{K}} \operatorname{Ker} f < \cdots < \dim_{\mathbb{K}} \operatorname{Ker} f^{i_0},$$

und die Ungleichung folgt. Aus (e) haben wir

$$\chi_f = \chi_N \cdot \chi_B.$$

Weil B invertierbar ist, kann 0 nicht ein Eigenwert von B sein; also $x \nmid \chi_B$. Anderseits $N \in \operatorname{Mat}_r(\mathbb{K})$, mit $r = \dim_{\mathbb{K}} \operatorname{Ker} f^\infty$, ist nilpotent, also aus Korollar 10.67[a] ist $\chi_N = x^r$. Also $\mu_a(f, 0) = r = \dim_{\mathbb{K}} \operatorname{Ker} f^\infty$.

Q.E.D.

[a] Das ist eigentlich ganz einfach zu sehen, weil 0 der einzige Eigenwert ist

[10]23.5.'22

10.9 Jordan[12] Zerlegung

Wir haben in Satz 10.32 gesehen, dass ein Endomorphismus genau dann diagonalisierbar ist, wenn V die direkte Summe der Eigenräume ist. Wenn das nicht der Fall ist, dann gibt es mindestens einen Eigenwert mit $\mu_g < \mu_a$. Für trigonalisierbare Endomorphismen[13] werden wir "verallgemeinerte Eigenräume" einführen. Diese sind invariante Unterräume in deren direkten Summe sind der Vektorraum V zerlegen lässt.

Der Eigenraum von f zum Eigenwert λ ist $\operatorname{Eig}(f, \lambda) = \operatorname{Ker}(f - \lambda \operatorname{id}_V)$. Der erweitere Eigenraum (oder Hauptraum) zum Eigenwert λ wird diesen $\mathbb{K}$-Unterraum enthalten.

Definition 10.57. Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, $f \in \operatorname{End}_{\mathbb{K}}(V)$ und $\lambda \in \mathbb{K}$ ein Eigenwert von f. Der **Hauptraum** von f zum Eigenwert λ ist der $\mathbb{K}$-Unterraum von V:

$$\operatorname{Hau}(f, \lambda) := \operatorname{Ker}(f - \lambda \operatorname{id}_V)^\infty.$$

Bemerkung 10.58. Wir haben

$$\chi_{\lambda \operatorname{id}_V - f}(x) = \det(x \operatorname{id}_V - (f - \lambda \operatorname{id}_V)) = \det((x + \lambda) \operatorname{id}_V - f) = \chi_f(x + \lambda).$$

Insbesondere, wenn $\lambda \in \mathbb{K}$ ein Eigenwert von f ist, dann ist 0 ein Eigenwert von $f - \lambda \operatorname{id}_V$, und weiterhin

$$\mu_a(f, \lambda) = \mu_a(f - \lambda \operatorname{id}_V, 0)$$

Bemerkung 10.59. Seien $f \in \operatorname{End}_{\mathbb{K}} V$, $U \subseteq_{UVR} V$, und $\lambda \in \mathbb{K}$. Es gilt:

$$U \text{ ist } f\text{−invariant} \iff U \text{ ist } (f - \lambda \operatorname{id}_V)\text{−invariant}.$$

[12] Camille Jordan, französischer Mathematiker, 1838-1922.

[13] Weil jeder Körper $\mathbb{K}$ in einem algebraisch abgeschlossenen Körper $\overline{\mathbb{K}}$ enthalten ist, kann man voraussetzen, dass jeder Endomorphismus trigonalisierbar ist. Die Existenz von $\overline{\mathbb{K}}$ geht aber über die Ziele dieser Vorlesung hinaus, und wir werden es nicht beweisen.

Beweis-Skizze: Es reicht eine Richtung zu zeigen. Die andere folgt wenn man λ durch $-\lambda$ ersetzt. Sei also U ein $(f - \lambda \operatorname{id}_V)$-invarianter $\mathbb{K}$-VR. Dann haben wir für $u \in U$:

$$f(u) = f(u) - \lambda u + \lambda u = (f - \lambda \operatorname{id}_V)(u) + \lambda u \in U.$$

Q.E.D.

Satz 10.60. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \operatorname{End}_{\mathbb{K}}(V)$ ein Endomorphismus, sodass*

$$\chi_f(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m_i},$$

wobei $\lambda_i \neq \lambda_j$ wenn $i \neq j$ und $m_i = \mu_a(f, \lambda_i)$ für alle $i = 1, \ldots, r$. Es gelten

(a) $\operatorname{Hau}(f, \lambda_i)$ *ist f-invariant für alle $i = 1, \ldots, r$.*

(b) $\dim_{\mathbb{K}} \operatorname{Hau}(f, \lambda_i) = \mu_a(f, \lambda_i)$ *für alle $i = 1, \ldots, r$.*

(c) $V = \operatorname{Hau}(f, \lambda_1) \oplus \cdots \oplus \operatorname{Hau}(f, \lambda_r)$.

(d) *Es gibt $f_N, f_D \in \operatorname{End}_{\mathbb{K}}(V)$ mit f_N nilpotent, f_D diagonalisierbar, und $f_N \circ f_D = f_D \circ f_N$, sodass*

$$f = f_D + f_N.$$

Beweis-Skizze: (a) Aus Lemma von Fitting (a) gilt $(f - \lambda_i \operatorname{id}_V)(\operatorname{Hau}(f, \lambda_i)) \subseteq \operatorname{Hau}(f, \lambda_i)$. Dieser Punkt folgt also aus der Bemerkung 10.59.
(b) Folgt direkt aus Lemma von Fitting Punkt (f) für $\lambda_i \cdot \operatorname{id}_V - f$ und Bemerkung 10.58.
(c) Wir zeigen das durch Induktion über r - die Anzahl verschiedener Eigenwerte. Wir bezeichnen für $i = 1, \ldots, r$ mit

$$g_i := f - \lambda_i \operatorname{id}_V .$$

Es gilt also, dass jeder Eigenvektor $v \in V$ von f zum Eigenwert λ_i, ist ein Eigenvektor von g_i zum Eigenwert 0, und

$$\operatorname{Hau}(f, \lambda_i) = \operatorname{Ker} g_i^{\infty}. \tag{10.11}$$

$\boxed{r = 1}$ Es folgt aus Lemma 10.56 (f), dass $\dim_{\mathbb{K}} \operatorname{Ker} g_1^{\infty} = \dim_{\mathbb{K}} V$, also $V = \operatorname{Hau}(f, \lambda_1)$.
$\boxed{r - 1 \Rightarrow r}$ Aus Lemma 10.56 (d) für g_1 haben wir

$$V = \operatorname{Hau}(f, \lambda_1) \oplus \operatorname{Bild} g^{\infty}.$$

Aus Lemma 10.56 (a) ist $W := \operatorname{Bild} g^{\infty}$ ein g_1-invarianter $\mathbb{K}$-Unterraum von V. Also, nach Bemerkung 10.59, ist W auch f-invariant. Wir betrachten nun die Einschränkung $f|_W \in \operatorname{End}_{\mathbb{K}}(W)$. Aus Lemma 10.56 (e) gilt

$$\chi_{f|_W} = \prod_{i=2}^{r} (x - \lambda_i)^{m_i}.$$

Insbesondere hat $f|_W$ genau $r - 1$ verschiedene Eigenwerte: $\lambda_2, \ldots, \lambda_r$. Aus der Induktiven Voraussetzung haben wir also,

$$W = \operatorname{Hau}(f|_W, \lambda_2) \oplus \cdots \oplus \operatorname{Hau}(f|_W, \lambda_r),$$

und (c) folgt.
(d) Wir wissen aus Satz 10.39, dass f trigonalisierbar ist. Es gibt also eine Basis $\mathcal{B}$ von V, sodass $M^{\mathcal{B}}(f)$ eine obere Dreiecksmatrix ist, mit der Diagonale $\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_r, \dots, \lambda_r$, wobei jeder λ_i genau m_i Mal vorkommt. Wir definieren dann $f_D \in \mathrm{End}_{\mathbb{K}}(V)$, sodass

$$M^{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \quad \dots\dots \quad, \lambda_r, \dots, \lambda_r)$$

und $f_N \in \mathrm{End}_{\mathbb{K}}(V)$, sodass

$$M^{\mathcal{B}}(f_N) = M^{\mathcal{B}}(f) - M^{\mathcal{B}}(f).$$

Nach Korollar 10.67 ist f_N nilpotent, und es gilt $f = f_D + f_N$.

$\boxed{f_D \circ f_N = f_N \circ f_D}$ Aus Bemerkung 10.42 folgt, dass die Zerlegung $V = \mathrm{Hau}(f, \lambda_1) \oplus \dots \oplus \mathrm{Hau}(f, \lambda_r)$ in invariante $\mathbb{K}$-Unterraum, zu einer diagonalen Blockform von $M^{\mathcal{B}}(f)$ führt. Die Basis $\mathcal{B}$ ist die Vereinigung der Basen $\mathcal{B}_i$ von $\mathrm{Hau}(f, \lambda_i)$. Also sind die Blöcke auf der Diagonale die obere Dreiecksmatrizen von $g_i|_{\mathrm{Hau}(f,\lambda_i)}$. Es reicht also uns auf die entsprechenden Einschränkungen der g_i zu konzentrieren. Wir bezeichnen mit $\widetilde{g_i} := g_i|_{\mathrm{Hau}(f,\lambda_i)}$. Es gilt also

$$M^{\mathcal{B}_i}(\widetilde{g_i}) = \lambda_i I_{m_i} + N_i,$$

wobei $N_i \in \mathrm{Mat}_{m_i}(\mathbb{K})$ eine nilpotente obere Dreiecksmatrix ist. Weil λI_n mit allen Matrizen kommutiert, gilt

$$(\lambda_i I_{m_i}) \cdot N_i = N_i \cdot (\lambda_i I_{m_i}) = \lambda_i N_i,$$

und somit auch $f_D \circ f_N = f_N \circ f_D$. Q.E.D.

Eine Zerlegung $f = f_D + f_N$ wie im Punkt (d) in Satz 10.60 heißt **Jordan Zerlegung**.

Bemerkung 10.61. Wir haben mehr bewiesen, nämlich, dass es eine Basis $\mathcal{B}$ von V gibt, sodass

$$M_f^{\mathcal{B}} = \begin{pmatrix} \boxed{\lambda_1 I_{m_1} + N_1} & 0 & \dots & 0 \\ 0 & \boxed{\lambda_2 I_{m_2} + N_2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \boxed{\lambda_r I_{m_r} + N_r} \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ die verschiedene Eigenwerte von f, mit $m_i = \mu_a(f, \lambda_i)$ für alle $i = 1, \dots, r$, und mit $N_i \in \mathrm{Mat}_{m_i}(\mathbb{K})$ nilpotente obere Dreiecksmatrizen.

[11]25.5.'22

10.10 Nilpotente Endomorphismen

Definition 10.62. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $n \in \mathbb{N}_{>0}$.

(a) Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist **nilpotent**, wenn es eine positive ganze Zahl $m \in \mathbb{N}_{>0}$ mit $f^m = 0$ gibt.

(b) Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ ist **nilpotent**, wenn es eine positive ganze Zahl $m \in \mathbb{N}_{>0}$ mit $A^m = 0$ gibt.

Beispiel 10.63. 1. Sei $f \in \mathrm{End}_Q(\mathbb{Q}^2)$ mit $f(x,y) = (6x - 3y, 12x - 6y)$, dann gilt

$$f^2(x,y) = f(6x-3y, 12x-6y) = (36x - 18y - 36x + 18y, 72x - 36y - 72x + 36y) = (0,0).$$

Es existiert also $m = 2 \in \mathbb{N}_{>0}$ mit $f^m = 0$.

2. Die Matrix $A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C})$ ist auch nilpotent, weil $A^3 = 0$.

Bemerkung 10.64. 1. Die Determinante einer nilpotenten Matrix A ist Null. Das gilt weil, wenn $A^m = 0$, dann gilt $(\det A)^m = \det A^m = \det 0 = 0$, und weil $\mathbb{K}$ ein Körper ist, muss $\det A = 0$ sein.

2. Wenn A nilpotent ist, und wenn $B \sim A$, dann ist auch B nilpotent. Das gilt weil, wenn $B = T^{-1}AT$ und $A^m = 0$, dann ist

$$B^m = (T^{-1}AT)(T^{-1}AT)\cdots(T^{-1}AT) = T^{-1}A^mT = T^{-1}0T = 0.$$

Satz 10.65 (Struktur der Nilpotenten Endomorphismen)**.** *Seien $n \in \mathbb{N}_{>0}$, V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n$, und $f \in \mathrm{End}_{\mathbb{K}}(V)$ ein nilpotenter Endomorphismus. Es existiert eine geordnete Basis $\mathcal{B} = v_1, \ldots, v_n$ von V, sodass*

$$f(v_i) \in \{0, v_{i-1}\}, \qquad \forall\ i = 1, \ldots, n.$$

Beweis-Skizze: Sei $m \in \mathbb{N}$ minimal mit der Eigenschaft, dass $f^m = 0$. Wir bezeichnen

$$K^i := \mathrm{Ker}\, f^i.$$

Wir haben dann folgende Kette von $\mathbb{K}$-Unterraum von V:

$$\{0\} \subsetneq K^1 \subsetneq K^2 \subsetneq \cdots \subsetneq K^{m-1} \subsetneq K^m = V. \tag{10.12}$$

Die Inklusionen wurden in Abschnitt 10.8 bewiesen. Aus der Minimalität von m und aus (10.9) folgt, dass alle Inklusionen strikt sind.
Sei $v \in K^i$. Dann gilt $0 = f^i(v) = f^{i-1}(f(v))$, also

$$f(K^i) \subseteq K^{i-1}. \tag{10.13}$$

Wir bekommen aus (10.13) eine Kette in der alle Abbildungen $\hookrightarrow$ durch $\widehat{v} \mapsto \widehat{f(v)}$ gegeben sind:

$$K^m/K^{m-1} \hookrightarrow K^{m-1}/K^{m-2} \hookrightarrow \cdots \hookrightarrow K^2/K^1 \hookrightarrow K^1. \tag{10.14}$$

- **Die Abbildungen $\hookrightarrow$ sind wohl definiert:**
  Seien $v, v' \in K^{i+1}$ mit $\widehat{v} = \widehat{v'} \in K^{i+1}/K^i$. Es gilt also $v' - v := w \in K^i$. Wir haben

$$\widehat{v'} = \widehat{v+w} \quad \mapsto \quad \widehat{f(v+w)} = \widehat{f(v)} + \widehat{f(w)} = \widehat{f(v)} \quad \leftarrow\!\shortmid \quad \widehat{v}.$$

  Die dritte Gleichung gilt weil wir aus (10.13) $f(w) \in K^{i-1}$ haben und somit $\widehat{f(w)} = 0$.

- **Die Abbildungen $\hookrightarrow$ sind injektiv:**
  Sei $v \in K^{i+1}$ mit $\widehat{f(v)} = 0 \in K^i/K^{i-1}$, also $f(v) \in K^{i-1}$. Das heißt, dass $f^i(v) = f^{i-1}(f(v)) = 0$, also $v \in K^i$, und somit $\widehat{v} = 0 \in \mathbb{K}^{i+1}/K^i$. Wir haben also gezeigt, dass

$$\operatorname{Ker}\left(K^{i+1}/K^i \hookrightarrow K^i/K^{i-1}\right) = 0.$$

Wir wählen für $i = m, m-1, \ldots, 1$ (in dieser Reihenfolge) Vektoren $B^i_{neu} = u_{i,1}, \ldots, u_{i,p_i} \in K^i$ sodass[a]

$$\widehat{B^i} := \widehat{f(B^{i+1})} \cup \widehat{B^i_{neu}} \text{ eine Basis von } K^i/K^{i-1} \text{ ist.}$$

Die Wahl solcher B^i ist möglich, weil alle Abbildungen in der Kette (10.14) injektiv sind. Das heißt, dass wenn $\widehat{u_1}, \ldots, \widehat{u_j} \in K^{i+1}/K^i$ linear unabhängig sind, dann sind auch $\widehat{f(u_1)}, \ldots, \widehat{f(u_j)} \in K^i/K^{i-1}$ linear unabhängig. Wir bekommen also:

$$\begin{array}{rccccccccccc}
B^m: & u_{m,1} & \ldots & u_{m,p_m} & & & & & & & & \\
B^{m-1}: & f(u_{m,1}) & \ldots & f(u_{m,p_m}) & & u_{m-1,1} & \ldots & u_{m-1,p_{m-1}} & & & & \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & & & & \\
B^1: & f^{m-1}(u_{m,1}) & \ldots & f^{m-1}(u_{m,p_m}) & f^{m-2}(u_{m-1,1}) & \ldots & f^{m-2}(u_{m-1,p_{m-1}}) & \ldots & u_{1,1} & \ldots & u_{1,p_1}
\end{array}$$

Wir zeigen jetzt, dass $B = B^m \cup B^{m-1} \cup \cdots \cup B^1$ eine Basis von V ist.

- **Lineare Unabhängigkeit:**
  Seien ${}^\bullet\lambda_{\bullet,\bullet} \in \mathbb{K}$, sodass eine lineare Kombination der Vektoren in B gleich mit Null ist. Das ist nicht einfach in dieser allgemeiner Form aufzuschreiben. Wir brauchen das aber nicht, weil wenn wir f^{m-1} auf der ganzen linearen Kombination anwenden, dann werden alle Vektoren aus $B^{m-1}, \ldots, B^1$ auf Null abgebildet. Es überlebt also nur

$${}^0\lambda_{m,1}u_{m,1} + \cdots + {}^0\lambda_{m,p_m}u_{m,p_m} = 0$$

  und daraus folgt, dass alle λ die vorkommen Null sein müssen, weil B^m aus linear unabhängigen Vektoren besteht. Wir wenden dann auf der ursprünglichen linearen Kombination f^{m-2} an, und es überleben dieses Mal nur die Vektoren aus B^{m-1}. Das ganze geht völlig analog weiter, bis wir gezeigt haben, dass alle λ gleich Null sind.

- **Basis:**
  Wir haben eine linear unabhängige Menge B. Um zu zeigen, dass es eine Basis ist, reicht es zu zeigen, dass es $\dim_{\mathbb{K}} V$ Vektoren enthält. Wir haben

$$\begin{aligned}
\sharp B &= \sharp B^m + \cdots + \sharp B^1 \\
&= \dim K^m/K^{m-1} + \dim K^{m-1}/K^{m-2} + \cdots + \dim K^2/K^1 + \dim K^1 \\
&= \dim K^m - \dim K^{m-1} + \dim K^{m-1} - \dim K^{m-2} + \cdots + \dim K^2 - \dim K^1 + \dim K^1 \\
&= \dim K^m \\
&= \dim V
\end{aligned}$$

Wobei die letzte Gleichung gilt, weil $K^m = V$.

Wir haben also eine Basis. Wir müssen diese nur noch richtig ordnen, damit wir $f(v_i) \in \{0, v_{i-1}\}$ bekommen. Die Ordnung ist in der obigen Tafel: der Spalten nach von links nach rechts, und in jeder Spalte von oben nach unten. Also

$$v_1 = u_{m,1}, \quad v_2 = f^1(u_{m,1}), \quad \ldots, \quad v_m = f^{m-1}(u_{m,1}), \quad v_{m+1} = f^0(u_{m,2}) = u_{m,2}, \quad \ldots\ldots$$

Es ist jetzt einfach zu sehen, dass $f(v_i) = v_{i-1}$ wenn $v_i = f^j(u_{k,l})$ mit $j < k-1$, oder $f(v_i) = 0$ wenn $v_i = f^{k-1}(u_{k,l})$, weil $u_{k,l} \in \operatorname{Ker} f^k$.

Q.E.D.

[a] Hier bezeichnet $\widehat{B} := \{\widehat{u} \ : \ u \in B\}$.

Analog erhalten wir den Satz für nilpotente Matrizen:

Satz 10.66. *Seien $n \in \mathbb{N}_{>0}$ und $A \in \operatorname{Mat}_n(\mathbb{K})$ eine nilpotente Matrix. Dann ist $A \sim B = (b_{ij})$, wobei*

$$b_{ij} = 0 \quad \textit{wenn } j \neq i+1 \qquad \textit{und} \qquad b_{i,i+1} \in \{0,1\} \quad \forall\, i = 1, \ldots, n-1.$$

Korollar 10.67. Seien $n \in \mathbb{N}_{>0}$, ein $\mathbb{K}$-Vektorraum V mit $\dim_{\mathbb{K}} V = n$, ein Endomorphismus $f \in \operatorname{End}_{\mathbb{K}}(V)$ und eine Matrix $A \in \operatorname{Mat}_n(\mathbb{K})$.

(a) Folgende Aussagen sind äquivalent

(i) f ist nilpotent.
(ii) Es gibt $1 \leqslant m \leqslant n$ mit $f^m = 0$.
(iii) $\chi_f(x) = x^n$.
(iv) f ist trigonalisierbar mit $\lambda = 0$ als einziger Eigenwert.
(v) Es gibt eine geordnete Basis $\mathcal{B}$ von V, sodass

$$M^{\mathcal{B}}(f) = \begin{pmatrix} 0 & * & 0 & \ldots & 0 \\ 0 & 0 & * & \ldots & 0 \\ \vdots & & & \ddots & * \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}, \quad \text{mit} \quad * \in \{0,1\}.$$

(b) Folgende Aussagen sind äquivalent

(i) A ist nilpotent.
(ii) Es gibt $1 \leqslant m \leqslant n$ mit $A^m = 0$.
(iii) $\chi_A(x) = x^n$.
(iv) A ist trigonalisierbar mit $\lambda = 0$ als einziger Eigenwert.
(v) A ist ähnlich zu einer Matrix $B \in \operatorname{Mat}_n(\mathbb{K})$ der Form:

$$B = \begin{pmatrix} 0 & * & 0 & \ldots & 0 \\ 0 & 0 & * & \ldots & 0 \\ \vdots & & & \ddots & * \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}, \quad \text{mit} \quad * \in \{0,1\}.$$

[12]30.5.'22

10.11 Jordan Normalform

Wir haben jetzt alle Zutaten für die Jordan Normalform. Wir müssen diesen nur noch Namen geben.

Definition 10.68. Seien $d \in \mathbb{N}_{>0}$ und $\lambda \in \mathbb{K}$. Der **nilpotente Jordan-Block** ist die Matrix $J_d = (\gamma_{ij}^d) \in \mathrm{Mat}_d(\mathbb{K})$ mit $\gamma_{ij}^d = \delta_{i+1,j}$. Der **Jordan-Block** zum Skalar λ ist die Matrix $J_d(\lambda) := \lambda I_d + J_d \in \mathrm{Mat}_d(\mathbb{K})$. Also

$$J_d = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{pmatrix} \qquad J_d(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & \lambda \end{pmatrix}.$$

Wir sagen, dass eine Matrix $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ eine **Block-diagonale Form** hat, falls es $r \in \mathbb{N}_{>0}$ und, für jeden $k = 1, \dots, r$, eine quadratische Matrix $D_k = (d_{ij}^k) \in \mathrm{Mat}_{l_k}(\mathbb{K})$ gibt, sodass

$$a_{ij} = \begin{cases} d_{ij}^k & \text{falls } i,j \in \{1 + \sum_{s=1}^{k-1} l_s, \dots, \sum_{s=1}^{k} l_s\} \\ 0 & \text{sonst.} \end{cases}$$

In diesem Fall stellen wir A wie folgt dar:

$$A = \begin{pmatrix} \boxed{D_1} & & & \mathrm{O} \\ & \boxed{D_2} & & \\ & & \ddots & \\ \mathrm{O} & & & \boxed{D_r} \end{pmatrix}.$$

Definition 10.69. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ **ist** in **Jordan-Normalform** (JNF), falls es eine Block-diagonale Form hat, mit Jordan-Blöcke auf der Diagonale:

$$A = \begin{pmatrix} \boxed{J_{d_1}(\lambda_1)} & & & \mathrm{O} \\ & \boxed{J_{d_2}(\lambda_2)} & & \\ & & \ddots & \\ \mathrm{O} & & & \boxed{J_{d_r}(\lambda_r)} \end{pmatrix},$$

wobei $\lambda_i = \lambda_j$ und $d_i = d_j$ für $i \neq j$ erlaubt ist.

Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ **hat** eine **Jordan-Normalform**, falls es einer Matrix in JNF ähnlich ist.

Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ eines endlichdimensionalen $\mathbb{K}$-Vektorraum V hat eine Jordan-Normalform, falls es eine Basis $\mathcal{B}$ von V gibt, sodass $M^{\mathcal{B}}(f)$ in JNF ist.

Lemma 10.70. *Zwei Jordan-Blöcke $J_d(\lambda)$, $J_e(\mu)$ sind ähnlich genau dann, wenn $d = e$ und $\lambda = \mu$.*

Beweis-Skizze: Ähnliche Matrizen haben denselben Typ und gleiche charakteristische Polynome, also aus $J_d(\lambda) \sim J_e(\mu)$ folgt $(x-\lambda)^d = (x-\mu)^e$ und somit $\lambda = \mu$ und $d = e$. Q.E.D.

Lemma 10.71. *Zwei Matrizen $A, B \in \mathrm{Mat}_n(\mathbb{K})$ in JNF sind ähnlich genau dann, wenn sie die gleiche Anzahl von jedem Jordan-Block $J_d(\lambda)$ haben.*

Beweis-Skizze: $\boxed{\Leftarrow}$ ist klar: man braucht nur die Jordan-Blöcke permutieren, und dass macht man durch konjugieren mit einer Permutationsmatrix.

$\boxed{\Rightarrow}$ Seien $A \sim B$, beide in JNF. Das heißt beide sind Darstellungen desselben Endomorphismus $f \in \mathrm{End}(\mathbb{K}^n)$ bezüglich verschiedener Basen. Weil A und B in JNF sind, können wir beide in Block-diagonal Form schreiben, wobei jeder Block einem EW λ entspricht. Jeder solcher Block, sowohl von A als auch von B, ist vom Typ $\mu_a(f,\lambda) \times \mu_a(f,\lambda)$. Es entspricht der Einschränkung $f|_{\mathrm{Hau}(f,\lambda)} : \mathrm{Hau}(f,\lambda) \longrightarrow \mathrm{Hau}(f,\lambda)$. Es reicht also diese Blöcke separat zu betrachten, das heißt, wir können annehmen, dass A und B einen einzigen EW haben: λ.

Wenn $A = T^{-1}BT$, dann gilt auch $A - \lambda I_n = T^{-1}(B - \lambda I_n)T$. Wir können also A und B durch $A - \lambda I_n$, beziehungsweise $B - \lambda I_n$ ersetzen, und somit annehmen, dass $\lambda = 0$. Das heißt, dass A und B ähnliche nilpotente Matrizen in JNF sind.

Jetzt werden wir den Beweis von Satz 10.65 anwenden.

Behauptung: Die Anzahl von Blöcke J_d in A und B ist nur von f abhängig.

Um das zu zeigen, beweisen wir

$$\sharp\{\text{Jordanblöcke } J_i \text{ in A}\} = \dim K^{i-1}/K^{i-2} - \dim K^i/K^{i-1}, \tag{10.15}$$

und analog für B. Um das zu *sehen*, wiederholen wir hier das Tableau:

$\frac{K^m}{K^{m-1}}$	$u_{m,1}$	$\dots$	u_{m,p_m}							
$\frac{K^{m-1}}{K^{m-2}}$	$f(u_{m,1})$	$\dots$	$f(u_{m,p_m})$	$u_{m-1,1}$	$\dots$	$u_{m-1,p_{m-1}}$				
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$				
$\frac{K^1}{0}$	$f^{m-1}(u_{m,1})$	$\dots$	$f^{m-1}(u_{m,p_m})$	$f^{m-2}(u_{m-1,1})$	$\dots$	$f^{m-2}(u_{m-1,p_{m-1}})$	$\dots$	$u_{1,1}$	$\dots$	u_{1,p_1}
	J_m	$\dots$	J_m	J_{m-1}	$\dots$	J_{m-1}	$\dots$	J_1	$\dots$	J_1

Die entsprechenden Äquivalenzklassen in jeder Zeile sind eine Basis des Quotientenraumes. Die Vektoren in einer Spalte entsprechen einem Jordan-Block der genau so groß ist wie die Spalte hoch ist.

Wir wollen aber umgekehrt, die Dimensionen der Quotienten aus den Jordan-Blöcke lesen. Wir konzentrieren uns erstmals auf einem Jordan-Block J_d. Seien $v_1, \dots, v_d$ Basis-Vektoren die einem Jordan-Block J_d entsprechen[a]:

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

Das heißt, dass

$$\begin{array}{rcccccccl} v_1 & \neq & 0 & \text{und} & f(v_1) & = & 0 & \Longrightarrow & 0 \neq v_1 \in K^1 \\ f(v_2) = v_1 & \neq & 0 & \text{und} & f^2(v_2) & = & 0 & \Longrightarrow & 0 \neq \widehat{v_2} \in K^2/K^1 \\ & \vdots & & & & \vdots & & & \\ f^{d-1}(v_d) = v_1 & \neq & 0 & \text{und} & f^d(v_d) & = & 0 & \Longrightarrow & 0 \neq \widehat{v_d} \in K^d/K^{d-1} \end{array}$$

Also für jeden Jordan-Block J_{a_i} mit $a_i \geqslant d$ bekommen wir einen $\widehat{v_{d,i}} \neq 0 \in K^d/K^{d-1}$.

Wir zeigen jetzt, dass $\widehat{v_{d,1}}, \dots, \widehat{v_{d,r}}$ linear unabhängig in K^d/K^{d-1} sind. Seien $\lambda_i \in \mathbb{K}$ mit

$$\sum \lambda_i \cdot \widehat{v_{d,i}} = 0 \in K^d/K^{d-1}, \quad \text{also} \quad \sum \lambda_i \cdot v_{d,i} \in K^{d-1}.$$

Wir wenden f^{d-1} auf der zweiten linearen Kombination an und, aus der Linearität von f, bekommen

$$\sum \lambda_i \cdot v_{1,i} = 0$$

wobei die $v_{1,i}$ denselben Jordan-Blöcke J_{a_i} wie $v_{d,i}$ entsprechen. Somit sind die $v_{1,i}$, linear unabhängig[b]. Es folgt also, dass $\lambda_i = 0$ für alle i. Wir haben also gezeigt

$$\dim K^d/K^{d-1} \geqslant \sharp\{\text{Jordan Blöcke } J_a \text{ mit } a \geqslant d\}$$

Wenn wir auf beiden Seiten der obigen Ungleichung über $d = 1, \dots, n$ addieren, bekommen wir auf beiden Seiten n. Es muss also die Gleichheit gelten:

$$\sharp\{\text{Jordan Blöcke } J_a \text{ mit } a \geqslant d\} = \dim K^d - \dim K^{d-1}. \tag{10.16}$$

und somit auch

$$\begin{array}{rcl} \dim K^d/K^{d-1} - \dim K^{d+1}/K^d & = & \sharp\{\text{Jordan Blöcke } J_i \text{ mit } i \geqslant d\} - \\ & & -\sharp\{\text{Jordan Blöcke } J_i \text{ mit } i \geqslant d+1\} \\ & = & \sharp\{\text{Jordan Blöcke } J_d\}. \end{array}$$

Wir haben also (10.15) gezeigt, uns somit die Behauptung, und somit den Satz.

Q.E.D.

[a] das heißt, in den entsprechen Spalten werden die Koordinaten von $f(v_i)$ eingetragen

[b] weil sie eine Teilmenge einer Basis sind.

Bemerkung 10.72. Nach dem Beweis von Lemma 10.71 haben wir zu jedem Eigenwert λ, dass

$$\begin{aligned}\sharp\{\text{Jordan Blöcke } J_d(\lambda)\} &= \dim K^d/K^{d-1} - \dim K^{d+1}/K^d \\ &= \dim K^d - \dim K^{d-1} - \dim K^{d+1} + \dim K^d \\ &= -\dim K^{d+1} + 2\cdot\dim K^d - \dim K^{d-1},\end{aligned}$$

wobei wir in diesem Fall die Bezeichnung verwenden:

$$K^i := \operatorname{Ker}(f - \lambda\,\mathrm{id}_V)^i.$$

Wir formulieren jetzt den Hauptsatz dieses Kapitels. Wir haben es eigentlich schon bewiesen, wir fassen es hier nur zusammen.

Satz 10.73. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \operatorname{End}_{\mathbb{K}}(V)$. Wenn das charakteristische Polynom von f vollständig in Linearfaktoren zerfällt, dann hat f eine Jordan Normalform, die bis auf permutieren der Blöcke eindeutig ist.*

Beweis-Skizze: Aus Satz 10.60 und 10.61 haben wir die Block-Diagonalform. Aus Satz 10.66 für $f|_{\mathrm{Hau}(f,\lambda_i)} - \lambda_i\,\mathrm{id}_{\mathrm{Hau}(f,\lambda_i)}$ haben wir die JNF für jeden Block. Aus Lemma 10.71 haben wir die Eindeutigkeit bis auf Permutation der Blöcke. Q.E.D.

Eine äquivalente Aussage gilt auch für Matrizen.

Satz 10.74. *Seien $n \in \mathbb{N}_{>0}$ und $A \in \mathrm{Mat}_n(\mathbb{K})$. Wenn das charakteristische Polynom von A vollständig in Linearfaktoren zerfällt, dann hat A eine Jordan Normalform, die bis auf permutieren der Blöcke eindeutig ist.*

Korollar 10.75. Sei $f \in \operatorname{End}_{\mathbb{K}}(V)$ mit $\chi_f(x) = \prod_{i=1}^r (x-\lambda_i)^{m_i}$ und $\lambda_i \neq \lambda_j$ wenn $i \neq j$. Für jeden $i = 1, \ldots, r$ seien

$$\begin{aligned} d_i &:= \text{die Anzahl aller Jordan Blöcke } J_k(\lambda_i) \text{ in der JNF von } A \\ j_{\max}(i) &:= \max\{k \in \mathbb{N} \mid J_k(\lambda_i) \text{ ist ein JB von } A\}.\end{aligned}$$

Dann gilt

(a) $\mu_g(f,\lambda_i) := \dim_{\mathbb{K}} \operatorname{Eig}(f,\lambda_i) = d_i$.

(b) $\mathrm{mPol}_f = \prod_{i=1}^r (x-\lambda_i)^{j_{\max}(i)}$.

Die Analoge Aussage gilt auch für eine Trigonalisierbare Matrix.

Beweis-Skizze:

(a) Wir haben $\operatorname{Eig}(f,\lambda_i) \subseteq \operatorname{Hau}(f,\lambda_i)$. Wenn wir f auf $\operatorname{Hau}(f,\lambda_i)$ einschränken, dann entspricht $\operatorname{Eig}(f,\lambda_i)$ dem Kernel K^1. Aus (10.16) in dem Beweis von Lemma 10.71, haben wir also

$$d_i = \dim K^1 - \dim K^0 = \dim K^1 = \dim \operatorname{Ker}(f - \lambda_i\,\mathrm{id}_V).$$

(b) Für diesen Teil müssen wir erstmals bemerken, dass wenn wir zwei Matrizen in diagonaler Block-Form, mit Blöcke desselben Typs auf der Diagonale, multiplizieren, dann reicht es die

Blöcke auf der Diagonale zu multiplizieren:

$$\begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & D_r \end{pmatrix} \cdot \begin{pmatrix} C_1 & 0 & \dots & 0 \\ 0 & C_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & C_r \end{pmatrix} = \begin{pmatrix} D_1C_1 & 0 & \dots & 0 \\ 0 & D_2C_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & D_rC_r \end{pmatrix}, \tag{10.17}$$

wobei $D_i, C_i \in \mathrm{Mat}_{r_i}(\mathbb{K})$. Die zweite Bemerkung ist, dass wenn N eine nilpotente Matrix in JNF ist, dann

$$m = \min\{i \ : \ N^i = 0\} = \max\{d \ : \ J_d \text{ ist ein Jordan Block von } N\}.$$

Sei A eine zugeordnete Matrix von f in JNF, mit

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & A_r \end{pmatrix},$$

wobei A_i dem Hauptraum $\mathrm{Hau}(f, \lambda_i)$ entspricht. Für jeden EW λ_i bezeichnen wir $D_{i,k} := A_k - \lambda_i I_{m_k}$. Weil $\lambda_k \neq \lambda_i$ für $i \neq k$, haben wir, dass

$$D_{i,k} \text{ ist invertierbar} \qquad \text{wenn } i \neq k, \tag{10.18}$$

$$\min\{a \ : \ D_{i,i}^a = 0\} = j_{\max}(i) \qquad \text{wenn } i = k. \tag{10.19}$$

Es gilt also, dass alle $D_{i,k}^j$ invertierbar für alle j sind und

$$D_i^j := (A - \lambda_i I_n)^j = \begin{pmatrix} D_{i,1}^j & 0 & \dots & 0 & \dots & 0 \\ 0 & D_{i,2}^j & \dots & 0 & \dots & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & \dots & D_{i,i}^j & \dots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & D_{i,r}^j \end{pmatrix}, \tag{10.20}$$

und, wenn $j \geqslant j_{\max}(i)$, dann ist genau der i.Block auf der Diagonale Null. Aus (10.17) und (10.20) bekommen wir dann, dass

$$\mathrm{mPol}_f \mid \prod_{i=1}^{r} (x - \lambda_i)^{j_{\max}(i)} =: P.$$

Nehmen wir an, dass $\mathrm{mPol}_f \neq P$. Dann teilt mPol_f auch ein Produkt von $(x - \lambda_i)$ mit einem der Exponenten kleiner als $j_{\max}(i)$. Ohne die Allgemeinheit zu beschränken, nehmen wir an,

dass dieser der Exponent von $(x-\lambda_1)$ ist. Also, dass

$$\mathrm{mPol}_f \mid (x-\lambda_1)^{j_{\max}(1)-1} \cdot \prod_{i=2}^{r} (x-\lambda_i)^{j_{\max}(i)} := Q.$$

Das heißt, dass $Q(A) = 0$. Das ist ein Produkt von Block-diagonale Matrizen, und der erste Block auf der Diagonale ist

$$D_{1,1}^{j_{\max}(1)-1} \cdot \prod_{i=2}^{r} D_{1,i}^{j_{\max}(i)}$$

Aus (10.19) ist der erste Faktor nicht Null, und aus (10.18) sind die anderen Faktoren invertierbar. Also das gesamte Produkt ist nicht Null – ein Widerspruch zu $Q(A) = 0$. Also $\mathrm{mPol}_f = P$.

Q.E.D.

10.12 Beispiele und Algorithmen

Beispiele:

1. Wir schauen uns erst eine Nilpotente Matrix an:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}), \qquad \text{also} \qquad A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \text{und} \qquad A^3 = 0.$$

Wir wollen die JNF von A berechnen, und dafür gehen wir wie im Beweis von Satz 10.66 vor. Wir bezeichnen mit $K^1 := \mathrm{Ker}(A) = \mathrm{Ker}(f_A)$, $K^2 := \mathrm{Ker}(A^2)$, usw.
Der Rang von A ist 2 (weil es zwei linear unabhängige Spalten hat), also $\dim_{\mathbb{K}} K^1 = 1$. Wir haben

$$\begin{aligned} K^1 &= \mathrm{Span}_{\mathbb{K}}\{e_1\}, \\ K^2 &= \mathrm{Span}_{\mathbb{K}}\{e_1, e_2\}, \\ K^3 &= \mathrm{Span}_{\mathbb{K}}\{e_1, e_2, e_3\}. \end{aligned}$$

Wir bezeichnen jetzt mit $k_i := \dim K^i$, und haben also

$$k_1 = 1, \qquad k_2 = 2, \qquad k_3 = 3.$$

Wir suchen jetzt die (Repräsentanten der Elementen der) Basen B^3, B^2, B^1 von K^3/K^2, K^2/K^1, beziehungsweise K^1. Aus der Beschreibung der K^i hier oben finden wir:

$$\begin{aligned} B^3 &: \quad e_3 \\ B^2 &: \quad A \cdot e_3 \quad = e_2 + e_1 \\ B^1 &: \quad A^2 \cdot e_3 \quad = e_1 \end{aligned}$$

Wir haben also $\mathcal{B} = e_1,\ e_1 + e_2,\ e_3$. Somit ist die JNF von A

$$\mathrm{JNF}(A) := M^{\mathcal{B}}(f_A) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Damit wir die Matrix in $\mathrm{GL}_3(\mathbb{R})$ mit $T^{-1}AT = \mathrm{JNF}(A)$ finden, setzen wir die Vektoren von $\mathcal{B}$ als Spalten:

$$T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad T^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Wir schauen uns ein größeres Beispiel an:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{also} \quad A^2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad A^3 = 0.$$

Für jede Matrix M bezeichnen wir mit $\operatorname{Bild} M$ den Spaltenraum der Matrix M. Wir haben dann

$$K^1 = \operatorname{Ker} A = \operatorname{Bild} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{und} \quad K^2 = \operatorname{Ker} A^2 = \operatorname{Bild} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Damit bekommen wir $k_1 = 3$, $k_2 = 5$, $k_3 = 6$. Also die Dimensionen von K^3/K^2, K^2/K^1, K^1 sind 1, 2, bzw. 3. Wir haben per Definition von $k_i = \dim \operatorname{Ker} A^i$, dass

$$k_i = \begin{cases} 6 & \text{wenn } i \geqslant 4 \\ 0 & \text{wenn } i = 0. \end{cases}$$

Also nach Bemerkung 10.72 haben wir

$$\begin{aligned} \sharp\{3 \times 3 \text{ Jordan Blöcke}\} &= -k_4 + 2k_3 - k_2 = -6 + 12 - 5 = 1, \\ \sharp\{2 \times 2 \text{ Jordan Blöcke}\} &= -k_3 + 2k_2 - k_1 = -6 + 10 - 3 = 1, \\ \sharp\{1 \times 1 \text{ Jordan Blöcke}\} &= -k_2 + 2k_1 - k_0 = -5 + 6 - 0 = 1. \end{aligned}$$

Für alle anderen Typen bekommen wir 0. Wir müssen das aber nicht überprüfen weil wir mit einem J_3, J_2 und J_1 haben wir schon eine 6×6 Matrix.

$$\mathrm{JNF}(A) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Um die Basis zu finden die die JNF gibt (also die invertierbare Matrix $T \in \mathrm{GL}_6(\mathbb{K})$ mit $T^{-1}AT = JNF(A)$) müssen wir B^3, B^2, B^1 wie im Beweis von Satz 10.65 finden. Wir haben $\dim K^3/K^2 = 1$, und brauchen also nur einen nicht-Null Vektor in $K^3 \setminus K^2$:

$$B^3 = u_{3,1} = e_6.$$

Wir haben dann

$$A \cdot u_{3,1} = A \cdot e_6 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{und} \qquad A^2 \cdot u_{3,1} = A^2 \cdot e_6 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Damit wir B^2 finden brauchen wir zwei linear unabhängige Vektoren in $K^2 \setminus K^1$. Einer davon muss $f(u_{3,1})$ sein. Für den zweiten suchen erstmals eine Basis von $K^2 = \operatorname{Ker} A^2$ zu bestimmen, die eine Basis von $\operatorname{Ker} A$ ergänzt. Zum Beispiel $b_1, \ldots, b_5$ sind die Spalten folgender Matrix:

$$\operatorname{Ker} A^2 = \operatorname{Bild} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

Wir suchen also einen Vektor in $\operatorname{Ker} A^2$, der nicht in $\operatorname{Ker} A$ liegt, und der linear unabhängig von $f(u_{3,1})$ modulo $\operatorname{Ker} A$ ist. In diesem Fall ist e_2 **nicht** eine gute Wahl: $e_2 = f(u_{3,1}) - b_2$. Aber $u_{2,1} = b_5 = (0, 0, -1, 0, 1, 0)^\top$ ist. Wir berechnen jetzt

$$f(u_{2,1}) = A \cdot \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Wir brauchen jetzt also nur noch $\{f^2(u_{3,1}), f(u_{2,1})\}$ zu einer Basis von $\operatorname{Ker} A$ zu ergänzen. Das ist aber einfach: $u_{1,1} = b_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}^T$. Wir haben jetzt die geordnete Basis von $\mathbb{K}^6$ die wir suchen: $f^2(u_{3,1}), f(u_{3,1}), u_{3,1}, f(u_{2,1}), u_{2,1}, u_{1,1}$ also

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix} \qquad \text{und} \qquad T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Probe:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix} = \left(\begin{array}{ccc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

10.12.1 Algorithmus zur Bestimmung der Jordanschen Normalform und des Minimalpolynom

Aus den Beweisen die zum Satz 10.73 führen, kann man ein Verfahren zur Bestimmung der JNF zusammenstellen. Damit kann man auch entscheiden ob zwei gegebene Matrizen ähnlich sind. Wenn die Matrix nicht trigonalisierbar ist, kann man den Körper $\mathbb{K}$ zu einem algebraisch abgeschlossenem Körper $\overline{\mathbb{K}} \supset \mathbb{K}$ erweitern. In dieser Vorlesung werden wir uns auf dem Fall $\overline{\mathbb{R}} = \mathbb{C}$ einschränken.

