Lineare Algebra I

Bernd Ammann, WS 2007/08

Inhaltsverzeichnis

Literaturverzeichnis

[1] S. Bosch, Lineare Algebra, Springer
[2] H.D. Ebbinghaus, Einführung in die Mengenlehre, BI Wissenschaftsverlag
[3] Gerd Fischer, Lineare Algebra, Vieweg
[4] Paul R. Halmos, Finite-dimensional vector spaces
[5] Paul .R. Halmos, Naive Mengenlehre, Vandenhoeck und Ruprecht
[6] K. Jänich, Lineare Algebra, Springer
[7] M. Koecher, Lineare Algebra und Analytische Geometrie
[8] Falko Lorenz, Lineare Algebra I und II
[9] Serge Lang, Linear Algebra, Second Edition, Addison-Wesley
[10] Ulf Friedrichsdorf, Alexander Prestel, Mengenlehre für den Mathematiker, Vieweg Studium

Warnungen

Dies ist das Skript der Vorlesung Lineare Algebra 1, Regensburg, Wintersemester 2007/08. Es fehlen noch Diagramme und Zeichnungen, die teilweise zum Verständnis sehr nützlich wären. Es ist unwahrscheinlich, dass es ein Skript zur Linearen Algebra 2 gibt. Ich werde im Sommersemester parallel zur Linearen Algebra 2 eine Vorlesung für mittlere Semester lesen, über ein Gebiet, in dem nicht so viel Literatur existiert. Wenn ich also im Sommersemester ein Skript erstelle, dann zur anderen Vorlesung.

Und schließlich eine wichtige Bitte: Legen Sie das Skript nicht in eine Ecke mit dem ruhigen Gewissen, es ja später lesen und durcharbeiten zu können. Beginnen Sie sobald wie möglich, die Lücken zu schließen. Schwierige Beweise durchschauen Sie am besten, wenn Sie sich überlegen, was der Beweis in konkreten Beispielen macht.

KAPITEL 1

Zahlen und Kongruenzen

1. Vorbemerkungen zur Axiomatik

Die natürlichen Zahlen werden seit Jahrtausenden intuitiv benutzt und untersucht. Sie sind uns vertraut, ohne dass wir aber wirklich wissen, was sie charakterisiert. So ähnlich war es mit vielen mathematischen Konzepten, zum Beispiel dem mathematischen Konzept der unendlichen Summe. Man nutzte viele Konzepte lange in einer vagen Bedeutung, ohne sich zu überlegen, wie man sie definiert. Leider führte dies zu wachsenden Problemen. Es gab unter anderem viele Diskussionen, was denn der Wert der unendlichen Summe

$$1 + (-1) + 1 + (-1) + \ldots$$

sei, manche Mathematiker vertraten die Ansicht es sei 0: mit der Begündung

$$(1 + (-1)) + (1 + (-1)) + (1 + (-1)) + \ldots = 0 + 0 + 0 + \ldots = 0.$$

Dies erscheint überzeugend. Mit derselben Logik kann man aber auch begründen, dass man den Wert 1 erhält:

$$1 + ((-1) + 1) + ((-1) + 1) + \ldots = 1 + 0 + 0 + \ldots = 1.$$

Derartige Probleme motivierten die Mathematiker, die Mathematik auf solide Grundlagen zu stellen. Diese Bewegungen, die man Axiomatik nennen kann, begann im Bereich der Geometrie bereits mit Euklid von Alexandria (ca. 300 v. Chr.) und Aristoteles (384–322 v. Chr.). Wichtige Fortschritte in der Axiomatik der Geometrie und insgesamt der Axiomatik wurden im 19. Jahrhundert vollbracht.

Das Ziel der Axiomatik ist es, die gesamte Mathematik aus wenigen Grundaussagen, sogenannte Axiomen, herzuleiten. Im Prinzip soll alles auf den Axiomen der Mengenlehre aufbauen, oft versieht man Teilgebiete mit eigenen Axiomen, so zum Beispiel die natürlichen Zahlen, wie wir unten sehen werden. Die Axiome sind nicht mehr weiter beweisbar, sie werden einfach als gegeben hingenommen, entweder weil man sie als evident, also offensichtlich ansieht, oder weil man sie als Kennzeichen der Theorie ansieht. Aus diesen Axiomen werden dann Schlussfolgerungen gezogen. Durch erlaubte Kombination von bereits bekannten wahren Aussagen erhält man neue wahre Aussagen. Eine Sammlung solcher Aussagen nennt man Beweis. Falls diese Aussagen interessant erscheinen, nennt man solche hergeleiteten Aussagen dann Theorem, Lemma, Korollar, Proposition, Satz, Hilfssatz, Folgerung oder ähnlich. Hierbei ist im allgemeinen ein Satz oder ein Theorem eine wichtige Aussage, ein Lemma oder ein Hilfssatz eine Aussage, die nur als Zwischenschritt dient, und eine Proposition hat eine Mittelstellung. Erhält man eine Aussage nahezu unmittelbar aus einem Theorem oder Satz, so nennt man dies eine Folgerung oder ein Korollar.

Damit die Aussagen nicht immer länger und länger werden, macht man Definitionen. Hierbei gibt man mathematischen Objekten oder mathematischen Sachverhalten einen Namen. Die Mathematiker sind im Prinzip recht frei in der Wahl ihrer Definitionen. So könnte man die folgende Definition machen: Ein *Auto* ist eine Menge, in der die Elemente 1, 2 und 3 enthalten sind. Ein *Hund* ist eine Menge, in der die Elemente 1 und 2 enthalten sind. Man schließt daraus, dass jedes Auto einen Hund enthält. Diese Definitionen sind natürlich sehr irreführend, aber prinzipiell erlaubt. Wir bemühen uns die Dinge so zu benennen, dass sie etwas mit der „wirklichen Welt" zu tun haben. Um Koordinaten auf einer Kugel anzugeben, definiert man den Begriff einer „Karte", und ein „Atlas" ist dann definiert als Menge von Karten, so dass alles überdeckt wird. Diese Begriffe sind dann zwar nicht genau das, was man damit umgangssprachlich meint, aber auch nicht völlig ohne Zusammenhang. Die mathematischen Begriffe „Halm" oder „Garbe" der Mathematik haben aber keinerlei Anwendungen in der Landwirtschaft, die „Knoten" der Mathematik sind aber nahe an dem, was man alltagssprachlich als Knoten bezeichnet.

Viele Definitionen werden von allen Mathematikern gleich gemacht, es herrscht Konsens. Man ist sich aber nicht einig, ob die Definition der natürlichen Zahlen die Null einschließen soll oder nicht. Für unsere Vorlesung gilt: die natürlichen Zahlen sind

$$\mathbb{N} = \{1, 2, \ldots\}.$$

Die Menge $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ bezeichnen wir als natürliche Zahlen mit Null. Dies ist eine der üblichen Definitionen. Viele Mathematiker definieren hingegen die Menge der natürlichen Zahlen als $\{0, 1, 2 \ldots\}$. Das ist nicht weiter schlimm. Ob Null eine natürliche Zahl ist oder nicht, ist Definitionssache. Man schaut sich die Definition des Autors an und weiß, was er meint.

In zentralistischen Ländern wie Frankreich ist klar geregelt: Null ist eine natürliche Zahl. In Deutschland besagt DIN 5473 ebenfalls, dass Null eine natürliche Zahl ist. Lehrer in der Schule sollten sich an diese DIN-Norm halten. An den deutschen Universitäten definiert man aber zumeist die natürlichen Zahlen ohne Null, und daran halten wir uns hier auch.

Anders ist es bei der Zahl π. Dass der Wert dieser Zahl zwischen 3,1415 und 3,1416 liegt ist eine Aussage und keine Definition. Deswegen ist es lächerlich, dass der US-Bundesstaat Indiana 1897 den Wert von π auf $3,2$ gesetzlich festlegen wollte, um Berechnungen zu vereinfachen und um es den Schülern einfacher zu machen.

In der Vorlesung können wir aber nicht alles streng axiomatisch einführen, da dies viel zu lange dauern würde. Wir müssen einen Kompromiss zwischen der nötigen Strenge und angemessener Kürze finden. Wir wollen deswegen die natürlichen, ganzen, rationalen und reellen Zahlen durch Eigenschaften beschreiben. Diese Eigenschaften dienen dann als Axiome für die Theorie der jeweiligen Zahlensysteme, aus denen man dann die weiteren Eigenschaften herleiten kann. Man kann auch zeigen, dass man mit Hilfe der Mengenlehre Modelle für diese Zahlen konstruieren und somit die Existenz von diesen Zahlen beweisen. Diesen Teil wollen wir hier überspringen. Ein Modell der natürlichen Zahlen und der reellen Zahlen wird in der Analysis konstruiert werden.

2. Natürliche Zahlen

2.1. Die Peano-Axiome. Wir schreiben also $\mathbb{N} = \{1, 2, 3, \ldots\}$ für die Menge der natürlichen Zahlen und $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ für die natürlichen Zahlen mit Null.

Die natürlichen Zahlen wurden von Dedekind (1888) und Peano (1889) axiomatisiert. Siehe [**10**, Kapitel 3 und 4] oder [**2**, Kapitel V] für mehr Details.

Axiome der natürlichen Zahlen (Peano-Axiome)

(P1) 1 ist eine natürliche Zahl.
(P2) Zu jeder natürlichen Zahl n gibt es genau einen Nachfolger n^+, der auch natürliche Zahl ist.
(P3) Es gibt keine natürliche Zahl deren Nachfolger 1 ist.
(P4) Jede natürliche Zahl ist der Nachfolger von höchstens einer natürlichen Zahl.
(P5) Falls S eine Menge von natürlichen Zahlen ist, so dass
  (a) $1 \in S$,
  (b) Für jedes n in S ist auch n^+ in S enthalten,
  dann ist S gleich der Menge der natürlichen Zahlen.

In unserer Vorlesung verstehen wir unter dem Nachfolger einfach eine Abbildung $\mathbb{N} \to \mathbb{N}$ die die obigen Eigenschaften erfüllen soll. Dies weicht etwas von der Analysis-Vorlesung ab!

Das letzte der Axiome ist nötig, damit die natürlichen Zahlen so klein wie möglich sind. Man kann eine Menge M mit Nachfolger-Abbildung $M \to M$, $m \mapsto m^+$ konstruieren, die (P1)–(P4) erfüllt, aber nicht (P5).

Proposition 2.1. *Jede natürliche Zahl ungleich* 1 *ist der Nachfolger genau einer natürlichen Zahl.*

Beweis (Direkter Beweis). Wir wissen bereits, dass jede natürliche Zahl der Nachfolger höchstens einer natürlichen Zahl ist. Zu zeigen bleibt also, dass jede natürlichen Zahl n ungleich 1 Nachfolger einer natürlichen Zahl ist. Wir definieren:

$$T := \{n \in \mathbb{N} \,|\, n \text{ ist Nachfolger einer natürlichen Zahl}\}$$

und dann

$$S := T \cup \{1\}.$$

Die Menge S erfüllt die Eigenschaften im Peano-Axiom (P5): 1 ist in S, und wenn n in S enthalten ist, dann ist n eine natürliche Zahl, und somit ist n^+ eine natürliche Zahl, die Nachfolger einer natürlichen Zahl ist. Somit ist n^+ in S. Axiom 5 besagt also, dass S gleich $\mathbb{N}$ ist. Daraus folgt $T = \mathbb{N}$ oder $T = \mathbb{N} \setminus \{1\}$. Also ist jede Zahl ungleich 1 Nachfolger einer natürlichen Zahl. □

Man kann den Beweis auch anders führen, als sogenannten *Widerspruchsbeweis.*

Beweis (Widerspruchsbeweis). **Wir nehmen an, die Aussage des Lemmas ist falsch, und wollen daraus einen Widerspruch herleiten. Wenn die Aussage des Lemmas falsch ist, dann können wir annehmen:** *Es gebe* eine natürliche Zahl n ungleich 1, die nicht Nachfolger einer natürlichen Zahl ist. Wir definieren $S := \mathbb{N} \setminus \{n\}$. Die Menge S enthält 1, da $n \neq 1$. Ist $s \in S$, so ist s auch eine natürliche Zahl und somit ist auch der Nachfolger s^+ eine natürliche Zahl. Da n kein Nachfolger einer natürlichen Zahl ist, folgt $s^+ \neq n$ und somit $s^+ \in S$. Die Menge S erfüllt also die Eigenschaften in Peano-Axiom (P5) und somit gilt $S = \mathbb{N}$. Zusammen mit $n \in \mathbb{N}$ und $n \notin S$ ergibt sich ein Widerspruch. **Da die Existenz einer Zahl n mit den obigen Eigenschaften zu einem Widerpruch führen würde, wissen wir, dass es ein solches Gegenbeispiel nicht gibt, und die Aussage des Lemmas ist somit bewiesen.** □

In der bisherigen Art und Weise kommen wir aber leider nur langsam vorwärts. Man lässt deswegen Sätze, die sich aus dem Kontext ergeben, wie zum Beispiel die **fett gedruckten** Sätze weg. Wir benutzen oben den Konjunktiv Präsens „Es gebe" an Stelle vom Indikativ „Es gibt", um damit eine Annahme auszudrücken. Man benutzt den Konjunktiv auch oft für Definitionen.

Wir vereinigen nun $\mathbb{N}$ mit einem weiteren Element, das wir 0 oder „Null" nennen. Die Menge $\mathbb{N}_0$ *sei* die Menge $\mathbb{N} \cup \{0\}$, die Menge der natürlichen Zahlen mit Null.

2.2. Rekursive Definitionen. Es erscheint uns intuitiv klar, wie man natürliche Zahlen addiert. Wenn man die Addition aber sauber definieren möchte, muss man mehr Aufwand betreiben, als man zunächst denkt. Eine Möglichkeit ist, für gegebenes $n \in \mathbb{N}$ eine Abbildung $f_n : \mathbb{N} \to \mathbb{N}, m \mapsto f_n(m)$ rekursiv zu definieren.

Definitionsanfang

$$f_n(1) := n^+$$

Definitionsschritt Wir nehmen an, dass $f_n(m)$ definiert ist, dann setzen wir

$$f_n(m^+) := (f_n(m))^+.$$

Man muss mehrere Dinge überprüfen, um sicher zu sein, dass nun dies tatsächlich eine wohldefinierte Abbildung $f_n : \mathbb{N} \to \mathbb{N}, m \mapsto f_n(m)$ ergibt, was wir hier nicht in allen Einzelheiten machen wollen. Es ist aber eine gute Übung, sich zu überlegen, wie man dies im Detail macht. Man benötigt Peano-Axiome (P1) und (P2), damit die obige Definition überhaupt Sinn ergibt. Man braucht Peano-Axiom (P3) um sicherzustellen, damit dem Ausdruck $f_n(1)$, der bereits im Definitionsanfang definiert wurde, nicht nochmal im Definitionsschritt etwas anderes zugeordnet wird. Und schließlich benötigt man Peano-Axiom (P4), um zu zeigen, dass dem Ausdruck $f_n(m^+)$ nicht mehrere verschiedene Werte zugewiesen werden, was einen Widerspruch darstellen würde. Übrig bleibt die Frage:

Ist $f_n(m)$ nun bereits für alle $m \in \mathbb{N}$ definiert?

Wir definieren hierzu S_n als die Menge aller natürlicher Zahlen m, für die $f_n(m)$ durch die obige Definition definiert ist. Es gilt $1 \in S_n$ (Definitionsanfang), und falls m in S_n ist, so ist auch $m^+ \in S_n$ (Definitionsschritt). Peano-Axiom (P5) besagt also $S_n = \mathbb{N}$. Wir sehen also unter Benutzung aller Peano-Axiome, dass die Abbildung $f_n : \mathbb{N} \to \mathbb{N}$, $m \mapsto f_n(m)$ auf $\mathbb{N}$ wohldefiniert ist.

Man schreibt dann letztendlich

$$2 := 1^+ \qquad 3 := 2^+ \qquad \dots \qquad n + m := f_n(m).$$

Wir definieren außerdem $n + 0 := n$ und $0 + n := n$.

Mit genau derselben Argumentation sehen wir:

Satz 2.2 (Prinzip der rekursiven Definition)**.** *Sei $F : M \to M$ eine Funktion. Wähle $x \in M$. Dann wird die Funktion $f : \mathbb{N} \to M$ auf eindeutige Art und Weise durch die beiden folgenden Zuordnungen festgelegt.*

- Definitionsanfang
$$f(1) := x$$
- Definitionsschritt *Wir nehmen an, dass $f(m)$ definiert ist, dann setzen wir*
$$f(m^+) := F(f(m)).$$

Wenn wir $f_n(1) := n$ und $f_n(m + 1) := n + f_n(m)$ setzen, so erhalten wir mit

$$n \cdot m := f_n(m)$$

eine rekursive Definition der Multiplikation auf $\mathbb{N}$. Außerdem setzen wir $0 \cdot n := 0$ und $n \cdot 0 := 0$. Analog mit $f_n(1) := n$ und $f_n(m + 1) := n \cdot f_n(m)$, $n^m := f_n(m)$ erhalten wir das Exponieren.

Satz 2.3. $(\mathbb{N}_0, +, \cdot)$ *erfüllt die folgenden Eigenschaften:*

(Aa) ***Addition ist assoziativ***
Für alle $x, y, z \in \mathbb{N}_0$ gilt
$$(x + y) + z = x + (y + z).$$

(Ak) ***Addition ist kommutativ***
Für alle $x, y \in \mathbb{N}_0$ gilt
$$x + y = y + x.$$

(Ma) ***Multiplikation ist assoziativ***
Für alle $x, y, z \in \mathbb{N}_0$ gilt
$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(Mk) ***Multiplikation ist kommutativ***
Für alle $x, y \in \mathbb{N}_0$ gilt
$$x \cdot y = y \cdot x.$$

(AMd) ***Addition und Multiplikation erfüllen das Distributivgesetz***
Für alle $x, y, z \in \mathbb{N}_0$ gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z$$

$$(y + z) \cdot x = y \cdot x + z \cdot x$$

Den Beweis kann man mit vollständiger Induktion durchführen, die wir im nächsten Abschnitt kennenlernen werden. Es ist eine gute Übung, einmal die Kommutativität der Addition durch vollständige Induktion oder direkt aus den Peano-Axiomen herzuleiten. Man sieht schnell, dass dies etwas trickreicher ist, als man zunächst gedacht hat. Ein Beweis der Kommutativität geht einfacher und intuitiver mit etwas Mengenlehre. Deswegen wollen wir diesen Beweis hier überspringen.

Es gelten auch noch die folgenden Eigenschaften, die trivial erscheinen. Es wird später klar werden, wieso wir sie als Eigenschaft formulieren.

(An) **Addition hat neutrales Element**
Es gibt ein Element $0 \in \mathbb{N}_0$, so dass für alle $x \in \mathbb{N}_0$ gilt

$$x + 0 = 0 + x = x.$$

Man nennt 0 das *neutrale Element der Addition.*

(Mn) **Multiplikation hat neutrales Element**
Es gibt ein Element $1 \in \mathbb{N}_0$, so dass für alle $x \in \mathbb{N}_0$ gilt

$$x \cdot 1 = 1 \cdot x = x.$$

Man nennt 1 das *neutrale Element der Multiplikation.*

Summen- und Produktzeichen
Oft ist es sinnvoll über Ausdrücke der Form

$$a_1 + \ldots + a_n$$

zu reden. Um dies exakt zu machen, führen wir ein Symbol ein. Man definiert rekursiv

$$\sum_{j=1}^{1} a_j := a_1$$

und

$$\sum_{j=1}^{n+1} a_j := \left(\sum_{j=1}^{n} a_j \right) + a_{n+1}.$$

Analog hierzu

$$\prod_{j=1}^{1} a_j := a_1$$

und

$$\prod_{j=1}^{n+1} a_j := \left(\prod_{j=1}^{n} a_j\right) \cdot a_{n+1}.$$

Fakultät
Wir definieren

$$0! := 1$$
$$n! := (n-1)! \cdot n$$

Man nennt dies "n Fakultät".

2.3. Vollständige Induktion. Nehmen wir an, dass P_n eine Aussage ist, die von einer natürlichen Zahl n abhängt. Ein Beispiel ist:

$$P_n : \quad \Leftrightarrow \quad \sum_{j=1}^{n} j = n(n+1)/2$$

Solche Aussagen zeigt man am besten mit einem Beweisprinzip, das sich vollständige Induktion (oder manchmal auch rekursiver Beweis) nennt.

SATZ 2.4 (Vollständige Induktion). `Sei` P_n `eine Aussage, die von einem Parameter` $n \in \mathbb{N}$ `abhängt. Wir nehmen an, dass Induktionssanfang und Induktionsschritt erfüllt sind:`
Induktionsanfang: `Die Aussage` P_1 `ist wahr.`
Induktionsschritt: `Für alle` $n \in \mathbb{N}$ `gilt: Falls die Aussage` P_n `wahr ist, so ist auch` P_{n+1} `wahr.`
Dann ist die Aussage P_n für alle $n \in \mathbb{N}$ wahr.

Dieser Satz ist wie viele Sätze in der Mathematik klar in einen `so gedruckten Teil mit Voraussetzungen` und einen **fett gedruckten Teil mit den Folgerungen** geteilt. Es ist wichtig, in Aussagen immer klar zu machen, was zu den Voraussetzungen und was zu den Folgerungen gehört.

Der Beweis ergibt sich direkt aus dem Peano-Axiom (P5):

Beweis. Sei S die Menge aller natürlichen Zahlen n, für die P_n wahr ist. Auf Grund des Induktionsanfangs ist 1 in S. Der Induktionsschritt besagt: wenn $n \in S$, dann ist auch $n+1$ in S. Die Menge S erfüllt also die Eigenschaften in Peano-Axiom (P5) und somit gilt $S = \mathbb{N}$. □

Wenn man etwas für alle $n \in \mathbb{N}_0$ zeigen will, kann man dies genauso machen, wenn man mit 0 anstelle mit 1 beginnt.

Beispiel. Wir wollen die Induktionsaussage

$$P_n :\Leftrightarrow \quad \sum_{j=1}^{n} j = n(n+1)/2$$

für alle $n \in \mathbb{N}$ zeigen.

Induktionsanfang: Wir zeigen die Aussage für $n = 1$:

$$\sum_{j=1}^{1} j = 1 = 1(1+1)/2$$

Induktionsschritt: Wir nehmen P_n an, d. h.

$$\sum_{j=1}^{n} j = n(n+1)/2.$$

Wir erhalten

$$\left(\sum_{j=1}^{n+1} j\right) = \left(\sum_{j=1}^{n} j\right) + (n+1) = n(n+1)/2 + (n+1) = (n+2)(n+1)/2 = (n+1)((n+1)+1)/2$$

und somit P_{n+1}.

Durch vollständige Induktion gilt also P_n für alle $n \in \mathbb{N}$.

2.4. Ordnung der natürlichen Zahlen.

Eine Funktion

$$R : M \times M \to \{\text{wahr}, \text{falsch}\}, \qquad (n,m) \mapsto R(n,m) = nRm$$

nennt man Relation auf M.

Es gibt auf $\mathbb{N}$ eine Relation $\leq$ mit den Eigenschaften

(1) Reflexivität: Für alle m in $\mathbb{N}$ ist $m \leq m$ wahr.
(2) Antisymmetrie: Für alle n und m in $\mathbb{N}$ gilt

$$n \leq m \wedge m \leq n \Rightarrow n = m.$$

(3) Transitivität: Für alle n, m und k in $\mathbb{N}$ gilt:

$$n \leq m \wedge m \leq k \Rightarrow n \leq k.$$

(4) Totalität: Für alle n und m mit in $\mathbb{N}$ gilt

$$n \leq m \vee m \leq n \vee m = n$$

(5) Für alle $n \in \mathbb{N}$ gilt $n \leq n+1$.
(6) Für alle $n \in \mathbb{N}$ gilt $1 \leq n$.

Funktionen $R : M \times M \to \{\text{wahr}, \text{falsch}\}$, die (1) bis (3) erfüllen, nennt man *Ordnungsrelationen.* Mehr dazu in der Analysis. Jede Ordnungsrelation, die (5) erfüllt, stimmt mit $\leq$ überein, das heißt die Relation $\leq$ ist durch die Eigenschaften (1)–(3) und (5) eindeutig charakterisiert.

Für einen strengen Aufbau der natürlichen Zahlen müsste man all diese Aussagen mit vollständiger Induktion zeigen, was wir hier überspringen wollen.

DEFINITION 2.5. Sei R eine Ordnungsrelation auf M. Ein Minimum (beziehungsweise Maximum) ist ein Element $m \in M$, so dass mRn (bzw. nRm) für alle $n \in M$.

Bemerkung. Wegen der Antisymmetrie gibt es höchstens ein Minimum.

Beispiele: Das offene Intervall $(0, 1)$ in $\mathbb{R}$ hat kein Minimum bezüglich $\leq$.

Die Menge $M := \{\{a\}, \{b\}, \{a, b\}\}$ trägt die Ordnungsrelation $\subseteq$. Es existiert kein Minimum in M.

PROPOSITION 2.6. *Sei A eine nichtleere Teilmenge von $\mathbb{N}$, dann besitzt A ein Minimum.*

Beweis. Wir nehmen an, A besitze kein Minimum. Zu zeigen ist, dass A die leere Menge ist. Wir zeigen induktiv die Aussage

$$P_n :\Leftrightarrow \qquad \{1, 2, \dots, n\} \cap A = \emptyset,$$

woraus die Aussage folgt.
Induktionsanfang: Angenommen 1 wäre in A. Dann ist 1 das Minimum. Da es aber kein Minimum in A gibt, folgt $1 \notin A$, also P_1.
Induktionsschritt: Es gelte P_n. Falls $n + 1 \in A$, so ist $n + 1$ ein Minimum von A. Da es aber kein Minimum gibt, gilt $n + 1 \notin A$, und somit P_{n+1}. □

Mit der Ordnungsrelation kann man eine stärkere Version der vollständigen Induktion zeigen [1]

SATZ 2.7 (Erweiterte vollständige Induktion)**.** *Sei P_n eine Aussage, die von einem Parameter $n \in \mathbb{N}$ abhängt. Wir nehmen an, dass der Induktionssanfang und der schwache Induktionsschritt erfüllt sind:*
Induktionsanfang*: Die Aussage P_1 ist wahr.*
Schwacher Induktionsschritt*: Für alle $n \in \mathbb{N}$ gilt: Falls die Aussage für alle P_m mit $m \in \{1, 2, \dots, n\}$ wahr ist, so ist auch P_{n+1} wahr.*
Dann ist die Aussage P_n für alle $n \in \mathbb{N}$ wahr.

Beweis. Wir definieren die Aussage

$$Q_n :\Leftrightarrow \qquad P_m \text{ ist wahr für alle } m \in \{1, 2, \dots, n\}.$$

Offensichtlich gilt $Q_n \Rightarrow P_n$. Die Aussage Q_1 ist äquivalent zu P_1.

---

[1] Wir brauchen hier die Ordnungsrelation um die Menge

$$\{1, 2, \dots, n\} := \{k \in \mathbb{N} \,|\, k \leq n\}$$

zu definieren.

Wir nehmen an, dass P_n den Induktionsanfang und den schwachen Induktionsschritt erfüllt. Der schwache Induktionsschritt für P_n impliziert offensichtlich den Induktionsschritt für Q_n. Wir sehen also mit vollständiger Induktion, dass Q_n für alle $n \in \mathbb{N}$ wahr ist, somit ist auch P_n für alle $n \in \mathbb{N}$. □

3. Die ganzen Zahlen

Wir schreiben
$$\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\} \supseteq \mathbb{N}_0$$
für die ganzen Zahlen. Die Addition und die Multiplikation setzen sich zu Abbildungen
$$\begin{aligned} + : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad & (a, b) \mapsto a + b \\ \cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad & (a, b) \mapsto a \cdot b \end{aligned}$$
fort, und die Eigenschaften (Aa), (Ak), (An), (Ma), (Mk), (Mn) und (AMd) gelten weiterhin, wenn man $\mathbb{N}_0$ durch $\mathbb{Z}$ ersetzt. Außerdem gilt

(Ai) **Addition hat inverse Elemente**
Zu jedem $x \in \mathbb{Z}$ gibt es ein $y \in \mathbb{Z}$, so dass
$$x + y = y + x = 0.$$
Man nennt y das Inverse von x bezüglich der Addition und schreibt normalerweise $-x$ anstelle von y.

Die ganzen Zahlen sind die kleinste Erweiterung der natürlichen Zahlen, die diese Eigenschaften hat.

Auch die Ordnung setzt sich auf $\mathbb{Z}$ fort und die Ordnungseigenschaften (1)–(5) gelten.

Eine mit Addition und Mulitplikation versehene Abbildung, die die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), und (AMd) hat, nennt man *Ring*. Gilt zusätzlich (Mn), so spricht man von einem *Ring mit Eins*, und wenn zusätzlich (Mk) gilt, so ist es ein *kommutativer Ring*. Diese Eigenschaften bilden die Axiome der Ringtheorie.

Die ganzen Zahlen bilden somit einen kommutativen Ring mit Eins. Ein Ring, der die natürlichen Zahlen enthält und kleinstmöglich ist, stimmt „im wesentlichen“ mit den ganzen Zahlen überein.

4. Die rationalen Zahlen

Die rationalen Zahlen sind
$$\mathbb{Q} := \left\{ \frac{z}{n} \,\middle|\, z \in \mathbb{Z},\ n \in \mathbb{N} \right\}.$$
Hierbei gilt
$$\frac{z}{n} = \frac{y}{m} \Leftrightarrow zm = yn.$$

Die Addition und die Multiplikation setzen sich zu Abbildungen

$$+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \quad (a, b) \mapsto a + b$$
$$\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \quad (a, b) \mapsto a \cdot b$$

fort, und die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), (Mk), (Mn) und (AMd) gelten weiterhin für $(\mathbb{Q}, +, \cdot)$. Außerdem gilt

(Mi) **Multiplikation hat inverse Elemente**
Zu jedem $x \in X \setminus \{0\}$ gibt es ein $y \in X$, so dass

$$x \cdot y = y \cdot x = 1.$$

Man nennt y das Inverse von x bezüglich der Multiplikation und schreibt normalerweise x^{-1} anstelle von y.

Die rationalen Zahlen sind die kleinste Erweiterung von $\mathbb{Z}$, die diese Eigenschaften hat.

Auch die Ordnung setzt sich fort. Die Ordnung auf $\mathbb{Q}$ ist auch total.

Mit Addition und Multiplikation versehene Mengen mit mindestens 2 Elementen nennt man *Körper*, falls die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), (Mk), (Mn), (Mi) und (AMd) erfüllt sind. Diese Eigenschaften bilden die Axiome der Körpertheorie, aus denen alle weiteren Definitionen und Eigenschaften der Körpertheorie hergeleitet werden. Jeder Körper, der die natürlichen Zahlen enthält und „kleinstmöglich" ist, ist „im wesentlichen gleich" den rationalen Zahlen.

5. Die reellen Zahlen

Die reellen Zahlen $\mathbb{R}$ sind eine Erweiterung der rationalen Zahlen $\mathbb{Q}$. Addition, Multiplikation und die Ordnung setzen sich fort. Auch $(\mathbb{R}, +, \cdot)$ ist ein Körper. Um ihn zu charakterisieren, spielt die Ordnung eine zentrale Rolle.

(g) **Geordneter Körper**
Auf $\mathbb{R}$ existiert eine totale Ordnungsrelation $\leq$ für die gilt:
(a) Für alle $x, y, z\mathbb{R}$ gilt: $x \leq y \Rightarrow x + z \leq y + z$.
(b) Für alle $x, y, z\mathbb{R}$ gilt: $x \leq y \wedge 0 \leq z \Rightarrow xz \leq yz$.

Definition 5.1. Eine Teilmenge $A \subset \mathbb{R}$ heißt *nach oben beschränkt*, falls es ein $R \in \mathbb{R}$ gibt, so dass für alle $a \in A$ die Aussage $a \leq R$ gilt. Ein solches R heißt obere Schranke von A.

In den reellen Zahlen gilt das Axiom

(S) **Supremumseigenschaft**
Sei A eine nichtleere, nach oben beschränkte Menge. Dann enthält die Menge

$$\{t \in \mathbb{R} \,|\, t \text{ ist obere Schranke von } A\}$$

ein Minimum.

Dieses Minimum, die kleinste obere Schranke, nennt man das Supremum von A und notiert kurz $\sup A$. Analog erhält man die größte untere Schranke, die man Infimum $\inf A$ nennt. Ist A nicht nach oben beschränkt, so setzen wir $\sup A := \infty$, ist A nicht nach unten beschränkt, dann $\inf A := -\infty$. Wir setzen auch $\sup \emptyset := -\infty$ und $\inf \emptyset := \infty$.

Zum Beispiel ist $\sup(0,1) = 1$, $\sup[0,1] = 1$, $\sup(0,2) \cap \mathbb{Q} = 2$, $\sup\{-\frac{1}{n} | n \in \mathbb{N}\} = 0$, $\sup \mathbb{Q} = \infty$.

Ein Körper mit Eigenschaft (g) heißt geordneter Körper. Jeder geordnete Körper mit der Supremumseigenschaft ist „im wesentlichen gleich“ den reellen Zahlen.

Die rationalen Zahlen erfüllen die Supremumseigenschaft nicht. Die Menge

$$A := \{x \in \mathbb{Q} \,|\, x^2 \leq 2\}$$

ist nach oben und unten beschränkt. Die Menge der rationalen oberen Schranken von A ist $M := \{x \in \mathbb{Q} \,|\, x > 0 \text{ und } x^2 > 2\}$. Die kleinste obere Schranke wäre also eine Wurzel von 2, die es aber in $\mathbb{Q}$ nicht gibt.

Die bisher erwähnten Zahlensysteme haben wir durch ihre grundlegenden Eigenschaften, ihre Axiome charakterisiert. Wichtig wäre auch die Frage, ob es derartige Zahlensysteme überhaupt gibt. Mathematisch präzise sollte man fragen, ob man mit Hilfe der Axiome der Mengenlehre Modelle konstruieren kann, die diese Axiome erfüllen. Diese Frage wollen wir hier aber nicht näher erörtern. Bei den nun folgenden komplexen Zahlen und den Kongruenzen erscheint es uns aber didaktisch besser, aus den reellen Zahlen bzw. ganzen Zahlen heraus ein Modell für die komplexen Zahlen und Kongruenzen zu konstruieren anstelle sie axiomatisch zu beschreiben.

6. Die komplexen Zahlen

Die Zahl -1 besitzt keine Wurzel in $\mathbb{R}$. Wir werden in der linearen Algebra sehen, dass es oft sehr nützlich ist, derartige Wurzeln ziehen zu können. Auch die Physik des 20. Jahrhunderts ist unvorstellbar ohne die Möglichkeit, eine Wurzel aus -1 ziehen zu können.

Wir definieren die Menge der *komplexen Zahlen*

$$\mathbb{C} := \{(x,y) \,|\, x, y \in \mathbb{R}\}.$$

Auf dieser Menge definieren eine Addition und Multiplikation

$$\begin{array}{rcll} + : \mathbb{C} \times \mathbb{C} & \to & \mathbb{C}, & (x_1,y_1) + (x_2,y_2) \mapsto (x_1 + x_2, y_1 + y_2) \\ \cdot : \mathbb{C} \times \mathbb{C} & \to & \mathbb{C}, & (x_1,y_1) \cdot (x_2,y_2) \mapsto (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) \end{array}$$

Beachten Sie: Innerhalb der Klammern stehen die bereits definierte Addition und Multiplikation der reellen Zahlen. Zwischen den Klammern steht die neue Addition und Multiplikation.

Die Abbildung

$$I : \mathbb{R} \to \mathbb{C}, \qquad r \mapsto (r, 0)$$

ist offensichtlich injektiv. Sie bewahrt außerdem Addition und Multiplikation, das heißt

$$I(r+s) = I(r) + I(s) \qquad I(rs) = I(r) \cdot I(s).$$

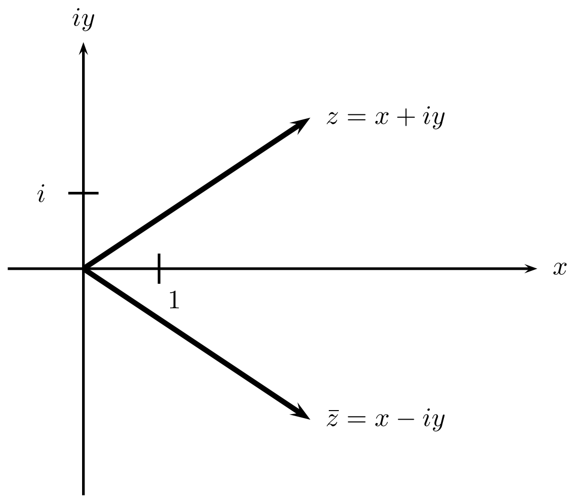


ABBILDUNG 1. Komplexe Konjugation $z \mapsto \bar{z}$

Wir wollen deswegen ab sofort $\mathbb{R}$ vermöge der Abbildung I mit dem Bild von I identifizieren. Mit dieser Identifikation gilt $\mathbb{R} = \{(r, 0) \,|\, r \in \mathbb{R}\}$, $r = (r, 0)$. Dann ist $(\mathbb{C}, +, \cdot)$ eine Erweiterung von $(\mathbb{R}, +, \cdot)$.

Man schreibt außerdem $i := (0, 1)$. In dieser Schreibweise gilt

$$x + iy = (x, y) \qquad i^2 = -1.$$

Man nennt nun x den *Realteil* $\boldsymbol{Re}(x + iy)$ *der komplexen Zahl* $x + iy$ und y den *Imaginärteil* $\mathbf{Im}(x + iy)$. Der Imaginärteil y von $z = x + iy$ verschwindet genau dann, wenn $z \in \mathbb{R} \subseteq \mathbb{C}$. Man nennt solche komplexe Zahlen deswegen *reell*. Komplexe Zahlen, deren Realteil verschwindet, nennt man *rein imaginär*.

Die Abbildung $\mathbb{C} \to \mathbb{C}$, $x + iy \mapsto \overline{x + iy} := x - iy$ heißt *komplexe Konjugation*. Die Größe $|z| := \sqrt{z\bar{z}}$ nennt man den *Betrag* oder *Absolutbetrag* von z. Schreiben wir $z = x + iy$ mit $x, y \in \mathbb{R}$, so gilt $|z| := \sqrt{x^2 + y^2}$.

Die geometrische Interpretation sehen wir in Abbildung 1.

PROPOSITION 6.1. *Die komplexen Zahlen* $\mathbb{C}$ *mit der oben definierten Addition und Multiplikation erfüllen die Eigenschaften (Aa), (Ak), (An), (Ai), (Ma), (Mk), (Mn), (Mi) und (AMd). In anderen Worten,* $(\mathbb{C}, +, \cdot)$ *ist ein Körper.*

Beweisskizze Das neutrale Element der Addition ist $0 = 0 + i \cdot 0 = (0, 0)$, das neutrale Element der Multiplikation ist $1 = (1, 0)$, somit gilt (An) und (Mn). Die Eigenschaften (Aa), (Ak), (Ai),(Mk) und (AMd) sind offensichtlich. Die Eigenschaft (Ma) überprüft man durch nachrechnen. Die Existenz des Inversen bezüglich der Multiplikation (Mi) wollen wir im Detail zeigen. Hier zeigen wir zuerst einen Hilfssatz.

Hilfssatz 6.2. *Für alle $z, w \in \mathbb{C}$ gilt*

$$|zw| = |z|\,|w|.$$

Beweis des Hilfssatzes.

$$|zw| = \sqrt{(zw)\overline{(zw)}} = \sqrt{zw\overline{z}\overline{w}} = \sqrt{(z\overline{z})(w\overline{w})} = |z|\,|w|.$$

□

Wir nehmen nun mal an, es gebe ein Inverses w von z, das heißt $wz = 1$. Multiplikation mit $\overline{z}$ ergibt dann

$$w|z|^2 = wz\overline{z} = 1 \cdot \overline{z} = \overline{z}.$$

Offensichtlich $|z|^2 \neq 0$, falls $z \neq 0$. Multiplikation mit der reellen Zahl $\frac{1}{|z|^2}$ ergibt

$$w = \frac{\overline{z}}{|z|^2}.$$

Wenn es also ein Inverses von z gibt, so ist es $\frac{\overline{z}}{|z|^2}$.

Beweis von (Mi). Sei $z \in \mathbb{C} \setminus \{0\}$. Wir definieren

$$w := \frac{\overline{z}}{|z|^2}.$$

Wir rechnen

$$wz = \frac{\overline{z}z}{|z|^2} = \frac{|z|^2}{|z|^2} = 1.$$

Und somit folgt (Mi). □

Wir definieren

$$z_1 \leq z_2 :\Leftrightarrow z_2 - z_1 \text{ ist eine positive reelle Zahl oder } z_1 = z_2.$$

Dies ist eine Ordnungsrelation, die aber nicht mehr total ist: es gilt weder $0 \leq i$ noch $0 \geq i$ noch $0 = i$. Die Zahl -1 hat nun zwei Quadratwurzeln: i und $-i$. Allerdings ist es unklar, ob man nun $\sqrt{-1} := i$ oder $\sqrt{-1} := -i$ definieren sollte. Die uns wohlvertraute Eigenschaft

$$\sqrt{rs} = \sqrt{r}\sqrt{s} \qquad \forall r, s \geq 0$$

geht bei beiden Definitionen verloren für $r = s = -1$.

Wichtig für die komplexen Zahlen ist:

Satz 6.3 (Fundamentalsatz der Algebra, Gauß 1799)**.** *Sei $P : \mathbb{C} \to \mathbb{C}$ eine Polynom-Funktion, d.h. es gebe $a_0, a_1, \ldots, a_n \in \mathbb{C}$ mit $P(z) = \sum_{j=0}^{n} a_j z^j$. Dann ist P konstant (d.h. $a_2 = \ldots = a_n = 0$) oder es gibt eine Nullstelle $z \in \mathbb{C}$ von P (d.h. ($P(z) = 0$).*

Der Beweis dieses Satzes ist etwas aufwändiger und wird normalerweise in Analysis 3 oder 4 gemacht.

7. Kongruenzen

Sei $n \in \mathbb{N}$ gegeben.

Definition 7.1. Wir sagen $x, y \in \mathbb{Z}$ sind kongruent modulo n, falls n die Differenz $x - y$ teilt. Wir schreiben dann $x \equiv y \mod n$.

Die Relation $\equiv$ erfüllt.

(1) Reflexivität: Für alle $x \in \mathbb{Z}$ gilt $x \equiv x \mod n$.
(2) Symmetrie: Für alle $x, y \in \mathbb{Z}$ gilt
$$(x \equiv y \mod n) \quad \Rightarrow \quad (y \equiv x \mod n).$$
(3) Transitivität: Für alle $x, y, z \in \mathbb{Z}$ gilt:
$$(x \equiv y \mod n) \wedge (y \equiv z \mod n) \quad \Rightarrow \quad (x \equiv z \mod n).$$

Solche Relationen nennt man Äquivalenzrelationen. Die Relation ist auch verträglich mit Addition und Multiplikation im folgenden Sinne.

Lemma 7.2. *Es gelte $x \equiv x' \mod n$ und $y \equiv y' \mod n$. Dann gilt auch $x + y \equiv x' + y' \mod n$ und $xy \equiv x'y' \mod n$.*

Beweis. Es gelte $x \equiv x' \mod n$ und $y \equiv y' \mod n$. Dann $x - x' = kn$ und $y - y' = mn$ für $k, m \in \mathbb{Z}$. Dann $(x + y) - (x - y) = (k + m)n$, also $x + y \equiv x' + y' \mod n$. Ferner gilt
$$xy - x'y' = x(y - y') + (x - x')y' = xmn + kny'.$$
□

Man kann nun mit Zahlen modulo n rechnen, das heißt anschaulich man ignoriert alle Vielfachen von n. Formal bildet man Restklassen
$$x + n\mathbb{Z} := \{y \in \mathbb{Z} \,|\, x \equiv y \mod n\}.$$
Die folgenden Aussagen sind äquivalent

(1) $x \equiv y \mod n$
(2) Es gibt $k \in \mathbb{Z}$ mit $x = kn + y$.
(3) $x \in y + n\mathbb{Z}$

(4) $y \in x + n\mathbb{Z}$
(5) $x + n\mathbb{Z} = y + n\mathbb{Z}$

Man setzt $\mathbb{Z}/n\mathbb{Z} := \{x + n\mathbb{Z} \mid x \in \mathbb{Z}\}$. Dies ist eine Menge mit n Elementen.

Beispiel: $n = 2$

$0 + n\mathbb{Z} = \{\dots, -2, 0, 2, \dots\}$, $1 + n\mathbb{Z} = \{\dots, -1, 1, 3, \dots\}$,

$$\mathbb{Z}/n\mathbb{Z} := \{0 + n\mathbb{Z}, 1 + n\mathbb{Z}\} = \Big\{\{\dots, -2, 0, 2, \dots\}, \{\dots, -1, 1, 3, \dots\}\Big\}$$

Wenn die Zahl n aus dem Kontext klar ist, schreibt man oft auch $\overline{x}$ anstelle von $x + n\mathbb{Z}$.

Für $n = 3$:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

$\bullet$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Für $n = 4$:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

$\bullet$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Man definiert $\overline{x} + \overline{y} := \overline{x + y}$ und $\overline{x} \cdot \overline{y} := \overline{xy}$. Es ist offensichtlich, dass $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ein kommutativer Ring mit 1 ist, d.h. es gelten (Aa), (Ak), (An), (Ai), (Ma), (Mk), (Mn) und (AMd).

