Lineare Algebra I

Wintersemester 2021/2022
Albert-Ludwigs-Universität Freiburg
Heike Mildenberger

Fassung vom 13.12.2021

Warnung, das Skript kann Fehler enthalten.

Inhaltsverzeichnis

Kapitel 1

Vektorräume

Quellen: [1], [2], [4], [5], [7], [8], [9] [10], [11], [17].

1.1 Gruppen, Funktionen, endliche Mengen

Wir beginnen mit einem Axiomensystem für Gruppen in einer Notation mit dem neutralen Element:

Definition 1.1. (a) Eine Struktur $(G, \circ, e)$ heißt Gruppe, wenn sie die Gruppenaxiome erfüllt. Diese sind:

(G1) das Assoziativgesetz:
Für all $x, y, z \in G$ gilt $(x \circ y) \circ z = x \circ (y \circ z)$.

(G2) das Gesetz vom neutralen Element e:
Das Element $e \in G$ ist linksneutral, d.h.: Für alle $x \in G$ gilt $e \circ x = x$.

(G3) die Existenz vom Inversen:
Jedes Element hat ein Linksinverses bezüglich e, d.h.: Zu jedem $x \in G$ gibt es $y \in G$, so dass $y \circ x = e$.

(b) Eine Struktur $(G, \circ)$ heißt abelsche[1] Gruppe, wenn sie die Gruppenaxiome erfüllt und zusätzlich kommutativ (auch abelsch genannt) ist. D.h., das Kommutativgesetz gilt: Für alle $x, y \in G$ ist $x \circ y = y \circ x$.

Alternativ kann man e aus der Notation weglassen und nur durch einen Quantor fordern. Dieses Axiomensystem für Gruppen lautet dann so und mutet sich auf den ersten Blick vielleicht schwächer an.

[1] Nils Henrik Abel, 1802 – 1829

Definition 1.2. (a) Eine Struktur $(G, \circ)$ heißt Gruppe, wenn sie die Gruppenaxiome erfüllt. Diese sind:

(G1) das Assoziativgesetz:
Für all $x, y, z \in G$ gilt $(x \circ y) \circ z = x \circ (y \circ z)$.

$(\text{G2})_{\text{alternativ}}$ das Gesetz vom neutralen Element e:
Es gibt ein e, so dass e linksneutral ist, d.h. Für alle $x \in G$ gilt $e \circ x = x$.

$(\text{G3})_{\text{alternativ}}$ die Existenz vom Inversen:
Es gibt ein linksneutrales Element e, so dass
jedes Element ein Linksinverses bezüglich e hat, d.h. zu jedem $x \in G$ gibt es $y \in G$, so dass $y \circ x = e$.

(b) Eine Struktur $(G, \circ)$ heißt abelsche Gruppe, wenn sie die Gruppenaxiome erfüllt und zusätzlich kommutativ (auch abelsch genannt) ist. D.h., das Kommutativgesetz gilt: Für alle $x, y \in G$ ist $x \circ y = y \circ x$.

Wir zeigen, dass e eindeutig ist, wenn in $(G, \circ)$ (G1), $(\text{G2})_{\text{alternativ}}$ und $(\text{G3})_{\text{alternativ}}$ gelten. Meistens quantifiziert man e nicht ab.

Lemma 1.3. *Sei $(G, \circ)$ eine Gruppe und sei e ein linksneutrales Element wie im Gesetz $(G3)_{\text{alternativ}}$, oder sei $(G, \circ, e)$ eine Gruppe und e das Element aus (G3). Dann gilt*

$$\forall a, b \in G,\ a \circ b = e \to b \circ a = e.$$

Beweis: Sei $a \circ b = e$. Wir nehmen ein e-linksinverses Element ℓ_a von a und haben dann

$$\begin{aligned}
e &= \ell_a \circ a \\
&= \ell_a \circ (e \circ a) && \text{(da } e \text{ linksneutral)} \\
&= (\ell_a \circ e) \circ a && \text{(Assoziativges.)} \\
&= (\ell_a \circ (a \circ b)) \circ a && \text{(Voraussetzung eingesetzt)} \\
&= ((\ell_a \circ a) \circ b) \circ a && \text{(assoz.)} \\
&= b \circ a.
\end{aligned}$$