Algorithmus A

`Eingabe:` $A \in \mathrm{Mat}_n(\mathbb{K})$

`Schritt 1:` Bestimme $\chi_A(x) \in \mathbb{K}[x]$

`Schritt 2:` Zerlege $\chi_A(x)$ in Linearfaktoren (eventuell in $\overline{\mathbb{K}}[x]$):

$$\chi_A(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m_i}, \qquad \text{mit } \lambda_i \neq \lambda_j \text{ wenn } i \neq j.$$

`Schritt 3:` Für jeden Index $i = 1, \ldots, r$ und jeden $j = 1, \ldots, m_i$ berechne

$$k_{ij} := \dim_{\mathbb{K}} \mathrm{Ker}(A - \lambda_i \cdot I_n)^j.$$

`Schritt 4:` Für jeden Index $i = 1, \ldots, r$ und jeden $j = 1, \ldots, m_i$ berechne

$$d_{ij} := -k_{i,j+1} + 2k_{ij} - k_{i,j-1},$$

wobei $k_{i,0} = 0$ und $k_{i,m_i+1} = k_{i,m_i}$.

`Schritt 5:` Für jeden $i = 1, \ldots, r$ berechne

$$j_{\max}(i) = \max\{j \mid d_{ij} \neq 0\}$$

`Ausgabe:`

- Die Anzahl von Jordan Blöcke der Form $J_j(\lambda_i)$ in $\mathrm{JNF}(A)$ ist d_{ij}.
- Das Minimalpolynom von A ist

$$\mathrm{mPol}_A(x) = \prod_{i=1}^{r} (x - \lambda_i)^{j_{\max}(i)}$$

- $\dim_{\mathbb{K}} \mathrm{Eig}(A, \lambda_i) = k_{i1}$ = Anzahl der Jordan Blöcke zum Eigenwert λ_i.

Bemerkung 10.76. Für das obige Verfahren muss man keine Basis ausrechnen. Es reicht die entsprechenden Matrizen mit dem Gaußschen Algorithmus in Zeilenstufenform zu bringen und somit die Dimensionen der Kerne zu bestimmen. Das ist aber nicht die einzige Methode. Zum Beispiel, wenn $k_{i,1} = 1$, dann kann man zeigen, dass $k_{ij} = 1$, für alle $j = 1, \ldots, m_i$ (Übung). Es gibt also in diesem Fall einen einzigen Jordan Block.

Wenn χ_A in Linearfaktoren zerfällt, kann man algorithmisch auch eine Matrix $T \in \mathrm{GL}_n(\mathbb{K})$ finden, sodass $T^{-1}AT$ in JNF ist. Dafür muss man das Verfahren aus dem Satz 10.66 über nilpotente Matrizen für jeden Hauptraum anwenden. Wir nehmen an, dass in `Algorithmus A, Schritt 3` oben die Haupträume $\mathrm{Hau}(A, \lambda_i)$ bestimmt worden sind. Das Ziel ist für jeden i eine Basis wie im Satz 10.66 zu finden.

[13]1.6.'22

`Algorithmus B`

`Eingabe :` $M = A - \lambda I_n \in \mathrm{Mat}_n(\mathbb{K})$, mit $\lambda \in \mathbb{K}$ ein Eigenwert von A.

`Schritt 0:` Setze $k_0 = 0$ und für $i = 1, \ldots, m+1$

$$\begin{aligned} k_i &:= \dim_{\mathbb{K}} \operatorname{Ker} M^i, \\ d_i &:= -k_{i+1} + 2k_i - k_{i-1}. \end{aligned}$$

`Schritt 1:`

1. Finde linear unabhängige Vektoren $u_{m,1}, \ldots, u_{m,d_m} \in \operatorname{Ker} M^m \setminus \operatorname{Ker} M^{m-1}$.
2. Berechne
$$C_m := \bigcup_{k=1}^{m-1} \{M^k \cdot u_{m,1}, \ldots, M^k \cdot u_{m,d_m}\}.$$

$\vdots$

`Schritt` i`:` Wir nehmen an $u_{k,1}, \ldots, u_{k,d_k}$ und C_k wurden für $k = m, \ldots, i+1$ berechnet.

1. Finde linear unabhängige Vektoren
$$u_{i,1}, \ldots, u_{i,d_i} \in \operatorname{Ker} B^i \setminus \big(\operatorname{Span}_{\mathbb{K}} C_m + \cdots + \operatorname{Span}_{\mathbb{K}} C_{i+1} + \operatorname{Ker} B^{i-1} \big).$$
2. Berechne
$$C_i := \bigcup_{k=1}^{i-1} \{M^k \cdot u_{i,1}, \ldots, M^k \cdot u_{i,d_i}\}.$$

$\vdots$

`Schritt` m`:` Wir nehmen an $u_{k,1}, \ldots, u_{k,d_k}$ und C_k wurden für $k = m, \ldots, 2$ berechnet.

1. Ergänze $C_m \cup \cdots \cup C_2$ mit $\{u_{1,1}, \ldots, u_{1,d_1}\}$ zu einer Basis von $\operatorname{Ker} M$.

`Ausgabe :` $\mu_a(A, \lambda)$-viele Spalten der Matrix $T \in \mathrm{GL}_n(\mathbb{K})$ die $T^{-1}AT = \mathrm{JNF}(A)$ erfüllt:

$$M^{m-1} \cdot u_{m,1} \;\Big|\; \ldots \;\Big|\; M^0 \cdot u_{m,1} \;\Big|\; \ldots\ldots \;\Big|\; M^{i-1} \cdot u_{i,j} \;\Big|\; \ldots \;\Big|\; M^0 \cdot u_{i,j} \;\Big|\; \ldots\ldots \;\Big|\; u_{1,1} \;\Big|\; \ldots \;\Big|\; u_{1,d_1}$$

`Bemerkung:` Um die Matrix T zu finden, muss man also `Algorithmus B` für alle Eigenwerte λ von A anwenden.

Beispiele: 1. Als erstes, wenden wir `Algorithmus A` und `Algorithmus B` für folgende Matrix an.

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

`Algorithmus A`

`Schritt 1:` Wir berechnen

$$\begin{aligned}
\chi_A(x) &= \det \begin{pmatrix} x & -1 & -2 & 0 & 2 \\ 1 & x+2 & -1 & 0 & 1 \\ -1 & -1 & x+2 & 0 & -1 \\ 1 & 4 & 0 & x-2 & 0 \\ -1 & -1 & 4 & 0 & x-3 \end{pmatrix} = (x-2) \cdot \det \begin{pmatrix} x & -1 & -2 & 2 \\ 1 & x+2 & -1 & 1 \\ -1 & -1 & x+2 & -1 \\ -1 & -1 & 4 & x-3 \end{pmatrix} \\
&= (x-2) \cdot \det \begin{pmatrix} x & -1 & -2 & 2 \\ 1 & x+2 & -1 & 1 \\ 0 & x+1 & x+1 & 0 \\ 0 & x+1 & 3 & x-2 \end{pmatrix} = (x-2) \cdot \det \begin{pmatrix} x & -1 & -2 & 0 \\ 1 & x+2 & -1 & 0 \\ 0 & x+1 & x+1 & x+1 \\ 0 & x+1 & 3 & x+1 \end{pmatrix} \\
&= (x-2)(x+1) \cdot \det \begin{pmatrix} x & -1 & -2 & 0 \\ 1 & x+2 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & x+1 & 3 & x+1 \end{pmatrix} = (x-2)(x+1) \cdot \det \left(\begin{array}{cc|cc} x & -1 & -2 & 0 \\ 1 & x+2 & -1 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & x+1 \end{array} \right) \\
&= (x-2)(x+1) \cdot [(x(x+2)+1) \cdot (x+1-3)] \\
&= (x-2)^2(x+1)^3.
\end{aligned}$$

`Schritt 2:` Also $\lambda_1 = -1$, mit $m_1 = \mu_a(A, -1) = 3$, und $\lambda_2 = 2$, mit $m_2 = \mu_a(A, 2) = 2$.

`Schritt 3:` $\boxed{i = 1}$ Wir bringen $A - \lambda_1 I_5$ auf reduzierte Zeilenstufenform:

$$\begin{aligned}
A - (-1) \cdot I_5 &= \begin{pmatrix} 1 & 1 & 2 & 0 & -2 \\ -1 & -1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 & 1 \\ -1 & -4 & 0 & 3 & 0 \\ 1 & 1 & -4 & 0 & 4 \end{pmatrix} \\
\begin{matrix} \mathbf{2} \to \mathbf{2} + (1) \cdot \mathbf{1} \\ \mathbf{3} \to \mathbf{3} + (-1) \cdot \mathbf{1} \\ \mathbf{4} \to \mathbf{4} + (1) \cdot \mathbf{1} \\ \mathbf{5} \to \mathbf{5} + (-1) \cdot \mathbf{1} \end{matrix} \quad &\rightsquigarrow \begin{pmatrix} 1 & 1 & 2 & 0 & -2 \\ 0 & 0 & 3 & 0 & -3 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & -3 & 2 & 3 & -2 \\ 0 & 0 & -6 & 0 & 6 \end{pmatrix} \\
\begin{matrix} \mathbf{3} \to \mathbf{3} + (1) \cdot \mathbf{2} \\ \mathbf{5} \to \mathbf{5} + (2) \cdot \mathbf{2} \\ \mathbf{2} \leftrightarrow \mathbf{4} \\ \mathbf{4} \leftrightarrow \mathbf{3} \end{matrix} \quad &\rightsquigarrow \begin{pmatrix} 1 & 1 & 2 & 0 & -2 \\ 0 & -3 & 2 & 3 & -2 \\ 0 & 0 & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

Also $\dim_{\mathbb{K}}(\mathrm{Ker}(A-(-1)\cdot I_5)) = k_{1,1} = 2$. Da

$$2 = k_{1,1} < k_{1,2} \leqslant 3 = \mu_a(A,-1),$$

folgt $k_{1,2} = k_{1,i} = 3$ für alle $i \geqslant 2$.

$\boxed{i=2}$ Wir bringen $A - \lambda_2 I_5$ auf reduzierte Zeilenstufenform:

$$\begin{aligned}
A-(2)\cdot I_5 &= \begin{pmatrix} -2 & 1 & 2 & 0 & -2 \\ -1 & -4 & 1 & 0 & -1 \\ 1 & 1 & -4 & 0 & 1 \\ -1 & -4 & 0 & 0 & 0 \\ 1 & 1 & -4 & 0 & 1 \end{pmatrix} \\
\mathbf{1} \leftrightarrow \mathbf{4} &\rightsquigarrow \begin{pmatrix} -1 & -4 & 0 & 0 & 0 \\ -1 & -4 & 1 & 0 & -1 \\ 1 & 1 & -4 & 0 & 1 \\ -2 & 1 & 2 & 0 & -2 \\ 1 & 1 & -4 & 0 & 1 \end{pmatrix} \\
\begin{matrix} \mathbf{2} \to \mathbf{2} + (-1)\cdot \mathbf{1} \\ \mathbf{3} \to \mathbf{3} + (1)\cdot \mathbf{1} \\ \mathbf{4} \to \mathbf{4} + (-2)\cdot \mathbf{1} \\ \mathbf{5} \to \mathbf{5} + (1)\cdot \mathbf{1} \end{matrix} &\rightsquigarrow \begin{pmatrix} -1 & -4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & -3 & -4 & 0 & 1 \\ 0 & 9 & 2 & 0 & -2 \\ 0 & -3 & -4 & 0 & 1 \end{pmatrix} \\
\begin{matrix} \mathbf{4} \to \mathbf{4} + (3)\cdot \mathbf{3} \\ \mathbf{5} \to \mathbf{5} + (-1)\cdot \mathbf{3} \end{matrix} &\rightsquigarrow \begin{pmatrix} -1 & -4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & -3 & -4 & 0 & 1 \\ 0 & 0 & -10 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
\begin{matrix} \mathbf{4} \to \mathbf{4} + (10)\cdot \mathbf{2} \\ \mathbf{3} \leftrightarrow \mathbf{2} \end{matrix} &\rightsquigarrow \begin{pmatrix} -1 & -4 & 0 & 0 & 0 \\ 0 & -3 & -4 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

Also $k_{2,1} = \dim_{\mathbb{K}} \mathrm{Ker}(A - \lambda_2 I_5) = 1$. Da

$$1 = k_{2,1} < k_{2,2} \leqslant 2 = \mu_a(A,2),$$

folgt $k_{2,2} = k_{2,i} = 2$ für alle $i \geqslant 2$.

`Schritt 4:` Für $i = 1, 2$ berechnen wir $d_{i,j}$ mit $j = 1, \ldots, \mu_a(A, \lambda_i)$.

$$\begin{aligned}
d_{1,1} &= -k_{1,2} + 2k_{1,1} - k_{1,0} = -3 + 4 - 0 = 1 \\
d_{1,2} &= -k_{1,3} + 2k_{1,2} - k_{1,1} = -3 + 6 - 2 = 1 \\
d_{1,3} &= -k_{1,4} + 2k_{1,3} - k_{1,2} = -3 + 6 - 3 = 0 \\
d_{2,1} &= -k_{2,2} + 2k_{2,1} - k_{2,0} = -2 + 2 - 0 = 0 \\
d_{2,2} &= -k_{2,3} + 2k_{2,2} - k_{2,1} = -2 + 4 - 1 = 1
\end{aligned}$$

`Schritt 5:`

$$\begin{aligned}
j_{\max}(1) &= \max\{j \mid d_{1,j} \neq 0\} = 2 \\
j_{\max}(2) &= \max\{j \mid d_{2,j} \neq 0\} = 2
\end{aligned}$$

`Ausgabe:`

$$\mathrm{JNF}(A) = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

$$\mathrm{mPol}_A(x) = (x-\lambda_1)^{j_{\max}(1)}(x-\lambda_2)^{j_{\max}(2)} = (x+1)^2(x-2)^2$$

$$\dim_{\mathbb{K}} \mathrm{Eig}(A,-1) = 2 \qquad \dim_{\mathbb{K}} \mathrm{Eig}(A,2) = 1$$

Nun zu der Basiswechselmatrix $T \in \mathrm{GL}_5(\mathbb{R})$ mit $T^{-1}AT = \mathrm{JNF}(A)$.

`Algorithmus B`

Wir führen den Algorithmus erst für $M = A - (-1)I_5$, dann für $M = A - 2I_5$ durch.

$\boxed{M = A - (-1)I_5}$

$$M = \begin{pmatrix} 1 & 1 & 2 & 0 & -2 \\ -1 & -1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 & 1 \\ -1 & -4 & 0 & 3 & 0 \\ 1 & 1 & -4 & 0 & 4 \end{pmatrix} \qquad M^2 = \begin{pmatrix} 0 & 0 & 9 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -9 & -6 & 9 & 6 \\ 0 & 0 & -9 & 0 & 9 \end{pmatrix}$$

Also M^2 hat Rang 2. Wir müssen nicht weiter rechnen, weil $k_{1,2} = \mu_a(A,-1) = 3$, also $\mathrm{Hau}(A,-1) = \mathrm{Ker}\, M^2$ und $m = 2$. Wir bekommen also `Schritt 0:`

$$k_0 = 0, \quad k_1 = 2, \quad k_2 = k_3 = 3, \qquad d_1 = -3+4-0 = 1, \quad d_2 = -3+6-2 = 1.$$

`Schritt 1:`

1. Wir suchen $u_{2,1} \in \mathrm{Ker}\, M^2$ mit $[u_{2,1}]$ linear unabhängig in $\mathrm{Ker}\, M^2 / \mathrm{Ker}\, M$. Das heißt einfach, $[u_{2,1}] \neq [0]$, also wir suchen $u_{2,1} \in \mathrm{Ker}\, M^2 \setminus \mathrm{Ker}\, M$. Da die erste Spalte von M^2 Null ist, aber die erste Spalte von M nicht, wählen wir:

$$u_{2,1} = e_1.$$

2. Wir berechnen

$$C_1 = \{M \cdot e_1\} = \{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \}.$$

`Schritt 2:` Wir suchen jetzt $u_{2,1} \in \operatorname{Ker} M$, sodass $[u_{2,1}]$ in $\operatorname{Ker} M / \operatorname{Span}_{\mathbb{K}}(C_1)$ linear unabhängig ist. Wir brauchen also einen Vektor in $\operatorname{Ker} M$ der nicht ein Vielfaches von $(1,-1,1,-1,1)^\top$ ist. Weil die 3te Spalte in M gleich mit (-1) mal die 5te Spalte in M ist, wählen wir

$$u_{2,1} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

`Ausgabe:`

$$T = \begin{pmatrix} 1 & 1 & 0 & ? & ? \\ -1 & 0 & 0 & ? & ? \\ 1 & 0 & 1 & ? & ? \\ -1 & 0 & 0 & ? & ? \\ 1 & 0 & 1 & ? & ? \end{pmatrix}$$

$\boxed{M = A - 2I_5}$

$$M = \begin{pmatrix} -2 & 1 & 2 & 0 & -2 \\ -1 & -4 & 1 & 0 & -1 \\ 1 & 1 & -4 & 0 & 1 \\ -1 & -4 & 0 & 0 & 0 \\ 1 & 1 & -4 & 0 & 1 \end{pmatrix} \qquad M^2 = \begin{pmatrix} 3 & -6 & -3 & 0 & 3 \\ 6 & 15 & -6 & 0 & 6 \\ -6 & -6 & 15 & 0 & -6 \\ 6 & 15 & -6 & 0 & 6 \\ -6 & -6 & 15 & 0 & -6 \end{pmatrix}$$

`Schritt 0:`

$$k_1 = 1, \quad k_2 = k_3 = 2, \qquad d_1 = -2 + 2 - 0 = 0, \quad d_2 = -2 + 4 - 1 = 1.$$

Wir brauchen also nur `Schritt 1` in diesem Fall.

`Schritt 1:`

1. Wir suchen, wie im Fall $\lambda = -1$, ein Vektor $u_{2,1} \in \operatorname{Ker} M^2 \setminus \operatorname{Ker} M$. Man betrachtet die zwei Matrizen, und merkt, dass die erste und fünfte Spalte von M^2 gleich sind, aber die erste und fünfte Spalte von M nicht gleich sind. Wir wählen also

$$u_{2,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

2.

$$C_1 = \{M \cdot u_{2,1}\} = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \}$$

`Ausgabe:` Die nächsten (und letzten) zwei Spalten von T.

$$T = \left(\begin{array}{ccc|cc} 1 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 \end{array}\right)$$

Wir berechnen jetzt

$$T^{-1} = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}$$

Probe:

$$\begin{aligned} T^{-1}AT &= \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 2 & 0 & -2 \\ -1 & -2 & 1 & 0 & -1 \\ 1 & 1 & -2 & 0 & 1 \\ -1 & -4 & 0 & 2 & 0 \\ 1 & 1 & -4 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 2 & -1 & 0 & 1 \\ -1 & -1 & 1 & 0 & -1 \\ 0 & -1 & -1 & 0 & 0 \\ 0 & 2 & 1 & -2 & -1 \\ 0 & 0 & 2 & 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 \end{pmatrix} = \\ &= \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} = \mathrm{JNF}(A) \end{aligned}$$

2. Wir betrachten jetzt wie viele verschiedene JNF Formen es für das charakteristische Polynom $\chi_A(x) = (x+1)^3(x-2)^2$ gibt. Für $\lambda_1 = -1$ haben wir folgende drei Möglichkeiten:

- Drei 1-Blöcke $J_1(-1)$.
- Ein 2-Block $J_2(-1)$ und ein 1-Block $J_1(-1)$.
- Ein 3-Block $J_3(-1)$.

Für den Eigenwert $\lambda_2 = 2$ haben wir nur zwei Möglichkeiten:

- Zwei 1-Blöcke $J_1(2)$.
- Ein 2-Block $J_2(2)$.

Insgesamt haben wir also $6 = 3 \times 2$ Möglichkeiten:

$d_{1,1}, d_{1,2}, d_{1,3}$	$d_{2,1}, d_{2,2}$	JNF A	mPol_A
$3, 0, 0$	$2, 0$	$\left(\begin{array}{ccc\|cc} -1 & & & & \\ & -1 & & & \\ & & -1 & & \\ \hline & & & 2 & \\ & & & & 2 \end{array}\right)$	$(x+1)(x-2)$
$3, 0, 0$	$0, 1$	$\left(\begin{array}{ccc\|cc} -1 & & & & \\ & -1 & & & \\ & & -1 & & \\ \hline & & & 2 & 1 \\ & & & & 2 \end{array}\right)$	$(x+1)(x-2)^2$
$1, 1, 0$	$2, 0$	$\left(\begin{array}{ccc\|cc} -1 & 1 & & & \\ & -1 & & & \\ & & -1 & & \\ \hline & & & 2 & \\ & & & & 2 \end{array}\right)$	$(x+1)^2(x-2)$
$1, 1, 0$	$0, 1$	$\left(\begin{array}{ccc\|cc} -1 & 1 & & & \\ & -1 & & & \\ & & -1 & & \\ \hline & & & 2 & 1 \\ & & & & 2 \end{array}\right)$	$(x+1)^2(x-2)^2$
$0, 0, 1$	$2, 0$	$\left(\begin{array}{ccc\|cc} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & & \\ \hline & & & 2 & \\ & & & & 2 \end{array}\right)$	$(x+1)^3(x-2)$
$0, 0, 1$	$0, 1$	$\left(\begin{array}{ccc\|cc} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & & \\ \hline & & & 2 & 1 \\ & & & & 2 \end{array}\right)$	$(x+1)^3(x-2)^2$

10.13 Das Matrixexponential

Eine wichtige Anwendung der Jordanschen Normalform ist im Gebiet der Differentialgleichungen. Ein homogenes lineares System von Differentialgleichungen des ersten Grades ist eine Matrixgleichung der Form:

$$\mathbf{y}'(t) = A \cdot \mathbf{y}(t),$$

wobei $A \in \mathrm{Mat}_n(\mathbb{C})$ und $\mathbf{y} : \mathbb{C} \longrightarrow \mathbb{C}^n$ eine differenzierbare Abbildung, mit Ableitung $\mathbf{y}'$ ist:

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} \qquad \text{und} \qquad \mathbf{y}'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_n'(t) \end{pmatrix}.$$

Die Lösungsmenge ist ein $\mathbb{C}$-Vektorraum. Man kann zeigen, dass der Lösungsraum von der Spalten folgender Matrix erzeugt ist:

$$e^{At}.$$

Wir werden uns in diesem Teil um e^A kümmern. Dafür werden wir nur Matrizen mit Einträgen in $\mathbb{R}$ oder in $\mathbb{C}$ betrachten. Wir werden aber alles nur für $\mathbb{C}$ aufschreiben.
Die Exponentialfunktion $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ ist durch eine unendliche, aber konvergente, Summe definiert:

$$x \longmapsto e^x := 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots.$$

Das Exponential einer Matrix ist durch einsetzen von A an der Stelle von x erhalten. Man muss aber sicher sein, dass die neue Summe immer noch konvergiert. Wir werden das aber ohne Beweis anwenden.

Definition 10.77. Sei $A \in \mathrm{Mat}_n(\mathbb{C})$. Das **Exponential der Matrix** A ist die Matrix

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Allgemein ist es schwierig e^A mit der Definition direkt zu berechnen. Für diagonale Matrizen ist es aber einfach.

Bemerkung 10.78. Wenn $A = \mathrm{diag}(a_1, \dots, a_n) \in \mathrm{Mat}_n(\mathbb{C})$ diagonal ist, dann ist

$$e^A = \mathrm{diag}(e^{a_1}, \dots, e^{a_n}).$$

Auch die Ähnlichkeit der Matrizen ist mit dem Exponential gut verträglich.

Bemerkung 10.79. Seien $A \sim B \in \mathrm{Mat}_n(\mathbb{C})$ ähnliche Matrizen mit $A = T^{-1}BT$. Dann haben wir $A^k = T^{-1}B^kT$ für alle k, also

$$e^A = T^{-1}e^BT.$$

Das heißt, dass wenn A diagonalisierbar ist, dann können wir das zu erst diagonalisieren, dann potenzieren, und dann wieder mit der Übergangsmatrix konjugieren: Wenn $A = T^{-1}\,\mathrm{diag}(a_1, \dots, a_n)T$, dann

$$e^A = T^{-1}\,\mathrm{diag}(e^{a_1}, \dots, e^{a_n})T.$$

Für nilpotente Matrizen ist es auch einfacher, weil die Summe endlich ist. Zum Beispiel

$$e^{\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \frac{1}{1!}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} + \frac{1}{2!}\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.$$

Unser Ziel ist jetzt die JNF anzuwenden, um das Exponential beliebiger Matrizen in $\mathrm{Mat}_n(\mathbb{C})$ zu berechnen. Zu erst, eine Bemerkung zu einer bekannten Formel für das Exponential komplexer Zahlen:

Bemerkung 10.80. Die Formel $e^{x+y} = e^x e^y$ beweist man durch vergleichen der unendlichen Summen

$$\begin{aligned} e^{x+y} &= 1 + \frac{x+y}{1!} + \frac{(x+y)^2}{2!} + \frac{(x+y)^3}{3!} + \dots \\ e^x e^y &= \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\right) \cdot \left(1 + \frac{y}{1!} + \frac{y^2}{2!} + \frac{y^3}{3!} + \dots\right) \end{aligned}$$

Die Gleichheit folgt nur wenn $xy = yx$. Zum Beispiel, wenn man den Grad 2 Teil vergleicht, dann haben wir

$$\frac{1}{2}\left(x^2 + xy + yx + y^2\right) = \frac{x^2}{2} + xy + \frac{y^2}{2} \qquad \Longleftrightarrow \qquad xy = yx.$$

Also für Matrizen gilt nicht allgemein $e^{A+B} = e^A e^B$. Es gilt nur dann, wenn $AB = BA$.

Korollar 10.81. Für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$, ist e^A invertierbar, mit Inverse e^{-A}.

Bemerkung 10.80 ist der Grund warum wir in Satz 10.60 Punkt (d) gezeigt haben. Nämlich, wir haben gezeigt, dass es für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ eine diagonalisierbare Matrix D und eine nilpotente Matrix N

$$A = D + N \qquad \text{und} \qquad DN = ND.$$

Es gilt also, nach Bemerkung 10.80, dass

$$e^A = e^{D+N} = e^D \cdot e^N.$$

Insbesondere, wenn man die Jordansche Normalform einer Matrix kennt, und die Übergangsmatrix $T \in \mathrm{GL}_n(\mathbb{C})$ kennt mit $T^{-1}AT = \mathrm{JNF}(A)$, dann kann man e^A schnell berechnen.

Beispiele:

1. Wenn $A = 0 \in \mathrm{Mat}_n(\mathbb{C})$, dann $e^A = I_n$.
2. Wenn $A = I_n$, dann $e^{I_n} = e \cdot I_n$.
3. $e^{\lambda A} = e^\lambda \cdot e^A$.
4. Wenn $A = J_3(2) = \begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix}$, dann bezeichnen wir mit $B = A - 2I_3$ die nilpotente Matrix,

und bekommen

$$\begin{aligned} e^A &= e^{2I_3+B} = e^{2I_3}e^B \\ &= e^2 \cdot I_3 \cdot \left[I_3 + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \\ &= \begin{pmatrix} e^2 & 1 & \frac{1}{2} \\ 0 & e^2 & 1 \\ 0 & 0 & e^2 \end{pmatrix}. \end{aligned}$$

Jacobi's Formel: $\det(e^A) = e^{\mathrm{Spur}(A)}$.

10.14 Sehr kurze Zusammenfassung

Die Jordansche Normalform hat verschiedene wichtige Anwendungen. Unser Ziel war einen "kanonischen" Repräsentanten in jeder Äquivalenzklasse für die Ähnlichkeit der Matrizen zu finden. Das gibt uns eine Methode zu entscheiden, dass zwei Matrizen A und B ähnlich sind, ohne eine Übergangsmatrix explizit zu berechnen. Auf dem Weg zu der JNF müssen mehrere Zahlen übereinstimmen, damit $A \sim B$. Dieses Verhalten nennen wir invariant (unter Ähnlichkeit). Wir machen das etwas genauer.

Definition 10.82. Eine Abbildung $j : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow X$, wobei X eine Menge ist, heißt **Invariante** einer Matrix (bis auf Ähnlichkeit), falls

$$A \sim B \quad \Rightarrow \quad j(A) = j(B).$$

Also, falls j konstant auf Äquivalenzklassen ist.

Aus der universellen Eigenschaft der Faktormenge, heißt das, dass wir eine Abbildung

$$\overline{j} : \mathrm{Mat}_n(\mathbb{K})/\sim \longrightarrow X$$

definieren können.

Satz 10.83. *Folgende Abbildungen sind Invarianten von Matrizen bis auf Ähnlichkeit.*

(a) *Der Rang einer Matrix*

$$\mathrm{Rang} : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{N}.$$

(b) *Die Spur einer Matrix*

$$\mathrm{Spur} : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{K}.$$

(c) *Die Determinante einer Matrix*

$$\det : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{K}.$$

(d) *Die Menge der Eigenwerte einer Matrix*

$$EW : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \{\textit{endliche Teilmengen von } \mathbb{K}\}.$$

(e) *Das charakteristische Polynom*

$$\chi : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{K}[x].$$

(f) *Das Minimalpolynom*

$$\mathrm{mPol} : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{K}[x].$$

(g) *Die algebraische Vielfachheit*

$$\mu_a(-, \lambda) : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{N}.$$

(h) *Die geometrische Vielfachheit*

$$\mu_g(-, \lambda) : \mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathbb{N}.$$

(i) *Die Anzahl der Jordan-Blöcke*[14] *der Form* $J_j(\lambda)$ *in der JNF*

$$d_{\lambda,j} : \mathrm{Mat}_n(\mathbb{C}) \longrightarrow \mathbb{N}.$$

[14]8.6.'22

Alle diese Invarianten kann man leicht von der Jordanschen Normalform ablesen.

[14]Hier brauchen wir $\mathbb{K}$ algebraisch abgeschlossen. Wir schreiben es aber nur für $\mathbb{K}$=$\mathbb{C}$.

Kapitel 11

Bilinearformen

11.1 Definition und zugeordnete Matrizen

Seien V, W, U drei endlichdimensionale $\mathbb{K}$-Vektorräume. Das kartesische Produkt $V \times W$ ist auch ein endlichdimensionaler $\mathbb{K}$-Vektorraum, mit komponentenweise Vektoraddition und Multiplikation mit Skalaren. Es gilt

$$\dim_{\mathbb{K}} V \times W = \dim_{\mathbb{K}} V + \dim_{\mathbb{K}} W.$$

Bilineare Abbildungen haben ein Kartesisches Produkt als Definitionsbereich. Die Kompatibilität ist aber nicht mit der $\mathbb{K}$-Vektorraumstruktur des Produktes, sonder mit der der einzelnen Faktoren. Das heißt, eine Bilineare Abbildung nimmt zwei Argumente und ist linear in jedem.

Definition 11.1. Eine Abbildung $\varphi : V \times W \longrightarrow U$ ist eine **bilineare Abbildung** wenn für alle $v, v' \in V$, $w, w' \in W$ und $\lambda, \mu \in \mathbb{K}$ gilt

(a) $\varphi(\lambda \cdot v + \mu \cdot v', w) = \lambda \cdot \varphi(v, w) + \mu \cdot \varphi(v', w)$,

(b) $\varphi(v, \lambda \cdot w + \mu \cdot w') = \lambda \cdot \varphi(v, w) + \mu \cdot \varphi(v, w')$.

Eine **Paarung** zwischen V und W ist eine bilineare Abbildung $\varphi : V \times W \longrightarrow \mathbb{K}$.
Eine **Bilinearform** auf V ist eine bilineare Abbildung $\varphi : V \times V \longrightarrow \mathbb{K}$.

Wenn der Wertebereich der bilinearen Abbildung der 1-dimensionale $\mathbb{K}$-Vektorraum $\mathbb{K}$ist, dann kann man Matrizen zuordnen.

Definition 11.2. Seien $\mathcal{B} = v_1, \ldots, v_n$ eine Basis von V und $\mathcal{C} = w_1, \ldots, w_m$ eine Basis von W. Sei $\varphi : V \times W \longrightarrow \mathbb{K}$ eine Paarung. Die **darstellenden Matrix** der Paarung φ bezüglich der geordneten Basen $\mathcal{B}$ und $\mathcal{C}$ ist die Matrix:

$${}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}(\varphi) = \big(\varphi(v_i, w_j)\big) \in \mathrm{Mat}_{n,m}(\mathbb{K}).$$

Wenn wir eine Bilinearform $\varphi : V \times V \longrightarrow \mathbb{K}$ auf V betrachten, dann betrachten wir (wie im Fall der Endomorphismen) nur darstellende Matrizen bezüglich einer einzigen Basis, also

$$\mathcal{M}^{\mathcal{B}}(\varphi) := {}^{\mathcal{B}}\mathcal{M}^{\mathcal{B}}(\varphi) \in \mathrm{Mat}_n(\mathbb{K}).$$

Beispiele:

1 $\varphi : \mathbb{K}^2 \times \mathbb{K}^2 \longrightarrow \mathbb{K}$ mit $\varphi\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$. Die zugeordnete Matrix ist $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

2 Die kanonische Bilinearform $\langle \ | \ \rangle : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}$ gegeben durch

$$\langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$$

für alle $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$.

3 Für den unendlich-dimensionalen $\mathbb{R}$-Vektorraum $V = \mathcal{C}([a,b]; \mathbb{R}) = \{f : [a,b] \longrightarrow \mathbb{R} \ : \ f \text{ ist stetig}\}$ ist folgende Bilinearform wichtig:

$$V \times V \longrightarrow \mathbb{R} \qquad (f,g) \longmapsto \int_a^b f(t) \cdot g(t) dt.$$

4 Die Paarung $\langle \ | \ \rangle : V^* \times V \longrightarrow \mathbb{K}$, wobei $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ der Dualraum von V ist, gegeben durch

$$\langle f \mid v \rangle := f(v) \qquad \forall \ f \in V^*, v \in V.$$

Bemerkung 11.3. Seien V und W zwei $\mathbb{K}$-Vektorräume mit Basen $\mathcal{B} = v_1, \ldots, v_n$, beziehungsweise $\mathcal{C} = w_1, \ldots, w_m$. Seien $v \in V$, $w \in W$ und $\varphi : V \times W \longrightarrow \mathbb{K}$ eine Paarung. Wir bezeichnen die Koordinaten von v und w bezüglich $\mathcal{B}$, beziehungsweise $\mathcal{C}$, durch $a_i, b_i \in \mathbb{K}$. Das heißt, diese sind die eindeutig bestimmten Skalare mit

$$v = \sum_{i=1}^{n} a_i v_i \ \text{ und } w = \sum_{j=1}^{m} b_j w_j.$$

Dann gilt

$$\varphi(v,w) = \varphi\left(\textstyle\sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{m} b_j w_j\right) = \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \varphi(v_i, w_j) = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \cdot {}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}(\varphi) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Umgekehrt, für jede Matrix $A \in \mathrm{Mat}_{n,m}(\mathbb{K})$ definiert man die Paarung ${}^{\mathcal{B}}\varphi_A^{\mathcal{C}} : V \times W \longrightarrow \mathbb{K}$ durch

$${}^{\mathcal{B}}\varphi_A^{\mathcal{C}}\left(\textstyle\sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{m} b_j w_j\right) := \begin{pmatrix} a_1 & \ldots & a_n \end{pmatrix} \cdot A \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Eine direkte Überprüfung zeigt, dass

$${}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}({}^{\mathcal{B}}\varphi_A^{\mathcal{C}}) = A.$$

Beispiel 11.4. Seien $V = W = \mathbb{R}^2$ mit der Standardbasis $\mathcal{B} = \mathcal{C} = e_1, e_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und sei $\varphi : V \times V \longrightarrow \mathbb{K}$ eine Bilinearform über der wir wissen, dass

$$\begin{array}{lllllll} \varphi(e_1, e_1) & = & 1 & \varphi(e_1, e_2) & = & 2 \\ \varphi(e_2, e_1) & = & 3 & \varphi(e_2, e_2) & = & 4. \end{array}$$

Also

$$\mathcal{M}^{\mathcal{B}}(\varphi) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Dann haben wir für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, dass

$$\begin{aligned}
\varphi(\mathbf{x}, \mathbf{y}) &= \varphi(x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) \\
&= x_1 \cdot \varphi(e_1,\ y_1 e_1 + y_2 e_2) + x_2 \cdot \varphi(e_2,\ y_1 e_1 + y_2 e_2) \\
&= x_1 \cdot \varphi(e_1, e_1) \cdot y_1 + x_1 \cdot \varphi(e_1, e_2) \cdot y_2 + x_2 \cdot \varphi(e_2, e_1) \cdot y_1 + x_2 \cdot \varphi(e_2, e_2) \cdot y_2 \\
&= x_1 \cdot 1 \cdot y_1 + x_1 \cdot 2 \cdot y_2 + x_2 \cdot 3 \cdot y_1 + x_2 \cdot 4 \cdot y_2 \\
&= x_1 y_1 + 2 x_1 y_2 + 3 x_2 y_1 + 4 x_2 y_2 \\
&= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} 1 \cdot y_1 + 2 \cdot y_2 \\ 3 \cdot y_1 + 4 \cdot y_2 \end{pmatrix} \\
&= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{\top} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.
\end{aligned}$$

Wir bezeichnen[1] die Menge der Paarungen zwischen zwei $\mathbb{K}$-Vektorräume V und W mit $\mathrm{Bil}_{\mathbb{K}}(V, W)$. Wir bezeichnen die Menge der Bilinearformen auf V mit $\mathrm{Bil}_{\mathbb{K}}(V)$.

Bemerkung 11.5. Für alle $\varphi_1, \varphi_2, \varphi \in \mathrm{Bil}_{\mathbb{K}}(V, W)$ und alle $\lambda \in \mathbb{K}$ definieren wir die Abbildungen $\varphi_1 + \varphi_2$, beziehungsweise $\lambda \cdot \varphi$ durch

$$\begin{aligned}
(\varphi_1 + \varphi_2)(v, w) &= \varphi_1(v, w) + \varphi_2(v, w), \quad \forall\ (v, w) \in V \times W, \\
(\lambda \cdot \varphi)(v, w) &= \lambda \cdot (\varphi(v, w)), \quad \forall\ (v, w) \in V \times W.
\end{aligned}$$

Allein mit der Definition kann man überprüfen, dass

$$\begin{aligned}
\varphi_1 + \varphi_2 &\in \mathrm{Bil}_{\mathbb{K}}(V, W), \quad \forall\ \varphi_1, \varphi_2 \in \mathrm{Bil}_{\mathbb{K}}(V, W), \\
\lambda \cdot \varphi &\in \mathrm{Bil}_{\mathbb{K}}(V, W), \quad \forall\ \varphi \in \mathrm{Bil}_{\mathbb{K}}(V, W) \text{ und } \forall\ \lambda \in \mathbb{K}.
\end{aligned}$$

Es ist eine einfache Übung zu beweisen, dass $\mathrm{Bil}_{\mathbb{K}}(V, W)$ mit den obigen Operationen ein $\mathbb{K}$-Vektorraum ist. Insbesondere ist $\mathrm{Bil}_{\mathbb{K}}(V)$ ein $\mathbb{K}$-Vektorraum.

Aus den obigen Bemerkungen bekommen wir den folgenden Satz.

Satz 11.6. *Seien V und W zwei $\mathbb{K}$-Vektorräume mit $\dim_{\mathbb{K}} V = n < \infty$ und $\dim_{\mathbb{K}} W = m < \infty$. Seien $\mathcal{B}$ und $\mathcal{C}$ geordnete Basen von V, beziehungsweise W. Die Abbildung ${}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}} : \mathrm{Bil}_{\mathbb{K}}(V, W) \longrightarrow \mathrm{Mat}_{n,m}(\mathbb{K})$, gegeben durch*

$$\varphi \longmapsto {}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}(\varphi)$$

ist ein Isomorphismus von $\mathbb{K}$-Vektorräume.

[1] $\mathrm{Paar}(V, W)$ würde besser passen mit dem Namen, aber letztendlich ist Bilinearität die Eigenschaft die wichtig ist.

Beweis-Skizze: Übung Q.E.D.