Proposition 7.3. *(Mi) gilt genau dann wenn n eine Primzahl ist. In anderen Worten: $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ist ein Körper genau dann, wenn n eine Primzahl ist.*

Beweis kommt später (Kapitel 2 Folgerung 2.7).

KAPITEL 2

Gruppen, Ringe, Körper

1. Gruppen

DEFINITION 1.1. Eine *Menge mit Verknüpfung* ist ein Paar $(X, \circ)$, bestehend aus einer Menge X und einer Abbildung

$$\circ : X \times X \to X, \qquad (x, y \mapsto x \circ y)$$

Beispiele: $(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$, $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{Z}/n\mathbb{Z}, +)$ sind Mengen mit Verknüpfung.

DEFINITION 1.2. Eine Menge mit Verknüpfung $(X, \circ)$ nennt man Gruppe, falls gilt

(Va) **Assoziativität**
Für alle $x, y, z \in X$ gilt

$$(x \circ y) \circ z = x \circ (y \circ z).$$

(Vn) **Verknüpfung hat neutrales Element**
Es gibt ein Element $e \in X$, so dass für alle $x \in X$ gilt

$$x \circ e = e \circ x = x.$$

Man nennt e das *neutrale Element.*

(Vi) **Verknüpfung hat inverse Elemente**
Zu jedem $x \in X$ gibt es ein $y \in X$, so dass

$$x \circ y = y \circ x = e.$$

Man nennt y das Inverse von x normalerweise x^{-1} anstelle von y.

Gilt zusätzlich

(Vk) **Kommutativität**
Für alle $x, y \in X$ gilt

$$x \circ y = y \circ x,$$

so nennt man $(X, \circ)$ eine *kommutative Gruppe* oder eine *abelsche Gruppe.*

BEISPIELE.

(1) $(\mathbb{N}_0, +)$ erfüllt (Aa) und (An), aber nicht (Ai), ist also keine Gruppe.
(2) Wenn wir $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ und $\mathbb{Z}/n\mathbb{Z}$ mit der Addition versehen, dann sind dies abelsche Gruppen.

(3) Wenn wir $\{1, -1\}$, $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$, $\mathbb{C} \setminus \{0\}$ und $\mathbb{Q}^+ := \{q \in \mathbb{Q} \,|\, q > 0\}$ mit der Multiplikation versehen, so sind dies abelsche Gruppen.

Es ist sinnvoll, die Theorie der Gruppen zu studieren, da Gruppen oft in der Mathematik und Physik vorkommen und wir diese Eigenschaften nicht immer wieder, sondern für alle Situationen simultan studieren und beweisen wollen. Historisch wurde der Begriff einer Gruppe wichtig in den Werken von Abel (1802–1829, Norwegen) und Galois (1811–1832, Paris). Diese Mathematiker konnten u.a. zeigen, dass man die Nullstellen von Polynomen von Grad 5 oder größer im Allgemeinen nicht durch die Grundrechenarten (+,−,· und /) und durch Wurzelziehen berechnen kann.

LEMMA 1.3. *Sei $(X, \circ)$ eine Menge mit Verknüpfung, die (Vn) erfüllt. Dann ist das neutrale Element eindeutig.*

Beweis. Seien e_1 und e_2 zwei neutrale Elemente. Es gilt dann

$$e_1 = e_1 \circ e_2 = e_2,$$

und somit ist das neutrale Element eindeutig. □

Dieses Lemma ist streng genommen bereits nötig, damit wir von **dem** neutralen Element e reden dürfen. Wir benötigen das Lemma auch, um der Eigenschaft (Vi) Bedeutung zu geben.

NICHT-KOMMUTATIVES BEISPIEL. Sei M eine Menge und $\mathrm{Bij}(M)$ die Menge aller Bijektionen von M nach M. Als Verknüpfung auf $X = \mathrm{Bij}(M)$ wählen wir die Verkettung von Abbildungen, d.h. für $f, g \in \mathrm{Bij}(M)$ und $m \in M$ gelte

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

Dann ist $(\mathrm{Bij}(M), \circ)$ eine Gruppe, die Permutationsgruppe von M. Das neutrale Element ist die Identität und das Inverse einer Abbildung f ist ihre Umkehrabbildung f^{-1}. Die Menge $\mathcal{S}_n = \mathrm{Bij}(\{1, \ldots, n\})$ nennt auch *die symmetrische Gruppe zum Index n*. Die Gruppe $\mathcal{S}_3$ kann man auch als Symmetrie-Gruppe eines gleichseitigen Dreiecksverstehen.

Spiegelung an der Mittelsenkrechten von (12) und anschließende Spiegelung an der Mittelsenkrechten von (23) ergibt eine Drehung um 120 Grad im Uhrzeigersinn. Vertauscht man die Reihenfolge, erhält man eine Drehung um 120 Grad im Gegen-Uhrzeigersinn, siehe auch Abbildung 1.

LEMMA 1.4 (Kürzungsregeln)**.** *Sei $(G, \circ)$ eine Gruppe und $x, y, z \in G$. Dann folgt aus $x \circ y = x \circ z$ bereits $y = z$ (von links kürzen). Ebenso folgt aus $x \circ z = y \circ z$ bereits $x = y$ (von rechts kürzen).*

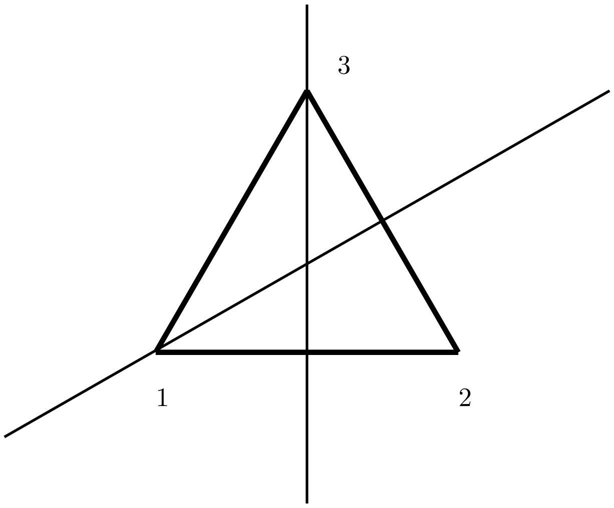


ABBILDUNG 1. Ein gleichseitige Dreieck, mit den Mittelsenkrechten zu (12) und zu (23)

Beweis. Wenn wir die Gleichung $x \circ y = x \circ z$ von links mit x^{-1} multiplizieren, so ergibt sich

$$y = e \circ y = x^{-1} \circ x \circ y = x^{-1} \circ x \circ z = e \circ z = z.$$

Dies ergibt die Linkskürzungsregel. Die Rechtskürzungsregel zeigt man analog. $\square$

Aus den Kürzungsregeln folgt unter anderem, dass das Inverse eindeutig ist. Somit ist die Schreibweise x^{-1} nicht mehrdeutig.

PROPOSITION 1.5. *Sei $(X, \circ)$ eine endliche Menge mit Verknüpfung, die (Va), (Vn) und die Rechtskürzungsregel erfüllt. Dann ist $(X, \circ)$ bereits eine Gruppe.*

Beweis. Sei $x \in X$. Dann ist die Abbildung

$$r_x : X \to X, \qquad y \mapsto y \circ x$$

auf Grund der Rechtskürzungsregel injektiv. Da X eine endliche Menge ist, ist r_x auch surjektiv. Somit gibt es ein $z \in X$ mit $e = r_x(z) = z \circ x$. Das Element z ist also Linksinverses zu x. Da x beliebig gewählt wurde, hat jedes Element ein Linksinverses. Mit der Übungsaufgabe 4 auf Blatt 2 sieht man, dass $(X, \circ)$ eine Gruppe ist. $\square$

Die Proposition gilt auch, wenn man an Stelle der Rechtskürzungsregel die Linkskürzungsregel zur Verfügung hat.

Die Bedingung „endlich“ kann nicht weggelassen werden. Man sieht zum Beispiel, dass $(\mathbb{N}_0, +)$ auch (Va), (Vn) und die Rechtskürzungsregel erfüllt, aber keine Gruppe ist.

Definition 1.6. Eine Teilmenge Y einer Menge mit Verknüpfung $(X, \circ)$ heißt *abgeschlossen (bezüglich $\circ$)*, falls für alle $x, y \in Y$ auch $x \circ y \in Y$.

Die Menge $\{0, 1\}$ ist nicht abgeschlossen in $(\mathbb{N}_0, +)$, aber abgeschlossen in $(\mathbb{N}_0, \cdot)$. Die Menge $\mathbb{N}$ ist abgeschlossen in $(\mathbb{N}_0, +)$ und $(\mathbb{N}_0, \cdot)$. Die Menge der ganzen Zahlen $\mathbb{Z}$ ist abgeschlossen in $(\mathbb{Q}, +)$ und $(\mathbb{Q}, \cdot)$.

Definition 1.7. Sei $(G, \circ)$ eine Gruppe. Sei H eine Teilmenge und $\times$ eine Verknüpfung auf H. Man sagt $(H, \times)$ ist eine Untergruppe von $(G, \circ)$ falls $(H, \times)$ eine Gruppe ist und

$$x \circ y = x \times y \qquad \forall x, y \in H.$$

Wir schreiben dann meistens auch $\circ$ an Stelle von $\times$.

Beispiele. $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{Q}, +)$, dies eine Untergruppe von $(\mathbb{R}, +)$ und dies von $(\mathbb{C}, +)$. Für $n \in \mathbb{N}_0$ ist die Menge $n\mathbb{Z} := \{nk \,|\, k \in \mathbb{Z}\}$ mit der Addition eine Untergruppe von $(\mathbb{Z}, +)$. Ist $(K, \circ)$ eine Untergruppe von $(H, \circ)$ und $(H, \circ)$ eine Untergruppe von $(G, \circ)$, so ist $(K, \circ)$ eine Untergruppe von $(G, \circ)$. Die Relation „ist Untergruppe von“ ist eine Ordnungsrelation auf den Gruppen. Die Gruppe $(\{-1, 1\}, \cdot)$ ist keine Untergruppe von $(\mathbb{Z}, +)$.

Das neutrale Element e_G von G und das neutrale Element e_H sind gleich. Denn es gilt $e_H \circ e_H = e_H = e_G \circ e_H$ und aus der Rechtskürzungsregel in G folgt $e_H = e_G$. Da in G die Kürzungsregel gilt, stimmt für jedes $x \in H$ das Inverse von x in G mit dem Inversen von x in H überein.

Proposition 1.8. *Sei $(G, \circ)$ eine Gruppe, und H eine Teilmenge. Äquivalent sind:*

(1) *H mit der eingeschränkten Verknüpfung $\circ : H \times H \to H$ ist eine Untergruppe.*
(2) (a) *$e_G \in H$, und*
 (b) *H ist abgeschlossen bezüglich $\circ$, und*
 (c) *ist $x \in H$, so ist auch x^{-1} in H.*
(3) (a) *H ist nichtleer, und*
 (b) *sind $x, y \in H$, so ist auch $x^{-1} \circ y \in H$.*

Beweis.
“(1) $\Rightarrow$ (2)”: klar.
“(2) $\Rightarrow$ (1)”: Die eingeschränkte Verknüpfung ist offensichtlich assoziativ, das Element e_G ist auch das neutrale Element von H und Bedingung (c) garantiert ein Inverses. Die Menge H ist mit der eingeschränkten Verknüpfung also eine Gruppe und somit eine Untergruppe von $(G, \circ)$.
“(2) $\Rightarrow$ (3)”: klar.
“(3) $\Rightarrow$ (2)”: Da H nicht leer ist, existiert ein $x \in H$. Mit Aussage (3b) folgt $e_G = x^{-1} \circ x \in H$

also (2a). Wenden wir nun (3b) mit $y = e_G$ an, so ergibt sich (2c). Um die Abgeschlossenheit von $\circ$ zu zeigen, nehmen wir $a, b \in H$ her und wollen $a \circ b \in H$ zeigen. Zunächst folgern wir aus dem bereits bewiesenen (2c) die Aussage $a^{-1} \in H$. Wir wenden nochmal (3b) für $x := a^{-1}$ und $y := b$ an und sehen

$$a \circ b = (a^{-1})^{-1} \circ b \in H.$$

Und alles zu zeigende ist gezeigt. □

Raum der Funktionen. Sei M eine Menge, $(G, \times)$ eine Menge mit Verknüpfung. Wir definieren auf der Menge $\mathrm{Abb}(M, G)$ aller Abbildungen eine Verknüpfung $\boxtimes$. Seien $f, g \in \mathrm{Abb}(M, G)$ dann definieren wir $f \boxtimes g \in \mathrm{Abb}(M, G)$ durch

$$(f \boxtimes g)(m) \quad = \quad f(m) \times g(m)$$

für alle $m \in M$.

Ist $(G, \times)$ eine Gruppe, so ist $(\mathrm{Abb}(M, G), \boxtimes)$ ebenfalls eine Gruppe. Das neutrale Element ist die Abbildung $M \to G, \quad m \mapsto e_G$ und das Inverse von f ist die Abbildung $m \mapsto (f(m))^{-1}$. Die konstanten Funktionen, d.h. Funktionen der Form $M \to G$ $m \mapsto x$ für festes x bilden eine Untergruppe von $\mathrm{Abb}(M, G)$.

Beispiel:

(1) Sei $M = \mathbb{R}$, $(G, \times) = (\mathbb{R}, +)$ und $f(x) = x$, $g(x) = x^2$. Dann ist
$$(f \boxplus g)(x) = f(x) + g(x) = x^2 + x$$
(2) Sei $M = \mathbb{R}$, $(G, \times) = (\mathbb{R}, \cdot)$ und $f(x) = x$, $g(x) = x^2$. Dann ist
$$(f \boxdot g)(x) = f(x) \cdot g(x) = x^3$$

Man schreibt an Stelle von f oft $(f(m))_{m \in M}$. Wenn man diese Notation verwendet, nennt man $(f(m))_{m \in M}$ ist eine *Familie von Elementen von* G und nennt M die *Indexmenge* der Familie.

Im Fall $M = \{1, 2, \ldots, n\}$ interpretieren wir f als das n-Tupel, d.h. als geordnete Menge mit n Elementen und schreiben $(f(1), f(2), \ldots, f(n))$. Die 2-Tupeln interpretiert man auch als Paare. Die Menge aller n-Tupeln schreiben wir als G^n oder

$$\underbrace{G \times \ldots \times G}_{n\text{-mal}}.$$

Beispiel: $(5, 7, 9) \in \mathbb{R}^3$ steht für die Funktion $\{1, 2, 3\} \to \mathbb{R}$, $1 \mapsto 5$, $2 \mapsto 7$, $3 \mapsto 9$. Man kann $(5, 7, 9)$ aber auch als Vektor im drei-dimensionalen Raum $\mathbb{R}^3$ ansehen.

Additive und multiplikative Notation. Häufig notiert man die Verknüpfung additiv, also $+$, oder multiplikativ, also $\cdot$. Man nennt die Gruppe dann eine *additive bzw. multiplikative Gruppe.* Das Adjektiv „additiv" ist somit im Gegensatz zu „abelsch" oder „assoziativ" keine Eigenschaft einer Menge mit Verknüpfung, sondern bedeutet, dass die Verknüpfung $+$ notiert wird. Man hält sich normalerweise an die Konvention, dass die Verknüpfung $+$ normalerweise abelsch (kommutativ) ist. Nicht-kommutative Gruppen werden immer multiplikativ notiert. Aber es gibt auch abelsche multiplikative Gruppen, z. B. $(\mathbb{R}^*, \cdot)$.

Bei additiver Notation bezeichnet 0 das neutrale Element und $-x$ das Inverse zu x. Für $n \in \mathbb{N}$ und $x \in G$ definiert man

$$n \cdot x := \underbrace{x + x + \cdots + x}_{n-\text{mal}} \qquad 0 \cdot x := 0 \qquad (-n) \cdot x := -(n \cdot x).$$

Bei multiplikativer Notation wird der Multiplikationspunkt oft weggelassen, das neutrale Element wird mit 1 bezeichnet und das Inverse zu x mit x^{-1}. Für $n \in \mathbb{N}$ und $x \in G$ definiert man

$$x^n := \underbrace{x \cdot x \cdots x}_{n-\text{mal}} \qquad x^0 := 1 \qquad x^{-n} := (x^n)^{-1}.$$

2. Ringe

Definition 2.1. Ein Ring ist eine Menge R mit zwei Verknüpfungen $+$ und $\cdot$, so dass $(R, +)$ eine kommutative Gruppe ist ((Aa), (An), (Ai) und (Ak) gelten) und außerdem (Ma) und (AMd) gelten. Der Ring $(R, +, \cdot)$ ist *kommutativ*, falls (Mk) gilt, und er ist ein *Ring mit* 1, falls (Mn) gilt.

Beispiele.

$(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ sind kommutative Ringe mit 1.

$(2\mathbb{Z}, +, \cdot)$ ist der Ring der geraden Zahlen. Er ist kommutativ, aber hat keine 1.

Eine Menge mit einem Element sei mit den einzigen möglichen Operationen Addition und Multiplikation versehen. Dies ist ein kommutativer Ring mit 1. Es gilt hier $0 = 1$. Wir nennen „diesen“ Ring den Nullring.

$n \times n$-Matrizen (nächstes Kapitel) bilden einen nicht-kommutativen Ring mit 1.

Es ist bereits bekannt, dass die Null und die Inversen der Addition eindeutig sind.

Proposition 2.2. *Es gilt in einem Ring* $(R, +, \cdot)$

(1) $\forall r \in R : 0 \cdot r = 0 = r \cdot 0$,
(2) *Es existiert höchstens eine Eins in R,*
(3) *Falls eine Eins existiert und* $0 = 1$*, dann ist R der Nullring.*

Beweis.
„(1)“:

$$0 + 0 \cdot r = 0 \cdot r = (0 + 0) \cdot r = 0 \cdot r + 0 \cdot r$$

Wir dürfen in der Gruppe $(R, +)$ kürzen und erhalten

$$0 = 0 \cdot r.$$

„(2)“: Seien e_1 und e_2 zwei Einsen in R.

$$e_1 = e_1 \cdot e_2 = e_2.$$

„(3)“: Wenn $0 = 1$ gilt, dann sehen wir für jedes $r \in R$:

$$r = 1 \cdot r = 0 \cdot r = 0.$$

Dies heißt $R = \{0\}$. □

DEFINITION 2.3. Sei $(R, +, \cdot)$ ein Ring. Sei S eine Teilmenge und $\oplus$ und $\otimes$ zwei Verknüpfungen auf S. Man sagt $(S, \oplus, \otimes)$ ist ein *Unterring* $(R, +, \cdot)$ falls $(S, \oplus, \otimes)$ ein Ring ist und

$$x + y = x \oplus y \qquad \forall x, y \in H.$$

$$x \cdot y = x \otimes y \qquad \forall x, y \in H.$$

Falls R und S Ringe mit Einsen 1_R und 1_S sind und falls $1_R = 1_S$, so nennt man ihn einen *Unterring mit* 1.

Wir schreiben dann meistens auch $+$ und $\cdot$ an Stelle von $\oplus$ und $\otimes$.

BEISPIELE. $(\mathbb{Z}, +)$ ist ein Unterring mit 1 von $(\mathbb{Q}, +)$, dies ein Unterring mit 1 von $(\mathbb{R}, +)$. Die Relation „ist Unterring von“ ist eine Ordnungsrelation auf den Ringen.

Raum der Funktionen. Sei $(R, +, \cdot)$ ein Ring und M eine Menge. Mit derselben Konstruktion wie oben erhalten wir eine Addition $\boxplus$ und eine Multiplikation $\boxdot$ auf $\mathrm{Abb}(M, R)$. Dann ist $(\mathrm{Abb}(M, R), \boxplus, \boxdot)$ ebenfalls ein Ring. Er ist kommutativ bzw. mit 1, falls R kommutativ bzw. mit 1 ist. Die konstanten Abbildungen bilden einen Unterring.

Achtung: Wir haben zwei 2 Multiplikationen auf $\mathbb{R}^2 = \mathrm{Abb}(\{1, 2\}, \mathbb{R})$: die soeben definierte und die Multiplikation komplexer Zahlen.

Polynom-Funktionen. Sei R ein **kommutativer** Ring mit 1. Eine Funktion $P : R \to R$ heißt Polynom-Funktion, falls es ein $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in R$ gibt, so dass

$$P(r) := \sum_{j=0}^{n} a_j r^j.$$

Wir bezeichnen die Menge der Polynom-Funktionen auf R mit $\mathrm{Abb}_{\mathrm{Pol}}(R, R)$. Die Polynom-Funktionen kann man mit $\boxplus$ addieren, mit $\boxdot$ multiplizieren und mit $\circ$ verketten. Man sieht leicht, dass $(\mathrm{Abb}_{\mathrm{Pol}}(R, R), \boxplus, \boxdot)$ ein Unterring von $(\mathrm{Abb}(R, R), \boxplus, \boxdot)$ ist.

Invertierbare Elemente Ein Element x eines Rings mit 1 $(R, +, \cdot)$ heißt *invertierbar*, falls es ein $y \in R$ gibt mit $yx = xy = 1$. Falls a und b Inverse a^{-1} und b^{-1} besitzen, so ist $b^{-1}a^{-1}$ ein Inverses von ab. Da a ein Inverses von a^{-1} ist und 1 invertierbar ist, ist

$$R^* := \{r \in R \,|\, r \text{ invertierbar}\}$$

eine multiplikative Gruppe, die sogenannte *Einheitengruppe*.

Beispiele. $(\{-1, 1\}, \cdot)$ ist die Einheitengruppe von $(\mathbb{Z}, +, \cdot)$. $(\mathbb{Q}^*, \cdot)$ ist die Einheitengruppe von $(\mathbb{Q}, +, \cdot)$.

Nullteiler.

Definition 2.4. Sei $(R, +, \cdot)$ ein Ring. Ein Element $r \in R \setminus \{0\}$ heißt Nullteiler von R, falls es ein $s \in R \setminus \{0\}$ gibt mit $rs = 0$ oder $sr = 0$. Ein Ring ohne Nullteiler heißt *nullteilerfrei*.

Beispiele.
(1) $\overline{2}$ ist ein Nullteiler von $\mathbb{Z}/6\mathbb{Z}$, denn $\overline{2} \cdot \overline{3} = \overline{0}$.
(2) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 1$ für $x > 0$ und $f(x) = 0$ für $x \leq 0$ ist ein Nullteiler von $(\mathrm{Abb}(\mathbb{R}, \mathbb{R}), \boxplus, \boxdot)$. Wenn wir $g(x) := f(-x)$ setzen, so gilt $f(x)g(x) =$ für alle $x \in \mathbb{R}$ und somit $f \boxdot g = 0$.
(3) Sei p Primzahl. Angenommen $\overline{k} \cdot \overline{m} = \overline{0}$ in $\mathbb{Z}/p\mathbb{Z}$. Dann ist also km ein Vielfaches von p. Somit teilt p entweder k oder m, also $\overline{k} = \overline{0}$ oder $\overline{m} = \overline{0}$. Es gibt also keine Nullteiler in $\mathbb{Z}/p\mathbb{Z}$.
(4) Ein Nullteiler in einem Ring mit 1 ist nie invertierbar, falls $0 \neq 1$. Denn für invertierbares x folgt aus $xy = 0$ die Bedingung
$$y = x^{-1}xy = x^{-1}0 = 0.$$

Lemma 2.5. *Falls x kein Nullteiler ist, dann gilt für $y, z \in R$*
$$(yx = zx) \Rightarrow (y = z) \qquad \textit{(x von rechts kürzen)}$$
und
$$(xy = xz) \Rightarrow (y = z) \qquad \textit{(x von links kürzen)}$$

Beweis. Aus $yx = zx$ folgt $(y - z)x = 0$ und somit $y - z = 0$, d.h. $y = z$. Analog von links kürzen. □

Für Elemente x, y in einem **kommutativen** Ring gilt: sind x und y keine Nullteiler, dann ist auch xy kein Nullteiler. Die Menge
$$R_\emptyset := \{x \in R \setminus \{0\} \mid x \text{ ist kein Nullteiler}\}$$
ist also abgeschlossen bezüglich der Multiplikation. Falls eine 1 existiert, haben wir $1 \in R^* \subset R_\emptyset$.

Satz 2.6. *Sei $(R, +, \cdot)$ ein endlicher kommutativer Ring mit 1, $1 \neq 0$. Dann ist $(R_\emptyset, \cdot)$ eine Gruppe. Insbesondere ist jeder endliche, kommutative, nullteilerfreie Ring mit 1 und $1 \neq 0$ bereits ein Körper.*

Beweis. Die Menge $R_\emptyset$ ist endlich, nichtleer, bezüglich $\cdot$ abgeschlossen. Die Multiplikation ist assoziativ, und wir können von rechts kürzen. Proposition 1.5 besagt, dass $(R_\emptyset, \cdot)$ eine Gruppe ist. Falls $(R, +, \cdot)$ nullteilerfrei ist, so ist also jedes Element ungleich Null invertierbar. Somit sind alle Körperaxiome erfüllt. □

FOLGERUNG 2.7 (Proposition 7.3)**.** *Sei $n \in \mathbb{N}_0$. Dann ist $\mathbb{Z}/n\mathbb{Z}$ genau dann ein Körper, wenn n prim ist.*

Beweis. Falls $n = 0$, dann ist $\mathbb{Z}/0\mathbb{Z}$ im wesentlichen dasselbe wie $\mathbb{Z}$, also kein Körper.

Falls $n = 1$, dann besitzt $\mathbb{Z}/1\mathbb{Z}$ ein Element, also kein Körper.

Falls n prim ist, so ist $\mathbb{Z}/n\mathbb{Z}$ ein endlicher, kommutativer, nullteilerfreier Ring mit 1, also ein Körper.

Falls $n \geq 4$ nicht prim ist, so schreiben wir $n = ab$ mit $a, b \geq 2$. Dann ist $ab \equiv 0 \mod n$, aber $a \not\equiv 0 \mod n$ und $b \not\equiv \mod n$. Somit gibt es Nullteiler und $\mathbb{Z}/n\mathbb{Z}$ ist kein Körper. □

3. Körper

In der folgenden Proposition präsentieren wir noch einmal die Definition und äquivalente Versionen davon.

PROPOSITION 3.1. *Sei $\mathbb{K}$ eine Menge mit zwei Verknüpfungen $+$ und $\cdot$. Äquivalent sind:*

(1) $(\mathbb{K}, +, \cdot)$ *ist Körper*
(2) $\mathbb{K}$ *hat mindestens zwei Elemente und es gelten (Aa), (An), (Ai), (Ak), (Ma), (Mn), (Mi), (Mk), (AMd).*
(3) $(\mathbb{K}, +)$ *und* $(\mathbb{K} \setminus \{0\}, \cdot)$ *sind kommutative Gruppen und es gilt (AMd).*

(1) $\Leftrightarrow$ (2) ist Definition, (2) $\Leftrightarrow$ (3) fasst einfach nur zusammen.

BEISPIELE. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}/p\mathbb{Z}$, p prim mit $+$ und $\cdot$.

DEFINITION 3.2. Sei $(\mathbb{K}, +, \cdot)$ ein Körper. Wir sagen $(\mathbb{K}', +, \cdot)$ ist Unterkörper von $(\mathbb{K}, +, \cdot)$, wenn $(\mathbb{K}', +, \cdot)$ ein Unterring ist und selbst Körper ist.

Beispiel: $\mathbb{Q}$ ist Unterkörper von $\mathbb{R}$.

Wenn $(\mathbb{K}', +, \cdot)$ Unterkörper ist, so ist insbesondere $(\mathbb{K}' \setminus \{0\}, \cdot)$ eine Untergruppe von $(\mathbb{K} \setminus \{0\}, \cdot)$. Also ist die 1 von $\mathbb{K}$ auch die 1 von $\mathbb{K}'$.

KAPITEL 3

Matrizen

In diesem Kapitel sei R immer ein kommutativer Ring mit 1 mit Addition $+$ und Multiplikation $\cdot$. Man denke an $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{C}$ oder $\mathbb{Z}/n\mathbb{Z}$.

1. Definition

Sei $n, m \in \mathbb{N}$. Eine $n \times m$-Matrix über R ist eine Abbildung

$$\begin{array}{rcl} \{1,2,\ldots,n\} \times \{1,2,\ldots,m\} & \to & R \\ (i,j) & \mapsto & a_{ij}. \end{array}$$

Die Menge aller $n \times m$-Matrizen über R nennen wir $\mathrm{Mat}(n,m;R)$. Wir schreiben Matrizen normalerweise in einer der folgenden Formen

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} = (a_{ij})_{i \in \{1,\ldots,n\}, j \in \{1,\ldots,m\}} = (a_{ij})$$

Spezialfälle:
Eine $n \times m$-Matrix mit $n = 1$, nennt man *Zeilenvektor.*

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \end{pmatrix}.$$

Offensichtlich ist die Abbildung

$$\mathcal{M}_Z : R^m \to \mathrm{Mat}(1,m;R) \qquad (r_1, r_2, \ldots, r_m) \mapsto \begin{pmatrix} r_1 & r_2 & \cdots & r_m \end{pmatrix}$$

bijektiv.

Eine $n \times m$-Matrix mit $m = 1$, nennt man *Spaltenvektor.*

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$$

Offensichtlich ist die Abbildung

$$\mathcal{M}_S : R^n \to \mathrm{Mat}(n,1;R) \qquad (r_1, r_2, \ldots, r_n) \mapsto \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$$

bijektiv.

In den meisten Büchern identifiziert man n-Tupel mit Spaltenvektoren vermöge $\mathcal{M}_S$, also kurz: $R^n = \mathrm{Mat}(n,1;R)$. Wir wollen dies später auch tun, aber zunächst noch unterscheiden.

Eine $n \times m$-Matrix heißt *quadratische Matrix*, falls $n = m$.

2. Addition und Multiplikation von Matrizen

$n \times m$-Matrizen werden wie Abbildungen $\mathrm{Abb}(\{1,2,\ldots,n\} \times \{1,2,\ldots,m\}, R)$ addiert. Man kann sie nur addieren wenn die Zahl der Zeilen n und der Spalten m übereinstimmt.

$$+ : \mathrm{Mat}(n,m;R) \times \mathrm{Mat}(n,m;R) \to \mathrm{Mat}(n,m;R)$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2m}+b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nm}+b_{nm} \end{pmatrix}$$

Auf Spaltenvektoren erhalten wir wieder die Addition auf R^n.

Die Multiplikation ist anders: Man kann eine $n_1 \times m_1$-Matrix mit einer $n_2 \times m_2$-Matrix multiplizieren, falls

$$m_1 = n_2.$$

Wir definieren dann das Produkt

$$(c_{ij})_{i\in\{1,\ldots,n_1\}, j\in\{1,\ldots,m_2\}} = (a_{ij})_{i\in\{1,\ldots,n_1\}, j\in\{1,\ldots,m_1\}} \cdot (b_{ij})_{i\in\{1,\ldots,n_2\}, j\in\{1,\ldots,m_2\}}$$

durch die Formel

$$c_{ij} := \sum_{k=1}^{m_1} a_{ik} b_{kj}.$$

Beispiele: $R = \mathbb{Z}$:

$$\begin{pmatrix} 1 & 0 & 4 \\ 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} 6 \\ 2 \\ 7 \end{pmatrix} = \begin{pmatrix} 1\cdot 6 + 0\cdot 2 + 4\cdot 7 \\ 2\cdot 6 + 3\cdot 2 + 5\cdot 7 \end{pmatrix} = \begin{pmatrix} 34 \\ 53 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 7 & 9 \end{pmatrix} = \begin{pmatrix} 1 \cdot 7 & 1 \cdot 9 \\ 1 \cdot 3 & 1 \cdot 2 \end{pmatrix} = \begin{pmatrix} 7 & 9 \\ 3 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 \\ 7 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 9 & 7 \end{pmatrix}$$

Es gelten die Distributivgesetze und die Multiplikation ist assoziativ. Die Matrizen-Multiplikation ist aber nicht kommutativ.

Wir definieren

$$\mathbb{1}_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathrm{Mat}(n, n; R).$$

Für eine $n \times m$-Matrix A gilt dann

$$\mathbb{1}_n A = A \mathbb{1}_m = A.$$

3. Multiplikation mit Skalaren

Um Elemente von R von Matrizen $\in \mathrm{Mat}(n, m; R)$ und Zeilen- und Spaltenvektoren zu unterscheiden, nennt man die Elemente von R oft *Skalare*. Man kann Skalare mit Matrizen multiplizieren. Sei $\lambda \in R$, $(a_{ij}) \in \mathrm{Mat}(n, m; R)$. Dann definieren wir

$$\lambda \cdot (a_{ij}) := \Big((\lambda a_{ij}) \Big)$$

Beispiel:

$$2 \cdot \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 0 \end{pmatrix}$$

Achtung: Diese Multiplikation darf man nicht als Skalarmultiplikation bezeichnen, da dieser Begriff etwas anderes bedeutet.

Offensichtlich gilt für $\lambda, \mu \in R$, $A, B \in \mathrm{Mat}(n, m; R)$, $C \in \mathrm{Mat}(m, k; R)$:

$$\begin{aligned} \lambda(AC) &= (\lambda A)C = A(\lambda C) \\ (\lambda\mu)A &= \lambda(\mu A) \\ (\lambda + \mu)A &= \lambda A + \mu A \\ \lambda(A + B) &= \lambda A + \lambda B \end{aligned}$$

4. Transposition von Matrizen

Die Abbildung $\mathrm{Mat}(n,m;R) \to \mathrm{Mat}(m,n;R)$, die die Rolle von Zeilen und Spalten vertauscht, d.h. die Matrix $A = (a_{ij})_{i\in\{1,\ldots,n\},j\in\{1,\ldots,m\}}$ auf die Matrix $A^T := (a_{ji})_{i\in\{1,\ldots,m\},j\in\{1,\ldots,n\}}$ abbildet, nennen wir die Transposition von Matrizen. Wir nennen A^T die transponierte Matrix von A.

Beispiele. (1)

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

(2) Für $v \in R^n$ gilt $(\mathcal{M}_S(v))^T = \mathcal{M}_Z(v)$ und $(\mathcal{M}_Z(v))^T = \mathcal{M}_S(v)$.

Es gilt für Matrizen $A, B \in \mathrm{Mat}(n,m;R)$ und $C \in \mathrm{Mat}(m,k;R)$:

(1) $(A^T)^T = A$
(2) $(A+B)^T = A^T + B^T$
(3) $(AC)^T = C^T A^T$

5. Matrizen und lineare Abbildungen

Sei $A = (a_{ij})$ eine $n \times m$-Matrix. Dann definieren wir

$$\begin{aligned} \mathcal{L}_A : \mathrm{Mat}(m,1;R) = R^m &\to \mathrm{Mat}(n,1;R) = R^n \\ v &\mapsto \mathcal{L}_A(v) := Av \end{aligned}$$

Es gilt dann für $v, w \in R^m$, $\mu \in R$

$$\mathcal{L}_A(v+w) = \mathcal{L}_A(v) + \mathcal{L}_A(w),$$

$$\mathcal{L}_A(\lambda v) = \lambda \mathcal{L}_A(v).$$

Abbildungen mit diesen beiden Eigenschaften heißen *linear*. Wir nennen $\mathcal{L}_A$ die *zu A assoziierte lineare Abbildung*. Ob der Buchstabe $\mathcal{L}$ hier für „Linskmultiplikation“ oder für „linear“ steht, darf der Leser selbst entscheiden.

Umgekehrt sei $\mathcal{L} : R^m \to R^n$ eine lineare Abbildung. Wir definieren die Matrix $A \in \mathrm{Mat}(n,m;R)$ durch

$$A := \left(\mathcal{L}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) \mathcal{L}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) \cdots \mathcal{L}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right) \right).$$

Man rechnet leicht nach, dass $\mathcal{L}_A = \mathcal{L}$. Also ist jede lineare Abbildung $R^m \to R^n$ durch eine Matrix beschrieben.

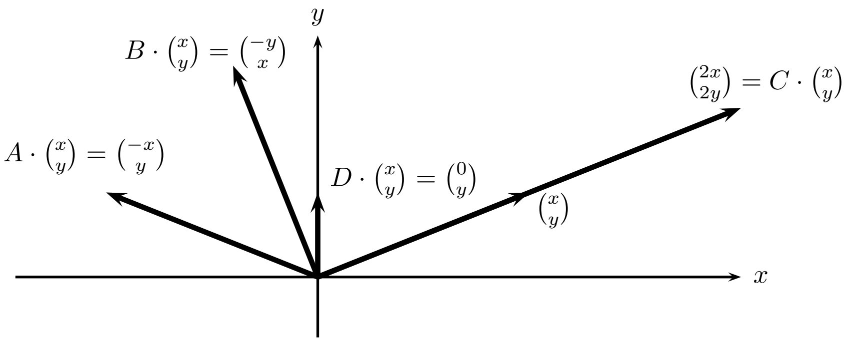


ABBILDUNG 1. Die Abbildungen $\mathcal{L}_A$, $\mathcal{L}_B$, $\mathcal{L}_C$ und $\mathcal{L}_D$

BEISPIELE.

(1) Gilt $A = \mathbb{1}_n \in \mathrm{Mat}(n, n; R)$, so ist $\mathcal{L}_A$ die Identität von R^n.
(2) $R = \mathbb{R}$, $n = m = 2$,

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

$\mathcal{L}_A$ ist die Spiegelung an der y-Achse, $\mathcal{L}_B$ ist die Drehung um 90 Grad im Gegenuhrzeigersinn um den Ursprung, $\mathcal{L}_C$ ist die zentrische Streckung am Ursprung um den Faktor 2, und $\mathcal{L}_D$ ist die Orthogonalprojektion auf die y-Achse, siehe Abbildung 2.
(3) Ist $A \in \mathrm{Mat}(3, 1; \mathbb{R}) = \mathbb{R}^3$, dann ist $\mathcal{L}_A : \mathbb{R} \to \mathbb{R}^3$ die Parametrisierung einer Ursprungsgeraden in Richtung des Spaltenvektors A. Falls $A = 0$, dann ist die parametrisierte Gerade degeneriert: es ist ein Punkt, der Ursprung.
(4) Ist $A \in \mathrm{Mat}(3, 2; \mathbb{R})$, dann ist $\mathcal{L}_A : \mathbb{R}^2 \to \mathbb{R}^3$ die Parametrisierung einer Ursprungsebene in $\mathbb{R}^3$. Falls $A = 0 \in \mathrm{Mat}(3, 2; \mathbb{R})$ so degeneriert die Ebene zu einem Punkt. Im Fall $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}$ ist das Bild von $\mathcal{L}_A$ eine Gerade in Richtung $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

Verkettung und Matrizen-Multiplikation: Es gilt für $A \in \mathrm{Mat}(n,m;R)$ und $B \in \mathrm{Mat}(m,k;R)$:

$$\mathcal{L}_{AB} = \mathcal{L}_A \circ \mathcal{L}_B.$$

6. Lineare Gleichungssysteme

Gegeben sei ein lineares Gleichungssystem über R:

$$\begin{array}{ccccc} a_{11}x_1 & +a_{12}x_2 & +\cdots & +a_{1m}x_m & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\cdots & +a_{2m}x_m & = & b_2 \\ . & . & & . & . & . \\ . & . & & . & . & . \\ . & . & & . & . & . \\ . & . & & . & . & . \\ a_{n1}x_1 & +a_{n2}x_2 & +\cdots & +a_{nm}x_m & = & b_n \end{array}$$

Hierbei sind im Normalfall die $a_{ij} \in R$ und die $b_i \in R$ gegeben und die Gesamtheit der möglichen $x_i \in R$ ist zu bestimmen.

Wir fassen zusammen $A := (a_{ij}) \in \mathrm{Mat}(n,m;R)$, $x = \mathcal{M}_S(x_1,\ldots,x_m)$ und $b = \mathcal{M}_S(b_1,\ldots,b_n)$. Dann ist das Gleichungssystem äquivalent zur Gleichung

$$Ax = b.$$

Ein Gleichungssystem heißt *homogen*, falls $b = 0$, und ansonsten *inhomogen*. Falls $Ax = b$ ein Gleichungssystem ist, so heißt $Ax = 0$ das *zugehörige homogene Gleichungssystem*.

Proposition 6.1. *Sei $Ax = 0$ ein lineares Gleichungssystem, $A \in \mathrm{Mat}(n,m;R)$ gegeben, $x \in R^m$ gesucht. Dann gilt für den Raum der Lösungen $\mathrm{Lös}(Ax = 0)$:*

(1) *$\mathrm{Lös}(Ax = 0)$ ist eine Untergruppe von $(R^m, +)$.*
(2) *Falls $x \in \mathrm{Lös}(Ax = 0)$ und $\lambda \in R$, dann gilt auch $\lambda x \in \mathrm{Lös}(Ax = 0)$.*

Insbesondere ist also $0 \in R^m$ immer eine Lösung.

Die konkrete Lösungsmenge kann man mit dem Gaußschen Verfahren berechnen.

Frage: Sei L eine Teilmenge, die (1) und (2) erfüllt. Gibt es ein homogenes Gleichungssystem über R, so dass L die Menge der Lösungen ist?

Die Anwort ist Nein für $R = \mathbb{Z}$, $m = 1$, $L = 2\mathbb{Z}$. Wir werden sehen, dass die Antwort Ja ist, falls ein R Körper ist. Falls R ein Körper ist, so nennt man eine Teilmenge, die die Eigenschaft (1) und (2) besitzt, einen *Untervektorraum* von R^m.

Beweis der Proposition. $A0 = 0$, also ist $0 \in \mathrm{Lös}(Ax = 0)$, und somit ist $\mathrm{Lös}(Ax = 0)$ nicht leer. Falls $x, y \in \mathrm{Lös}(Ax = 0)$, dann gilt $Ax = Ay = 0$ und somit auch $A(x - y) = Ax - Ay = 0 - 0 = 0$, das heißt $x - y \in \mathrm{Lös}(Ax = 0)$. Also ist $\mathrm{Lös}(Ax = 0)$ eine Untergruppe von $(R^m, +)$.

Ist $x \in \mathrm{Lös}(Ax = 0)$ und $\lambda \in R$, so gilt $A(\lambda x) = \lambda(Ax) = \lambda 0 = 0$ und somit $\lambda x \in \mathrm{Lös}(Ax = 0)$. □

Inhomogene Gleichungssysteme können gar keine Lösungen besitzen: Man sieht mit dem Gaußschen Verfahren, dass das Gleichungssystem über $\mathbb{R}$

$$\begin{aligned} x_1 + 2x_2 &= 1 \\ 2x_1 + 4x_2 &= 0 \end{aligned}$$

keine Lösungen besitzt.

PROPOSITION 6.2. *Angenommen $\tilde{x}$ sei eine Lösung des inhomogenen Systems $Ax = b$. Dann gilt für die Menge aller Lösungen von $Ax = b$*

$$\textit{Lös}(Ax = b) = \{\tilde{x} + x \mid x \in \textit{Lös}(Ax = 0)\}.$$

Beweis. Seien $\tilde{x}$ und $\hat{x}$ Lösungen des inhomogenen Systems, dann ist $A(\tilde{x} - \hat{x}) = A\tilde{x} - A\hat{x} = b - b = 0$. Also ist $\tilde{x} - \hat{x}$ eine Lösung des homogenen Systems $Ax = 0$. Es folgt

$$\text{Lös}(Ax = b) \subseteq \{\tilde{x} + x \mid x \in \text{Lös}(Ax = 0)\}.$$

Umgekehrt sei x eine Lösung des homogenen Systems und $\tilde{x}$ eine Lösung des inhomogenen. Dann folgt $A(\tilde{x} + x) = b + 0 = b$, und deswegen ist $\tilde{x} + x$ auch eine Lösung des inhomogenen Systems. Es folgt

$$\text{Lös}(Ax = b) \supseteq \{\tilde{x} + x \mid x \in \text{Lös}(Ax = 0)\}.$$

□

Ob ein inhomogenes System eine Lösung besitzt und falls ja, wieviele und welche kann man wiederum mit dem Gaußschen Verfahren berechnen.

BEISPIEL. Eine Ebene in $\mathbb{R}^3$ wird durch ein Gleichungssystem mit einer Zeile beschrieben, d. h. durch eine Gleichung der Form $Ax = b$, wobei $A \neq 0$ ein Zeilenvektor ist, $x \in \mathbb{R}^3$ bzw. Spaltenvektor, $b \in \mathbb{R}$. Man nennt A einen Normalenvektor der Ebene. Das zugehörige homogene Gleichungssystem ist $Ax = 0$. Lösungen hiervon sind alle Vektoren, die senkrecht auf A^T stehen. Man sieht leicht, dass

$$\tilde{x} = \frac{b}{\|A\|^2} A^T$$

eine spezielle Lösung von $Ax = b$ ist. Somit liegt x auf der Ebene, genau dann wenn $x - \tilde{x}$ senkrecht auf A^T steht.

BEISPIEL. Schnitt zweier Geraden im Raum. Wir nehmen an, dass

$$t \mapsto p + tv \qquad s \mapsto q + sw$$

zwei Geraden in $\mathbb{R}^3$ sind, $p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \in \mathbb{R}^3$, $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in \mathbb{R}^3 \setminus \{0\}$. Gesucht ist die Menge der Schnittpunkte. Hierzu ist das System

$$p + tv = q + sw$$

zu lösen. In Matrix-Schreibweise ergibt dies

$$\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \begin{pmatrix} t \\ -s \end{pmatrix} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix}$$

Dieses System kann überhaupt keine Lösungen haben. Dies tritt zum Beispiel auf, wenn die Geraden parallel, aber nicht identisch sind. Parallel besagt, dass es $\lambda \in \mathbb{R}^*$ gibt mit $v = \lambda w$. Es kann aber auch keine Lösungen geben, obwohl die Geraden nicht parallel sind. Beispiel $p = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $q = 0$, $p = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $w = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. In diesem Fall sagt man, die Geraden sind *windschief* zueinander.