□

Mit ähnlichem Gleichungsjonglieren zeigt man:

Lemma 1.4. *Sei $(G, \circ)$ eine Gruppe.*

(a) Es gibt nur ein linksneutrales Element. Wenn man also e in die Kennzeichnung der Gruppe aufnehmen möchte, hat man nur eine Wahl. Wegen der Eindeutigkeit kann die Kennzeichnung aber auch weggelassen werden. Dieses ist auch rechtsneutrales Element.

(b) Zu jedem a gibt es nur ein Linksinverses b und dieses b ist auch rechtsinvers zu a.

Beweis: (a) Wir zeigen: Jedes linksneutrale Element ist auch rechtsneutral. Sei e linksneutral. Dann ist $a \circ e = a \circ (\ell_a \circ a) = a \circ (r_a \circ a) = (a \circ r_a) \circ a = e \circ a = a$. Seien nun e' und e beides neutrale Elemente. Dann ist $e' = e' \circ e = e$. Die erste Gleichung gilt, da e rechtsneutral ist, und die zweite folgt aus der Linksneutralität von e'.

(b) Seien $b \circ a = c \circ a = e$. Dann multipliziert man beide Gleichungen von rechts mit r_a. Die zweite Behauptung wurde schon im vorigen Lemma gezeigt. □

Wir schreiben von nun an a^{-1} für das Inverse zu a. Da e eindeutig ist, sind beide Schreibweisen $(G, \circ)$ und $(G, \circ, e)$ gleichberechtigt.

Beobachtung 1.5. *(a)* $(a^{-1})^{-1} = a$.

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

Beispiele 1.6. für Gruppen.

(1) Die einelementige Gruppe.

(2) Die zweielementige Gruppe.

(3) Die Gruppen $\mathbb{Z}_n$. $\mathbb{Z}_n$ ist die Menge $\{0, \ldots, n-1\}$ mit der Addition modulo n als Gruppenoperation. Man schreibt für $x \in \mathbb{Z}$ das Zeichen $[x]_n$ für den kanonischen Vertreter der Restklasse von x modulo n, also für das $r \in \{0, \ldots, n-1\}$, so dass es ein $y \in \mathbb{Z}$ gibt mit $x = y \cdot n + r$.

Man definiert $+_n$ durch $[x]_n +_n [y]_n = [x + y]_n$. Dies ist wohldefiniert, und alle Gruppengesetze von $(\mathbb{Z}, +)$ vererben sich auf $(\mathbb{Z}_n, +_n)$. Die Abbildung $x \mapsto [x]_n$ ist ein Gruppenhomomorphismus von $(\mathbb{Z}, +)$ auf $(\mathbb{Z}_n, +_n)$, d.h. eine surjektive Funktion (s.u.), die $[x]_n +_n [y]_n = [x + y]_n$ erfüllt.

In der Tat ist die Relation $\sim_n$ mit $x \sim_n y$, wenn n die Differenz $x - y$ teilt (in Zeichen $n|x - y$) , also wenn $\exists z n \cdot z = x - y$, eine Äquivalenzrelation auf den ganzen Zahlen. Die Restklasse von x modulo n ist $\{y \in \mathbb{Z} : n \text{ teilt} x - y\}$. Genaueres zu Äquivalenzrelationen und -klassen findet man in Abschnitt 1.7. Die Relation $\sim_n$ repektiert $+$, d.h., $z_1 \sim_n z_2$ und $y_1 \sim_n y_2$ impliziert $z_1 + y_1 \sim_n z_2 + y_2$.