Korollar 11.7. Für endlichdimensionale $\mathbb{K}$-Vektorräume V und W gilt

$$\dim_{\mathbb{K}} \mathrm{Bil}_{\mathbb{K}}(V,W) = (\dim_{\mathbb{K}} V)(\dim_{\mathbb{K}} W) \qquad \text{und} \qquad \dim_{\mathbb{K}} \mathrm{Bil}_{\mathbb{K}}(V) = (\dim_{\mathbb{K}} V)^2.$$

Satz 11.8. *Die Abbildung* $\Psi : \mathrm{Bil}_{\mathbb{K}}(V,W) \longrightarrow \mathrm{Hom}_{\mathbb{K}}(W, V^*)$ *gegeben durch*

$$\begin{array}{rccc} \Psi(\varphi): & W & \longrightarrow & V^* \\ & \cup & & \cup \\ & w & \longmapsto & \varphi(w,\ \cdot\ ) : V \longrightarrow \mathbb{K} \end{array}, \qquad \forall\ \varphi \in \mathrm{Bil}_{\mathbb{K}}(V,W)$$

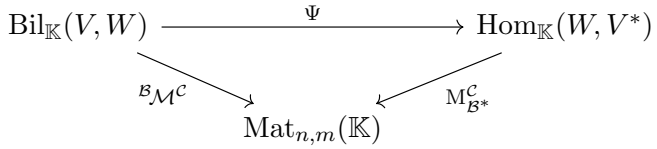
ist ein Isomorphismus von $\mathbb{K}$*-Vektorräume.*

Beweis-Skizze: Die Linearität von Ψ ist eine direkte Überprüfung der Axiome. Aus Korollar 11.7, $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V^*$ und der Dimensionsformel für $\dim \mathrm{Hom}_{\mathbb{K}}(W, V^*)$ haben wir

$$\dim_{\mathbb{K}} \mathrm{Bil}_{\mathbb{K}}(V,W) = \dim_{\mathbb{K}} \mathrm{Hom}_{\mathbb{K}}(W, V^*).$$

Also, um zu zeigen, dass Ψ ein Isomorphismus ist, reicht es zu zeigen, dass es injektiv ist, also dass $\mathrm{Ker}\,\Psi = \{0\}$. Sei $\varphi \in \mathrm{Ker}\,\Psi$. Das heißt, dass für alle $w \in W$ gilt $\varphi(w,\ \cdot\ ) = 0 \in V^*$. Das heißt $\varphi(w,v) = 0$ für alle $v \in V$ und $w \in W$, also $\varphi = 0$. Q.E.D.

Korollar 11.9. Wenn $\mathrm{M}_{\mathcal{C}}^{\mathcal{B}} : \mathrm{Hom}_{\mathbb{K}}(V,W) \longrightarrow \mathrm{Mat}_{n,m}$ den Isomorphismus der einer linearen Abbildung f die Matrix $\mathrm{M}_{\mathcal{C}}^{\mathcal{B}}(f)$ zuordnet ist, dann haben wir folgendes kommutative Diagramm:



In anderen Worten gilt

$${}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}(\varphi) = \mathrm{M}^{\mathcal{C}}_{\mathcal{B}^*}(\Psi(\varphi)).$$

Satz 11.10. *Seien* V, W *endlichdimensionale* $\mathbb{K}$*-Vektorräume,* $\mathcal{B}, \mathcal{B}'$ *Basen von* V *und* $\mathcal{C}, \mathcal{C}'$ *Basen von* W*. Dann gilt:*

$${}^{\mathcal{B}'}\mathcal{M}^{\mathcal{C}'}(\varphi) = \left(\mathrm{M}^{\mathcal{B}'}_{\mathcal{B}}\right)^{\top} \cdot {}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}(\varphi) \cdot \mathrm{M}^{\mathcal{C}'}_{\mathcal{C}}$$

Wir erinnern kurz wie Basiswechselmatrizen funktionieren. Seien $\mathcal{B} = v_1, \ldots, v_n$ und $\mathcal{B}' = v_1', \ldots, v_n'$ zwei geordnete Basen von V. Die Basiswechselmatrix ist definiert als

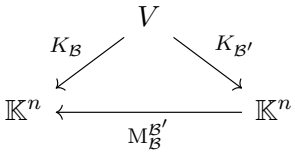
$$\mathrm{M}^{\mathcal{B}'}_{\mathcal{B}} := \mathrm{M}^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V).$$

Also die Spalten von $\mathrm{M}^{\mathcal{B}'}_{\mathcal{B}}$ sind die Koordinaten der Vektoren $v_1', \ldots, v_n'$ bezüglich der Basis $\mathcal{B}$. Die Wirkung von $\mathrm{M}^{\mathcal{B}'}_{\mathcal{B}}$ ist aber auf der Koordinaten umgekehrt: man bekommt die Koordinaten bezüglich $\mathcal{B}$ wenn man die Koordinaten bezüglich $\mathcal{B}'$ links mit $\mathrm{M}^{\mathcal{B}'}_{\mathcal{B}}$ multipliziert. Um das genauer zu machen,

bezeichnen wir mit $K_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n$ den Isomorphismus den man der Basiswahl zuordnet, das heißt die $\mathbb{K}$-lineare Abbildung mit Inverse

$$\begin{array}{rccc} K_{\mathcal{B}}^{-1}: & \mathbb{K}^n & \longrightarrow & V \\ & \cup & & \cup \\ & \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} & \longmapsto & a_1 v_1 + \cdots + a_n v_n. \end{array}$$

Dann haben wir das kommutative Diagramm



Direkt aus der Definition folgt

$$\left(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}'}\right)^{-1} = \mathrm{M}_{\mathcal{B}'}^{\mathcal{B}}.$$

Wenn $\mathcal{B}^*$ und $(\mathcal{B}')^*$ die Dualbasen des Dualraumes V^* bezüglich $\mathcal{B}$ und $\mathcal{B}'$ bezeichnen, dann gilt

$$\mathrm{M}_{\mathcal{B}^*}^{(\mathcal{B}')^*} = \left(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}'}\right)^{\top}.$$

Wir erinnern noch, dass für Homomorphismen $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ haben wir folgende Gleichheit:

$$\mathrm{M}_{C'}^{B'}(f) = \mathrm{M}_{C'}^{C} \cdot \mathrm{M}_{C}^{B}(f) \cdot \mathrm{M}_{B}^{B'}.$$

Beweis-Skizze: [von Satz 11.8]

Wir bringen einfach alle obige Gleichheiten zusammen.

$$\begin{aligned} {}^{\mathcal{B}'}\mathcal{M}^{\mathcal{C}'}(\varphi) &= \mathrm{M}_{(\mathcal{B}')^*}^{\mathcal{C}'}(\Psi(\varphi)) \\ &= \mathrm{M}_{(\mathcal{B}')^*}^{\mathcal{B}^*} \cdot \mathrm{M}_{\mathcal{B}^*}^{\mathcal{C}}(\Psi(\varphi)) \cdot \mathrm{M}_{\mathcal{C}}^{\mathcal{C}'} \\ &= \left(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}'}\right)^{\top} \cdot \mathrm{M}_{\mathcal{B}^*}^{\mathcal{C}}(\Psi(\varphi)) \cdot \mathrm{M}_{\mathcal{C}}^{\mathcal{C}'} \\ &= \left(\mathrm{M}_{\mathcal{B}}^{\mathcal{B}'}\right)^{\top} \cdot {}^{\mathcal{B}}\mathcal{M}^{\mathcal{C}}(\varphi) \cdot \mathrm{M}_{\mathcal{C}}^{\mathcal{C}'}. \end{aligned}$$

Q.E.D.

Korollar 11.11. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, seien $\mathcal{B}, \mathcal{B}'$ zwei Basen von V mit Basiswechselmatrix $S = \mathrm{M}_{\mathcal{B}}^{\mathcal{B}'}$ und sei $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine Bilinearform auf V. Es gilt

$$\mathcal{M}^{\mathcal{B}'}(\varphi) = S^{\top} \cdot \mathcal{M}^{\mathcal{B}}(\varphi) \cdot S,$$

wobei $S^{\top}$ die Transponierte Matrix von S ist.

Um zu untersuchen wann zwei Matrizen dieselbe Bilinearform darstellen können, führen wir folgende Äquivalenzrelation ein.

Definition 11.12. Zwei quadratische Matrizen $A, B \in \mathrm{Mat}_n(\mathbb{K})$ sind **kongruent** wenn es eine Matrix $S \in \mathrm{GL}_n(\mathbb{K})$ existiert, sodass

$$A = S^\top \cdot B \cdot S.$$

Wir bezeichnen diese Relation mit $A \approx B$.

Bemerkung 11.13. Zwei Matrizen A und B sind also kongruent wenn es eine Bilinearform φ und zwei Basen $\mathcal{A}$ und CB gibt, sodass $\mathcal{M}^{\mathcal{A}}(\varphi) = A$ und $\mathcal{M}^{\mathcal{B}}(\varphi) = B$. Die Kongruenz der Matrizen ist also eine Äquivalenzrelation auf $\mathrm{Mat}_n(\mathbb{K})$.

11.2 Symmetrische, antisymmetrische und alternierende Bilinearformen

Definition 11.14. Sei V ein $\mathbb{K}$-Vektorraum.

(a) Eine Bilinearform φ auf V heißt **symmetrische Bilinearform** wenn

$$\varphi(v, v') = \varphi(v', v), \quad \forall\, v, v' \in V.$$

Die Menge aller symmetrischen Bilinearformen auf V ist mit $\mathrm{Bil}_{\mathbb{K}}^{\mathrm{sym}}(V)$ bezeichnet.

(b) Eine Bilinearform φ auf V heißt **antisymmetrische Bilinearform** (oder **schiefsymmetrische Bilinearform**) wenn

$$\varphi(v, v') = -\varphi(v', v), \quad \forall\, v, v' \in V.$$

Die Menge aller antisymmetrischen Bilinearformen auf V ist mit $\mathrm{Bil}_{\mathbb{K}}^{\text{a-sym}}(V)$ bezeichnet.

(c) Eine Bilinearform $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$ heißt **alternierende Bilinearform** wenn

$$\varphi(v, v) = 0, \quad \forall\, v \in V.$$

Die Menge aller antisymmetrischen Bilinearformen auf V ist mit $\mathrm{Bil}_{\mathbb{K}}^{\mathbf{alt}}(V)$ bezeichnet.

Definition 11.15. Sei $n \in \mathbb{N}_{>0}$.

(a) Eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ heißt **symmetrische Matrix** wenn

$$A = A^\top.$$

Die Menge aller symmetrischen $n \times n$ Matrizen mit Einträgen in $\mathbb{K}$ ist mit $\mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K})$ bezeichnet.

(b) Eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ heißt **antisymmetrische Matrix** (oder **schiefsymmetrische Matrix**) wenn

$$A = -A^\top.$$

Die Menge aller antisymmetrischen $n \times n$ Matrizen mit Einträgen in $\mathbb{K}$ ist mit $\mathrm{Mat}_n^{\text{a-sym}}(\mathbb{K})$ bezeichnet.

(c) Eine quadratische Matrix $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ heißt **alternierende Matrix** wenn

$$A = -A^\top \quad \text{und } a_{ii} = 0 \quad \text{für alle } i = 1, \ldots, n.$$

Die Menge aller alternierenden $n \times n$ Matrizen mit Einträgen in $\mathbb{K}$ ist mit $\mathrm{Mat}_n^{\text{a-sym}}(\mathbb{K})$ bezeichnet.

Bemerkung 11.16. (a) Die Mengen $\mathrm{Bil}_{\mathbb{K}}^{\mathrm{sym}}(V)$ und $\mathrm{Bil}_{\mathbb{K}}^{\text{a-sym}}(V)$ sind $\mathbb{K}$-Untervektorräume von $\mathrm{Bil}_{\mathbb{K}}(V)$.

(b) Die Mengen $\mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K})$ und $\mathrm{Mat}_n^{\text{a-sym}}(\mathbb{K})$ sind $\mathbb{K}$-Untervektorräume von $\mathrm{Mat}_n(\mathbb{K})$.

(c) Wenn char $\mathbb{K} = 2$, das heißt wenn im Körper $\mathbb{K}$ die Gleichheit $2 := 1 + 1 = 0$ gilt, das heißt $1 = -1$, dann sind die Begriffe *symmetrisch* und *antisymmetrisch* sowohl für Bilinearformen als auch für Matrizen gleich.

(d) Wenn char $\mathbb{K} \neq 2$, das heißt wenn $1 \neq -1$, dann haben wir

$$\mathrm{Mat}_n(\mathbb{K}) = \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K}) \oplus \mathrm{Mat}_n^{\text{a-sym}}(\mathbb{K}).$$

Das gilt weil $\mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K}) \cap \mathrm{Mat}_n^{\text{a-sym}}(\mathbb{K}) = \{0\}$ und für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ gilt

$$A = \tfrac{1}{2}(A + A^\top) + \tfrac{1}{2}(A - A^\top),$$

wobei $\frac{1}{2}(A + A^\top)$ symmetrisch und $\frac{1}{2}(A - A^\top)$ antisymmetrisch ist.

[15]13.6.'22

Beispiele:

1. Die kanonische Bilinearform auf $\mathbb{K}^n$ ist symmetrisch.
2. Die Determinante $\det : \mathbb{K}^2 \times \mathbb{K}^2 \longrightarrow \mathbb{K}$ ist antisymmetrisch.

Lemma 11.17. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, $\mathcal{B}$ eine Basis von V, und $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$. Es gelten*

(a) *φ ist symmetrisch $\Leftrightarrow$ $\mathcal{M}^{\mathcal{B}}(\varphi)$ ist symmetrisch.*

(b) *φ ist antisymmetrisch $\Leftrightarrow$ $\mathcal{M}^{\mathcal{B}}(\varphi)$ ist antisymmetrisch.*

Beweis-Skizze: Man setzt einfach die Definition ein. Q.E.D.

Korollar 11.18. Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $\mathcal{B}$ eine Basis von V. Die Einschränkungen des $\mathbb{K}$-Isomorphismus $\mathcal{M}^{\mathcal{B}} : \mathrm{Bil}_{\mathbb{K}}(V) \longrightarrow \mathrm{Mat}_n(\mathbb{K})$ auf $\mathrm{Bil}_{\mathbb{K}}^{\mathrm{sym}}(V)$, beziehungsweise $\mathrm{Bil}_{\mathbb{K}}^{\text{a-sym}}(V)$, induzieren Isomorphismen der entsprechenden Unterräume:

$$\mathcal{M}^{\mathcal{B}} : \mathrm{Bil}_{\mathbb{K}}^{\mathrm{sym}}(V) \longrightarrow \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K}),$$

$$\mathcal{M}^{\mathcal{B}} : \mathrm{Bil}_{\mathbb{K}}^{\text{a-sym}}(V) \longrightarrow \mathrm{Mat}_n^{\text{a-sym}}(\mathbb{K}).$$

Proposition 11.19. *Sei V ein $\mathbb{K}$-Vektorraum und $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine Bilinearform.*

(a) *Wenn φ alternierend ist, dann ist φ antisymmetrisch.*

(b) *Wenn $2 \neq 0$ in $\mathbb{K}$ und φ ist antisymmetrisch, dann ist φ alternierend.*

Beweis-Skizze:

(a) Seien $v, v' \in V$ beliebig. Dann gilt

$$0 = \varphi(v+v', v+v') = \varphi(v,v) + \varphi(v,v') + \varphi(v',v) + \varphi(v',v') = \varphi(v,v') + \varphi(v',v).$$

(b) Sei $v \in V$. Es gilt

$$\varphi(v,v) = -\varphi(v,v) \quad \Rightarrow \quad 2\varphi(v,v) = 0 \quad \Rightarrow \quad \varphi(v,v) = 0.$$

Q.E.D.

Wir fassen formulieren hier die Aussage aus Proposition 11.19 neu:
Wenn char $\mathbb{K} \neq 2$, dann

$$\textit{alternierend} = \textit{antisymmetrisch} \qquad \text{und} \qquad \{\textit{antisymmetrich}\} \cap \{\textit{symmetrisch}\} = 0.$$

Wenn char $\mathbb{K} = 2$, dann

$$\textit{alternierend} \Rightarrow \textit{antisymmetrisch} \qquad \text{und} \qquad \textit{antisymmetrich} = \textit{symmetrisch}.$$

11.3 Orthogonalität und nicht ausgeartete Bilinearformen

In diesem Teil ist weiterhin V ein beliebiger endlich-dimensionaler $\mathbb{K}$-Vektorraum, und jede Bilinearform $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$ ist symmetrisch, antisymmetrisch oder alternierend. Wir werden das (manchmal) als "φ ist SAMOA" abkürzen[2].Der Grund dieser Einschränkung wird bald (nach Satz 11.21) klar sein. Wir fangen mit einer Definition an.

Definition 11.20. Sei $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine Bilinearform auf einem $\mathbb{K}$-Vektorraum V. Ein Vektor $v \in V$ ist **orthogonal** zu einem Vektor $w \in V$ bezüglich φ wenn

$$\varphi(v,w) = 0.$$

Satz 11.21. *Seien V ein $\mathbb{K}$-Vektorraum und $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$. Es gilt*

$$[\textit{Für alle } v, w \in V \textit{ gilt } v \perp w \Rightarrow w \perp v] \quad \Leftrightarrow \quad \varphi \textit{ ist } \text{SAMOA}.$$

Beweis-Skizze: $\boxed{\Leftarrow}$ Das ist offensichtlich für symmetrische und antisymmetrische Bilinearformen. Für alternierende gilt die Symmetrie der Orthogonalität weil alternierend antisymmetrisch impliziert.

$\boxed{\Rightarrow}$ Das ist nicht so einfach zu beweisen. Wir lassen das hier aus. Siehe [Rom08, Theorem 11.4] für den vollständigen Beweis. Q.E.D.

[2] **S**ymmetrisch, **A**nti-sy**M**metrisch **O**der **A**lternierend.

Weil wir die nur symmetrische Orthognoalität betrachten wollen, nehmen wir ab jetzt an, dass φ SAMOA ist. Wir schreiben dann für $\varphi(v,w)=0$:

$$v \perp w,$$

oder $v \perp^{\varphi} w$ wenn φ nicht klar aus dem Kontext ist. Für eine Teilmenge $X \subseteq V$ bezeichnen wir mit $X^{\perp}$ die Menge aller Vektoren die orthogonal auf X bezüglich φ sind:

$$X^{\perp} := \{v \in V \mid x \perp v,\ \forall\, x \in X\}.$$

Zwei Teilmengen $X, Y \subseteq V$ sind orthogonal wenn $x \perp y$ für alle $(x,y) \in X \times Y$.

Beispiele:

1. Für die kanonische Bilinearform auf $\mathbb{K}^2$ und $X = \operatorname{Span} e_1$ gilt

$$X^{\perp} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{K}^2 \mid \left\langle \begin{pmatrix} x \\ y \end{pmatrix} \mid \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \right\rangle = \lambda x = 0,\ \forall\, \lambda \in \mathbb{K} \right\} = \operatorname{Span} e_2.$$

   Weiterhin gilt

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \perp \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \Leftrightarrow \quad \operatorname{char} \mathbb{K} = 2.$$

2. Für die kanonische Bilinearform auf $\mathbb{K}^3$ und $X = \{e_1 + e_2, e_3 + e_1\}$ gilt

$$\begin{aligned} X^{\perp} &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 \ : \ \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 0 \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{K}^3 \ : \ x = -y, \ \ x = -z \right\} \\ &= \operatorname{Span}_{\mathbb{K}}\{e_1 - e_2 - e_3\}. \end{aligned}$$

Lemma 11.22. *Sei X eine Teilmenge eines $\mathbb{K}$-Vektorraumes und sei φ eine SAMOA Bilinearform auf V. Die Menge $X^{\perp}$ ist ein $\mathbb{K}$-Untervektorraum von V.*

Beweis-Skizze: Wir haben $0 \in X^{\perp}$, also $X^{\perp} \neq \emptyset$. Wenn $u, w \in X^{\perp}$ und $\lambda, \mu \in \mathbb{K}$, dann gilt für alle $x \in X$

$$\varphi(x, \lambda u + \mu w) = \lambda\varphi(x,u) + \mu\varphi(x,w) = 0 + 0 = 0,$$

also $\lambda u + \mu w \in X^{\perp}$. Q.E.D.

Definition 11.23. Eine SAMOA Bilinearform φ ist **nicht ausgeartet** (oder nicht entartet) wenn $V^{\perp} = 0$. Sonst, heißt die SAMOA Bilinearform **ausgeartet** (oder entartet).

go to Kap.8
[16]15.6.'22
return Kap.8

Der $\mathbb{K}$-Untervektorraum $V^{\perp}$ wir deswegen auch **Ausartungsraum** oder **Entartungsraum** von φ genannt. Für den nächsten Satz wiederholen wir die Definition von Ψ aus Satz 11.8:

$$\Psi(\varphi) : V \longrightarrow V^*, \qquad \Psi(\varphi)(v) := \varphi(v, \ \cdot\ ).$$

Lemma 11.24. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und φ eine SAMOA Bilinearform. Es gilt*

$$\operatorname{Ker}\Psi(\varphi) = V^{\perp}.$$

Beweis-Skizze:

$$v \in \operatorname{Ker}\Psi(\varphi) \Leftrightarrow \Psi(\varphi)(v) = 0 \in V^* \Leftrightarrow \varphi(v, v') = 0, \forall v' \in V \Leftrightarrow v \in V^{\perp} = 0.$$

Q.E.D.

Satz 11.25. *Seien V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $\varphi \in \mathrm{Bil}_{\mathbb{K}}^{\mathrm{SAMOA}}(V)$. Folgende Aussagen sind äquivalent.*

(a) *φ ist nicht ausgeartet.*

(b) *Die lineare Abbildung $\Psi(\varphi) : V \longrightarrow V^*$ ist injektiv.*

(c) *Die lineare Abbildung $\Psi(\varphi) : V \longrightarrow V^*$ ist ein Isomorphismus.*

(d) *Jede darstellende Matrix von φ ist invertierbar.*

(e) *Eine darstellende Matrix von φ ist invertierbar.*

Beweis. $\boxed{(a) \Leftrightarrow (b)}$ Das folgt aus Lemma 11.24.
$\boxed{(b) \Leftrightarrow (c)}$ Das folgt weil $\dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V < \infty$.
$\boxed{(c) \Rightarrow (d)}$ Aus Korollar 11.9 ist $\mathcal{M}^{\mathcal{B}}(\varphi) = \mathrm{M}^{\mathcal{B}}_{\mathcal{B}^*}(\Psi(\varphi))$ für jede Basis $\mathcal{B}$ von V, also invertierbar.
$\boxed{(d) \Rightarrow (e)}$ Trivial.
$\boxed{(e) \Rightarrow (a)}$ Sei $\mathcal{B} = v_1, \ldots, v_n$ die Basis bezüglich welcher $\mathcal{M}^{\mathcal{B}}(\varphi)$ invertierbar ist. Wir werden zeigen, dass $v \in V^{\perp} \Rightarrow v = 0$.
Sei also $v = \sum_{i=1}^{n} x_i v_i \in V^{\perp}$ beliebig. Das heißt insbesondere, dass

$$\varphi(v, v_i) = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \cdot \mathcal{M}^{\mathcal{B}}(\varphi) \cdot e_i = 0, \quad \forall\, i = 1, \ldots, n.$$

Daraus folgt

$$\begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \cdot \mathcal{M}^{\mathcal{B}}(\varphi) \cdot I_n = \begin{pmatrix} 0 & \ldots & 0 \end{pmatrix}. \tag{11.1}$$

Wir multiplizieren rechst mit $\left(\mathcal{M}^{\mathcal{B}}(\varphi)\right)^{-1}$ und bekommen

$$\begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} = \begin{pmatrix} 0 & \ldots & 0 \end{pmatrix} \cdot (\mathcal{M}^{\mathcal{B}}_{\varphi})^{-1} = \begin{pmatrix} 0 & \ldots & 0 \end{pmatrix},$$

also $v = 0$. □

[17]20.6.'22

Korollar 11.26. Wenn $\mathcal{B}$ eine Basis des n-dimensionalen $\mathbb{K}$-Vektorraum V ist und wenn φ eine SAMOA Bilinearform ist, dann gilt

$$v \in V^{\perp} \Leftrightarrow K_{\mathcal{B}}(v) \in \mathcal{L}(\mathcal{M}^{\mathcal{B}}(\varphi) \mid \mathbf{0}),$$

wobei $K_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n$ die Koordinatenabbildung bezüglich der Basis $\mathcal{B}$ ist.

Beweis-Skizze: Der Beweis folgt aus (11.1) wenn man die Matrizen transponiert. Q.E.D.

Der $\mathbb{K}$-Untervektorraum $V^\perp$ wird auch **Kern der bilinearen Abbildung** φ genannt. Wir werden von dem Rang einer Bilinearform sprechen, aber nicht als Dimension des Bildes, sonder als:

$$\operatorname{Rang}\varphi := \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} V^\perp.$$

Aus Korollar 11.26 folgt dann gleich

Korollar 11.27. Sei V eine endlichdimensionaler $\mathbb{K}$-Vektorraum mit Basis $\mathcal{B}$ und $\varphi \in \operatorname{Bil}_{\mathbb{K}}^{\text{SAMOA}}(V)$ eine symmetrische Bilinearform auf V. Es gilt

$$\operatorname{Rang}\varphi = \operatorname{Rang}\mathcal{M}^{\mathcal{B}}(\varphi).$$

Nach Korollar 11.9 gilt $\operatorname{Rang}\varphi = \dim_{\mathbb{K}} \operatorname{Bild}\Psi(\varphi)$. Das Bild von der Bilinearform φ selbst ist eher uninteresent. Das wäre 0, wenn φ die Nullabbildung ist, und $\mathbb{K}$ sonst.

In Kapitel 8 hatten wir für einen $\mathbb{K}$-Untervektorraum $U \subseteq V$ den Annullator als $U^0 := \{w \in V^* \mid w(u) = 0,\ \forall\ u \in U\} \subseteq V^*$ definiert. In Satz 8.13 wurde dann bewiesen, dass

$$\dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^0 = \dim_{\mathbb{K}} V. \tag{11.2}$$

Durch die Korrespondenz $\Psi(\varphi) : V \longrightarrow V^*$ ist es einfach zu sehen, dass

$$(\Psi(\varphi))(U^\perp) \subseteq U^0.$$

Wenn die Bilinearform nicht ausgeartet ist, dann ist $\Psi(\varphi)$ bijektiv und die andere Inklusion gilt auch. Wir können dann alle Sätze aus Kapitel 8 für $U^\perp$ übernehmen. Insbesondere, gilt

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^\perp.$$

Für eine beliebige (i.e. möglicher weise ausgeartete) SAMOA Bilinearform auf V gilt folgender Satz.

Satz 11.28. *Sei U ein $\mathbb{K}$-Untervektorraum des endlichdimensionalen $\mathbb{K}$-Vektorraumes V und sei φ eine SAMOA Bilinearform. Es gilt*

$$\dim_{\mathbb{K}} U + \dim_{\mathbb{K}} U^\perp = \dim_{\mathbb{K}} V + \dim_{\mathbb{K}}(U \cap V^\perp).$$

Beweis-Skizze: Die Strategie ist den Dimensionssatz für die Einschränkung $\alpha = \Psi(\varphi)|_U : U \longrightarrow V^*$ anzuwenden. Es gilt

$$\operatorname{Ker}\alpha = U \cap (\operatorname{Ker}\Psi(\varphi)) = U \cap V^\perp.$$

Um die Dimension von $\operatorname{Bild}\alpha$ zu bestimmen, betrachten wir den Annullator von $\operatorname{Bild}\alpha \subseteq V^*$. Diser ist ein $\mathbb{K}$-Untervektorraum von $(V^*)^* = V$:

$$\begin{aligned}(\operatorname{Bild}\alpha)^0 &= \{v \in V \mid f(v) = 0 \text{ für alle } f \in \operatorname{Bild}\alpha\}\\ &= \{v \in V \mid \varphi(u,v) = 0 \text{ für alle } u \in U\}\\ &= U^\perp.\end{aligned}$$

Aus (11.2) folgt

$$\dim_{\mathbb{K}} \operatorname{Bild}\alpha = \dim_{\mathbb{K}} V - \dim_K (\operatorname{Bild}\alpha)^0 = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} U^\perp.$$

Aus dem Dimensionssatz für α haben wir $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \operatorname{Ker} \alpha + \dim_{\mathbb{K}} \operatorname{Bild} \alpha$, es folgt also

$$\dim_{\mathbb{K}} U = \dim_{\mathbb{K}}(U \cap V^{\perp}) + \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} U^{\perp},$$

und die Aussage folgt. Q.E.D.

Korollar 11.29. Sei $U \subseteq V$ ein $\mathbb{K}$-Untervektorraum des endlichdimensionalen $\mathbb{K}$-Vektorraumes V und sei $\varphi \in \operatorname{Bil}_{\mathbb{K}} V$ SAMOA. Folgende Aussagen sind äquivalent:

(a) $\varphi|_{U \times U}$ ist nicht ausgeartet.

(b) $V = U \oplus U^{\perp}$.

Beweis-Skizze: Übung Q.E.D.

Korollar 11.30. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $\varphi \in \operatorname{Bil}^{\mathrm{sym}}_{\mathbb{K}}(V)$. Die induzierte Bilinearform $\overline{\varphi} : V/V^{\perp} \times V/V^{\perp} \longrightarrow \mathbb{K}$ ist symmetrisch und nicht ausgeartet.

Beweis-Skizze: Übung Q.E.D.

11.4 Orthogonalbasen und der erste Trägheitssatz von Sylvester

Definition 11.31. Sei V ein $\mathbb{K}$-Vektorraum und φ eine SAMOA Bilinearform. Eine **Orthogonalbasis** von V bezüglich φ ist eine Basis $\mathcal{B}$ von V mit der Eigenschaft, dass

$$\varphi(v_i, v_j) = 0, \quad \forall\, v_i, v_j \in \mathcal{B} \ \text{ mit } v_i \neq v_j.$$

Bemerkung 11.32. Eine Basis $\mathcal{B}$ ist eine Orthogonalbasis bezüglich der SAMOABilinearform φ genau dann, wenn die darstellende Matrix $\mathcal{M}^{\mathcal{B}}(\varphi)$ eine Diagonalmatrix ist. Das heißt, dass wenn φ alternierend ist, dann gibt es eine Orthogonalbasis nur wenn $\varphi = 0$. Deswegen werden wir über Orthogonalbasen nur für *symmetrische* Bilinearformen sprechen.

Für den übrigen Teil dieses Abschnittes werden wir nur symmetrische Bilinearformen betrachten.

!Vorsicht! Ortho*gonal*basen sind nicht mit Ortho*normal*basen zu verwechseln (cf. Definition 12.13). Bei orthogonalen Basen ist für $\varphi(v_i, v_i)$ alles erlaubt, inklusive Null.

Beispiele:

1. Die Standardbasis von $\mathbb{K}^n$ ist eine Orthogonalbasis für die Standard Bilinearform auf $\mathbb{K}^n$.
2. Für das Standard Skalarprodukt $\varphi = \langle \cdot\,, \cdot \rangle \in \operatorname{Bil}^{\mathrm{sym}}_{\mathbb{K}}(\mathbb{R}^2)$, also für

$$\langle (x_1, x_2) \mid (y_1, y_2) \rangle = x_1 y_1 + x_2 y_2,$$

ist $\{(1,1), (-1,1)\}$ eine Orthogonalbasis, aber $\{(2,1), (-2,1)\}$ nicht.

Lemma 11.33. *Seien V ein endlichdimensionaler$\mathbb{K}$-Vektorraum, $\varphi \in \operatorname{Bil}^{\mathrm{sym}}_{\mathbb{K}}(V)$, und $v \in V$ mit $\varphi(v,v) \neq 0$. Dann gilt*

$$V = \operatorname{Span} v \oplus \{v\}^{\perp}.$$

Beweis-Skizze: Zu zeigen ist also, dass $\operatorname{Span} v \cap \{v\}^\perp = 0$ und $\operatorname{Span} v + \{v\}^\perp = V$. Für die erste Gleichheit, sei $w \in \operatorname{Span} v$, also $w = \lambda v$ für irgendwelches $\lambda \in \mathbb{K}$. Es gilt

$$\lambda v \in \{v\}^\perp \Leftrightarrow \varphi(\lambda v, v) = 0 \Leftrightarrow \lambda \cdot \varphi(v, v) = 0 \Leftrightarrow \varphi(v, v) \neq 0 \lambda = 0.$$

Es gilt also $\dim_{\mathbb{K}}(\operatorname{Span} v + \{v\}^\perp) = \dim_{\mathbb{K}} \operatorname{Span} v + \dim_{\mathbb{K}} \{v\}^\perp$.
Wenn V endlichdimensional ist, dann folgt aus Satz 11.28, dass

$$\dim_{\mathbb{K}}(\operatorname{Span} v + \{v\}^\perp) = \dim_{\mathbb{K}} \operatorname{Span} v + \dim_{\mathbb{K}} \{v\}^\perp \geq \dim_{\mathbb{K}} V,$$

also $\operatorname{Span} v + \{v\}^\perp = V$ folgt.

Wenn V nicht endlichdimensional ist[a], dann zeigen wir direkt, dass jeder Vektor $w \in V$ als $\lambda v + v'$ geschrieben werden kann, mit $v' \perp v$. Wir definieren

$$\lambda := \frac{\varphi(w, v)}{\varphi(v, v)} \in \mathbb{K} \quad \text{und} \quad v' := w - \lambda v.$$

Offensichtlich gilt $w = \lambda v + v'$. Wir müssen nur noch zeigen, dass $v' \in \{v\}^\perp$. Das geht so:

$$\varphi(v', v) = \varphi(w - \lambda v, v) = \varphi(w, v) - \lambda \varphi(v, v) = \varphi(w, v) - \frac{\varphi(w, v)}{\varphi(v, v)} \varphi(v, v) = 0.$$

Q.E.D.

[a] Das Lemma gilt auch ohne der Voraussetzung, dass $\dim_{\mathbb{K}} V < \infty$, deswegen fügen wir diesen Teil hinzu.

Satz 11.34 (Existenz von Orthogonalbasen)**.** *Sei V ein endlichdimensionaler Vektorraum über den Körper $\mathbb{K}$ mit* $\operatorname{char} \mathbb{K} \neq 2$*, und sei $\varphi \in \operatorname{Bil}^{\mathrm{sym}}_{\mathbb{K}}(V)$ eine symmetrische Bilinearform auf V. Dann existiert eine Orthogonalbasis von V bezüglich φ.*

In anderen Worten, ist jede symmetrische Matrix über einen Körper von Charakteristik verschieden von zwei kongruent zu einer Diagonalmatrix.

Beweis-Skizze: Wir beweisen den Satz durch vollständige Induktion nach $n = \dim_{\mathbb{K}} V$.
$\boxed{n = 1}$ Jede 1x1 Matrix ist diagonal, also jede Basis ist eine Orthogonalbasis.
$\boxed{n \Rightarrow n + 1}$ Die induktive Voraussetzung ist, dass jeder n-dimensionale $\mathbb{K}$-Vektorraum eine Orthogonalbasis besitzt. Sei $\dim_{\mathbb{K}} V = n + 1$. Wenn φ die Nullabbildung ist, dann ist wieder jede Basis orthogonal. Wir nehmen also an, dass $\varphi \neq 0$ und wollen Lemma 11.33 anwenden. Dafür müssen wir zu erst zeigen, dass es $v \in V$ gibt mit $\varphi(v, v) \neq 0$. Aus $\varphi \neq 0$ folgt es existieren v_1, v_2 mit $\varphi(v_1, v_2) \neq 0$. Es gilt, weil φ symmetrisch ist, dass

$$\varphi(v_1 + v_2, v_1 + v_2) - \varphi(v_1, v_1) - \varphi(v_2, v_2) = 2\varphi(v_1, v_2).$$

Weil $2 \neq 0$ und $\varphi(v_1, v_2) \neq 0$ folgt, dass wenigstens ein Vektor der Menge $\{v_1, v_2, v_1 + v_2\}$ die gesuchte Eigenschaft $\varphi(v, v) \neq 0$ hat. Sei also $v \in V$ mit $\varphi(v, v) \neq 0$. Aus Lemma 11.33 folgt

$$V = \operatorname{Span} v \oplus \{v\}^\perp.$$

Aus der induktiven Voraussetzung, weil $\dim_{\mathbb{K}}\{v\}^{\perp} = n$, hat $\{v\}^{\perp}$ eine Orthogonalbasis $\mathcal{B}$. Weil alle Vektoren aus $\{v\}^{\perp}$ per Definition orthogonal zu v sind, ist $\{v\} \cup \mathcal{B}$ die gesuchte Basis. Q.E.D.

Korollar 11.35. Wenn $\operatorname{char} \mathbb{K} \neq 2$, dann ist jede symmetrische Matrix in $\operatorname{Mat}_n^{\mathrm{sym}}(\mathbb{K})$ kongruent zu einer Diagonalmatrix.

Bemerkung 11.36. In Charakteristik 2 ist das nicht mehr so, und die Klassifikation der Kongruenzklassen ist nicht so einfach. Eine gute Übung ist ein kleinstmögliches Beispiel von symmetrischer Matrix, über $\mathbb{F}_2$ zu finden, die nicht kongruent zu einer Diagonalmatrix ist.

Korollar 11.37. Wenn $\mathbb{K} = \mathbb{C}$ und V ein endlichdimensionaler $\mathbb{C}$-Vektorraum ist, dann existiert für jede symmetrische Bilinearform $\varphi \in \operatorname{Bil}_{\mathbb{C}}^{\mathrm{sym}}(V)$ eine Basis $\mathcal{B}$ von V, sodass

$$\mathcal{M}^{\mathcal{B}}(\varphi) = \begin{pmatrix} \boxed{I_r} & \\ & \mathrm{O} \end{pmatrix}.$$

Beweis-Skizze: Nach Satz 11.34 existiert eine Orthogonalbasis $\mathcal{B}' = \{v_1, \ldots, v_n\}$, sodass $\mathcal{M}_{\varphi}^{\mathcal{B}'} = \operatorname{diag}(d_1, \ldots, d_n)$. Nach eventuelles Umordnen, können wir annehmen, dass es $r \in \{0, \ldots, n\}$ existiert, sodass $d_i = 0 \Leftrightarrow i > r$. Weil der Körper der komplexen Zahlen algebraisch abgeschlossen ist, haben wir, dass für jeden Index $i = 1, \ldots, r$ existiert eine komplexe Zahl $c_i \in \mathbb{C}$ mit $c_i^2 = d_i$ (also $c_i = \sqrt{d_i}$). Wir setzen dann

$$\mathcal{B} = \{{}^1\!/_{c_1} \cdot v_1, \ldots, {}^1\!/_{c_r} \cdot v_r, v_{r+1}, \ldots, v_n\}.$$

Q.E.D.

Korollar 11.38 (Trägheitssatz I von Sylvester[3])**.** Wenn $\mathbb{K} = \mathbb{R}$ und V ein endlichdimensionaler $\mathbb{R}$-Vektorraum ist, dann existiert für jede symmetrische Bilinearform $\varphi \in \operatorname{Bil}_{\mathbb{R}}^{\mathrm{sym}}(V)$ eine Basis $\mathcal{B}$ von V, und $r, s \in \mathbb{N}$, sodass

$$\mathcal{M}^{\mathcal{B}}(\varphi) = \begin{pmatrix} \boxed{I_r} & & \\ & \boxed{-I_s} & \\ & & \mathrm{O} \end{pmatrix}.$$

Beweis-Skizze: Nach Satz 11.34 existiert eine Orthogonalbasis $\mathcal{B}' = \{v_1, \ldots, v_n\}$, sodass $\mathcal{M}_{\varphi}^{\mathcal{B}'} = \operatorname{diag}(d_1, \ldots, d_n)$. Nach eventuelles Umordnen, können wir annehmen, dass es $r, s \in \{0, \ldots, n\}$ existieren, sodass

$$d_i \begin{cases} = 0 & \Leftrightarrow r + s < i \\ < 0 & \Leftrightarrow r < i \leq s \\ > 0 & \Leftrightarrow i \leq r \end{cases}$$

[3] Eine Erklärung warum es "Trägheitssatz" heißt finden Sie auf `math.stackexchange`.

Weil positive reelle Zahlen immer eine Wurzel haben, existiert für jeden Index $i = 1, \dots, r+s$ existiert eine reelle Zahl $c_i \in \mathbb{R}$ mit $c_i^2 = |d_i| \neq 0$. Wir setzen dann

$$\mathcal{B} = \{1/c_1 \cdot v_1, \dots, 1/c_{r+s} \cdot v_{r+s}, v_{r+s+1}, \dots, v_n\}.$$

Q.E.D.

[18]22.7.'22

11.5 Das Verfahren zu der Diagonalisierung der symmetrischen Bilinearformen

In diesem Teil nehmen wir an, dass char $\mathbb{K} \neq 2$.

Aus Satz 11.34 existieren für jede symmetrische Matrix $A \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K})$ eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ und $d_1, \dots, d_n \in \mathbb{K}$ sodass

$$S^\top \cdot A \cdot S = \mathrm{diag}(d_1, \dots, d_n).$$

Wir beschreiben hier ein Verfahren, das wir den **symmetrischen Gaußschen Algorithmus** nennen, das sowohl S als auch die d_is bestimmt. Die Idee ist sehr einfach: man geht wie im Gaußschen Algorithmus zur Bestimmung einer Zeilen-Stufen-Form vor, aber zusätzlich, *nach jeder* Zeilenumformung, wird dieselbe Umformung auch auf der Menge der Spalten durchgeführt. Zum Beispiel: wenn wir zu der 3-ten Zeile 2 Mal die 1-te Zeile addieren, dann gleich danach addieren wir zu der 3-ten Spalte (der neuen Matrix) 2 Mal die 1-te Spalte (der neuen Matrix). Und so weiter, bis wir eine Diagonalform erreichen.

Die Zeilenumformung die hier beschrieben wurde entspricht der links-Multiplikation mit der Zeilenumformungsmatrix $U_{\mathbf{3}\to\mathbf{3}+(2)\cdot\mathbf{1}}$:

$$A \longmapsto U_{\mathbf{3}\to\mathbf{3}+(2)\cdot\mathbf{1}} \cdot A.$$

Es ist einfach und direkt zu überprüfen, dass die gespiegelte Spaltenumformung der neuen Matrix durch rechts-Multiplikation mit der transponierten Zeilenumformungsmatrix erhalten ist:

$$\left(U_{\mathbf{3}\to\mathbf{3}+(2)\cdot\mathbf{1}} \cdot A\right) \mapsto \left(U_{\mathbf{3}\to\mathbf{3}+(2)\cdot\mathbf{1}} \cdot A\right) \cdot \left(U_{\mathbf{3}\to\mathbf{3}+(2)\cdot\mathbf{1}}\right)^\top.$$

Also jeder Schritt: Zeilenumformung + gleiche Spaltenumformung gibt uns:

$$A \mapsto U^\top \cdot A \cdot U$$

wobei $U^\top$ eine Zeilenumformungsmatrix ist. Das heißt, dass die neue Matrix kongruent zu der alten Matrix ist. Die gesuchte Matrix S finden nach endlich-viele Schritte als $U_1^\top, \dots, U_r^\top$:

$$A \longmapsto U_r^\top \cdot \dots \cdot U_1^\top \cdot A \cdot U_1 \cdot \dots \cdot U_r = (U_1 \cdot \dots \cdot U_r)^\top \cdot A \cdot (U_1 \cdot \dots \cdot U_r).$$

Also, damit wir S finden, müssen wir *nur* die Spaltenumformungen, in derselben Reihenfolge wie auf A, auf der Identitätsmatrix durchführen:

$$I_n \cdot U_1 \cdot \dots \cdot U_r = S.$$

Beispiele dazu kann man in den Notizen zu der Zentralübung 11 und in [Fis09, S.359-360] finden.