Wenn es mindestens eine Lösung t_0, s_0 gibt, dann kann man das zugehörige homogene System betrachten:

$$\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \begin{pmatrix} t \\ -s \end{pmatrix} = 0$$

Wenn die Geraden (also v und w) parallel sind, dann hat dieses homogene System einen 1-dimensionalen Lösungsraum. Die Geraden sind gleich, lediglich die Parametrisierungen sind evtl. verschieden. Wenn v und w nicht parallel sind, so hat das homogene System nur die Null als Lösung. Der Schnittpunkt ist also eindeutig.

Um diese vier Fälle zu unterscheiden, betrachtet man also am besten

$$V := \{tv + sw \mid t, s \in \mathbb{R}\} \qquad d := p - q.$$

V ist wieder ein Untervektorraum, er ist 1-dimensional (v, w parallel) oder sonst 2-dimensional. Lösungen gibt es gdw $d \in V$.

Zusammenfassung

	$d \in V$	$d \notin V$
$\dim V = 1$	1-dim Lösungsraum Geraden gleich	keine Lösung Geraden parallel
$\dim V = 2$	eindeutige Lösung Geraden schneiden sich	keine Lösung Geraden windschief

7. Quadratische Matrizen

Sei $n \in \mathbb{N}$. Die Menge $\mathrm{Mat}(n, n; R)$ mit Matrizenaddition und Matrizenmultiplikation versehen, bildet einen Ring: (Aa), (An), (Ai), (Ak) sind klar, (Ma) und (AMd) haben wir nachgerechnet. Die Matrix $\mathbb{1}_n$ ist eine 1 in diesem Ring von Matrizen. Er ist nicht kommutativ für $n > 1$.

Es gibt viele Nullteiler, z. B.

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

$$\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} = 0$$

Frage 7.1. Wenn A und B keine Nullteiler in $\mathrm{Mat}(n, n, ; R)$ sind, ist dann auch AB kein Nullteiler? Wenn A Nullteiler in $\mathrm{Mat}(n, n; R)$ ist und $B \in \mathrm{Mat}(n, n; R)$ beliebig, ist dann auch AB Null oder ein Nullteiler?

Wiederholung: Eine Matrix $A \in \mathrm{Mat}(n, n; R)$ nennen wir invertierbar, falls es ein $B \in \mathrm{Mat}(n, n; R)$ gibt mit $AB = BA = \mathbb{1}_n$. Die invertierbaren Matrizen bilden eine multiplikative Gruppe, also sind die Inversen eindeutig. Man nennt B kurz *das Inverse von* A oder *die inverse Matrix zu* A. Alle Nullteiler sind nicht invertierbar.

Frage 7.2. Ist jede nicht invertierbare Matrix ein Nullteiler?

Die Antwort ist „Nein“, wenn $R = \mathbb{Z}$, denn $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}(2, 2; \mathbb{Z})$ ist weder invertierbar noch Nullteiler in $\mathrm{Mat}(2, 2; \mathbb{Z})$.

Frage 7.3. Sei $A \in \mathrm{Mat}(n, n; R)$. Sind folgenden Aussagen äquivalent?

(1) A ist invertierbar,
(2) A besitzt ein Rechtsinverses, d. h. es gibt ein $B \in \mathrm{Mat}(n, n; R)$ mit $AB = \mathbb{1}_n$,
(3) A besitzt ein Linksinverses, d. h. es gibt ein $B \in \mathrm{Mat}(n, n; R)$ mit $BA = \mathbb{1}_n$?

Frage 7.4. Wenn ein Rechtsinverses existiert, ist es eindeutig?

Im folgenden betrachten wir vor allem Matrizen über Körpern. Wir werden sehen, dass dann die Antwort auf all diese Fragen „Ja“ sein wird.

Geometrische Interpretation: Falls eine lineare Abbildung $\mathcal{L}_A : R^n \to R^n$ durch eine Matrix $A \in \mathrm{Mat}(n, n; R)$ beschrieben wird, dann ist die Suche nach einer Inversen B gleichbedeutend zu der Suche nach einer Umkehrabbildung $\mathcal{L}_B$.

KAPITEL 4

Vektorräume

In diesem Kapitel werden Vektorräume definiert und untersucht. Ziele: systematische Beschreibung von Gleichungssystemen, Beschreibung von Untervektorräumen ohne den umgebenden Raum, Bereitstellung grundlegender Konzepte der mathematischen Physik (QM, PDG, etc).

In diesem Kapitel sei $\mathbb{K}$ mit Addition und Multiplikation versehen immer ein Körper.

1. Definition und elementare Eigenschaften

Definition 1.1. Ein *Vektorraum* über $\mathbb{K}$ (oder $\mathbb{K}$-Vektorraum) ist eine Menge V versehen mit einer Addition

$$+ : V \times V \to V, \qquad (v, w) \mapsto v + w$$

und einer Multiplikation mit Skalaren

$$\cdot : \mathbb{K} \times V \to V, \qquad (\lambda, v) \mapsto \lambda \cdot v$$

mit den folgenden Eigenschaften

(1) $(V, +)$ ist eine abelsche Gruppe.
(2) Für alle $v \in V$ gilt $1 \cdot v = v$.
(3) (Assoziativität): Für alle $\lambda, \mu \in \mathbb{K}$ und $v \in V$ gilt

$$(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v).$$

(4) (Distributivität) Für alle $\lambda, \mu \in \mathbb{K}$ und $v, w \in V$ gilt

$$\begin{aligned} (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot v \\ \lambda \cdot (v + w) &= \lambda \cdot v + \lambda \cdot w \end{aligned}$$

Es versteht sich von selbst, dass mit 1 die 1 von $\mathbb{K}$ gemeint ist.

Beachten Sie, dass $(V, \cdot)$ keine Menge mit Verknüpfung im obigen Sinne ist: die Multiplikation ist auf $\mathbb{K} \times V$ und nicht auf $V \times V$ definiert

Elemente eines Vektorraums nennt man *Vektoren*, Elemente des Körpers *Skalare*.

Beispiele.

(1) Auf $\mathbb{K}^n$ führen wir die folgende Multiplikation mit Skalaren ein. Für $\lambda, x_1, x_2, \dots, x_n \in \mathbb{K}$ definieren wir
$$\lambda \cdot (x_1, x_2, \dots, x_n) := (\lambda x_1, \lambda x_2, \dots, \lambda x_n).$$
Dann ist $\mathbb{K}^n$ mit Addition und mit Multiplikation mit Skalaren versehen ein $\mathbb{K}$-Vektorraum.
(2) Die Lösungsmenge eines homogenen linearen Gleichungssystems über $\mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum.
(3) $\mathrm{Mat}(n, m; \mathbb{K})$ mit der Addition von Matrizen und der Multiplikation mit Skalaren ist ein $\mathbb{K}$-Vektorraum.
(4) Falls $\mathbb{K}$ Unterkörper von $\mathbb{L}$ ist, so ist $\mathbb{L}$ ein $\mathbb{K}$-Vektorraum. Insbesondere ist $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum und $\mathbb{Q}$-Vektorraum und $\mathbb{R}$ ein $\mathbb{Q}$-Vektorraum.
(5) Eine Menge mit genau einem Element kann auf genau eine Art zu einem $\mathbb{K}$-Vektorraum gemacht werden. Man nennt dies „den“ Nullvektorraum.

Lemma 1.2. *Es gilt:*

(1) *Für alle $\lambda \in \mathbb{K}$ gilt $\lambda \cdot 0_V = 0_V$.*
(2) *Für alle $v \in V$ gilt $0_K \cdot v = 0_V$.*
(3) *Für alle $\lambda \in \mathbb{K}$ und $v \in V$ gilt $(-\lambda) \cdot v = \lambda \cdot (-v) = -(\lambda \cdot v)$.*
(4) *Aus $\lambda \cdot v = 0_V$, $\lambda \in \mathbb{K}$, $v \in V$ folgt $\lambda = 0_K$ oder $v \in 0_V$.*

Beweis.
(1):
$$0_V + \lambda \cdot 0_V = \lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V.$$
Durch Rechtskürzen von $\lambda \cdot 0_V$ folgt die Behauptung.
(2): Analog.
(3): $(-\lambda) \cdot v + \lambda \cdot v = (-\lambda + \lambda) \cdot v = 0_K \cdot v = 0_V$. Also $(-\lambda) \cdot v = -(\lambda \cdot v)$. Die Gleichung $\lambda \cdot (-v) = -(\lambda \cdot v)$ zeigt man analog.
(4) Falls $\lambda \cdot v$ und $\lambda \neq 0$, dann ist λ invertierbar. Es gilt dann
$$0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda \cdot v) = (\lambda^{-1} \cdot \lambda) \cdot v = 1_K \cdot v = v.$$
□

2. Untervektorräume und Erzeugendensysteme

Definition 2.1. Sei $(V, +, \cdot)$ ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $U \subseteq V$ versehen mit einer Addition $\boxplus : U \times U \to U$ und einer Multiplikation mit Skalaren $\boxdot : \mathbb{K} \times U \to U$ heißt *Untervektorraum* von $(V, +, \cdot)$ falls

(1) $(U, \boxplus, \boxdot)$ ist ein $\mathbb{K}$-Vektorraum.

(2) Für alle $\lambda \in \mathbb{K}$ und $v, w \in U$ gilt

$$\lambda \cdot v = \lambda \boxdot v \qquad v + w = v \boxplus w.$$

Wiederum bestimmen $+$ und $\cdot$ durch Einschränkung auf U eindeutig $\boxplus$ und $\boxdot$, und wir schreiben auch $+$ und $\cdot$ für $\boxplus$ und $\boxdot$.

Beispiele: $\{0\}$ ist ein Untervektorraum von V.
Ursprungsgeraden und Ursprungsebenen in $\mathbb{R}^3$ sind Untervektorräume von $\mathbb{R}^3$.
(Affine) Ebenen in $\mathbb{R}^3$, die nicht 0 enthalten, sind keine Untervektorräume.

Lemma 2.2. *Sei $(V, +, \cdot)$ ein $\mathbb{K}$-Vektorraum. Eine Teilmenge mit der auf U eingeschränkten Addition und Multiplikation mit Skalaren ist ein Untervektorraum, genau dann wenn*

(1) *$U \neq \emptyset$,*
(2) *Falls $u, v \in U$, dann gilt auch $u + v \in U$, d. h. U ist abgeschlossen bezüglich der Addition.*
(3) *Falls $\lambda \in \mathbb{K}$ und $u \in U$, so ist $\lambda \cdot u \in U$, d. h. U ist abgeschlossen bezüglich der Multiplikation mit Skalaren.*

Beweis. Offensichtlich erfüllt ein Untervektorraum die Eigenschaften (1) bis (3).

Sei U nun eine Menge, die (1) bis (3) erfüllt. Auf Grund der Eigenschaften (2) und (3) definiert die Einschränkung der Addition und der Multiplikation mit Skalaren auf U Abbildungen $+ : U \times U \to U$ und $\cdot : \mathbb{K} \times U \to U$. Um zu zeigen, dass $(U, +)$ eine Untergruppe ist, nutzen wir zunächst $U \neq \emptyset$ und zu zeigen bleibt dann:

$$\forall u, v \in U : u - v \in U.$$

Wegen $u - v = u + (-1) \cdot v$ folgt dies direkt aus (2) und (3). Die Eigenschaften $1 \cdot v = v$, Assoziativität und Distributivität gelten in U, da sie auch in V gelten. $\square$

Sei $\mathcal{P}(V)$ die Potenzmenge von V, d. h. die Menge aller Teilmengen von V. Eine *Familie von Teilmengen von V* ist eine Abbildung $I \to \mathcal{P}(V)$, $i \mapsto M_i \subseteq V$, die wir $(M_i)_{i \in I}$ oder $(M_i \mid i \in I)$ notieren. Die Menge I nennt man *Indexmenge*. Man kann dann den Schnitt und die Vereinigung definieren:

$$\begin{aligned} \bigcap_{i \in I} M_i \quad &:= \quad \bigcap (M_i \mid i \in I) := \{x \in V \mid x \in M_i \ \forall i \in I\} \\ \bigcup_{i \in I} M_i \quad &:= \quad \bigcup (M_i \mid i \in I) := \{x \in V \mid x \in M_i \ \exists i \in I\} \end{aligned}$$

Beispiel: Für $I := \{1, 2\}$ gilt $\bigcap_{i \in I} M_i = M_1 \cap M_2$.

Wir sagen $(M_i)_{i \in I}$ ist eine Familie von Untervektorräumen, wenn jedes M_i zusätzlich ein Untervektorraum ist (mit der auf M_i eingeschränkten Addition und Multiplikation mit Skalaren).

Proposition 2.3. *Sei $(U_i)_{i \in I}$ eine Familie von Untervektorräumen. Dann ist $\bigcap_{i \in I} U_i$ ebenfalls ein Untervektorraum.*

Beispiel: Sei $I = \mathbb{R}$. Für $\lambda \in I$ definieren wir die Ursprungsebene

$$U_\lambda := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, x_1 + \lambda x_2 = 0 \right\}$$

Jedes U_λ definiert eine Ursprungsebene. Die 3. Achse

$$A_3 := \left\{ \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} \,\middle|\, x_3 \in \mathbb{R} \right\}$$

ist in jedem U_λ enthalten, und man sieht leicht:

$$\bigcap_{\lambda \in I} U_\lambda = A_3 = U_0 \cap U_1.$$

Man überlegt sich leicht, dass $\bigcup_{\lambda \in I} U_\lambda$ **kein** Untervektorraum ist.

Beweis. Da 0 in allen U_i enthalten ist, gilt $0 \in \bigcap_{i \in i} U_i \ \forall i$, also $0 \in \bigcap_{i \in I} U_i$.
Sind $u, v \in \bigcap_{i \in I} U_i$ und $\lambda \in \mathbb{K}$, so gilt $u, v \in U_i \ \forall i \in I$, und somit $u + v \in U_i$ und $\lambda u \in U_i \ \forall i \in I$.
Es ergibt sich $u + v, \lambda u \in \bigcap_{i \in I} U_i$. □

Definition 2.4. Sei $A \subseteq V$. Sei

$$I_{A,V} := \{U \mid U \text{ Untervektorraum von } V \text{ und } A \subseteq U\}.$$

Dann heißt

$$\operatorname{span} A := \bigcap (U \mid U \in I_{A,V})$$

der *von A erzeugte Untervektorraum* oder der *von A aufgespannte Untervektorraum.*

Offensichtlich gilt $A \subseteq \operatorname{span} A$. Also ist $\operatorname{span}(A) \in I_{A,V}$. Der Raum $\operatorname{span} A$ ist also der kleinste Untervektorraum von V, der A umfasst, im folgenden Sinne: Die Relation $\subseteq$ ist eine Ordnungsrelation auf $I_{A,V}$ und $\operatorname{span} A$ ist das Minimum der geordneten Menge $(I_{A,V}, \subseteq)$.

Beispiele. In $V = \mathbb{K}^3$:

(1)
$$\operatorname{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$
ist die erste Koordinaten-Achse.

(2)
$$\operatorname{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$
ist die Ebene, die die ersten beiden Koordinatenachsen enthält.

(3)

$$\operatorname{span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$$

ist $\mathbb{K}^3$.

Sind U und W Untervektorräume eines Vektorraums, so definiert man

$$U+W := \{u+w \mid u\in U,\ w\in W\}.$$

Dann ist $U+W$ ein Untervektorraum und $U\cup W\subseteq U+W$. Andrerseits gilt für jeden Untervektorraum X mit $U\cup W\subseteq X$ auch $U+W\subseteq X$. Somit ist

$$\operatorname{span}(U\cup W)=U+W.$$

Wir sagen $A\subseteq V$ *erzeugt* V oder *ist ein Erzeugendensystem von* V, falls $\operatorname{span}A=V$.

Falls $A\subseteq B\subseteq V$, so gilt auch $\operatorname{span}A\subseteq\operatorname{span}B$. Wenn also A bereits V erzeugt, so tut dies B ebenfalls.

Notation: M Menge, $(M_i|i\in I)$ Familie von Teilmengen

$$\bigcap_{i\in I}M_i=\bigcap(M_i|i\in I)=\bigcap\{M_i|i\in I\}=\{x\in M|\forall i\in I: x\in M_i\}\subseteq M$$

Spezialfall: $Q\subseteq\mathcal{P}(M)$, dann ist $(B|B\in Q)$ eine Familie von Teilmengen, also $\{B|B\in Q\}=Q$ und somit ist es konsistent auch $\bigcap Q$ für $\bigcap(B|B\in Q)$ zu schreiben.

Achtung: für $J\neq\emptyset$: $\bigcap_{j\in J}Q=Q\subset\mathcal{P}(M)$.

Wiederholung: $A\subseteq V$. $I_{A,V}=\{U\subseteq V|A\subseteq U,\quad U \text{ Untervektorraum von } V\}$. Dann

$$\operatorname{span}A:=\bigcap I_{A,V}.$$

Lemma 2.5. *Sei V Vektorraum und $A\subseteq V$ endlich. Dann ist $v\in\operatorname{span}A$ genau dann, wenn es Skalare λ_a, $a\in A$ gibt mit*

$$v=\sum_{a\in A}\lambda_a a.$$

Einen derartigen Summenausdruck nennt man *Linearkombination von Elementen aus A*.

Beweis. Sei U_0 die Menge aller v der Form $\sum_{a\in A}\lambda_a v_a$. Man sieht leicht, dass U_0 ein Untervektorraum von V ist. Es gilt $A\subseteq U_0$, somit $U_0\in I_{A,V}$. Also $\operatorname{span}A=\bigcap I_{A,V}\subseteq U_0$. Jeder Untervektorraum, der A enthält, enthält auch beliebige Linearkombinationen von Elementen aus A und somit gilt $U_0\subseteq U$ für alle $U\in I_{A,V}$. Es folgt $U_0\subseteq\operatorname{span}A=\bigcap I_{A,V}$. □

Wenn A unendlich ist, stellt sich das Problem, dass ein Ausdruck der Form $\sum_{a\in A}\lambda_a a$ eine unendliche Summe sein kann und deren Wert somit nicht definiert ist.

Definition 2.6. Eine Familie von Vektoren in V ist eine Abbildung $I \to V$, $i \mapsto v_i$ die wir in der Form $(v_i)_{i\in I}$ oder $(v_i|i \in I)$ schreiben. Man kann n-Tupeln, d.h. Elemente von

$$\underbrace{V \times \ldots \times V}_{n\text{-mal}}$$

als Familien über der Indexmenge $I = \{1, \ldots, n\}$ formalisieren.

Die Familie $(v_i|i \in I)$ spielt eine ähnliche Rolle wie die Menge $\{v_i|i \in I\} = \bigcup(\{v_i\}|i \in I)$. Aber es gilt $\{v, v\} = \{v\}$ und $\{v, w\} = \{w, v\}$, wohingegen $(v, v) \neq (v)$ und $(v, w) \neq (w, v)$ (falls $v \neq w$).

Analog definieren wir eine Familie $(\lambda_i|i \in I)$ von Skalaren. Eine solche Familie $(\lambda_i|i \in I)$ heißt *quasi-endlich*, wenn die Menge

$$\{i \in I|\lambda_i \neq 0\}$$

endlich ist.

Definition 2.7. Sei $A \subseteq V$. Der Vektor $v \in V$ ist eine *Linearkombination von Vektoren aus* A, falls es eine quasi-endliche Familie $(\lambda_a|a \in A)$ gibt mit

$$v = \sum_{a\in A}\lambda_a a.$$

Sei $(v_i|i \in I)$ eine Familie von Vektoren. Der Vektor $v \in V$ ist eine *Linearkombination von* $(v_i|i \in I)$, falls es eine quasi-endliche Familie von Skalaren $(\lambda_i|i \in I)$ gibt mit

$$v = \sum_{i\in I}\lambda_i v_i.$$

Die λ_i nennt man *Koeffizienten*.

Lemma 2.8. *Sei V Vektorraum und $A \subseteq V$. Dann ist $v \in \operatorname{span} A$ genau dann, wenn es eine quasi-endliche Familie von Skalaren $(\lambda_a|a \in A)$ gibt mit*

$$v = \sum_{a\in A}\lambda_a a.$$

Der Beweis geht wie der Beweis des letzten Lemmas.

Für eine Familie von Vektoren $(v_i|i \in I)$ definieren wir den *von $(v_i|i \in I)$ erzeugten bzw. aufgespannten Unterraum* als

$$\langle v_i|i \in I\rangle = \operatorname{span}\{v_i|i \in I\}.$$

$(v_i|i \in I)$ nennt man *eine V erzeugende Familie*, falls $\langle v_i|i \in I\rangle = V$. Eine V erzeugende Familie nennt man *minimal*, falls

$$\forall J \subsetneq I : \langle v_i|i \in J\rangle \subsetneq V.$$

3. Lineare Unabhängigkeit

Eine Familie von Skalaren $(\lambda_i | i \in I)$ nennen wir trivial, falls $\lambda_i = 0$ für alle $i \in I$.

DEFINITION 3.1. Eine Familie von Vektoren $(v_i | i \in I)$ heißt *linear abhängig (über* $\mathbb{K}$*)*, falls es nicht-triviale quasi-endliche Familien von Skalaren $(\lambda_i | i \in I)$ gibt mit

$$\sum_{i \in I} \lambda_i v_i = 0.$$

Falls die Familie nicht linear abhängig über $\mathbb{K}$ ist, so nennen wir sie *linear unabhängig über* $\mathbb{K}$.

Beispiele:

(1) Ist $I = \{1, 2, \dots, n\}$, so ist $(v_1, \dots, v_n)$ linear abhängig genau dann, wenn es $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ gibt mit $\sum_{i=1}^n \lambda_i v_i = 0$.

(2) $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^2$, $I = \{1, 2\}$

$$v_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad v_2 := \begin{pmatrix} \pi \\ 2\pi \end{pmatrix}.$$

Das Paar (v_1, v_2) ist linear abhängig über $\mathbb{R}$, denn

$$\pi v_1 + (-1) v_2 = 0.$$

Wenn wir aber $V = \mathbb{R}^2$ als $\mathbb{Q}$-Vektorraum anschauen, so (v_1, v_2) über $\mathbb{Q}$ linear unabhängig. Um dies zu zeigen, nehmen wir an

$$\lambda_1 v_1 + \lambda_2 v_2 = 0$$

für $\lambda_1, \lambda_2 \in \mathbb{Q}$. Dann folgt insbesondere $\lambda_1 \cdot 1 + \lambda_2 \pi = 0$. Falls $\lambda_2 \neq 0$, so ergibt sich $\pi = -\frac{\lambda_1}{\lambda_2} \in \mathbb{Q}$ und somit ein Widerspruch. Also ist $\lambda_2 = 0$ und somit $\lambda_1 = 0$.

(3) Wir führen das *Kronecker-Symbol* δ_{ij} ein:

$$\delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j\,. \end{cases}$$

Wir definieren $e_i := (\delta_{ij})_{j \in \{1,2\dots,n\}} \in \mathbb{K}^n$. Also

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Dann ist $(e_1, \dots, e_n)$ linear unabhängig. Wir nennen diese Familie die *kanonische „Basis“ von* $\mathbb{K}^n$.

(4) Die drei Vektoren $e_1, e_2, e_1 + e_2 \in \mathbb{K}^2$ sind linear abhängig, aber paarweise linear unabhängig.

(5) Sei $(v_i| \in I)$ eine Familie von Vektoren, sei $j \in I$ und sei $v_j = 0$. Dann ist $(v_i| \in I)$ linear abhängig. Sei $\lambda_i := \delta_{ij}$, wobei δ_{ij} wieder das Kronecker-Symbol ist. Dann ist $(\lambda_i|i \in I)$ nichttrivial und quasi-endlich, und wir haben

$$\sum_{i\in I} \lambda_i v_i = 0.$$

(6) Sei (v, w) linear unabhängig. Die Familie (v, v, w) ist linear abhängig, denn $-1v + 1v + 0w = 0$. Da andrerseits $\{v, w\} = \{v, v, w\}$ gilt, ist es wichtig, die Familie $(v_i|i \in I)$ von der Menge $\{v_i|i \in I\}$ zu unterscheiden.

Ist $(v_i|i \in I)$ linear unabhängig und $J \subseteq I$, dann ist auch $(v_i|i \in J)$ linear unabhängig.

Proposition 3.2. *Sei $(v_i|i \in I)$ eine Familie von Vektoren in V und $v \in \langle v_i|i \in I\rangle$. Dann sind die Koeffizienten in der Linearkombination*

$$v = \sum_{i\in I} \lambda_i v_i$$

genau dann eindeutig, wenn $(v_i|i \in I)$ linear unabhängig ist.

Beweis. Wegen $v \in \langle v_i|i \in I\rangle$ gibt es eine quasi-endliche Familie von Skalaren $(\lambda_i|i \in I)$ mit

$$v = \sum_{i\in I} \lambda_i v_i.$$

Angenommen $(v_i|i \in I)$ sei linear abhängig, d. h. es existiere eine nichttriviale quasi-endliche Familie von Skalaren $(\mu_i|i \in I)$ mit $\sum_{i\in I} \mu_i v_i = 0$. Also gilt auch

$$\sum_{i\in I} (\lambda_i + \mu_i) v_i = \sum_{i\in I} \lambda_i v_i + \sum_{i\in I} \mu_i v_i = v + 0 = v.$$

Die Familie $(\lambda_i + \mu_i|i \in I)$ ist ebenfalls quasi-endlich. Wir haben also eine Darstellung von v als Linearkombination von $(v_i|i \in I)$ mit anderen Koeffizienten gefunden.

Angenommen

$$v = \sum_{i\in I} \kappa_i v_i$$

für quasi-endliches $(\kappa_i|i \in I) \neq (\lambda_i|i \in I)$. Dann folgt

$$\sum_{i\in I} (\lambda_i - \kappa_i) v_i = \sum_{i\in I} \lambda_i v_i - \sum_{i\in I} \kappa_i v_i = v - v = 0.$$

Da $(\lambda_i - \kappa_i|i \in I)$ eine nichttriviale quasi-endliche Familie von Skalaren ist, ist $(v_i|i \in I)$ linear abhängig. $\square$

PROPOSITION 3.3. *Sei $(v_i|i \in I)$ eine Familie von Vektoren in V. Sie ist linear abhängig genau dann, wenn es ein $i_0 \in I$ gibt, so dass v_{i_0} eine Linearkombination von $(v_i|i \in I \setminus \{i_0\})$ ist.*

Beweis. Sei $(v_i|i \in I)$ linear abhängig. Dann existiert eine nichttriviale quasi-endliche Familie $(\lambda_i|i \in I)$ mit $\sum_{i\in I} \lambda_i v_i = 0$. Wähle ein $i_0 \in I$ mit $\lambda_{i_0} \neq 0$. Dann gilt

$$v_{i_0} = \sum_{i\in I\setminus\{i_0\}} -\frac{\lambda_i}{\lambda_{i_0}} v_i.$$

Umgekehrt sei

$$v_{i_0} = \sum_{i\in I\setminus\{i_0\}} \mu_i v_i.$$

Wir setzen $\mu_{i_0} := -1$ und erhalten

$$0 = \sum_{i\in I} \mu_i v_i, \qquad (\mu_i|i \in I) \neq 0.$$

□

KOROLLAR 3.4. *Eine V erzeugende Familie ist genau dann minimal, wenn sie linear unabhängig ist.*

Beweis. Sei $(v_i|i \in I)$ eine V erzeugende Familie. Angenommen $(v_i|i \in I)$ sei linear abhängig. Dann ist ein v_{i_0} Linearkombination von $(v_i|i \in I \setminus \{i_0\})$. Somit erzeugt bereits $(v_i|i \in I \setminus \{i_0\})$ den Raum V und somit ist $(v_i|i \in I)$ nicht minimal.

Sei nun $(v_i|i \in I)$ nicht minimal. Dies bedeutet, dass es $J \subsetneq I$ gibt, so dass $\langle v_i|i \in J\rangle = V$. Sei also $i_0 \in I \setminus J$. Dann ist $v_{i_0} \in V = \langle v_i|i \in J\rangle \subseteq \langle v_i|i \in I \setminus \{i_0\}\rangle$. Der Vektor v_{i_0} ist also eine Linearkombination der anderen und somit ist $(v_i|i \in I)$ linear abhängig. □

DEFINITION 3.5. Eine maximale linear unabhängige Familie $(v_i|i \in I)$ von Vektoren in V ist eine linear unabhängige Familie, die nicht zu einer größeren linear unabhängigen Familie $(v_i|i \in J)$, $J \supsetneq I$ erweitert werden kann.

KOROLLAR 3.6. *Sei $(v_i|i \in I)$ eine linear unabhängige Familie. Sie erzeugt V genau dann, wenn sie eine maximale linear unabhängige Familie ist.*

Beweis. Sei $(v_i|i \in I)$ eine linear unabhängige Familie, die V erzeugt. Um zu zeigen, dass diese Familie maximal ist, betrachten wir eine Erweiterung der Familie zu $(v_i|i \in J)$, $J \supsetneq I$. O.B.d.A. $J = I \cup \{i_0\}$. Da V bereits von $(v_i|i \in I)$ erzeugt wird, ist v_{i_0} eine Linearkombination von $(v_i|i \in I)$. Nach der Proposition 3.3 ist deswegen $(v_i|i \in I \cup \{i_0\})$ linear abhängig. Es folgt, dass $(v_i|i \in I)$ maximal ist.

Wenn die Familie V nicht erzeugt, so gibt es ein $v \in V \setminus \langle v_i | i \in I \rangle$. Sei $i_0 \notin I$, $v_{i_0} := v$. Wir wollen zeigen, dass $(v_i | i \in I \cup \{i_0\})$ linear unabhängig ist. Wir nehmen dazu an, dass

$$0 = \sum_{i \in I \cup \{i_0\}} \lambda_i v_i$$

für eine quasi-endliche Familie $(\lambda_i | i \in I \cup \{i_0\})$. Falls $\lambda_{i_0} \neq 0$, so folgt

$$v_{i_0} = \sum_{i \in I} -\frac{\lambda_i}{\lambda_{i_0}} v_i \in \langle v_i | i \in I \rangle,$$

also ein Widerspruch. Wir schließen $\lambda_{i_0} = 0$ und somit

$$0 = \sum_{i \in J} \lambda_i v_i.$$

Da $(v_i | i \in I)$ linear unabhängig ist, folgt $\lambda_i = 0 \forall i$. Also ist die linear unabhängige Familie $(v_i | i \in I)$ nicht maximal. □

Wir haben somit gezeigt:

Satz 3.7. *Es sind äquivalent:*

(1) $(v_i | i \in I)$ *ist eine minimale V erzeugende Familie,*
(2) $(v_i | i \in I)$ *ist eine maximale linear unabhängige Familie von Vektoren in V,*
(3) $(v_i | i \in I)$ *ist eine V erzeugende und linear unabhängige Familie.*

Definition 3.8. Sei V ein Vektorraum. Eine *Basis von V* ist eine V erzeugende und linear unabhängige Familie von Vektoren in V.

Beispiel: $(e_1, \dots, e_n)$ ist eine Basis von $\mathbb{K}^n$.

Bemerkung: Die Familie $(v_i | i \in \emptyset)$ ist linear unabhängig: denn $\mathrm{Abb}(\emptyset, \mathbb{K})$ hat genau ein Element und somit gibt es genau eine Familie $(\lambda_i | i \in \emptyset)$, so dass $0 = \sum_{i \in \emptyset} \lambda_i v_i$.

$$\langle v_i | i \in \emptyset \rangle = \{0\}.$$

Also ist $(v_i | i \in \emptyset)$ Basis von $\{0\}$.

4. Minimale und maximale Elemente und das Lemma von Zorn

Besitzt jeder Vektorraum eine Basis?

Wiederholung: Eine Relation $\leq$ auf $\mathcal{M}$ heißt Ordnungsrelation, falls sie reflexiv, antisymmetrisch und transitiv ist. Das Paar $(\mathcal{M}, \leq)$ nennt man dann eine *(partiell) geordnete Menge*. Die Ordnungsrelation $\leq$ ist total, falls $\forall x, y \in \mathcal{M}$: $x \leq y$ oder $y \leq x$.

Beispiele geordneter Mengen:

(1) Sei M eine Menge, dann ist $(\mathcal{P}(M), \subseteq)$ eine geordnete Menge.
(2) Sei $\mathrm{Fam}(M)$ die Menge der Familien in M, d. h. die Elemente von $\mathrm{Fam}(M)$ sind Familien $(m_i|i \in I)$, $m_i \in M$, I beliebige Menge. Wir definieren für $(m_i|i \in I), (n_i|i \in I) \in \mathrm{Fam}(M)$:
$$(m_i|i \in I) \leq (n_i|i \in J) :\Leftrightarrow I \subseteq J \text{ und } \forall i \in I : \quad m_i = n_i.$$
Dann ist $(\mathrm{Fam}(M), \leq)$ eine geordnete Menge.
(3) Ist $(\mathcal{M}, \leq)$ eine geordnete Menge und $\mathcal{N} \subseteq \mathcal{M}$, dann definiert $\leq$ auch eine Ordnung auf $\mathcal{N}$.

Definition 4.1. Sei $\mathcal{M}$ eine Menge mit Ordnungsrelation $\leq$. Wir sagen $m \in \mathcal{M}$ ist

(1) ein *Maximum* von $(\mathcal{M}, \leq)$, falls $\forall x \in \mathcal{M}$: $x \leq m$,
(2) ein *Minimum* von $(\mathcal{M}, \leq)$, falls $\forall x \in \mathcal{M}$: $m \leq x$,
(3) ein *maximales Element* von $(\mathcal{M}, \leq)$, falls $\forall x \in \mathcal{M}$: $(m \leq x \Rightarrow x = m)$,
(4) ein *minimales Element* von $(\mathcal{M}, \leq)$, falls $\forall x \in \mathcal{M}$: $(x \leq m \Rightarrow x = m)$.

Falls ein Minimum $m \in \mathcal{M}$ existiert, so ist m auch minimales Element und es gibt keine weiteren minimalen Elemente (und somit keine weiteren Minima). In einer total geordneten Menge ist umgekehrt ein minimales Element ein Minimum. Analoges gilt für Maxima und maximale Elemente.

Beispiele:

(1) Die geordnete Menge $(\{\{1\}, \{2\}, \{1, 2\}\}, \subseteq)$ hat kein Minimum, aber zwei minimale Elemente $\{1\}$ und $\{2\}$.
(2) Sei V ein Vektorraum.
$$\mathcal{E}_V := \{(v_i|i \in I) \,|\, (v_i|i \in I) \text{ ist eine } V \text{ erzeugende Familie}\} \subseteq \mathrm{Fam}(V)$$
Ein minimales Element von $(\mathcal{E}_V, \leq)$ ist eine Basis von V, und jede Basis von V ist minimales Element von $(\mathcal{E}_V, \leq)$.
(3) Sei V ein Vektorraum.
$$\mathcal{U}_V := \{(v_i|i \in I) \,|\, (v_i|i \in I) \text{ ist eine linear unabhängige Familie von Vektoren in } V\} \subseteq \mathrm{Fam}(V)$$
Ein maximales Element von $(\mathcal{U}_V, \leq)$ ist eine Basis von V, und jede Basis von V is maximales Element von $(\mathcal{U}_V, \leq)$.

Hilfssatz 4.2. *Ist $\mathcal{M}$ endlich und nichtleer, so besitzt $(\mathcal{M}, \leq)$ maximale (und minimale) Elemente.*

Beweis. (Existenz maximaler Elemente) Induktion über $n = \#\mathcal{M}$: $n = 1$ ist klar. Die Aussage gelte nun für alle Mengen mit höchstens $n-1$ Elementen. Sei $\#\mathcal{M} = n$. Wähle $m \in \mathcal{M}$. Falls m kein maximales Element von $\mathcal{M}$ ist, so ist die Menge $\mathcal{M}_m := \{x \in \mathcal{M} \,|\, m \leq x, m \neq x\} \subsetneq M$ endlich und nichtleer. Nach Induktionsvoraussetzung gibt es ein maximales Element x von $\mathcal{M}_m$. Angenommen $x \leq y$ für ein $y \in \mathcal{M}$. Dann folgt $m \leq y$ und somit $y \in \mathcal{M}_m$, also $y = x$. Somit ist x auch maximales Element von $\mathcal{M}$. Der Hilfssatz folgt durch Induktion. $\square$

Satz 4.3. *Sei V ein Vektorraum, $n \in \mathbb{N}$ und $(v_i | i \in \{1, \ldots, n\})$ eine V erzeugende Familie. Dann besitzt V eine Basis.*

Beweis. Wir definieren

$$\mathcal{U}_V^{(v_i)} := \{(v_i | i \in J) | J \subseteq I \text{ und } (v_i | i \in J) \text{ ist linear unabhängig}\} \subseteq \mathrm{Fam}(V).$$

Wegen $(v_i | i \in \emptyset) \in \mathcal{U}_V^{(v_i)}$ ist $\mathcal{U}_V^{(v_i)} \neq \emptyset$. Die Menge $\mathcal{U}_V^{(v_i)}$ besitzt höchstens 2^n Elemente, ist also endlich. Der Hilfssatz ergibt, dass ein maximales Element $(v_i | i \in J)$, $J \subseteq \{1, \ldots, n\}$ in $(\mathcal{U}_V^{(vi)}, \leq)$ existiert.

Um zu zeigen, dass die linear unabhängige Familie $(v_i | i \in J \cup \{i_0\})$ den Vektorraum V erzeugt, kann man nahezu wörtlich den Beweis von Korollar 3.6 übernehmen. Die Familie ist also eine Basis. $\square$

Ein Vektorraum heißt *endlich-dimensional*, falls er von einer endlichen Familie erzeugt wird. Jeder endlich-dimensionale Vektorraum besitzt also eine Basis. Viele Vektorräume sind nicht endlich-dimensional.

Beispiel (Quantenmechanik): Ein ruhendes Elektron wird durch einen Vektor im $\mathbb{C}$-Vektorraum $\mathrm{Abb}(\mathbb{R}^3, \mathbb{C}^2)$ beschrieben.

Sei nun $\mathcal{M}$ eine unendliche Menge. Dann ist unklar, ob minimale und maximale Elemente existieren, siehe zum Beispiel $\mathcal{M} = \mathbb{Z}$ mit der üblichen Ordnung.

Definition 4.4. Sei $Q \subseteq \mathcal{M}$. Man sagt $m \in \mathcal{M}$ ist *obere (untere) Schranke* von Q, falls $\forall q \in Q$: $q \leq m$ $(m \leq q)$. Wir sagen Q ist eine *total geordnete Teilmenge* oder *Kette*, falls die Einschränkung von $\leq$ auf Q total ist, d.h. falls $\forall x, y \in Q$: $x \leq y$ oder $y \leq x$.

Lemma 4.5 (von Zorn)**.** *Sei $(\mathcal{M}, \leq)$ eine geordnete Menge, derart, dass jede total geordnete Teilmenge $Q \subseteq \mathcal{M}$ eine obere Schranke in $\mathcal{M}$ besitzt. Dann enthält $\mathcal{M}$ ein maximales Element.*

Der Beweis ist aufwändig und nutzt das Auswahlaxiom der Mengenlehre (siehe [**5**, Kapitel 15 und 16] oder [**10**, Kapitel 9 und 10]). Das Auswahlaxiom besagt:

Auswahlaxiom:
Sei $(M_i)_{i \in I}$ eine Familie nichtleerer Mengen, $I \neq \emptyset$. Dann ist das kartesische Produkt aller M_i, notiert $X_{i \in I} M_i$ ebenfalls nichtleer.

Hierbei ist das kartesische Produkt definiert als

$$X_{i \in I} M_i := \{f \in \mathrm{Abb}(I, \bigcup_{i \in I} M_i) \,|\, f(i) \in M_i\}.$$

Beispiele: $I = \{1, 2, \ldots, n\}$, dann ist $X_{i \in I} M_i = M_1 \times M_2 \times \ldots M_n$.
Falls $M_i = M_j \forall i, j \in I$, dann ist $X_{i \in I} M_i = \mathrm{Abb}(I, M_i)$.

Unter der Annahme aller anderen Axiome der Mengenlehre ist das Auswahlaxiom äquivalent zum Lemma von Zorn. Man könnte also das Lemma von Zorn auch als Axiom der Mengenlehre betrachten. Manche Mathematiker machen Mathematik ohne Auswahlaxiom.

Satz 4.6 (Basis-Ergänzungssatz)**.** *Sei V ein Vektorraum, $(v_i|i \in I)$ eine linear unabhängige Familie und $(v_i|i \in J)$, $J \supseteq I$, eine V erzeugende Familie. Dann gibt es eine Indexmenge K mit $I \subseteq K \subseteq J$, so dass $(v_i|i \in K)$ eine Basis ist.*

Beweis. Sei

$$\mathcal{U} := \{(v_i|i \in L) \,|\, I \subseteq L \subseteq J, \quad (v_i|i \in L) \text{ linear unabhängig}\}.$$

Wir wollen zeigen, dass $(\mathcal{U}, \leq)$ mindestens ein maximales Element $(v_i|i \in K)$ enthält. Analog zu oben sieht man, dass dies dann die gesuchte Basis ist.

Vereinfachung: Alle Element in $u \in \mathcal{U}$ sind von der Form $(v_i|i \in I_u)$. Die v_i sind immer dieselben, nur I_u hängt von u ab. Wir müssen also maximale Elemente von

$$U := \{L \subseteq J \,|\, I \subseteq L, \quad (v_i|i \in L) \text{ linear unabhängig}\}$$

bezüglich der Relation $\subseteq$ bestimmen.

Um die Existenz eines maximalen Elements zu zeigen, wenden wir das Lemma von Zorn an.

Sei Q eine total geordnete Teilmenge von U. Wir definieren

$$\tilde{I} := \bigcup Q = \bigcup_{A \in Q} A \subseteq J.$$

Behauptung: $(v_i|i \in \tilde{I})$ ist linear unabhängig.

Um die Behauptung zu zeigen, nehmen wir an, dass

$$0 = \sum_{i \in \tilde{I}} \lambda_i v_i$$

für eine quasi-endliche Familie $(\lambda_i|i \in \tilde{I})$. Sei $\hat{I} := \{i \in \tilde{I} \,|\, \lambda_i \neq 0\}$, also

$$0 = \sum_{i \in \hat{I}} \lambda_i v_i$$

Zu jedem $i \in \hat{I}$ wählen wir ein $I_i \in Q$ mit $i \in I_i$. Die endliche Menge $\{I_i|i \in \hat{I}\}$ hat ein bezüglich $\subseteq$ maximales Element, das wir $\bar{I}$ nennen. Es ist ein Maximum dieser Menge, da diese Menge Teilmenge der total geordneten Menge Q ist. Es folgt $\hat{I} \subseteq \bar{I}$. Da die Familie $(v_i|i \in \bar{I})$ linear unabhängig ist, folgt $\lambda_i = 0$ für alle i, und die Behauptung ist gezeigt.

Die Familie $(v_i|i \in \tilde{I})$ ist somit eine obere Schranke von Q in U.

Das Lemma von Zorn kann angewendet werden, und die Existenz eines maximalen Elements folgt. □

KOROLLAR 4.7. *Jeder Vektorraum besitzt eine Basis.*

Beweis. Sei V ein Vektorraum. Wir setzen $J := V$, $v_i := i$ für alle $i \in J$. Dann ist $V = \langle v_i | i \in J\rangle$. Außerdem ist $\langle v_i | i \in \emptyset\rangle$ linear unabhängig. Die Voraussetzungen des Satzes sind also mit $I := \emptyset$ erfüllt. Und die Existenz einer Basis folgt. □

BEISPIELE. (1) Ist die Familie $(v_1, \ldots, v_r)$ linear unabhängig in $\mathbb{K}^n$, so kann sie zu einer Basis ergänzt werden.
(2) Erzeugt die Familie $(v_1, \ldots, v_r)$ den Vektorraum $\mathbb{K}^n$, so erhält man durch geschicktes Weglassen von Vektoren eine Basis.

Obwohl wir wissen, dass Basen immer existieren. Dies bedeutet aber nicht, dass man in konkreten Fällen eine Basis explizit angeben kann.

Beispiel: Der $\mathbb{R}$-Vektorraum $\mathrm{Abb}(\mathbb{N}, \mathbb{R})$ besitzt eine Basis. Man kann aber keine Basis explizit angeben.

Wichtiger als die Existenz einer Basis ist der Aufbau des Beweises. Ähnliche Schlüsse kommen oft an entscheidenden Stellen der Mathematik vor.

5. Koordinaten in einem Vektorraum

Für eine beliebige Menge I bezeichnen wir mit $\mathrm{Abb}_e(I, \mathbb{K})$ die Menge der quasi-endlichen Familien von Skalaren mit Indexmenge I, d.h.

$$\mathrm{Abb}_e(I, \mathbb{K}) := \left\{ (v_i | i \in I) \in \mathrm{Abb}(I, \mathbb{K}) \;\middle|\; \#\{\lambda_i | \lambda_i \neq 0\} < \infty \right\}.$$

Man sieht leicht, dass $\mathrm{Abb}_e(I, \mathbb{K})$ ein Untervektorraum von $\mathrm{Abb}(I, \mathbb{K})$ ist. Falls $\#I < \infty$, dann $\mathrm{Abb}_e(I, \mathbb{K}) = \mathrm{Abb}(I, \mathbb{K})$.

Sei $(v_i | i \in I) = (v_i)$ eine Familie von Vektoren in V. Dann definieren wir die *Linearkombinationsabbildung zur Familie* (v_i) als

$$\begin{aligned} \mathcal{V}_{(v_i)} : \mathrm{Abb}_e(I, \mathbb{K}) &\to V \\ (\lambda_i) &\mapsto \sum_{i \in I} \lambda_i v_i \end{aligned}$$

Offensichtlich gilt für $(\lambda_i), (\mu_i) \in \mathrm{Abb}_e(I, \mathbb{K})$ und $\alpha \in \mathbb{K}$

$$\mathcal{V}_{(v_i)}((\lambda_i) + (\mu_i)) = \mathcal{V}_{(v_i)}((\lambda_i + \mu_i)) = \mathcal{V}_{(v_i)}((\lambda_i)) + \mathcal{V}_{(v_i)}((\mu_i)),$$

$$\mathcal{V}_{(v_i)}(\alpha(\lambda_i)) = \mathcal{V}_{(v_i)}((\alpha\lambda_i)) = \alpha\mathcal{V}_{(v_i)}((\lambda_i)),$$

d.h. $\mathcal{V}_{(v_i)}$ ist linear. Wir haben gesehen:

$$\mathrm{Bild}(\mathcal{V}_{(v_i)}) = \{\text{Linearkombinationen von } (v_i)\} = \langle v_i | i \in I\rangle.$$

Die Abbildung $\mathcal{V}_{(v_i)}$ ist also surjektiv, gdw (v_i) den Raum V erzeugt.

Proposition 3.2 besagt, dass die Abbildung $\mathcal{V}_{(v_i)}$ injektiv ist, gdw (v_i) linear unabhängig ist.

Also ist $\mathcal{V}_{(v_i)}$ bijektiv gdw $(v_i | i \in I)$ eine Basis.

Jeder Vektorraum V besitzt eine Basis $(v_i | i \in I)$. Es gibt somit eine lineare, bijektive Abbildung $\mathrm{Abb}_e(I, \mathbb{K}) \to V$. Wir nennen $(\lambda_i) := (\mathcal{V}_{(v_i)})^{-1}(v)$ die Koordinaten des Vektors $v \in V$ bezüglich der Basis (v_i). Für $i \in I$ nennen wir dann λ_i die i-te Koordinate.

Beispiel: λ_i ist die i-te Koordinate von $\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$ bezüglich der Basis $(e_1, \ldots, e_n)$, denn

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \sum_{i \in I} \lambda_i e_i.$$

$\mathcal{V}_{(e_i)} : \mathbb{K}^n \to \mathbb{K}^n$ ist die Identität.

6. Dimension

Wir wollen zunächst einmal eine in den Saalübungen gemachte Beobachtung wiederholen und begründen.

Hilfssatz 6.1. *Sei $Ax = y$ ein lineares Gleichungssystem, $A \in \mathrm{Mat}(n, m; \mathbb{K})$ und $y \in \mathbb{K}^n$ gegeben, $x \in \mathbb{K}^m$ gesucht. Die folgenden Zeilenumformungen des Gaußschen Verfahrens sind Äquivalenzumformungen, das heißt die Menge der Lösungen bleibt erhalten.*

(1) Man multipliziert eine Zeile mit einer Zahl $\lambda \in \mathbb{K} \setminus \{0\}$
(2) Sei $\lambda \in \mathbb{K}$, $j, k \in \{1, \ldots, n\}$, $j \neq k$. Man addiert das λ-fache der j-ten Zeile zu der k-ten Zeile. Hierbei ersetzt man die k-te Zeile durch diese neue Zeile.
(3) Man vertauscht zwei Zeilen.
(4) Man streicht eine Zeile, die nur aus Nullen besteht.

Beweis. Zu zeigen ist, dass (a) jede Lösung des Systems durch die Umformungen (1)-(4) erhalten bleibt und (b) keine neuen entstehen. Aussage (a) ist offensichtlich.

Aussage (b):
Umformung (4): Eine Zeile, die nur aus Nullen besteht ist immer wahr. Deswegen kann man sie weglassen, ohne die Lösungsmenge zu vergrößern.
Umformungen (1)-(3): Wir geben zu jeder Umformung U_1 eine andere Umformung U_2 vom Typ

(1)-(4) an, die sie rückgängig macht. Da diese Rück-Umformung U_2 ebenfalls alle Lösungen erhält, erzeugt die Umformung $U-1$ keine Lösungen.

(1) Wenn U_1 die Multiplikation einer Zeile mit $\lambda \in \mathbb{K} \setminus \{0\}$ ist, so ist U_2 die Multiplikation mit λ^{-1}.
(2) Ist U_1 die Addition des λ-fachen der j-ten Zeile zur k-ten Zeile, so ist U_2 die Addition des $-\lambda$-fachen der j-ten Zeile zur k-ten Zeile. (3) Wenn U_1 vom Typ (3) ist, wählen wir $U_2 = U_1$. □

Mit dem Gaußschen Verfahren erreicht man nach endlich vielen Zeilenumformungen Zeilenstufenform, aus der man rekursiv die Menge aller Lösungen ablesen kann.

Hilfssatz 6.2. *Sei $r, k \in \mathbb{N}$. Sei $(v_1, v_2, \ldots, v_r)$ ein linear unabhängiges r-Tupel von Vektoren in $\mathbb{K}^k$. Dann gilt $r \leq k$.*

Beweis. Angenommen es gibt ein r-Tupel $(v_1, v_2, \ldots, v_r)$ von Vektoren in $\mathbb{K}^k$ mit $r > k$. Wir wollen zeigen, dass dieses Tupel linear abhängig ist. Wir suchen ein n-Tupel von Skalaren $(\lambda_1, \lambda_2, \ldots, \lambda_r) \neq 0$, so dass $\sum_{i=1}^r \lambda_i v_i = 0$ oder äquivalent dazu

$$(v_1 \quad v_2 \quad \cdots \quad v_r) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_r \end{pmatrix} = 0 \in \mathbb{K}^k,$$

wobei $(v_1 \quad v_2 \quad \cdots \quad v_r)$ als $k \times r$-Matrix zu verstehen ist. Wir lösen dieses homogene lineare Gleichungssystem mit dem Gaußschen Verfahren. Nach endlich vielen Zeilenumformungen haben wir ein Matrix in Zeilenstufenform mit k Zeilen und $r > k$ Spalten. Wenn wir also in jeder Zeile einen Pivot markieren, gibt es mindestens eine Spalten ohne Pivot, sagen wir die j-Spalte. Wir können dann $\lambda_j = 1$ zu einer Lösung $(\lambda_1, \ldots, \lambda_r)$ des Systems ergänzen, die offensichtlich nichttrivial ist. Das Tupel von Vektoren $(v_1, v_2, \ldots, v_r)$ ist somit linear abhängig. □

Lemma 6.3. *Sei V ein Vektorraum, $(w_1, \ldots, w_k)$ eine Familie von Vektoren in V. Ist $(v_1, \ldots, v_r)$ eine linear unabhängige Familie von Vektoren in $\langle w_1, \ldots, w_k \rangle$, dann ist $r \leq k$.*

Beweis. $U := \langle w_1, \ldots, w_k \rangle$. Der Basisergänzungssatz besagt, dass es eine Teilmenge $I \subseteq \{1, 2, \ldots, k\}$ gibt, so dass $(w_i | i \in I)$ eine Basis von U ist. Die Abbildung $\mathcal{V}_{(w_i|i\in I)} : \mathbb{K}^{\#I} \to U$ ist also bijektiv und linear.

Sei nun $(v_1, \ldots, v_r)$ eine Familie von Vektoren in U mit $r > k \geq \#I$. Zu zeigen ist, dass (v_i) linear abhängig. Wir setzen $x_j := \mathcal{V}^{-1}_{(w_i|i\in I)}(v_j) \in \mathbb{K}^{\#I}$. Nach dem Hilfssatz ist das Tupel von Vektoren linear abhängig, d.h. es existiert $(\lambda_1, \ldots, \lambda_r) \neq 0$, s.d. $\sum_{j=1}^r \lambda_i x_i = 0$. Es folgt

$$\mathcal{V}^{-1}_{(w_i|i\in I)} \left(\sum_{i=1}^r \lambda_i v_i \right)) = \sum_{i=1}^r \lambda_i \mathcal{V}^{-1}_{(w_i|i\in I)}(v_i) = 0$$

und somit $\sum_{i=1}^r \lambda_i v_i = 0$. Somit ist $(v_1, \ldots, v_r)$ linear abhängig. □

Beweis (Diagrammatische Version). Sei wie oben $(w_i|i \in I)$ eine Basis von $U = \langle w_i|i \in \{1, \ldots, k\}\rangle$, und $x_j := \mathcal{V}^{-1}_{(w_i|i\in I)}(v_j)$. Man überlegt sich

$$\mathcal{V}_{(w_i|i\in I)} \circ \mathcal{V}_{(x_j|j\in\{1,\ldots,r\})} = \mathcal{V}_{(v_j|j\in\{1,\ldots,r\})}.$$

$$\begin{tikzcd}
 & U \subseteq V \\
\mathbb{K}^r \arrow[ur, "\mathcal{V}_{(v_i|i\in\{1,\ldots,r\})}"] \arrow[r, "\mathcal{V}_{(x_i|i\in\{1,\ldots,r\})}"'] & \mathbb{K}^{\#I} \arrow[u, "\mathcal{V}_{(w_i|i\in I)}"']
\end{tikzcd}$$

Die lineare Unabhängigkeit von (v_i) bewirkt, dass $\mathcal{V}_{(v_j|j\in\{1,\ldots,r\})}$ injektiv und da $\mathcal{V}_{(w_i|i\in I)}$ bijektiv, folgt $\mathcal{V}_{(x_j|j\in\{1,\ldots,r\})}$ injektiv, d.h. $(x_1, \ldots, x_r)$ linear unabhängig in $\mathbb{K}^{\#I}$, und somit $r \leq \#I \leq k$. $\square$

Satz 6.4. *Sei V ein Vektorraum. Seien $(v_i|i \in I)$ und $(w_i|i \in J)$ zwei Basen von V. Dann gibt es eine Bijektion $I \to J$.*

Beweis. Wir zeigen den Satz nur für den Fall, dass mindestens eine der Basen endlich ist. Sei o.B.d.A. J endlich. Wir zeigen zunächst, dass dann auch I endlich.

Hierzu nehmen wir an I sei unendlich. Wir nehmen eine Teilmenge $K \subseteq I$ mit $\#K > \#J$. Da $(v_i|i \in I)$ linear unabhängig ist, ist auch $(v_i|i \in K)$ eine linear unabhängige Familie von Vektoren in $V = \langle w_j|j \in J\rangle$. Das vorhergehende Lemma besagt $\#J \geq \#K > \#J$. Widerspruch. Also ist I endlich.

Wenn nun I und J endlich sind, so besagt das obige Lemma $\#I \leq \#J$ und wenn wir die Rolle der v_i und w_i vertauschen auch $\#J \leq \#I$. $\square$

Definition 6.5. Sei V ein Vektorraum mit Basis $(v_i|i \in I)$.
Falls $V = \{0\}$, so definieren wir $\dim_{\mathbb{K}} V := 0$.
Falls $I \neq \emptyset$ endlich, so sagen wir $\dim_{\mathbb{K}} V := \#I$.
Falls I unendlich, so sagen wir $\dim_{\mathbb{K}} V := \infty$ und V ist *unendlich-dimensional*. Wir sagen V ist *endlich-dimensional* gdw $\dim_{\mathbb{K}} V \in \mathbb{N}_0$. Falls aus dem Zusammenhang klar ist, welcher Körper $\mathbb{K}$ gemeint ist, schreiben wir auch $\dim V$ für $\dim_{\mathbb{K}} V$.

Beispiele: $\dim \mathbb{K}^n = n$, $\dim_{\mathbb{C}} \mathbb{C} = 1$, $\dim_{\mathbb{R}} \mathbb{C} = 2$, $\dim_{\mathbb{Q}} \mathbb{R} = \infty$, $\dim_{\mathbb{K}} \mathrm{Abb}(I, \mathbb{K}) = \#I$.

Bemerkung: Man könnte $\dim V := \infty$ noch verfeinern, indem man $\dim V$ als die Mächtigkeit der Basis definiert.

7. Direkte Summen von Untervektorräumen

Definition 7.1. Sei V ein Vektorraum und $(U_i|i \in I)$ eine Familie von Untervektorräumen. Die Summe $\sum_{i\in I} U_i$ der Untervektorräume $(U_i|i \in I)$ ist definiert als

$$\sum_{i\in I} U_i := \left\{ \sum_{i\in I} v_i \,\middle|\, (v_i|i \in I) \text{ ist quasi-endliche Familie von Vektoren in } V \text{ und } \forall i \in I : v_i \in U_i \right\}.$$

Die Summe $\sum_{i\in I} U_i$ ist wieder ein Untervektorraum, er enthält alle U_i und jeder Untervektorraum, der alle U_i enthält, enthält auch $\sum\limits_{i\in I} U_i$. Somit gilt

$$\sum_{i\in I} U_i = \operatorname{span} \bigcup_{i\in I} U_i.$$

Falls $I = \{1, 2, \dots, n\}$, so schreibt man auch $U_1 + U_2 + \cdots + U_n$ an Stelle von $\sum_{i=1}^n U_i$. Es gilt offensichtlich $U_1 + U_2 = U_2 + U_1$ und $U_1 + U_2 + U_3 = (U_1 + U_2) + U_3 = U_1 + (U_2 + U_3)$ und analog alle assoziativen und kommutativen Relationen mit mehr Untervektorräumen.

Beispiele. (1) $V = \mathbb{R}^3$, $U_1 := \operatorname{span} e_1$, $U_2 := \operatorname{span} e_2$, $U_3 := \operatorname{span}(e_1 + e_2)$, $U_4 := \operatorname{span} e_3$. Dann ist $U_1 + U_2 = \operatorname{span}\{e_1, e_2\}$ die durch $x_3 = 0$ beschriebene Ursprungsebene. $U_1 + U_2 + U_3 = \operatorname{span}\{e_1, e_2, e_1 + e_2\} = \operatorname{span}\{e_1, e_2\} = U_1 + U_2 = U_1 + U_3 = U_2 + U_3$. Und es gilt $U_1 + U_2 + U_3 + U_4 = U_1 + U_2 + U_4 = V$.

(2) $V = \mathbb{K}^n$, $n \geq 3$. $\tilde{U}_1 = \operatorname{span}\{e_1, e_2\}$, $\tilde{U}_2 = \operatorname{span}\{e_2, e_3\}$. Dann ist

$$\tilde{U}_1 + \tilde{U}_2 = \operatorname{span}\{e_1, e_2, e_3\}.$$

(3) $\sum_{i\in I} \operatorname{span} v_i = V \Leftrightarrow (v_i|i \in I)$ erzeugt V.

Satz 7.2. *Sei $U = \sum_{i\in I}$ eine Summe von Unterräumen, $u \in U$. Dann sind äquivalent:*

(1) *Ist $0 = \sum_{i\in I} u_i$ für eine quasi-endliche Familie $(u_i|i \in I)$ von Vektoren in V mit $u_i \in U_i \quad \forall i \in I$, so gilt $u_i = 0 \quad \forall i \in I$.*
(2) *Ist $v = \sum_{i\in I} u_i = \sum_{i\in I} v_i$ für quasi-endliche Familien $(u_i|i \in I)$ und $(v_i|i \in I)$ von Vektoren in V mit $u_i, v_i \in U_i \quad \forall i \in I$, so gilt $u_i = v_i \quad \forall i \in I$.*
(3) *Für alle $j \in I$ gilt*

$$U_j \cap \left(\sum_{i\in I\setminus\{j\}} U_i \right) = \{0\}.$$

Ist I endlich, so ist auch äquivalent hierzu:
(4) *Die Abbildung $X_{i\in I} U_i \to U$, $(u_i|i \in I) \mapsto \sum_{i\in I} u_i$ ist bijektiv.*

Falls eine (und somit alle) der Bedinungen erfüllt sind, so sagen wir, U ist die *direkte Summe von* $(U_i|i \in I)$. In diesem Fall schreiben wir auch

$$U = \bigoplus_{i\in I} U_i,$$

im Fall $I = \{1, \ldots, n\}$ auch $U_1 \oplus \cdots \oplus U_n$.

Beweis.
„(2)⇒(1)“: Wir setzen $v_i := 0$, so ergibt sich (1) aus (2).
„(1)⇒(2)“: Es gelte (1). Es seien zwei Zerlegungen $v = \sum_{i \in I} u_i$ und $v = \sum_{i \in I} v_i$ wie oben gegeben. Wir setzen $\tilde{u}_i := u_i - v_i$. Dann ist $(\tilde{u}_i | i \in I)$ eine quasi-endliche Familie und $\tilde{u}_i \in U_i$. Wegen (1) gilt also $\tilde{u} = 0 \forall i$, also $u_i = v_i \forall i$.
„(1)⇒(3)“: Es gelte (1). Fixiere ein $j \in I$. Sei $w \in V_j$ und $w \in \sum_{i \in I \setminus \{j\}} U_i$. Wir schreiben

$$w = \sum_{i \in I \setminus \{j\}} u_i.$$

für eine quasi-endliche Familie $(u_i | i \in I \setminus \{j\})$, $u_i \in U_i$. Wir setzen $u_j := -w$. Es gilt dann

$$\sum_{i \in I} u_i = 0, \quad (u_i | i \in I) \text{ quasi-endlich und } u_i \in U_i \forall i \in I.$$

Also nach (1): $u_i = 0 \forall i \in I$, insbesondere $w = 0$. Es folgt (3).
„(3)⇒(1)“: Wir beweisen durch Widerspruch. Es gelte (1) nicht. Dann gibt es eine nicht-triviale quasi-endliche Familie $(u_i | i \in I)$, mit $u_i \in U_i \forall i \in I$ und mit $\sum_{i \in I} u_i = 0$. Wähle $j \in I$ mit $u_j \neq 0$. Dann gilt $u_j = \sum_{i \in I \setminus \{j\}} -u_i \in \sum_{i \in I \setminus \{j\}} U_i$ und somit ist $u_j \neq 0$ ein Element von $U_j \cap \left(\sum_{i \in I \setminus \{j\}} U_i \right)$. Es gilt also auch nicht (3).
„(2)⇔(4)“: Die obige Abbildung $X_{i \in I} U_i \to U$ ist offensichtlich surjektiv. Eigenschaft (2) ist äquivalent zur Injektivität. □

Beispiele. (Nummerierung wie zuvor)

(1) $U_1 \oplus U_2 \oplus U_4 = \mathbb{R}^3 = U_1 \oplus U_3 \oplus U_4$. $U_1 \oplus U_2 = U_2 \oplus U_3$. Die Summe $U_1 + U_2 + U_3$ ist nicht direkt.
(2) Die Summe $\tilde{U}_1 + \tilde{U}_2$ ist nicht direkt.
(3) $\bigoplus_{i \in I} \operatorname{span} v_i = U \Leftrightarrow (v_i | i \in I)$ ist linear unabhängig und $U = \langle v_i | i \in I \rangle$.

Definition 7.3. Sei U ein Untervektorraum von V. Ein Vektorraum W heißt *Komplement von* U, falls

$$V = U \oplus W.$$

Übungsaufgabe 7.4. Jeder Untervektorraum hat ein Komplement.

Übungsaufgabe 7.5. Es gelte $U = U_1 + U_2$ für endlich-dimensionale Untervektorräume U_1 und U_2. Dann gilt $\dim U \leq \dim U_1 + \dim U_2$. Falls $U = U_1 \oplus U_2$, dann gilt sogar $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$.

8. Basiswechsel

Sei $(b_1, \dots, b_n) \in V^n = V \times \dots \times V$ eine endliche Familie von Vektoren in V.
Konvention: solch eine Familie betrachten wir ab sofort immer als einen Zeilenvektor, dessen Einträge Vektoren sind.

Für $A = (a_{ij}) \in Mat(n, m; \mathbb{K})$ definieren wir

$$(b_1 \quad b_2 \quad \dots \quad b_n) \cdot A := \left(\sum_{i=1}^{n} a_{i1} b_i \quad \sum_{i=1}^{n} a_{i2} b_i \quad \dots \quad \sum_{i=1}^{n} a_{im} b_i \right) \in V^m.$$

Wir erhalten also eine Abbildung

$$V^n \times \mathrm{Mat}(n, m; \mathbb{K}) \to V^m,$$

und für jedes $A \in \mathrm{Mat}(n, m; \mathbb{K})$ eine Abbildung $\mathcal{W}_A : V^n \to V^m$. Offensichtlich gilt $\mathcal{W}_{AB} = \mathcal{W}_B \circ \mathcal{W}_A$.

Sei nun $(b_1, \dots, b_n)$ eine Basis von V und $(v_1, \dots, v_m)$ eine Familie von Vektoren in V. Definiere $a_{ij} \in \mathbb{K}$ so, dass

$$v_j = \sum_{i=1}^{n} a_{ij} b_i.$$

(Diese a_{ij} existieren, da $(b_1, \dots, b_n)$ den Raum V erzeugt, und sie sind eindeutig, da $(b_1, \dots, b_n)$ linear unabhängig ist.) Dann gilt für die Matrix $A = (a_{ij})_{i \in \{1,\dots,n\} j \in \{1,\dots,m\}}$

$$(b_1 \quad \dots \quad b_n) \cdot A = (v_1 \quad \dots \quad v_m).$$

Definition 8.1. Sei $A \in \mathrm{Mat}(n, m; \mathbb{K})$. Ein Links- (bzw. Rechts-)Inverses ist eine Matrix $B \in \mathrm{Mat}(m, n; \mathbb{K})$ mit $BA = \mathbb{1}_m$ (bzw. $AB = \mathbb{1}_n$).

Besitzt A ein Linksinverses B_l und ein Rechts-inverses B_r, so gilt:

$$B_l = B_l \mathbb{1}_n = B_l A B_r = \mathbb{1}_m B_r = B_r.$$

Satz 8.2 (Basiswechsel). *Sei $(b_1, \dots, b_n)$ eine Basis und $(v_1 \quad \dots \quad v_m) = (b_1 \quad \dots \quad b_n) \cdot A$. Dann besitzt A genau dann ein Links- und ein Rechtsinverses, wenn $(v_1, \dots, v_m)$ eine Basis ist.*

Da alle Basen gleich viele Elemente haben, folgt: Falls $A \in \mathrm{Mat}(n, m; \mathbb{K})$ ein Links- und Rechtsinverses B besitzt, so gilt $n = m$. Die Matrix A ist also ein invertierbares Element des Rings $(\mathrm{Mat}(n, n; \mathbb{K})$.

Definition 8.3. Sei $\mathbb{K}$ ein Körper. Die Menge der invertieren Elemente in $\mathrm{Mat}(n, n; \mathbb{K})$ nennen wir die *allgemeine lineare Gruppe* $\mathrm{GL}(n, \mathbb{K})$.

Wiederholung: Die invertierbaren Elemente eines Rings, versehen mit der Multiplikation, bilden eine Gruppe, die sogenannte Einheitengruppe des Rings.

Für alle invertierbaren Matrizen gilt

$$\mathcal{W}_{A^{-1}} = (\mathcal{W}_A)^{-1}.$$

Koordinaten unter Basiswechsel. Seien $\mathcal{B} = (b_1 \quad \dots \quad b_n)$ und $\mathcal{V} = (v_1 \quad \dots \quad v_n) = (b_1 \quad \dots \quad b_n) \cdot A$ zwei Basen. Ein Vektor $v \in V$ habe bezüglich der Basis $\mathcal{V}$ die Koordinaten $(\lambda_1, \dots, \lambda_n)$. Dann gilt

$$v = (v_1 \quad \dots \quad v_n) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = (b_1 \quad \dots \quad b_n) A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = (b_1 \quad \dots \quad b_n) \begin{pmatrix} \sum_{j=1}^n a_{1j}\lambda_j \\ \vdots \\ \sum_{j=1}^n a_{nj}\lambda_j \end{pmatrix}.$$

Die i-te Koordinate von v bezüglich der Basis $\mathcal{B}$ ist also $\sum_{j=1}^n a_{ij}$. Will man $\mathcal{B}$-Koordinaten in $\mathcal{V}$-Koordinaten umrechnen, so muss man hierzu A invertieren.

Beispiel: Sei $V = \mathbb{R}^2$, $b_1 = e_1$, $b_2 = e_2$, $v_1 = e_1 + e_2$, $v_2 = 2e_1 - 2e_2$. Ein Vektor $v = \sqrt{\pi} e_1 + 17 e_2$ wollen wir in der Basis (v_1, v_2) darstellen, d.h. wir suchen $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $v = \lambda_1 v_1 + \lambda_2 v_2$.

Die Transformations-Matrix von (e_1, e_2) nach (v_1, v_2) ist $A = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}$, denn $(e_1 \quad e_2) A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A = A = (v_1 \quad v_2)$. Also

$$\begin{pmatrix} \sqrt{\pi} \\ 17 \end{pmatrix} = A \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Wir multiplizieren von links mit

$$A^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} \end{pmatrix}$$

und erhalten:

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = A^{-1} \begin{pmatrix} \sqrt{\pi} \\ 17 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{\pi}+17}{2} \\ \frac{\sqrt{\pi}-17}{4} \end{pmatrix}.$$

Zum Beweis von Satz 8.2.

Hilfssatz 8.4. *Sei $(b_1, \dots, b_n)$ eine Basis und $(v_1 \quad \dots \quad v_m) = (b_1 \quad \dots \quad b_n) \cdot A$.*
(1) Falls $A \in \mathrm{Mat}(n, m; \mathbb{K})$ ein Rechtsinverses hat, so erzeugt $(v_1 \quad \dots \quad v_m)$ den Raum V.
(2) Falls $A \in \mathrm{Mat}(n, m; \mathbb{K})$ ein Linkssinverses hat, so ist $(v_1 \quad \dots \quad v_m)$ linear unabhängig.

Beweis. (1) Sei $AC = \mathbb{1}_n$ für $C = (c_{ij})_{ij} \in \mathrm{Mat}(m, n; \mathbb{K})$. Dann gilt $(b_1 \quad \dots \quad b_n) = (b_1 \quad \dots \quad b_n) AC = (v_1 \quad \dots \quad v_m) C$, also $b_j = \sum_{i=1}^m c_{ij} v_i \in \langle v_1, \dots, v_m \rangle$. Somit

$$V = \langle b_1, \dots, b_n \rangle \subseteq \langle v_1, \dots, v_m \rangle \subseteq V.$$

Also erzeugt $(v_1, \dots, v_m)$ den Raum V.

(2) Wir nehmen nun an, $A = (a_{ij})_{ij}$ besitze ein Linksinverses $D \in \mathrm{Mat}(m, n; \mathbb{K})$. Wir wollen zeigen, dass $(v_1, \dots, v_m)$ linear unabhängig. Wir nehmen an, $\sum_{i=1}^m \lambda_i v_i = 0$. Es folgt

$$0 = \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} \begin{pmatrix} \sum_{j=1}^m a_{1j}\lambda_j \\ \vdots \\ \sum_{j=1}^m a_{nj}\lambda_j \end{pmatrix}.$$

Da $(b_1, \dots, b_n)$ linear unabhängig ist, folgt

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}\lambda_j \\ \vdots \\ \sum_{j=1}^m a_{nj} \end{pmatrix} = 0.$$

Anwendung von D ergibt

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = DA \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = D0 = 0.$$

Somit ist $(v_1, \dots, v_m)$ linear unabhängig. □

Beweis von Satz 8.2. Wenn $(b_1, \dots, b_n)$ und $\begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} A$ Basen sind, so gibt es eine Matrix $C \in \mathrm{Mat}(m, n; \mathbb{K})$ mit $(b_1, \dots, b_n) = \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} C$. Es gilt dann $(b_1, \dots, b_n)\mathbb{1}_n = (b_1, \dots, b_n)AC$, woraus $AC = \mathbb{1}_n$ folgt. Ebenso gilt dann $\begin{pmatrix} v_1 & \dots & v_m \end{pmatrix}\mathbb{1}_m = \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} CA$, also $CA = \mathbb{1}_m$. Die Matrix C ist also Links- und Rechtsinverses.

Sei nun C Links- und Rechtsinverses von A. Auf Grund des Hilfssatzes ist dann $\begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} A$ eine Basis von V. □

KAPITEL 5

Lineare Abbildungen

1. Definitionen und erste Eigenschaften

In diesem Kapitel sei $\mathbb{K}$ ein Körper, U, V, W, X, Y Vektorräume über $\mathbb{K}$, u, v, w, x, y, z Vektoren und α, β, γ, λ, μ, κ, η Skalare.

Definition 1.1. Eine Abbildung $f : V \to W$ zwischen zwei Vektorräumen V und W heißt *($\mathbb{K}$-)linear* falls

(a) $f(v + w) = f(v) + f(w)$ für alle $v, w \in V$ und
(b) $f(\alpha v) = \alpha f(v)$ für alle $v \in V$, $\alpha \in \mathbb{K}$.

Unter anderem gilt dann $f(0) = 0$ (setze $\alpha = 0$).

Definition 1.2. Eine lineare Abbildung nennt man auch *($\mathbb{K}$-)(Vektorraum-)Homomorphismus*. Einen surjektiven Homomorphismus nennt man *Epimorphismus*, einen injektiven Homomorphismus nennt man *Monomorphismus*, einen bijektiven Homomorphismus nennt man *Isomorphismus*. Einen Homomorphismus $V \to V$ nennt man *Endomorphismus von V* und einen Isomorphismus $V \to V$ einen *Automorphismus von V*.

Zwei Vektorräume V und W heißen *isomorph*, falls es einen Isomorphismus $V \to W$ gibt, und wir schreiben $V \cong W$.

Beispiele:

(1) Die Spiegelung an der x-Achse (Abbildung 1)
$$\mathbb{R}^2 \to \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y \end{pmatrix}$$
ist linear.
(2) Eine Drehung um den Ursprung (Abbildung 2)
$$\mathbb{R}^2 \to \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$
ist linear.
(3) $A \in \operatorname{Mat}(n, m; \mathbb{K})$. Dann ist $\mathcal{L}_A : \mathbb{K}^m \to \mathbb{K}^n$, $v \mapsto Av$ linear.

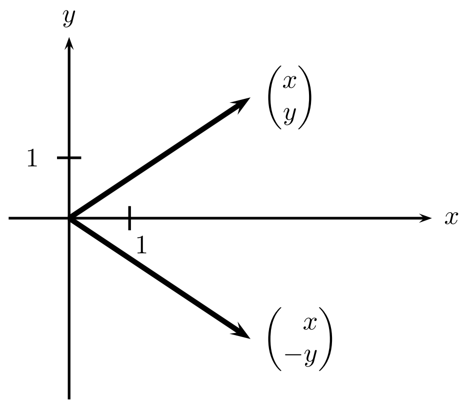


ABBILDUNG 1. Spiegelung an der x-Achse

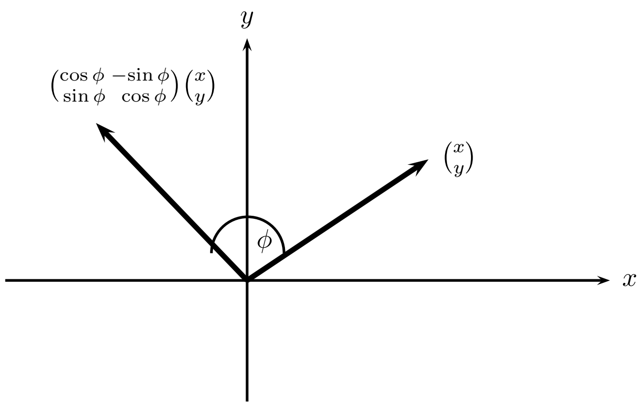


ABBILDUNG 2. Drehung um den Winkel ϕ

(4) Die Identität $\mathrm{Id}_V : V \to V$, $v \mapsto v$ ist linear.

(5) Sei $(v_1, \dots, v_n)$ eine Familie von Vektoren in V.

$$\mathcal{V}_{(v_i)} : \mathbb{K}^n \to V \qquad (\lambda_1, \dots, \lambda_n) \mapsto \sum_{i=1}^{n} \lambda_i v_i$$

Die Abbildung ist linear. Sie ist Monomorphismus, wenn (v_i) linear unabhängig, Epimorphismus, wenn (v_i) den Raum erzeugt V, und Isomorphismus, wenn (v_i) Basis ist. Jeder n-dimensionale $\mathbb{K}$-Vektorraum ist also isomorph zu $\mathbb{K}^n$.

(6) Die Konjugation $\mathbb{C} \to \mathbb{C}$, $z \mapsto \bar{z}$ ist $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.

(7) Sind V, W, X Vektorräume und sind $f : V \to W$ und $g : W \to X$ lineare Abbildungen, dann ist $g \circ f : V \to X$ ebenfalls eine lineare Abbildung.

Lemma 1.3. *Wenn $f : V \to W$ ein Isomorphismus ist, so ist die Umkehrabbildung $f^{-1} : W \to V$ auch ein Isomorphismus.*

Beweis. Zu zeigen ist nur, dass $f^{-1} : W \to V$ linear ist.
(1) Sei also $x, y \in W$.

$$f(f^{-1}(x+y)) = x + y = f(f^{-1}(x)) + f(f^{-1}(y)) = f(f^{-1}(x) + f^{-1}(y))$$

Da f injektiv ist, folgt

$$f^{-1}(x+y) = f^{-1}(x) + f^{-1}(y).$$

(2) ähnlich □

Isomorphie ist also eine Äquivalenzrelation auf der Klasse aller $\mathbb{K}$-Vektorräume.

Die Menge aller Automorphismen von V notieren wir $\mathrm{Aut}(V)$. Diese Menge, versehen mit der Komposition von Abbildungen, ist eine Gruppe.

Sei $f : V \to W$ eine lineare Abbildung. Die Menge

$$\text{Kern } f := \{v \in V \mid f(v) = 0\} = f^{-1}(\{0\})$$

ist ein Untervektorraum von V, denn:
(1) wenn $v, w \in \text{Kern } f$, dann $f(v) = f(w) = 0$ und somit

$$f(v + w) = f(v) + f(w) = 0 + 0 = 0.$$

(2) wenn $v \in \text{Kern } f$ und $\alpha \in \mathbb{K}$, dann $f(\alpha v) = \alpha f(v) = \alpha 0 = 0$.

Die Menge

$$\text{Bild} f := \{f(v) \mid v \in V\} = f(V)$$

ist ein Untervektorraum von W, denn:
(1) wenn $x, y \in \text{Bild} f$, dann gibt es $v, w \in V$ mit $f(v) = x$ und $f(w) = y$, also $x+y = f(v)+f(w) = f(v + w) \in \text{Bild} f$.
(2) wenn $x \in \text{Bild} f$, $\alpha \in \mathbb{K}$, dann gibt es $v \in V$ mit $f(v) = x$, also $\alpha f(v) = f(\alpha v) \in \text{Bild} f$.

Offensichtlich ist f ein Epimorphismus gdw $\text{Bild} f = W$.

Lemma 1.4. *Sei f linear. Dann ist f injektiv, gdw* $\ker f = \{0\}$.

Beweis. Falls f injektiv ist, dann hat 0 genau ein Urbild. $0 \in \text{Kern } f$. Also Kern $f = \{0\}$.

Sei nun Kern $f = \{0\}$. Falls $f(v) = f(w)$ für $v, w \in V$ gilt, dann haben wir auch $f(v - w) = f(v) - f(w) = 0$, also $v - w \in \text{Kern } f$. Es folgt $v = w$. Somit ist f injektiv. □

Übungsaufgabe 1.5. Sei $f : V \to W$ linear, $A \subseteq V$. Dann gilt

(1) Sei U ein Untervektorraum von V. Dann ist $f(U)$ Untervektorraum von W.
(2) Sei X ein Untervektorraum von W. Dann ist $f^{-1}(X)$ ein Untervektorraum von V.
(3) Für $A \subseteq V$ gilt $f(\operatorname{span} A) = \operatorname{span} f(A)$.
(4) Sei $(v_i | i \in I)$ eine linear abhängige Familie von Vektoren in V. Dann ist $(f(v_i)|i \in I)$ eine linear abhängige Familie von Vektoren in W.
(5) Sei $(v_i | i \in I)$ eine linear unabhängige Familie von Vektoren in V, und sei $\langle v_i | i \in I\rangle \cap \ker f = \{0\}$, dann ist auch $(f(v_i)|i \in I)$ linear unabhängig.

Sei also $f : V \to W$ ein Isomorphismus, und $(b_i | i \in I)$ eine Basis von V. Dann ist $(f(b_i)|i \in I)$ eine Basis von W. Also gilt für isomorphe $V \cong W$, dass $\dim V = \dim W$. Andrerseits, sei $(b_i | i \in I)$ eine Basis von V und $(v_i | i \in I)$ eine Basis von W. Dann sind die Abbildung $\mathcal{V}_{(b_i)} : \text{Abb}_e(I, \mathbb{K}) \to V$ und $\mathcal{V}_{(v_i)} : \text{Abb}_e(I, \mathbb{K} \to W$ Isomorphismen, also $V \cong \text{Abb}_e(I, \mathbb{K}) \cong W$. Also sind dann auch V und W isomorph. Insbesondere sind endlich dimensionale Vektorräume genau dann isomorph, wenn sie dieselbe Dimension haben.

Satz 1.6 (Rangformel)**.** *Sei $f : V \to W$ linear. Dann gilt*

$$\dim V = \dim \text{Kern } f + \dim \text{Bild} f.$$

Hierbei gilt die Konvention:

$$\infty + n := \infty \quad \forall n \in \mathbb{N}_0 \qquad \infty + \infty := \infty.$$

Beweis. Sei $(b_i | i \in I)$ eine Basis von $\ker f$. Wir ergänzen diese Familie zu einer Basis $(b_i | i \in J)$, $J \supseteq I$ von V. Wir wollen zeigen, dass $(f_i | i \in J \setminus I)$ eine Basis von Bild$f$ ist.

$$\langle f(b_i) | i \in J \setminus I\rangle = \langle f(b_i) | i \in J\rangle = \operatorname{span} f(\{b_i | i \in J\}) = f(\operatorname{span}\{b_i | i \in J\}) = f(V) = \text{Bild} f.$$

Außerdem ist $(b_i | i \in J \setminus I)$ linear unabhängig, und

$$\langle b_i | i \in J \setminus I\rangle \cap \text{Kern } f = \langle b_i | i \in J \setminus I\rangle \cap \langle b_i | i \in I\rangle = \{0\}.$$

Also ergibt sich aus der Übungsaufgabe die lineare Unabhängigkeit von $(f(b_i)|i \in J \setminus I)$. Ist $\dim \text{Kern } f$ und somit I unendlich, dann auch J und somit $\dim V$. Analog kann man aus der Unendlichkeit von $\dim \text{Bild} f$ die Unendlichkeit von $\dim V$ folgern. Wenn $\ker f$ und Bild$f$ endlich-dimensional sind, dann haben wir

$$\dim V = \#J = \#I + \#(J \setminus I) = \dim \text{Kern } f + \dim \text{Bild} f.$$

□

Insbesondere gilt immer $\dim f(V) \leq \dim V$. Man nennt $\mathrm{Rang}(f) := \dim f(V)$ den *Rang* von f.

BEISPIEL. (1) $v \in \mathbb{R}^3 \setminus \{0\}$, $f : U = \mathbb{R} \to V = \mathbb{R}^3$, $\alpha \mapsto \alpha v$. $\dim U = 1$, $\dim \mathrm{Kern}\ f = 0$, also $\dim \mathrm{Bild} f = 1$.

(2) $v, w \in \mathbb{R}^3 \setminus \{0\}$, $F : \mathbb{R}^2 \to \mathbb{R}^3$, $(\alpha, \beta) \mapsto \alpha v + \beta w$. $\dim \mathbb{R}^2 = 2$. Falls v und w linear unabhängig, dann ist Kern $f = \{0\}$ und somit $\dim \mathrm{Bild} f = 2$. Falls v und w linear abhängig sind, so ist $\dim \mathrm{Kern}\ f = 1$ und somit $\dim \mathrm{Bild} f = 1$.

(3) Sei $\dim V = \dim W = n \in \mathbb{N}$ und $f : V \to W$ linear. Dann ist f surjektiv gdw $\mathrm{Rang} f = n$ gdw $\dim \ker f = 0$ gdw f injektiv gdw f bijektiv.

PROPOSITION 1.7.

(a) *Seien $f, g : V \to W$ linear und $(v_i | i \in I)$ eine V erzeugende Familie. Es gelte $f(v_i) = g(v_i)$ für alle $i \in I$. Dann gilt bereits $f = g$.*

(b) *Sei $(v_i | i \in I)$ eine Familie von Vektoren in V und $(w_i | i \in I)$ eine Familie von Vektoren in W mit gleicher Indexmenge I. Falls $(v_i | i \in I)$ linear unabhängig ist, dann gibt es eine lineare Funktion $f : V \to W$ mit $f(v_i) = w_i$.*

(c) *Sei $(v_i | i \in I)$ eine Basis von V, und $(w_i | i \in I)$ eine beliebige Familie von Vektoren in W mit gleicher Indexmenge I. Dann gibt es eine eindeutige lineare Abbildung $f : V \to W$ mit*

$$f(v_i) = w_i.$$

Beweis.

„(a)“: Die Abbildung $f - g : V \to W$, $v \mapsto f(v) - g(v)$ ist linear, denn es gilt für $v, w \in V$:

$$(f-g)(v+w) = f(v+w) - g(v+w) = f(v) + f(w) - g(v) - g(w) = (f-g)(v) + (f-g)(w),$$

und analog $(f-g)(\alpha v) = \alpha (f-g)(v)$. Insbesondere ist $U := \mathrm{Kern}\ (f-g) = \{v \in V \mid f(v) = g(v)\}$ ein Untervektorraum. Wegen $f(v_i) = g(v_i)$ sind alle v_i in U enthalten und deswegen auch $\langle v_i | i \in I\rangle \subseteq U$. Da $(v_i | i \in I)$ den Raum V erzeugt, haben wir $U = V$ und somit $f = g$.

„(b)“: Wir ergänzen zunächst $(v_i | i \in I)$ zu einer Basis $(v_i | i \in J)$, $J \subseteq I$. Für $i \in J \setminus I$ setzen wir $w_i = 0$. Dann definieren wir

$$f(\sum_{i \in J} \lambda_i v_i) = \sum_{i \in J} \lambda_i w_i,$$

wobei $(\lambda_i | i \in J)$ eine quasi-endliche Familie von Skalaren ist. Um zu zeigen, dass f wohldefiniert ist, ist zu zeigen, dass diese Abbildung jedem $v \in V$ ein Element zuordnet und dies eindeutig ist. Das erste folgt aus $\langle v_i | i \in J\rangle = V$, das zweite (die Eindeutigkeit) aus der linearen Unabhängigkeit der $(v_i | i \in J)$. Man überprüft leicht, dass f linear ist und die gewünschte Eigenschaft hat. [1]

„(c)“: folgt direkt aus (a) und (b). □

Eine lineare Abbildung ist also durch die Bilder einer Basis eindeutig bestimmt.

[1] Kurzer Beweis von (b): $\mathcal{V}_{(v_i | i \in J)} : \mathrm{Abb}_e(J, \mathbb{K}) \to V$ ist ein Isomorphismus, und $\mathcal{V}_{(w_i | i \in J)} : \mathrm{Abb}_e(J, \mathbb{K}) \to W$ ein Homomorphismus, also $\mathcal{V}_{(w_i | i \in J)} \circ (\mathcal{V}_{(v_i | i \in J)})^{-1} : V \to W$ ein Homomorphismus. DIAGRAMM

2. Matrix einer linearen Abbildung, Basiswechsel

In diesem Abschnitt seien V, W, X endlich-dimensionale $\mathbb{K}$-Vektorräume.

Lemma 2.1. *Sei $f : \mathbb{K}^m \to \mathbb{K}^n$ eine lineare Abbildung. Dann gibt es eine Matrix $A \in \mathrm{Mat}(n, m; \mathbb{K})$, so dass $f = \mathcal{L}_A : \mathbb{K}^m \to \mathbb{K}^n$, $v \mapsto Av$.*

Beweis. Wir betrachten die Matrix

$$A := \begin{pmatrix} \mathcal{L}(e_1) & \mathcal{L}(e_2) & \cdots & \mathcal{L}(e_m) \end{pmatrix},$$

wobei die Notation so zu verstehen ist, dass wir von jedem Zeilenvektor $f(e_j)$ die Klammern weglassen und den Gesamtausdruck dann als $n \times m$-Matrix anschauen. Man sieht leicht $Ae_j = f(e_j)$. Somit gilt $f(e_j) = \mathcal{L}_A(e_j)$ und deswegen nach der letzten Proposition $f = \mathcal{L}_A$. $\square$

Sei nun $(v_1, \ldots, v_m)$ eine Basis von V und $(w_1, \ldots, w_n)$ eine Basis von W. Es gibt dann zu jeder Abbildung $f : V \to W$ eine eindeutig bestimmte Matrix $A \in \mathrm{Mat}(n, m; \mathbb{K})$, so dass

$$f \circ \mathcal{V}_{(v_1,\ldots,v_m)} = \mathcal{V}_{(w_1,\ldots,w_n)} \circ \mathcal{L}_A.$$

Die letzte Gleichung wird auch oft so beschrieben: das Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{K}^m \arrow[u, "\cong"', "\mathcal{V}_{(v_1,\ldots,v_m)}"] \arrow[r, "\mathcal{L}_A"'] & \mathbb{K}^n \arrow[u, "\cong"', "\mathcal{V}_{(w_1,\ldots,w_n)}"] \end{tikzcd}$$

kommutiert. Wir notieren $A = \mathrm{Mat}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(f)$. Man nennt A die *zur Abbildung f gehörige Matrix bezüglich der Basen $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$.*

Umgekehrt gibt es auch zu jeder Matrix $A \in \mathrm{Mat}(n, m; \mathbb{K})$ eine eindeutig bestimmte lineare Abbildung $f : V \to W$, so dass das obige Diagramm kommutiert. Wir notieren

$$f = \mathrm{Lin}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(A).$$

Man nennt f die *zur Matrix A gehörige lineare Abbildung bezüglich der Basen $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$*. Für $A = (a_{ij})_{ij}$ gilt dann

$$\mathrm{Lin}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(A)(v_i) = \sum_{k=1}^{n} a_{ki} w_k.$$

Sei nun außerdem X ein Vektorraum mit Basis $(x_1, \ldots, x_k)$ und $g : W \to X$ ebenfalls linear. Dann haben wir

$$\mathrm{Mat}^{(v_1,\ldots,v_m)}_{(x_1,\ldots,x_k)}(g \circ f) = \mathrm{Mat}^{(w_1,\ldots,w_n)}_{(x_1,\ldots,x_k)}(g)\mathrm{Mat}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(f)$$

oder anders ausgedrückt:

$$\mathrm{Lin}^{(v_1,\ldots,v_m)}_{(x_1,\ldots,x_k)}(BA) = \mathrm{Lin}^{(w_1,\ldots,w_n)}_{(x_1,\ldots,x_k)}(B)\mathrm{Lin}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(A)$$

für $A \in \mathrm{Mat}(n, m; \mathbb{K})$ und $B \in \mathrm{Mat}(k, n; \mathbb{K})$.