11.6 Positive und negative Teile reeller Bilinearformen

Für den nächsten Abschnitt werden wir nur $\mathbb{R}$-Vektorräume betrachten. Die für uns wichtige Eigenschaften von $\mathbb{R}$ sind, dass es eine totale Ordnung $\geq$ auf $\mathbb{R}$ gibt die verträglich[4] mit der Körper Struktur ist, und, dass es eine Betragabbildung $|\cdot| : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$ gibt. Die totale Ordnung teilt $\mathbb{R} \setminus \{0\}$ in positive (> 0) und negative (< 0) Zahlen.

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

Definition 11.39. Eine symmetrische Bilinearform $\varphi \in \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{R}}(V)$ ist **positiv-semi-definit** (bzw. **positiv-definit**) wenn

$$\varphi(v,v) \geq 0 \quad (\text{bzw. } \varphi(v,v) > 0) \quad \forall\, v \in V \setminus \{0\}.$$

Eine symmetrische Bilinearform $\varphi \in \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{R}}(V)$ ist **negativ-(semi-)definit** falls $-\varphi$ positiv-(semi-)definit ist. Eine symmetrische Matrix $A \in \mathrm{Mat}^{\mathrm{sym}}_n(\mathbb{R})$ ist positiv/negativ-(semi-)definit wenn die zugeordnete $\varphi_A \in \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^n)$ positiv/negativ-(semi-)definit ist.

Bemerkung 11.40. Seien V ein endlichdimensionaler $\mathbb{R}$-Vektorraum von Dimension n, $\mathcal{B}$ eine Basis von V, und $\varphi \in \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{R}}(V)$ eine symmetrische Bilinearform.

(a) φ ist positiv-(semi-)definit genau dann, wenn $\mathcal{M}^{\mathcal{B}}_{\varphi} \in \mathrm{Mat}^{\mathrm{sym}}_n(\mathbb{R})$ positiv-(semi-)definit ist.

(b) Seien $\mathcal{B}$ die Basis und $r, s \in \{1, \ldots, n\}$ mit $r + s \leq n$ die ganze Zahlen aus Korollar 11.38 für φ. Es gilt also

$$\mathcal{M}^{\mathcal{B}}_{\varphi} = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & \mathrm{O} \end{pmatrix}.$$

Es gelten

(i) φ ist positiv-semi-definit $\iff s = 0$.

(ii) φ ist positiv-definit $\iff s = 0$ und $r = n$.

(iii) φ ist negativ-semi-definit $\iff r = 0$.

(iv) φ ist negativ-definit $\iff r = 0$ und $s = n$.

Insbesondere, ist eine positiv/negativ-definite Bilinearform nicht ausgeartet.

(c) Eine symmetrische Matrix $A \in \mathrm{Mat}^{\mathrm{sym}}_n(\mathbb{R})$ ist positiv-definit genau dann, wenn es kongruent zur Einheitsmatrix I_n ist.

Bemerkung 11.41. Wenn eine Basis $v_1, \ldots, v_n$ existiert mit $\varphi(v_i, v_i) > 0$ für alle i, heißt es nicht, dass φ positiv-definit ist. Zum Beispiel: $q_\varphi(x,y) = x^2 - y^2$ ist positiv für $v_1 = (1,0)$ und $v_2 = (2,1)$ und diese sind eine Basis. Aber φ ist nicht positiv-definit.

[4] Das heißt:

$+$ Wenn $z_1 > z_2$, dann $z_1 + z_3 > z_2 + z_3 \quad \forall z_3 \in \mathbb{C}$

$\cdot$ Wenn $z_1 > z_2$ und $z_3 > 0$, dann $z_1 z_3 > z_2 z_3$.

Bemerkung 11.42. Wenn $A = (a_{ij})$ positiv-definit ist, dann gilt

$$a_{ii} > 0 \quad \text{und} \quad \det A > 0.$$

Die erste Ungleichung erhalten wir wenn wir die ensprechende Bilinearform $\varphi_A A$ auf der kanonischen Basis anwenden. Die zweite, weil positiv-definit äquivalent zu $\exists\, S \in \mathrm{GL}_n(\mathbb{R})$ mit $S^\top AS = I_n$ ist. Also

$$\det A \cdot (\det S)^2 = 1.$$

Weil $(\det S)^2 > 0$, folgt auch $\det A > 0$.

Satz 11.43 (Kriterium für positiv-definit)**.** *Sei $A = (a_{ij})_{i,j=1,\dots,n} \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{R})$ eine symmetrische Matrix. Für jeden $k = 1, \dots, n$ definieren wir die Untermatrix*

$$A_k = (a_{ij})_{i,j=1,\dots,k} \in \mathrm{Mat}_k^{\mathrm{sym}}(\mathbb{R}).$$

Es gilt dann

$$A \text{ ist positiv-definit} \iff \det A_k > 0 \ \ \forall\ k = 1, \dots, n.$$

Beweis-Skizze:
[Fis09, S.385] Q.E.D.

Satz 11.44 (Trägheitssatz II von Sylvester)**.** *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $\varphi \in \mathrm{Bil}_{\mathbb{R}}^{\mathrm{sym}}(V)$. Dann gibt es eine Zerlegung von V als direkte Summe von $\mathbb{R}$-Untervektorräume*

$$V = V_+ \oplus V_- \oplus V_0$$

mit folgenden Eigenschaften

(a) *Die $\mathbb{R}$-Untervektorräume V_+, V_- und V_0 sind paarweise orthogonal.*

(b) *$\varphi|_{V_+ \times V_+}$ ist positiv-definit.*

(c) *$\varphi|_{V_- \times V_-}$ ist negativ-definit.*

(d) *$\varphi|_{V_0 \times V_0} = 0$.*

Weiterhin, der $\mathbb{R}$-Untervektorraum V_0 und die Dimensionen $\dim_{\mathbb{R}} V_+$ und $\dim_{\mathbb{R}} V_-$ sind eindeutig bestimmt.

Beweis-Skizze: Aus den Trägheitssatz I von Sylvester (Korollar 11.38) folgt die Existenz einer Basis $\mathcal{B} = \{v_1, \dots, v_n\}$, sodass

$$\mathcal{M}_\varphi^{\mathcal{B}} = \begin{pmatrix} \boxed{I_r} & & \\ & \boxed{-I_s} & \\ & & \mathrm{O} \end{pmatrix}.$$

Man definiert dann

$$\begin{aligned} V_+ &= \operatorname{Span} v_1, \dots, v_r \\ V_- &= \operatorname{Span} v_{r+1}, \dots, v_{r+s} \\ V_0 &= \operatorname{Span} v_{r+s+1}, \dots, v_n. \end{aligned}$$

Die Eindeutigkeit von V_0 folgt, weil $V_0 = V^\perp$ der Ausartungsraum von φ ist. Wenn $V = V'_+ \oplus V'_- \oplus V_0$, dann, weil $V'_+ \cap (V_- \oplus V_0) = 0$, folgt $\dim_{\mathbb{R}} V'_+ \leq \dim_{\mathbb{R}} V_+$. Durch Symmetrie, folgt dann auch die andere Ungleichung. Aus der direkten Summe folgt dann auch $\dim_{\mathbb{R}} V_- = \dim_{\mathbb{R}} V'_-$. Q.E.D.

[19]27.6.'22

Definition 11.45. x Die **Signatur** einer symmetrischen Bilinearform φ auf einem endlichdimensionalen $\mathbb{R}$-Vektorraum V ist

$$\operatorname{sgn}(\varphi) := \dim_{\mathbb{K}} V_+ - \dim_{\mathbb{K}} V_-,$$

wobei $V = V_+ \oplus V_- \oplus V_0$ eine Zerlegung wie im Trägheitssatz II von Sylvester ist. Die Signatur einer symmetrischen Matrix $A \in \operatorname{Mat}_n^{\mathrm{sym}}(\mathbb{R})$ ist die Signatur der zugehörigen Bilinearform auf $\mathbb{R}^n$.

Bemerkung 11.46. Seien $A, B \in \operatorname{Mat}_n^{\mathrm{sym}}(\mathbb{R})$. Es gilt

$$A \approx B \iff \operatorname{Rang}(A) = \operatorname{Rang}(B) \text{ und } \operatorname{sgn}(A) = \operatorname{sgn}(B).$$

Wenn der Rang maximal ist, also wenn die Bilinearform nicht ausgeartet ist, dann wird manchmal[5] die Signatur auch durch einem Vektor von $+$ und $-$ bezeichnet: Zum Beispiel $(+,+,+,-)$.

Beispiel 11.47 (Minkowski[6] Raum)**.** Wir werden in Kapitel **??** Skalarprodukte als symmetrische (oder hermitesche) positiv-definite nicht-ausgeartete Bilinearformen definieren. Diese Bilinearformen haben also Signatur gleich mit der Dimension des Raumes, und erlauben uns eine Metrik zu definieren. Es gibt aber auch Anwendungen von nicht-ausgearteten Bilinearformen mit Signatur kleiner als die Dimension: Der **Minkowski Raum**. Dieser ist der 4-dimensionale reelle Vektorraum $\mathbb{R}^4$ zusammen mit der nicht ausgearteten Bilinearform definiert durch

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Dieser Raum spielt eine zentrale Rolle in der allgemeinen Relativitätstheorie. Dieser entspricht dem Raum-Zeit-Kontinuum, mit e_1, e_2, e_3 als Raum-Koordinaten und e_4 entspricht der Zeit-Koordinate (es sollte eigentlich ct entsprechen, wobei c die Lichtgeschwindigkeit und t die Zeit ist). In der Teilchenphysik wird eine (physisch-äquivalente) Bilinearform mit Signatur -2 verwendet.

[5] Insbesondere in der Physik.
[6] Nach Hermann Minkowski (1864-1909), ein russisch-deutscher Mathematiker und Physiker.

Kapitel 12

Euklidische und unitäre Vektorräume

In diesem Kapitel werden wir nur Vektorräume über dem Körper $\mathbb{R}$ der reellen Zahlen oder dem Körper $\mathbb{C}$ der komplexen Zahlen betrachten. In beiden Fällen wollen wir ein Instrument einführen, das uns erlaubt Längen und Winkeln zu messen. Dieses Instrument heißt Skalarprodukt. Im reellen Fall ist es eine symmetrische, bilineare, und positiv-definite Form. Im komplexen Fall erlauben Symmetrie und Bilinearität keinen guten Begriff von Positivität einzuführen. Deswegen werden diese zwei Voraussetzungen zu *sesquilinear* und *hermitesch* "geschwächt". Wir fangen mit dem reellen Fall an.

12.1 Skalarprodukte

12.1.1 Auf $\mathbb{R}$-Vektorräume

Falls ein $\mathbb{K}$ in diesem Abschnitt vorkommt, dann ist es aus Versehen. Wir betrachten hier nur Vektorräume V über $\mathbb{R}$.

Definition 12.1. Ein **Skalarprodukt** auf dem $\mathbb{R}$-Vektorraum V ist eine Abbildung

$$\begin{aligned} \langle \cdot\,, \cdot \rangle : V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \langle v, w \rangle \end{aligned}$$

die bilinear, symmetrisch und positiv-definit ist.

Wir geben auch eine äquivalente Charakterisierung an, die den Zusammenhang mit Skalarprodukte auf $\mathbb{C}$-Vektorräume klarer macht.

Bemerkung 12.2. Eine Abbildung $\langle \cdot\,, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ ist ein Skalarprodukt genau dann, wenn folgende drei Axiome erfüllt sind:

(SP 1) Für alle $v_1, v_2, w \in V$ und alle $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt $\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \lambda_1 \langle v_1, w \rangle + \lambda_2 \langle v_2, w \rangle$.

(SP 2) Für alle $v, w \in V$ gilt $\langle v, w \rangle = \langle w, v \rangle$.

(SP 3) Für alle $0 \neq v \in V$ gilt $\langle v, v \rangle > 0$.

Die ersten zwei *zusammen* sind äquivalent zu bilinear und symmetrisch, während das dritte ist die Definition von positiv-definit.

Definition 12.3. Ein **euklidischer Vektorraum** ist ein Paar $(V, \langle \cdot\,, \cdot \rangle)$ wobei V ein $\mathbb{R}$-Vektorraum und $\langle \cdot\,, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ ein Skalarprodukt ist.

Wir werden oft einfach sagen, dass V ein euklidischer Vektorraum ist, ohne das Skalarprodukt explizit zu schreiben. Man kann immer annehmen, dass es mit $\langle \cdot\,, \cdot \rangle$ bezeichnet wird.

Beispiele:

1. $V = \mathbb{R}^n$, für $n \geqslant 1$, wobei wenn $v = (x_1, \ldots, x_n)$ und $w = (y_1, \ldots, y_n)$, dann ist
$$\langle v, w \rangle := x_1 y_1 + \cdots + x_n y_n.$$
Dieser Raum wird auch der **euklidische Standardraum** von Dimension n genannt, und das oben definierte Skalarprodukt ist das **Standardskalarprodukt** (oder **kanonische Skalarprodukt**, oder **euklidisches Skalarprodukt**).

2. Jede symmetrische, positiv-definite Matrix $A \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{R})$ definiert ein Skalarprodukt auf $\mathbb{R}^n$ durch
$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \cdot A \cdot \mathbf{y}.$$

3. Hier ist ein konkretes Beispiel in $\mathbb{R}^3$:
$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix},$$
Die Bilinearität folgt weil aus der Matrixmultiplikation, die Symmetrie aus der Symmetrie von A.Die positiv-definit Eigenschaft folgt nach Satz 11.43 weil die drei Determinanten ($\det A_1 = 2,\ \det A_2 = 3,\ \det A_3 = 4$) positiv sind. Man könnte diese Eigenschaft auch direkt überprüfen:
$$\begin{aligned} (x_1, x_2, x_3) \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 2x_1 - x_2, & -x_1 + 2x_2 - x_3, & -x_2 + 2x_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= 2x_1^2 - x_1x_2 - x_1x_2 + 2x_2^2 - x_2x_3 - x_2x_3 + 2x_3^2 \\ &= x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2. \end{aligned}$$
Also $\langle (x_1, x_2, x_3), (x_1, x_2, x_3) \rangle \geqslant 0$ für alle $\mathbf{x} \in \mathbb{R}^3$ und es ist Null nur für $(0, 0, 0)$.

4. Seien $a, b \in \mathbb{R}$ mit $a < b$, und sei $V = \{f : [a, b] \longrightarrow \mathbb{R} \ : \ f \text{ ist stetig}\}$. Man definiert das Skalarprodukt:
$$\langle f, g \rangle := \int_a^b f(t)g(t)dt. \quad \forall\ f, g \in V.$$

Bemerkung 12.4. Wenn $U \subseteq V$ ein $\mathbb{R}$-Untervektorraum eines euklidischen Vektorraumes ist, dann ist U selbst mit der Einschränkung $\langle \cdot\,, \cdot \rangle\,|_{U \times U}$ ein euklidischer Vektorraum.

12.1.2 Auf $\mathbb{C}$-Vektorräume

In diesem Abschnitt gilt $\mathbb{K} = \mathbb{C} = \{a + ib \ : \ a, b \in \mathbb{R} \text{ und } i^2 = -1\}$, der Körper der komplexen Zahlen. Es gibt auch in diesem Fall, wie im reellen Fall, einen Betrag $|\cdot| : \mathbb{C} \longrightarrow \mathbb{R}_{\geqslant 0}$ gegeben durch

$$|a + ib| = \sqrt{a^2 + b^2}.$$

Was aber nicht mehr gilt, ist das $|z| \in \{\pm z\}$. Nicht nur das, es kann keine totale Ordnung kompatibel[1] mit der Körperstruktur geben. Das heißt, man kann die komplexen Zahlen nicht vernünftig in positiv oder negativ teilen. Für $\mathbb{C}$-Bilinearformen hat also "positiv definit" keinen Sinn. Was funktionieren könnte ist eine Abbildung $\varphi : V \times V \longrightarrow \mathbb{C}$ die die Diagonale $\{(v, v) \ : \ v \in V\}$ auf $\mathbb{R}$ abbildet. Dann würde "positiv-definit" Sinn haben. Wenn aber φ bilinear wäre, dann hätten wir

$$\varphi(iv, iv) = (i)^2\varphi(v, v) = -\varphi(v, v).$$

Also Bilinearformen können Positivität auf der Diagonale nicht liefern. Wir werden ein Skalarprodukt auf $\mathbb{C}$-Vektorräume nicht mit Hilfe von Bilinearformen, sonder mit so-genannte *Sesquilinearformen*, das heißt Formen die "anderthalb"[2] linear sind, definieren. Dazu verwenden wir eine Eigenschaft von $\mathbb{C}$ die $\mathbb{R}$ nicht hat[3] anwenden: es gibt einen Körper-Automorphismus von $\mathbb{C}$ der nicht die Identität ist. Dieser ist die Konjugation $\overline{\cdot} : \mathbb{C} \longrightarrow C$, gegeben für alle $z = a + ib \in \mathbb{C}$ durch

$$z = a + ib \longmapsto \overline{z} := a - ib.$$

Es gilt also für alle $z \in \mathbb{C}$:

$$z\overline{z} = |z|^2 \in \mathbb{R}, \qquad \overline{\overline{z}} = z, \qquad \overline{z} = z \Leftrightarrow z \in \mathbb{R}.$$

Definition 12.5. Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $h : V \times V \longrightarrow \mathbb{C}$ heißt **Sesquilinearform** auf V wenn

(SQ1) h ist $\mathbb{C}$-linear im zweiten Argument, also

$$h(v, \lambda w_1 + \mu w_2) = \lambda \cdot h(v, w_1) + \mu \cdot h(v, w_2), \quad \forall\ \lambda, \mu \in \mathbb{C} \text{ und } \forall\ v, w_1, w_2 \in V.$$

(SQ2) h ist $\mathbb{C}$-*semilinear* (oder *konjugiert-linear*) im ersten Argument, also

$$h(\lambda v_1 + \mu v_2, w) = \overline{\lambda} \cdot h(v_1, w) + \overline{\mu} \cdot h(v_2, w), \quad \forall\ \lambda, \mu \in \mathbb{C} \text{ und } \forall\ v_1, v_2, w \in V.$$

Linearität in einem Argument zusammen mit Symmetrie würde automatisch zur Bilinearität führen. Das wollen wir nicht. Deswegen werden wir statt "symmetrische Sesquilinearformen", was nur die Nullabbildung sein könnte, über *Hermitesche*[4] Formen sprechen.

[1] Das heißt wir folgende Kompatibilität mit $+$ und $\cdot$:
$\boxed{+}$ Wenn $z_1 > z_2$, dann $z_1 + z_3 > z_2 + z_3 \ \ \forall z_3 \in \mathbb{C}$
$\boxed{\cdot}$ Wenn $z_1 > z_2$ und $z_3 > 0$, dann $z_1 z_3 > z_2 z_3$.
Daraus folgt für alle z, dass $z > 0 \Leftrightarrow -z < 0$ und $z^2 > 0$. Gäbe es eine solche Ordnung auf $\mathbb{C}$, dann hätten wir $-1 = (i^2) > 0$, also $1 < 0$, aber $1 = (-1)^2 > 0$ - eine Widerspruch.

[2] *sesquialter* (lat.) = anderthalb. *sesqui* (lat.) = mit ein Halb mehr.

[3] Können Sie beweisen, dass wenn $f : \mathbb{R} \longrightarrow \mathbb{R}$ ein Körperisomorphismus ist, dann gilt $f = \mathrm{id}_{\mathbb{R}}$?

[4] Nach dem französischen Mathematiker Charles Hermite (1822-1901).

Definition 12.6. Eine Sesquilinearform auf dem $\mathbb{C}$-Vektorraum V heißt **Hermitesch** (beziehungsweise **schief-Hermitesch**) wenn

$$h(v_1, v_2) = \overline{h(v_2, v_1)} \qquad (\text{bzw. } h(v_1, v_2) = -\overline{h(v_2, v_1)}) \qquad \forall\, v_1, v_2 \in V.$$

Genau wie bei Bilinearformen, wenn $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V ist, definieren wir die darstellende Matrix einer Sesquilinearform h bezüglich der Basis $\mathcal{B}$ durch

$$\mathcal{M}^{\mathcal{B}}(h) := (h(v_i, v_j))_{i,j=1,\ldots,n} \in \mathrm{Mat}_n(\mathbb{C}).$$

Definition 12.7. Sei $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{C})$. Die **(komplexe) adjungierte Matrix** von A ist

$$A^{\mathsf{H}} := \overline{A}^{\top} = (\overline{a_{ji}}).$$

Die Matrix A heißt **Hermitesch** (beziehungsweise **schief-Hermitesch**) wenn

$$A = A^{\mathsf{H}} \quad (\text{beziehungsweise } A = -A^{\mathsf{H}}).$$

Beispiel 12.8. Die **kanonische Hermitesche Form** auch $\mathbb{C}^n$ ist $h : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$ gegeben durch

$$h(\mathbf{z}, \mathbf{w}) = \sum_{k=1}^{n} \overline{z_k} w_k.$$

Die darstellende Matrix von h bezüglich der Standardbasis ist die Einheitsmatrix $I_n \in \mathrm{Mat}_n(\mathbb{C})$.

Die nächste Bemerkung sagt uns einfach, dass, solang man die Transponierung mit Adjugierung ersetzt, die Zuordnung Sesquilinearform $\mapsto$ Matrix analog zu der Zuordnung Bilinearform $\mapsto$ Matrix funktioniert.

Bemerkung 12.9. Seien V ein n-dimensionaler $\mathbb{C}$-Vektorraum mit $n \in \mathbb{N}$ und $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V. Sei $K_{\mathcal{B}} : V \longrightarrow \mathbb{C}^n$ die Koordinatenabbildung definiert durch $K_{\mathcal{B}}(v_i) = e_i$.

1. Für jede Sesquilinearform $h : V \times V \longrightarrow \mathbb{C}$ gilt
$$h(v, v') = K_{\mathcal{B}}(v)^{\mathsf{H}} \cdot \mathcal{M}^{\mathcal{B}}(h) \cdot K_{\mathcal{B}}(v').$$
2. Jede Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ definiert eine Sesquilinearform $h_A^{\mathcal{B}} : V \times V \longrightarrow \mathbb{C}$ durch
$$h_A^{\mathcal{B}}(v, v') = K_{\mathcal{B}}(v)^{\mathsf{H}} \cdot A \cdot K_{\mathcal{B}}(v'),$$
dessen darstellende Matrix bezüglich $\mathcal{B}$ wieder A ist.
3. Die Abbildung $\mathcal{M}^{\mathcal{B}} : \mathrm{SLF}(V) \longrightarrow \mathrm{Mat}_n(\mathbb{C})$ ist ein $\mathbb{C}$-linearer Isomorphismus.
4. Wenn $\mathcal{B}, \mathcal{B}'$ Basen von V sind, $S = \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}$ die Basiswechselmatrix ist, und $h \in \mathrm{SLF}(V)$, dann gilt
$$\mathcal{M}^{\mathcal{B}'}(h) = S^{\mathsf{H}} \cdot \mathcal{M}^{\mathcal{B}}(h) \cdot S$$

Bemerkung 12.10. Sei V ein $\mathbb{C}$-Vektorraum und $h \in \mathrm{SLF}(V)$.

1. Wenn h Hermitesch ist, dann gilt $h(v, v) \in \mathbb{R}$ für alle $v \in V$.

2. Wenn h schief-Hermitesch ist, dann gilt $h(v,v) \in i \cdot \mathbb{R}$ für alle $v \in V$.

Definition 12.11. Sei V ein $\mathbb{C}$-Vektorraum.

1. Eine Hermitesche Form auf V heißt **positiv-definit** wenn $h(v,v) > 0$ für alle $v \in V \setminus \{0\}$.
2. Positiv semi-definit, negativ (semi-)definit sind analog zum reellen Fall definiert.
3. Ein **Skalarprodukt** auf V ist eine positiv definite Hermitesche Form.
4. Ein **unitärer Vektorraum** ist ein $\mathbb{C}$-Vektorraum mit einem Skalarprodukt.

Beispiel 12.12. Der unitäre Standardraum ist $\mathbb{C}^n$ zusammen mit der kanonischen Hermiteschen Form $\langle \cdot \mid \cdot \rangle$. Es gilt also

$$\langle \mathbf{z}, \mathbf{z} \rangle = \sum_{k=1}^{n} z_k \overline{z_k} = \sum_{k=1}^{n} |z_k|^2 > 0, \quad \forall\, \mathbf{z} \neq 0.$$

12.2 Orthogonale Zerlegung

Wir werden ab jetzt euklidische und unitäre Vektorräume gemeinsam behandeln. Das heißt V ist ein $\mathbb{K}$-Vektorraum wobei der Körper $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$ ist, und $\langle \cdot, \cdot \rangle$ wird immer das Skalarprodukt bezeichnen.

Orthogonalität wird wie in Definition 11.20 gegeben:

$$v \perp w \quad \Longleftrightarrow \quad \langle v, w \rangle = 0.$$

Definition 12.13. Ein **Orthogonalsystem** ist eine Menge $S \subseteq V \setminus \{0_V\}$ von nicht-trivialen Vektoren aus V, sodass

$$\langle v, w \rangle = 0, \quad \forall v, w \in S \quad \text{mit} \quad v \neq w.$$

Ein Orthogonalsystem S ist **vollständig** wenn es kein anderes Orthogonalsystem S' gibt, mit $S \subsetneq S'$.

Lemma 12.14. *Wenn S ein Orthogonalsystem ist, dann ist S linear unabhängig. Insbesondere, wenn V endlichdimensional ist, dann $|S| \leqslant \dim_{\mathbb{K}} V$.*

Beweis-Skizze: Seien $v_1, \ldots, v_n \in S$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, sodass $\sum_{i=1}^n \lambda_i \cdot v_i = 0_V$. Für jeden $j \in \{1, \ldots, n\}$ haben wir

$$0 = \langle v_j, 0_V \rangle = \langle v_j, \textstyle\sum_{i=1}^n \lambda_i v_i \rangle = \sum_{i=1}^n \lambda_i \langle v_j, v_i \rangle = \lambda_j \langle v_j, v_j \rangle .$$

Weil $v_j \neq 0_V$ und das Skalarpordukt positiv-definit ist, haben wir $\langle v_j, v_j \rangle \neq 0$, also $\lambda_j = 0$. Q.E.D.

Bemerkung 12.15. Ein Orthogonalsystem S ist vollständig genau dann, wenn $S^\perp = \{0_V\}$. Sonst kann man es mit einem $0 \neq v \in S^\perp$ ergänzen.

Satz 12.16. *Sei $U \subseteq V$ eine $\mathbb{K}$-UVR eines endlichdimensionalen euklidischen/unitären Vektorraumes V. Sei $U^{\perp\perp} := (U^\perp)^\perp$. Dann gelten*

(i) $V = U \oplus U^\perp$.

(ii) $U^{\perp\perp} = U$.

Beweis-Skizze: (i) Die Aussage ist trivial wenn $U = 0$. Weil $\langle \cdot\,, \cdot \rangle$ positiv-definit ist, ist die Einschränkung auf jedem nicht-trivialen Untervektorraum nicht ausgeartet, und (i) folgt somit aus Korollar 11.29.
(ii) Wir haben offensichtlich $U \subseteq U^{\perp\perp}$, und aus Teil (i) gilt $\dim U = \dim U^{\perp\perp}$. Q.E.D.

Weil $V = U \oplus U^\perp$, wird $U^\perp = \{v \in V \;:\; v \perp u,\ \forall\, u \in U\}$ das **orthogonale Komplement** von U in V genannt.

Korollar 12.17. Sei $U \subseteq V$ ein $\mathbb{K}$-Untervektorraum eines endlichdimensionalen euklidischen/unitären Vektorraumes. Dann ist jeder $v \in V$ auf eindeutiger Weise als $v = u + u^\perp$ zerlegbar, mit $u \in U$ und $\langle u, u^\perp \rangle = 0$.

Die kanonische Projektion $p : V = U \oplus U^\perp \longrightarrow U$ gegeben durch

$$p(v) := u, \quad \text{wobei } u \in U \text{ der eindeutige Vektor mit } v = u + u^\perp \text{ ist,}$$

heißt die **orthogonale Projektion** auf dem Untervektorraum U, oder die orthogonale Projektion entlang $U^\perp$.

12.3 Die Norm

Um das Skalarprodukt für Messungen von Längen und Winkeln anzuwenden, definieren wir zuerst eine Norm, und beweisen eine fundamentale Ungleichung dafür: die Cauchy-Schwarz Ungleichung. Sei weiterhin V ein euklidischer $\mathbb{R}$-Vektorraum oder ein unitärer $\mathbb{C}$-Vektorraum. Wir schreiben $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

Definition 12.18. Die **Länge** (oder **Norm**) eines Vektors $v \in V$ ist

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

Bemerkung 12.19. Für jeden $v \in V$ und $\lambda \in \mathbb{K}$ gelten

(i) $\|v\| \geqslant 0$ und $\|v\| = 0 \iff v = \mathbf{0}$.

(ii) $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$.

Ein **Orthonormalsystem** ist ein Orthogonalsystem S für das zusätzlich

$$\|v\| = 1, \quad \forall\, v \in S \;\text{ gilt.}$$

Die Menge $S = \{v_1, \ldots, v_n\} \subseteq V$ ist also ein Orthonormalsystem genau dann, wenn $\langle v_i, v_j \rangle = \delta_{ij}$.
Wir fangen mit einer Ungleichung an, die für Orthonormalsysteme gilt. Mit Hilfe dieser Ungleichung beweisen wir dann die Cauchy-Schwarzsche Ungleichung.

Satz 12.20 (Besselsche Ungleichung)**.** *Sei $S = \{v_1, \ldots, v_m\}$ ein (nicht unbedingt vollständiges) Orthonormalsystem in dem euklidischen/unitären Vektorraum V. Für alle $v \in V$ gilt*

$$\sum_{i=1}^{m} |\, \langle v_i, v \rangle \,|^2 \leqslant \|v\|^2. \tag{12.1}$$

Weiterhin, gilt

$$v - \sum_{i=1}^{m} \langle v_i, v \rangle\, v_i \in S^\perp.$$

Beweis-Skizze: Wir bezeichnen mit $\lambda_i = \langle v_i, v\rangle$ und definieren $w := v - \sum_{i=1}^m \lambda_i v_i$. Wir haben

$$\begin{aligned} 0 \leqslant \|w\|^2 &= \langle w, w\rangle \\ &= \langle v - \textstyle\sum_{i=1}^m \lambda_i v_i, v - \sum_{i=1}^m \lambda_i v_i\rangle \\ &= \langle v, v\rangle - \textstyle\sum_{i=1}^m \overline{\lambda_i}\,\langle v_i, v\rangle - \sum_{i=1}^m \lambda_i \langle v, v_i\rangle + \sum_{i,j=1}^m \overline{\lambda_i}\lambda_j \langle v_i, v_j\rangle \\ &= \langle v, v\rangle - \textstyle\sum_{i=1}^m \overline{\lambda_i}\lambda_i - \sum_{i=1}^m \lambda_i\overline{\lambda_i} + \sum_{i,j=1}^m \overline{\lambda_i}\lambda_j\delta_{ij} \\ &= \|v\|^2 - 2\textstyle\sum_{i=1}^m |\lambda_i|^2 + \sum_{i=1}^m |\lambda_i|^2 \\ &= \|v\|^2 - \textstyle\sum_{i=1}^m |\lambda_i|^2. \end{aligned}$$

Die Orthonormalität haben wir beim Übergang von der 3. zu der 4. Zeile, in der letzten Summe verwendet.

Für die zweite Aussage haben wir $\langle v_j, w\rangle = \langle v_j, v\rangle - \sum_{i=1}^n \lambda_i \langle v_j, v_i\rangle = \lambda_j - \sum_{i=1}^n \lambda_i\delta_{ij} = 0$.

Q.E.D.

Satz 12.21 (Cauchy-Schwarzsche Ungleichung)**.** *Sei V ein euklidischer/unitärer Vektorraum. Für alle Vektoren $v, w \in V$ gilt*

$$|\langle v, w\rangle| \leqslant \|v\| \cdot \|w\|. \tag{12.2}$$

Beweis-Skizze: Wenn $w = \mathbf{0}$, dann gilt sogar Gleichheit. Wenn $w \neq \mathbf{0}$, dann ist die Menge $S = \{\frac{1}{\|w\|} \cdot w\}$ ein Orthonormalsystem. Die besselsche Ungleichung (12.1) für $v \in V$ und S gibt uns

$$|\langle v, \tfrac{1}{\|w\|} \cdot w\rangle|^2 \leqslant \|v\|^2 \quad \Longleftrightarrow \quad \tfrac{1}{\|w\|^2} \cdot |\langle v, w\rangle|^2 \leqslant \|v\|^2.$$

Die Cauchy-Schwarzsche Ungleichung folgt, weil die Norm und der Betrag positive reelle Zahlen geben.

Q.E.D.

Bemerkung 12.22. In der Cauchy-Schwarzschen Ungleichung gilt Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis-Skizze: Wenn $v = \lambda w$, dann gilt offensichtlich die Gleichung. Für die andere Richtung, nehmen wir an, dass $|\langle v, w\rangle| = \|v\| \cdot \|w\|$. Das heißt

$$\|v\|^2 = \frac{|\langle v, w\rangle|^2}{\|w\|^2}. \tag{12.3}$$

Allgemein gilt auch $|\langle v,w\rangle| = |\overline{\langle w,v\rangle}| = |\langle w,v\rangle|$. Für $\lambda = \frac{\langle w,v\rangle}{\|w\|^2}$ gilt also, dass

$$\begin{aligned}
\langle v-\lambda w,\ v-\lambda w\rangle &= \|v\|^2 - \overline{\lambda}\langle w,v\rangle + \lambda\langle v,w\rangle + |\lambda|^2\|w\|^2 \\
&= \|v\|^2 - \frac{\overline{\langle w,v\rangle}}{\|w\|^2}\cdot\langle w,v\rangle - \frac{\langle w,v\rangle}{\|w\|^2}\cdot\langle v,w\rangle + \frac{|\langle v,w\rangle|^2}{\|w\|^2} \\
&= \|v\|^2 - \frac{\overline{\langle w,v\rangle}\langle w,v\rangle}{\|w\|^2} - \frac{\langle w,v\rangle\langle v,w\rangle}{\|w\|^2} + \frac{|\langle v,w\rangle|^2}{\|w\|^2} \\
&= \frac{|\langle v,w\rangle|^2}{\|w\|^2} - \frac{|\langle w,v\rangle|^2}{\|w\|^2} - \frac{|\langle w,v\rangle|^2}{\|w\|^2} + \frac{|\langle v,w\rangle|^2}{\|w\|^2} \\
&= 0.
\end{aligned}$$

Also, weil $\langle\cdot\,,\cdot\rangle$ positiv-definit ist, muss $v-\lambda w = 0$ gelten, also $v = \lambda w$. Q.E.D.

Bemerkung 12.23. Für zwei beliebige Vektoren $v, w \in V$ gilt

$$\begin{aligned}
\|v+w\|^2 &= \langle v+w,\ v+w\rangle \\
&= \|v\|^2 + \langle v,w\rangle + \langle w,v\rangle + \|w\|^2 \\
&= \|v\|^2 + \langle v,w\rangle + \overline{\langle v,w\rangle} + \|w\|^2 \\
&\leqslant \|v\|^2 + 2|\langle v,w\rangle| + \|w\|^2 \\
\text{aus (12.2)} \quad &\leqslant \|v\|^2 + 2\|v\|\cdot\|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Der Übergang von der 3. zu der 4. Zeile gilt, weil $\forall z = a+ib \in \mathbb{C}$ gilt $z+\overline{z} = 2a$ und $2a \leqslant 2|a| = 2\sqrt{a^2} \leqslant 2\sqrt{a^2+b^2}$. Also, weil beide Seiten nicht-negativ sind, folgt auch die so-genannte Dreiecksungleichung:

$$\|v+w\| \leqslant \|v\| + \|w\|. \tag{12.4}$$

Die Cauchy-Schwarzsche Ungleichung ist wichtig, weil diese uns eine *Metrik* (Abstandsfunktion) auf dem Vektorraum zu definieren erlaubt.

12.3.1 Allgemeine Normen, Distanzen, und metrische Räume

Wir erinnern erstmals kurz was ein Metrischer Raum ist.

Definition 12.24. Sei V ein $\mathbb{R}$-Vektorraum. Eine **Norm** ist eine Abbildung $\|\cdot\|\colon V \longrightarrow \mathbb{R}$ die folgende Axiome erfüllt:

(N 1) $\|v\| \geqslant 0$ für alle $v \in V$ und $\|v\| = 0 \iff v = 0_V$.

(N 2) $\|\lambda v\| = |\lambda|\cdot\|v\|$ für alle $\lambda \in \mathbb{R}$ und $v \in V$.

(N 3) $\|v+w\| \leqslant \|v\| + \|w\|$ für alle $v, w \in V$.

Die ersten zwei Axiome gelten offensichtlich für die Länge die wir in Definition 12.18 eingeführt haben. Das Axiom (N3), das auch als Dreiecksungleichung bekannt ist, haben wir in (12.4) bewiesen.

Definition 12.25. Sei X eine Beliebige Menge. Eine **Metrik** (oder **Abstandsfunktion**) auf X ist eine Abbildung $d\colon X \times X \longrightarrow \mathbb{R}$ die folgende Axiome erfüllt:

(M 1) $d(x,y) \geqslant 0$ für alle $x, y \in X$ und $d(x,y) = 0 \iff x = y$.

(M 2) $d(x,y) = d(y,x)$ für alle $x, y \in X$.

(M 3) $d(x,z) \leqslant d(x,y) + d(y,z)$ für alle $x, y, z \in X$.

Ein **Metrischer Raum** ist ein Paar (X,d) wobei X eine Menge und d eine Metrik auf X ist.

Wichtig zu bemerken hier ist, dass X keine weitere Struktur oder Verknüpfung haben muss.

12.3.2 Die euklidische/unitäre Räume haben eine Metrik

Definition 12.26. Sei V ein euklidischer/unitärer Vektorraum. Die **Distanz** (oder der **Abstand**, oder die **Entfernung**) zweier Vektoren $v, w \in V$ ist die reelle Zahl:

$$d(v,w) := \|v - w\| = \sqrt{\langle v - w, v - w\rangle}.$$

Bemerkung 12.27 (Dreiekcsungliechung)**.** Seien $v, w, u \in V$ drei Vektoren im euklidischen Vektorraum V. Es gilt

$$d(v,w) \leqslant d(v,u) + d(u,w). \tag{12.5}$$

Beweis-Skizze: Wir setzen in (12.4) $v \mapsto v-u$ und $w \mapsto u-w$ ein und bekommen die erwünschte Ungleichung. Q.E.D.

Bemerkung 12.28. Jeder euklidische/unitäre Vektorraum V ist also ein metrischer Raum mit der Metrik aus Definition 12.26. Weiterhin, ist diese Distanz unter Translationen invariant; das heißt

$$d(v,w) = d(v+u, w+u), \quad \forall\, v, w, u \in V.$$

Beispiele:

1. Wenn $V = \mathbb{R}^n$ der euklidische Standardraum von Dimension n ist, dann gibt uns die Cauchy-Schwarzsche Ungleichung:

   $$(x_1y_1 + \cdots + x_ny_n)^2 \leqslant (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2), \qquad \forall\, x_i, y_i \in \mathbb{R}.$$

   Die Abstandsfunktion die von dem euklidischen Skalarprodukt $\langle (x_1,\ldots,x_n), (y_1,\ldots,y_n)\rangle = \sum_{i=1}^n x_iy_i$ auf der Menge $\mathbb{R}^n$ definiert wird **euklidische Metrik** auf $\mathbb{R}^n$ genannt, und ist durch folgende Formel gegeben:

   $$d(\mathbf{x},\mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

2. Für den $\mathbb{R}$-Vektorraum der stetigen Funktionen folgt aus der Cauchy-Schwarzschen Ungleichung, dass

   $$\left(\int_a^b f(t)g(t)dt\right)^2 \leqslant \int_a^b f(t)^2dt \int_a^b g(t)^2dt.$$

12.3.3 Winkelmessung in euklidische Räume

Man kann den Winkel zwischen zwei Vektoren auch für komplexe Vektorräume definieren. Das Maß wäre dann eine komplexe Zahl, und weniger intuitiv. Wir werden uns hier nur um das Maß des Winkels zwischen zwei Vektoren in einem euklidischen, und somit reellem, Vektorraum.

Nur in diesem Abschnitt 12.3.3 sei also V ein euklidischer, aber kein unitärer, Vektorraum.

Bemerkung 12.29. Für zwei nicht-triviale Vektoren $v, w \in V$ gilt nach (12.2)

$$-1 \leqslant \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \leqslant 1.$$

Es gibt also ein eindeutiges $\theta \in [0, \pi]$, sodass

$$\cos\theta = \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}.$$

Definition 12.30. Der **Winkel zwischen zwei nicht-triviale Vektoren** v, w eines euklidischen Vektorraumes ist die reelle Zahl $\theta \in [0, \pi]$ gegeben durch

$$\theta := \arccos\left(\frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}\right).$$

12.4 Orthonormierte Basen

Wir fangen mit vollständige Charakterisierungen von endlichen vollständigen Orthonormalsystemen an.

Satz 12.31. *Sei $S = \{v_1, \ldots, v_n\}$ ein endliches Orthonormalsystem des euklidischen/unitären Vektorraumes V. Folgende Aussagen sind äquivalent:*

(a) *S ist ein vollständiges Orthonormalsystem.*

(b) *Wenn für $v \in V$ gilt $\langle v_i, v\rangle = 0,\ \forall\ i = 1, \ldots, n$, dann ist $v = 0_V$.*

(c) $\mathrm{Span}_{\mathbb{K}}\, S = V$.

(d) *Für jeden $v \in V$ gilt $v = \sum_{i=1}^{n} \langle v_i, v\rangle \cdot v_i$.*

(e) *Für jede zwei Vektoren $v, w \in V$ gilt* Parseval's Gleichung

$$\langle v, w\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle \cdot \langle v_i, w\rangle.$$

(f) *Für alle $v \in V$ gilt $\|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i\rangle|^2$.*

Beweis-Skizze: (a)⇒(b) Sei $v \in V$ mit $\langle v, v_i\rangle = 0,\ \forall\ i = 1, \dots, n$. Wenn $v \neq 0_V$, dann ist $S \cup \{\frac{1}{\|v\|} \cdot v\}$ ein Orthonormalsystem - ↯ Widerspruch zur Vollständigkeit von S.

(b)⇒(c) Sei $v \in V$. Aus Satz 12.20 folgt $w := v - \sum_{i=1}^n \langle v_i, v\rangle \cdot v_i \in S^\perp$. Also $\langle w, v_i\rangle = 0$ für alle i, und somit aus (b) $w = 0$. Das ist aber äquivalent zu $v = \sum_{i=1}^n \lambda_i v_i$, also $v \in \mathrm{Span}_{\mathbb{K}} S$.