Die lineare Abbildung $f : V \to W$ ist genau dann ein Isomorphismus, wenn $A := \mathrm{Mat}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}$ ein Links-und Rechtsinverses hat, d.h. falls $n = m$ gilt und A invertierbar im Ring $\mathrm{Mat}(n, n; \mathbb{K})$ ist.

Spezialfall $V = W$*:* Es gilt dann für $n = \dim V$.

$$\mathrm{Mat}^{(v_1,\ldots,v_n)}_{(v_1,\ldots,v_n)}(\mathrm{Id}_V) = \mathbb{1}_n.$$

Die Matrix A ist genau dann invertierbar, wenn $\mathrm{Lin}^{(v_1,\ldots,v_n)}_{(v_1,\ldots,v_n)}(A)$ ein Isomorphismus ist.

Wir haben im letzten Abschnitt von Kapitel 4 gezeigt: $\mathrm{Mat}^{(v_1,\ldots,v_m)A}_{(w_1,\ldots,w_n)}(\mathrm{Id}) = A$ bzw. $\mathrm{Lin}^{(v_1,\ldots,v_m)A}_{(w_1,\ldots,w_n)}(A) = \mathrm{Id}$

Basiswechsel. Wir wollen studieren, wie sich die Matrix einer linearen Abbildung unter Basiswechsel transformiert. Wir lassen $V \neq W$ nun wieder zu. Seien $(v_1, \ldots, v_m)$ und $(v_1, \ldots, v_m)A$ Basen von V, $(w_1, \ldots, w_n)$ und $(w_1, \ldots, w_n)B$ Basen von W. $f : V \to W$ linear. Mit dem kommutativen Diagramm

$$\begin{tikzcd}
\mathbb{K}^m \arrow[rrr, "\mathcal{L}_{\mathrm{Mat}^{(v_i)}_{(w_i)}(f)}"] \arrow[dr, "\mathcal{V}_{(v_i)}"] & & & \mathbb{K}^n \arrow[dl, "\mathcal{V}_{(w_i)}"'] \\
& V \arrow[r, "f"] & W & \\
\mathbb{K}^m \arrow[uu, "\mathcal{L}_A"] \arrow[ur, "\mathcal{V}_{(v_i)\cdot A}"'] \arrow[rrr, "\mathcal{L}_{\mathrm{Mat}^{(v_i)\cdot A}_{(w_i)\cdot B}(f)}"'] & & & \mathbb{K}^n \arrow[uu, "\mathcal{L}_B"'] \arrow[ul, "\mathcal{V}_{(w_i)\cdot B}"]
\end{tikzcd}$$

sieht man sofort

$$\mathrm{Mat}^{(v_1,\ldots,v_m)A}_{(w_1,\ldots,w_n)B}(f) = B^{-1}\mathrm{Mat}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(f)A.$$

Einfache Matrix durch geschickte Basen-Wahl.

PROPOSITION 2.2. *Sei $V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Vektorräumen. Dann gibt es eine Basis $(v_1, \ldots, v_m)$ von V und $(w_1, \ldots, w_n)$ von W, so dass*

$$\mathrm{Mat}_{(w_1,\ldots,w_n)}^{(v_1,\ldots,v_m)}(f) = \begin{pmatrix} & & & 0 & 0 & \ldots & 0 \\ & \mathbb{1}_r & & 0 & 0 & \ldots & 0 \\ & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix} \in \mathrm{Mat}(n, m; \mathbb{K}),$$

und r ist der Rang von f.

Beweis. Wir wählen die Basen im Beweis der Rangformel. □

3. Homomorphismen als Vektorräume

Die Menge aller linearen Abbildungen (Homomorphismen) von V nach W bezeichnet man als $\mathrm{Hom}_{\mathbb{K}}(V, W)$ oder $\mathrm{Hom}(V, W)$. Im Fall $V = W$ handelt es sich um Endomorphismen und wir schreiben dann $\mathrm{End}_{\mathbb{K}}(V)$ oder $\mathrm{End}(V)$ für $\mathrm{Hom}_{\mathbb{K}}(V, V)$.

Die Menge $\mathrm{Hom}_{\mathbb{K}}(V, W)$, versehen mit der Addition

$$\begin{aligned} + : \mathrm{Hom}_{\mathbb{K}}(V, W) \times \mathrm{Hom}_{\mathbb{K}}(V, W) &\to \mathrm{Hom}_{\mathbb{K}}(V, W) \\ (f, g) &\mapsto \Big(v \mapsto (f + g)(v) := f(v) + g(v)\Big) \end{aligned}$$

und der Multiplikation mit Skalaren

$$\begin{aligned} \mathbb{K} \times \mathrm{Hom}_{\mathbb{K}}(V, W) &\to \mathrm{Hom}_{\mathbb{K}}(V, W) \\ (\alpha, f) &\mapsto \Big(v \mapsto (\alpha f)(v) := \alpha(f(v))\Big) \end{aligned}$$

ist ein $\mathbb{K}$-Vektorraum.

Beispiel: Sei $V = \mathbb{K}^3$ und $W = \mathbb{K}^2$. Jede lineare Abbildung $f : \mathbb{K}^3 \to \mathbb{K}^2$ ist von der Form $\mathcal{L}_A$ mit $A \in \mathrm{Mat}(2, 3; \mathbb{K})$. Der $\mathbb{K}$-Vektorraum $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^3, \mathbb{K}^2)$ ist 6-dimensional: $(\mathcal{L}_{A_1}, \mathcal{L}_{A_2}, \ldots, \mathcal{L}_{A_6})$ ist eine Basis, wobei

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A_5 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad A_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Abbildung

$$\begin{array}{rcl} \mathrm{Hom}(V,W) & \stackrel{\cong}{\to} & \mathrm{Mat}(n,m;\mathbb{K}) \\ f & \mapsto & \mathrm{Mat}^{(v_1,\dots,v_m)}_{(w_1,\dots,w_n)}(f) \end{array}$$

ist linear.

Spezialfall: $V = W$, $n = \dim V$. Die Abbildung

$$\begin{array}{rcl} \mathrm{End}(V,V) & \stackrel{\cong}{\to} & \mathrm{Mat}(n,n;\mathbb{K}) \\ f & \mapsto & \mathrm{Mat}^{(v_1,\dots,v_n)}_{(v_1,\dots,v_n)}(f) \end{array}$$

ist linear und erhält außerdem die Multiplikation in dem Sinne:

$$\mathrm{Mat}^{(v_1,\dots,v_n)}_{(v_1,\dots,v_n)}(f)\,\mathrm{Mat}^{(v_1,\dots,v_n)}_{(v_1,\dots,v_n)}(g) \;=\; \mathrm{Mat}^{(v_1,\dots,v_n)}_{(v_1,\dots,v_n)}(f\circ g).$$

Durch Einschränkung erhalten wir also eine Bijektion $\mathrm{Aut}(V) \to \mathrm{GL}(n,\mathbb{K})$, der die Verkettung in $\mathrm{Aut}(V)$ in die Matrizen-Multiplikation überführt, sogenannter Gruppenisomorphismus von $(\mathrm{Aut}(V),\circ)$ nach $(\mathrm{GL}(n,\mathbb{K}),\cdot)$.

4. Dualraum

Sei V ein $\mathbb{K}$-Vektorraum. Dann ist auch $\mathrm{Hom}(V,\mathbb{K})$ ein $\mathbb{K}$-Vektorraum. Wir nennen $V' := \mathrm{Hom}(V,\mathbb{K})$ den Dualraum zu V. Vektoren in V' nennt man auch *Linearformen auf* V. Ist $a \in V'$, $a \neq 0$, $\dim V = n$, dann ist Kern a ein $n-1$-dimensionaler Untervektorraum.

Ist V endlich-dimensional, und $(b_1, b_2, \dots, b_n)$ eine Basis von V, dann definieren wir für $i \in \{1, 2, \dots, n\}$ die Linearform $b_i' : V \to \mathbb{K}$ durch die Angabe der Bilder der Basis $(b_1, \dots, b_n)$.

$$b_i'(b_j) = \delta_{ij}.$$

Proposition 4.1. *Ist $(b_1, \dots, b_n)$ eine Basis von V, so ist die wie oben definierte Familie $(b_1', \dots, b_n')$ eine Basis von V'.*

Beweis. Lineare Unabhängigkeit: sei $\sum_{i=1}^n \lambda_i b_i' = 0_{V'} \in V'$. Wir setzen b_j ein.

$$0_{\mathbb{K}} = 0_{V'}(b_j) = \sum_{i=1}^n \lambda_i b_i'(b_j) = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j.$$

Also sind alle $\lambda_j = 0$ und deswegen ist $(b_1', \dots, b_n')$ linear unabhängig.

Erzeugendensystem: Sei $a \in V'$. Unser Ziel ist es λ_i zu finden, so dass

$$a = \sum_{i=1}^n \lambda_i b_i'. \tag{4.2}$$

Um die Formel zu motivieren, nehmen wir zunächst an, wir hätten bereits (4.2) und werten auf b_j aus und erhalten

$$a(b_j) = \sum_{i=1}^{n} \lambda_i b'_i(b_j) = \sum_{i=1}^{n} \lambda_i \delta_{ij} = \lambda_j.$$

Wenn es also λ_j mit den gewünschten Eigenschaften gibt, dann $\lambda_j = a(b_j)$. Wir probieren also aus:

$$\lambda_i := a(b_i) \qquad \tilde{a} = \sum_{i=1}^{n} \lambda_i b'_i.$$

Es folgt dann wie oben $a(b_j) = \lambda_j = \tilde{a}(b_j)$. Somit stimmen a und $\tilde{a}$ auf der Basis $(b_1, \ldots, b_n)$ überein, also $a = \tilde{a}$. Es folgt, dass $(b'_1, \ldots, b'_n)$ den Raum V' erzeugt. □

Es gilt insbesondere dann $\dim V = \dim V'$.

Achtung: Ist V unendlich-dimensional mit Basis $(b_i | i \in I)$, dann ist die analog definierte Familie $(b'_i | i \in I)$ zwar linear unabhängig, aber keine V' erzeugende Familie. Eine Basis von V' ist dann mächtiger als eine Basis von V.

Sei nun $f : V \to W$ ein Homomorphismus zwischen endlich-dimensionalen Vektorräumen V und W. Sei $(v_1, \ldots, v_m)$ eine Basis von V und $(w_1, \ldots, w_n)$ eine Basis von W,

$$(a_{ij})_{ij} = \mathrm{Mat}^{(v_1,\ldots,v_m)}_{(w_1,\ldots,w_n)}(f).$$

Dann also $f(v_j) = \sum_{i=1}^{n} a_{ij} w_i$. Und somit

$$w'_k(f(v_j)) = \sum_{i=1}^{n} a_{ij} w'_k(w_i) = \sum_{i=1}^{n} a_{ij} \delta_{ki} = a_{kj}.$$

Der Bidualraum ist $V'' = \mathrm{Hom}(\mathrm{Hom}(V, \mathbb{K}), \mathbb{K})$. Es gibt eine natürliche Abbildung

$$\Phi : V \to V'',$$

die wie folgt definiert ist: Zu $v \in V$ definieren wir die lineare Abbildung

$$\begin{array}{rcl} \Phi(v) : \mathrm{Hom}(V, \mathbb{K}) & \to & \mathbb{K} \\ \alpha & \mapsto & \alpha(v) \end{array}$$

Offensichtlich ist Φ linear. Wir wollen zeigen, dass Φ injektiv ist. Sei $v \in V \setminus \{0\}$. Setze $b_1 := v$ und ergänze zu Basis $(b_i | i \in I)$. Die wie oben definierte Linearform b'_1 erfüllt insbesondere $b'_1(v) = 1$.

$$\Phi(v)(b'_1) = b'_1(v) = 1.$$

Also $\Phi(v) \neq 0$. Es folgt Kern $\Phi = \{0\}$, d.h. Φ ist Monomorphismus. Es folgt:

PROPOSITION 4.3. *Ist V ein endlich-dimensionaler Vektorraum, so ist $\Phi : V \to V''$ ein Isomorphismus.*

Beweis. $\dim V = \dim V' = \dim V''$. Da $\Phi : V \to V''$ ein Monomorphismus ist, folgt mit der Rangformel $\dim \mathrm{Bild}\Phi = \dim V = \dim V''$ und somit ist die Abbildung ein Isomorphismus. □

Sei $f : V \to W$ eine lineare Abbildung. Für jedes $a \in W'$ ist dann

$$f'(a) := a \circ f \in V'.$$

$$V \xrightarrow{f} W \xrightarrow{a} \mathbb{K}.$$

Die Abbildung $f' : W' \to V'$ ist ebenfalls linear. Wir nennen f' die zu f *duale Abbildung*. Man kann zeigen, dass $f'' \circ \Phi_V = \Phi_W \circ f$.

$$\begin{tikzcd}
V'' \arrow[r] & W'' \\
V \arrow[u, "\Phi_V"] \arrow[r] & W \arrow[u, "\Phi_W"']
\end{tikzcd}$$

Es gilt offensichtlich für $f, f_1 : V \to W$, $\alpha \in \mathbb{K}$ und $g : W \to X$:

$$(f + f_1)' = f' + f_1'$$
$$(\alpha f)' = \alpha f'$$
$$(g \circ f)' = f' \circ g'$$

Satz 4.4. *Sei $(v_1, \ldots, v_m)$ eine Basis von V, und $(w_1, \ldots, w_n)$ eine Basis von W, und $f : V \to W$ linear. Dann gilt*

$$\mathrm{Mat}^{(w_1', \ldots, w_n')}_{(v_1', \ldots, v_m')}(f') = \left(\mathrm{Mat}^{(v_1, \ldots, v_m)}_{(w_1, \ldots, w_n)}(f)\right)^T.$$

Beweis. Sei wie oben

$$(a_{ij})_{ij} = \mathrm{Mat}^{(v_1, \ldots, v_m)}_{(w_1, \ldots, w_n)}(f)$$

und somit

$$a_{ij} = w_i'(f(v_j)) = w_i' \circ f \circ v_j.$$

Sei analog definiert

$$(b_{ij})_{ij} = \mathrm{Mat}^{(w_1', \ldots, w_n')}_{(v_1', \ldots, v_m')}(f')$$

und wir erhalten

$$b_{ij} = v_i''(f'(w_j')).$$

Nun ist aber $v_i'' = \Phi(v_i)$ und somit

$$v_i''(f'(w_j')) = (f'(w_j'))(v_i) = w_j' \circ f \circ v_i = a_{ji}.$$

$\square$

PROPOSITION 4.5. *Sei $f : V \to W$ ein Homomorphismus zwischen beliebigen Vektorräumen. Dann gilt*

$$\mathrm{Bild}(f') \cong (\mathrm{Bild} f)'$$

ÜBUNGSAUFGABE 4.6. Seien X, Y und Z Vektorräume und $g : X \to Y$ und $h : X \to Z$ Epimorphismen mit Kern $g =$ Kern h. Dann gibt es einen eindeutigen Isomorphismus $I : Y \to Z$ mit

$$h = I \circ g.$$

$$\begin{tikzcd}
 & X \arrow[dl, "g"'] \arrow[dr, "h"] & \\
Y \arrow[rr, "\cong"] & & Z
\end{tikzcd}$$

Beweis der Proposition. Die Abbildung f' ist ein Epimorphismus $W' \to \mathrm{Bild}(f')$, $\mathrm{Bild}(f') \subseteq V'$. Für eine Linearform $a \in W'$ gilt

$$a \in \text{Kern } f' \Leftrightarrow f'(a) = 0 \Leftrightarrow a \circ f = 0 \Leftrightarrow a(f(x)) = 0 \quad \forall x \in V \Leftrightarrow a(y) = 0 \quad \forall y \in \mathrm{Bild} f.$$

Eine Linearform $a \in W'$, $a : W \to \mathbb{K}$ kann auch auf den Unterraum Bild$f$ eingeschränkt werden. Wir erhalten einen Homomorphismus $r : W' \to (\mathrm{Bild} f)'$, $r(a) := a \to a|_{Bildf} \in \mathrm{Hom}(\mathrm{Bild} f, \mathbb{K}) = (\mathrm{Bild} f)'$. Offensichtlich ist $a \in$ Kern r genau dann, wenn $a(y) = 0$ $\forall y \in$ Bild$f$. Es folgt Kern $f' =$ Kern r. Wenn wir noch die Surjektivität von r zeigen, dann folgt die Proposition also aus der Übungsaufgabe ($X = W'$, $Y = \mathrm{Bild}(f')$, $g = f'$, $Z = (\mathrm{Bild} f)'$, $h = r$).

$$\begin{tikzcd}
 & W' \arrow[dl, "f'"'] \arrow[dr, "r"] & \\
\mathrm{Bild}(f') \arrow[rr, "\cong"] & & (\mathrm{Bild} f)'
\end{tikzcd}$$

Sei also $d \in (\mathrm{Bild} f)'$, $d : \mathrm{Bild} f \to \mathbb{K}$. Wir wählen eine Basis $(b_i | i \in I)$ von $\mathrm{Bild} f$ und ergänzen zu einer Basis $(b_i | i \in J)$ von W. Wir definieren $a \in \mathrm{Hom}(W, \mathbb{K})$, indem wir die Werte auf einer Basis angeben:

$$a(b_i) = d(b_i) \quad \forall i \in I \qquad a(b_i) = 0 \quad \forall i \in J \setminus I.$$

Es folgt $r(a) = d$ und somit ist r surjektiv. □

5. Zeilenrang und Spaltenrang

Lemma 5.1. *Sei* $A = (a_{ij})_{ij} \in \mathrm{Mat}(n, m; \mathbb{K})$, $\mathcal{L}_A : \mathbb{K}^m \to \mathbb{K}^n$. *Sei*

$$s_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \in \mathbb{K}^n = \mathrm{Mat}(n, 1; \mathbb{K})$$

die j*-te Spalte. Dann gilt* $\mathrm{Bild}\mathcal{L}_A = \langle s_1, s_2, \ldots, s_n \rangle$.

Wir nennen $\dim \mathrm{Bild}\mathcal{L}_A = \mathrm{Rang}\mathcal{L}_A$ den *Spaltenrang* von A.

Beweis. Es gilt $s_j = \mathcal{L}_A(e_j) = Ae_j$. Wir haben somit

$$\mathrm{Bild}\mathcal{L}_A = \mathcal{L}_A(\mathbb{K}) = \mathcal{L}_A(\langle e_1, e_2, \ldots, e_m \rangle) = \langle \mathcal{L}_A(e_1), \mathcal{L}_A(e_2), \ldots, \mathcal{L}_A(e_m) \rangle = \langle s_1, s_2, \ldots, s_m \rangle.$$

□

Sei $f : V \to W$ eine lineare Abbildung zwischen dem endlich-dimensionalen Vektorraum V mit Basis (v_i) und dem endlich-dimensionalen Vektorraum W mit Basis (w_i), und sei $A := \mathrm{Mat}^{(v_i)}_{(w_i)}(f)$ die zugehörige Matrix. Da $\mathcal{V}_{(w_i)}$ das Bild von $\mathcal{L}_A$ isomorph auf das Bild von f abbildet, gilt $\dim \mathrm{Bild} f = \dim \mathrm{Bild}\mathcal{L}_A$. Der Rang von f ist also gleich dem Spaltenrang von A.

Analog sei

$$z_i := \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{pmatrix} \in \mathrm{Mat}(1, m; \mathbb{K}) \cong \mathrm{Hom}(\mathbb{K}^m, \mathbb{K}) = (\mathbb{K}^m)'$$

die i-te Spalte. Dann bezeichnen wir

$$\dim \langle z_1, z_2, \ldots, z_n \rangle$$

als den *Zeilenrang* von A.

Offensichtlich ist der Zeilenrang von A gleich dem Spaltenrang von A^T, d.h. der Zeilenrang von A ist gleich $\dim \mathrm{Bild}\mathcal{L}_{A^T}$.

Theorem 5.2. *Für jede Matrix* $A \in \mathrm{Mat}(n, m; \mathbb{K})$ *stimmen Zeilenrang und Spaltenrang überein.*

Wir schreiben dann einfach $\mathrm{Rang} A$ für den Zeilen- bzw. Spaltenrang von A.

Beweis mit Dualräumen. Sei S der Spaltenrang von A und Z der Zeilenrang von A. Es gilt nach Definition $S = \dim \mathrm{Bild}\mathcal{L}_A$ und Z ist der Spaltenrang von A^T.

Wir drücken die Abbildung $(\mathcal{L}_A)' : (\mathbb{K}^n)' \to (\mathbb{K}^m)'$ in den Basen $(e_1', \ldots, e_n')$ und $(e_1', \ldots, e_m')$ aus.

$$\mathrm{Mat}^{(e_1',\ldots,e_n')}_{(e_1',\ldots,e_m')}\Big((\mathcal{L}_A)'\Big) = A^T$$

Es folgt $Z = \mathrm{Rang}\big((\mathcal{L}_A)'\big) = \dim \mathrm{Bild}.\big((\mathcal{L}_A)'\big)$.

Andererseits ist $\mathrm{Bild}\big((\mathcal{L}_A)'\big)$ isomorph zu $\big(\mathrm{Bild}\mathcal{L}_A\big)'$ und somit

$$Z = \dim\big(\mathrm{Bild}\mathcal{L}_A\big)' = \dim \mathrm{Bild}\mathcal{L}_A = S.$$

□

6. Beweis von Zeilenrang=Spaltenrang mit elementaren Zeilenumformungen

Notation: $E_{rs} = (\delta_{ri}\delta_{sj})_{ij} \in \mathrm{Mat}(n, m; \mathbb{K})$.

$$E_{rs} = \begin{array}{c} \begin{pmatrix} 0 & \ldots & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots & 0 & \ldots & 0 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow r\text{-te Zeile} \\ \\ \\ \end{array} \\ \begin{array}{ccccc} & & \uparrow & & \\ & & s\text{-te Zeile} & & \end{array} \end{array}$$

Wir betrachten wieder die folgenden elementaren Zeilenumformungen des Gaußschen Verfahrens:

(1) Man multipliziert die j-te Zeile mit einer Zahl $\lambda \in \mathbb{K} \setminus \{0\}$
(2) Sei $\lambda \in \mathbb{K}$, $j, k \in \{1, \ldots, n\}$, $j \neq k$. Man addiert das λ-fache der j-ten Zeile zu der k-ten Zeile. Hierbei ersetzt man die k-te Zeile durch diese neue Zeile.
(3) Man vertauscht die j-te mit der k-ten Zeile, $j \neq k$.
(4) Man streicht eine Zeile, die nur aus Nullen besteht.

Wenn wir die Umformungen (1) bis (3) auf eine Matrix A anwenden, dann entpricht dies der Links-Multiplikation mit einer $n \times n$-Matrix $B \in GL(n, \mathbb{K})$, wobei:

(1) $$B = B^{(1)}_{j\lambda} = \mathbb{1}_n + (\lambda - 1)E_{jj} = \begin{array}{cl} \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & & & \vdots \\ 0 & 0 & \dots & \lambda & \dots & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \dots & 1 \end{pmatrix} . & \begin{matrix} \\ \\ \\ \leftarrow \text{j-te Zeile} \\ \\ \\ \end{matrix} \\ \begin{matrix} \uparrow \\ \text{j-te Spalte} \end{matrix} & \end{array}$$

Der Eintrag λ steht in der t-ten Zeile der j-te Spalte.

(2) $B = B^{(2)}_{jk\lambda} = \mathbb{1}_n + \lambda E_{kj}$.

$$\mathbb{1}_n + \lambda E_{kj} = \begin{array}{cl} \begin{pmatrix} 1 & 0 & \dots & & & 0 \\ 0 & 1 & & & & \\ \vdots & & \ddots & & & 0 \\ \vdots & & & 1 & \lambda & 0 \\ \vdots & & & & \ddots & \\ 0 & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ \leftarrow \quad k\text{-te Zeile} \\ \\ \\ \end{matrix} \\ \begin{matrix} \uparrow \\ j\text{-te Spalte} \end{matrix} & \end{array}$$

(3) $B = B^{(3)}_{jk} = E_{jk} + E_{kj} + \sum_{i \in \{1,2,\dots,n\} \setminus \{j,k\}} E_{ii}$

$$\dots = \begin{array}{cl} \begin{pmatrix} 1 & 0 & & & & & & & & & 0 & \\ 0 & 1 & & & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & & \ddots & & & & & & & & \\ & & & & 0 & & & & 1 & & & \\ & & & & & 1 & & & & & & \\ & & & & & & \ddots & & & & & \\ & & & & & & & 1 & & & & \\ & & & & 1 & & & & 0 & & & \\ & & & & & & & & & 1 & & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \leftarrow k\text{-te Zeile} \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \uparrow & \qquad & \uparrow \\ j\text{-te} & & k\text{-te} \\ \text{Spalte} & & \text{Spalte} \end{matrix} & \end{array}$$

Lemma 6.1. *Sei $A \in \mathrm{Mat}(n,m;\mathbb{K})$ und $B \in GL(n,\mathbb{K})$. Dann ist der Spaltenrang von A gleich dem Spaltenrang von BA. Elementare Zeilenumformungen vom Typ (1)–(3) bewahren also den Spaltenrang.*

Beweis. Die Abbildung $\mathcal{L}_B : \mathbb{K}^n \to \mathbb{K}^n$ ist ein Isomorphismus.

$$\mathrm{Bild}\mathcal{L}_{BA} = \mathrm{Bild}\mathcal{L}_B \circ \mathcal{L}_A = \mathcal{L}_B(\mathcal{L}_A(\mathbb{K}^n)) = \mathcal{L}_B(\mathrm{Bild}\mathcal{L}_A).$$

Die Einschränkung $\mathcal{L}_B|_{\mathrm{Bild}\mathcal{L}_A} : \mathrm{Bild}\mathcal{L}_A \to \mathrm{Bild}\mathcal{L}_B \circ \mathcal{L}_A$ ist also auch ein Isomorphismus. Somit

$$\dim \mathrm{Bild}\mathcal{L}_{BA} = \dim \mathrm{Bild}\mathcal{L}_B \circ \mathcal{L}_A = \dim \mathrm{Bild}\mathcal{L}_A.$$

Also haben A und BA denselben Spaltenrang.

$$\begin{tikzcd}
\mathbb{K}^n \arrow[r, "\mathcal{L}_B", "\cong"'] & \mathbb{K}^n \\
\mathrm{Bild}\mathcal{L}_A \arrow[u] \arrow[r, "\mathcal{L}_{B|\mathrm{Bild}\mathcal{L}_A}", "\cong"'] & \mathrm{Bild}\mathcal{L}_{BA} \arrow[u]
\end{tikzcd}$$

□

Man sieht auch sofort, dass Zeilenumformungen von Typ (4) den Spaltenrang erhalten.

Lemma 6.2. *Elementare Zeilenumformungen lassen den Zeilenrang unverändert. Sei lassen sogar den von den Zeilen aufgespannten Vektorraum unverändert.*

Beweis. Die Behauptung ist offensichtich für Typ (4).
Seien $z_1, \dots, z_n$ die Zeilen vor einer elementaren Zeilenumformung von Typ (1)-(3) und seien $z_1', \dots, z_n'$ die Zeilen nach der Umformung. Man erhält jedes z_i' als Linearkombination von z_1, z_2,… z_n, also

$$\langle z_1', z_2', \dots, z_n' \rangle \subseteq \langle z_1, z_2, \dots, z_n \rangle.$$

Jede elementare Zeilenumformung von Typ (1)–(3) kann durch eine Umformung gleichen Typs rückgängig gemacht werden. Daraus ergibt sich deswegen auch

$$\langle z_1, z_2, \dots, z_n \rangle \subseteq \langle z_1', z_2', \dots, z_n' \rangle.$$

□

Achtung: Der von den Spalten erzeugte Vektorraum wird im allgemeinen durch elementare Zeilenumformungen verändert, nur die Dimension bleibt unverändert.

Proposition 6.3. *Sei $A \in \mathrm{Mat}(n,m;\mathbb{K})$. Dann erhalten wir aus A durch endlich viele elementare Zeilenumformungen vom Typ (1)–(3) eine Matrix in Zeilenstufenform*

$$Z = \begin{pmatrix} & & & 0 & 0 & \dots & 0 \\ & \mathbb{1}_r & & 0 & 0 & \dots & 0 \\ & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \in \mathrm{Mat}(n,m;\mathbb{K}).$$

In anderen Worten es gibt endlich viele $B_1, B_2, \dots, B_k \in \mathrm{GL}(n,\mathbb{K})$ vom Typ (1)–(3), so dass

$$B_k B_{k-1} \cdots B_2 B_1 A$$

Zeilenstufenform hat.

Der Beweis folgt direkt aus der Tatsache, dass das Gaußsche Verfahren immer erfolgreich durchgeführt werden kann.

Beweis von Spaltenrang=Zeilenrang. Sei r die Zahl der nicht-verschwindenden Zeilen. Dann erzeugen die Spalten den Raum $\{sum_{i=1}^r \lambda_i e_i \,|\, \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{K}\}$, der offensichtlich r-dimensional ist. Also ist der Spaltenrang von Z gleich r.

Die nicht-verschwindenden Zeilen sind linear unabhängig, also erzeugen sie einen r-dimensionalen Raum, also ist auch der Zeilenrang von Z gleich r. Durch elementare Zeilenumformungen, d.h. durch Linksmultiplikation mit den B_i ändert sich weder Zeilenrang noch Spaltenrang. Also hat auch A Zeilenrang r und Spaltenrang r. Zeilenrang und Spaltenrang von A sind also gleich. $\square$

KAPITEL 6

Determinanten

1. Motivation

Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ eine quadratische Matrix. Wir definieren:

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

Diese Abbildung erfüllt:

(1) $\det A \neq 0 \Leftrightarrow A$ ist invertierbar $\Leftrightarrow \mathrm{Rang} A = 2$
(2) $\det A = \det A^T$
(3) *Bilinearität in den Spalten:* $\forall (a_{ij}) \in \mathrm{Mat}(2,2;\mathbb{K}),\ \lambda, a'_1, a'_2 \in \mathbb{K}$:

$$\det \begin{pmatrix} a_{11}+a'_1 & a_{12} \\ a_{21}+a'_2 & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \det \begin{pmatrix} a'_1 & a_{12} \\ a'_2 & a_{22} \end{pmatrix}$$

$$\det \begin{pmatrix} \lambda a_{11} & a_{12} \\ \lambda a_{21} & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \lambda a_{12} \\ a_{21} & \lambda a_{22} \end{pmatrix} = \lambda \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

(4) *Bilinearität in den Zeilen:* $\forall (a_{ij}) \in \mathrm{Mat}(2,2;\mathbb{K}),\ \lambda, a'_1, a'_2 \in \mathbb{K}$:

$$\det \begin{pmatrix} a_{11}+a'_1 & a_{12}+a'_2 \\ a_{21}+ & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \det \begin{pmatrix} a'_1 & a'_2 \\ a_{21} & a_{22} \end{pmatrix}$$

$$\det \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix} = \lambda \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

(5) *Alternierend in den Spalten*

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{pmatrix}$$

(6) *Alternierend in den Zeilen*

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix}$$

(7) $\det AB = \det A \det B$, $\det \mathbb{1}_2 = 1$ und $\det A^{-1} = (\det A)^{-1}$

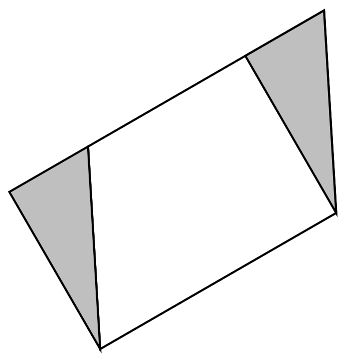


ABBILDUNG 1. Berechnung des Fächeninhalts

(8) $\det \begin{pmatrix} a_{11} & a_{12} \\ \lambda a_{11} & \lambda a_{12} \end{pmatrix} = 0$

LEMMA 1.1. *Sei V ein 2-dimensionaler Vektorraum mit einer Basis (v_1, v_2) und einer weiteren Basis (w_1, w_2). Sei $f \in End(V)$. Dann gilt*

$$\det \mathrm{Mat}^{(v_1,v_2)}_{(v_1,v_2)}(f) = \det \mathrm{Mat}^{(w_1,w_2)}_{(w_1,w_2)}(f).$$

Beweis. Sei $(w_1, w_2) = (v_1, v_2)B$. Dann haben wir gesehen, dass

$$\mathrm{Mat}^{(w_1,w_2)}_{(w_1,w_2)}(f) = B^{-1}\mathrm{Mat}^{(v_1,v_2)}_{(v_1,v_2)}(f)B.$$

Wir rechnen

$$\det \mathrm{Mat}^{(w_1,w_2)}_{(w_1,w_2)}(f) = (\det B^{-1})(\det \mathrm{Mat}^{(v_1,v_2)}_{(v_1,v_2)}(f))(\det B) = \det \mathrm{Mat}^{(v_1,v_2)}_{(v_1,v_2)}(f).$$

□

Geometrische Interpretation ($\mathbb{K} = \mathbb{R}$).

Ein Parallelogramm P werde von den Vektoren $a_1 := \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ und $a_2 := \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ aufgespannt. Dann ist der Flächeninhalt von P gegeben durch

$$\left| \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right|.$$

Begründung: Wähle $b_1 := a_1/\|a_1\| = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ für ein geeignetes $\alpha \in \mathbb{R}$, und dann $b_2 := \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$.

Wir drücken a_1 und a_2 in der Basis (b_1, b_2) aus:

$$a_1 = \|a_1\| b_1 \qquad a_2 = \lambda a_1 + \mu a_2,$$

wobei $\|a_1\| := \sqrt{a_{11}^2 + a_{21}^2}$. Es gilt $\|b_1\| = \|b_2\| = 1$. Durch Abschneiden des gestreiften Dreiecks und Ankleben an der anderen Seite (Abbildung 1) sehen wir, dass der Flächeninhalt von P gleich dem Flächeninhalt des von a_1 und μa_2 aufgespannten Rechtecks ist, also gleich $\|a_1\|\,|\mu|$.

Andererseits gilt

$$\det(a_1, a_2) = \det(\|a_1\| b_1, \lambda b_1 + \mu b_2) = \|a\| \det(b_1, \lambda b_1 + \mu b_2)$$

$$= \|a\| \left(\underbrace{\det(b_1, \lambda b_1)}_{=0} + \det(b_1, \mu b_2) \right) = \|a\|\, \mu((\cos\alpha)^2 + (\sin\alpha)^2) = \|a\|\, \mu.$$

Analog: Berechnung des Flächeninhalts eines Dreiecks.

Frage 1.2. Gibt es auch auf $\mathrm{Mat}(n, n; \mathbb{K})$ eine Determinante?

2. Die symmetrischen Gruppen

Wiederholung: Wir versehen die Menge $\mathcal{S}_n := \mathrm{Bij}(\{1, \ldots, n\})$, $n \in \mathbb{N}$ mit der Verkettung von Abbildungen. Dies ist eine Gruppe mit neutralem Element $\mathrm{Id}_{\{1,\ldots,n\}}$. Man nennt sie die *Permutationsgruppe* oder die *symmetrische Gruppe zum Index n*. Permutationen $\sigma \in \mathcal{S}_n$ wollen wir in der Form

$$\begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix}$$

schreiben. Schreibweise für $k \in \mathbb{N}$

$$\sigma^k = \underbrace{\sigma \circ \cdots \circ \sigma}_{k\text{-mal}} \quad \sigma^{-k} = \underbrace{\sigma \circ \cdots \circ \sigma}_{k\text{-mal}} \quad \sigma^0 = \mathrm{Id}.$$

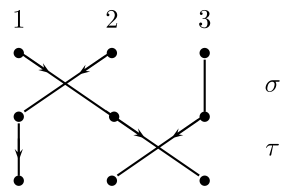
Es gilt $\#\mathcal{S}_n = n!$. Die Gruppen $\mathcal{S}_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ und $\mathcal{S}_2 = \left\{ \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \right\}$ sind abelsch, alle anderen nicht.

Beispiel: $n = 3$:

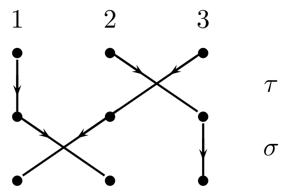
$$\sigma := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \qquad \tau := \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}.$$

Dann

$$\tau \circ \sigma = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix} \qquad \sigma \circ \tau = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}.$$



$\tau \circ \sigma$



$\sigma \circ \tau$

DEFINITION 2.1. Eine Permutation $\sigma \in \mathcal{S}_n$ heißt *Transposition von i und j*, $i \neq j$, falls $\sigma(i) = j$ und $\sigma(j) = i$ und $\sigma(k) = k \quad \forall k \in \{1, \ldots, n\} \setminus \{i, j\}$. Alternativ zur obigen Schreibweise schreiben wir dann $\sigma = \big[i, j\big]$.

Sei $m \geq 2$. Gibt es m verschiedene Elemente $i_1, i_2, \ldots, i_m \in \{1, 2, \ldots, n\}$ und gilt $\sigma(i_j) = \sigma(i_{j+1}) \quad \forall j \in \{1, 2, \ldots, m-1\}$, $\sigma(i_m) = \sigma(i_1)$, und $\sigma(k) = k \quad \forall k \in \{1, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_m\}$, dann nennt man σ einen *Zykel der Länge m*. (Zykel der Länge 2 sind dasselbe wie Transpositionen.) Wir schreiben diesen Zykel in der Form

$$\big[i_1, i_2, \ldots, i_m\big] .$$

Man sagt, die Elemente i_1, i_2, $\ldots$, i_m werden zyklisch vertauscht.

Zwei Zykel σ_1 und σ_2 heißen disjunkt, wenn kein Element sowohl von σ_1 als auch von σ_2 verschoben wird, in anderen Worten, falls $\sigma_1(i) = i$ oder $\sigma_2(i) = i$ für alle $i \in \{1, 2, \ldots, n\}$.

Beispiel: $n = 3$, σ, τ wie oben.
$\sigma = [1, 2]$, $\tau = [2, 3]$, $\tau \circ \sigma = [1, 2, 3] = [2, 3, 1] = [3, 1, 2]$ und $\sigma \circ \tau = [2, 1, 3] = [1, 3, 2] = [3, 2, 1]$.
$\mathcal{S}_3 = \{\mathrm{Id}, [1, 2], [2, 3], [1, 3], [1, 2, 3], [1, 3, 2]\}$.

BEISPIEL. $n = 4$. Die Permutation $[1, 2] \circ [2, 3]$ ist gleich $[1, 2, 3]$ und somit ein Zykel. Die Permutation $[1, 2] \circ [3, 4]$ ist kein Zykel.

Sei $\sigma \in \mathcal{S}_n$, $i \in \{1, 2, \ldots, n\}$. Wir definieren den *$\sigma$-Orbit von i* als die Menge

$$O_\sigma(i) := \{\sigma^k(i) \,|\, k \in \mathbb{Z}\}.$$

ÜBUNGSAUFGABE 2.2. (1)

$$O_\sigma(i) = \{\sigma^k(i) \,|\, k \in \mathbb{N}\}$$

(2) $i \sim j :\Leftrightarrow j \in O_\sigma(i)$ ist eine Äquivalenzrelation auf der Menge $\{1, 2, \ldots, n\}$.

Offensichtlich ist $\sigma|_{O_\sigma(i)} : O_\sigma(i) \to O_\sigma(i)$ bijektiv und ist gleich $[i, \sigma(i)\sigma^2(i), \ldots, \sigma^{r-1}]|_{O_\sigma(i)} : O_\sigma(i) \to O_\sigma(i)$, wobei $r := \#O_\sigma(i)$.

PROPOSITION 2.3. *Jede Permutation ist die Verkettung von disjunkten Zykeln.*

Beweis.
1. Schritt: Setze $I_1 := \{1, 2, \dots, n\}$. Wir wählen zunächst ein $i_1 \in I_1$ aus, z.B. $i_1 = 1$. Es gilt dann $\sigma(O_\sigma(i_1)) = O_\sigma(i_1)$ und $\sigma(I_1 \setminus O_\sigma(i_1)) = I_1 \setminus O_\sigma(i_1)$. Falls $O_\sigma(i_1) = I_0$, dann ist der 1. Schritt beendet.
Ansonsten wählen wir nun ein $i_2 \in I_2 := I_1 \setminus O_\sigma(i_1)$, $I_3 := I_2 \setminus O_\sigma(i_2)$. Wiederum gilt $\sigma(O_\sigma(i_2)) = O_\sigma(i_2)$ und $\sigma(I_3) = I_3$. Falls $I_3 = \emptyset$, dann ist der 1. Schritt beendet.
Ansonsten gehen wir analog weiter, wählen rekursiv $i_k \in I_k$, setzen $I_{k+1} := I_k \backslash O_\sigma(i_k)$ bis schließlich $I_{m+1} = 0$.

2.Schritt: Wir erhalten nun eine disjunkte Zerlegung

$$\{1, 2, \dots, n\} = O_\sigma(i_1) \dot\cup O_\sigma(i_2) \dot\cup \dots \dot\cup O_\sigma(i_m).$$

Wir können (nach eventueller Umordnung) annehmen, dass die ein-elementigen Orbiten rechts liegen, sagen wir es gelte $\#O_\sigma(i_j) = 1$ genau dann, wenn $j > \ell$. Wir setzen $r_j := O_\sigma(i_j)$. Dann sieht man sofort, dass

$$\sigma = [i_1, \sigma(i_1), \sigma^2(i_1), \dots, \sigma^{r_1}(i_1)] \circ [i_2, \sigma(i_2), \sigma^2(i_2), \dots, \sigma^{r_2}(i_2)] \circ \dots \circ [i_\ell, \sigma(i_\ell), \sigma^2(i_\ell), \dots, \sigma^{r_\ell}(i_\ell)].$$

□

Bemerkung: Disjunkte Zykel kommutieren.

Beispiel:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 6 & 3 & 5 & 1 & 2 \end{bmatrix} = [1, 4, 5] \circ [2, 6]$$

$$\{1, 2, 3, 4, 5, 6\} = O(1) \dot\cup O(2) \dot\cup O(3) = O(4) \dot\cup O(6) \dot\cup O(3)$$

Proposition 2.4. *Jede Permutation ist die Verkettung von Transpositionen*

Beweis. Wir zerlegen eine gegebene Permutation zunächst in Zykel. Wir zerlegen jeden Zykel weiter:

$$[i, \sigma(i), \sigma^2(i), \dots, \sigma^{r-1}(i)] = [i, \sigma(i)] \circ [\sigma(i), \sigma^2(i)] \circ \dots \circ [\sigma^{r-2}(i), \sigma^{r-1}(i)].$$

□

Beispiel:

$$[1, 2, 3, 4] = [1, 2] \circ [2, 3] \circ [3, 4].$$

Ein *Fehlstand* einer Permutation $\sigma \in \mathcal{S}_n$ ist eine Teilmenge $\{i, j\} \subseteq \{1, \dots, n\}$, $i < j$ mit $\sigma(i) > \sigma(j)$.

Lemma 2.5. *Sei* $\sigma \in \mathcal{S}_n$.

$$\prod_{1 \le i < j \le n} (\sigma(j) - \sigma(i)) = (-1)^p \prod_{1 \le i < j \le n} (j - i),$$

wobei p *die Anzahl der Fehlstände von* σ *ist.*

Beweis. Sei $\mathcal{P}_2^n$ die Menge der 2-elementigen Teilmengen von $\{1,\ldots,n\}$. Zu σ erhalten wir eine Abbildung $\sigma_* : \mathcal{P}_2^n \to \mathcal{P}_2^n$, $\sigma_*(\{i,j\}) = \{\sigma(i), \sigma(j)\}$. Da $(\sigma^{-1})_*$ die Umkehrabbildung von σ_* ist, ist σ_* eine Bijektion.

Es folgt

$$\prod_{1\le i<j\le n} |\sigma(j)-\sigma(i)| = \prod_{\{i,j\}\in\mathcal{P}_2^n} |\sigma(j)-\sigma(i)| = \prod_{s,t\in\mathcal{P}_2^n} |t-s| = \prod_{1\le i<j\le n} |j-i|.$$

Wenn wir analog $\prod_{1\le i<j\le n}(\sigma(j)-\sigma(i))$ in $\prod_{1\le i<j\le n}(j-i)$ umformen, dann erhalten wir für jeden Fehlstand ein Minuszeichen. □

Definition 2.6. Wir definieren das Signum von σ

$$\operatorname{sgn}(\sigma) := \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i} = \prod_{\{i,j\}\in\mathcal{P}_2^n} \frac{\sigma(j)-\sigma(i)}{j-i} \in \{-1,1\}.$$

Das rechte Produkt ist wohldefiniert, da $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$. Wir nennen σ *gerade*, wenn $\operatorname{sgn}(\sigma) = 1$, und *ungerade*, wenn $\operatorname{sgn}(\sigma) = -1$.

Beispiel: Sei $\sigma = [p,q]$ eine Transposition, $1 \le p < q \le n$. Dann ist $\{i,j\}$, $i<j$, genau dann ein Fehlstand, wenn

(a) $i = p$ und $j = q$ oder
(b) $i = p$ und $i < q < j$ oder
(c) $j = q$ und $i < p < j$.

Also haben wir $2(q-p)-1$ Fehlstände, und somit $\operatorname{sgn}(\sigma) = -1$.

Satz 2.7. *Für $\sigma, \tau \in \mathcal{S}_n$ gilt*

$$\operatorname{sgn}(\tau\circ\sigma) = \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma).$$

Beweis.

$$\operatorname{sgn}(\tau\circ\sigma) = \prod_{\{i,j\}\in\mathcal{P}_2^n} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{j-i} = \prod_{\{i,j\}\in\mathcal{P}_2^n} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \underbrace{\prod_{\{i,j\}\in\mathcal{P}_2^n} \frac{\sigma(j)-\sigma(i)}{j-i}}_{\operatorname{sgn}(\sigma)}$$

Wir setzen wieder $s = \sigma(i)$ und $t = \sigma(j)$ und erhalten

$$\prod_{\{i,j\}\in\mathcal{P}_2^n} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} = \prod_{\{s,t\}\in\mathcal{P}_2^n} \frac{\tau(t)-\tau(s)}{t-s} = \operatorname{sgn}\tau.$$

□

KOROLLAR 2.8. *Eine Permutation ist genau dann gerade, wenn sie die Verkettung einer geraden Anzahl von Transpositionen ist. Sie ist genau dann ungerade, wenn sie eine Verkettung einer ungeraden Anzahl von Transpositionen ist.*

□

3. Multilineare Abbildungen

DEFINITION 3.1. Seien V_1, V_2, $\ldots$, V_r, W Vektorräume über $\mathbb{K}$. Eine Abbildung $F : V_1 \times V_2 \ldots \times V_r \to W$ heißt *multilinear* oder *r-linear* (über $\mathbb{K}$), wenn für alle $j \in \{1, \ldots, n\}$ und alle $v_1 \in V_1$, $v_2 \in V_2$,$\ldots v_r \in V_r$ ist die Abbildung

$$V_j \to W \qquad v \mapsto F(v_1, v_2, \ldots, \underbrace{\cancel{v_j}}_{v}, \ldots, v_n)$$

linear ist. Für 2-linear sagt man normalerweise *bilinear*.

BEISPIELE. (1) Jede lineare Abbildung ist 1-linear.

(2) Die Abbildung

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right) \mapsto \sum_{i=1}^n x_i y_i =: \left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle$$

ist bilinear. Man nennt sie das *kanonische Skalarprodukt* auf $\mathbb{R}^n$.

(3) Die 2×2-Determinante, aufgefasst als Abbildung auf den Spaltenvektoren ist eine bilineare Abbildung

$$\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}.$$

(4) Die 2×2-Determinante ist auch bilinear in den Zeilenvektoren.

(5) Das Kreuzprodukt oder Spatprodukt

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \qquad \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) \mapsto \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

ist bilinear.

DEFINITION 3.2. Sei $\mathbb{K}$ ein Körper mit $2 \neq 0 \in \mathbb{K}$. (Für Körper mit $2 = 0 \in \mathbb{K}$, zum Beispiel $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$ siehe unten.) Eine Abbildung $F : V \times V \ldots V \to W$ heißt *alternierend*, falls für alle $v_1, \ldots, v_r \in V$ und $i \neq j$ gilt:

$$F(v_1, v_2, \ldots, \underbrace{\cancel{v_i}}_{v_j}, v_{i+1}, \ldots, \underbrace{\cancel{v_j}}_{v_i}, \ldots v_n) = -F(v_1, v_2, \ldots, v_n) \qquad (*)$$

Die Abbildung wechselt also das Vorzeichen, wenn wir die i-te und j-te Spalte vertauschen für $i \neq j$.

Beispiel. Die Abbildung

$$\mathbb{K}^{2n} \times \mathbb{K}^{2n} \to \mathbb{K}, \qquad \left(\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \right) \mapsto \sum_{i=1}^{n} x_{2i-1} y_{2i} - x_{2i} y_{2i-1}$$

ist eine alternierende bilineare Abbildung. Sie heißt *kanonische symplektische Struktur auf* $\mathbb{K}^{2n}$ und ist zum Beispiel in der klassischen Mechanik wichtig.

Lemma 3.3. *Für eine alternierende r-lineare Abbildung F und $\sigma \in \mathcal{S}_r$ gilt*

$$F(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(r)}) = (\operatorname{sgn} \sigma) F(v_1, v_2, \dots, v_r).$$

Beweis. Wir schreiben σ als Verkettung von p Transpositionen. Wir erhalten $F(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(r)})$ aus $F(v_1, v_2, \dots, v_r)$, indem wir nacheinander diese Transpositionen anwenden. Für jede Transposition erhalten wir ein Minuszeichen, also $F(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(r)}) = (-1)^p F(v_1, v_2, \dots, v_r)$. □

Im folgenden Lemma benutzen wir die Sprechweise: Eine Familie $(v_1, \dots, v_m)$ hat mindestens zwei gleiche Elemente, wenn es $i,\, j \in \{1, \dots, n\}$, $i \neq j$, gibt, so dass $v_i = v_j$.

Lemma 3.4. *Sei $2 \neq 0$. Für eine r-lineare Abbildung $F : V \times V \dots V \to \mathbb{K}$ sind äquivalent:*

(1) *F ist alternierend.*
(2) *Für alle Familien $(v_1, v_2, \dots, v_n)$ mit mindestens zwei verschiedenen Elementen gilt*

$$F(v_1, v_2, \dots, v_r) = 0.$$

Beweis. Es gelte (1) und $v_i = v_j$. Dann haben wir

$$F(v_1, \dots, v_r) = -F(v_1, \dots, \underbrace{\cancel{v_i}}_{v_j}, \dots, \underbrace{\cancel{v_j}}_{v_i}, \dots v_n) = -F(v_1, v_2, \dots, v_n).$$

Also

$$2F(v_1, \dots, v_n) = 0.$$

Da $2 \neq 0$, ergibt sich die Aussage (2).

Es gelte nun (2):

$$\begin{aligned} 0 &= F(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) \\ &= \underbrace{F(v_1, \dots, v_i, \dots, v_i, \dots, v_n)}_{0} + F(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + F(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\ &\quad + \underbrace{F(v_1, \dots, v_j, \dots, v_j, \dots, v_n)}_{0}, \end{aligned}$$

und daraus folgt (1). □

Bemerkung für Körper mit $2 = 0$.
Wenn wir wie oben vorgehen, impliziert auch in diesem Fall (2) die Eigenschaft (*), aber nicht umgekehrt. Da wir aber (2) später brauchen, definiert man alternierend besser wie folgt: Eine r-lineare Abbildung heißt alternierend, wenn für alle Familien $(v_1, v_2, \dots, v_n)$ mit mindestens zwei verschiedenen Elementen gilt

$$F(v_1, v_2, \dots, v_r) = 0.$$

Diese Definition stimmt im Fall $2 \neq 0$ mit der obigen Definition überein und erfüllt in allen Fällen die Eigenschaft (*). äquivalenten Eigenschaft.

Lemma 3.5. *Sei* $F : V \times V \dots V \to \mathbb{K}$ *eine alternierende* r*-lineare Abbildung, und sei* $(v_1, v_2, \dots, v_r)$ *linear abhängig, dann ist*

$$F(v_1, v_2, \dots, v_r) = 0.$$

Beweis. Auf Grund der linearen Abhängigkeit finden wir ein v_i, das sich als Linearkombination der anderen darstellen lässt. O.B.d.A. $i = 1$, also

$$v_1 = \sum_{i=2}^{n} \lambda_i v_i.$$

Dann gilt

$$F(v_1, v_2, \dots, v_n) = F(\sum_{i=2}^{n} \lambda_i v_i, v_2, \dots, v_n) = \sum_{i=2}^{n} \lambda_i \underbrace{F(v_i, v_2, \dots, v_n)}_{0} = 0.$$

□

Korollar 3.6. *Ist* $F : V \times \dots \times V \to W$ *eine* r*-lineare Abbildung und* $r > \dim V$*, dann bildet* F *alles auf* 0 *ab.*

4. Alternierende r-Formen, Determinantenformen, Determinanten

Definition 4.1. Eine alternierende r-lineare Abbildung mit Zielvektorraum $W = \mathbb{K}$ nennt man auch *(alternierende) r-Form.* Gilt $n = \dim V$, dann nennt man eine alternierende n-Form auf V auch Determinantenform.

Lemma 4.2. *Die Funktion*

$$\det_n : \underbrace{\mathbb{K}^n \times \mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n-mal} \to \mathbb{K}$$

$$\left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} \right) \mapsto \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

ist eine Determinantenform auf $\mathbb{K}^n$. *Außerdem ist*

$$\det(e_1, e_2, \ldots, e_n) = 1.$$

Zumeist schreiben wir einfach det für $\det_n$.

Beweis im Fall $2 \neq 0$..

Setze $a_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$. Die Abbildung det ist multilinear: wir zeigen dazu die Linearität im j-ten Eintrag:

$$\det(a_1, a_2, \ldots, \lambda a_j, \ldots, a_n) = \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots (\lambda a_{\sigma(j)j}) \cdots a_{\sigma(n)n} = \lambda \det(a_1, a_2, \ldots, a_n)$$

$$\begin{aligned} \det(a_1, a_2, \ldots, a_j + \tilde{a}_j, \ldots, a_n) &= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots (a_{\sigma(j)j} + \tilde{a}_{\sigma(j)j}) \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(j)j} \cdots a_{\sigma(n)n} \\ &\quad + \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots \tilde{a}_{\sigma(j)j} \cdots a_{\sigma(n)n} \\ &= \det(a_1, a_2, \ldots, a_j, \ldots, a_n) + \det(a_1, a_2, \ldots, \tilde{a}_j, \ldots, a_n) \end{aligned}$$

für einen Vektor $\tilde{a}_j := \begin{pmatrix} \tilde{a}_{1j} \\ \tilde{a}_{2j} \\ \vdots \\ \tilde{a}_{nj} \end{pmatrix}$.

Die Abbildung ist also n-linear.

Wir betrachten nun die Vertauschung von a_i und a_j. Sei hierzu τ die Transposition $[i,j]$. Für $\sigma \in \mathcal{S}_n$ setze $\tilde{\sigma} := \sigma \circ \tau$. Dann gilt auch $\sigma = \tilde{\sigma} \circ \tau$.

$$
\begin{aligned}
&\det(a_1, a_2, \dots, \underbrace{\cancel{a_i}}_{a_j}, a_{i+1}, \dots, \underbrace{\cancel{a_j}}_{a_i}, \dots a_n) \\
&= \det(a_{\tau(1)}, a_{\tau(2)}, \dots, a_{\tau(n)}) \\
&= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1),\tau(1)} a_{\sigma(2),\tau(2)} \cdots a_{\sigma(n)\tau(n)} \\
&= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\tilde{\sigma} \circ \tau) a_{\tilde{\sigma}\circ\tau(1),\tau(1)} a_{\tilde{\sigma}\circ\tau(2),\tau(2)} \cdots a_{\tilde{\sigma}\circ\tau(n),\tau(n)} \\
&= \sum_{\tilde{\sigma} \in \mathcal{S}_n} (\operatorname{sgn} \tilde{\sigma})(\operatorname{sgn} \tau) a_{\tilde{\sigma}(1),1} a_{\tilde{\sigma}(2),2} \cdots a_{\tilde{\sigma}(n),n} \\
&= -\det(a_1, a_2, \dots, a_n)
\end{aligned}
$$

□

Siehe [**1**, Abschnitt 4.2] für den Fall $2 = 0$.

Geometrische Interpretation ($\mathbb{K} = \mathbb{R}$).
$|\det(v_1, v_2, \dots, v_n)|$ ist das Volumen des von v_1, v_2, …, v_n aufgespannten Parallelotops P in $\mathbb{R}$.

$$P := \left\{ \sum_{i=1}^{n} \lambda_i v_i \mid \lambda_i \in [0,1] \right\}.$$

Definition 4.3. Die Menge aller alternierenden r-Formen auf V bezeichnen wir mit $\mathcal{A}^r(V)$. Konvention: $\mathcal{A}^0(V) := \mathbb{K}$.

Häufig sieht man in Büchern $\Lambda^r V'$ an Stelle von $\mathcal{A}^r(V)$.

$\mathcal{A}^r(V)$ ist ein Vektorraum mit der folgenden Addition und Multiplikation mit Skalaren: für alle $F_1, F_2 \in \mathcal{A}^r(V)$, $v_1, \dots, v_r \in V$, $\lambda \in \mathbb{K}$:

$$
\begin{aligned}
(F_1 + F_2)(v_1, \dots, v_r) &= F_1(v_1, \dots, v_r) + F_2(v_1, \dots, v_r), \\
(\lambda F_1)(v_1, \dots, v_r) &= \lambda F_1(v_1, \dots, v_r).
\end{aligned}
$$

Beispiele. (1) $\mathcal{A}^1(V) = V'$.
(2) $\det \in \mathcal{A}^n(\mathbb{K}^n)$.
(3) Sei $a \in (\mathbb{R}^3)'$ und $\times$ das Kreuzprodukt. Sei

$$F(v, w) := a(v \times w)$$

für $v, w \in \mathbb{R}^3$. Dann ist $F \in \mathcal{A}^2(\mathbb{R}^3)$. Z. B. für $a = e_1'$ gilt

$$F\left(\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right) = v_2 w_3 - v_3 w_2.$$

(4) Sei $v \in V = \mathbb{K}^r$,
$$F(v_1, v_2, \ldots, v_{r-1}) = \det(v, v_1, v_2, \ldots, v_{r-1}).$$
Dann ist $F \in \mathcal{A}^{r-1}(\mathbb{K}^r)$.
(5) $\mathcal{A}^r(V) = \{0\}$, falls $r > \dim V$.

Definition 4.4. Seien $a_1, a_2, \ldots, a_r \in V'$. Dann ist $a_1 \wedge a_2 \wedge \cdots \wedge a_r \in \mathcal{A}^r(V)$ wie folgt definiert
$$(a_1 \wedge a_2 \wedge \cdots \wedge a_r)(v_1, v_2, \ldots, v_r) := \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) a_{\sigma(1)}(v_1) a_{\sigma(2)}(v_2) \cdots a_{\sigma(r)}(v_r).$$

Übung: Überprüfen Sie, dass dies tatsächlich eine alternierende r-Form ist.

Beispiele.
$$(a_1 \wedge a_2)(v_1, v_2) = a_1(v_1)a_2(v_2) - a_1(v_2)a_2(v_1)$$
$$\det_n = e_1' \wedge e_2' \wedge \ldots \wedge e_n'$$

Physikalisches Beispiel: In der Speziellen Relativitätstheorie ist die Raumzeit ein reeller 4-dimensionaler Vektorraum V. Das elektromagnetische Feld ist eine Funktion
$$F : V \to \mathcal{A}^2(V).$$
Die Maxwell-Gleichungen werden dann ganz einfach: $dF = 0$ und $\delta F = *J$.

Schreibweise: Wir zerlegen $A \in \mathrm{Mat}(n, n; \mathbb{K})$ in seine Spalten: $A = (a_1 \quad a_2 \quad \cdots \quad a_n)$. Wir schreiben dann einfach $\det_n A$ für $\det(a_1, a_2, \ldots, a_n)$.

Proposition 4.5. *Sei $(v_1, v_2, \ldots, v_r)$ eine Familie von Vektoren in V, und $A \in \mathrm{Mat}(r, r; \mathbb{K})$, $(w_1, w_2, \ldots, w_r) = (v_1, v_2, \ldots, v_r) \cdot A$. Dann gilt für $F \in \mathcal{A}^r(V)$:*
$$F(w_1, \ldots, w_r) = F(v_1, \ldots, v_r)(\det_r A)$$

Beweis. Wir schreiben $A = (a_{ij})_{ij}$. Dann gilt $w_j = \sum_{i=1}^r a_{ij} v_i$. Somit ergibt sich
$$\begin{aligned} F(w_1, w_2, \ldots, w_r) &= F\left(\sum_{i_1=1}^r a_{i_1 1} v_{i_1}, \sum_{i_2=1}^r a_{i_2 2} v_{i_2}, \ldots, \sum_{i_r=1}^r a_{i_r r} v_{i_r}\right) \\ &= \sum_{i_1=1}^r \sum_{i_2=1}^r \cdots \sum_{i_r=1}^r a_{i_1 1} a_{i_2 2} \cdots a_{i_r r} F(v_{i_1}, v_{i_2}, \ldots, v_{i_r}) \end{aligned}$$
Summanden mit $i_j = i_k$ für $j \neq k$ ergeben $F(v_{i_1}, v_{i_2}, \ldots, v_{i_r}) = 0$. Nach Streichen aller solcher Summanden bleiben die Indextupel $(i_1, i_2, \ldots, i_r)$ übrig, für die $\{1, 2, \ldots, r\} \to \{1, 2, \ldots, r\}$, $j \mapsto i_j$ injektiv (und somit surjektiv) ist. In anderen Worten: durch $\sigma(j) := i_j$ ist eine Permutation $\sigma \in \mathcal{S}_r$ definiert. Wir erhalten somit
$$F(w_1, w_2, \ldots, w_r) = \sum_{\sigma \in \mathcal{S}_r} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(r)r} F(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(r)}).$$

Da F alternierend ist, haben wir

$$F(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(r)}) = (\operatorname{sgn}\sigma)\, F(v_1, \ldots, v_r).$$

Insgesamt also

$$F(w_1, w_2, \ldots, w_r) = \underbrace{\sum_{\sigma \in \mathcal{S}_r} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(r)r}\, (\operatorname{sgn}\sigma)}_{=\det_r A}\; F(v_1, \ldots, v_r).$$

□

Korollar 4.6. *Für $A, B \in \operatorname{Mat}(n, n; \mathbb{K})$ gilt*

$$\det{}_n(BA) = (\det{}_n B)\,(\det{}_n A). \tag{4.7}$$

Beweis. Wir zerlegen B und BA in Spalten, die wir v_i und w_i nennen:

$$B = (v_1 \quad v_2 \quad \cdots \quad v_n) \qquad \text{und} \qquad BA = (w_1 \quad w_2 \quad \cdots \quad w_n).$$

Wir können nun die Proposition für $r := n$, $V := \mathbb{K}^n$, $F := \det_n \in \mathcal{A}^n(V)$ anwenden und erhalten (4.7). □

Korollar 4.8. *Eine Matrix $A \in \operatorname{Mat}(n, n; \mathbb{K})$ ist genau dann invertierbar, wenn $\det_n A \neq 0$. In diesem Fall gilt dann $\det(A^{-1}) = (\det A)^{-1}$.*

Beweis. Falls A ein Inverses B besitzt, dann gilt

$$1 = \det{}_n \mathbb{1}_n = \det{}_n(BA) = (\det{}_n B)(\det{}_n A),$$

also $\det_n A \neq 0$. Falls A nicht invertierbar ist, dann ist $\operatorname{Rang} A < n$, die Spalten von A sind also linear abhängig, und deswegen $\det_n A = 0$. □

Korollar 4.9. *Für einen endlich-dimensionalen Vektorraum V gilt*

$$\dim \mathcal{A}^{\dim V}(V) = 1.$$

Beweis. $n := \dim V$. Sei $(b_1, \ldots, b_n)$ eine Basis von V. Dann ist

$$\Delta(v_1, \ldots, v_n) := \det{}_n\left((\mathcal{V}_{(b_i)})^{-1}(v_1), (\mathcal{V}_{(b_i)})^{-1}(v_2), \ldots, (\mathcal{V}_{(b_i)})^{-1}(v_n)\right)$$

eine Determinantenform auf V.

DIAGRAMM

$$\Delta(b_1, \ldots, b_n) = \det{}_n(e_1, \ldots, e_n) = 1,$$

also $\Delta \neq 0$. Sei nun $F \in \mathcal{A}^n(V)$. Es gelte nun

$$(b_1, b_2, \ldots, b_n)A = (v_1, \ldots, v_n)$$

für $A \in \mathrm{Mat}(n,n;\mathbb{K})$, das heißt die j-te Spalte von A ist $(\mathcal{V}_{(b_i)})^{-1}(v_j)$. Also $\Delta(v_1,\dots,v_n) = \det_r A$. Andererseits folgt aus der vorangehenden Proposition

$$F(v_1,\dots,v_n) = F(b_1,\dots,b_n)\det_n A = F(b_1,\dots,b_n)\,\Delta(v_1,\dots,v_n),$$

also

$$F = \lambda\Delta$$

mit $\lambda := F(b_1,\dots,b_n) \in \mathbb{K}$. Also wird $A^n(V)$ von Δ aufgespannt. □

Übungsaufgabe 4.10. Zeigen Sie, dass die Konstruktion der obigen Determinantenform $\Delta \in \mathcal{A}^n(V)$ von der Wahl der Basis abhängt. Geben Sie eine Formel an, die die analog definierte Determinantenform $\widetilde{\Delta}$ zur Basis $(\tilde{b}_1,\dots,\tilde{b}_n)$ mit Hilfe von Δ ausdrückt.

Korollar 4.11. *Sei $n = \dim V$, $F \in \mathcal{A}^n(V)$, $F \neq 0$ und $w_1,\dots,w_n \in V$. Dann gilt*

$$F(w_1,\dots,w_n) \neq 0$$

genau dann, wenn $(w_1,w_2,\dots,w_n)$ linear unabhängig ist.

Beweis. Wir haben bereits gesehen, dass $F(w_1,\dots,w_n) = 0$ falls die Vektoren linear abhängig sind. Falls $(w_1,\dots,w_n)$ linear unabhängig sind, so ist diese Familie eine Basis. Analog zu oben definieren wir $\Delta \in \mathcal{A}^n(V)$ durch

$$\Delta(v_1,\dots,v_n) := \det_r\left((\mathcal{V}_{(w_i)})^{-1}(v_1),(\mathcal{V}_{(w_i)})^{-1}(v_2),\dots,(\mathcal{V}_{(w_i)})^{-1}(v_n)\right).$$

Wir erhalten wie oben

$$F = F(w_1,\dots,w_n)\,\Delta.$$

Gilt $F(w_1,\dots,w_n) = 0$, so folgt also $F = 0$. □

Korollar 4.12. *Die einzige n-Form $F \in \mathcal{A}^n(\mathbb{K}^n)$ mit $F(e_1,\dots,e_n) = 1$ ist die Determinante.*

Beweis. Wir wissen bereits, dass die Determinante diese Eigenschaften erfüllt. Sei nun $F \in \mathcal{A}^n(\mathbb{K}^n)$ mit $F(e_1,\dots,e_n) = 1$. Dann gibt es ein $\lambda \in \mathbb{K}$ mit $F = \lambda \det_n$ und

$$1 = F(e_1,\dots,e_n) = \lambda\det_n(e_1,\dots,e_n) = \lambda 1 = \lambda.$$

Also $F = \det_n$. □

Proposition 4.13. *Für alle quadratischen Matrizen A gilt*

$$\det A = \det A^T.$$

Beweis. Sei $A = (a_{ij})_{ij}$. Dann gilt

$$\det A = \sum_{\sigma\in\mathcal{S}_n} (\operatorname{sgn}\sigma) a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$$

und

$$\begin{aligned}\det A^T &= \sum_{\sigma\in\mathcal{S}_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \sum_{\tau\in\mathcal{S}_n} (\operatorname{sgn}\sigma) a_{1\tau^{-1}(1)} a_{2\tau^{-1}(2)} \cdots a_{n\tau^{-1}(n)} \\ &= \sum_{\tau\in\mathcal{S}_n} (\operatorname{sgn}\sigma) a_{\tau(1)\tau\circ\tau^{-1}(1)} a_{\tau(2)\tau\circ\tau^{-1}(2)} \cdots a_{\tau(n)\tau\circ\tau^{-1}(n)} \\ &= \det A\end{aligned}$$

Hierbei ergibt sich die erste Gleichheit, indem wir die Definition der Determinante auf die transponierte Matrix anwenden. Die zweite Gleichheit erhalten wir, indem wir $\tau := \sigma^{-1}$ setzen und nutzen, dass $\sigma \mapsto \sigma^{-1}$ eine Bijektion $\mathcal{S}_n \to \mathcal{S}_n$ ist. Um die dritte Gleichheit zu erhalten permutieren wir die Faktoren des Produkts gemäß der Permutation τ. Wenn wir nun τ durch σ substituieren, erhalten wir die Definition von $\det A$. □

5. Determinanten von Endomorphismen

Definition 5.1. Sei V ein endlich-dimensionaler Vektorraum, $f : V \to V$ ein Endomorphismus. Sei $\mathcal{B} := (b_1, \ldots, b_n)$ eine Basis von V. Dann definieren wir die Determinante von f als

$$\det^{\mathcal{B}}(f) := \det_n\Big(\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f)\Big).$$

Lemma 5.2. *Die Determinante hängt nicht von der Wahl von $\mathcal{B}$ ab.*

Beweis. Sei $\mathcal{B}' = \mathcal{B} \cdot A$ eine andere Basis. Dann gilt

$$\mathrm{Mat}^{\mathcal{B}'}_{\mathcal{B}'}(f) = A^{-1}\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f)A.$$

DIAGRAMM

Somit

$$\begin{aligned}\det^{\mathcal{B}'}(f) &= \det\mathrm{Mat}^{\mathcal{B}'}_{\mathcal{B}'}(f) = \det\left(A^{-1}\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f)A\right) \\ &= (\det A^{-1})(\det\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f))(\det A) = (\det A)^{-1}(\det\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f))(\det A) \\ &= \det\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f) = \det^{\mathcal{B}}(f)\end{aligned}$$

□

Wir schreiben nun einfach $\det(f)$.

Aus den uns bekannten Eigenschaften der Determinante von Matrizen erhalten wir sofort die folgenden Eigenschaften:

(1) $\det\mathrm{Id} = 1$.

(2) $\det(\alpha \mathrm{Id}_V) = \alpha^n$ für $\mathrm{Id}_V : V \to V$, $n := \dim V$.
(3) $\det(f \circ g) = (\det f)(\det g)$ für $f, g \in \mathrm{End}(V)$.
(4) $\det f \neq 0 \Leftrightarrow f$ invertierbar. Falls f invertierbar ist, dann gilt $\det(f^{-1}) = (\det f)^{-1}$.
(5) $\det f' = \det f$.

6. Berechnung von Determinanten und Cramersche Regel

6.1. Diagonalmatrizen.

$$\det \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \lambda_1 \det \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \lambda_1 \lambda_2 \cdots \lambda_n \det \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Insbesondere gilt für Matrizen vom Typ (1) in Kap. 5 Abschn. 6:

$$\det B^{(1)}_{j\lambda} = \lambda.$$

Elementare Zeilen- oder Spaltenumformungen vom Typ (1) multiplizieren die Determinante mit λ.

6.2. Typ (2) Umformungen und Typ (2) Matrizen.

$$\det \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = \det \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda \\ 1 \end{pmatrix} \right) = \lambda \underbrace{\det \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)}_{0} + \underbrace{\det \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)}_{1} = 1$$

!!!!!!!

$$\det B^{(2)}_{jk\lambda} = ... = 1$$

Elementare Zeilen- oder Spaltenumformungen vom Typ (2) erhalten die Determinante..

6.3. Typ (3) Umformungen und Typ (3) Matrizen.

$B^{(3)}_{jk}$ entsteht aus $\mathbb{1}_n$ durch Vertauschen der j-ten und k-ten Zeile. also folgt $\det B^{(3)}_{jk} = -1$. Elementare Zeilen- oder Spaltenumformungen vom Typ (3) wechseln das Verzeichen der Determinante..
!!!!!!!!!!!!!!!!!

6.4. Berechnung mit Gaußschem Verfahren.

Formulierung 1: Wir formen eine Metrik unter elementaren Zeilen- (bzw. Spalten-)Umformungen in eine Matrix um, deren Determinante bereits bekannt ist, z. B. durch elementare Zeilenumformungen in die Matrix $\mathbb{1}_n$ oder in eine Matrix mit mindestens einer Zeile aus Nullen. Da wir wissen, wie sich die Determinante unter elementaren Zeilen- und Spaltenumformungen ändert, können wir daraus die Determinante der urspünglich gegebenen Matrix berechnen.

Formulierung 2: Bestimme zu der gegebenen Matrix A Matrizen B_j vom Typ (1) bis (3) wie in Kap. 5 Abschn. 6, so dass

$$Z := B_k B_{k-1} \cdots B_1 A$$

Zeilenstufenform hat. Dann folgt also

$$\det A = \frac{\det Z}{(\det B_1)(\det B_2)\cdots(\det B_k)}$$

6.5. Blockmatrizen.

Proposition 6.1. *Sei $A \in \mathrm{Mat}(n,n;\mathbb{K})$, $B \in \mathrm{Mat}(n,m;\mathbb{K})$ und $C \in \mathrm{Mat}(m,m;\mathbb{K})$. Dann gilt*

$$\det_{n+m}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = (\det_n A)(\det_m C).$$

Beweis. $A = (a_{ij})_{ij}$, $C = (c_{ij})$, $D := \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = (d_{ij})$.

$$\det_{n+m} D = \sum_{\sigma \in \mathcal{S}_{n+m}} d_{\sigma(1)1} d_{\sigma(2)2} \cdots d_{\sigma(n+m)n+m}$$

Es gilt $d_{ij} = 0$ falls $i > n$ und $j \leq n$. Falls $\sigma(j) > n$ für ein $j \leq n$, dann verschwindet der Summand zu diesem σ. Wir betrachten deswegen nur noch die $\sigma \in \mathcal{S}_{n+m}$ mit $\sigma(\{1,\ldots,n\}) \subseteq \{1,\ldots,n\}$. Da σ bijektiv ist, folgt $\sigma(\{1,\ldots,n\}) = \{1,\ldots,n\}$ und $\sigma(\{n+1,\ldots,n+m\}) = \{n+1,\ldots,n+m\}$. Also $\sigma|_{\{1,\ldots,n\}} \in \mathcal{S}_n$ und $\sigma_{\{n+1,\ldots,n+m\}}$ ist eine Permutation von $\{n+1,\ldots,n+m\}$, also bis auf Verschieben der Indizes ein Element von $\mathcal{S}_m$. Wir erhalten

$$\begin{aligned} \det_{n+m} D &= \sum_{\sigma \in \mathcal{S}_n} \sum_{\tau \in \mathcal{S}_m} d_{\sigma(1)1} d_{\sigma(2)2} \cdots d_{\sigma(n)n} d_{n+\tau(1),n+1} d_{n+\tau(2),n+2} \cdots d_{n+\tau(m),n+m} \\ &= \underbrace{\sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}}_{=\det_n A} \underbrace{\sum_{\tau \in \mathcal{S}_m} c_{\tau(1)1} \cdots c_{\tau(m)m}}_{=\det_m C} \end{aligned}$$

□

6.6. Laplacescher Entwicklungssatz und Cramersche Regel.

Sei $A = (a_{ij})_{ij} = (a_1 \quad a_2 \quad \cdots \quad a_n) \in \mathrm{Mat}(n,n;\mathbb{K})$, $n \geq 2$. Wir definieren die Matrizen $A_{ij} \in \mathrm{Mat}(n,n;\mathbb{K})$ und $\hat{A}_{ij} \in \mathrm{Mat}(n-1,n-1;\mathbb{K})$ wie folgt....!!!!!!

Lemma 6.2.

$$\det A_{ij} = \det(a_1,\ldots,a_{j-1},e_i,a_{j+1},\ldots,a_n)$$

Beweis. !!! □

Lemma 6.3.

$$\det_n A_{ij} = (-1)^{i+j}\det_{n-1}\hat{A}_{ij}.$$

Beweis. !!! □

Da $a_k = \sum_{i=1}^n a_{ik}e_i$ rechnen wir nach

$$\begin{aligned}
\sum_{i=1}^n a_{ik}\det A_{ij} &= a_{ik}\det(a_1,\dots,a_{j-1},e_i,a_{j+1},\dots,a_n)\\
&= \det(a_1,\dots,a_{j-1},\sum_{i=1}^n a_{ik}e_i,a_{j+1},\dots,a_n)\\
&= \begin{cases}\det A, & \text{falls } k=j\\ 0, & \text{falls } k\neq j\end{cases}\\
&= \delta_{jk}\det A
\end{aligned}$$

Dies ergibt im Fall $j = k$:

Korollar 6.4 (Laplacesche Entwicklung nach der j-ten Spalte)**.**

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij}\det\hat{A}_{ij}$$

Beispiel. !!!!!!!!

Korollar 6.5 (Laplacesche Entwicklung nach der j-ten Zeile)**.**

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ji}\det\hat{A}_{ji}$$

Beweis. Wenn wir im vorangehenden Korollar A durch A^T ersetzen, so haben wir die Formel

$$\det A^T = \sum_{i=1}^n (-1)^{i+j} a_{ji}\det(\hat{A}^T)_{ij}.$$

Es gilt nun $(\hat{A}^T)_{ij} = (\hat{A}_{ji})^T$ und deswegen $\det(\hat{A}^T)_{ij} = \det(\hat{A}_{ji})$. Mit $\det A^T = \det A$ erhalten wir

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ji}\det(\hat{A}_{ji}).$$

Beispiel. !!!!!!!!

Sei nun
$$\alpha_{ij} := \det A_{ij} = (-1)^{i+j}\det\hat{A}_{ij}$$
und $A^{\mathrm{ad}} := (\alpha_{ji})_{ij} \in \mathrm{Mat}(n,n;\mathbb{K})$.

Satz 6.6. *Cramersche Regel Es gilt* $A^{\mathrm{ad}} \cdot A = (\det A)\,\mathbb{1}_n$ *und* $A \cdot A^{\mathrm{ad}} = (\det A)\,\mathbb{1}_n$.

Beweis. Der Koeffizient der Matrix $A^{\mathrm{ad}} \cdot A$ in der i-ten Zeile und k-ten Spalte ist
$$\sum_{j=1}^{n} \alpha_{ji} a_{jk} = \sum_{j=1}^{n} (\det A_{ji}) a_{jk} = \delta_{ik}\det A$$
und dies ist der Koeffizient der Matrix $(\det A)\,\mathbb{1}_n$ in der i-ten Zeile und k-ten Spalte. Die erste Gleichung folgt.

Wir studieren nun, wie sich unsere Ausdrücke unter Transponieren ändern: $(A^T)_{ij} = (A_{ji})^T$, also $\det(A^T)_{ij} = \det(A_{ji})$. Somit $(A^T)^{\mathrm{ad}} = (A^{\mathrm{ad}})^T$. Wir ersetzen nun in der ersten Gleichung A durch A^T und erhalten
$$(A^T)^{\mathrm{ad}} \cdot A^T = (\det A^T)\,\mathbb{1}_n.$$
Die rechte Seite ist offensichtlich gleich $(\det A)\,\mathbb{1}_n$, die linke formt sich in $(A^{\mathrm{ad}})^T \cdot A^T = (A \cdot A^{\mathrm{ad}})^T$ um. Wir erhalten
$$(A \cdot A^{\mathrm{ad}})^T = (\det A)\,\mathbb{1}_n$$
und durch Transponieren und Nutzung von $\mathbb{1}_n^T = \mathbb{1}_n$ erhalten wir die zweite Gleichung. □

Anwendung: Ist $\det A \neq 0$, so sehen wir:
$$A^{-1} := \frac{1}{\det A} A^{\mathrm{ad}}.$$
Die Berechnung einer Inversen mit dieser Formel ist allerdings für Berechnungen zumeist aufwändiger als mit dem Gaußschen Verfahren.

Beispiel. $n = 2$: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\hat{A}_{11} = (d)$, $\alpha_{11} = d$, $\hat{A}_{12} = (c)$, $\alpha_{12} = -c$, $\hat{A}_{21} = (b)$, $\alpha_{21} = -b$, $\hat{A}_{22} = (a)$, $\alpha_{22} = a$.
$$A^{\mathrm{ad}} := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$
.

Anwendung: Lösung eines linearen Gleichungssystems.

Seien $A \in \mathrm{GL}(n,\mathbb{K})$ und $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{K}^n$ gegeben. Wir suchen $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ mit
$$Ax = y.$$

Wir erhalten für die eindeutige Lösung

$$x = A^{-1}y = \frac{1}{\det A} A^{\mathrm{ad}} y$$

und somit

$$\begin{aligned} x_i \det A &= \sum_{j=1}^{n} \alpha_{ji} y_j = \sum_{j=1}^{n} (\det A_{ji}) y_j \\ &= \sum_{j=1}^{n} \det(a_1, a_2, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n)\, y_j \\ &= \det(a_1, \ldots, a_{i-1}, y, a_{i+1}, \ldots, a_n). \end{aligned}$$

Es folgt

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

Bemerkung 6.7. Wenn wir $\mathbb{K}$ durch den kommutativen Ring R mit 1 ersetzen, dann kann man ebenfalls die Determinante von $A = (a_{ij})$ durch die Formel

$$\det A := \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

definieren. Manche der zuvor gemachten Aussagen und Beweise bleiben richtig, u.a. die Cramersche Regel mit Beweis. Eine Matrix $A \in \mathrm{Mat}(n, n; R)$ ist also genau dann invertierbar, falls $\det A$ invertierbar in R ist.

Zwei Fälle sind für uns wichtig:
(1) R ist der Ring aller $\mathbb{K}$-Polynome (später) und
(2) $R = \mathbb{Z}$. Im Fall $R = \mathbb{Z}$ sehen wir also, dass eine Matrix $A \in \mathrm{Mat}(n, n; \mathbb{Z})$ genau dann in $\mathrm{Mat}(n, n; \mathbb{Z})$ invertierbar ist, wenn $\det A = \pm 1$.

KAPITEL 7

Eigenwerte und Eigenvektoren

1. Definition

Definition 1.1. Sei V ein $\mathbb{K}$-Vektorraum und $F : V \to V$ ein $\mathbb{K}$-Endomorphismus. Ein Skalar $\lambda \in \mathbb{K}$ heißt *Eigenwert von F*, wenn es ein $v \in V \setminus \{0\}$ gibt, so dass

$$F(v) = \lambda v.$$

Der Vektor v heißt ein *zum Eigenwert λ gehöriger Eigenvektor von F*. Der Vektorraum $E_\lambda :=$ Kern $F - \lambda \mathrm{Id}_V$ heißt der *zu λ gehörige Eigenvektorraum.* Die Dimension on E_λ heißt *Multiplizität* des Eigenwerts λ. Die Menge aller Eigenwerte von F nennt man das *(Punkt-)Spektrum von F*.

Für $A \in \mathrm{Mat}(n, n; \mathbb{K})$ nennt man die Eigenwerte, Eigenvektoren und Eigenvektorräume von $\mathcal{L}_A$ auch die *Eigenwerte, Eigenvektoren und Eigenvektorräume von A*, ebenso ist das *Spektrum von A* definiert als das Spektrum von $\mathcal{L}_A$.

Beispiel. Sei $A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ eine Diagonalmatrix. Dann ist e_i Eigenvektor von A zum Eigenwert λ_i. Der zu $\lambda \in \mathbb{K}$ gehörige Eigenvektorraum von A ist

$$E_\lambda = \text{Kern } (A - \lambda \mathbb{1}_n) = \begin{pmatrix} \lambda_1 - \lambda & 0 & \cdots & 0 \\ 0 & \lambda_2 - \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n - \lambda \end{pmatrix}.$$

Man sieht leicht, dass Kern $(A - \lambda \mathbb{1}_n) = \{0\}$ für $\lambda \in \mathbb{K} \setminus \{\lambda_1, \ldots, \lambda_n\}$. Es gibt also nur die Eigenwerte λ_1, …, λ_n. Alle Vektoren in $E := \bigcup_{i=1}^n \operatorname{span} e_i \setminus \{0\}$ sind Eigenvektoren, und wenn die $\lambda_1, \ldots, \lambda_n$ alle verschieden sind, dann ist jeder Eigenvektor in E. Falls mindestens zwei Eigenwerte übereinstimmen, gibt es weitere Eigenvektoren. [1] Sei zum Beispiel $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. Ein Vektor

[1] Ein Vektor $v \neq 0$ ist genau dann ein Eigenvektor zum Eigenvektor λ, wenn es eine Familie $(\mu_i | i \in I)$ von Skalaren mit Indexmenge $I := \{i \in \{1, \ldots, n\} \,|\, \lambda_i = \lambda\}$ gibt, so dass $v = \sum_{i \in I} \mu_i e_i$.

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ ist Eigenvektor genau dann, wenn entweder $((x,y) \neq (0,0)$ und $z = 0)$ oder $((x,y) = (0,0)$ und $z \neq 0)$.

2. Motivation, Beispiele und Anwendungen

2.1. Ellipse. Auf $V = \mathbb{R}^2$ betrachten wir die Funktion $f : V \to \mathbb{R}$, $v = \begin{pmatrix} x \\ y \end{pmatrix} \mapsto ax^2 + 2bxy + cy^2$. Was wissen wir über die Menge $N = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \,\middle|\, f \begin{pmatrix} x \\ y \end{pmatrix} = 1 \right\}$. Für $a = c > 0$, $b = 0$ ist es ein Kreis von Radius $1/\sqrt{a}$ um 0. Für $a > 0$, $c > 0$, $b = 0$ eine Ellipse mit Symmetrie-Geraden span $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und span $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Die Halbachsenlängen sind $1/\sqrt{a}$ und $1/\sqrt{c}$.

SKIZZE

Was geschieht für $a > 0$, $c > 0$, b beliebig?

Lösungsmethode: Setze $A := \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Dann ist $f(v) = v^T A v$.

Wir werden unter anderem sehen:

(1) A besitzt einen oder zwei Eigenwerte $\lambda_1, \lambda_2 \in \mathbb{R}$ (evtl. $\lambda_1 = \lambda_2$) Sei $Av_1 = \lambda_1 v_1$ und $Av_2 = \lambda_2 v_2$, $v_1, v_2 \in V$.
(2) N ist eine Ellipse mit Mittelpunkt 0, Symmetrieachsen span v_1 und span v_2 und Halbachsenlängen $1/\sqrt{\lambda_1}$ und $1/\sqrt{\lambda_2}$.

SKIZZE

2.2. Trägheitstensor der Mechanik. $A \in \mathrm{Mat}(3,3;\mathbb{R})$ siehe Saalübung vom 7.1.2008

2.3. Schwingende Saite. ZEICHNUNG

$$V = \{f : [0,L] \to \mathbb{R} \text{ unendlich oft differenzierbar}\}$$

Eine reine Eigenschwingung der Saite entspricht einer Lösung der Gleichung

$$-c^2 f''(x) = \lambda f(x), \qquad f(0) = f(L) = 0.$$

Die Konstante c hängt nur von den physikalischen Parametern Saite ab. $\sqrt{\lambda}/2\pi$ ist die Frequenz der Schwingung (=Tonhöhe)

Wir definieren $F \in \mathrm{End}(V)$ als $F(f) := -cf''$ und such nun nach Eigenvektoren von F. Beispiele $f(x) = \sin(k\pi x/L)$, $\lambda = k^2\pi^2c^2/L^2$.

2.4. Quantenmechanik. Wasserstoffartiges Atom

$U : \mathbb{R}^3 \to \mathbb{R}$ Energiepotential

$$V = \{f : \mathbb{R}^3 \to \mathbb{C} \mid \text{ unendlich oft differenzierbar und } \int_{\mathbb{R}^3} |f(x)|^2\, dx^3 < \infty \quad + \text{Zusatzbedingungen}\}$$

Unendlich-dimensionaler komplexer Vektorraum.

$F : V \to V$, $F(f) = -\left(\frac{\partial^2 f}{\partial_x^2} + \frac{\partial^2 f}{\partial_y^2} + \frac{\partial^2 f}{\partial_z^2}\right) + Uf$

Eigenwerte von F entsprechen möglichen Energien von gebundenen Zuständen eines Elektrons. Dies führt zum Orbitalmodell der Chemie.

3. Grundlegende Eigenschaften

Proposition 3.1. *Sei $f : V \to V$ ein Endomorphismus und $g : W \to V$ ein Isomorphismus. Der Vektor $v \in V$ ist genau dann ein Eigenvektor von f, wenn $g^{-1}(v)$ ein Eigenvektor von $\hat{f} := g^{-1} \circ f \circ g$ zum selben Eigenwert ist. Insbesondere ist λ genau dann Eigenwert von f, wenn λ Eigenwert von $\hat{f}$ ist, und die Multiplizitäten stimmen überein.*

DIAGRAMM

Beweis. v ist Eigenvektor von f zum Eigenwert $\lambda \Leftrightarrow f(v) = \lambda v \Leftrightarrow g^{-1} \circ f(v) = g^{-1}(\lambda v) \Leftrightarrow (g^{-1} \circ f \circ g)(g^{-1}(v)) = \lambda g^{-1}(v) \Leftrightarrow g^{-1}(v)$ ist Eigenvektor von $\hat{f}$ zum Eigenwert λ. □

Sei V nun ein endlich-dimensionaler Vektorraum mit Basis $(b_1, b_2, \ldots, b_n)$, $W = \mathbb{K}^n$ und $g = \mathcal{V}_{(b_i)}$. Dann ist $\hat{f} = \mathcal{L}_{\mathrm{Mat}^{(b_i)}_{(b_i)}(f)}$.