(c)⇒(d) Sei $v \in V = \mathrm{Span}_{\mathbb{K}} S$. Es gibt also $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $v = \sum_{i=1}^n \lambda_i v_i$. Es gilt dann für jeden j:

$$\langle v_j, v\rangle = \textstyle\sum_{i=1}^n \lambda_i \langle v_j, v_i\rangle = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j.$$

(d)⇒(e) Seien $v, w \in V$. Aus (d) gelten $v = \sum_{i=1}^n \langle v_i\ , v\rangle \cdot v_i$ und $w = \sum_{i=1}^n \langle v_i, w\rangle \cdot v_i$. Dann

$$\begin{aligned}
\langle v, w\rangle &= \langle \textstyle\sum_{i=1}^n \langle v_i, v\rangle \cdot v_i,\ \sum_{j=1}^n \langle v_j, w\rangle \cdot v_j\rangle \\
&= \textstyle\sum_{i,j=1}^n \overline{\langle v_i, v\rangle} \cdot \langle v_j, w\rangle \cdot \langle v_i, v_j\rangle \\
&= \textstyle\sum_{i,j=1}^n \overline{\langle v_i, v\rangle} \cdot \langle v_j, w\rangle \cdot \delta_{ij} \\
&= \textstyle\sum_{i=1}^n \langle v, v_i\rangle \cdot \langle v_i, w\rangle .
\end{aligned}$$

(e)⇒(f) Man muss nur $v = w$ in (e) einsetzen.

(f)⇒(a) Aus Bemerkung 12.15 reicht es zu zeigen, dass $S^\perp = 0$. Sei also $v \in S^\perp$. Aus (f) haben wir $\|v\|^2 = \sum_{i=1}^n \langle v, v_i\rangle = \sum_{i=1}^n 0 = 0$.

Q.E.D.

Bemerkung 12.32. Aus Lemma 12.14 und der Äquivalenz (a) ⇔(c) in Satz 12.31 folgt, dass ein vollständiges Orthonormalsystem eine Basis ist. Wir nennen eine solche Basis **orthonormierte Basis**.

Der Beweis des nächsten Satzes bring eigentlich mehr als dessen Aussage: ein Algorithmus der eine orthonormierte Basis produziert.

Satz 12.33. *Jeder endlichdimensionale euklidische/unitäre Vektorraum besitzt eine orthonormierte Basis.*

Beweis-Skizze: Wir wissen schon, dass es immer eine Basis $\mathcal{B} = \{v_1, \dots, v_n\}$ von V existiert.
Behauptung: $\forall\ i = 1, \dots, n$ *existiert ein Orthonormalsystem* $\{w_1, \dots, w_i\}$ *in* V*, sodass*

$$w_k \in \mathrm{Span}_{\mathbb{K}}\{v_1, \dots, v_k\}, \quad \forall\ k = 1, \dots, i.$$

Wir konstruieren die w_is Schritt-weise.
$i = 1$ Weil $\mathcal{B}$ eine Basis ist, ist $v_1 \neq 0$. Wir setzen dann $w_1 := \frac{1}{\|v_1\|} \cdot v_1$.
$i \Rightarrow i+1$ mit $i < n$. Sei $\{w_1, \dots, w_i\}$ das Orthogonalsystem mit $w_k \in \mathrm{Span}_{\mathbb{K}}\{v_1, \dots, v_k\},\ \forall\ k = 1, \dots, i$, das aus der induktiven Voraussetzung existiert. Wir setzen

$$w'_{i+1} := v_{i+1} - \textstyle\sum_{k=1}^i \langle w_k, v_{i+1}\rangle\, w_k.$$

Wenn $w'_{i+1} = \mathbf{0}$, dann, weil $w_k \in \mathrm{Span}_{\mathbb{K}}\{v_1, \dots, v_k\}\ \forall\ k$, haben wir eine nicht-triviale Line-

arkombination der Vektoren $v_1, \ldots, v_{i+1}$ die den Nullvektor gibt - Widerspruch ↯ zur linearen Unabhängigkeit der Elementen aus $\mathcal{B}$.
Also $w'_{i+1} \neq \mathbf{0}$. Aus der Aussage 12.20 in Satz 12.20 haben wir $w'_{i+1} \in \{w_1, \ldots, w_i\}^{\perp}$. Also, wenn wir

$$w_{i+1} := \frac{1}{\|w'_{i+1}\|} \cdot w'_{i+1}$$

setzen, dann haben wir das erwünschte Orthonormalsystem gefunden. Unsere Behauptung ist also wahr, und dessen Aussage für $i = n$ zusammen mit Lemma 12.14 gibt uns den Satz. Q.E.D.

Aus dem Beweis des obigen Satzes bekommen wir ein Verfahren das aus einer beliebigen Basis eine orthonormierte Basis produziert.

`Der Gram-Schmidt`[5] `Algorithmus:`

`Eingabe:` $\{v_1, \ldots, v_n\}$ eine Basis des euklidischen Raum V.

`Verfahren:` Für $i = 1 \ldots n$ setze:

$$\begin{aligned} w'_i &:= v_i - \sum_{k=1}^{i-1} \langle w_k, v_i \rangle \, w_k \\ w_i &:= \frac{1}{\|w'_i\|} \cdot w'_i, \end{aligned}$$

(wobei die leere Summe $\sum_{k=1}^{0}$ gleich mit Null ist und für unitäre Räume das Skalarprodukt konjugiert-linear im ersten Argument ist.)
`Ausgabe:` Die orthonormierte Basis $\{w_1, \ldots, w_n\}$.

12.5 Orthogonale und unitäre Endomorphismen

Auch in diesem Teil werden wir so weit wie möglich euklidische und unitäre Vektorräume zusammen behandeln. Die Standard Bezeichnung für den zentralen Begriff ist aber unterschiedlich in den zwei Fällen. Jedes Mal wenn es günstig ist werden wir die zwei Fälle mit "/", statt mit "beziehungsweise" trennen. Zum Beispiel in der ersten Definition gleich hier unten.

Definition 12.34. Sei V ein euklidischer/unitärer Vektorraum mit Skalarprodukt $\langle \cdot\,, \cdot \rangle : V \times V \longrightarrow \mathbb{K}$. Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{R}}(V)/f \in \mathrm{End}_{\mathbb{C}}(V)$ heißt **orthogonal**, beziehungsweise **unitär** (bezüglich $\langle \cdot\,, \cdot \rangle$) wenn

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \qquad \forall\, v, w \in V.$$

Die Menge aller orthogonalen/unitären Endomorphismen auf V wird mit $O(V)$, beziehungsweise $U(V)$ bezeichnet.

Folgende Proposition sagt uns eigentlich, dass wenn unter einer Abbildung Längen erhalten bleiben, dann bleiben auch das Winkelmaß erhalten. Das ist etwas überraschend.

[5] Nach dem dänischen Mathematiker Jørgen Pedersen Gram (1850-1916) und dem deutschen Mathematiker Erhard Schmidt (1876-1959).

Proposition 12.35. *Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist orthogonal/unitär genau dann, wenn $\|f(v)\| = \|v\|$ für alle $v \in V$.*

Beweis-Skizze: Eine Richtung ist klar: $\|f(v)\| = \sqrt{\langle f(v), f(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|$. Für die andere Richtung, wenn $\mathbb{K}{=}\mathbb{R}$ dann haben wir aus der Polarisierungsformel (13.1) auf Seite 295, dass

$$\begin{aligned}\langle f(v), f(w)\rangle &= \frac{1}{4}\big(\|f(v)+f(w)\|^2 + \|f(v)-f(w)\|^2\big)\\ &= \frac{1}{4}\big(\|f(v+w)\|^2 + \|f(v-w)\|^2\big)\\ &= \frac{1}{4}\big(\|v+w\|^2 + \|v-w\|^2\big)\\ &= \langle v, w\rangle\end{aligned}$$

Die Polarisierung Formel für unitäre Räume ist komplizierter:

$$\langle v, w\rangle = \frac{1}{4}\big(\|v+w\|^2 - \|v-w\|^2\big) + \frac{1}{4}i\big(\|v+iw\|^2 - \|v-iw\|^2\big),$$

aber daraus folgt die Äquivalenz auch für unitäre Vektorräume. Q.E.D.

Bemerkung 12.36. Wenn f orthogonal/unitär ist, dann ist f injektiv. Sonst, existiert $0 \neq v \in \mathrm{Ker}\, f$, mit $\langle v, v\rangle = \langle f(v), f(v)\rangle = \langle 0, 0\rangle = 0$ - ein Widerspruch. Das heißt insbesondere, dass wenn V endlich dimensional ist, dann sind orthogonale Abbildungen automatisch auch bijektiv, also Automorphismen von V. Im unendlich dimensionalen Fall kann aber Surjektivität scheitern, wie uns das nächste Beispiel zeigt.

Beispiel 12.37. Sei wieder $\mathbb{K}{=}\mathbb{R}$ oder $\mathbb{C}$, und sei ℓ^2 der $\mathbb{K}$-Vektorraum aller *quadratisch addierbaren* reellen/komplexen Folgen:

$$\ell^2 = \big\{(s_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \;:\; \textstyle\sum_{i\in\mathbb{N}} |s_n|^2 < \infty\big\}.$$

Dieser $\mathbb{K}$-Raum ist euklidisch/unitär mit folgendem Skalarprodukt:

$$\langle (s_i), (t_i)\rangle := \textstyle\sum_{i\in\mathbb{N}} \overline{s_i}\cdot t_i.$$

Das ist konvergent, weil für jeden i gilt

$$0 \leqslant (|s_i| - |t_i|)^2 = |s_i|^2 - 2|s_i||t_i| + |t_i|^2,$$

also $2|s_i t_i| = 2|\overline{s_i} t_i| \leqslant |s_i|^2 + |t_i|^2$. Es ist einfach zu überprüfen, dass das ein Skalarprodukt definiert. Die Abbildung $f : \ell^2 \longrightarrow \ell^2$ gegeben durch

$$f(x_0, x_1, x_2, \dots) = (0, x_0, x_1, \dots)$$

ist orthogonal/unitär, aber offensichtlich nicht surjektiv.

12.5.1 Der endlich-dimensionale Fall

Wir konzentrieren uns ab jetzt auf dem Fall $\dim V < \infty$. Das heißt, dass orthogonale/unitäre Abbildungen immer bijektiv sind.

Lemma 12.38. *Sei $\mathcal{B} = v_1, \ldots, v_n$ eine Orthonormalbasis von V. Eine $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow V$ ist orthogonal genau dann, wenn $f(\mathcal{B})$ eine Orthonormalbasis ist.*

Beweis-Skizze: Wenn f orthogonal ist, dann gilt $\langle f(v_i), f(v_i)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$.
Für die Umkehrung seine $v, w \in V$ mit $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{i=1}^n \mu_i v_i$. Dann gilt $\langle v, w\rangle = \sum_{i=1}^n \lambda_i\mu_i$. Anderseits, weil $f(\mathcal{B})$ orthonormiert ist, gilt das auch für $\langle f(v), f(w)\rangle$:

$$\langle f(v), f(w)\rangle = \langle \textstyle\sum_{i=1}^n \lambda_i f(v_i), \sum_{i=1}^n \mu_i f(v_i)\rangle = \sum_{i=1}^n \lambda_i\mu_i.$$

Q.E.D.

Definition 12.39. Eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) heißt **orthogonal/unitär** wenn

$$A^{\mathsf{H}} A = I_n.$$

Im Fall $\mathbb{K}= \mathbb{R}$ ist $A^{\mathsf{H}} = A^{\top}$.

Korollar 12.40. Eine Abbildung $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist orthogonal/unitär genau dann, wenn für eine Onrthonormalbasis $\mathcal{B}$ die zugeordnete Matrix $M^{\mathcal{B}}(f)$ orthogonal/unitär ist.

Beweis-Skizze: Aus dem Beweis von Lemma 12.38 reicht das für A, f_A und das Standard Skalarprodukt zu zeigen.
Sei also $A \in \mathrm{Mat}_n(\mathbb{K})$ und $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ der zugeordnete Endomorphismus gegeben durch $f_A(\mathbf{x}) = A \cdot \mathbf{x}$. Wir betrachten $\mathbb{K}$ mit dem standard Skalarprodukt auf $\mathbb{K}^n$: $\langle \mathbf{x}, \mathbf{y}\rangle := \mathbf{x}^{\mathsf{H}} \cdot \mathbf{y}$. Dann gilt

$$\langle f_A(\mathbf{x}), f_A(\mathbf{y})\rangle = \langle A\mathbf{x}, A\mathbf{y}\rangle = (A\mathbf{x})^{\mathsf{H}}(A\mathbf{y}) = \mathbf{x}^{\mathsf{H}} \cdot A^{\mathsf{H}} \cdot A \cdot \mathbf{y}.$$

Also A ist orthogonal/unitär genau dann, wenn f_A orthogonal bezüglich des standard Skalarprodukt ist. Q.E.D.

Bemerkung 12.41.

1. Die Menge aller *orthogonalen* $n \times n$ Matrizen wird mit $O(n)$ bezeichnet. $\mathrm{O}(n)$ ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$. Diese ist als die **orthogonale Gruppe** bekannt.
2. Die Menge aller *unitären* $n \times n$ Matrizen wird mit $U(n)$ bezeichnet. $\mathrm{U}(n)$ ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{C})$. Diese ist als die **unitäre Gruppe** bekannt.
3. Wenn $h \in \mathrm{End}_{\mathbb{K}}(V)$ orthogonal/unitär ist, und wenn λ ein Eigenwert von h ist, dann gilt $|\lambda| = 1$.
4. Wenn $A \in \mathrm{O}(n)$ oder $A \in \mathrm{U}(n)$, dann gilt

$$1 = \det I_n = \det(A^{\mathsf{H}} \cdot A) = \det A^{\mathsf{H}} \cdot \det A = \overline{\det(A)} \cdot \det A = |\det A|^2,$$

also $|\det A| = 1$. Insbesondere, wenn $\mathbb{K} = \mathbb{R}$ ist $\det A \in \{\pm 1\}$.

5. Der Homomorphismus $\det : \mathrm{O}(n) \longrightarrow \mathrm{O}(1) = \{\pm 1\}$ ist surjektiv. Der Kern dieses Homomorphismus heißt die **spezielle Orthogonalgruppe** und wird mit $\mathrm{SO}(n)$ bezeichnet. Also

$$\mathrm{SO}(n) := \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid A^\top A = I_n \text{ und } \det A = 1\}.$$

Wenn $n = 2$ oder $n = 3$, dann entspricht diese Gruppe der "gewöhnlichen" Drehungen um den Ursprung. In $\mathbb{R}^3$ ist jede Drehung um den Ursprung eine Drehung um einer Achse, das heißt um einer Geraden durch den Ursprung. Man kann also zeigen, dass

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \;:\; \theta \in [0, 2\pi) \right\}.$$

6. $U(1) = \{z \in \mathbb{C} \mid |z| = 1\}$.

7. Die Abbildung $\det : U(n) \longrightarrow U(1)$ ist ein surjektiver Gruppenhomomorphismus. Der Kernel davon ist als die **spezielle unitäre Gruppe** bekannt, und wird mit

$$\mathrm{SU}(n) = \{A \in \mathrm{Mat}_n(\mathbb{C}) \mid AA^{\mathsf{H}} = I_n \text{ und } \det(A) = 1\}$$

bezeichnet.

Beispiele:

1 Wenn $h_\lambda \in \mathrm{End}_{\mathbb{R}}(V)$ durch $h_\lambda(v) := \lambda \cdot v \;\; \forall\, v \in V$ gegeben ist, dann ist $h_\lambda \in O(V) \iff \lambda \in \{\pm 1\}$.

2 Wenn $A \in \mathrm{Mat}_n(\mathbb{R})$ orthogonal ist, und wenn $\lambda \in \mathbb{R}$ ein Eigenwert von A ist, dann $\lambda \in \{\pm 1\}$.

3 Sei $\mathbb{R}^2$ die euklidische Standardebene. Drehungen D_θ um Winkel θ sind orthogonal. Die zugeordnete Matrix bezüglich der Standardbasis ist

$$A_\theta := M^{\{e_1,e_2\}}(D_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

4 Sei $\mathbb{R}^2$ die euklidische Standardebene. Die Spiegelung σ_{Ox} an der x-Achse ist orthogonal. Die zugeordnete Matrix bezüglich der Standardbasis ist

$$M^{\{e_1,e_2\}}(\sigma_{Ox}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

5 Sei $\mathbb{R}^2$ die euklidische Standardebene. Sei $g \subset \mathbb{R}^2$ eine Gerade durch den Ursprung die mit der x-Achse einen Winkel θ bildet. Die Spiegelung σ_g an der Gerade g ist orthogonal. Die zugeordnete Matrix bezüglich der Standardbasis ist

$$M^{\{e_1,e_2\}}(\sigma_g) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

6 Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$ eine orthogonale Matrix. Das heißt

$$a^2 + c^2 = b^2 + d^2 = 1 \quad \text{und } ab + cd = 0.$$

Seien $\alpha := a + ic \in \mathbb{C}$ und $\beta := b + id \in \mathbb{C}$. Dann gilt

$$|\alpha| = |\beta| = 1 \quad \text{und } \mathrm{Re}(\alpha\overline{\beta}) = 0.$$

Es folgt $|\alpha\overline{\beta}| = |\alpha||\overline{\beta}| = |\alpha||\beta| = 1$. Also $\alpha\overline{\beta}$ ist eine komplexe Zahl mit reeller Teil 0 und Betrag 1. Also $\alpha\overline{\beta} \in \{\pm i\}$, also $\beta = \pm i\alpha$. Aus der Polarform haben wir $\alpha = \cos\theta + i\sin\theta$ mit $\theta \in [0, 2\pi)$. Es folgt also, dass f_A entweder eine Drehung oder eine Spiegelung an der Gerade durch 0 und $\sqrt{\alpha}$ ist.

12.6 Spektralsätze

Die Menge der Eigenwerte eines Endomorphismus ist auch als **Spektrum** bekannt. In diesem letzten Teil des Kapitels werden wir unsere Aufmerksamkeit zurück auf Eigenwerte, Eigenvektoren und Diagonalisierbarkeit lenken. Nicht allgemein, nur für hermitesche komplexe Matrizen und symmetrische reelle Matrizen.
Wir mit einem einfachen und direkten Beweis der Aussage: alle Eigenwerte einer symmetrischen Matrix mit reellen Einträgen sind reell. Sei $A \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{R})$. Weil $\mathbb{R} \subset \mathbb{C}$, können wir diese Matrix als komplexe Matrix betrachten. Weil $\mathbb{C}$ algebraisch abgeschlossen ist, hat diese Matrix, mit Multiplizität gezählt, genau n komplexe Eigenwerte. Wir zeigen jetzt, dass alle diese komplexe Eigenwerte reell sein müssen. Sei also $\lambda \in \mathbb{C}$, sodass $\exists \mathbf{x} \in \mathbb{C}^n \setminus \{0\}$ mit

$$A \cdot \mathbf{x} = \lambda \cdot \mathbf{x}. \tag{12.6}$$

Wir konjugieren alles und, weil Konjugation ein Automorphismus von $\mathbb{C}$ ist (also verträglich mit der Addition und der Multiplikation der komplexen Zahlen), bleibt die Gleichheit erhalten:

$$\overline{A} \cdot \overline{\mathbf{x}} = \overline{\lambda} \cdot \overline{\mathbf{x}}.$$

Wir können auch beide Seiten transponieren, und dann bekommen wir:

$$\overline{\mathbf{x}}^\top \cdot \overline{A}^\top = \overline{\mathbf{x}}^\top \cdot \overline{\lambda}. \tag{12.7}$$

Wir multiplizieren jetzt (12.6) links mit $\overline{\mathbf{x}}^\top$ und (12.7) rechts mit $\mathbf{x}$ und verwenden die Symmetrie der reellen Matrix A: Das heißt $A = \overline{A}^\top$:

$$\overline{\mathbf{x}}^\top \cdot \lambda \cdot \mathbf{x} = \overline{\mathbf{x}}^\top \cdot A \cdot \mathbf{x} = \overline{\mathbf{x}}^\top \cdot \overline{A}^\top \cdot \mathbf{x} = \overline{\mathbf{x}}^\top \cdot \overline{\lambda} \cdot \mathbf{x}.$$

Weil $\mathbf{x} \neq 0$, gilt auch $\overline{\mathbf{x}} \cdot \overline{\mathbf{x}} = |x_1|^2 + \cdots + |x_n|^2 \neq 0$. Also aus der obigen Gleichung folgt dann $\lambda = \overline{\lambda}$, und das heißt genau, dass der Eigenwert λ reell ist.

Satz 12.42 (von Schur)**.** *Für jede Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ existiert eine unitäre Matrix $U \in U(n)$, sodass $U^{\mathsf{H}}AU$ eine obere Dreiecksmatrix ist.*

Beweis-Skizze: Weil $\mathbb{C}$ algebraisch abgeschlossen ist, zerfällt jedes Polynom in einer in $\mathbb{C}[x]$ in Linearfaktoren. In Satz 10.39 haben wir gezeigt, dass wenn das charakteristische Polynom in Linearfaktoren zerfällt, dann ist die Matrix/ der Endomorphismus, trigonalisierbar. Der Beweis war induktiv, durch ergänzen eines Eigenvektors zu einer Basis. In Satz 11.34 haben wir induktiv gezeigt, dass man eine Orthonormalbasis so produzieren kann. Q.E.D.

Beispiel 12.43. Sei $A = \begin{pmatrix} 1+i & 1 \\ 1 & 1-i \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$. Das ist eine symmetrische Matrix, aber nicht eine hermitesche Matrix, weil

$$A^{\mathsf{H}} = \begin{pmatrix} 1-i & 1 \\ 1 & 1+i \end{pmatrix} \neq A.$$

Wir werden sehen, dass diese auch nicht orthogonal diagonalisierbar sein wird. Das charakteristische Polynom ist $\chi_A(x) = (x-1)^2$, also A hat nur 1 als Eigenwert. Es gilt $(A-I_2)^2 = 0$, also $\mathrm{Hau}(A,1) = \mathbb{C}^2$. Wir wählen dann $v_2 = (0,1)$ und finden $v_1 = (A - I_2) \cdot v_2$. Wir haben v_1 ein Eigenvektor von A. Wir bekommen also

$$\begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix} \cdot \begin{pmatrix} 1+i & 1 \\ 1 & 1-i \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Das ist also die $\mathrm{JNF}(A)$, aber die Matrix $S = \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix}$ die uns dazu führt ist nicht Unitär:

$$S^{\mathsf{H}} = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \neq S^{-1}.$$

Wenn wir aber den Gram-Schmidt Algorithmus auf S anwenden, dann bekommen wir eine Orthonormalbasis:

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix},$$

also mit $U^{\mathsf{H}} \cdot U = I_2$ und $U^{\mathsf{H}} A U = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Der Grund warum das eine obere Dreiecksmatrix bleibt, ist dass $w_1 \in \mathrm{Span}_{\mathbb{C}} v_1$ und $w_2 \in \mathrm{Span}_{\mathbb{C}}\{v_1, v_2\}$. Es mag also nicht mehr die JNF sein, aber es ist mit sicherheit eine obere Dreiecksmatrix. Und das wird hinreichend sein.

Satz 12.44 (Spektralsatz für Hermitesche Matrizen)**.** *Sei $A \in \mathrm{Mat}_n(\mathbb{C})$ eine Hermitesche Matrix. Es gelten:*

(a) *Es gibt eine unitäre Matrix U, sodass*

$$U^{\mathsf{H}} A U = \mathrm{diag}(d_1, \ldots, d_n) \qquad \text{mit } d_i \in \mathbb{R} \ \ \forall\, i = 1, \ldots, n.$$

(b) *Alle Eigenwerte von A sind reell.*

(c) *$\mathbb{C}^n$ hat eine orthonormierte Basis die aus Eigenvektoren von A besteht.*

Beweis-Skizze: (a) Nach dem Satz 12.42 von Schur gibt es $U \in U(n)$ mit $U^{\mathsf{H}}AU = T$, mit $T \in \mathrm{Mat}_n(\mathbb{C})$ eine obere Dreiecksmatrix. Weil A Hermitesch ist, also $A^{\mathsf{H}} = A$, folgt

$$T^{\mathsf{H}} = (U^{\mathsf{H}}AU)^{\mathsf{H}} = U^{\mathsf{H}}A^{\mathsf{H}}U = T.$$

Also, weil T eine obere Dreiecksmatrix ist, muss diese diagonal mit reellen Einträgen sein.

(b) Die Spalten von U bilden die orthonormale Basis von Eigenvektoren von A.

(c) Man muss nur bemerken, dass A und T ähnlich sind, und somit dieselben Eigenwerte haben.

Q.E.D.

Satz 12.45 (Spektralsatz für reelle symmetrische Matrizen)**.** *Sei $A \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{R})$ eine symmetrische Matrix. Es gelten:*

(a) *Das charakteristische Polynom von A zerfällt in Linearfaktoren. (Also A hat, mit Vielfachheit gezählt, n reelle Eigenwerte.)*

(b) *$\mathbb{R}^n$ hat eine orthogonale (sogar orthonormale) Basis die aus Eigenvektoren von A besteht.*

(c) *Es gibt eine orthogonale Matrix P, sodass*

$$P^{\top}AP = \mathrm{diag}(d_1, \ldots, d_n) \in \mathrm{Mat}_n(\mathbb{R}).$$

Weil P orthogonal ist haben wir $P^{\top} = P^{-1}$ also A ist insbesondere diagonalisierbar.

Beweis-Skizze:

(a) Wir betrachten erstmals A als Matrix in $\mathrm{Mat}_n(\mathbb{C})$, obwohl die Einträge aus $\mathbb{R}$ sind. Das heißt, dass $a_{ij} = \overline{a_{ij}}$ für alle i, j. Also

$$A^{\mathsf{H}} = A^{\top} = A.$$

Insbesondere ist A eine hermitsche Matrix. Dann können wir Satz 12.44 anwenden, und A wird nur reelle Eigenwerte haben. Also das charakteristische Polynom χ_A zerfällt in Linearfaktoren in $\mathbb{R}[x]$.

(b) + (c) Nach Punkt (a) und Satz 10.39 ist A trigonalisierbar. Analog zu dem Beweis von Satz 12.42 (siehe Beispiel 12.43) wenden wir das Gram-Schmidt Verfahren an, und finden eine orthonormale Basis P die die Matrix trigonalisiert:

$$P^{\top}AP = T.$$

Wegen der Symmmetrie aber, genau wie im Beweis von Satz 12.44 haben wir

$$T^{\top} = (P^{\top}AP) = P^{\top}A^{\top}(P^{\top})^{\top}) = P^{\top}AP = T.$$

Also T ist eine Diagonalmatrix und die Spalten von P sind Eigenvektoren.

Q.E.D.

Kapitel 13

Quadratische Formen

13.1 Polynome in mehrere Variablen

Wir nehmen an, wir haben den Polynomring in einer Variable mit Koeffizienten in einem kommutativen unitären Ring R definiert:

$$R[x] = \left\{ \sum_{i=0}^{d} c_i x^i \quad : \quad d \in \mathbb{N} \text{ und } c_i \in R \right\}$$

mit den offensichtlichen Addition, Multiplikation und Skalarmultiplikation. Den Polynomring in n Variablen $x_1, \ldots, x_n$ ($n \in \mathbb{N}_{>0}$) definieren wir induktiv als

$$R[x_1, \ldots, x_n] := \big(R[x_1, \ldots, x_{n-1}]\big)[x_n].$$

Zum Beispiel, für $n = 2$ und $d = 2$ wählen wir $c_0 = x_1^2 + 1$, $c_1 = 5x_1 - 1$, $c_2 = x_1^3 + 2x_1^2 + 3x_1 + 4$, und bekommen ein Polynom

$$\begin{aligned} f &= c_0 + c_1 x_2 + c_2 x_2^2 \\ &= (x_1^2 + 1) + (5x_1 - 1)x_2 + (x^3 + 2x^2 + 3x + 4)x_2^2 \\ &= 1 - x_2 + x_1^2 + 5x_1 x_2 + 4x_2^2 + 3x_1 x_2^2 + 2x_1^2 x_2^2 + x_1^3 x_2^2. \end{aligned}$$

Wir haben die Klammern aufgemacht, und die Summanden nach der Summe der Exponenten im Produkt geordnet. Allgemein können wir jedes Polynom in mehrere Variablen als eine endliche R-lineare Kombination unterschiedlicher Produkte von Variablen schreiben.

Ein **Monom** in den Variablen $x_1, \ldots, x_n$ ist ein formales Produkt $x_1^{i_1} \cdots x_n^{i_n}$. Der Grad des Monoms ist die Summe der Exponenten:

$$\deg x_1^{i_1} \cdots x_n^{i_n} = i_1 + \cdots + i_n.$$

Für ein n-Tupel $\mathbf{i} = (i_1, \ldots, i_n) \in \mathbb{N}$ schreiben wir $\mathbf{x}^{\mathbf{i}} := x_1^{i_1} \ldots x_n^{i_n}$. Ein Polynom p in n Variablen $x_1, \ldots, x_n$ über den Ring R ist eine endliche lineare Kombination von Monome, also ein formaler Ausdruck der Form

$$p(x_1, \ldots, x_n) = \sum_{\mathbf{i}=(i_1,\ldots,i_n)\in\mathbb{N}^n} c_{\mathbf{i}}\, \mathbf{x}^{\mathbf{i}}$$

wobei nur endlich viele $c_{\mathbf{i}} \in R$ nicht Null sind. Der Grad eines Polynoms ist

$$\deg p := \max\{\deg \mathbf{x}^{\mathbf{i}} \mid c_{\mathbf{i}} \neq 0\}.$$

Ein Polynom ist **homogen von Grad** d wenn wenn es eine lineare Kombination von Monome von Grad d ist, also wenn

$$c_{\mathbf{i}} \neq 0 \Rightarrow \operatorname{Grad} \mathbf{x}^{\mathbf{i}} = d.$$

Insbesondere ist das Nullpolynom homogen von jedem Grad d. Wir bezeichnen die Menge aller homogenen Polynome von Grad d in Variablen $x_1, \ldots, x_n$ mit $R[x_1, \ldots, x_n]_d$. Polynome kann man wie erwartet mit Skalare multiplizieren, miteinander Addieren und Multiplizieren. Der Polynomring $R[x_1, \ldots, x_n]$ ist also eine R-Algebra.

Wenn R ein *Integritätsbereich*[1] ist, dann gilt für $f, g \in R[x_1, \ldots, x_n]$ und $\lambda \in R \setminus \{0\}$:

$$\begin{aligned} \deg fg &= \deg f + \deg g, \\ \deg f + g &\leqslant \max\{\deg f, \deg g\}, \\ \deg \lambda f &= \deg f. \end{aligned}$$

Wenn aber f und g homogen von Grad d sind, dann ist auch ihre Summe homogen von Grad d.

Wenn $R = \mathbb{K}$ ein Körper ist, dann ist die Menge aller *homogenen* Polynome von Grad d ist$\mathbb{K}[x_1, \ldots, x_n]_d$ ist auch ein $\mathbb{K}$-Vektorraum.

Jedes Polynom definiert eine so-genannte *polynomielle* Abbildung $p : \mathbb{K}^n \longrightarrow \mathbb{K}$, indem wir das Polynom p in jedem Tupel $(t_1, \ldots, t_n)$ evaluieren. Für unsere Zwecke hier werden wir Polynome und die entsprechenden polynomielle Abbildungen identifizieren[2].

Bemerkung 13.1. Für ein homogenes Polynom p von Grad d und jedes $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{K}^n$ gilt

$$p(\lambda \cdot \mathbf{t}) = \lambda^d \cdot p(\mathbf{t}).$$

13.2 Quadratische Formen

Definition 13.2. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Eine **quadratische Form** auf V ist eine Abbildung $q : V \longrightarrow \mathbb{K}$ mit den Eigenschaften

(Q1) $q(\lambda v) = \lambda^2 q(v)$ für alle $\lambda \in \mathbb{K}$ und alle $v \in V$.

(Q2) Die Abbildung $\beta_q : V \times V \longrightarrow \mathbb{K}$ gegeben durch

$$\beta_q(v, v') := q(v + v') - q(v) - q(v')$$

ist eine Bilinearform auf V.

Die Abbildung β_q ist die zu q **assoziierte Bilinearform**. Eine quadratische Form ist nicht ausgeartet wenn die assoziierte Bilinearform nicht ausgeartet ist. Wir bezeichnen die Menge aller quadratischen Formen auf V mit $\mathrm{QF}_{\mathbb{K}}(V)$.

Bemerkung 13.3. 1. Außer den Fällen $q = 0$ oder $V = \mathbb{K} = \mathbb{F}_2$, ist eine quadratische Form q nicht eine lineare Abbildung.

[1] Das heißt: wenn für $a, b \in R$ gilt $ab = 0$, dann $a = 0$ oder $b = 0$.

[2] Das soll man aber allgemein nicht machen! Ein Polynom ist keine Abbildung! Das dient hier nur der Vereinfachung der Sprache wenn wir über die Korrespondenz zwischen Quadratische Formen und homogene Polynome von Grad 2 sprechen werden.

2. Die assoziierte Bilinearform β_q ist symmetrisch.

3. Die Menge $\mathrm{QF}_{\mathbb{K}}(V) \subseteq \mathrm{Abb}(V, \mathbb{K})$ ist ein $\mathbb{K}$-Vektorraum.

Satz 13.4. *Sei $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis des $\mathbb{K}$-Vektorraumes V. Sei $q : V \longrightarrow \mathbb{K}$ eine Abbildung, und sei $p_q^{\mathcal{B}} : \mathbb{K}^n \longrightarrow \mathbb{K}$ die Abbildung definiert durch*

$$p_q^{\mathcal{B}}(\mathbf{x}) := q(x_1 v_1 + \cdots + x_n v_n).$$

Die Abbildung q ist eine quadratische Form genau dann, wenn $p_q^{\mathcal{B}}$ ein homogenes Polynom von Grad 2 ist.

Beweis-Skizze: $\boxed{\Rightarrow}$ Sei $q \in \mathrm{QF}_{\mathbb{K}}(V)$. Aus der Gleichung

$$q(v + v') = q(v) + q(v') + \beta_q(v, v') \qquad \forall v, v' \in V$$

folgt durch vollständige Induktion, dass

$$q(u_1 + \cdots + u_m) = \sum_{i=1}^{m} q(u_i) + \sum_{1 \le i < j \le m} \beta_q(u_i, u_j) \quad \forall\, m \in \mathbb{N}, \forall u_1, \ldots, u_m \in V.$$

Es folgt also

$$\begin{aligned} p_q^{\mathcal{B}}(\mathbf{x}) &= q(x_1 v_1 + \ldots x_n v_n) \\ &= \sum_{i=1}^{n} q(x_i v_i) + \sum_{1 \le i < j \le m} \beta_q(x_i v_i, x_j v_j) \\ &= \sum_{i=1}^{n} q(v_i) \cdot x_i^2 + \sum_{1 \le i < j \le m} \beta_q(v_i, v_j) \cdot x_i x_j. \end{aligned}$$

$\boxed{\Leftarrow}$ Sei $p_q^{\mathcal{B}} = \sum_{1 \le i \le j \le n} a_{ij} x_i x_j \in \mathbb{K}[x_1, \ldots, x_n]_2$. Seien $v, v' \in V$ beliebig mit Koordinaten $(x_1, \ldots, x_n)$, bzw $(x'_1, \ldots, x'_n)$ bezüglich $\mathcal{B}$. Es gilt dann für alle $\lambda \in \mathbb{K}$, dass

$$q(\lambda v) = q(\lambda \cdot (\sum_{i=1}^{n} x_i v_i)) = p_q^{\mathcal{B}}(\lambda x_1, \ldots, \lambda x_n) = \lambda^2 p_q^{\mathcal{B}}(x_1, \ldots, x_n) = \lambda^2 q(v),$$

also (Q1) gilt.

Es gilt

$$\begin{aligned} q(v+v') &= p_q^{\mathcal{B}}(x_1+x_1',\ldots,x_n+x_n') \\ &= \sum_{1\le i\le j\le n} a_{ij}(x_i+x_i')(x_j+x_j') \\ &= \sum_{1\le i\le j\le n} a_{ij}x_ix_j + \sum_{1\le i\le j\le n} a_{ij}x_i'x_j' + \sum_{1\le i\le j\le n} a_{ij}(x_ix_j'+x_jx_i') \\ &= q(v)+q(v')+\sum_{1\le i\le j\le n} a_{ij}(x_ix_j'+x_jx_i'). \end{aligned}$$

Also $\beta_q(v,v') = \sum_{1\le i\le j\le n} a_{ij}(x_ix_j'+x_jx_i')$. Eine direkte Überprüfung gibt uns, dass β_q die Bilinearform dessen darstellende Matrix bezüglich der Basis $\mathcal{B}$ die symmetrische Matrix $B=(b_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ ist, wobei

$$b_{ij} = \begin{cases} a_{ij} & \text{wenn } i<j \\ 2a_{ij} & \text{wenn } i=j \\ a_{ji} & \text{wenn } i>j \end{cases}.$$

Q.E.D.

Bemerkung 13.5. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum.

1. Wenn $\varphi \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine nicht unbedingt symmetrische Bilinearform auf V ist, dann ist $q_\varphi : V \longrightarrow V$, gegeben durch
$$q_\varphi(v) := \varphi(v,v), \quad \forall\, v \in V$$
eine quadratische Form auf V. Die assoziierte Bilinearform $\beta_{q_\varphi} \in \mathrm{Bil}_{\mathbb{K}}^{\mathrm{sym}}(V)$ ist gegeben durch
$$\beta_{q_\varphi}(v,v') = q_\varphi(v+v') - q_\varphi(v) - q_\varphi(v') = \varphi(v,v') + \varphi(v',v).$$
Wenn $\mathcal{B}$ eine Basis von V ist, gilt also
$$\mathcal{M}^{\mathcal{B}}_{\beta_{q_\varphi}} = \mathcal{M}^{\mathcal{B}}_{\varphi} + (\mathcal{M}^{\mathcal{B}}_{\varphi})^{\top}.$$

2. Wenn $\varphi \in \mathrm{Bil}_{\mathbb{K}}^{\mathrm{sym}}(V)$ symmetrisch ist, dann gilt die so genannte *Polarisierungs-Identität* oder *Polarisierungsformel*:
$$\begin{aligned} 2\varphi(v,v') &= q_\varphi(v+v') - q_\varphi(v) - q_\varphi(v') \\ &= \varphi(v+v',v+v') - \varphi(v,v) - \varphi(v',v') \\ &= \varphi(v,v) + \varphi(v,v') + \varphi(v',v) + \varphi(v',v') - \varphi(v,v) - \varphi(v',v') \\ &= \varphi(v,v') + \varphi(v',v) \end{aligned}$$

   Analog bekommen wir auch
$$\begin{aligned} 2\varphi(v,v') &= q_\varphi(v) + q_\varphi(v') - q_\varphi(v-v') \\ &= \varphi(v,v) + \varphi(v',v') - \varphi(v-v',v-v') \\ &= \varphi(v,v) + \varphi(v',v') - \varphi(v,v) - \varphi(v,-v') - \varphi(-v',v) - \varphi(-v',-v') \\ &= \varphi(v,v) + \varphi(v',v') - \varphi(v,v) + \varphi(v,v') + \varphi(v',v) - \varphi(v',v') \end{aligned}$$

Wenn wir die zwei Gleichungen für $2\varphi(v, v')$ addieren, dann bekommen wir:

$$4\varphi(v, v') = q_\varphi(v + v') - q_\varphi(v - v'). \tag{13.1}$$

Insbesondere gilt für die assoziierte Bilinearform von q_φ, dass $\beta_{q_\varphi} = 2\varphi$.

3. Wenn $2 \neq 0$ in $\mathbb{K}$ gilt und $\varphi \in \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{K}}(V)$ ist, dann kann man immer φ aus q_φ zurückgewinnen durch
$$\varphi = \frac{1}{2}\beta_{q_\varphi}.$$
Wenn $2 = 0$, dann läuft viel schief. Zum Beispiel, wenn $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ und $\varphi = \varphi_A \in \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{K}}(\mathbb{K}^2)$, also
$$\varphi(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) = x_1x_2 + x_1y_2 + x_2y_1 + x_2y_2.$$
dann gilt $q_\varphi(\begin{pmatrix} x \\ y \end{pmatrix}) = x^2 + 2xy + y^2 = x^2 + y^2$. Es ist also die standard quadratische Form auf $\mathbb{K}$. Also zwei verschiedene symmetrische Bilinearformen geben die selbe quadratische Form. Weiterhin, es gilt, wie erwartet aus der Polarisierungsformel, dass
$$\beta_{q_\varphi}(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) = (x_1 + x_2)^2 + (y_1 + y_2)^2 - x_1^2 - y_1^2 - x_2^2 - y_2^2 = 0.$$

Beispiel 13.6. Allgemein, auch wenn $\mathrm{char}\,\mathbb{K} \neq 2$, gibt es verschiedene Bilinearformen die dieselbe quadratische Form definieren: $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ und $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. In Charakteristik 2 haben wir gesehen, dass sogar zwei symmetrische Bilinearformen dieselbe quadratische Form definieren können.

Satz 13.7. *Sei $\mathbb{K}$ ein Körper mit* $\mathrm{char}\,\mathbb{K} \neq 2$ *und V ein $\mathbb{K}$-Vektorraum. Dann ist die Abbildung*

$$\begin{array}{ccc} \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{K}}(V) & \longrightarrow & \mathrm{QF}_{\mathbb{K}}(V) \\ \cup & & \cup \\ \varphi & \longmapsto & q_\varphi \end{array},$$

gegeben durch $q_\varphi(v) := \varphi(v, v)$ für alle $v \in V$, ein linearer Isomorphismus.

Beweis-Skizze: Aus Bemerkung 13.5 ist die Abbildung wohl definiert. Die Linearität folgt direkt aus der Definition. Aus der Polarisierungs-Identität ist die Umkehrabbildung durch

$$q \longmapsto \frac{1}{2}\beta_q$$

gegeben. Q.E.D.

Korollar 13.8. Wenn $\dim_{\mathbb{K}} V = n$ und $\mathrm{char}\,\mathbb{K} \neq 2$, dann gilt

$$\mathrm{Mat}^{\mathrm{sym}}_n(\mathbb{K}) \simeq \mathrm{Bil}^{\mathrm{sym}}_{\mathbb{K}}(V) \simeq \mathrm{QF}_{\mathbb{K}}(V).$$

Bemerkung 13.9. Wenn char $\mathbb{K} \neq 2$ und $q(x) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j \in \mathrm{QF}_{\mathbb{K}}(\mathbb{K}^n)$ ist $C = (c_{ij})$ mit

$$c_{ij} = \begin{cases} \frac{a_{ij}}{2}, & i < j \\ a_{ii}, & i = j \\ \frac{a_{ji}}{2}, & i > j \end{cases}$$

Das folgt, weil $\beta_q(e_i, e_j) = 2c_{ij}$.