DIAGRAMM

Wir erhalten also das Korollar:

Korollar 3.2. *v ist Eigenvektor von f genau dann, wenn $(\mathcal{V}_{(b_i)})^{-1}(v)$ Eigenvektor von $\mathrm{Mat}^{(b_i)}_{(b_i)}(f)$ zum selben Eigenwert ist.*

□

Sei nun weiterhin $V = \mathbb{K}^n$ mit beliebiger Basis $(b_1, \ldots, b_n) = (e_1, \ldots, e_n)A$. Dann ist $\mathcal{V}_{(b_i)} = \mathcal{L}_A$. Wir erhalten somit.

Korollar 3.3. *Sei $A \in \mathrm{GL}(n, \mathbb{K})$ und $C \in \mathrm{Mat}(n, n; \mathbb{K})$. Der Vektor $Av \in \mathbb{K}^n$ ist genau dann ein Eigenvektor von C, wenn v ein Eigenvektor von $\hat{C} := A^{-1} \cdot C \cdot A$ zum selben Eigenwert ist.*

□

DIAGRAMM

Definition 3.4. Zwei Matrizen $C, \tilde{C} \in \mathrm{Mat}(n, n; \mathbb{K}^n)$ heißen *ähnlich*, falls es ein $A \in \mathrm{GL}(n, \mathbb{K})$ gibt mit $\tilde{C} = A^{-1}CA$.

Ähnlich zu sein, ist eine Äquivalenzrelation.

Angenommen (b_i) ist eine Basis von V, dann gilt für $f \in \mathrm{End}(V)$:

$$\mathrm{Mat}^{(b_i)A}_{(b_i)A}(f) = A^{-1} \,\mathrm{Mat}^{(b_i)}_{(b_i)}(f)\, A.$$

Also sind C und $\tilde{C}$ genau dann ähnlich, falls sie denselben Endomorphismus, aber evtl. in einer anderen Basis beschreiben.

Definition 3.5. Eine Matrix $C = (c_{ij})$ heißt *diagonal*, falls $c_{ij} = 0$ für $i \neq j$. Sie heißt *trigonal*, falls $c_{ij} = 0$ für $i > j$. Ein Endomorphismus $F \in \mathrm{End}(V)$ heißt *diagonalisierbar (trigonalisierbar)*, falls es eine Basis (b_i) von V gibt, so dass $\mathrm{Mat}^{(b_i)}_{(b_i)}(f)$ diagonal (trigonal) ist. Eine Matrix $C \in \mathrm{Mat}(n, n; \mathbb{K})$ nennen wir *diagonalisierbar (trigonalisierbar)*, falls $\mathcal{L}_C$ diagonalisierbar (trigonalisierbar) ist.

Achtung: „diagonalisierbar“ ist nicht dasselbe wie „diagonal“.

Jede diagonale Matrix ist natürlich auch trigonal, und jede diagonalisierbare Matrix ist auch trigonalisierbar.

Beispiel. Die Matrix $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ ist nicht diagonal, aber sie ist diagonalisierbar. Denn

$$\begin{aligned} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} &= 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

Wenn wir also $\mathcal{L}_C$ in der Basis $b_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $b_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ausdrücken, dann folgt aus $\mathcal{L}(b_1) = 4b_1$ und $\mathcal{L}(b_2) = 2b_2$, also

$$\mathrm{Mat}^{(b_1,b_2)}_{(b_1,b_2)}(\mathcal{L}_C) = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

LEMMA 3.6. *Eine Matrix $C \in \mathrm{Mat}(n, n; \mathbb{K})$ ist diagonalisierbar (trigonalisierbar) genau dann, wenn C zu einer diagonalen (trigonalen) Matrix ähnlich ist.*

Beweis. Wenn C diagonalisierbar ist, dann gibt es eine Basis $(b_1, b_2, \dots, b_n) = (e_1, e_2, \dots, e_n)A$ von $\mathbb{K}^n$, so dass $D := \mathrm{Mat}^{(b_1,b_2)}_{(b_1,b_2)}(\mathcal{L}_C)$ diagonal ist. Aus dem Verhalten der Matrix einer linearen Abbildung unter Basistransformation folgt $D = A^{-1}CA$. Somit C ist ähnlich zur Diagonalmatrix D.

Ist umgekehrt C ähnlich zur diagonalen Matrix D, sagen $C = ADA^{-1}$, dann berechnet sich die Matrix von $\mathcal{L}_C$ in der Basis $(b_1, b_2, \dots, b_n) := (e_1, e_2, \dots, e_n)A$ wie folgt:

$$\mathrm{Mat}^{(b_1,b_2)}_{(b_1,b_2)}(\mathcal{L}_C) := A^{-1}CA = D.$$

Also ist C diagonalisierbar.

Der Beweis, dass „trigonalisierbar“ dasselbe ist wie „ähnlich zu einer trigonalen Matrix“, geht völlig analog. □

SATZ 3.7. *Seien $\lambda_1, \lambda_2, \dots, \lambda_k$ verschiedene Eigenwerte von $f \in End(V)$ und $E_{\lambda_1}, \dots, E_{\lambda_k}$ die zugehörigen Eigenvektorräumen. Dann ist die Summe*

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$$

eine direkte Summe von Untervektorräumen.

Beweis. Wir zeigen den Satzt durch Induktion über k. Die Aussage ist offensichtlich für $k = 1$. Wir wollen nun die Aussage für $k-1$ annehmen und für k, $k \geq 2$ daraus folgern. Seien also k verschiedene Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_k$ gegeben. Sei $\sum_{i=1}^k v_i = 0$ mit $v_i \in E_{\lambda_i}$ gebeben, d.h. $f(v_i) = \lambda_i v_i$. Zu zeigen ist $v_i = 0$ für alle i.

Hierzu berechnen wir für

$$0 = (f - \lambda_1 \mathrm{Id}_V)(0) = (f - \lambda_1 \mathrm{Id}_V)\left(\sum_{i=1}^k v_i\right) = \sum_{i=1}^k \Big(f(v_i) - \lambda_1 v_i\Big) = \sum_{i=1}^k (\lambda_i - \lambda_1) v_i = \sum_{i=2}^k (\lambda_i - \lambda_1) v_i.$$

Nach Induktionsvoraussetzung gilt also $(\lambda_i - \lambda_1)v_i = 0$ für alle $i \in \{2, \dots, n\}$ und somit $v_i = 0$ für solche i. Daraus folgt auch $v_1 = 0$, also das zu zeigende. □

FOLGERUNG 3.8. *Ist $n = \dim V < \infty$, dann gibt es höchstens n verschiedene Eigenwerte. Es gilt sogar: die Summe der Multiplizitäten ist höchstens n.*

Die Summe der Multiplizitäten kann auch kleiner als n sein: Die Matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ besitzt nur den Eigenwert 0, und dieser Eigenwert hat die Multiplizität $1 < 2$. Die Matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ hat (über dem Körper $\mathbb{K} = \mathbb{R}$) keine Eigenwerte.

Satz 3.9. *Sei V endlich-dimensional und $f \in End(V)$. Wir schreiben das Spektrum von f als $\{\lambda_1, \dots, \lambda_k\}$, wobei alle λ_i verschieden seien. Dann sind äquivalent:*

(a) *f ist diagonalisierbar*
(b) $V = E_{\lambda_1} + \cdots + E_{\lambda_k}$
(c) $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$

Beweis.
„(a)⇒(b)“ Falls f diagonalisierbar ist, dann gibt es eine Basis $(b_1, \dots, b_n)$ von V mit $f(b_j) = \mu_j b_j$. Sei $J_i := \{j \in \{1, 2, \dots, n\} \mid \mu_j = \lambda_i\}$, somit $E_{\lambda_i} = \langle b_j \mid j \in J_i \rangle$. Es folgt offensichtlich

$$V = \operatorname{span}\langle b_1, \dots, b_n \rangle = E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_k}.$$

„(b)⇒(c)“ folgt direkt aus dem letzten Satz
„(c)⇒(a)“ Wähle eine $(b_{1,1}, \dots, b_{m_1,1}$ von E_{λ_1}, eine Basis $(b_{1,2}, \dots, b_{m_2,2})$ von E_{λ_2}, …, und eine Basis $(b_{1,k}, \dots, b_{m_k,k})$ von E_{λ_k}. Wir vereinigen diese Vektoren zur Familie $\mathcal{B} = (b_i | i \in I)$, $I = \{(1,1), \dots, (m_k, k)\}$, die eine Basis von V ist. Es folgt

$$\operatorname{Mat}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_k \end{pmatrix}$$

Die Charakteristische Polynom-Funktion. Für einen Endomorphismus $f \in \operatorname{End}(V)$, $\dim V < \infty$, definieren wir die *charakteristische Polynom-Funktion von f als $\chi_f : \mathbb{K} \to \mathbb{K}$*, $\chi_f(\lambda) := \det(\lambda \mathrm{Id}_V - f)$. Analog definieren wir für eine Matrix $A \in \operatorname{Mat}(n, n; \mathbb{K})$ die *charakteristische Polynom-Funktion von A als $\chi_A : \mathbb{K} \to \mathbb{K}$*, $\chi_A(\lambda) = \det_n(\lambda \mathbb{1}_n - A)$.

Proposition 3.10. *$\chi_f(\lambda) = 0$ genau dann, wenn λ ein Eigenwert von f ist. Analoges gilt für eine Matrix A.*

Beweis. λ Eigenwert von f gdw $0 \neq E_\lambda = \text{Kern } (f - \lambda \mathrm{Id}_V) = \text{Kern } (\lambda \mathrm{Id}_V - f)$ gdw $\lambda \mathrm{Id}_V - f$ nicht invertierbar gdw $\det(\lambda \mathrm{Id}_V - f) = 0$. □

KAPITEL 8

Euklidische und unitäre Vektorräume

1. Bilinear-Formen

Definition 1.1. Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $F : V \times V \to \mathbb{K}$ heißt *Bilinearform auf V (über $\mathbb{K}$)*, falls für alle $w \in V$ die Abbildungen $V \to \mathbb{K}$, $v \mapsto F(v,w)$ und $V \to \mathbb{K}$, $v \mapsto F(w,v)$ linear (über $\mathbb{K}$) sind.

Den Raum aller Bilinearformen auf V bezeichnen wir mit $\mathrm{Bilin}(V)$. Eine Bilinearform $F \in \mathrm{Bilin}(V)$ heißt symmetrisch, falls $F(v,w) = F(w,v)$ für alle v, w in V. Den Raum aller symmetrischen Bilinearformen auf V bezeichnen wir mit $\mathrm{Sym}^2(V)$.

Mit der Addition von $\mathbb{K}$-wertigen Funktionen und mit der Multiplikation mit Skalaren von Funktionen ist $\mathrm{Bilin}(V)$ ein $\mathbb{K}$-Vektorraum. Die Menge $\mathrm{Sym}^2(V)$ und $\mathcal{A}^2(V)$ sind Untervektorräume von $\mathrm{Bilin}(V)$.

Beispiele. (1) Die Abbildung $V \times V \to \mathbb{K}$, $(v,w) \mapsto 0$ ist eine Bilinearform.

(2) $\mathbb{K} = \mathbb{R}$. Auf $\mathbb{R}^n$ definieren wir

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle_n := \sum_{i=1}^{n} x_i y_i.$$

Die Abbildung $\langle\,\cdot\,,\,\cdot\,\rangle$ ist eine symmetrische Bilinearform. Es heißt das *kanonische Skalarprodut auf* $\mathbb{R}^n$.

(3) $\mathbb{K} = \mathbb{R}$. Auf $\mathbb{R}^4$ definieren wir

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right\rangle_{3,1} := x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4.$$

Die Abbildung $\langle\,\cdot\,,\,\cdot\,\rangle_{3,1}$ ist eine symmetrische Bilinearform. Die Raumzeit der Speziellen Relativitätstheorie ist $\mathbb{R}^4$ versehen mit der symmetrischen Bilinearform $\langle\,\cdot\,,\,\cdot\,\rangle_{3,1}$.

(4) Seien $a, b \in V'$. Dann definieren wir $a \otimes b \in \mathrm{Bilin}(V)$ durch:

$$(a \otimes b)(v,w) = a(v) \cdot b(w).$$

Auf $\mathbb{R}^n$ gilt also $\langle\,\cdot\,,\,\cdot\,\rangle_n = \sum_{i=1}^n e_i' \otimes e_i'$.

Beschreibung von Bilinearformen in Basen:

Sei $F \in \mathrm{Bilin}(V)$ und $(b_1,\ldots,b_n)$ eine Basis von V. Dann definieren wir die Matrix

$$\mathrm{mat}_{(b_i)}(F) = (F(b_i,b_j))_{i,j\in\{1,\ldots,n\}} \in \mathrm{Mat}(n,n;\mathbb{K}).$$

Lemma 1.2. *Sei $(b_1,\ldots,b_n)$ eine Basis von V. Die Abbildung $\mathrm{mat}_{(b_i)} : \mathrm{Bilin}(V) \to \mathrm{Mat}(n,n;\mathbb{K})$ ist ein Isomorphismus.*

Beweis. Es ist klar, dass diese Abbildung $\mathbb{K}$-linear ist. Zu zeigen ist also die Injektivität und die Surjektivität.

Sei $(b_1',\ldots,b_n')$ die duale Basis zu $(b_1,\ldots,b_n)$. Für $A = (a_{ij}) \in \mathrm{Mat}(n,n;\mathbb{K})$ definieren wir

$$\mathrm{Bilin}_{(b_i)}(A) := \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_i' \otimes b_j'.$$

Wir rechnen nach:

$$\begin{aligned}
\left(\mathrm{Bilin}_{(b_i)}(A)\right)(b_k,b_l) &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_i' \otimes b_j'(b_k,b_l)\\
&= \sum_{i=1}^n \sum_{j=1}^n a_{ij}\, \underbrace{b_i'(b_k)}_{\delta_{ik}}\,\underbrace{b_j'(b_l)}_{\delta_{jl}}\\
&= a_{kl}.
\end{aligned}$$

Also ist

$$\mathrm{mat}_{(b_i)}\left(\mathrm{Bilin}_{(b_i)}(A)\right) = (a_{kl})_{kl} = A. \tag{1.3}$$

Also ist $\mathrm{mat}_{(b_i)} : \mathrm{Bilin}(V) \to \mathrm{Mat}(n,n;\mathbb{K})$ surjektiv. Wir bestimmen den Kern dieser Abbildung: Sei also $F \in \mathrm{Kern}\,(\mathrm{mat}_{(b_i)})$ in anderen Worten F ist eine Bilinearform auf V mit $F(b_i,b_j) = 0$ für alle i,j. Sei $v = \sum_{i=1}^n v_i b_i$ und $w = \sum_{i=1}^n w_i b_i$. Dann sehen wir

$$F(v,w) = F\left(\sum_{i=1}^n v_i b_i, \sum_{j=1}^n w_j b_j\right) = \sum_{i=1}^n \sum_{j=1}^n v_i w_j\, \underbrace{F(b_i,b_j)}_{=0} = 0.$$

also ist der Kern gleich $\{0\}$, und somit die Abbildung injektiv. Aus (1.3) folgt außerdem, dass $\mathrm{Bilin}_{(b_i)}$ die Umkehrabbildung von $\mathrm{mat}_{(b_i)}$ ist. □

Man sieht nun leicht, dass eine Bilinearform genau dann symmetrisch ist, falls die zugehörige Matrix symmetrisch ist. (Eine symmetrische Matrix ist eine quadratische Matrix (a_{ij}) mit $a_{ij} = a_{ji}$, d.h. eine quadratische Matrix A mit $A^T = A$.)

ÜBUNGSAUFGABE 1.4. Die Abbildung $\mathrm{Bilin}(V) \to \mathrm{Hom}(V, V')$, $F \mapsto (v \mapsto F(v, \cdot\,) \in V')$ ist ein Isomorphismus.

Basistransformation. Seien $(b_1, \ldots, b_n)$ und $(d_1, \ldots, d_n) = (b_1, \ldots, b_n)C$ zwei Basen von V. Welche Relation besteht zwischen $\mathrm{mat}_{(b_i)}(F)$ und $\mathrm{mat}_{(d_i)}(F)$?

Wir schreiben dazu $C = (c_{ij})_{ij}$, also $d_j = \sum_{i=1}^n c_{ij} b_i$.

$$\begin{aligned} F(d_i, d_j) &= F\left(\sum_{k=1}^n c_{ki} b_k, \sum_{l=1}^n c_{lj} b_j\right) \\ &= \sum_{k=1}^n \sum_{l=1}^n c_{ki} c_{lj} F(b_k, b_l). \end{aligned}$$

Dies bedeutet

$$\mathrm{mat}_{(d_i)}(F) = C^T \left(\mathrm{mat}_{(b_i)}(F)\right) C. \tag{1.5}$$

Wieso ist diese Transformationsformel anders als die von $\mathrm{Hom}(V, V)$?

Ein $F \in \mathrm{Bilin}(V)$ kann man mit der obigen Aufgabe als $\mathrm{Hom}(V, V')$ auffassen. Dann gilt

$$\mathrm{mat}_{(b_i)}(F) = \mathrm{Mat}^{(b_i)}_{(b'_i)}(F),$$

wobei $(b'_1, \ldots, b'_n)$ die duale Basis zu $(b_1, \ldots, b_n)$ ist. Sei nun $(d_1, \ldots, d_n) = (b_1, \ldots, b_n)C$ eine andere Basis. Man sieht dann leicht, dass für die dualen Basen die Relation $(d'_1, \ldots, d'_n) = (b'_1, \ldots, b'_n)(C^T)^{-1}$ gilt. Also folgt (1.5) aus der Transformationsformel für Homomorphismen.

2. Reelle Skalarprodukte und Euklidische Vektorräume

In diesem Abschnitt $\mathbb{K} = \mathbb{R}$.

DEFINITION 2.1. Eine symmetrische Bilinearform $G \in \mathrm{Bilin}(V)$ heißt *positiv definit*, falls $G(v, v) > 0$ für alle $v \in V \setminus \{0\}$. Eine positiv definite symmetrische Bilinearform bezeichnen wir auch als *Skalarprodukt (auf V)*. Ein *Euklidischer Vektorraum* ist ein Paar (V, G), wobei V ein reeller Vektorraum und G ein Skalarprodukt auf V ist.

BEISPIELE. (a) Das kanonische Skalarprodukt auf $\mathbb{R}^n$ ist positiv definit, denn

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle_n = 0 \quad \Leftrightarrow \quad \sum_{i=1}^n x_i^2 = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

(b) Die symmetrische Bilinearform $\langle\,\cdot\,,\,\cdot\,\rangle_{3,1}$ ist nicht positiv definit, denn

$$\left\langle \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\rangle_{3,1} := -1.$$

(c) Sei $(b_i | i \in I)$ eine Basis von V. Dann definiert

$$G = \sum_{i\in I} b_i' \otimes b_i'$$

eine positiv definite symmetrische Bilinearform auf V. Diese Summe ist selbst dann wohldefiniert, wenn V unendlich-dimensional ist. Jeder reelle Vektorraum besitzt also ein Skalarprodukt, aber es ist natürlich nicht eindeutig bestimmt.

(d) Sei $V = C^0([0,1],\mathbb{R})$ der Raum der reellwertigen, stetigen Funktionen auf dem Intervall $[0,1]$. Dies ist ein reeller Vektorraum. Man definiert nun für $f, g \in V$:

$$G(f,g) := \int_{[0,1]} f(x)g(x)\,dx.$$

Diese Abbildung definiert ein Skalarprodukt auf V.

Sei (V,G) ein Euklidischer Vektorraum (reeller Vektorraum mit Skalarprodukt). Wir sagen $x, y \in V$ sind orthogonal, falls $G(x,y) = 0$. Wir schreiben dann $x \perp y$.

Ist A eine Teilmenge eines Euklidischen Vektorraums (V,G), dann definieren wir

$$A^\perp := \{w \in V \mid w \perp v \quad \forall v \in A\}.$$

Übungsaufgabe 2.2. Ist V endlich-dimensional und ist U ein Untervektorraum, so gilt

$$(U^\perp)^\perp = U, \qquad U \oplus U^\perp = V.$$

Man nennt dann $U^\perp$ das orthogonale Komplement von U.

Satz 2.3 (Cauchy-Schwarz-Ungleichung)**.** *Auf einem Euklidischen Vektorraum (V,G) gilt für alle $x, y \in V$*

$$G(x,y)^2 \le G(x,x)G(y,y)$$

und Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Beweis. Die Aussage ist offensichtlich richtig, wenn $y = 0$.

Im Fall $y \neq 0$ betrachten wir

$$0 \le G(x-\lambda y, x-\lambda y) = G(x,x) - 2\lambda G(x,y) + \lambda^2 G(y,y) =: P(\lambda).$$

Wir setzen dann $\lambda := G(x,y)/G(y,y)$. (Dies ist das Minimum von $P(\lambda)$.)

GRAPH VON $P(\lambda)$.

Dann ergibt sich
$$0 \leq G(x,x) - G(x,y)^2/G(y,y),$$
woraus sich die Ungleichung direkt ergibt.

Gleichheit gilt genau dann, wenn $G(x - \lambda y, x - \lambda y) = 0$, also wenn $x = \lambda y$. Wir haben also Gleichheit genau dann, wenn x und y linear abhängig sind. $\square$

DEFINITION 2.4. Für einen Euklidischen Vektorraum (V, G) definieren wir die *(von G induzierte) Norm* als
$$\|x\|_G := \sqrt{G(x,x)}$$
für alle $x \in V$.

Dann schreibt sich die Cauchy-Schwarz-Ungleichung als
$$|G(x,y)| \leq \|x\|_G \|y\|_G.$$

Die Norm $\|\cdot\|_G$ erfüllt die Eigenschaften:

(1) $\|x\|_G > 0$ für alle $x \in V \setminus \{0\}$,
(2) $\|\lambda x\|_G = |\lambda| \, \|x\|_G$ für alle $x \in V$ und alle $\lambda \in \mathbb{R}$,
(3) $\|x + y\|_G \leq \|x\|_G + \|y\|_G$ für alle $x, y \in V$ (Dreiecksungleichung).

Einen reellen Vektorraum V zusammen mit einer Abbildung $\|\cdot\| : V \to [0, \infty)$, der diese drei Eigenschaften erfüllt, bezeichnet man als normierten Vektorraum.

Beweis. Eigenschaft (1) und (2) sind offensichtlich. Wir rechnen:

$$\begin{aligned} \|x+y\|_G^2 &= G(x+y, x+y) = G(x,x) + G(x,y) + G(y,x) + G(y,y) \\ &= \|x\|_G^2 + \|y\|_G^2 + 2G(x,y) \leq \|x\|_G^2 + \|y\|_G + 2\|x\|_G\|y\|_G = (\|x\|_G + \|y\|_G)^2, \end{aligned}$$
und hieraus folgt die Dreiecksungleichung unmittelbar.

Winkel

DEFINITION 2.5. Sind $v, w \in V \setminus \{0\}$, so definiert man den Winkel zwischen v und w als die Zahl $\alpha \in [0, \pi]$, so dass
$$\cos \alpha = \frac{G(v,w)}{\|v\| \, \|w\|}.$$

Orthogonale und orthonormale Basen

DEFINITION 2.6. Sei (V, G) ein Euklidischer Vektorraum. Eine Familie $(v_i | i \in I)$ von Vektoren in V heißt *orthogonal*, falls $v_i \perp v_j$ für $i, j \in I$, $i \neq j$. Sie heißt *orthonormal*, wenn sie orthogonal ist und zusätzlich gilt $\|v_i\|_G = 1$ für alls $i \in I$.

LEMMA 2.7. *Jede G-orthogonale Familie $(v_i|i \in I)$ mit $v_i \neq 0 \quad \forall i \in I$ ist linear unabhängig.*

Beweis. Sei $(\lambda_i|i \in I)$ eine quasi-endliche Familie von Skalaren und $\sum_{i\in I} \lambda_i v_i = 0$. Wir rechnen für beliebiges j:

$$0 = G(v_j, \sum_{i\in I} \lambda_i v_i) = \sum_{i\in I} \lambda_i \underbrace{G(v_j, v_i)}_{=0,\text{ falls } i\neq j} = \lambda_j \|v_j\|_G^2.$$

Da $\|v_j\|_G \neq 0$, folgt $\lambda_j = 0$. $\square$

Gram-Schmidtsche Orthogonalisierungsverfahren Aufgabe: Eine Basis $(v_1, \dots, v_n)$ von V sei gegeben. Man bestimme eine Orthonormalbasis $(b_1, \dots, b_n)$ von V.

Wir geben eine rekursive Definition an und zeigen simultan induktiv, dass die bisher definierten Vektoren eine orthonormale Familie bilden.

$$b_1 := \frac{v_1}{\|v_1\|}.$$

Offensichtlich ist die einelementige Familie (b_1) orthonormal.

$$\tilde{b}_2 := v_2 - G(v_2, b_1)b_1.$$

Wir rechnen $G(\tilde{b}_2, b_1) = G(v_2, b_1) - G(v_2, b_1)G(b_1, b_1) = 0$. Nun normieren wir

$$b_2 := \frac{\tilde{b}_2}{\|\tilde{b}_2\|}.$$

Nun $G(b_2, b_2) = 1$. Also ist (b_1, b_2) orthonormal.

Sei nun $(b_1, \dots, b_k)$ orthonormal $k < n$. Wir setzen

$$\tilde{b}_{k+1} := v_{k+1} - \sum_{i=1}^{k} G(v_{k+1}, b_i)b_i.$$

Wir rechnen

$$G(\tilde{b}_{k+1}, b_j) = G(v_{k+1}, b_j) - \sum_{i=1}^{k} G(v_{k+1}, b_i) \underbrace{G(b_i, b_j)}_{\delta_{ij}} = 0.$$

Nun normieren wir

$$b_{k+1} := \frac{\tilde{b}_{k+1}}{\|\tilde{b}_{k+1}\|}.$$

Also ist nun $(b_1, \dots, b_{k+1})$ orthonormal.

Bemerkung: Die so konstruierten Vektoren erfüllen außerdem

$$\operatorname{span}\{b_1, \dots, b_k\} = \operatorname{span}\{v_1, \dots, v_k\}$$

für alle $k \in \{1, \ldots, k\}$ und $G(v_k, b_k) > 0$. Zu gegebener Basis $(v_1, \ldots, v_n)$ gibt es genau eine orthonormale Basis $(b_1, \ldots, b_n)$ mit diesen Zusatzbedingungen. Dies kann man nutzen, um die obige Formel herzuleiten, wenn man sie nicht auswendig weiß.

3. Sesquilinearformen, Komplexe Skalarprodukte und unitäre Vektorräume

In diesem Abschnitt ist immer $\mathbb{K} = \mathbb{C}$, und V, W, ... sind $\mathbb{C}$-Vektorräume.

Ziel: Wir hätten gerne auch ein Skalarprodukt auf einem endlich-dimensionalen $\mathbb{C}$-Vektorraum. Daraus wollen wir wie im reellen eine Norm definieren.

Problem: Ist B eine Bilinearform auf V. Es gelte $B(v, v) \in \mathbb{R}^+ = (0, \infty)$. Dann gilt $B(iv, iv) = i^2 B(v, v) \in \mathbb{R}^-$. Somit ist $\sqrt{B(v, v)}$ nicht immer definiert! Es gibt also keine positiv definiten Bilinearformen auf komplexen Vektorräumen.

Abhilfe: Sesquilinearformen. sesqui (latein.) bedeutet anderthalb.

Zuvor wollen wir klären, was „semi-linear" bedeutet: Eine Abbildung $f : V \to W$ heißt *semilinear (oder konjugiert linear)*, falls für alle $v, w \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$f(v + w) = f(v) + f(w) \qquad f(\lambda v) = \bar{\lambda} f(v).$$

BEISPIELE. (1) Die Konjugation $\mathbb{C} \to \mathbb{C}$, $z \mapsto \bar{z}$ ist semilinear.

(2) Die Konjugation von Vektoren in $\mathbb{C}^n$ ist semilinear: $\mathbb{C}^n \to \mathbb{C}^n$, $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \mapsto \begin{pmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_n \end{pmatrix} =: \overline{\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}}$.

(3) Die Konjugation von Matrizen ist semilinear: Dann ist

$$\mathrm{Mat}(n, m; \mathbb{C}) \to \mathrm{Mat}(n, m; \mathbb{C}), \qquad A = (a_{ij}) \mapsto \bar{A} := (\overline{a_{ij}})$$

semilinear.

Alle semi-linearen Abbildungen sind $\mathbb{R}$-linear, aber sie sind im allgemeinen nicht ($\mathbb{C}$-)linear. Eine Abbildung $f : V \to W$ die linear und semi-linear ist, erfüllt

$$f(iv) = if(v) = -if(v),$$

also $f(v) = 0$ für alle $v \in V$. Die Verkettung von zwei semilinearen Abbildungen ist linear.

Viele Aussagen, die für lineare Abbildungen gelten, gelten auch für semilineare. Muss man alles neu zeigen? Nein, es wird einfacher, wenn wir einen Trick anwenden. Wir definieren den zu V komplex konjugierten Vektorraum $\bar{V}$. Als Menge und als additive Gruppe definieren wir $\bar{V} = V$. Wir versehen aber $\bar{V}$ mit einer *anderen* Multiplikation mit Skalaren. Wenn $\cdot$ die Multiplikation mit Skalaren in V ist, dann definieren wir die Multiplikation $\times$ von $\bar{V}$ als $\lambda \times v := \bar{\lambda} \cdot v$. Man überprüfe die Vektorraumaxiome! Nun ist $\mathrm{Id}_V : V = (V, +, \cdot) \to \bar{V} = (V, +, \times)$ semi-linear und nicht mehr linear. Es folgt, dass $f : V \to W$ genau dann semi-linear ist, wenn $f : \bar{V} \to W$ linear ist, und dies gilt genau dann, wenn $f : V \to \bar{W}$ linear ist.

Wir sehen: Der Kern und das Bild einer semi-linearen Abbildung sind (komplexe) Untervektorräume. Bijektive semi-lineare Abbildungen bilden Basen auf Basen ab, etc..

DEFINITION 3.1. Sei V ein komplexer Vektorraum. Eine Abbildung $G : V \times V \to \mathbb{C}$ heißt *sesquilinear*, falls für alle $w \in V$ die Abbildung $V \to \mathbb{C}$, $v \mapsto G(v, w)$ linear ist, und die Abbildung $V \to \mathbb{C}$, $v \mapsto G(w, v)$ semi-linear ist. Den Raum aller Sesquilinearformen auf V bezeichnen wir mit $\mathrm{Sesqui}(V)$.

Eine Sesquilinearform G heißt *hermitesch*, falls

$$G(w, v) = \overline{G(v, w)}.$$

Eine Sesquilinearform G heißt *positiv definit*, falls für alle $v \in V \setminus \{0\}$ gilt:

$$G(v, v) \in \mathbb{R} \text{ und } G(v, v) > 0.$$

BEISPIELE.

(1) Für $v = \sum_{i=1}^n v_i e_i, w = \sum_{i=1}^n w_i e_i \in \mathbb{C}^n$ definieren wir

$$\langle v, w \rangle_n^{\mathbb{C}} := \sum_{i=1}^n v_i \bar{w}_i = v^T \bar{w}.$$

Die Abbildung $(v, w) \mapsto \langle v, w \rangle_n^{\mathbb{C}}$ ist sesquilinear, hermitesch und positiv definit. Man nennt dies das kanonische Skalarpordukt auf $\mathbb{C}^n$.

(2) Sei $A = (a_{ij}) \in \mathrm{Mat}(n, n; \mathbb{R})$. Notation $\bar{A} = (\bar{a}_{ij})$. Wir definieren $G_A(v, w) := v^T A \bar{w}$. Diese Abbildung ist sesquilinear.

$$G_A(w, v) = w^T A \bar{v} = \bar{v}^T A^T w = \overline{v^T \bar{A}^T \bar{w}} = \overline{G_{\bar{A}^T}(v, w)}.$$

Wenn also $\bar{A}^T = A$ gilt, so ist G_A eine hermitesche Sesquilinearform auf $\mathbb{C}^n$. Umgekehrt, wenn $G_A(w, v)$ hemitesch ist, dann haben wir auch

$$a_{ij} = e_i^T A e_j = G_A(e_i, e_j) = \overline{G_A(e_j, e_i)} = \bar{a}_{ji}.$$

Also gilt dann auch $A = \bar{A}^T$. Konkretes Beispiel:

$$A = \begin{pmatrix} 1 & i & 0 \\ -i & 2 & \alpha \\ 0 & \bar{\alpha} & 1 \end{pmatrix}$$

DEFINITION 3.2. Wir sagen $A \in \mathrm{Mat}(n, n; \mathbb{C})$ ist *hermitesch (oder selbstadjungiert)*, falls $\bar{A}^T = A$.

LEMMA 3.3. *Jede positiv definite Sesquilinearform ist hermitesch.*

Beweis. Sei G eine positiv definite Sesquilinearform auf V. Dann gilt für $x, y \in V$

$$\mathbb{R} \ni G(x + y, x + y) = \underbrace{G(x, x)}_{\in\mathbb{R}} + G(x, y) + G(y, x) + \underbrace{G(y, y)}_{\in\mathbb{R}}$$

also $G(x,y) + G(y,x) \in \mathbb{R}$. Dies gilt auch dann noch, wenn wir x durch ix ersetzen, d.h. $\mathbb{R} \ni G(ix,y)+G(y,ix) = iG(x,y)-iG(y,x)$. Dies impliziert $\mathbf{Im}G(x,y) = -\mathbf{Im}G(y,x)$ und $\mathbf{Re}G(x,y) = \mathbf{Re}G(y,x)$, d.h. $G(y,x) = \overline{G(y,x)}$. □

Definition 3.4. Ein (komplexes oder unitäres) Skalarprodukt auf V ist eine positiv definite Sesquilinearform auf V. Ein *unitärer Vektorraum* ist ein Paar (V,G), wobei V ein komplexer Vektorraum und G ein komplexes Skalarprodukt auf V ist. Wir setzen auch

$$\|v\|_G := \sqrt{G(v,v)}.$$

Der Vektor v heißt *normiert*, falls $\|v\|_G = 1$. Zwei Vektoren v und w sind (G-)orthogonal (in Formeln: $v \perp w$), wenn $G(v,w) = 0$.

Wenn $(v,w) \mapsto G(v,w)$ ein komplexes Skalarprodukt ist, dann ist $(v,w) \mapsto \mathbf{Re}G(v,w)$ ein reelles Skalarprodukt, das wir als $\mathbf{Re}G$ notieren. Es gilt $\|v\|_G = \|v\|_{\mathbf{Re}G}$, das heißt $\|\cdot\|_G$ ist wiederum eine Norm auf dem reellen Vektorraum V. Es gilt sogar ein bisschen mehr:

(1) $\|x\|_G > 0$ für alle $x \in V \setminus \{0\}$,
(2) $\|\lambda x\|_G = |\lambda|\, \|x\|_G$ für alle $x \in V$ und alle $\lambda \in \mathbb{C}$,
(3) $\|x+y\|_G \leq \|x\|_G + \|y\|_G$ für alle $x, y \in V$ (Dreiecksungleichung).

Die Relation (2) gilt nicht nur für alle $\lambda \in \mathbb{R}$, sondern sogar für alle $\lambda \in \mathbb{C}$:

$$\|\lambda x\|_G := \sqrt{\langle \lambda x, \lambda x\rangle_G} = \sqrt{\lambda\bar{\lambda}\langle x, x\rangle_G} = \sqrt{\lambda\bar{\lambda}}\|x\|_G = |\lambda|\, \|x\|_G.$$

Satz 3.5 (Cauchy-Schwarzsche Ungleichung)**.**

$$|G(v,w)| \leq \|v\|_G\|w\|_G$$

Beweis. Schreibe $G(v,w) = rz$ mit $r \in \mathbb{R}^+$, $|z| = 1$. Also $\bar{z} = z^{-1}$. Also

$$r = G(z^{-1}v, w) = \mathbf{Re}G(z^{-1}v, w) \leq \|z^{-1}v\|_{\mathbf{Re}G}\|w\|_{\mathbf{Re}G} = \|z^{-1}v\|_G\|w\|_G.$$

Da $G(z^{-1}v, z^{-1}v) = z^{-1}\overline{z^{-1}}G(v,v) = G(v,v)$ folgt $\|z^{-1}v\|_G = \|v\|_G$. □

Eine $\mathbb{C}$-Basis $(b_1, \ldots, b_n)$ eines unitären Vektorraums (V,G) ist *orthonormal* (bezüglich G), falls

$$G(b_i, b_j) = \delta_{ij}.$$

Die kanonische Basis von $\mathbb{C}^n$ ist orthonormal bezüglich G.

Gram-Schmidtsches Verfahren Gegeben sei eine Basis $(v_1, \ldots, v_n)$ von (V,G). Wir bestimmen eine Orthonormalbasis wie folgt:

$$b_1 := \frac{v_1}{\|v_1\|}.$$

$$\tilde{b}_{k+1} := v_{k+1} - \sum_{i=1}^{k} G(v_{k+1}, b_k)b_k.$$

$$b_{k+1} := \frac{\tilde{b}_{k+1}}{\|\tilde{b}_{k+1}\|}.$$

4. Isometrien und orthogonale Matrizen

In diesem Abschnitt $\mathbb{K} = \mathbb{R}$, alle Vektorräume sind Euklidische Vektorräume. Wir schreiben das Skalarprodukt auf V als $\langle\,\cdot\,,\,\cdot\,\rangle_V$.

Definition 4.1. Seien V und W Euklidische Vektorräume. Ein Abbildung $f : V \to W$ heißt Isometrie, falls f linear ist und

$$\langle f(v), f(\tilde{v})\rangle_W = \langle v, \tilde{v}\rangle_V \qquad \forall v, \tilde{v} \in V.$$

Ist f zusätzlich ein Isomorphismus, so nennt man f einen *isometrischen Isomorphismus.*

Isometrien sind immer injektiv, denn sei f eine Isometrie und $v \in \operatorname{Kern} f$. Dann folgt also

$$0 = \langle f(v), f(v)\rangle_W = \langle v, v\rangle_V$$

und somit $v = 0$.

Beispiel. Die Abbildung $\mathbb{R} \mapsto \mathbb{R}^2$, $x \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$ ist eine Isometrie von $(\mathbb{R}, \langle\,\cdot\,,\,\cdot\,\rangle_1)$ nach $(\mathbb{R}^2, \langle\,\cdot\,,\,\cdot\,\rangle_2)$, aber kein Isomorphismus.

Definition 4.2. Zwei Euklidische Vektorräume V und W heißen *isometrisch isomorph*, falls es einen isometrischen Isomorphismus $f : V \to W$ gibt.

Die Relation *isometrisch isomorph* ist eine Äquivalenzrelation.

Korollar 4.3. *Sei $n \in \mathbb{N}$. Jeder n-dimensionale Euklidische Vektorraum ist isometrisch isomorph zu $(\mathbb{R}^n, \langle\,\cdot\,,\,\cdot\,\rangle_n)$.*

Beweis. Sei $(V, \langle\,\cdot\,,\,\cdot\,\rangle_V)$ ein endlich-dimensionaler Euklidischer Vektorraum, $n = \dim V$. Wähle zunächst eine Basis von V. Wir konstruieren mit dem Gram-Schmidtschen Verfahren eine Orthonormalbasis $(b_1, \ldots, b_n)$. Dann ist $\mathcal{V}_{(b_i)}$ ein isometrischer Isomorphismus von $(\mathbb{R}^n, \langle\,\cdot\,,\,\cdot\,\rangle)$ nach (V, G). □

Proposition 4.4. *Seien $(V, \langle\,,\,\rangle_V)$ und $(W, \langle\,,\,\rangle_W)$ Euklidische Vektorräume, $f \in \operatorname{Hom}(V, W)$ und $(b_i | i \in I)$ eine orthonormale Basis von V. Dann ist f eine Isometrie genau dann, wenn $(f(b_i) | i \in I)$ eine orthonormale Familie ist.*

Beweis. Wenn f eine Isometrie ist, so gilt $\langle f(b_i), f(b_j)\rangle_W = \langle b_i, b_j\rangle_V = \delta_{ij}$, also ist $(f(b_i) | i \in I)$ orthonormal. Sei nun umgekehrt $(f(b_i) | i \in I)$ orthonormal. Wir zerlegen gegebene $v, w \in V$ in

$v = \sum_{i \in I} v_i b_i$ und $w = \sum_{i \in I} w_i b_i$ und erhalten dann

$$\begin{aligned}
\langle f(v), f(w) \rangle_W &= \langle f(\sum_{i \in I} v_i b_i), f(\sum_{i \in I} w_i b_i) \rangle_W \\
&= \langle \sum_{i \in I} v_i f(b_i), \sum_{j \in I} w_j f(b_j) \rangle_W \\
&= \sum_{i \in I} \sum_{j \in I} v_i w_j \underbrace{\langle f(b_i), f(b_j) \rangle_W}_{=\delta_{ij}} \\
&= \sum_{i \in I} \sum_{j \in I} v_i w_j \underbrace{\langle b_i, b_j \rangle_V}_{=\delta_{ij}} \\
&= \langle \sum_{i \in I} v_i b_i, \sum_{j \in I} w_j b_j \rangle_V \\
&= \langle v, w \rangle_V
\end{aligned}$$

Definition 4.5. Eine Matrix $A \in \mathrm{Mat}(n, n; \mathbb{R})$ heißt *orthogonal*, falls $\mathcal{L}_A : \mathbb{R}^n \to \mathbb{R}^n$ eine Isometrie ist.

Wenn $\mathcal{L}_A$ eine Isometrie ist, so ist diese Abbildung auch injektiv. Alle injektiven Endomorphismen eines endlich-dimensionalen Vektorraums sind bijektiv. Also ist $\mathcal{L}_A$ ein isometrischer Isomorphismus (sogar ein isometrischer Automorphismus).

Satz 4.6. *Äquivalent sind für* $A \in \mathrm{Mat}(n, n; \mathbb{R})$

(1) *A ist orthogonal,*
(2) *A^T ist orthogonal,*
(3) $AA^T = \mathbb{1}_n$,
(4) $A^T A = \mathbb{1}_n$,
(5) *die Spalten von A bilden eine orthonormale Basis,*
(6) *die Zeilen von A bilden eine orthonormale Basis.*

Beweis. (1) $\Leftrightarrow$ (5): Wenn $\mathcal{L}_A$ eine Isometrie ist, dann ist $(\mathcal{L}_A e_1, \ldots, \mathcal{L}_A e_n)$ eine orthonormale Familie, dies sind aber gerade die Spalten von A.
(2) $\Leftrightarrow$ (6): genauso, wenn wir A durch A^T ersetzen.
(5) $\Leftrightarrow$ (4): Bezeichne die i-te Spalte von A mit a_i. Der Koeffizient der Matrix $A^T A$ zum Indexpaar (i, j) ist $(A^T A)_{ij} = \langle a_i, a_j \rangle_n$. Dies ist genau dann gleich δ_{ij}, wenn $(a_1, \ldots, a_n)$ orthonormal ist.
(3) $\Leftrightarrow$ (6): genauso, wenn wir A durch A^T ersetzen.
(3) $\Leftrightarrow$ (4): $A^T A = \mathbb{1}_n$ gdw A ist Linksinverses zu A^T gdw A ist Rechtsinverses zu A^T gdw $A^T A = \mathbb{1}_n$. $\square$

Die orthogonalen $n \times n$-Matrizen, versehen mit der Matrizenmultiplikation, bilden eine Gruppe, die sogenannte *orthogonale Gruppe* $O(n)$. Es ist eine Untergruppe von $\mathrm{GL}(n, \mathbb{R})$.

5. Isometrien von unitären Vektorräumen und unitäre Matrizen

In diesem Abschnitt $\mathbb{K} = \mathbb{C}$. Alle Vektorräume sind unitäre Vektorräume. Wir schreiben das Skalarprodukt auf V als $\langle \cdot\,, \cdot \rangle_V$. Viele Beweise sind fast genauso wie im reellen Fall, viele Aussagen folgen sogar direkt aus dem reellen Fall.

Definition 5.1. Seien V und W unitäre Vektorräume. Ein Abbildung $f : V \to W$ heißt Isometrie, falls f linear ist und

$$\langle f(v), f(\tilde{v})\rangle_W = \langle v, \tilde{v}\rangle_V \qquad \forall v, \tilde{v} \in V.$$

Ist f zusätzlich ein Isomorphismus, so nennt man f einen *isometrischen Isomorphismus*.
Zwei unitäre Vektorräume V und W heißen *isometrisch isomorph*, falls es einen isometrischen Isomorphismus $f : V \to W$ gibt.

Korollar 5.2. *Sei $n \in \mathbb{N}$. Jeder n-dimensionale unitäre Vektorraum ist isometrisch isomorph zu $(\mathbb{C}^n, \langle \cdot\,, \cdot \rangle_n^{\mathbb{C}})$.*

Beweis wie im reellen Fall.

Proposition 5.3. *Seien $(V, \langle\ ,\ \rangle_V)$ und $(W, \langle\ ,\ \rangle_W)$ unitäre Vektorräume, $f \in \mathrm{Hom}_{\mathbb{C}}(V, W)$ und $(b_i|i \in I)$ eine orthonormale Basis von V. Dann ist f eine Isometrie genau dann, wenn $(f(b_i)|i \in I)$ eine orthonormale Familie ist.*

Definition 5.4. Eine Matrix $A \in \mathrm{Mat}(n, n; \mathbb{C})$ heißt *unitär*, falls $\mathcal{L}_A : \mathbb{C}^n \to \mathbb{C}^n$ eine Isometrie ist.

Wenn $\mathcal{L}_A$ eine Isometrie ist, so ist diese Abbildung auch injektiv. Alle injektiven Endomorphismen eines endlich-dimensionalen Vektorraums sind bijektiv. Also ist $\mathcal{L}_A$ ein isometrischer Isomorphismus (sogar ein isometrischer Automorphismus).

Satz 5.5. *Äquivalent sind für $A \in \mathrm{Mat}(n, n; \mathbb{C})$*

(1) *A ist unitär,*
(2) *$\bar{A}$ ist unitär,*
(3) *A^T ist unitär,*
(4) *$\bar{A}^T$ ist unitär,*
(5) *$A\bar{A}^T = \mathbb{1}_n$,*
(6) *$\bar{A}^T A = \mathbb{1}_n$,*
(7) *die Spalten von A bilden eine orthonormale Basis,*
(8) *die Zeilen von A bilden eine orthonormale Basis.*

Beweis. (1) $\Leftrightarrow$ (2): $(a_1, \ldots, a_n)$ ONB $\Leftrightarrow$ $\langle a_i, a_j\rangle_n^{\mathbb{C}} = \delta_{ij}$ $\Leftrightarrow$ $\langle \bar{a}_i, \bar{a}_j\rangle_n^{\mathbb{C}} = \overline{\delta_{ij}} = \delta ij$ $\Leftrightarrow$ $(\bar{a}_1, \ldots, \bar{a}_n)$ ONB.
(3) $\Leftrightarrow$ (4): wie (1) $\Leftrightarrow$ (2)
(1) $\Leftrightarrow$ (7) und (3) $\Leftrightarrow$ (8): wie im reellen

(7) $\Leftrightarrow$ (6): Bezeichne die i-te Spalten von A mit a_i. Der Koeffizient der Matrix $\bar{A}^T A$ zum Indexpaar (i,j) ist

$$(\bar{A}^T A)_{ij} = \bar{a}_i^T a_j = \langle \bar{a}_i, \bar{a}_j \rangle_n^{\mathbb{C}} = \overline{\langle a_i, a_j \rangle_n^{\mathbb{C}}} = \langle a_j, a_i \rangle_n^{\mathbb{C}}.$$

Dies ist genau dann gleich δ_{ij}, wenn $(a_1, \ldots, a_n)$ orthonormal ist.
(8) $\Leftrightarrow$ (5): Folgt aus (7) $\Leftrightarrow$ (6) wenn wir A durch A^T ersetzen und (5) konjugieren.
(5) $\Leftrightarrow$ (6): wie im reellen. □

Die unitären $n \times n$-Matrizen, versehen mit der Matrizenmultiplikation, bilden eine Gruppe, die sogenannte *unitäre Gruppe* $U(n)$. Es ist eine Untergruppe von $\mathrm{GL}(n, \mathbb{C})$.

6. Die Topologie von Euklidischen und unitären Vektorräumen

Hier $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Sei $(V, \langle \cdot\,, \cdot \rangle)$ ein Euklidischer oder unitärer Vektorraum. Dann ist $\| \cdot \|_V$ eine Norm auf V. Wir definieren hieraus eine Metrik

$$d_V : V \times V \to \mathbb{R}, \qquad d_V(x,y) := \|x - y\|_V.$$

Der reelle Vektorraum $\mathbb{R}^n$ und der komplexe Vektorraum $\mathbb{C}^n$ erhalten die aus der Analysis bekannten Standardtopologien.

Proposition 6.1. *Sind V und W endlich-dimensionale Euklidische bzw. unitäre Vektorräume, so ist jeder Homomorphismus $f : V \to W$ stetig.*

Beweis. Sei $(b_1, \ldots, b_n)$ eine Orthonormalbasis von V. Wir schreiben einen Vektor $v \in V$ als $v = \sum_{i=1}^n v_i b_i$. Mit der Cauchy-Schwarz-Ungleichung (Saalübung [1]) sehen wir
$\sum_{i=1}^n |v_i| \leq \sqrt{n} \|v\|_V$. Andererseits

$$\|f(v)\|_W = \|f(\sum_{i=1}^n v_i b_i)\|_W$$

$$\leq \sum_{i=1}^n |v_i| \|f(b_i)\|_W \leq \max_{i \in \{1,\ldots,n\}} \|f(b_i)\|_W \sum_{i=1}^n |v_i| \leq \sqrt{n} \max_{i \in \{1,\ldots,n\}} \|f(b_i)\|_W \|v\|_V = C\|v\|_V$$

mit $C = \sqrt{n} \max_{i \in \{1,\ldots,n\}} |\|f(b_i)\|_W$.

Sei $\epsilon > 0$ gegeben. Wir setzen $\delta := \epsilon / C$. Dann gilt für $x, y \in V$ mit $d_V(x,y) < \delta$:

$$d_W(f(x), f(y)) = \|f(x) - f(y)\|_W = \|f(v - w)\|_W \leq C\|v - w\|_V \leq C\delta = \epsilon.$$

---

[1] In der Basis $(b_1, \ldots, b_n)$ rechnet man nach, vgl. Saalübung: $\sum_{i=1}^n |v_i| = \langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} \pm 1 \\ \vdots \\ \pm 1 \end{pmatrix} \rangle \leq \|v\| \, \| \begin{pmatrix} \pm 1 \\ \vdots \\ \pm 1 \end{pmatrix} \| \leq \sqrt{n} \|v\|$.

Also ist die Abbildung stetig. □

Topologie von $V \times W$. Wir versehen $V \times W$ mit der Norm und Metrik

$$\|(v,w)\|_{V\times W} := \sqrt{\|v\|_V^2 + \|w\|_W^2} \qquad v \in V, w \in W.$$

$$d_{V\times W}((v,w),(\tilde v,\tilde w)) = \|(v-\tilde v, w-\tilde w\|_{V\times W} \qquad v,\tilde v \in V, w,\tilde w \in W.$$

Lemma 6.2. *Seien V, W, Z endlich-dimensionale Euklidische bzw. unitäre Vektorräume mit Normen $\|\cdot\|_V$, $\|\cdot\|_W$ und $\|\cdot\|_Z$ und $F : V \times W \to Z$ eine bilineare (oder sesquilineare) Abbildung. Dann gibt es eine Konstante $C \in \mathbb{R}$, so dass für alle $v \in V$ und $w \in W$ gilt:*

$$\|F(v,w)\|_Z \leq C\|v\|_V\,\|w\|_W.$$

Beweis. Wir wählen Orthonormalbasen $(b_1,\dots,b_n)$ von V und $(c_1,\dots,c_m)$ von W und schreiben $v = \sum_{i=1}^n v_i b_i$ und $w = \sum_{j=1}^n w_j c_j$. Mit der Cauchy-Schwarz-Ungleichung sieht man

$$\sum_{i=1}^n |v_i| \leq \sqrt{n}\|v\|_V \qquad \sum_{j=1}^m |w_i| \leq \sqrt{m}\|w\|_W.$$

Dann gilt mit der Dreiecksungleichung der Norm auf Z

$$\||F(v,w)\|_Z = \|\sum_{i=1}^n\sum_{j=1}^m v_i w_j F(b_i,c_j)\|_Z$$

$$\leq \sum_{i=1}^n\sum_{j=1}^m |v_i|\,|w_j|\,\|F(b_i,c_j)\|_Z \leq \sqrt{nm}\|v\|_V\,\|w\|_W \max_{i\in\{1,\dots,n\}}\max_{j\in\{1,\dots,m\}}\|F(b_i,c_j)\|_Z.$$

Das Lemma folgt also für $C := \sqrt{nm}\max_{i\in\{1,\dots,n\}}\max_{j\in\{1,\dots,m\}}\|F(b_i,c_j)\|_Z$. □

Korollar 6.3. *Seien V, W und Z endlich-dimensionale Euklidische bzw. unitäre Vektorräume und sei $F : V \times W \to Z$ bilinear (oder sesqulinear). Dann ist F stetig.*

Beweis. Für alle $v, a \in V$ und $w, b \in W$ gilt

$$\|F(v+a,w+b)-F(v,w)\|_Z = \|F(a,w)+F(v,b)+F(a,b)\|_Z \leq \|F(a,w)\|_Z+\|F(v,b)\|_Z+\|F(a,b)\|_Z$$
$$\leq C(\|v\|_V\|b\|_W + \|a\|_V\|w\|_W + \|a\|_V\|b\|_W).$$

Zu gegebenem $\epsilon > 0$ und festem $v \in V$ und $w \in W$ definieren wir

$$\delta := \frac{1}{C}\min\left\{\frac{\epsilon}{4\|v\|_V+1}, \frac{\epsilon}{4\|w\|_W+1}, \frac{\sqrt{\epsilon}}{4}\right\}.$$

Sei $d_{V\times W}((v+a,w+b),(v,w) < \delta$, d.h. $\delta > \|(a,b)\|_{V\times W} = \sqrt{\|a\|_V^2 + \|b\|_W^2}$, also $\|a\|_V < \delta$ und $\|b\|_W < \delta$. Dies impliziert

$$d_Z(F(v+a,w+b),F(v,w)) = \|F(v+a,w+b)-F(v,w)\|_Z \leq C\{(\|v\|_V\delta + \delta\|w\|_W + \delta^2) < \epsilon.$$

□

7. Reelle Hauptachsentransformation

In diesem Abschnitt sei $\mathbb{K} = \mathbb{R}$. Wir schreiben das kanonische Skalarprodukt auf $\mathbb{R}^n$ als $\langle \cdot\,, \cdot \rangle$ und die zugehörige Norm als $\| \cdot \|$. Ein Vektor v heißt *normiert*, falls $\|v\| = 1$. Sei S^{n-1} die Menge der normierten Vektoren $x \in \mathbb{R}^n$. Diese Menge S^{n-1} ist beschränkt und abgeschlossen, also kompakt und folgenkompakt (Analysis).

Satz 7.1. *Sei V ein endlich-dimensionaler Euklidischer Vektorraum und $F \in \mathrm{Sym}^2(V)$. Dann gibt es einen Vektor $x_F \in S^{n-1}$ so, dass*

$$F(x_F, x_F) \leq \frac{F(w,w)}{\|w\|^2}$$

für alle $w \in V \setminus \{0\}$.

Beweis. Da $(V, \langle \cdot\,, \cdot \rangle_V)$ isometrisch isomorph zu $(\mathbb{R}^n, \langle \cdot\,, \cdot \rangle_n)$ ist, können wir o.B.d.A. $V = \mathbb{R}^n$ und $\langle \cdot\,, \cdot \rangle_V = \langle \cdot\,, \cdot \rangle_n$ annehmen. Die Abbildungen

$$\begin{aligned} i : S^{n-1} &\to \mathbb{R}^n, & x &\mapsto x \\ \Delta : \mathbb{R}^n &\to \mathbb{R}^n \times \mathbb{R}^n, & x &\mapsto (x,x) \\ F : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \end{aligned}$$

sind stetig. Wir verketten sie und erhalten eine stetige Abbildung

$$\hat{F} : S^{n-1} \to \mathbb{R}, \qquad x \mapsto F(x,x).$$

Da S^{n-1} kompakt ist, ist das Bild von $\hat{F}$ beschränkt und es gibt ein $x_F \in S^{n-1}$ mit $\hat{F}(x_F, x_F) = \min_{x \in S^{n-1}} \hat{F}(x)$ (Analysis!).

Sei nun $w \in V \setminus \{0\}$. Dann ist $w/\|w\| \in S^{n-1}$ und somit gilt

$$F(x_F, x_F) \leq F\left(\frac{w}{\|w\|}, \frac{w}{\|w\|}\right) = \frac{F(w,w)}{\|w\|^2}.$$

$\square$

Proposition 7.2. *Seien V, F und x_F wie oben. Dann gilt für $y \in V$:*

$$y \perp x_F \quad \Rightarrow \quad F(y, x_F) = 0.$$

Beweis. Angenommen es gebe ein y mit $y \perp x_F$ und $F(y, x_F) \neq 0$. O.B.d.A. $\|y\| = 1$. Dann gilt

$$\|x_F + ty\|^2 = \langle x_F + ty, x_F + ty \rangle = \underbrace{\langle x_F, x_F \rangle}_{=1} + 2t \underbrace{\langle x, y \rangle}_{=0} + t^2 \underbrace{\langle y, y \rangle}_{=1} = 1 + t^2 \geq 1.$$

Sei μ wie oben. Dann gilt

$$f(t) := F(x_F + ty, x_F + ty) \geq \mu \|x_F + ty\|^2 \geq \mu$$

und andererseits

$$f(t) = F(x_F + ty, x_F + ty) = F(x_F, x_F) + 2tF(y, x_F) + t^2 F(y,y) = \mu + 2tF(y, x_F) + t^2 F(y,y).$$

Also nimmt die quadratische Polynom-Funktion $t \mapsto f(t)$ ihr Minimum in 0 an, d.h. $0 = f'(0) = 2F(y, x_F)$. □

THEOREM 7.3 (Orthonormale Hauptachsentransformation von reellen Bilinearformen)**.** *Sei* $(V, \langle\cdot\,,\cdot\rangle)$ *ein* n*-dimensionaler Euklidischer Vektorraum,* $F \in \mathrm{Sym}^2(V)$*. Dann gibt es eine Orthonormalbasis* $(b_1, \ldots, b_n)$ *und* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *so dass*

$$\mathrm{mat}_{(b_i)}(F) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Anders ausgedrückt:

$$F(b_i, b_j) = \delta_{ij}\lambda_i.$$

Bemerkung: Die verschiedenen Versionen der Hauptachsentransformation werden oft auch als Sylvesterscher Trägheitssatz bezeichnet.

Beweis durch Induktion über n*..*
Induktionsanfang: Die Behauptung ist offensichtlich für $n = 1$.
Induktionsschritt von $n-1$ **auf** n**:** Das Theorem gelte für $(n-1)$-dimensionale Euklidische Vektorräume.

Sei nun $F \in \mathrm{Sym}^2(V)$, $\dim V = n$. Der vorangehende Satz liefert ein b_n, $\|b_n\| = 1$ mit

$$\lambda_n := F(b_n, b_n) \leq \frac{F(x,x)}{\|x\|^2}$$

für alle $x \in V \setminus \{0\}$. Wir definieren

$$W := b_n^\perp = \{w \in V \,|\, w \perp b_n\}.$$

Dann gilt nach der Proposition $F(w, b_n) = 0$ für $w \in W$. Die Einschränkung von F auf $W \times W$ ergibt $F|_{W\times W} \in \mathrm{Sym}^2(W)$. Da $\dim W = n-1$ ist, haben wir nach Induktionsvoraussetzung eine Orthonormalbasis $(b_1, \ldots, b_{n-1})$ von W und reelle Zahlen $\lambda_1, \ldots, \lambda_{n-1}$ mit

$$F(b_i, b_j) = \delta_{ij}\lambda_i \qquad \forall i, j \in \{1, \ldots, n-1\}.$$

Da auch $F(b_n, b_i) = 0$ für $i < n$, folgt

$$F(b_i, b_j) = \delta_{ij}\lambda_i \qquad \forall i, j \in \{1, \ldots, n\}.$$

□

Wenn wir nicht unbedingt eine Orthonormalbasis wollen, können wir noch eine einfachere Gestalt erhalten.

Korollar 7.4 (Hauptachsentransformation von reellen Bilinearformen)**.** *Sei V ein endlich-dimensionaler Vektorraum, und sei $F \in \mathrm{Sym}^2(V)$. Dann gibt es eine Basis $(b_1, \ldots, b_n)$ von V, so dass*

$$\mathrm{mat}_{(b_i)}(F) = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & -1 & & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & 0 & \cdots & 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 & & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \cdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} . = \begin{pmatrix} \mathbb{1}_r & 0 & 0 \\ 0 & -\mathbb{1}_s & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

V ist also eine direkte Summe $V = V_+ \oplus V_- \oplus V_0$, wobei F auf V_+ positiv definit ist, auf V_- negativ definit ist, und auf V_0 verschwindet, und $F(v, w) = 0$, wenn v und w in verschiedenen Summanden liegen.

Wenn zusätzlich ein Skalarprodukt $\langle \cdot\, , \cdot \rangle$ gewählt ist, dann kann die Basis orthogonal, aber nicht immer orthonormal gewählt werden.

Beweis. Wir wählen ein beliebiges Skalarprodukt $\langle \cdot\, , \cdot \rangle$ auf V. Wir bestimmen zunächst eine Basis $(\tilde{b}_1, \ldots, \tilde{b}_n)$, so dass die Matrix von F diagonal ist, sagen wir $F(\tilde{b}_i, \tilde{b}_j) = \lambda_i \delta_{ij}$. Durch Umordnen können wir erreichen, dass $\lambda_1, \ldots, \lambda_r$ positiv sind, $\lambda_{s+1}, \ldots, \lambda_{r+s}$ negativ sind und $\lambda_{r+s+1} = \ldots = \lambda_n = 0$ für geeignetes r und s. Wir setzen nun

$$b_i := \frac{1}{\sqrt{|\lambda_i|}} \tilde{b}_i$$

für $i \leq r + s$ und $b_i := \tilde{b}_i$ für $i > r + s$. Die Transformationsformel liefert die gewünschte Form. □

Bemerkung: Die Zahlen r und s sind durch F bereits bestimmt, sie hängen nicht von der Wahl der Basis ab.

Wiederholung: Eine *symmetrische Matrix* ist eine Matrix A mit $A^T = A$. Dann ist $\mathrm{Bilin}_{(e_i)}(A) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(v, w) \mapsto v^T A w$ eine symmetrische Bilinearform.

Korollar 7.5 (Orthonormale Hauptachsentransformation von symmetrischen Matrizen)**.** *Sei A eine symmetrische $n \times n$-Matrix. Dann gibt es eine Matrix $S \in O(n)$, so dass $S^T A S = S^{-1} A S$ eine Diagonalmatrix ist.*

Wiederholung: Die Skalare auf der Diagonalen sind die dann die Eigenwerte von A (mit Multiplizität).

Achtung: Wir können i.a. nicht erreichen, dass auf der Diagonale nur 0, 1 und -1 steht. Beispiel: $A = 2\mathbb{1}_n$.

Beweis. Zu der symmetrischen Bilinearform $F := \mathrm{Bilin}_{(e_i)}(A)$ gibt es eine orthonormale Basis $(b_1, \ldots, b_n)$, so dass $\mathrm{mat}_{(b_i)}(F)$ diagonal ist. Sei S die Matrix, deren Spalten die b_i sind, d.h. $(e_1, \ldots, e_n)S = (b_1, \ldots, b_n)$. Da $(b_1, \ldots, b_n)$ orthonormal ist, ist S orthogonal, also $S^T = S^{-1}$. Die Transformationsformel besagt dann

$$\mathrm{mat}_{(b_i)}(F) = S^T \mathrm{mat}_{(e_i)}(F) S = S^T AS.$$

□

Definition 7.6. Zwei quadratische Matrizen A, $B \in \mathrm{Mat}(n, n; \mathbb{R})$ sind *orthogonal ähnlich*, falls es ein $S \in O(n)$ gibt mit $B = S^{-1}AS$. Eine Matrix A ist *orthogonal diagonalisierbar* (*orthogonal trigonalisierbar*), falls A orthogonal ähnlich zu einer diagonalen (trigonalen) Matrix ist.

Orthogonal ähnlich ist eine Äquivalenzrelation.

Korollar 7.7. *Eine Matrix $A \in \mathrm{Mat}(n, n; \mathbb{R})$ ist genau dann orthogonal diagonalisierbar, wenn A symmetrisch ist.*

Beweis. Dass symmetrische Matrizen orthogonal diagonalisierbar sind, besagt das letzte Korollar.

Andrerseits sei A orthogonal diagonalisierbar, d.h. $S^{-1}AS = D$ für ein $S \in O(n)$. Mit $S^{-1} = S^T$ haben wir dann $A = SDS^T$ und deswegen $A^T = (SDS^T)^T = (S^T)^T D^T S^T = SDS^T = A$. □

8. Komplexe $n \times m$-Matrizen als reelle $2n \times 2m$-Matrizen

Es ist ziemlich offensichtlich, dass man jede reelle Matrix $A \in \mathrm{Mat}(n, m; \mathbb{R})$ auch als komlexe Matrix in $\mathrm{Mat}(n, m; \mathbb{C})$ auffassen kann. Jede reelle Zahl ist ja auch eine komplexe Zahl (mit verschwindendem Imaginärteil). Man kann also jeden reellen Koeffizienten von A einfach als komplexe Zahl (mit verschwindendem Imaginärteil) intepretieren, und dann ist auch $A \in \mathrm{Mat}(n, m; \mathcal{M})$. In diesem Sinne gilt $\mathrm{Mat}(n, m; \mathbb{R}) \subset \mathrm{Mat}(n, m; \mathbb{C})$. Eine Matrix $B \in \mathrm{Mat}(n, m; \mathbb{C})$ ist genau dann in $\mathrm{Mat}(n, m; \mathbb{R})$, wenn $\overline{B} = B$. Reelle Matrizen sind also spezielle komplexe Matrizen.

Diese Einbettung von reellen Matrizen in komplexe Matrizen ist aber nicht der Gegenstand dieses Abschnitt. Wir wollen vielmehr komplexe Matrizen als Teilmenge von reellen Matrizen verstehen: Wir wollen eine Matrix $B \in \mathrm{Mat}(n, m; \mathbb{C})$ als reelle $2n \times 2m$-Matrix auffassen können. $\mathrm{Mat}(n, m; \mathbb{C}) \hookrightarrow \mathrm{Mat}(2n, 2m; \mathbb{R})$.

Sind V und W $\mathbb{C}$-Vektorräume, so notieren wir mit $\mathrm{Hom}_{\mathbb{C}}(V, W)$ die $\mathbb{C}$-linearen Abbildungen von V nach W, und mit $\mathrm{Hom}_{\mathbb{R}}(V, W)$ die $\mathbb{R}$-linearen. Jede $\mathbb{C}$-lineare Abbildung ist $\mathbb{R}$-linear, aber die

Umkehrung ist falsch: nichttriviale semilineare Abbildungen sind zum Beispiel $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.

Die Abbildung $\mathrm{Mat}(n,m;\mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^m, \mathbb{K}^n)$, $A \to \mathcal{L}_A$ ist ein Isomorphismus von $\mathbb{K}$-Vektorräumen für jeden Körper $\mathbb{K}$, insbesondere für $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$. Der komplexe Vektorraum $\mathbb{C}^n$ ist auch ein reeller Vektorraum der Dimension $2n$ mit Basis $\mathcal{B}_n^{\mathbb{R}} = (e_1, e_2, \ldots, e_n, ie_1, e_2, \ldots, ie_n)$.

Sei nun also $A \in \mathrm{Mat}(n,m;\mathbb{C})$. Wir definieren dann

$$A_{\mathbb{R}} := \mathrm{Mat}^{\mathcal{B}_m^{\mathbb{R}}}_{\mathcal{B}_n^{\mathbb{R}}}(\mathcal{L}_A) \in \mathrm{Mat}(2n, 2m; \mathbb{R}).$$

Beispiele.

(1) $n = m = 1$, $A = (z) \in \mathrm{Mat}(1,1;\mathbb{C})$, $z = x + iy$, $e_1 = 1$. Dann gilt $\mathcal{L}_A(e_1) = xe_1 + yie_1$ und $\mathcal{L}_A(ie_1) = -ye_1 + xie_1$, also

$$A_{\mathbb{R}} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \in \mathrm{Mat}(2,2;\mathbb{R}).$$

(2) $n = m = 2$, $A = \begin{pmatrix} 2 & 0 \\ 0 & 1+i \end{pmatrix}$. Es gilt

$$\begin{aligned} \mathcal{L}_A(e_1) &= 2e_1 = 2e_1 + 0e_2 + 0ie_1 + 0ie_2 \\ \mathcal{L}_A(e_2) &= (1+i)e_2 = 0e_1 + 1e_2 + 0ie_1 + 1ie_2 \\ \mathcal{L}_A(ie_1) &= 2ie_1 = 0e_1 + 0e_2 + 2ie_1 + 0ie_2 \\ \mathcal{L}_A(ie_2) &= (1+i)ie_2 = 0e_1 - 1e_2 + 0ie_1 + 1ie_2, \end{aligned}$$

also

$$A_{\mathbb{R}} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}(4,4;\mathbb{R}).$$

(3) $n = m$, $A = i\mathbb{1}_n$. Dann gilt

$$A_{\mathbb{R}} = \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix}.$$

Allgemein gilt

$$A_{\mathbb{R}} = \begin{pmatrix} \mathbf{Re}A & -\mathbf{Im}A \\ \mathbf{Im}A & \mathbf{Re}A \end{pmatrix}.$$

Die Abbildung $I_{\mathbb{R}} : \mathrm{Mat}(n,m;\mathbb{C}) \to \mathrm{Mat}(2n,2m;\mathbb{R})$, $A \mapsto A_{\mathbb{R}}$ ist offensichtlich eine injektive $\mathbb{R}$-lineare Abbildung (ein $\mathbb{R}$-Monomorphismus).

Lemma 8.1. *Eine Matrix $B \in \mathrm{Mat}(2n, 2m; \mathbb{R})$ ist im Bild von $I_\mathbb{R}$ genau dann, wenn*
$$\begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} B = B \begin{pmatrix} 0 & -\mathbb{1}_m \\ \mathbb{1}_m & 0 \end{pmatrix}.$$

Beweis. Sei B im Bild, d.h. $B = A_\mathbb{R}$ für eine Matrix $A \in \mathrm{Mat}(n, m; \mathbb{C})$. Da $\mathcal{L}_A$ komplex linear ist, gilt $A(iv) = iA(v)$, also $A_\mathbb{R} \begin{pmatrix} 0 & -\mathbb{1}_m \\ \mathbb{1}_m & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} A_\mathbb{R}$.

Sei umgekehrt $B \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} B$. Wir zerteilen B in $n \times m$-Blöcke:
$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$
Dann gilt
$$B \begin{pmatrix} 0 & -\mathbb{1}_m \\ \mathbb{1}_m & 0 \end{pmatrix} = \begin{pmatrix} B_{12} & -B_{11} \\ B_{22} & -B_{21} \end{pmatrix}$$
und
$$\begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} B = \begin{pmatrix} -B_{21} & -B_{22} \\ B_{11} & B_{12} \end{pmatrix}.$$
Es ergibt sich also $B_{12} = -B_{21}$ und $B_{11} = B_{22}$. Wir setzen nun $A := B_{11} + iB_{21}$. Dann gilt $A_\mathbb{R} = B$, also ist B im Bild. $\square$

Kurzschreibweise
$$\mathrm{Mat}(n, m; \mathbb{C}) \hat{=} \{A \in \mathrm{Mat}(2n, 2m; \mathbb{R}) \mid A \begin{pmatrix} 0 & -\mathbb{1}_m \\ \mathbb{1}_m & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} A\}.$$

Lemma 8.2. *Eine Matrix $A \in \mathrm{Mat}(n, n; \mathbb{C})$ ist genau dann hermitesch, falls $A_\mathbb{R}$ symmetrisch ist. Eine Matrix $A \in \mathrm{Mat}(n, n; \mathbb{C})$ ist genau dann unitär, wenn $A_\mathbb{R}$ orthogonal ist.*

Beweis. Es gilt
$$(A^T)_\mathbb{R} = \begin{pmatrix} \mathbf{Re}A^T & -\mathbf{Im}A^T \\ \mathbf{Im}A^T & \mathbf{Re}A^T \end{pmatrix}.$$
$$(\bar{A})_\mathbb{R} = \begin{pmatrix} \mathbf{Re}A & \mathbf{Im}A \\ -\mathbf{Im}A & \mathbf{Re}A \end{pmatrix}.$$
$$(\bar{A}^T)_\mathbb{R} = \begin{pmatrix} \mathbf{Re}A^T & \mathbf{Im}A^T \\ -\mathbf{Im}A^T & \mathbf{Re}A^T \end{pmatrix} = \begin{pmatrix} \mathbf{Re}A & -\mathbf{Im}A \\ \mathbf{Im}A & \mathbf{Re}A \end{pmatrix}^T = (A_\mathbb{R})^T.$$
also: A hermitesch gdw $\bar{A}^T = A$ gdw $(A_\mathbb{R})^T = A_\mathbb{R}$ gdw $A_\mathbb{R}$ symmetrisch.

Es gilt $(AB)_\mathbb{R} = A_\mathbb{R} B_\mathbb{R}$ und $(\mathbb{1}_n)_\mathbb{R} = \mathbb{1}_{2n}$, also auch $(A^{-1})_\mathbb{R} = (A_\mathbb{R})^{-1}$.

A unitär gdw $\bar{A}^T A = \mathbb{1}_n$ gdw $(\bar{A}^T A)_\mathbb{R} = (\mathbb{1}_n)_\mathbb{R}$ gdw $(A_\mathbb{R})^T A_\mathbb{R} = \mathbb{1}_{2n}$ gdw $A_\mathbb{R}$ orthogonal. $\square$

Kurzschreibweise für die zweite Behauptung:

$$U(n) \hat{=} O(2n) \cap \mathrm{Mat}(n,n;\mathbb{C}).$$

9. Komplexe Hauptachsentransformation

Wir notieren das (komplexe) kanonische Skalarprodukt auf $\mathbb{C}^n$ mit $\langle\,\cdot\,,\,\cdot\,\rangle_n^{\mathbb{C}}$. Dann ist $\langle\,\cdot\,,\,\cdot\,\rangle_n^{\mathbb{R}} := \mathbf{Re}\langle\,\cdot\,,\,\cdot\,\rangle_n^{\mathbb{C}}$ das reelle kanonische Skalarprodukt auf $\mathbb{C}^n$. Wir setzen

$$J := \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix}.$$

Satz 9.1 (Unitäre Hauptachsentransformation von hermiteschen Matrizen)**.** *Sei $A \in \mathrm{Mat}(n,n;\mathbb{C})$ hermitesch. Dann gibt es eine Matrix $U \in U(n)$, so dass $\bar{U}^T AU = U^{-1}AU$ eine Diagonalmatrix mit reellen Diagonaleinträgen ist.*

Lemma 9.2. *Für alle $v \in \mathbb{C}^n = \mathbb{R}^{2n}$ gilt*

$$\langle v, Jv\rangle_n^{\mathbb{R}} = 0.$$

Außerdem ist $\mathcal{L}_J : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ eine Isometrie.

Beweis.

$$\langle v, v\rangle_n^{\mathbb{C}} \in \mathbb{R},$$

also $\langle v, \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} v\rangle_n^{\mathbb{C}} = \langle v, iv\rangle_n^{\mathbb{C}} = -i\langle v, v\rangle_n^{\mathbb{C}} \in i\mathbb{R}$ und somit

$$\langle v, \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} v\rangle_n^{\mathbb{R}} = \mathbf{Re}(\langle v, \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix} v\rangle_n^{\mathbb{C}}) = 0.$$

Die Isometrie-Eigenschaft folgt aus

$$\langle iv, iv\rangle_n^{\mathbb{C}} = i(-i)\langle v, v\rangle_n^{\mathbb{C}} = \langle v, v\rangle_n^{\mathbb{C}}.$$

□

Beweis des Satzes. Sei A hermitesch, dann ist $A_{\mathbb{R}}$ symmetrisch. Es gibt deswegen eine Matrix $S \in O(2n)$, so dass $D := S^{-1}A_{\mathbb{R}}S$ diagonal ist.

(Achtung: Man weiß jetzt noch nicht, dass $S = U_{\mathbb{R}}$ für eine unitäre Matrix $U \in U(n)$.)

Seien $\lambda_1, \dots, \lambda_r$ die Eigenwerte von $A_{\mathbb{R}}$. Wir wissen durch Satz 3.9, dass die Eigenräume E_{λ_i} von $A_{\mathbb{R}}$ erfüllen

$$\mathbb{C}^n \cong \mathbb{R}^{2n} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r} \tag{9.3}$$

im Sinne von reellen Untervektorräumen.

Ist $v \in E_{\lambda_i}$, dann gilt $A_{\mathbb{R}} Jv \cong Aiv = iAv = JA_{\mathbb{R}}v = J\lambda_i v = \lambda_i Jv$. Also ist auch $Jv \in E_{\lambda_i}$. Die Räume E_{λ_i} sind also auch komplexe Unterräume und die direkte Summenzerlegung in 9.3 gilt auch im Sinne von komplexen Unterräumen. Wir wählen nun induktiv eine Orthonormalbasis

$$(b_1, b_2, \dots, b_n, -Jb_1, \dots, -Jb_n)$$

aus Eigenvektoren von $A_{\mathbb{R}}$. Zunächst wählen wir einen normierten Eigenvektor b_1. Dann ist $-Jb_1$ auch normiert und senkrecht auf b_1, also $(b_1, -Jb_1)$ orthonormal. Die Abbildung $\mathcal{L}_{A_{\mathbb{R}}}$ bildet $(b_1, -Jb_1)^{\perp}$ auf sich selbst ab. Wir wählen einen Eigenvektor b_2 in $(b_1, -Jb_1)^{\perp}$ und erhalten ein orthonormale Familie $(b_1, b_2, -Jb_1, -Jb_2)$, u.s.w..

Wir setzen

$$U := (b_1 \quad \dots \quad b_n).$$

$\mathbf{Re}\langle b_i, b_j\rangle_n^{\mathbb{C}} = \delta_{ij}$ und $\mathbf{Im}\langle b_i, b_j\rangle_n^{\mathbb{C}} = \mathbf{Re} - i\langle b_i, b_j\rangle_n^{\mathbb{C}} = \langle -Jb_i, b_j\rangle_{\mathbb{R}} = 0$. Also ist $U \in U(n)$. Ferner

$$U_{\mathbb{R}} = (b_1 \quad \dots \quad b_n \quad -Jb_1 \quad \dots \quad -Jb_n).$$

Da diese Zeilenvektoren Eigenvektoren von $A_{\mathbb{R}}$ sind, ist

$$(U_{\mathbb{R}})^{-1} A_{\mathbb{R}} U_{\mathbb{R}} = (U^{-1}AU)_{\mathbb{R}}$$

diagonal. Somit ist auch $U^{-1}AU$ diagonal und hat reelle Koeffizienten auf der Diagonalen. □

10. Adjungierte Homomorphismen und selbstadjungierte Endomorphismen

In diesem Abschnitt sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, und $(V, \langle\,\cdot\,,\,\cdot\,\rangle_V)$, $(W, \langle\,\cdot\,,\,\cdot\,\rangle_W)$ seien endlich-dimensionale unitäre bzw. Euklidische Vektorräume.

Definition 10.1. Sei $f \in \mathrm{Hom}(V, W)$. Ein Operator $g \in \mathrm{Hom}(W, V)$ heißt adjungiert zu f, falls für alle $v \in V$ und $w \in W$ gilt

$$\langle w, f(v)\rangle_W = \langle g(w), v\rangle_V.$$

Offensichtlich gilt f adjungiert zu g gdw g adjungiert zu f.

Satz 10.2. *Zu $f \in \mathrm{Hom}(V, W)$ gibt es genau einen adjungierten Homomorphismus, den wir ab sofort mit $f^* \in \mathrm{Hom}(W, V)$ bezeichnen. Es gilt $(f^*)^* = f$. Die Abbildung $\mathrm{Hom}(V, W) \to \mathrm{Hom}(W, V)$, $f \mapsto f^*$ ist eine semilinear Bijektion.*

Beweis der Eindeutigkeit. Seien g_1 und g_2 adjungierte Abbildungen zu f. Dann gilt für alle $v \in V$ und $w \in W$:

$$\langle g_1(w) - g_2(w), v\rangle_W = \langle g_1(w), v\rangle_W - \langle g_2(w), v\rangle_W = \langle w, f(v)\rangle_V - \langle w, f(v)\rangle_V = 0.$$

Dies gilt insbesondere für $v = g_1(w) - g_2(w)$, also $\langle g_1(w) - g_2(w), g_1(w) - g_2(w)\rangle_W = 0$, also $g_1(w) = g_2(w)$. Da w beliebig ist, folgt $g_1 = g_2$. □

Beweis der Existenz mit Hilfe der dualen Abbildung $f' : W' \to V'$. Man betrachte die semilineare Abbildung

$$b_V : V \to V', \qquad v \mapsto b_V := \langle \cdot\,, v\rangle_V .$$

Wir bestimmen den Kern. Sei $v \in \operatorname{Kern} b_V$, also $\langle \tilde v, v\rangle_V = 0$ für alle $\tilde v \in V$. Dies gilt insbesondere für $\tilde v = v$, also $\langle v, v\rangle_V = 0$ und somit $v = 0$. Somit $\operatorname{Kern} b_V = \{0\}$, d.h. b_V ist injektiv. Mit $\infty > \dim V = \dim V'$ sehen wir dann, dass b_V eine semilineare Bijektion ist. Analog definieren wir die semilineare Bijektion $b_W : W \to W'$.

Wir setzen nun

$$g := (b_V)^{-1} \circ f' \circ b_W .$$

Diese Abbildung ist nun eine lineare Abbildung, also $g \in \operatorname{Hom}(W, V)$.

DIAGRAMM

Wir zeigen, dass g zu f adjungiert ist.

$$\begin{aligned}
\langle g(w), v\rangle_V &= \overline{\langle v, g(w)\rangle_W} = \overline{b_V(g(w))(v)} \\
&= \overline{(f'(b_W(w)))(v)} = \overline{\underbrace{b_W(w)}_{\in W'}(f(v))} \\
&= \overline{\langle f(v), w\rangle_W} = \langle w, f(v)\rangle_W
\end{aligned}$$

Da $f \mapsto f'$ linear ist, folgt die Semilinearität von $f \mapsto f^*$ aus der Semilinearität von $(b_V)^{-1}$. Die restlichen Aussagen sind nun klar. □

Beweis der Existenz mit Hilfe von Basen. Sei $(b_1, \ldots, b_n)$ eine Orthonormalbasis von V und $(d_1, \ldots, d_m)$ eine Orthonormalbasis von W. Sei $A = (a_{ij})_{ij} = \operatorname{Mat}^{(b_i)}_{(d_i)}(f)$, d.h. $f(b_j) = \sum_{i=1}^m a_{ij} d_i$. Es folgt

$$\langle d_k, f(b_j)\rangle_W = \langle d_k, \sum_{i=1}^m a_{ij} d_i\rangle_W = \sum_{i=1}^m \overline{a_{ij}} \delta_{ik} = \overline{a_{kj}}$$

□

Sei nun $g : \operatorname{Hom}(W, V)$ der Homomorphismus, der durch die Matrix $\bar A^T$ beschrieben wird, d. h. $\bar A^T = (\bar a_{ji})_{ij} = \operatorname{Mat}^{(d_i)}_{(b_i)}(g)$, dann gilt ebenso

$$\langle g(d_k), b_j\rangle_V = \langle \sum_{i=1}^n \bar a_{ki} b_i, b_j\rangle_V = \sum_{i=1}^n \bar a_{ki}\delta_{ij} = \bar a_{kj} .$$

Durch Linearkombination folgt für beliebige $v \in V$ und $w \in W$:

$$\langle w, f(v)\rangle_W = \langle g(w), v\rangle_V ,$$

d.h. g ist zu f adjungiert. Die restlichen Aussagen wie oben. □

Dieser Beweis besagt auch, dass für Orthonormalbasen (v_i) und (w_i) gilt

$$\mathrm{Mat}_{(v_i)}^{(w_i)}(f^*) = \overline{\mathrm{Mat}_{(w_i)}^{(v_i)}(f)}^T .$$

Man schreibt oft $A^* := \bar{A}^T$, dann ergibt dies $\mathrm{Mat}_{(v_i)}^{(w_i)}(f^*) = \mathrm{Mat}_{(w_i)}^{(v_i)}(f)^*$.

Lemma 10.3. Kern $f^* = (\mathrm{Bild} f)^\perp$ *und* $\mathrm{Bild} f^* = (\mathrm{Kern}\ f)^\perp$.

Beweis.

$$\begin{aligned} y \in \mathrm{Kern}\ f^* \quad &\Leftrightarrow \quad f^*(y) = 0 \quad \Leftrightarrow \quad \langle f^*(y), x\rangle_V = 0 \quad \forall x \in V \\ &\Leftrightarrow \quad \langle y, f(x)\rangle_W = 0 \quad \forall x \in V \Leftrightarrow y \perp \mathrm{Bild} f. \end{aligned}$$

Also Kern $f^* = (\mathrm{Bild} f)^\perp$ für alle $f \in \mathrm{Hom}(V, W)$.

Sei nun $\tilde{f} \in \mathrm{Hom}(\tilde{V}, \tilde{W})$ gegeben. Wir setzen $V := \tilde{W}$, $W := \tilde{V}$ und $f := (\tilde{f})^* \in \mathrm{Hom}(V, W)$, also $f^* = \tilde{f}$. Die erste Aussage ergibt dann

$$\mathrm{Kern}\ \tilde{f} = \mathrm{Kern}\ f^* = (\mathrm{Bild} f)^\perp = (\mathrm{Bild} \tilde{f}^*)^\perp.$$

Somit

$$\mathrm{Bild} \tilde{f}^* = (\mathrm{Kern}\ \tilde{f})^\perp$$

für alle $\tilde{f} \in \mathrm{Hom}(\tilde{V}, \tilde{W})$. Wenn wir nun $\tilde{f}$ durch f, $\tilde{V}$ durch V und $\tilde{W}$ durch W ersetzen, erhalten wir die zweite Aussage. □

f^* injektiv $\Leftrightarrow$ f surjektiv

f^* surjektiv $\Leftrightarrow$ f injektiv

Definition 10.4. Sein V ein endlich-dimensionaler unitärer bzw. Euklidischer Vektorraum. Ein Endomorphismus $f \in \mathrm{End}(V)$ heißt *selbstadjungiert*, wenn $f^* = f$.

Im Fall $\mathbb{K} = \mathbb{R}$ (und $\dim V < \infty$) sagt man manchmal auch symmetrisch an Stelle von selbstadjungiert, und im Falle $\mathbb{K} = \mathbb{C}$ auch hermitesch an Stelle von selbstadjungiert.

In Orthonormalbasen ausgedrückt, entprechen sich also die folgenden Objekte

$\mathbb{K} = \mathbb{R}$: symmetrische Matrizen, symmetrische Bilinearformen, symmetrische Endomorphismen

$\mathbb{K} = \mathbb{C}$: hermitesche Matrizen, hermitesche Sesquilinearformen, hermitesche Endomorphismen

Wir habe auch eine Version der Hauptachsentransformation für selbstadjungierte Endomorphismen.

Satz 10.5 (Orthonormale Hauptachsentransformationen von selbstadjungierten Endomorphismen)**.** *Sei V ein endlich-dimensionaler unitärer bzw. Euklidischer Vektorraum, und $f \in End(V)$. Dann ist f genau dann selbstadjungiert, wenn es eine Orthonormalbasis $\mathcal{B}$ von V, so dass*

$$\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(f)$$

eine reelle Diagonalmatrix ist.

Beweis. Es gebe zunächst eine Orthonormalbasis $\mathcal{B}$ von V, so dass

$$A := \mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f)$$

eine reelle Diagonalmatrix ist. Dann gilt $\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f^*) = A^* = A$, also $f^* = f$.

Sei nun umgekehrt $f = f^*$, und $\mathcal{D} = (d_1, \ldots, d_n)$ eine Orthonormalbasis von V. Dann ist $A := \mathrm{Mat}^{\mathcal{D}}_{\mathcal{D}}(f)$ hermitsch (bzw. symmetrisch). Es gibt also eine unitäre (bzw. orthogonale) Matrix U, so dass $\tilde{A} := U^{-1}AU$ eine reelle Diagonalmatrix ist.

Wir setzen $\mathcal{B} := \mathcal{D}U$. Dann ist $\mathcal{B}$ wieder orthogonal und die Basistransformationsformel besagt

$$\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(f) = U^{-1}AU = \tilde{A}.$$

□

1. Überblick über algebraische Strukturen

(Aa) **Addition ist assoziativ**
Für alle $x, y, z \in X$ gilt

$$(x + y) + z = x + (y + z).$$

(An) **Addition hat neutrales Element**
Es gibt ein Element $0 \in X$, so dass für alle $x \in X$ gilt

$$x + 0 = 0 + x = x.$$

Man nennt 0 das *neutrale Element der Addition*.

(Ai) **Addition hat inverse Elemente**
Zu jedem $x \in X$ gibt es ein $y \in X$, so dass

$$x + y = y + x = 0.$$

Man nennt y das Inverse von x bezüglich der Addition und schreibt normalerweise $-x$ anstelle von y.

(Ak) **Addition ist kommutativ**
Für alle $x, y \in X$ gilt

$$x + y = y + x.$$

(Ma) **Multiplikation ist assoziativ**
Für alle $x, y, z \in X$ gilt

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

(Mn) **Multiplikation hat neutrales Element**
Es gibt ein Element $1 \in X$, so dass für alle $x \in X$ gilt

$$x \cdot 1 = 1 \cdot x = x.$$

Man nennt 1 das *neutrale Element der Multiplikation*.

(Mi) **Multiplikation hat inverse Elemente**
Zu jedem $x \in X \setminus \{0\}$ gibt es ein $y \in X$, so dass

$$x \cdot y = y \cdot x = 1.$$

Man nennt y das Inverse von x bezüglich der Multiplikation und schreibt normalerweise x^{-1} anstelle von y.

(Mk) **Multiplikation ist kommutativ**
Für alle $x, y \in X$ gilt

$$x \cdot y = y \cdot x.$$

(AMd) **Addition und Multiplikation erfüllen das Distributionsgesetz**
Für alle $x, y, z \in X$ gilt

$$x \cdot (y + z) = x \cdot y + x \cdot z$$
$$(y + z) \cdot x = y \cdot x + z \cdot x$$