Korollar 13.10. Wenn char $\mathbb{K} \neq 2$ und $q \in \mathbb{K}[x_1, \ldots, x_n]_2$, dann gibt es $d_i, b_{ij} \in \mathbb{K}$ für alle $i, j = 1, \ldots, n$ sodass

$$q(x_1, \ldots, x_n) = \sum_{i=1}^{n} d_i (b_{i1} x_1 + \cdots + b_{in} x_n)^2.$$

Beweis-Skizze: Sei $q = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$. Nach Bemerkung 13.9 ist $q(\mathbf{x}) = \mathbf{x}^\top \cdot C \cdot \mathbf{x}$, wobei $c_{ij} = \frac{1}{2}(a_{ij} + a_{ji})$. Aus Korollar 11.34 existiert eine invertierbare Matrix $B = (b_{ij}) \in \mathrm{GL}_n(\mathbb{K})$ sodass $B^\top C B = D = \mathrm{diag}(d_1, \ldots, d_n)$. Es gilt also

$$q(\mathbf{x}) = \mathbf{x}^\top C \mathbf{x} = \mathbf{x}^\top B^\top D B \mathbf{x} = (B\mathbf{x})^\top D (B\mathbf{x}) = \sum_{i=1}^{n} d_i Z_i(B\mathbf{x})^2,$$

wobei $Z_i(B\mathbf{x})$ die i-Zeile von $B\mathbf{x}$ bezeichnet. Q.E.D.

Korollar 13.11. (a) Für jede quadratische Form auf einem n-dimensionalen $\mathbb{C}$-Vektorraum V gibt es eine Zahl $0 \leq r \leq n$ und eine geordnete Basis $\mathcal{B}$ von V, sodass das zugeordnete homogene Polynom von q aus Satz 13.4 folgende Form hat

$$p_q^{\mathcal{B}}(x_1, \ldots, x_n) = x_1^2 + \cdots + x_r^2.$$

(b) Für jede quadratische Form auf einem n-dimensionalen $\mathbb{R}$-Vektorraum V gibt es zwei Zahlen $0 \leq r, s \leq n$, mit $r + s \leq n$, und eine geordnete Basis $\mathcal{B}$ von V, sodass das zugeordnete homogene Polynom von q aus Satz 13.4 folgende Form hat

$$p_q^{\mathcal{B}}(x_1, \ldots, x_n) = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2.$$

Kapitel 14

Lineare Affine Geometrie

14.1 Affine Räume

Definition 14.1. Sei $\mathbb{K}$ ein beliebiger Körper. Ein *affiner Raum* über $\mathbb{K}$ ist ein Tripel $(\mathbb{A}, V, \varphi)$, wobei

- $\mathbb{A}$ ist eine nicht-leere Menge
- V ist ein endlich dimensionaler $\mathbb{K}$-Vektorraum
- $\varphi : \mathbb{A} \times \mathbb{A} \longrightarrow V$ ist eine Abbildung die folgende Axiome erfüllt:

(AR 1) Für alle $A, B, C \in \mathbb{A}$, wenn wir $\varphi(A, B)$ mit $\overrightarrow{AB}$ bezeichnen, gilt

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$$

(AR 2) Es gibt einen Punkt $O \in \mathbb{A}$ so dass die Abbildung $\varphi_O : \mathbb{A} \longrightarrow V$, gegeben durch $\varphi_O(A) = \overrightarrow{OA}$ $\forall A \in \mathbb{A}$, ist bijektiv.

Die Elementen von $\mathbb{A}$ heißen *Punkte*, und V heißt der *zugeordnete Vektorraum* von $\mathbb{A}$.

Definition 14.2. Sei $(\mathbb{A}, V, \varphi)$ ein affiner Raum. Die *Dimension* von $\mathbb{A}$ ist die Dimension von V als $\mathbb{K}$-Vektorraum, d.h. $\dim(\mathbb{A}) = \dim_{\mathbb{K}}(V)$.

Bemerkung 14.3. Aus (AR 1) folgt, dass für alle Punkte $A, B \in \mathbb{A}$ gilt:

1. $\overrightarrow{AA} = 0_V$
2. $-\overrightarrow{AB} = \overrightarrow{BA}$

Bemerkung 14.4. Seien $A, B, C, D \in \mathbb{A}$ vier Punkte eines affines Raumes, mit der Eigenschaft, dass $\overrightarrow{AB} = \overrightarrow{CD}$. Dann gilt auch $\overrightarrow{AC} = \overrightarrow{BD}$.

Beweis-Skizze: Wenn wir (AR 1) anwenden, bekommen wir:

$$\overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD} = \overrightarrow{AB} + \overrightarrow{BD}$$

Q.E.D.

Bemerkung 14.5. Aus (AR 1) und (AR 2) folgt, dass für jeder Punkt $O' \in \mathbb{A}$ die Abbildung $\varphi_{O'} : \mathbb{A} \longrightarrow V$, gegeben durch $\varphi_{O'}(A) = \overrightarrow{O'A}$, $\forall A \in \mathbb{A}$, bijektiv ist.

Das erlaubt den Übertrag der Vektorraumstruktur von V auf $\mathbb{A}$, sodass O der Nullvektor wird:

$$\begin{aligned} A + B &= \varphi^{-1}(\overrightarrow{OA} + \overrightarrow{OB}) \quad \forall A, B \in \mathbb{A} \\ \lambda \cdot A &= \varphi^{-1}(\lambda \cdot \overrightarrow{OA}) \quad \forall \lambda \in \mathbb{K}, \quad A \in \mathbb{A} \end{aligned}$$

Auf dieser Weise erhalten wir die einzige $\mathbb{K}$-VR Struktur auf $\mathbb{A}$ so dass φ_O ein VR-Isomorphismus ist. Wir bezeichnen diesen VR mit $\overrightarrow{O\mathbb{A}}$.

Beispiel 14.6.

1. Seien $a, b, c \in \mathbb{R}$ drei reelle Zahlen mit $a^2 + b^2 > 0$, dann ist $\{(x, y) \mid ax + by = c\} \subseteq \mathbb{R}^2$ ein 1-dimensionaler affiner Raum.
2. (V, V, φ), wobei $\varphi(v, w) = w - v$. Dieser wird mit $\mathbb{A}(V)$ bezeichnet. Wenn $V = \mathbb{K}^n$, dann heißt dieser *der n-dimensionale affine Standardraum über* $\mathbb{K}$.
3. Sei V ein $\mathbb{K}$-VR, und $f \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ mit $f \neq 0$. Dann ist $f^{-1}(1) \subset V$ ein affiner Raum.

14.1.1 Gruppenwirkung von V auf $\mathbb{A}$

Definition 14.7. Es seien $(G, \circ)$ eine Gruppe mit neutralem Element e und M eine beliebige Menge. Eine **Gruppenwirkung** ist eine Abbildung

$$\begin{aligned} G \times M &\longrightarrow M \\ (g, m) &\longmapsto g \cdot m \end{aligned}$$

die folgende Axiome erfüllt:

(GW 1) $(g_1 \circ g_2) \cdot m = g_1 \cdot (g_2 \cdot m), \quad \forall\, g_1, g_2 \in G, \quad m \in M.$

(GW 2) $e \cdot m = m, \quad \forall\, m \in M.$

Beispiele:

1. $G = S_n$ und $M = \{1, \ldots, n\}$ mit
$$\sigma \cdot i := \sigma(i).$$
2. Dregungen auf $\mathbb{R}^2$.
3. $\mathrm{GL}(V)$ auf V.

Definition 14.8. Eine Gruppenwirkung $\cdot : G \times M \longrightarrow M$ heißt **transitive Gruppenwirkung**, wenn

$$\forall x, y \in M, \ \exists\, g \in G \text{ sodass } g \cdot x = y.$$

Eine Gruppenwirkung $G \times M \longrightarrow M$ heißt **freie Gruppenwirkung**, wenn für alle $x \in M$ und $g, h \in G$ gilt

$$g \cdot x = h \cdot x \Rightarrow g = h.$$

Bemerkung 14.9. Eine Gruppenwirkung τ ist frei und transitiv wenn und nur wenn $\tau(-,x): G \longrightarrow X$ bijektiv für alle $x \in X$ ist.

Lemma 14.10. *Auf einer nicht-leeren Menge $\mathbb{A}$ gibt es eine affiner Raum Struktur mit V als zugeordnetem Vektorraum $\iff$ es eine transitive und freie Gruppenwirkung von $(V,+)$ auf $\mathbb{A}$ gibt.*

Beweis-Skizze: $\boxed{\Rightarrow}$ Sei $(\mathbb{A}, V, \varphi)$ die affiner-Raum-Struktur. Dann ist die Wirkung $\tau : V \times \mathbb{A} \longrightarrow \mathbb{A}$ durch

$$\tau(v, A) := \varphi_A^{-1}(v).$$

Anders gesagt: $\tau(v, A) = A'$, wobei A' der einzige Punkt mit $\overrightarrow{AA'} = v$ ist.
Transitivität: $\forall A, B \in \mathbb{A}, \quad \exists\, \overrightarrow{AB}$ so dass $A + \overrightarrow{AB} = B$
Freiheit: $A + v = A + w \Rightarrow v = w$ weinφ_A wohl-definiert und bijektiv ist.
$\boxed{\Leftarrow}$ Sei $\tau : \mathbb{A} \times V \longrightarrow \mathbb{A}$ eine transitive + freie Gruppenwirkung. Dann gibt $\varphi : \mathbb{A} \times \mathbb{A} \longrightarrow \mathbb{A}$,

$$\varphi(A, B) = \tau(A, -)^{-1}(B), \text{ wobei } \tau(A, -) : V \longrightarrow \mathbb{A}$$

die affiner-Raum-Struktur.
$\boxed{\text{(AR 1)}}$ weil $\tau(\varphi(B,C)+\varphi(A,B), A) = \tau(\varphi(B,C), \tau(\varphi(A,B), A)) = \tau(\varphi(B,C), B) = C$. $\boxed{\text{(AR 2)}}$
weil $\varphi(A, -) = \tau(A, -)^{-1} : \mathbb{A} \longrightarrow V$ + Bemerkung 14.9. Q.E.D.

Mit der Bezeichnug $\overrightarrow{OA} := \varphi(O, A)$ schreiben wir auch

$$\begin{aligned} \overrightarrow{OA} &= A - O \\ A &= O + \overrightarrow{OA} \end{aligned}$$

wobei mit dem +-Zeichen in der zweiten Gelichung die Gruppenwirkung gemeint wird.

14.1.2 Gewichtetes Baryzentrum

Seien $(\mathbb{A}, V, \varphi)$ ein affiner Raum über $\mathbb{K}$, $n \geq 1$ eine ganze Zahl, $(n+1)$ Punkte $P_0, \ldots, P_n \in \mathbb{A}$, und $(n+1)$ Skalare $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$ mit $\lambda_0 + \cdots + \lambda_n = 1$. Sei $O \in \mathbb{A}$ ein beliebiger Punkt. Wir definieren P als der einzige Punkt mit der Eigenschaft

$$\overrightarrow{OP} = \lambda_0 \overrightarrow{OP_0} + \cdots + l a_n \overrightarrow{OP_n}.$$

Dieser ist eindeutig bestimmt, weil ϕ_O bijektiv ist, und heißt das *gewichtete Baryzentrum* (oder der *gewichtete Schwerpunkt*) der Punkte $P_0, \ldots, P_n$ mit den Gewichten $\lambda_0, \ldots, \lambda_n$.

Lemma 14.11. *(i) Mit der obigen Notation, haben wir für jeden anderen Punkt O' gilt*

$$\overrightarrow{O'A} = \lambda_0 \overrightarrow{O'P_0} + \cdots + l a_n \overrightarrow{O'P_n},$$

insbesondere ist der Punkt A von der Wahl von O unabhängig.

(ii) Falls $\lambda_0 + \cdots + \lambda_n = 0$, dann ist der Vektor $\lambda_0 \overrightarrow{OP_0} + \cdots + \lambda_n \overrightarrow{OP_n}$ von der Wahl von O unabhängig.

Beweis-Skizze: (i) Aus (AR 1) + $\lambda_0 + \cdots + \lambda_n = 1$
(ii) Aus (AR 1) + $\lambda_0 + \cdots + \lambda_n = 0$
Q.E.D.

Definition 14.12. Sei $n \geq 0$ eine ganze Zahl, und seinen $P_0, \ldots, P_n \in \mathbb{A}$ Punkte und $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$ Skalare mit $\lambda_0 + \cdots + \lambda_n = 1$. Wir definieren $\lambda_0 P_0 + \cdots + \lambda_n P_m$ als der einzige Punkt $P \in \mathbb{A}$, so dass es ein Punkt $O \in \mathbb{A}$ gibt für dem

$$\overrightarrow{OP} = \lambda_0 \overrightarrow{OP_0} + \cdots + \lambda_n \overrightarrow{OP_n}.$$

Dieser Punkt heißt das *gewichtete Baryzentrum* (oder der *gewichtete Schwerpunkt*) der Punkte $P_0, \ldots, P_n$ mit den Gewichten $\lambda_0, \ldots, \lambda_n$.

Die physikalische Interpretation: wenn jeder Punkt P_i die Masse λ_i hat, dann ist das Baryzentrum, der physikalische Schwerpunkt des Systems.
Für $n+1$ Skalare $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$ mit $\lambda_0 + \cdots + \lambda_n = 0$ schreiben wir, dass $\lambda_0 P_0 + \cdots + \lambda_n P_n = 0$ wenn es einen Punkt $O \in \mathbb{A}$ gibt, so dass $\lambda_0 \overrightarrow{OP_0} + \cdots + \lambda_n \overrightarrow{OP_n} = 0_V$.

Beispiel 14.13.
1. Wenn $\mathrm{char}\, \mathbb{K} \nmid (n+1)$, können wir $\lambda_0 = \cdots = \lambda_n = \frac{1}{n+1}$ wählen, und dann heißt $\sum_i \lambda_i P_i$ einfach *das Baryzentrum* (oder *der Schwepunkt*) von $P_0, \ldots, P_n$.
2. Wenn $n = 1$ und $\mathrm{char}\, \mathbb{K} \neq 2$, dann ist $\frac{1}{2}P_0 + \frac{1}{2}P_1$ die Mitte von P_0 und P_1.
3. Wenn $n = 1$ und $\mathrm{char}\, \mathbb{K} \neq 2$, dann ist $Q := (-1)P_0 + 2P_1$ der symmetrische Punkt von P_0 bezüglich P_1.

14.1.3 Affine Unterräume

Definition 14.14. Sei $(\mathbb{A}, V, \varphi)$ ein affiner Raum über $\mathbb{K}$. Eine Teilmenge $U \subseteq \mathbb{A}$ heißt *affiner $\mathbb{K}$-Unterraum* von $\mathbb{A}$ wenn entweder $\mathbb{A} = \emptyset$ oder wenn es einen Punkt $O \in U$ gibt so dass $\varphi_O(U) = \{\overrightarrow{OP} \mid P \in U\}$ ein $\mathbb{K}$-Untervektorraum von V ist.

Bemerkung 14.15. Sei U ein affiner $\mathbb{K}$-Unterraum von $\mathbb{A}$, sei $W = \varphi_O(U)$ der $\mathbb{K}$-UVR aus Definition 14.14, und sei $O' \in U$ ein beliebiger Punkt. Dann gilt $\varphi_{O'}(U) = W$, und W heißt der *zugeordnete liniare Unterraum* (oder die *Richtung*) von U.

Beweis-Skizze: $\varphi_{O'}(U) = \{\overrightarrow{O'O} + \overrightarrow{OP} \mid P \in U\} = \overrightarrow{O'O} + W = W$, wo die letzte Gleichung folgt weil $\overrightarrow{O'O} \in W$. Q.E.D.

Satz 14.16. *Sei $(\mathbb{A}, V, \varphi)$ ein affiner Raum über $\mathbb{K}$, und $U \subseteq \mathbb{A}$ eine nichtleere Teilmenge. Folgende Aussagen sind äquivalent:*

(i) *$U \subseteq \mathbb{A}$ ist ein affiner $\mathbb{K}$-Unterraum*

(ii) *$\exists\, O \in U$ und $W \subseteq V$ ein $\mathbb{K}$-UVR so dass $U = O + W$.*

(iii) *Für jede endliche Teilmenge $\{P_0, \ldots, P_n\} \subseteq U$ und für alle Skalare $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$ mit $\lambda_0 + \cdots + \lambda_n = 1$ gilt $\lambda_0 P_0 + \cdots + \lambda_n P_n \in U$.*

Beweis-Skizze: (i)⇒(ii) Folgt aus der Definition 14.14 + Bemerkung 14.15.
(ii)⇒(iii) $\lambda_0 P_0 + \cdots + \lambda_n P_n = (\sum_{i=0}^n \lambda_i) O + \lambda_0 \overrightarrow{OP_0} + \cdots + \lambda_n \overrightarrow{OP_n} \in O + W = U$
(iii)⇒(i) Sei $O \in U$. Wir wollen zeigen, dass $W := \varphi_O(U)$ (UVR1) und (UVR2) erfüllt.
Es seien also $w_1, w_2 \in W$, also $\exists\ P_1, P_2 \in U$ mit $w_i = \overrightarrow{OP_i}$ für $i = 1, 2$. Wir setzen noch $P_0 := O$, $\lambda_1 = \lambda_2 = 1$ und $\lambda_0 = -1$ (also $\lambda_0 + \lambda_1 + \lambda_2 = 1$). Aus (iii) folgt also $P = \lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2 = P_1 + P_2 - O \in U$, und dann auch $\overrightarrow{OP} = \overrightarrow{OP_1} + \overrightarrow{OP_2} - \overrightarrow{OO} = w_1 + w_2 \in W$ (UVR1).
Sei $\lambda \in \mathbb{K}$ und $w = \overrightarrow{OP_1} \in W$. Aus (iii) folgt $Q := (1-\lambda)O + \lambda P \in U$, also $\lambda w = (1-\lambda)\overrightarrow{OO} + \lambda\overrightarrow{OP} = \overrightarrow{OQ} \in W$ (UVR2). Q.E.D.

Korollar 14.17. Sei $\{U_i\}_{i \in I}$ eine nicht-leere Familie affiner $\mathbb{K}$-UR von $(\mathbb{A}, V, \mathbb{K})$. Dann ist $\cap_{i \in I} U_i$ ein affiner $\mathbb{K}$-UR (eventuell leer) von $(\mathbb{A}, V, \mathbb{K})$.

Definition 14.18. Die *Dimension* eines nicht-leeren affinen UR $U = O + W$ ist die Dimension von W als $\mathbb{K}$-VR. Wenn U ist der leere affine UR, dann ist die Konvention, dass dessen Dimension -1 ist.

Bemerkung 14.19. Wenn $(\mathbb{A}, V, \mathbb{K})$ ein endlich-dimensionaler affiner Raum und $U \subseteq \mathbb{A}$ ein affiner UR ist, dann gilt $\dim(U) \leq \dim(\mathbb{A})$, und wenn $\dim(U) = \dim(\mathbb{A})$, dann $U = \mathbb{A}$.

Beispiel 14.20.
1. Punkte von $\mathbb{A}$ sind 0-dimensionale UR.
2. Wenn $U \subseteq \mathbb{A}^n(\mathbb{K})$, dann ist U ein UR $\iff \exists W \subseteq \mathbb{K}^n$ ein UVR und $\exists\ v \in \mathbb{K}^n$ mit $U = v + W := \{v + w \mid w \in W\}$.

Definition 14.21. Zwei affine UR $U, U' \subseteq \mathbb{A}$ heißen *parallel* wenn sie die selbe Richtung haben. D.h. wenn es $W \subseteq V$, und $O, O' \in \mathbb{A}$ gibt mit $U = O + W$ und $U' = O' + W$.

Wir sollten das aber allgemeiner definieren, als $W \subseteq W'$ oder umgekehrt. Dann muss man in der folgenden Bemerkung hinzufügen, dass die Dimensionen gleich sein sollen.

Bemerkung 14.22. Affine Räume erfüllen das Parallelenaxiom (oder Axiom von Playfair): zu jedem aUR U und jedem Punkt P außerhalb von U gibt es genau ein aUR U' der P enthält und parallel zu U ist.

Definition 14.23. Sei $S \subseteq \mathbb{A}$ eine Menge von Punkten im affinen Raum $(\mathbb{A}, V, \varphi)$ über $\mathbb{K}$. Die affine Hülle von S (oder der von S-erzeugte affine $\mathbb{K}$-UR) ist

$$\mathrm{Aff}(S) = \bigcap_{S \subseteq U} U, \qquad \text{wobei alle } U \text{ affine UR sind.}$$

Aus Korollar 14.17 folgt, dassAff(S) der kleinste affine UR der S enthält ist. Wenn $S = \emptyset$ dann $\mathrm{Aff}(S) = \emptyset$.

Satz 14.24. *Sei $\{P_0, \ldots, P_m\} \subseteq \mathbb{A}$ eine endliche Menge von Punkten im affinen Raum $(\mathbb{A}, V, \varphi)$. Dann*

$$\mathrm{Aff}(P_0, \ldots, P_m) = \left\{ \sum_{i=0}^m \lambda_i P_i \mid \lambda_i \in \mathbb{K}\ \forall\ 1 \leq i \leq m \text{ und } \sum_{i=0}^m \lambda_i = 1 \right\}.$$

Außerdem, ist die Richtung von $\mathrm{Aff}(P_0, \ldots, P_n)$ der $\mathbb{K}$-VR $\mathrm{Span}\, \overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_m}$. Insbesondere, haben wir $\dim \mathrm{Aff}(P_0, \ldots, P_m) \leq m$.

Beweis-Skizze: Bezeichnen wir mit $U := \{\sum_{i=0}^m \lambda_i P_i \mid \lambda_i \in \mathbb{K} \; \forall \; 1 \leq i \leq m \text{ und } \sum_{i=0}^m \lambda_i = 1\}$. Wir wollen also $\mathrm{Aff}(P_0, \ldots, P_m) = U$.
$\boxed{\supseteq}$ Folgt aus Satz 14.16.
$\boxed{\subseteq}$ Weil $P_i \in U$ für alle i, reicht es z.z., dass U ein affiner UR ist. Wir zeigen, dass $U = P_0 + W$, wobei $W = \mathrm{Span}\, \overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_m}$. Wenn $P \in U$, dann ist per Definition $P = P_0 + (\lambda_1 \overrightarrow{P_0P_1} + \cdots + \lambda_m \overrightarrow{P_0P_m})$. Umgekehrt, sei $w \in W$, dann ist $w = \lambda_1 \overrightarrow{P_0P_1} + \cdots + \lambda_m \overrightarrow{P_0P_m}$. Wenn wir $\lambda_0 = 1 - \sum_{i=1}^m \lambda_i$ setzen, dann ist $P = P_0 + w = \sum_{i=0}^m \lambda_i P_i$. Q.E.D.

14.1.4 Affine Unabhängigkeit

Definition 14.25. Sei $(\mathbb{A}, V, \varphi)$ ein affiner Raum über $\mathbb{K}$. Die Punkte $P_0, \ldots, P_m \in \mathbb{A}$ heissen *Punkte in allgemeiner Lage* (oder *affin unabhängig*), wenn die Vektoren $\{\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_m}\}$ linear unabhängig in V sind.

Satz 14.26. *Sei $\{P_0, \ldots, P_m\} \subseteq \mathbb{A}$ eine endliche Menge von Punkten im affinen Raum $(\mathbb{A}, V, \varphi)$. Folgende Aussagen sind äquivalent:*

(i) $P_0, \ldots, P_m$ *sind Punkte in algemeiner Lage*

(ii) $\dim \mathrm{Aff}(P_0, \ldots, P_m) = m$

(iii) Für alle $(\lambda_0, \ldots, \lambda_m) \in \mathbb{K}^{m+1}$ *mit* $\lambda_0 + \cdots + \lambda_m = 0$ *und* $\lambda_0 P_0 + \cdots + \lambda_m P_m = 0$ *folgt dass* $\lambda_0 = \cdots = \lambda_m = 0$.

(iv) Für jeder $P \in \mathrm{Aff}(P_0, \ldots, P_m)$ *existiert ein einziger* $(\lambda_0, \ldots, \lambda_m) \in \mathbb{K}^{m+1}$, *so dass* $\lambda_0 + \cdots + \lambda_m = 1$ *und* $P = \lambda_0 P_0 + \cdots + \lambda_m P_m$.

(v) Für jeder $i = 0, \ldots, m$ *gilt* $P_i \notin \mathrm{Aff}(P_0, \ldots, \hat{P_i}, \ldots, P_m)$.

Beweis-Skizze: $\boxed{\text{(i)}\Leftrightarrow\text{(ii)}}$ Folgt aus Definition 14.25 + Satz 14.24
$\boxed{\text{(i)}\Rightarrow\text{(iii)}}$ Seien $\lambda_0, \ldots, \lambda_m \in \mathbb{K}$ mit $\sum_i \lambda_i = 0$ und $\sum_i \lambda_i P_i = 0$. Wenn wir $O = P_0$ wählen, haben wir $\lambda_1 \overrightarrow{P_0P_1} + \cdots + \lambda_m \overrightarrow{P_0P_m} = 0_V$. Aus (i) $\Rightarrow \lambda_1 = \cdots = \lambda_m = 0$, und aus $\sum_i \lambda_i = 0$ folgt auch $\lambda_0 = 0$.
$\boxed{\text{(i)}\Leftarrow\text{(iii)}}$ Wenn $\lambda_1 \overrightarrow{P_0P_1} + \cdots + \lambda_m \overrightarrow{P_0P_m} = 0_v$, dann setzen wir $\lambda_0 = -(\lambda_1 + \cdots + \lambda_m)$ und aus *(iii)* folgt, dass alle null sind.
$\boxed{\text{(i)}\Rightarrow\text{(iv)}}$ Sei $P \in \mathrm{Aff}(P_0, \ldots, P_m)$ mit $P = \sum_i \lambda_i P_i = \sum_i \mu_i P_i$, wobei $\sum_i \lambda_i = \sum_i \mu_i = 1$. Dann $\sum_{i=1}^m (\lambda_i - \mu_i) \overrightarrow{P_0P_i} = O_V$, und aus (i) folgt $\lambda_i = \mu_i \; \forall \; i = 1, \ldots, m$, also auch $\lambda_0 = 1 - \lambda_1 - \cdots - \lambda_m = 1 - \mu_1 - \cdots - \mu_m = \mu_0$.
$\boxed{\text{(i)}\Leftarrow\text{(iv)}}$ Wenn $\lambda_1 \overrightarrow{P_0P_1} + \cdots + \lambda_m \overrightarrow{P_0P_m} = 0_V$, ist äquivalent zu $P_0 = \lambda_0 P_0 + \lambda_1 P_1 + \cdots + \lambda_m P_m$, wobei $\lambda_0 = 1 - \sum_{i=1}^m \lambda_i$. Aber $P_0 = 1P_0 + 0P_1 + \cdots + 0P_m$, also aus (iv) folgt, dass $\lambda_i = 0 \; \forall \; i = 1, \ldots, m$.
$\boxed{\text{(i)}\Leftrightarrow\text{(v)}}$ Für $m = 0$ ist das klar. Für $m \geq 1$, haben wir aus (i)$\Leftrightarrow$(ii), dass affine Unabhängigkeit nicht auf der Rheienfolge der Punkte ankommnt. Sei $i \in \{0, \ldots, m\}$. Dann sind $P_0, \ldots, P_m$ sind in allgemeiner Lage $\Longleftrightarrow$ $\overrightarrow{P_jP_0}, \ldots, \overrightarrow{P_jP_{j-1}}, \overrightarrow{P_jP_{j+1}}, \ldots, \overrightarrow{P_jP_m}$ für ein $j \in \{1, \ldots, m\} \setminus \{i\}$ linear

unabhängig sind. Alles folgt jetzt aus der ähnlichen Aussage für Vektoren $(v_1, \dots, v_m)$ sind l.u. $\Leftrightarrow$ $v_i \notin \operatorname{Span} v_1, \dots, \hat{v_i}, \dots, v_m$. Q.E.D.

Satz 14.27. *Sei $U \subseteq \mathbb{A}$ ein affiner UR von $(\mathbb{A}, V, \varphi)$, mit $\dim U = m \leq n = \dim \mathbb{A} < \infty$. Dann existieren $P_0, \dots, P_m \in U$ Punkte in allgemeiner Lage, so dass $U = \operatorname{Aff}(P_0, \dots, P_m)$.*

Beweis-Skizze: $U = P_0 + W$, und $W = \operatorname{Span} w_1, \dots, w_m$, dann setzen wir $P_i := \varphi_{P_0}^{-1}(w_i)$. Q.E.D.

Definition 14.28. Sei $(\mathbb{A}, V, \varphi)$ ein n-dimensionaler affiner Raum über $\mathbb{K}$. Ein $(n+1)$-Tupel von Punkten $(P_0, \dots, P_n)$ heißt *affiner n-Bein* wenn $\operatorname{Aff}(P_0, \dots, P_n) = \mathbb{A}$.

Bemerkung 14.29. Für ein affiner Raum $(\mathbb{A}, V, \varphi)$, ein n-Bein zu geben ist äquivalent zu ein Ursprung $O \in \mathbb{A}$ und eine $\mathbb{K}$-Basis von V zu geben.

Beweis-Skizze: $(P_0, \dots, P_n) \Rightarrow O = P_0$ und die $\mathbb{K}$-Basis ist $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n}$. Umgekehrt, wenn O und $v_1, \dots, v_n$ gegeben sind, dann ist $O, \varphi_O^{-1}(v_1), \dots, \varphi_O^{-1}(v_n)$ das n-Bein. Q.E.D.

Bemerkung 14.30. Aus Satz 14.26 folgt, dass ein affiner n-Bein aus Punkte in allgemeiner Lage besteht. Aus Satz 14.27 folgt, dass ein n-Bein immer existiert.

Beispiel 14.31. Für $\mathbb{A} = \mathbb{K}^n$: $E_0 = (0, \dots, 0)$ und $E_i = e_i$

14.1.5 Baryzentrische Koordinaten

Bemerkung 14.32. Aus Satz 14.26 Teil (iv) haben wir, dass wenn $P_0, \dots, P_n$ ein n-Bein von $\mathbb{A}$ ist, dann sind für jeder $P \in \mathbb{A}$ die Skalare $\lambda_0, \dots, \lambda_n \in \mathbb{K}$ mit $\sum \lambda_i = 1$ und $P = \sum \lambda_i P_i$ eindeutig bestimmt.

Definition 14.33.

1. Die Skalare aus Bemerkung 14.32 heißen *baryzentrische Koordinaten* von P bezüglich dem n-Bein $P_0, \dots, P_n$.
2. Die Koordinaten von $\overrightarrow{P_0P} \in V$ bezüglich der Basis $\{\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n}\}$ heißen *affine Koordinaten* von P (bez. dem n-Bein $P_0, \dots, P_n$).

Bemerkung 14.34. Wenn $(\lambda_0, \dots, \lambda_n)$ bar. Koord. sind $\Rightarrow$ $(\lambda_1, \dots, \lambda_n)$ sind die aff. Koord.
Wenn $(a_1, \dots, a_n)$ sind aff. Koord. $\Rightarrow$ $(1 - \sum_{i=1}^n a_i, a_1, \dots, a_n)$ sind die bar. Koord.

Beispiel 14.35. Seien $\mathbb{A}$ ein affiner Raum über $\mathbb{R}$. Dann haben wir folgendes Bild.
Warum brauche ich $\mathbb{R}$ hier? Was passiert wenn $\mathbb{K} = \mathbb{F}_2$?

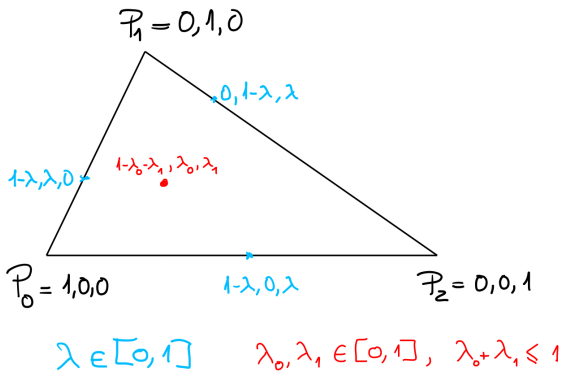
14.1.6 Das Verhältnis dreier kollineare Punkte

Seien $A, B \in \mathbb{A}$ zwei verschiedene Punkte. Dann sind A, B affin unabhängig. Wir bezeichnen mit

$$AB = \operatorname{Aff}(A, B) = \{(1-\lambda)A + \lambda B \mid \lambda \in \mathbb{K}\}$$

den 1-dimensionalen affiner UR, und nennen es *die Gerade* durch A und B. Umgekehrt, jeder 1-dimensionaler affiner UR U hat die Form AB (wähle $A, B \in U$ mit $A \neq B$, und wende Bemerkung 14.19 an).

Bemerkung 14.36. Zwei verschiedene Geraden schneiden sich höchstens in einem Punkt.



Beweis-Skizze: Wenn $A \neq B$ existieren mit $\{A, B\} \subseteq G_1 \cap G_2$, dann $G_1 = G_2 = AB$. Q.E.D.

Definition 14.37. Drei Punkte $A, B, C \in \mathbb{A}$ heißen *kollinear*, wenn zwei davon gleich sind, oder wenn alle 3 verschieden sind und $C \in AB$.

Übung: Diese Definition ist von der Ordnung der Punkte unabhängig, und 3 Punkte sind nicht kollinear wenn und nur wenn diese affin unabhängig sind.

Bemerkung 14.38. A, B, C sind kollinear $\iff \exists (a, b, c) \in \mathbb{K}^3 \setminus (0, 0, 0)$, mit $a + b + c = 0$ und $aA + bB + cC = 0$. Weiterhin, wenn die 3 Punkte verschieden sind, dann gilt auch $abc \neq 0$.

Beweis: Übung.

Definition 14.39. Sei V ein 1-dimensionaler $\mathbb{K}$ - VR, und $v, w \in V \setminus 0$. Der eindeutige Skalar $\lambda \in \mathbb{K}$ mit $v = \lambda w$ heißt *Proportionalitätsfaktor* zwischen v und w, und wird mit $\frac{v}{w}$ bezeichnet.

Definition 14.40. Das *Teilungsverhältnis* dreier verschiedener kollinearer Punkte A, B, C ist

$$\frac{\overrightarrow{AC}}{\overrightarrow{AB}}.$$

Bemerkung 14.41. Wenn A, B, C kollinear und verschieden sind, dann haben wir $C = (1-t)A + tB$, für $t \in \mathbb{K}$, also

$$\frac{\overrightarrow{AC}}{\overrightarrow{AB}} := t \in \mathbb{K}.$$

Da $C \neq A, B$ (das steht für "$C \neq A$ und $C \neq B$"), dann haben wir $t \neq 1$, und auch $A = \frac{t}{t-1}B + \frac{-1}{t-1}C$, also $\overrightarrow{CA} = \frac{t}{t-1}\overrightarrow{CB}$ und wir haben das Verhältnis:

$$\frac{\overrightarrow{CA}}{\overrightarrow{CB}} := \frac{t}{t-1} \in \mathbb{K}.$$

Beispiel 14.42. Wenn C die Mitte von A und B ist, dann ist $C = \frac{1}{2}A + \frac{1}{2}B$, und also

$$\frac{\overrightarrow{AC}}{\overrightarrow{AB}} = \frac{1}{2} \qquad \text{und} \qquad \frac{\overrightarrow{CA}}{\overrightarrow{CB}} = -1.$$

Verhältnisse können also auch negativ sein! Und

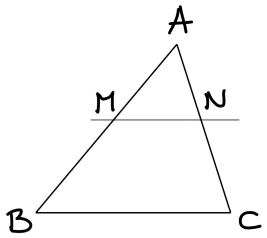
$$\frac{\overrightarrow{CA}}{\overrightarrow{CB}} = -\frac{\overrightarrow{CA}}{\overrightarrow{BC}}$$

$\frac{\overrightarrow{BC}}{\overrightarrow{AB}} = t - 1.$

Mit *Dreieck in* $\mathbb{A}$ verstehen wir (im Moment) 3 Punkte A, B, C in allgemeiner Lage. Um über Dreiecke in $\mathbb{A}$ zu sprechen, muss also $\dim \mathbb{A} > 1$.
Auf dem Übungsblatt 2 haben wir gesehen, dass zwei Geraden g_1, g_2 in der Ebene sind parallel (wir schreiben $g_1 \parallel g_2$) wenn und nur wenn $g_1 = g_2$ oder $g_1 \cap g_2 = \emptyset$.

Satz 14.43 (Strahlensatz/des Thales 600 v.Chr.)**.** *Sei ABC ein Dreieck in $\mathbb{A}^2(\mathbb{K})$, und seien $M \in AB \setminus \{A, B\}$ und $N \in AC \setminus \{A, C\}$. Dann gilt*

$$MN \parallel BC \iff \frac{\overrightarrow{MA}}{\overrightarrow{MB}} = \frac{\overrightarrow{NA}}{\overrightarrow{NC}}.$$



Beweis-Skizze: Seien $t, s \in \mathbb{K} \setminus \{0, 1\}$, so dass

$$M = (1-t)A + tB \qquad N = (1-s)A + sC \tag{14.1}$$

Sei $X \in MN$, das heißt $\exists\, u \in \mathbb{K}$ mit $X = (1-u)M + uN$. Wenn wir (14.1) einsetzen, bekommen wir

$$X = (1 - t + u(t-s))A + t(1-u)B + usC,$$

Also $X \in BC \iff$ der Koeffizient von A null ist. Da $MN \neq BC$ folgt also

$$MN \parallel BC \iff (1 - t + u(t-s)) \neq 0, \ \ \forall u \in K \iff t = s \iff \frac{\overrightarrow{MA}}{\overrightarrow{MB}} = \frac{\overrightarrow{NA}}{\overrightarrow{NC}}.$$

Q.E.D.

Bemerkung 14.44. Der Strahlensatz gibt uns auch

$$\frac{\overrightarrow{AM}}{\overrightarrow{AB}} = \frac{\overrightarrow{AN}}{\overrightarrow{AC}} \qquad \frac{\overrightarrow{BM}}{\overrightarrow{AB}} = \frac{\overrightarrow{CN}}{\overrightarrow{AC}} \qquad \frac{\overrightarrow{MN}}{\overrightarrow{BC}} = \frac{\overrightarrow{AN}}{\overrightarrow{AC}}$$

Beweis-Skizze: $\overrightarrow{MN} = \overrightarrow{AN} - \overrightarrow{AM}$... Q.E.D.

Satz 14.45 (des Menelaos 100.n.Chr)**.** *Sei ABC ein Dreieck in $\mathbb{A}^n(\mathbb{K})$ ($n \geq 2$), und seien M, N, P drei Punkte mit $\{M, N, P\} \cap \{A, B, C\} = \emptyset$. Wir nehmen an dass $M \in BC$, $N \in AC$ und $P \in AB$. Dann gilt*

$$M, N, P \text{ sind kollinear} \iff \frac{\overrightarrow{MB}}{\overrightarrow{MC}} \cdot \frac{\overrightarrow{NC}}{\overrightarrow{NA}} \cdot \frac{\overrightarrow{PA}}{\overrightarrow{PB}} = 1.$$

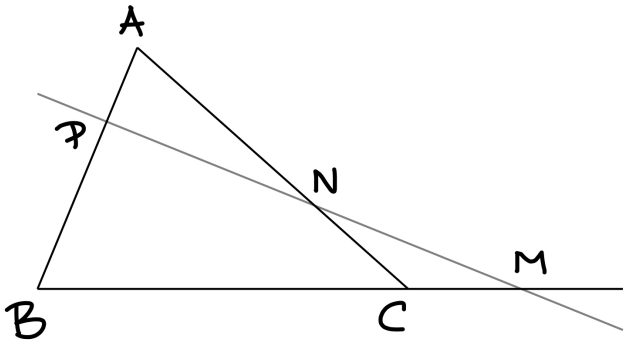


Abbildung 14.1: zur Datz des Menelaos

Beweis-Skizze: Seinen

$$M = (1 - t_M)B + t_M C \qquad N = (1 - t_N)C + t_N A \qquad P = (1 - t_P)A + t_P B \tag{14.2}$$

dann haben wir

$$\frac{\overrightarrow{MB}}{\overrightarrow{MC}} \cdot \frac{\overrightarrow{NC}}{\overrightarrow{NA}} \cdot \frac{\overrightarrow{PA}}{\overrightarrow{PB}} = \frac{t_M t_N t_P}{(t_M - 1)(t_N - 1)(t_P - 1)} \tag{14.3}$$

$\boxed{\Rightarrow}$ da M, N, P kollinear und verschieden sind, folgt aus Bemerekung 14.38, dass es $a, b, c \in \mathbb{K}$ mit $a + b + c = 0$ und $abc \neq 0$ gibt, so dass $aM + bN + cP = 0$. Wenn wir jetzt (14.2) einsetzen, bekommen wir

$$(bt_N + c(1 - t_P))A + (ct_P + a(1 - t_M))B + (at_M + b(1 - t_N))C = 0.$$

Die Summe der Skalare in der obigen Gleichung ist null, also, weil A, B, C nicht kollinear sind folgt aus Bemerkung 14.38, dass die 3 Skalare null sein müssen. Also

$$bt_N + c(1 - t_P) = 0 \iff \frac{t_N}{t_P - 1} = \frac{c}{b} \qquad \text{usw.}$$

und wir bekommen

$$\frac{\overrightarrow{MB}}{\overrightarrow{MC}} \cdot \frac{\overrightarrow{NC}}{\overrightarrow{NA}} \cdot \frac{\overrightarrow{PA}}{\overrightarrow{PB}} = \frac{t_M t_N t_P}{(t_M - 1)(t_N - 1)(t_P - 1)} = \frac{c}{b}\frac{a}{c}\frac{b}{a} = 1$$

$\boxed{\Leftarrow}$ Wir suchen $(a,b,c) \in \mathbb{K}^3 \setminus (0,0,0)$ mit $a+b+c=0$, so dass $aM+bN+cP=0$. Wir nehmen

$$a=1 \qquad b=\frac{t_M}{t_N-1} \qquad c=\frac{t_M-1}{t_P}.$$

Aus der Voraussetzung, dass (14.3)$=1$ folgt, dass $a+b+c=0$, und wir haben auch

$$\begin{aligned} aM+bN+cP &= M+\frac{t_M}{t_N-1}N+\frac{t_M-1}{t_P}P \\ \text{(aus (14.2))} &= (1-t_M)B+t_MC-t_MC+\frac{t_Mt_N}{t_N-1}A+\frac{(1-t_P)(t_M-1)}{t_P}A+(t_M-1)B \\ \text{(aus (14.3) =1)} &= \frac{t_Mt_N}{t_N-1}A-\frac{t_Mt_N}{t_N-1}A \\ &= 0 \end{aligned}$$

Q.E.D.

Die Geraden $G_1,\dots,G_i$ heißen *konkurrent* wenn es ein gemeinsamen Schnittpunkt gibt.

Satz 14.46 (von Giovanni Ceva 1678)**.** *Sei ABC ein Dreieck und seien A',B',C' drei Punkte mit $A \in BC\setminus\{B,C\}$,$B \in AC\setminus\{A,C\}$, und $C \in AB\setminus\{A,B\}$. Dann sind die Geraden AA', BB' und CC' konkurrent oder je zwei parallel wenn und nur wenn*

$$\frac{\overrightarrow{A'B}}{\overrightarrow{CA'}}\cdot\frac{\overrightarrow{B'C}}{\overrightarrow{AB'}}\cdot\frac{\overrightarrow{C'A}}{\overrightarrow{BC'}}=1 \tag{14.4}$$

Beweis-Skizze: $\boxed{\Rightarrow}$ **Fall 1:** $AA'\cap BB'\cap CC'=P$. Aus dem Satz des Menelaos für den Dreieck ABA' und der Geraden $C'P\in C$ haben wir

$$\frac{\overrightarrow{CB}}{\overrightarrow{CA'}}\cdot\frac{\overrightarrow{PA'}}{\overrightarrow{PA}}\cdot\frac{\overrightarrow{C'A}}{\overrightarrow{C'B}}=1$$

Aus dem Satz des Menelaos für den Dreieck $AA'C$ und der Geraden $B'P\in B$ haben wir

$$\frac{\overrightarrow{BA'}}{\overrightarrow{BC}}\cdot\frac{\overrightarrow{B'C}}{\overrightarrow{B'A}}\cdot\frac{\overrightarrow{PA}}{\overrightarrow{PA'}}=1$$

Und wir multiplizieren die beiden.
Fall 2. $AA'\parallel BB'\parallel CC'$. Aus dem Strahlensatz für das Dreieck CBB' mit der Geraden $A'A$, und dem Strahlensatz für das Dreieck BCC' mit der Geraden $A'A$ haben wir

$$\frac{\overrightarrow{AB'}}{\overrightarrow{CB'}}=\frac{\overrightarrow{A'B}}{\overrightarrow{CB}} \qquad \text{und} \qquad \frac{\overrightarrow{AC'}}{\overrightarrow{BC'}}=\frac{\overrightarrow{A'C}}{\overrightarrow{BC}}.$$

Also, nachdem wir die Rechte Gleichung invertieren, bekommen wir

$$\frac{\overrightarrow{AB'}}{\overrightarrow{CB'}}\cdot\frac{\overrightarrow{BC'}}{\overrightarrow{AC'}}=\frac{\overrightarrow{A'B}}{\overrightarrow{CB}}\cdot\frac{\overrightarrow{BC}}{\overrightarrow{A'C}}=-\frac{\overrightarrow{A'B}}{\overrightarrow{A'C}}=\frac{\overrightarrow{A'B}}{\overrightarrow{CA'}}.$$

$\boxed{\Leftarrow}$ **Fall 1:** $AA' \parallel BB'$. Aus dem Strahlensatz, für das Dreieck CBB' mit der Geraden AA', dass

$$\frac{\overrightarrow{A'C}}{\overrightarrow{A'B}} = \frac{\overrightarrow{AC}}{\overrightarrow{AB'}}.$$

Wenn wir das in (14.4) einsetzen bekommen wir

$$\frac{\overrightarrow{AB'}}{\overrightarrow{CA}} \cdot \frac{\overrightarrow{B'C}}{\overrightarrow{AB'}} \cdot \frac{\overrightarrow{C'A}}{\overrightarrow{BC'}} = 1 \qquad \Rightarrow \qquad \frac{\overrightarrow{C'A}}{\overrightarrow{C'B}} = \frac{\overrightarrow{CA}}{\overrightarrow{CB'}}$$

Also, wieder aus dem Strahlensatz folgt $BB' \parallel CC'$.
Fall 2: $AA' \cap BB' = P$. Die Geraden CP und AB können nicht parallel sein. Wären diese parallel, dann hätten wir

$$\frac{\overrightarrow{AB}}{\overrightarrow{CP}} = \frac{BA'}{A'C} \qquad \frac{\overrightarrow{CB'}}{\overrightarrow{B'A}} = \frac{\overrightarrow{CP}}{\overrightarrow{AB}},$$

und aus (14.4) auch $\frac{\overrightarrow{C'A}}{\overrightarrow{BC'}} = 1$ - ein Widerspruch. Es gibt also ein Schnittpunkt $C'' = CP \cap AB$, und aus dem $\boxed{\Rightarrow}$ Teil folgt

$$\frac{\overrightarrow{A'B}}{\overrightarrow{CA'}} \cdot \frac{\overrightarrow{B'C}}{\overrightarrow{AB'}} \cdot \frac{\overrightarrow{C''A}}{\overrightarrow{BC''}} = 1$$

und aus (14.4) bekommen wir $\frac{\overrightarrow{C''A}}{\overrightarrow{BC''}} = \frac{\overrightarrow{C'A}}{\overrightarrow{BC'}}$ und also $C' = C''$. Q.E.D.

14.2 Affine Abbildungen

Die Idee ist, dass Abbildungen zwischen affine Räume die affine Struktur bewahren sollen. Einerseits, werden wir das genau mit Hilfe der linearen Algebra ausdrücken, anderseits sollten wir uns diese als Abbildungen die Begriffe wie Unterraum, Kollinearität, Parallelität, Teilverhältnisse (wann das noch Sinn macht) bewahren.

Definition 14.47. Seien $(\mathbb{A}, V, \varphi)$ und $(\mathbb{B}, W, \psi)$ zwei affine Räume über $\mathbb{K}$, sei $f : \mathbb{A} \longrightarrow \mathbb{B}$ eine beliebige Abbildung, und $O \in \mathbb{A}$ ein Punkt. Die Abbildung

$$\mathrm{Tr}_O f := \psi_{f(O)} \circ f \circ \varphi_o^{-1} : V \longrightarrow W$$

heißt *Spurabbildung von f bezüglich O*. Also $\mathrm{Tr}_o f(\overrightarrow{OA}) = \overrightarrow{f(O)f(A)}$.

Definition 14.48. Eine Abbildung $f : \mathbb{A} \longrightarrow \mathbb{B}$ heißt *affine Abbildung* wenn es ein Punkt O gibt, so dass $\mathrm{Tr}_O(f)$ eine lineare Abbildung ist.

Insbesondere, für $\lambda \in \mathbb{K}$ und $v, w \in V$ haben wir

$$\begin{aligned} \mathrm{Tr}_O f(v+w) &= \mathrm{Tr}_O f(v) + \mathrm{Tr}_O f(w) \\ \mathrm{Tr}_O f(\lambda v) &= \lambda \,\mathrm{Tr}_O f(v) \end{aligned}$$

Lemma 14.49. *Mit der Voraussetzungen aus Definition 14.48 haben wir für jeden anderen Punkt $O' \in \mathbb{A}$*

$$\mathrm{Tr}_O f = \mathrm{Tr}_{O'} f.$$

Beweis-Skizze:

$$\begin{aligned}\overrightarrow{f(O)f(O')} + \overrightarrow{f(O')f(A)} &= \overrightarrow{f(O)f(A)} = \mathrm{Tr}_O\, f(\overrightarrow{OA}) = \mathrm{Tr}_O\, f(\overrightarrow{OO'} + \overrightarrow{O'A}) = \\ = \mathrm{Tr}_O\, f(\overrightarrow{OO'}) + \mathrm{Tr}_O\, f(\overrightarrow{O'A}) &= \overrightarrow{f(O)f(O')} + \mathrm{Tr}_O\, f(\overrightarrow{O'A}).\end{aligned}$$

Wir kürzen $\overrightarrow{f(O)f(O')}$ ab, und bekommen $\mathrm{Tr}_O\, f(A) = \overrightarrow{f(O')f(A)} = \mathrm{Tr}_{O'}\, f(A)$, $\forall\, A \in \mathbb{A}$. Q.E.D.

Wir bezeichnen also nur mit $\mathrm{Tr}\, f := \mathrm{Tr}_O\, f$ für irgendwelcher $O \in \mathbb{A}$, und nennen diese lineare Abbildung eifach *Spur von* f.

Satz 14.50. *Für jeder* $i = 1, 2, 3$ *sei* $\mathbb{A}_i$ *ein affiner Raum über* $\mathbb{K}$.

(i) $\mathrm{id}_{\mathbb{A}_1} : \mathbb{A}_1 \longrightarrow \mathbb{A}_1$ *ist eine affine Abbildung*

(ii) Wenn $f_i : \mathbb{A}_i \longrightarrow \mathbb{A}_{i+1}$ *affine Abbildungen für* $i = 1, 2$ *sind, dann ist auch* $f_2 \circ f_1 : \mathbb{A}_1 \longrightarrow \mathbb{A}_3$ *eine affine Abbildung.*

Beweis-Skizze: Übung Q.E.D.

Definition 14.51. Eine affine Abbildung $f : \mathbb{A} \longrightarrow \mathbb{B}$ heißt *affiner* $(\mathbb{K})$*-Isomorphismus* wenn es eine affine Abbildung $g : \mathbb{B} \longrightarrow \mathbb{A}$ gibt, so dass $g \circ f = \mathrm{id}_{\mathbb{A}}$ und $f \circ g = \mathrm{id}_{\mathbb{B}}$.

Definition 14.52. Ein affiner $\mathbb{K}$-Isomorphismus $f : \mathbb{A} \longrightarrow \mathbb{A}$ heißt *affiner* $\mathbb{K}$*-Automorphismus* von $\mathbb{A}$. Die Menge aller affinen $\mathbb{K}$-Automorphismen bildet eine Gruppe die wir mit $\mathrm{Aut}_{\mathbb{K}}(\mathbb{A})$ bezeichnen.

Satz 14.53. *Sei* $f : \mathbb{A} \longrightarrow \mathbb{B}$ *eine affine Abbildung. Dann haben wir*

$$f \text{ ist ein affiner Isomorphismus} \iff \mathrm{Tr}\, f \text{ ist ein VR Isomorphismus} \iff f \text{ ist bijektiv.}$$

Beweis-Skizze: Übung Q.E.D.

Satz 14.54. *Seien* $\mathbb{A}, \mathbb{B}$ *affine Räume über* $\mathbb{K}$ *und* $f : \mathbb{A} \longrightarrow \mathbb{B}$ *eine Abbildung. Dann ist* f *eine affine Abbildung wenn und nur wenn für jede endliche Menge von Punkten* $P_0, \ldots, P_m \in \mathbb{A}$ *und Skalare* $\lambda_0, \ldots, \lambda_m \in \mathbb{K}$ *mit* $\sum \lambda_i = 1$ *gilt*

$$f(\lambda_0 P_0 + \cdots + \lambda_m P_m) = \lambda f(P_0) + \cdots + \lambda f(P_m).$$

Beweis-Skizze: Sei $O = P_0$, $P = \sum_{i=0}^m \lambda_i P_i$ und $Q = \sum_{i=0}^m \lambda_i f(P_i)$.

$$\begin{aligned} f(P) = Q &\iff \overrightarrow{f(O)f(P)} = \overrightarrow{f(O)Q} \\ &\iff \mathrm{Tr}_O\, f(\overrightarrow{OP}) = \overrightarrow{f(O)Q} \\ &\iff \mathrm{Tr}_O\, f(\sum_i \lambda_i \overrightarrow{OP_i}) = \sum_{i=1}^m \lambda_i \overrightarrow{f(O)f(P_i)} \\ &\iff \mathrm{Tr}_O\, f(\sum_i \lambda_i \overrightarrow{OP_i}) = \sum_{i=1}^m \lambda_i\, \mathrm{Tr}_O\, f(\overrightarrow{OP_i}) \\ &\iff \mathrm{Tr}_O\, f \text{ ist linear.} \end{aligned}$$

Q.E.D.

Korollar 14.55. Wenn $P_0, \ldots, P_n$ ein affiner $n-$Bein von $\mathbb{A}$ ist, und wenn $f \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{A})$, dann ist $\{f(P_0), \ldots, f(P_n)\}$ ein affiner $n-$Bein.

Satz 14.56. *Sei $\mathbb{A}$ ein n-dimensionaler affiner Raum über $\mathbb{K}$, und seien $\mathcal{P} = \{P_0, \ldots, P_m\}$ und $\mathcal{Q} = \{Q_0, \ldots, Q_m\}$ zwei Mengen von affin-unabhängigen Punkten. Dann existiert $f \in \mathrm{Aut}_{\mathbb{K}} \mathbb{A}$ mit*

$$f(P_i) = Q_i \forall\ i = 0, \ldots, m.$$

Wenn $m = n$, dann ist f eindeutig bestimmt.

Beweis-Skizze: Wir können annehmen, dass $m = n$, sonst, wenn $m < n$, ergänzen wir beliebig $\mathcal{P}$ und $\mathcal{Q}$ zu affine n-Beine. Jeder Punkt $P \in \mathbb{A}$ schreibt man eindeutig als $P = \sum_{i=0}^{n} \lambda_i P_i$ mit $\sum \lambda_i = 1$. Dann, definieren wir $f : \mathbb{A} \longrightarrow \mathbb{A}$ durch $f(P) = \sum_i \lambda_i Q_i$. Insbesondere ist $f(P_i) = Q_i$, und aus Satz 14.54 ist es eine affine Abbildung. Da wir auch die Inverse ähnlicher Weise mit $f^{-1}(Q_i) = P_i$ definieren können, ist $f \in \mathrm{Aut}_{\mathbb{K}} \mathbb{A}$. Die Eindeutigkeit der λ_i gibt uns die Eindeutigkeit von f. (Wenn $m < n$ ist für jede Ergänzung zu einem n-Bein f eindeutig. Die Ergänzung ist aber nicht eindeutig.) Q.E.D.

Satz 14.57. *Seien $\mathbb{A}$ und $\mathbb{B}$ affine Räume über $\mathbb{K}$, mit* char $\mathbb{K} \neq 2$ *und sei $f : \mathbb{A} \longrightarrow \mathbb{B}$ eine Abbildung. Dann gilt*

$$f \text{ ist eine affine Abbildung } \iff f((1-t)A + tB) = (1-t)f(A) + tf(B), \forall\ A, B \in \mathbb{A} \text{ und } t \in \mathbb{K}.$$

Beweis-Skizze: Sonderfall von Satz 14.54 + Übungsaufgabe. Q.E.D.

Beispiel 14.58. Hier sind 3 wichtige Beispiele

1. Inklusion von affinen Unterräumen.
2. Wenn (A, V, φ) ein affiner Raum ist und $(\mathbb{A}(V), V, \psi)$ aus Beispiel 14.6.2,. Dann ist für jeder $O \in \mathbb{A}$ die Abbildung φ_O ein Isomorphismus, mit $\mathrm{Tr}\, \varphi_O = \mathrm{id}_V$. Deswegen darf man Sätze über affine Räume auf diesen Fall zurückführen. !Aber isomorph und gleich ist nicht dasselbe! Bei diesem Isomorphismus trifft man eine Wahl (O). Wenn man weiter zum Isomorphismus von V und $\mathbb{K}^n$ geht, dann wählt man auch eine Basis.
3. Sei $f : \mathbb{A} \longrightarrow \mathbb{B}$, wobei $(\mathbb{A}, V, \varphi)$ und $(\mathbb{B}, W, \psi)$ affine Räume sind. Dann ist $f' = \psi_{O'} \circ f \circ \varphi_O^{-1} : \mathbb{A}(V) \longrightarrow \mathbb{A}(W)$ auch eine affine Abbildung. Wir können also auch Aussagen über affine Abbildungen auf diesen Fall zurückführen.

Bemerkung 14.59. Sei $f : \mathbb{A}(V) \longrightarrow \mathbb{A}(W)$ mit V, W zwei $\mathbb{K}$-VR. Da $\varphi(v, v') = v' - v$ und $\psi(w, w') = w' - w$ die Abbildungen aus (V, V, φ) und (W, W, ψ) sind, wenn wir $O = 0_V$ wählen, und $O' := f(O)$, dann ist per Definition $\mathrm{Tr}\, f(A) = f(A) - f(0_V)$. Das heißt

$$f = f(0_V) + \mathrm{Tr}\, f.$$

Also f ist immer durch den Vektor $f(0_V)$ und die lineare Abbildung $\mathrm{Tr}\, f : V \longrightarrow W$ völlig bestimmt.

Wenn wir auch Basen für V und W wählen ($\dim_{\mathbb{K}} V = n, \dim_{\mathbb{K}} W = m$), $\operatorname{Tr} f$ als Matrix (a_{ij}) beschreiben und $f(0_V) = (b_1, \ldots, b_m)$ durch Koordinaten geben, dann haben wir folgenden Satz.

Satz 14.60. *Eine Abbildung $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ ist affin wenn und nur wenn es eine Matrix $A = (a_{ij}) \in M_{m\times n}(\mathbb{K})$ gibt, und $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{K}^m$ so das*

$$f(\mathbf{x}) = A \cdot \mathbf{x} + \mathbf{b}.$$

Für $\mathbb{A}^n(\mathbb{K})$, den affinen Standardraum über $\mathbb{K}$, bezeichnen wir die $\mathbb{K}$-Automorphismen von $\mathbb{A}^n(\mathbb{K})$ mit

$$\mathrm{AGL}_n(\mathbb{K}) := \mathrm{Aut}_{\mathbb{K}}(\mathbb{A}^n(\mathbb{K})).$$

Die natürliche Abbildung $\mathrm{Tr} : \mathrm{AGL}_n(\mathbb{K}) \longrightarrow \mathrm{GL}_n(\mathbb{K})$, die jeder f in $\operatorname{Tr} f$ schickt, ist ein surjektiver Gruppenhomomorphisums. Der Kern davon ist

$$\operatorname{Ker} \mathrm{Tr} = \{f \in \mathrm{AGL}_n(\mathbb{K}) \mid f(\mathbf{x}) = \mathbf{x} + \mathbf{b}, \quad \text{mit } \mathbf{b} \in \mathbb{K}^n.$$

Dieser ist also eine normale UG von $\mathrm{AGL}_n(\mathbb{K})$, die wir mit $T_n(\mathbb{K}) := \operatorname{Ker} \mathrm{Tr}$ bezeichnen. Für jeder Punkt $O \in \mathbb{A}^n(\mathbb{K})$, bezeichnen wir mit

$$\mathrm{AGL}_n(\mathbb{K}, O) := \{f \in \mathrm{AGL}_n(\mathbb{K}) \mid f(O) = O\}$$

die Untergruppe aller Automorphismen die O fix lassen. Dann haben wir, dass

$$\mathrm{Tr}\,|_{\mathrm{AGL}_n(K,O)} : \mathrm{AGL}_n(\mathbb{K}, O) \longrightarrow \mathrm{GL}_n(\mathbb{K})$$

ein Gruppenisomorphismus ist.

Satz 14.61. *Für jeder Punkt $O \in \mathbb{A}^n(\mathbb{K})$, die Gruppe $\mathrm{AGL}_n(\mathbb{K})$ ist das semidirekte Produkt von $T_n(\mathbb{K})$ und $\mathrm{AGL}_n(\mathbb{K}, O)$, d.h. für alle $f \in \mathrm{AGL}_n(\mathbb{K})$, existieren $\beta \in T_n(\mathbb{K})$ und $\alpha \in \mathrm{AGL}_n(\mathbb{K}, O)$ mit $f = \beta\circ\alpha$, und die Zerlegung ist eindeutig.*

Beweis-Skizze: $T_n(\mathbb{K})$ ist eine normale Untergruppe, da es ein Kern ist + Der Schnitt $\mathbb{T}_n(\mathbb{K}) \cap \mathrm{AGL}_n(\mathbb{K}, O) = \{id\}$ + für jedes f gibt es eine einzige Translation β mit $\beta(O) = f(O)$. Dann ist $\alpha = \beta^{-1} \circ f \in \mathrm{AGL}_n(\mathbb{K}, O)$, also $f = \beta \circ \alpha$. Q.E.D.

Beispiel 14.62. Projektionen

- Projektionen im ganz allgemeinem Fall sind Abbildungen $f : R \longrightarrow R$ mit $f^2 = f$, wobei R ein n-dimensionaler Raum ist (nicht unbedingt affin).
- e.g. Stereographische Projektion/ Ist die Kugel minus ein Punkt ein affiner Raum?
- Parallelprojektionen in $\mathbb{R}^2$ (entlang einer Geraden),im Raum $\mathbb{R}^3$ (entlang einer Geraden auf einer Ebene, oder entlang einer Ebene auf einer Gerade).
- "Zentralprojektion": $H_1 \parallel H_2$ zwei Ebenen in $\mathbb{A}^3(\mathbb{K})$, und $P \notin H_1 \cup H_2$. Wir definieren diese als $H_1 \ni QtoPQ \cap H_2$. Es ist also keine "echte" Projektion (im Sinne von Punkt 1 hier oben). Was passiert wenn $H_1 \nparallel H_2$? Ist diese die Einschränkung einer echten Projektion $\pi : \mathbb{A}^3(\mathbb{K}) \longrightarrow H_2$? (Teaser: projektive Räume)

14.3 Euklidische Affine Räume

In diesem Kapitel werden wir $\mathbb{K} = \mathbb{R}$ fixieren, also alle VR und AR werden über $\mathbb{R}$ sein.

Definition 14.63. Ein affiner Raum $(\mathbb{A}, V, \varphi)$ über $\mathbb{R}$ heißt *euklidisch* wenn $(V, \langle\ ,\ \rangle)$ ein Euklidischer VR ist. Wenn $\mathbb{A} = \mathbb{A}^n(\mathbb{R})$ und $\langle v, w\rangle = \sum v_i w_i$, dann heißt $\mathbb{A}^n(\mathbb{R})$ der *affine euklidische Standardraum* und wird mit $\mathbb{E}^n(\mathbb{R})$ (oder mit $\mathbb{E}^n_{\mathbb{R}}$, oder nur mit $\mathbb{E}^n$) bezeichnet. (In diesem Fall ist also $\varphi(v, w) = w - v$.)

Definition 14.64. Für zwei Punkte $A, B \in \mathbb{A}$ definieren die *Distanz* von A bis B

$$\mathrm{d}(A, B) := \left\| \overrightarrow{AB} \right\|.$$

Insbesondere, wenn $A, B \in \mathbb{E}^n(\mathbb{R})$, mit $A = (a_1, \ldots, a_n)$ und $B = (b_1, \ldots, b_n)$ dann ist

$$\mathrm{d}(A, B) = \sqrt{\sum (a_i - b_i)^2}$$

Bemerkung 14.65. Für alle $A, B, C \in \mathbb{A}$ haben wir

(MR1) $\mathrm{d}(A, B) \geq 0$ und $\mathrm{d}(A, B) = 0 \iff A = B$

(MR2) $\mathrm{d}(A, B) = \mathrm{d}(B, A)$

(MR3) $\mathrm{d}(A, C) \leq \mathrm{d}(A, B) + \mathrm{d}(B, C)$

Also $(\mathbb{A}, \mathrm{d})$ ist ein Metrischer Raum.

14.3.1 Unorientierte Winkel

Definition 14.66. Seien v, w zwei Vektoren in Euklidischen VR V, dann ist *der (unorientierte) Winkel* zwischen v und w

$$\angle(v, w) := \cos^{-1}\left(\frac{\langle v, w\rangle}{\|v\|\,\|w\|}\right).$$

Also $\theta \in [0, \pi]$.

Aus der Cauchy-Schwarz Ungleichung folgt, dass $\frac{\langle v,w\rangle}{\|v\|\|w\|} \in [-1, 1]$, also $\angle(v, w)$ ist richtig definiert. Wir sagen, dass zwei Vektoren v, w senkrecht zu einander sind, und schreiben $v \perp w$, wenn $\angle(v, w) = \frac{\pi}{2}$, also wenn $\langle v, w\rangle = 0$.

Satz 14.67 (Pythagoras)**.** *Seien O, A, B drei Punkte im Euklidischen affinen Raum $\mathbb{A}$. Die Vektoren $\overrightarrow{OA}$ und $\overrightarrow{OB}$ sind senkrecht zueinander wenn und nur wenn*

$$\left\| \overrightarrow{AB} \right\|^2 = \left\| \overrightarrow{OA} \right\|^2 + \left\| \overrightarrow{OB} \right\|^2.$$

Beweis-Skizze: Wenn $v = \overrightarrow{OA}$ und $w = \overrightarrow{OB}$, dann ist $\overrightarrow{AB} = w - v$, und

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

Q.E.D.

Seien F, G zwei Geraden in $\mathbb{E}^n$, also es gibt zwei Punkte $A, B \in \mathbb{E}^n$, und zwei Vektoren $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$, so dass

$$\begin{aligned} F &= A + \operatorname{Span} v \\ G &= B + \operatorname{Span} w \end{aligned}$$

Definition 14.68. Der Winkel zwischen F und G ist definiert als

$$\angle(F, G) := \cos^{-1}\left(\frac{|\langle v, w\rangle|}{\|v\|\,\|w\|}\right).$$

Also $\angle(F, G) \in [0, \frac{\pi}{2}]$. Zwei nicht unbedingt sich schneidende Geraden sind *senkrecht* zueinander wenn $v \perp w$.

Definition 14.69. Sei G eine Gerade in $\mathbb{E}^n$ und H eine Hyperebene. Wir sagen, dass $G \perp H$ wenn $G \perp F$ für alle Geraden $F \subseteq H$.

Bemerkung 14.70. Wenn $G = A + \operatorname{Span} v$ und $H = \{\mathbf{x} \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n = c\}$, mit $\mathbf{a} \neq 0$, dann

$$G \perp H \iff \exists \lambda \in \mathbb{R} \setminus \{0\} \text{ mit } a_i = \lambda v_i \;\forall\; i = 1, \ldots, n$$

Insbesondere, wenn $F \perp H$ und $G \perp H$ dann $F \parallel G$.

14.3.2 Entfernung zu einer Hyperenbene

Sei $H = \{\mathbf{x} \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n = c\}$ und $P = (p_1, \ldots, p_n) \in \mathbb{E}^n$, und sei G die Gerade durch P die senkrecht auf H ist ($G = P + \operatorname{Span} \mathbf{a}$). Sei $Q := G \cap H$. Aus Pythagoras Satz folgt, dass für alle $Q' \in H$ gilt $\mathrm{d}(P, Q) \leq \mathrm{d}(P, Q')$.

Definition 14.71. Mit der obigen Notation, ist *die Entfernung von P bis H*

$$\mathrm{d}(P, H) := \inf_{Q' \in H} \mathrm{d}(P, Q') = \mathrm{d}(P, Q).$$

Der Punkt Q heißt *Fußpunkt der Senkrechten durch P auf H*

14.3.3 Isometrien

Definition 14.72. Eine affiner Automorphismus $\sigma : \mathbb{E}^n \longrightarrow \mathbb{E}^n$ heißt *Isometrie* wenn

$$\mathrm{d}(A, B) = \mathrm{d}(\sigma(A), \sigma(B)), \quad \forall\; A, B \in \mathbb{E}^n.$$

Das ist äquivalent zu

$$\left\|\overrightarrow{AB}\right\| = \left\|\overrightarrow{\sigma(A)\sigma(B)}\right\| \quad \forall\; A, B \in \mathbb{E}^n.$$

Die Isometrien bilden eine Gruppe die wir mit $\mathrm{Aut}(\mathbb{E}^n(\mathbb{R}))$ bezeichnen.

Speziale Isometrien

Bemerkung 14.73. Für $\sigma \in \mathrm{AGL}_n(\mathbb{R})$, die durch $\sigma(\mathbf{x}) = A \cdot \mathbf{x} + \mathbf{c}$ gegeben ist, gilt

$$\sigma \in \mathrm{Aut}(\mathbb{E}^n) \iff A^T \cdot A = I_n \iff \mathrm{Tr}\,\sigma \in O(n).$$

Also $\sigma \in \mathrm{Aut}(\mathbb{E}^n) \Rightarrow \det \mathrm{Tr}\,\sigma \in \{-1, 1\}$.

Definition 14.74. Die Menge

$$\mathrm{SAut}(\mathbb{E}^n) := \{\sigma \in \mathrm{Aut}(E^n) \mid \det \mathrm{Tr}\,\sigma = 1\}$$

bildet eine Untergruppe von $\mathrm{Aut}(\mathbb{E}^n)$, die *Gruppe der Spezialen Isometrien* heißt.

Satz 14.75. *(i)* $\mathrm{SAut}(\mathbb{E}^n)$ *ist eine normale UG von* $\mathrm{Aut}(\mathbb{E}^n)$ *von Index 2.*

(ii) Symmetrien bezüglich Hyperebenen sind nicht spezial.

Beweis-Skizze: (i) - klar
(ii) Entweder konkret, oder:
Man gibt $H = \mathrm{Aff}(P_0, \ldots, P_{n-1})$ so dass $\overrightarrow{P_0P_i}$ eine orthonormale Basis ist, ergänzt diese zu einer orthonormalen Bassis von $\mathbb{E}^n$, und definiert die affine Abbildung τ mit $P_0 \mapsto 0$, und $P_i \mapsto e_i = (0, \ldots, 1, \ldots, 0)$. Diese schickt H in $\{\mathbf{x} \mid x_n = 0\}$, also $\mathrm{Tr}\,\tau^{-1} \circ \sigma \circ \tau$ ist durch die diagonale Matrix $(1, \ldots, 1, -1)$ gegeben, und hat also det $= -1$. Q.E.D.

14.3.4 Satz der 3 Senkrechten

Wir haben nur die Senkrechte von einem Punkt zu einer Hyperebene definiert. Man kann sehr einfach die Definition auf Senkrechte von einem Punkt P zu einem Unterraum $U \not\ni P$ definieren, in dem man in dem euklidischen Raum $\mathrm{Aff}(U \cup P)$ arbeitet, wo jetzt U eine Hyperebene ist (Übung: $\dim \mathrm{Aff}(U \cup P) = \dim U + 1$). Wenn $P \notin U$ also können wir auch über Fußpunkt der Senkrechten durch P auf U sprechen.

Bemerkung 14.76. Per Definition ist $Q \in U$ der Fußpunkt der Senkrechten durch P auf U wenn und nur wenn

$$\overrightarrow{PQ} \in \mathcal{D}(U)^\perp$$

Wenn es klar ist, dass $Q \in U$, dann schreiben wir einfach $PQ \perp U$, um zu sagen, dass Q der Fußpunkt der Senkrechten durch P auf U.

Satz 14.77 (Satz der 3 Senkrechten)**.** *Seien $T \subset U$ zwei affine Unterräume von $\mathbb{E}^n$, und sei P ein Punkt mit $P \notin U$. Seien $A \in U$ und $B \in T$ zwei Punkte.*

(i) Wenn $PA \perp U$ und $AB \perp T$, dann $PB \perp T$.

(ii) Wenn $PA \perp U$ und $PB \perp T$, dann $AB \perp T$.

Beweis-Skizze: Da $T \subset U$, haben wir auch $\mathcal{D}(T) \subseteq \mathcal{D}(U)$ und

$$\mathcal{D}(T)^\perp = \mathcal{D}(T)^\perp_{\mathcal{D}(U)} + \mathcal{D}(U)^\perp$$

(i) folgt aus $\overrightarrow{PB} = \overrightarrow{PA} + \overrightarrow{AB}$, Bemerkung 14.76, und die Gleichung hier oben.
(ii) aus $T \subset U \Rightarrow \mathcal{D}(U)^\perp \subset \mathcal{D}(T)^\perp$ + ähnlich wie (i). Q.E.D.

14.3.5 Senkrechte Unterräume

Wir haben bis jetzt nur über Geraden die Senkrecht auf Hyperebenen sind. Wir können das wie folgt verallgemeinern:

Definition 14.78. Seien $T, U \subseteq \mathbb{E}^n$ affine UR. Wir sagen, dass T senkrecht auf U ist, und schreiben $T \perp U$, wenn

(i) $\mathcal{D}(T) \subseteq \mathcal{D}(U)^\perp$, falls $\dim T + \dim U \leq n$

(ii) $\mathcal{D}(U)^\perp \subseteq \mathcal{D}(T)$, falls $\dim T + \dim U \geq n$.

Wenn $\dim T + \dim U = n$, dann sind die Bedingungen gleich.

Satz 14.79. *In der obigen Notation haben wir:*

(i) Wenn $T \perp U$, dann auch $U \perp T$.

(ii) Wenn $\dim T + \dim U \leq n$ und $T \perp U$ und $T' \subset T$ ist ein UR von T, dann gilt auch $T' \perp U$.

(iii) Wenn $\dim T + \dim U \geq n$ und $T \perp U$ und $T \subset T'$ ist ein UR von $\mathbb{E}^n$ der T enthält, dann gilt auch $T' \perp U$.

(iv) Wenn $P \in \mathbb{E}^n$ ein Punkt ist und $U \subseteq \mathbb{E}^n$ ein UR, dann gibt es ein einziger UR $T \subseteq \mathbb{E}^n$, mit $\dim T = n - \dim U$, $P \in T$ und $T \perp U$.

(v) Wenn $T, T', U \subseteq \mathbb{E}^n$ affine UR sind, mit $\dim T = \dim T' = n - \dim U$, dann, aus $T \perp U$ und $T' \perp U$ folgt $T \parallel T'$.

(vi) Wenn $U, T \subseteq \mathbb{E}^n$ affine UR sind, mit $\dim U + \dim T = n$ und $T \perp U$, dann ist $T \cap U$ ein Punkt.

Beweis-Skizze:

(i) Wenn $W_1 \subseteq W_2 \subseteq V$, dann $W_1^\perp \supseteq W_2^\perp$ per Definition.

(ii) Wenn $T' \subseteq T$, dann folgt auch $\mathcal{D}(T') \subseteq \mathcal{D}(T)$.

(iii) Wenn $T' \subseteq T$, dann folgt auch $\mathcal{D}(T') \subseteq \mathcal{D}(T)$.

(iv) $T = P + \mathcal{D}(U)^\perp$.

(v) $\mathcal{D}(T) = \mathcal{D}(U)^\perp = \mathcal{D}(T')$.

(vi) Wir haben, dass $\mathbb{R}^n = \mathcal{D}(T) \oplus \mathcal{D}(U)$. Seien $A \in T$ und $B \in U$, dann ist $\overrightarrow{AB} = v + u$ eindeutig, mit $v \in \mathcal{D}(T)$ und $u \in \mathcal{D}(U)$. Sei $C = A + v \in T$, d.h. $\overrightarrow{AC} = v$. Dann ist $\overrightarrow{BC} = \overrightarrow{AC} - \overrightarrow{AB} = -v \in \mathcal{D}(U)$. Also, $C \in U \cap T$. Also $\mathcal{D}(U \cap T) = \mathcal{D}(U) \cap \mathcal{D}(T)$. Aber in diesem Fall, ist $\mathcal{D}(U) \cap \mathcal{D}(T) = O_{\mathbb{R}^n}$, also ist der Schnitt $U \cap T$ null-dimensional und gleich mit dem Punkt C.

Q.E.D.

14.4 Dimensionsatz für affine Räume

Seien $U_1 = P_1 + W_1$ und $U_2 = P_2 + W_2$ zwei affine UR von $\mathbb{A}^n(\mathbb{K})$, wobei P_1, P_2 Punkte, und W_1, W_2 zwei $\mathbb{K}$-VR sind. Per Definition $\mathrm{Aff}(U_1 \cup U_2)$ ist der kleinste UR von $\mathbb{A}^n(\mathbb{K})$ der sowohl U_1 als auch U_2 enthält. Wir hatten mit $\mathcal{D}(U)$ den an U zugeordneten UVR bezeichnet, und wir haben im allgemein, dass

$$P, Q \in U \Rightarrow \overrightarrow{PQ} \in \mathcal{D}(U).$$

Lemma 14.80. *In der obigen Notation haben wir*

$$\mathrm{Aff}(U_1 \cup U_2) = P_1 + \left(\mathrm{Span}\,\overrightarrow{P_1P_2} + W_1 + W_2\right).$$

Wobei $\mathrm{Span}\, v$ *den* $\mathbb{K}$*-linearen UVR von* v *erzeugt bezeichnet.*

Beweis-Skizze: $\boxed{\subseteq}$ weil U_1 und U_2 im rechten Teil enthalten sind.
$\boxed{\supseteq}$ Da $P_1, P_2 \in U_1 \cup U_2$, haben wir auch $\overrightarrow{P_1P_2} \in \mathcal{D}(\mathrm{Aff}(U_1 \cup U_2))$. Da auch $W_i \subseteq \mathcal{D}(\mathrm{Aff}(U_1 \cup U_2))$ für $i = 1, 2$, folgt also

$$\mathrm{Span}\,\overrightarrow{P_1P_2} + W_1 + W_2 \subseteq \mathcal{D}(\mathrm{Aff}(U_1 \cup U_2)),$$

und das reicht. Q.E.D.

Lemma 14.81.

$$U_1 \cap U_2 \neq \emptyset \iff \mathrm{Span}\,\overrightarrow{P_1P_2} \subseteq W_1 + W_2$$

Beweis-Skizze:

$$\begin{aligned} \mathrm{Span}\,\overrightarrow{P_1P_2} \subseteq W_1 + W_2 &\iff \exists w_i \in W_i \text{ mit } \overrightarrow{P_1P_2} = w_1 + w_2 \\ &\iff P_2 + (-w_2) = P_1 + w_1 \in U_1 \cap U_2. \end{aligned}$$

Q.E.D.

Korollar 14.82. Wenn $P \in U_1 \cap U_2$, dann

$$\begin{aligned} U_1 \cap U_2 &= P + (W_1 \cap W_2) \\ \mathrm{Aff}(U_1 \cup U_2) &= P + (W_1 + W_2) \end{aligned}$$

Satz 14.83. *Wenn* $n \geq 2$ *und* $U_1, U_2 \subseteq \mathbb{A}^n(\mathbb{K})$ *affine Unterräume über* $\mathbb{K}$ *sind, dann,*

wenn $U_1 \cap U_2 \neq \emptyset$*, gilt*

$$\dim \mathrm{Aff}(U_1 \cup U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

wenn $U_1 \cap U_2 = \emptyset$*, gilt*

$$\dim \mathrm{Aff}(U_1 \cup U_2) = \dim U_1 + \dim U_2 - \dim_{\mathbb{K}}(W_1 \cap W_2) + 1.$$

Beweis-Skizze: 1. Folgt aus Korollar 14.82 + die Grassmann Formel für UVR.
2. Aus Lemma 14.80 und 14.81 folgt $\mathcal{D}(\mathrm{Aff}(U_1 \cup U_2)) = \mathrm{Span}\,\overrightarrow{P_1P_2} + W_1 + W_2$ und $\mathrm{Span}\,\overrightarrow{P_1P_2} \not\subseteq W_1 + W_2$, also wieder aus Grassmann's Formel folgt alles. Q.E.D.

14.5 Die Gleichungen von Unterräume

Ein Unterraum von $\mathbb{A}^n(\mathbb{K})$ ist durch $U = P + W$ gegeben, wobei $W \subseteq \mathbb{K}^n$ ist ein $\mathbb{K}$-UVR. Also, wenn $P = (p_1, \dots, p_n)$ und wenn wir eine Basis $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ von W wählen, mit $\mathbf{w}_i = (w_{i1}, \dots, w_{in})$, dann können wir U *parametrisch* beschreiben (oder durch *implizite Gleichungen* geben):

$$U = \{(p_1 + \sum_{j=1}^{m} w_{j1}t_j, \dots, p_n + \sum_{j=1}^{m} w_{jn}t_j) \in \mathbb{K}^n \mid t_j \in \mathbb{K}\ \forall\ j = 1, \dots, m\}.$$

Die Matrix (w_{ij}) hat maximalen Rang (weil Basis), und die $t_1, \dots, t_m$ sind die freie Parameter.

Beispiel 14.84. 1. In $\mathbb{R}^2$: x-Axis; Parallele zur x-Axis durch $(0, 1)$; die Gerade durch $(1, 0), (0, 1)$. Die Beschreibung ändert sich wenn wir verschiedene UVR Basen wählen.

2. Ebene in $\mathbb{K}^3$ (2 Parameter): Die x-Axis; die Ebene durch $(0, 0, 1), (1, 0, 1), (1, 1, 1)$.

Wir haben in Beispiel 14.6.1 gesehen, dass die Lösungsmenge einer nicht homogener linearer Gleichung ein affiner (Unter)raum ist. Das gilt ganz allgemein, nämlich:
Übung: Seien $n > m > 0$ ganze Zahlen, und $A = (a_{ij}) \in \mathrm{Mat}_{(n-m)\times n}(\mathbb{K})$ eine Matrix mit maximaler Rang, und sei $\mathbf{c} = (c_1, \dots, c_{n-m}) \in \mathbb{K}^{n-m}$. Mit der Bezeichnung $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{K}^n$, haben wir, dass die Lösungsmenge

$$U = \{\mathbf{x} \in \mathbb{K}^n \mid A \cdot \mathbf{x} = \mathbf{c}\}$$

ein affiner Unterraum ist.
(Hint: Satz 14.16 Teil (iii))
Es gilt auch die andere Richtung: jeder Unterraum kann man durch *implizite Gleichungen* beschreiben (d.h. als Lösungsmenge eines unhomogenes Gleichungssystems):

Wenn $U = P + W$, wählen wir eine Basis von W: $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ und ergänzen diese zu einer Basis von $\mathbb{K}^n$: $\{\mathbf{w}_1, \dots, \mathbf{w}_m, \mathbf{w}_{m+1}, \dots, \mathbf{w}_n\}$. Wenn $\mathbf{w}_i = (w_{1i}, \dots, w_{ni}) \in \mathbb{K}^n$, dann ist die Matrix $B = (w_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ invertierbar, und wir bezeichnen mit $A := (a_{ij}) := B^{-1}$. Für jeder $\mathbf{x} \in \mathbb{K}^n$, schreiben wir $\mathbf{x} = B \cdot \mathbf{y}$, wobei $\mathbf{y}$ die Koordinaten von $\mathbf{x}$ bezüglich der Basis B sind. Also auch $\mathbf{y} = A \cdot x$, und

$$\mathbf{x} \in W \iff y_{m+1} = \cdots = y_n = 0.$$

Das heißt, wenn $P = (p_1, \dots, p_n)$, dann ist

$$U = \{P + \mathbf{x} \mid \sum_{j=1}^{n} a_{ij}x_j = 0 \quad \forall\ i = m+1, \dots, n\}.$$

also, da die Matrix $(a_{ij}) \in M_{(n-m)\times n}$ maximalen Rang hat,

$$U = \{\mathbf{x} \mid \sum_{j=1}^{n} a_{ij}(x_i - p_i) = 0 \quad \forall\ i = m+1, \dots, n\}.$$

Beispiel 14.85. Alle Hyperebenen werden durch eine lineare Gleichung $a_1x_1 + \cdots + a_nx_n = c$, mit $\mathbf{a} \in \mathbb{K}^n \setminus \{0\}$ und $c \in \mathbb{K}$.

Satz 14.86. *Seien $H, H' \subset \mathbb{A}^n(\mathbb{K})$ zwei Hyperebenen die durch je eine Gleichung gegeben sind:*

$$\begin{aligned} H &= \{\mathbf{x} \in \mathbb{K}^n \mid a_1x_1 + \cdots + a_nx_n = c\} \\ H' &= \{\mathbf{x} \in \mathbb{K}^n \mid a'_1x_1 + \cdots + a'_nx_n = c'\}. \end{aligned}$$

Wenn $H = H'$, dann $\exists d \in \mathbb{K} \setminus \{0\}$ mit $a'_i = da_i$ und $c' = dc$.

Beweis-Skizze: Erstens $c = 0 \iff c' = 0$, (weil dann $(0, \ldots, 0) \in H = H' \Rightarrow c = 0$)
Fall 1: $\boxed{c \neq 0}$. Also auch $c' \neq 0$. Für alle i mit $a_i \neq 0$ ist $(0, \ldots, 0, ca_i^{-1}, 0, \ldots, 0) \in H = H'$, also

$$a'_i c a_i^{-1} = c'$$

Insbesondere ist auch $a'_i \neq 0$ und wir haben

$$d = \frac{c'}{c} = \frac{a'_i}{a_i} \quad \forall\ i \text{ mit } a_i \neq 0.$$

Fall 2: $\boxed{c = 0}$. Also auch $c' = 0$. Sei $a_i \neq 0$ und sei $j \neq i$. Dann ist $(0, \ldots, -a_i, \ldots, a_j, \ldots, 0)$ (wo $-a_i$ an der j-ten Stelle steht, und a_j an der i-ten). Da $H = H'$ folgt auch $a'_ia_j - a'_ja_i = 0$. Q.E.D.

Lemma 14.87. *Seien $P_1, \ldots, P_n \in \mathbb{A}^n(\mathbb{K})$ mit $P_i = (p_{i1}, \ldots, p_{in})$. Dann sind $P_1, \ldots, P_n$ in allgemeiner Lage wenn und nur wenn die $n \times (n+1)$ Matrix*

$$M(P_1, \ldots, P_n) := \begin{pmatrix} p_{11} & \ldots & p_{1n} & 1 \\ p_{21} & \ldots & p_{2n} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & \ldots & p_{nn} & 1 \end{pmatrix}$$

eine $n \times n$ Untermatrix enthält, die auch die letzte Spalte entält, und dessen Determinante nicht null ist.

Beweis-Skizze: Die Eigneschaft für $M(P_1, \ldots, P_n)$ ist äquivalent zur sleben Eigenschaft für

$$M'(P_1, \ldots, P_n) := \begin{pmatrix} p_{11} & \ldots & p_{1n} & 1 \\ p_{21} - p_{11} & \ldots & p_{2n} - p_{1n} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} - p_{11} & \ldots & p_{nn} - p_{1n} & 0 \end{pmatrix}$$

also zur maximalen Rang der Matrix $M'(P_1, \ldots, P_n)$ ohne die erste Reihe und die letzte Spalte, also zur l.u. der $\overrightarrow{P_1P_2}, \ldots, \overrightarrow{P_1P_n}$. Q.E.D.

Satz 14.88. *Seien $P_1, \ldots, P_n \in \mathbb{A}^n(\mathbb{K})$ Punkte in allgemeiner Lage.*
Die Hyperebene $\mathrm{Aff}(P_1, \ldots, P_n) \subset \mathbb{A}^n(\mathbb{K})$ *ist von der Gleichung* $\det M(\mathbf{x}, P_1, \ldots, P_n) = 0$ *gegeben, wo* $\mathbf{x} = (x_1, \ldots, x_n)$ *und*

$$M(\mathbf{x}, P_1, \ldots, P_n) := \begin{pmatrix} x_1 & \ldots & x_n & 1 \\ p_{11} & \ldots & p_{1n} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & \ldots & p_{nn} & 1 \end{pmatrix}$$

Beweis-Skizze: Wir wollen also

$$\mathrm{Aff}(P_1, \ldots, P_n) = \{\mathbf{x} \in \mathbb{K} \mid \det(M(\mathbf{x}, P_1, \ldots, P_n)) = 0.$$

Beide sind Hyperebenen, also es reicht eine Inklusion zu zeigen (Bemerkung 14.19). Aber P_i gehört zur rechten Seite für alle i, also $\boxed{\subseteq}$ gilt. Q.E.D.

Beispiel 14.89. $A, B \in \mathbb{A}^2(\mathbb{K})$ Punkte mit $A = (a_1, a_2)$ und $B = (b_1, b_2)$, dann ist

$$AB = \{(x, y) \in \mathbb{K}^2 \mid \det \begin{pmatrix} x & y & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{pmatrix} = 0\}$$

Axiomatic

David Hilbert hat 1899 Axiome von denen man die Euklidische Geometrie mit moderner Genauigkeit aufbauen kann vorgeschlagen. Diese kommen in 3 Gruppen: Inzidenz-Axiome, Anordnugs-Axiome, Kongruenz-Axiome, + das Parallelen-Axiom und das Kreis-Axiom. Wir werden noch das Archimedische Axiom, das Axiom von Dedekind, und die Axiomatic der projektiven Geometrie betrachten.

Leider kann man nicht immer zeigen, dass Axiom-Systeme (formale Systeme) wiederspruchfrei sind, d.h. dass man nicht sowohl ein Satz als auch seine Negation beweisen kann. Kurt Gödels Erste Unvollständigkeitssatz besagt, dass es in hinreichend starken widerspruchsfreien Systemen immer unbeweisbare Aussagen gibt. Der Zweite Unvollständigkeitssatz besagt, dass hinreichend starke widerspruchsfreie Systeme ihre eigene Widerspruchsfreiheit nicht beweisen können.
Relative Wiederspruchsfreiheit eines Axiom-Systems S gilt wenn es eine mathematische Theorie T gibt, in der man ein Model für S bauen kann. Dann, wenn T wiederspruchfrei ist, ist auch S wiederspruchfrei. D.h., wenn wir an die reelle Zahlen "glauben", dann müssen wir auch an Hilberts Axiome "glauben".

Punkte und Geraden sind nicht definiert, sondern "gegeben". Das heisst, es wird eine Menge $\mathcal{X}$ gegeben, dessen Elemente wir *Punkte* nennen, und eine Untermenge der Menge aller Teilmengen von $\mathcal{X}$: $\mathcal{G} \subseteq 2^{\mathcal{X}}$, dessen Elemente *Geraden* genannt werden. Wir sagen also nicht *was* diese sind, nur *dass* es diese gibt. Wir nennen so ein Paar $(\mathcal{X}, \mathcal{G})$ *Inzidenzstruktur*. Wenn ein Punkt P einer Geraden g gehört (also $P \in g$) sagen wir auch, dass P auf g *liegt*, oder, dass g durch P *läuft*.

Beispiel 14.90.
1. Punkte und Geraden im affinen Raum $\mathbb{A}^n(\mathbb{K})$;
2. Geraden und Ebenen im 3-dim VR;
3. Punkte auf der Kugel + große Kreise (Schnitt von Ebenen durch den Mittelpunkt mit der Kugel)
4. Fano Ebene (7 Punkte, 7 Geraden)

Kapitel 15

Kegelschnitte und Klassifikation der affinen Quadriken

Kegelschnitte sind ein klassischer Teil der Mathematik. Diese wurden seit der Antike studiert, insbesondere von Apollonios von Perge, Omar Khayyám, Johannes Kepler, Girard Desargues, Blaise Pascal, René Descartes, Pierre Fermat. Wir werden hier nur eine ganz kurze und oberflächliche Skizze zeigen. Eine gute Darstellung finden Sie in [Fis09, Teil 5.1] und [Fis01, Teil 1.4].

15.1 Motivation

Eine Gerade in der reellen Ebene ist durch eine Gleichung der Form $ax + by + c = 0$ gegeben. Zum Beispiel, eine Gerade ist die Teilmenge von $\mathbb{R}^2$ definiert als

$$\mathcal{G} = \{(x, y) \in \mathbb{R}^2 \ : \ 2x - y + 4 = 0\}.$$

Wenn wir das zeichnen wollen, ist es nicht schwer: wir berechnen zwei verschiedene Punkte die der Gerade gehören und zeichnen die "gerade Linie" dadurch. Konkret: Wenn $y = 0$, dann muss $2x+4 = 0$, also $x = -2$, somit ist ein Punkt $A = (-2, 0)$. Wenn $x = 0$, dann muss $-y + 4 = 0$, also ein zweiter Punkt ist $B = (0, 4)$. Das heißt $\mathcal{G}$ ist die einzige Gerade die durch A und B geht.

Was passiert aber wenn wir die Nullstellen einer komplizierter Gleichung betrachten wollen? Zum Beispiel

$$\{(x, y) \in \mathbb{R}^2 \ : \ 4\,x^2 + 4\,x\,y + y^2 - 7\,x - 4\,y - 1 = 0\}.$$

Man könnte in diesem Fall auch Punkte berechnen, aber wie viele brauchen wir um sicher zu sein, dass wir einen eindeutige "Linie" bestimmt haben? Und was für eine Figur muss man durch diese Zeichnen? Die Idee ist algebraisch die Gleichung zu bearbeiten um die Geometrie der Figur zu verstehen. In diesem Fall wenn man $x \mapsto 2a - b + 4$ und $y \mapsto -3a + 2b + 2$ in der Gleichung einsetzt bekommt man

$$\begin{aligned} &4(2a - b + 4)^2 + 4(2a - b + 4)(-3a + 2b + 2)+ \\ &+(-3a + 2b + 2)^2 - 7(2a - b + 4) - 4(-3a + 2b + 2) - 1 \quad = \quad \cdots = a^2 - b. \end{aligned}$$

Es handelt sich also um einer Parabel. Wir werden jetzt sehen, wie man genau auf diesem Koordinatenwechsel kommt. Die wichtigste Idee wird aber sein, dass wir Kurven in der Ebene nicht mehr als

"Mengen von Punkten" betrachten werden, und auch nicht als "Spuren der Bewegung eines Punktes"[1]. Kurven in der Ebene werden Äquivalenzklassen von Polynomen sein.

Mit der Lerserfreundlichkeit als Ziel, wiederholen wir zu erst Abschnitt 13.1 aus Kapitel 13.

15.2 Polynome in mehrere Variablen

Wir nehmen an, wir haben den Polynomring in einer Variable mit Koeffizienten in einem Körper $\mathbb{K}$definiert:

$$\mathbb{K}[x] = \left\{ \sum_{i=0}^{d} c_i x^i \quad : \quad d \in \mathbb{N} \text{ und } c_i \in \mathbb{K} \right\}$$

mit den offensichtlichen Addition, Multiplikation und Skalarmultiplikation. Analog wird der Polynomring in einer Variable mit Koeffizienten aus einem kommutativen Ring R definiert: $R[x]$. Den Polynomring in n Variablen $x_1, \ldots, x_n$ ($n \in \mathbb{N}_{>0}$) und Koeffizienten aus $\mathbb{K}$ definieren wir induktiv als

$$\mathbb{K}[x_1, \ldots, x_n] := \big(\mathbb{K}[x_1, \ldots, x_{n-1}]\big)[x_n].$$

Zum Beispiel, für $n = 2$ und $d = 2$ wählen wir Koeffizienten aus $\mathbb{K}[x_1]$:

$$c_0 = x_1^2 + 1, \quad c_1 = 5x_1 - 1, \quad c_2 = x_1^3 + 2x_1^2 + 3x_1 + 4$$

und definieren damit ein Polynom das wie folgt umgeschrieben werden kann:

$$\begin{aligned} f &= c_0 + c_1 x_2 + c_2 x_2^2 \\ &= (x_1^2 + 1) + (5x_1 - 1)x_2 + (x^3 + 2x^2 + 3x + 4)x_2^2 \\ &= 1 - x_2 + x_1^2 + 5x_1 x_2 + 4x_2^2 + 3x_1 x_2^2 + 2x_1^2 x_2^2 + x_1^3 x_2^2. \end{aligned}$$

Wir haben die Klammern aufgemacht, und die Summanden nach der Summe der Exponenten im Produkt geordnet. Allgemein können wir jedes Polynom in mehrere Variablen als eine endliche $\mathbb{K}$-lineare Kombination unterschiedlicher Produkte von Variablen schreiben.

Ein **Monom** in den Variablen $x_1, \ldots, x_n$ ist ein formales Produkt $x_1^{i_1} \cdots x_n^{i_n}$. Der Grad des Monoms ist die Summe der Exponenten:

$$\deg x_1^{i_1} \cdots x_n^{i_n} = i_1 + \cdots + i_n.$$

Für ein n-Tupel $\mathbf{i} = (i_1, \ldots, i_n) \in \mathbb{N}$ schreiben wir $\mathbf{x}^{\mathbf{i}} := x_1^{i_1} \ldots x_n^{i_n}$. Ein Polynom p in n Variablen $x_1, \ldots, x_n$ über den Körper $\mathbb{K}$ ist eine endliche lineare Kombination von Monome, also ein formaler Ausdruck der Form

$$p(x_1, \ldots, x_n) = \sum_{\mathbf{i}=(i_1,\ldots,i_n)\in\mathbb{N}^n} c_{\mathbf{i}}\, \mathbf{x}^{\mathbf{i}}$$

wobei nur endlich viele $c_{\mathbf{i}} \in \mathbb{K}$ nicht Null sind. Der Grad eines Polynoms ist

$$\deg p := \max\{\deg \mathbf{x}^{\mathbf{i}} \mid c_{\mathbf{i}} \neq 0\}.$$

Ein Polynom ist **homogen von Grad** d wenn wenn es eine lineare Kombination von Monome von Grad d ist, also wenn

$$c_{\mathbf{i}} \neq 0 \Rightarrow \operatorname{Grad} \mathbf{x}^{\mathbf{i}} = d.$$

[1] Zum Beispiel, der Kreis mit Mittelpunkt $(0, 0)$ könnte als $\{(\cos\theta, \sin\theta) \in \mathbb{R}^2 \quad : \quad \theta \in \mathbb{R}\}$ oder $\{(\cos\theta, \sin\theta) \in \mathbb{R}^2 \quad : \quad \theta \in [0, 2\pi)\}$ beschrieben werden.

Per Konvetion ist das Nullpolynom homogen von jedem Grad d; das brauchen wir damit die Menge der homogenen Polynome von einem fixierten Grad ein Untervektorraum ist. Wir bezeichnen die Menge aller homogenen Polynome von Grad d in Variablen $x_1, \ldots, x_n$ mit $\mathbb{K}[x_1, \ldots, x_n]_d$. Polynome kann man wie erwartet mit Skalare multiplizieren, miteinander Addieren und Multiplizieren. Der Polynomring $\mathbb{K}[x_1, \ldots, x_n]$ ist also eine $\mathbb{K}$-Algebra.

Allgemein gilt für $f, g \in \mathbb{K}[x_1, \ldots, x_n]$ und $\lambda \in \mathbb{K}$:

$$\begin{aligned} \deg fg &= \deg f + \deg g, \\ \deg f + g &\leqslant \max\{\deg f, \deg g\}, \\ \deg \lambda f &= \deg f. \end{aligned}$$

Wenn also f und g homogen von Grad d sind, dann ist auch jede $\mathbb{K}$-lineare Kombination der beiden homogen von Grad d. Dank der Konvention ist also die Menge aller *homogenen* Polynome von Grad d ein $\mathbb{K}$-Vektorraum, insbesondere ein $\mathbb{K}$-Untervektorraum von $\mathbb{K}[x_1, \ldots, x_n]$. Weiterhin gilt

$$\mathbb{K}[x_1, \ldots, x_n] = \bigoplus_{d \in \mathbb{N}} \mathbb{K}[x_1, \ldots, x_n]_d. \tag{15.1}$$

Jedes Polynom definiert eine so-genannte *polynomielle* Abbildung $p : \mathbb{K}^n \longrightarrow \mathbb{K}$, indem wir das Polynom p in jedem Tupel $(t_1, \ldots, t_n)$ evaluieren. Für unsere Zwecke hier werden wir Polynome und die entsprechenden polynomielle Abbildungen identifizieren[2].

Bemerkung 15.1. Für ein homogenes Polynom p von Grad d und jedes $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{K}^n$ gilt

$$p(\lambda \cdot \mathbf{t}) = \lambda^d \cdot p(\mathbf{t}).$$

15.3 Affine Hyperflächen

Wir fangen mit affine Hyperflächen im n-dimensionalen affinen Raum über einem (fast) beliebigen Körper $\mathbb{K}$. "Fast" heißt, dass wir $\operatorname{char} \mathbb{K} \neq 2$ annehmen werden. Wenn $n = 2$, dann sind diese genau die algebraischen affinen Kurven die uns interessieren.

Sei $n \in \mathbb{N}$. Wir bezeichnen mit $\mathbb{K}[\mathbf{x}]$ den Polynomring $\mathbb{K}[x_1, \ldots, x_n]$. Für jedes (nicht unbedingt homogenes) Polynom in $F \in \mathbb{K}[\mathbf{x}]$ von Grad $d \in \mathbb{N}$ existieren nach (15.1) für $i = 0, \ldots, d$ eindeutige homogene Polynome $F_i \in \mathbb{K}[\mathbf{x}]_i$ mit $F_d \neq 0$, sodass

$$F = F_d + F_{d-1} + \cdots + F_1 + F_0.$$

Beispiele:

$\boxed{d = 1}$ Ein allgemeines Polynom von Grad 1 sieht so aus:

$$F = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + b, \quad \text{mit } (a_1, \ldots, a_n) \in \mathbb{K}^n \setminus (0, \ldots, 0) \text{ und } b \in \mathbb{K}.$$

[2] Das soll man aber allgemein nicht machen! Ein Polynom ist keine Abbildung! Das dient hier nur der Vereinfachung der Sprache wenn wir über die Korrespondenz zwischen Quadratische Formen und homogene Polynome von Grad 2 sprechen werden.

$\boxed{d=2}$ Ein allgemeines Polynom von Grad 2 kann so dargestellt werden[3]:

$$F = F(x_1, \dots, x_n) = \sum_{i,j=1}^{n} c_{ij} x_i x_j + 2 \sum_{i=1}^{n} c_i x_i + c, \tag{15.2}$$

wobei $C = (c_{ij}) \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K}) \setminus \mathbf{0}$ und $c_i, c \in \mathbb{K}$ für $i = 1, \dots, n$.

Auf $\mathbb{K}[\mathbf{x}]$ definieren wir folgende Äquivalenzrelation

$$F \sim G \in \mathbb{K}[\mathbf{x}] \iff \exists \lambda \in \mathbb{K} \setminus \{0\} \text{ sodass } F = \lambda G.$$

Insbesondere, zwei äquivalente Polynome haben denselben Grad. Der Hintergrund ist nicht schwer zu raten: Für $(a_1, \dots, a_n) \in \mathbb{K}^n$ und $\lambda \in \mathbb{K} \setminus \{0\}$ gilt

$$F(a_1, \dots, a_n) = 0 \iff \lambda F(a_1, \dots, a_n) = 0.$$

Also äquivalente Polynome haben dieselben Nullstellen. Die Umkehrung gilt aber nicht. Auch wenn der Körper algebraisch abgeschlossen ist. Zum Beispiel, für $n = 1$ haben $F = x$ und $G = x^2$ dieselben Nullstellen, aber $F \not\sim G$.

Definition 15.2. Eine **affine Hyperfläche** von Grad d in $\mathbb{A}^n_{\mathbb{K}}$ ist die Äquivalenzklasse[4] $\widehat{F}$ eines Polynoms $F \in \mathbb{K}[x_1, \dots, x_n]$ von Grad d. Das Polynom F (oder die Relation $F = 0$) wird eine *Gleichung der Hyperfläche* genannt.

- Eine affine Hyperfläche in $\mathbb{A}^3_{\mathbb{K}}$ wird **affine algebraische Fläche** genannt.
- Eine affine Hyperfläche in $\mathbb{A}^2_{\mathbb{K}}$ wird **affine algebraische Kurve** genannt.
- Eine Hyperfläche von Grad 2 wird **Hyperquadrik** genannt.
- Eine Fläche von Grad 2 wird einfach **Quadrik** genannt.
- Eine affine Kurve von Grad 2 wird **Kegelschnitt**, oder konische Kurve (oder Konik?) genannt.

Beispiel 15.3. Beispiele von Kegelschnitte in $\mathbb{A}^2_{\mathbb{R}}$ sind

der Kreis	$x^2 + y^2 - 1$,	$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$c_1 = 0$	$c_2 = 0$	$c = -1$
die Hyperbel	$x^2 - y^2 - 1$,	$C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$c_1 = 0$	$c_2 = 0$	$c = -1$
die Parabel	$x^2 - y$,	$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$c_1 = 0$	$c_2 = -\frac{1}{2}$	$c = 0$
zwei sich-schneidende Geraden	$x^2 - y^2$,	$C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$c_1 = 0$	$c_2 = 0$	$c = -1$

[3] Hier ist char $\mathbb{K} \neq 2$ wichtig.
[4] bezüglich der obigen Äquivalenzrelation.

Für eine Hyperfläche $\widehat{F}$ bezeichnen wir die Nullstellenmenge (auch **geometrischer Ort** genannt) durch

$$\mathcal{Z}(\widehat{F}) := \mathcal{Z}(F) := \{(a_1, \ldots, a_n) \in \mathbb{A}^n_{\mathbb{K}} \ : \ F(a_1, \ldots, a_n) = 0\}.$$

Es ist wichtig nochmals zu betonen, dass der geometrischen Ort der Hyperfläche mit der Hyperfläche selbst nicht zu verwechseln sind. Auch wenn der Grad gleich ist, könnten unterschiedliche Hyperflächen denselben geometrischen Ort haben. Zum Beispiel in $\mathbb{A}^2_{\mathbb{R}}$ haben wir

$$\mathcal{Z}(x^2 + y^2 + 1) = \mathcal{Z}(x^2 + y^2 + 2) = \emptyset \quad \text{aber} \quad x^2 \widehat{+ y^2} + 1 \neq x^2 \widehat{+ y^2} + 2.$$

15.4 Die Wirkung von $\mathrm{AGL}_n(\mathbb{K})$ auf $\mathbb{K}[\mathbf{x}]$

Wir erinnern kurz, dass $\mathrm{AGL}_n(\mathbb{K}) = \{\sigma : \mathbb{A}^n_{\mathbb{K}} \longrightarrow \mathbb{A}^n_{\mathbb{K}} \ : \ \sigma = \text{ affiner Automorphismus}\}$. Also

$$\sigma(\mathbf{x}) = A \cdot \mathbf{x} + \mathbf{b}, \quad \text{mit } A \in \mathrm{GL}_n(\mathbb{K}) \ \text{ und } \mathbf{b} \in \mathbb{K}^n.$$

Die Inverse von σ ist gegeben durch

$$\sigma^{-1}(\mathbf{x}) = A^{-1} \cdot \mathbf{x} - A^{-1}\mathbf{b}.$$

Das ist einfach zu überprüfen:

$$(\sigma^{-1} \circ \sigma)(\mathbf{x}) = \sigma^{-1}(A \cdot \mathbf{x} + \mathbf{b}) = A^{-1}(A\mathbf{x} + \mathbf{b}) - A^{-1}\mathbf{b} = \mathbf{x}.$$

Die (rechts-)Wirkung $\star : \mathbb{K}[\mathbf{x}] \times \mathrm{AGL}_n(\mathbb{K}) \longrightarrow \mathbb{K}[\mathbf{x}]$ die wir suchen entspricht (links-)Wirkung von $\mathrm{AGL}_n(\mathbb{K})$ auf $\mathbb{A}^n_{\mathbb{K}}$, gegebnen durch $\mathbf{a} \mapsto \sigma(\mathbf{a})$. Wie müssen wir F ändern damit $\sigma(\mathbf{a})$ eine Nullstelle von $F \star \sigma$ genau dann ist, wenn $\mathbf{a}$ eine Nullstelle von F ist? Das ist erfüllt wenn

$$F(\mathbf{x}) \star \sigma := F(\sigma^{-1}(\mathbf{x})).$$

Es ist einfach zu überprüfen, dass $\star$ eine Gruppenwirkung (cf. Definition 14.7) definiert. Es ist noch einfacher zu sehen, dass wenn $F \sim G$, dann gilt auch $F \star \sigma \sim G \star \sigma$. Das heißt, dass wir eigentlich eine Wirkung von $\mathrm{AGL}_n(\mathbb{K})$ auf der Menge der Hyperflächen definiert haben:

$$(\mathbb{K}[\mathbf{x}]/\sim) \times \mathrm{AGL}_n(\mathbb{K}) \longrightarrow (\mathbb{K}[\mathbf{x}]/\sim).$$

Definition 15.4. Zwei Hyperflächen $\widehat{F}$ und $\widehat{G}$ sind **affin äquivalent** wenn es $\sigma \in \mathrm{AGL}_n(\mathbb{K})$ existiert, sodass

$$F \star \sigma = G.$$

Da $\deg F \star \sigma = F$, können wir die Wirkung auf jedem Grad einschränken:

$$\{\text{Hyperflächen von Grad } d\} \star \mathrm{AGL}_n(\mathbb{K}) \longrightarrow \{\text{Hyperflächen von Grad } d\}.$$

Das ist tatsächlich eine Äquivalenzrelation. Wir können also mathematisch genau formulieren, was es heißt affine Hyperflächen zu klassifizieren: Es heißt die Äquivalenzklassen der affinen Äquivalenz zu beschreiben. Insbesondere, kanonische Repräsentanten finden.

15.5 Affine Quadriken

Eine Quadrik gegeben durch eine Gleichung geschrieben in der Form (15.2) kann als

$$F(\mathbf{x}) = \mathbf{x}^\top \cdot C \cdot \mathbf{x} + 2\mathbf{c} \cdot \mathbf{x} + c, \qquad \text{wobei } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

zusammengefasst werden, wobei $C \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K})$ und $\mathbf{c} = \begin{pmatrix} c_1 & \dots & c_n \end{pmatrix} \in \mathrm{Mat}_{1,n}(\mathbb{K})$ und $c \in \mathbb{K}$. Man kann dass noch kompakter ausdrücken als

$$F(\mathbf{x}) = \widetilde{\mathbf{x}}^\top \cdot \widetilde{C} \cdot \widetilde{\mathbf{x}},$$

wobei

$$\widetilde{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} \qquad \widetilde{C} = \begin{pmatrix} c_{11} & \dots & c_{1n} & c_1 \\ \vdots & \ddots & \vdots & \vdots \\ c_{1n} & \dots & c_{nn} & c_n \\ c_1 & \dots & c_n & c \end{pmatrix}$$

Die Matrix $\widetilde{C}$ ist also auch symmetrisch. Wenn $n = 2$ dann haben wir also

$$\begin{aligned} F(x,y) &= \begin{pmatrix} x & y & 1 \end{pmatrix} \cdot \begin{pmatrix} c_{11} & c_{12} & c_1 \\ c_{12} & c_{22} & c_2 \\ c_1 & c_2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} x & y & 1 \end{pmatrix} \cdot \begin{pmatrix} c_{11}x + c_{12}y + c_1 \\ c_{12}x + c_{22}y + c_2 \\ c_1 x + c_2 y + c \end{pmatrix} \\ &= c_{11}x^2 + c_{12}xy + c_1 x + c_{12}xy + c_{22}y^2 + c_2 y + c_1 x + c_2 y + c \\ &= c_{11}x^2 + 2c_{12}xy + c_{22}y^2 + 2c_1 x + 2c_2 y + c. \end{aligned}$$

Wir haben also für eine Quadrik $\widehat{F}$ mit zugeordneten Matrizen $C \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{K})$ und $\widetilde{C} \in \mathrm{Mat}_{n+1}^{\mathrm{sym}}(\mathbb{K})$, und für $\sigma \in \mathrm{AGL}_2(\mathbb{K})$ mit $\sigma(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, dass

$$\begin{aligned} F \star \sigma &= (A^{-1}\mathbf{x} - A^{-1}\mathbf{b})^\top \cdot C \cdot (A^{-1}\mathbf{x} - A^{-1}\mathbf{b}) + 2\mathbf{c} \cdot (A^{-1}\mathbf{x} - A^{-1}\mathbf{b}) + c \\ &= \mathbf{x}^\top \cdot (A^{-1})^\top \cdot C \cdot A^{-1} \cdot \mathbf{x} + \dots \end{aligned}$$

Also die symmetrische Matrix C' von $F' := \sigma \star F$ ist $(A^{-1})^\top \cdot C \cdot A^{-1}$. Insbesondere gilt unter der Transformation, dass

$$\det C' = \frac{\det C}{(\det A)^2}.$$

Für den Fall $\mathbb{K} = \mathbb{R}$ heißt es, dass das Vorzeichen dieser Determinante sich nicht ändert.

Wenn wir

$$\widetilde{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & b_n \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

setzen, dann haben wir

$$\widetilde{A} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = \begin{pmatrix} A\mathbf{x} + \mathbf{b} \\ \\ 1 \end{pmatrix}.$$

In der kompakten Schreibweise von F haben wir dann, dass die $\widetilde{C'}$ Matrix für F' gleich mit

$$\widetilde{C'} = \left(\left(\widetilde{A}\right)^{-1}\right)^{\top} \cdot \widetilde{C} \cdot \left(\widetilde{A}\right)^{-1}.$$

Also wenn $\mathbb{K} = \mathbb{R}$ dann ist das Vorzeichen von $\widetilde{C}$ auch invariant unter affiner Äquivalenz.

Eine Quadrik heißt **nicht entartet** wenn $\det \widetilde{C} \neq 0$ und **entartet** wenn $\det \widetilde{C} = 0$.

15.6 Affine Klassifizierung Hyperquadriken in der komplexen

Sei F eine affine Hyperquadrik in $\mathbb{A}^n_{\mathbb{C}}$ mit zugeordneten symmetrischen Matrizen $C \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{C})$ und $\widetilde{C} \in \mathrm{Mat}_{n+1}^{\mathrm{sym}}(\mathbb{C})$. Wir können aber nur mit $\mathrm{AGL}_n(\mathbb{C})$ operieren, deswegen bearbeiten wir erstmals C. Aus Korolloar 11.37 ist jede symmetrische komplexe Matrix kongruent zu einer Matrix der Form:

$$D = \begin{pmatrix} \boxed{I_r} & \\ & \mathrm{O} \end{pmatrix}.$$

Das heißt, es gibt eine invertierbare Matrix $A \in \mathrm{GL}_n(\mathbb{C})$ mit $A^{\top} C A = D$. Wir wenden das auf der Matrix C der Hyperquardik an. Das heißt wir wirken auf F mit der en affinen Transformation

$$\sigma(\mathbf{x}) = A^{-1}\mathbf{x} + \mathbf{0}.$$

Wir können also die Matrix der Quadrik in der folgenden Form bringen

$$\widetilde{C} = \begin{pmatrix} 1 & & & & & & c_1 \\ & \ddots & & & & & \vdots \\ & & 1 & & & & c_r \\ & & & 0 & & & c_{r+1} \\ & & & & \ddots & & \vdots \\ & & & & & 0 & c_{r+1} \\ c_1 & \dots & c_r & c_{r+1} & \dots & c_n & c \end{pmatrix}.$$

Die Gleichung ist also jetzt:

$$F(\mathbf{x}) = x_1^2 + \cdots + x_r^2 + 2c_1x_1 + \cdots + 2c_rx_r + 2c_{r_1}x_{r+1} + \cdots + 2c_nx_n + c.$$

Wir können jetzt den $c_1, \ldots, c_r$ durch Quadratische Ergänzung los werden. Zum Beispiel $x_1^2 + 2c_1x1 + c_1^2 - c_1^2 = (x_1 + c_1)^2 - c_1^2$. Wir wollen also $x_1 + c_1$ als neue Variable "$\sigma(\mathbf{x})$" bezeichnen; usw. Die Affinität enspricht $\sigma(\mathbf{x}) = \mathbf{x} + (c_1, \ldots, c_r, 0, \ldots, 0)^\top$. Wir wirken also wie folgt:

$$\begin{aligned} F(\mathbf{x}) \mapsto F(\mathbf{x}) \star \sigma &= F(\sigma^{-1}(\mathbf{x})) \\ &= F(x_1 - c_1, \ldots, x_r - c_r, x_{r+1}, \ldots, x_n) \\ &= x_1^2 + \cdots + x_r^2 + 2c_{r+1}x_{r+1} + \cdots + 2c_nx_n + (c - c_1^2 + \cdots - c_r^2). \end{aligned}$$

Wir dürfen annehmen[5], dass für ein $s \in \{r, \ldots, n\}$ gilt

$$c_{r+1}, \ldots, c_s \neq 0 \text{ und } c_{s+1} = \cdots = c_n = 0.$$

Als nächstes *normalisieren*[6] wir die Koeffizienten des *linearen*[7] Teils. Zu diesem Zweck ersetzten wir, für $i = r+1, \ldots, s$, x_i mit $\frac{x_i}{2c_i}$. Das heißt, wir wirken mit $\sigma \in \mathrm{AGL}_n(\mathbb{K})$ gegeben durch der Matrix $A = \mathrm{diag}(1, \ldots, 1, 2c_{r+1}, \ldots, 2c_s, 1, \ldots, 1)$ und dem Vektor $\mathbf{b} = \mathbf{0}$. Wir bekommen somit

$$F(\sigma^{-1}(\mathbf{x})) = x_1^2 + \cdots + x_r^2 + x_{r+1} + \cdots + x_s + c.$$

Fall 1: *Wenn der Grad 1 Teil nicht Null ist.*
Wenn also $r < n$ und $r < s$, dann können wir noch x_s mit $x_s - c$ ersetzen, und somit dem konstanten Teil los werden.

Fall 2: *Wenn der Grad 1 Teil Null ist.*
Fall 2.1: *wenn* $c \neq 0$. Dann können wir noch $x_1, \ldots, x_r$ durch $\sqrt{c}x_1, \ldots, \sqrt{c}x_r$ ersetzen (das geht für alle $c \in \mathbb{C}$). Dann, weil $c \neq 0$, haben wir äquivalente Polynome[8]

$$cx_1^2 + \cdots + cx_r^2 + c \sim x_1^2 + \cdots + x_r^2 + 1.$$

Fall 2.2: *wenn* $c = 0$. Dann ist die Gleichung nicht mehr zu vereinfachern:

$$x_1^2 + \cdots + x_r^2.$$

Wir haben also folgenden Satz (fast) bewiesen:

Satz 15.5. *Jede Hyperquadrik in* $\mathbb{A}^n_{\mathbb{C}}$ *ist affin äquivalent zu* genau *einer der folgenden Hyperquadriken affin äquivalent:*

$$F = \begin{cases} x_1^2 + \cdots + x_r^2 + x_{r+1} + \cdots + x_s = 0 & \textit{mit } 1 \leqslant r < n,\ r+1 \leqslant s \leqslant n \text{ ODER} \\ x_1^2 + \cdots + x_r^2 = 0 & \textit{mit } 1 \leqslant r \leqslant n \text{ ODER} \\ x_1^2 + \cdots + x_r^2 + 1 = 0 & \textit{mit } 1 \leqslant r \leqslant n. \end{cases}$$

[5] Eventuell müssen wir die Indizes permutieren. Das enspricht der Wirkung einer Affinität gegeben durch einer Permutationsmatrix.
[6] Wir bringen sie also auf 1 durch die Wirkung einer Affinintät.
[7] Der Teil von Grand 1.
[8] Das heißt per Definition 15.2 dieselbe Hyperfläche.

Beweis-Skizze: Dass jede Hyperquadrik auf einer dieser Formen gebracht werden kann, haben wir oben bewiesen. Was fehlt noch, ist das keine zwei solche Hyperquadriken affin äquivalent sind.

Wir wissen, dass $\operatorname{Rang} C$ und $\operatorname{Rang} \widetilde{C}$ unter der Wirkung von $\mathrm{AGL}_n(\mathbb{K})$ konstant bleibt. Wir haben in allen drei Fällen

$$\operatorname{Rang} C = r.$$

Für $\widetilde{C}$ haben wir verschiedene Antworten in den drei Fällen:

$$\operatorname{Rang} \widetilde{C} = \begin{cases} s & \text{mit } 1 \leqslant r < n,\ r < s \leqslant n \text{ ODER} \\ r & \text{mit } 1 \leqslant r \leqslant n \text{ ODER} \\ r+1 & \text{mit } 1 \leqslant r \leqslant n. \end{cases}$$

Das zeigt, dass keine zwei der obigen Hyperflächen affin äquivalent sind. Q.E.D.

Korollar 15.6. Jeder Kegelschnitt in $\mathbb{A}^2_{\mathbb{C}}$ ist zu genau einem der folgenden äuqivalent:

- $x^2 + y^2 + 1 = 0$ mit Matrix $\widetilde{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,
- $x^2 + y^2 = 0$ mit Matrix $\widetilde{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,
- $x^2 + y = 0$ mit Matrix $\widetilde{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$,
- $x^2 + 1 = 0$ mit Matrix $\widetilde{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,
- $x^2 = 0$ mit Matrix $\widetilde{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Davon sind nur die erste und die dritte nicht ausgeartet.

15.7 Affine Klassifizierung reellen Kegelschnitte

Mit einem sehr ähnlichen Beweis, in dem wir nur die Vorzeichen berücksichtigen müssen, bekommen wir eine Klassifizierung in dem reellen Fall. Wir schreiben das aber hier nur für Kurven auf.

Satz 15.7. *Jeder Kegelschnitt in $\mathbb{A}^2_{\mathbb{R}}$ ist affin äquivalent zu genau einer der folgenden Kurven:*

(i) $x^2 + y^2 - 1 = 0$ *(**Ellipse**)*

(ii) $x^2 + y^2 + 1 = 0$ *(Ellipse ohne reellen Punkten)*

(iii) $x^2 + y^2 = 0$ *(Entartete Ellipse)*

(iv) $x^2 - y^2 - 1 = 0$ ***(Hyperbel)***

(v) $x^2 - y^2 = 0$ *(Entartete Hyperbel)*

(vi) $x^2 - y = 0$ ***(Parabel)***

(vii) $x^2 - 1 = 0$ *(entartete Parabel (parallele Geraden))*

(viii) $x^2 + 1 = 0$ *(entartete Parabel ohne reellen Punkten)*

(ix) $x^2 = 0$ *(doppelt entarteter Kegelschnitt (doppelte gerade))*

Beweis-Skizze: Sei F der Kegelschnitt mit zugeordneten symmetrischen Matrizen $C \in \mathrm{Mat}_2(\mathbb{R})$, beziehungsweise $\widetilde{C} \in \mathrm{Mat}_3(\mathbb{R})$. Aus dem Ersten Trägheitssatz von Sylvester (Korollar 11.38) existiert eine Matrix $A \in \mathrm{GL}_2(\mathbb{R})$, sodass

$$A^\top C A = \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{oder} \quad \pm \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Die Nullmaitrx kommt nicht vor, weil wir eine Gleichung des zweites Grades haben. Weil wir ohne die Kurve zu ändern Gleichungen mit -1 multiplizieren dürfen, können wir das "$\pm$" weglassen. Wir haben also drei Fälle. In den ersten beiden, können wir durch quadratische Ergänzung (also durch die Wirkung von $\sigma(x, y) = (x + c_1, y + c_2)$) dem linearen Teil los werden. In dem dritten Fall können wir nur dem linearen Teil in x los werden. Wir haben also die Fälle:

1. $x^2 + y^2 + c$

2. $x^2 - y^2 + c$

3. $x^2 + 2c_2 y + c$

Fall 1 & 2. $\boxed{c \neq 0}$

In dem komplexen Fall haben wir die Varieblen mit $\sqrt{c}$ skaliert um c auf 1 zu bringen. In $\mathbb{R}$ sind wir gezwungen mit $\sqrt{|c|}$ zu skalieren, das heißt $\sigma^{-1}(x, y) = (\sqrt{|c|}x, \sqrt{|c|}y)$. Das bringt uns auf

$$|c|x^2 \pm |c|y^2 + c \sim x^2 \pm y^2 \pm 1.$$

Fall 1 & 2. $\boxed{c = 0}$

In diesem Fall haben wir schon eine der vorgegebenen Formen (**1.3** oder **2.3** hier unten).

Wir haben also:

1.1. $x^2 + y^2 + 1$

1.2. $x^2 + y^2 - 1$

1.3. $x^2 + y^2$

2.1. $x^2 - y^2 + 1$

2.2. $x^2 - y^2 - 1$

2.3. $x^2 - y^2$

Fall **2.1.** und Fall **2.2.** sind aber affin äquivalent wenn wir x mit y tauschen und mit -1 multiplizieren.

Fall 3.1. $\boxed{c_2 \neq 0}$

Wir können mit $\sigma(x, y) = (x, -2c_2 - y + c)$ wirken, und bekommen

$$x^2 - y.$$

Fall 3.2.1. $\boxed{c_2 = 0, c \neq 0}$

Dann wirken wir wieder mit $\sigma^{-1}(x, y) = (\sqrt{|c|}x, y)$ und bekommen

3.2.1a. $x^2 + 1$

3.2.1b. $x^2 - 1$

Fall 3.2.2. $\boxed{c_2 = 0, c = 0}$

Dann ist die Gleichung schon in der affin kanonischer Form: x^2. Also alle 9 kanonischen Formen sind erreichbar.

Wir brauchen noch, dass diese paarweise nicht affin äquivalent sind. Dafür brauchen wir, neben dem Rang von C und von $\widetilde{C}$, noch die Signatur der Matrizen C und $\widetilde{C}$ ins Spiel zu bringen. Wir haben

Fall	Rang C	sign C	Rang $\widetilde{C}$	sign $\widetilde{C}$
(i)	2	2	3	1
(ii)	2	2	3	3
(iii)	2	2	3	2
(iv)	2	0	3	-1
(v)	2	0	2	0
(vi)	1	1	3	?
(vii)	1	1	2	0
(viii)	1	1	2	2
(ix)	1	1	1	1

Weil rang und signatur invariant unter kongruenz sind, sind alle diese neun Fälle paarweise nicht affin äquivalent. Q.E.D.

15.8 Metrische Klassifizierung der affinen Quadriken

DIESER TEIL IST UNVOLLSTÄNDIG

Aus Satz 12.45 ist jede symmetrische reelle Matrix durch eine Orthogonalmatrix diagonalisierbar. Sei

$A \in O(n)$ eine solche Matrix, und $\sigma \in \mathrm{AGL}_n(\mathbb{R})$ die entsprechende affine Transformation (mit $\mathbf{b} = \mathbf{0}$). Dann gilt

Satz 15.8. *Jede reelle Quadrik in dem euklidischen affinen Raum $\mathbb{A}^n_{\mathbb{R}}$ ist isometrisch äquivalent zu einer Quadrik mit Gleichung*

$$F = d_1 x_1^2 + \cdots + d_n x_n^2 + 2c_1 x_1 + \cdots + 2c_n x_n + c,$$

wobei $d_1, \ldots, d_n$ die Eigenwerte der Matrix C von F sind.

Weiterhin, wir können annehmen, dass es $1 \leqslant r \leqslant n$ gibt mit $d_i \neq 0$ wenn $i \leqslant r$ und $d_i = 0$ wenn $i > r$. Das heißt wir haben die Form

$$F = d_1 x_1^2 + \cdots + d_r x_r^2 + 2c_1 x_1 + \cdots + 2c_n x_n + c.$$

Durch quadratische Ergänzung, das heißt, durch die Affinität[9]

$$x_i \mapsto \begin{cases} x_i + c_i/d_i & \text{wenn } \ i \leq r \\ x_i & \text{wenn } \ i > r \end{cases}$$

$$F = d_1 x_1^2 + \cdots + d_r x_r^2 + 2c_{r+1} x_{r+1} + \cdots + 2c_n x_n + c,$$

mit $d_i \neq 0$.

[9] d.h. affiner Automorphismus

Index