(4) Symmetriegruppen von geometrischen Figuren im $\mathbb{R}^2$.

Definition 1.7. Die Ordnung einer Gruppe G ist die Mächtigkeit von G, $|G|$ oder auch $\#G$ geschrieben.

Im Falle endlicher G ist $|G|$ das eindeutig bestimmte n so dass es eine Bijektion von G auf $\{0,\dots,n-1\}$ gibt. Falls $n=0$, ist letzteres die leere Menge. Im Falle unendlicher G, ist $|G|$ eine Kardinalzahl (siehe z.B. [16]). Wir beschränken uns in dieser Vorlesung auf zwei unendliche Mächtigkeiten:

(a) abzählbar unendlich, d.h. es gibt eine Bijektion von G auf $\mathbb{N}$,

(b) und so mächtig wie $\mathbb{R}$.

Für eine exaktere Behandlung des Begriffs "Mächtigkeit" verweisen wir auf spätere Vorlesungen.

Definition 1.8. (und Bemerkungen) Seien X, Y Mengen. Das kartesische Produkt von X und Y ist $X\times Y=\{(x,y)\,:\,x\in X,y\in Y\}$. Sei $f\colon X\to Y$ eine Funktion, das heißt, $f\subseteq X\times Y$ und für jedes $u\in X$ gibt es genau ein $v\in Y$, so dass $(u,v)\in f$. Man schreibt für letzteres eher $f(u)=v$. Die Menge X heißt der Definitionsbereich von f, die Menge Y heißt Zielbereich. Man kann eine Funktion f mit ihren Graphen

$$\{(u,v)\in X\times Y\,:\,f(u)=v\}$$

identifizieren, verliert dabei jedoch die Information über den Zielbereich Y. Er wird ersetzt durch den minimalen Zielbereich

$$\mathrm{bild}(f)=f[X]=\{f(u)\,:\,u\in X\}\subseteq Y$$

oder irgendeine Obermenge der Bildmenge. $\mathrm{bild}(f)$ heißt die Bildmenge von f. Zu einer eindeutigen Zuordnung sagt man auch Operation (besonders auch bei mehrstelligen Funktionen, wenn also $X=X_1\times X_2$, Beispiel $\circ$ auf $G\times G$).

Wichtige Eigenschaften von Funktionen sind:

Definition 1.9. Sei $f\colon X\to Y$ eine Funktion, und sei $g\colon Y\to Z$ eine Funktion.

(1) $f\colon X\to Y$ heißt injektiv, wenn für je zwei $x,y\in X$, aus $f(x)=f(y)$ immer $x=y$ folgt.

(2) $f\colon X\to Y$ heißt surjektiv, wenn $\mathrm{bild}(f)=Y$.

(3) f heißt bijektiv, wenn f injektiv und surjektiv ist.

(4) $g\circ f$ ist die Funktion zuerst f, dann g. Also für $x\in X$, $(g\circ f)(x):=g(f(x))$. $g\circ f\colon X\to Z$.

Die Hintereinanderausführung $\circ$ liefert wichtige nicht-abelsche Gruppen.

Definition 1.10. Sei M eine Menge.

$$S(M) = \{f : M \to M \,:\, f \text{ bijektiv}\}.$$

$(S(M), \circ)$ heißt die symmetrische Gruppe auf M.

Für $n \in \mathbb{N}$ (mit 0) und $S(\{0, \dots, n-1\})$ schreibt man S_n.

Definition 1.11. $0! = 1$, $(n+1)! = n! \cdot (n+1)$. Man sagt "n Fakultät".

Lemma 1.12. *Für $n \in \mathbb{N}$ (mit 0) ist $|S_n| = n!$.*

Beweis: Wir zeigen dies durch vollständige Induktion über $\mathbb{N}$.

Anfang, $n = 0$: $|S(\varnothing)| = |\{\varnothing\}| = 1$.

Induktionsschritt. Seien $|M| = |M'| = n$. Dann gibt es auch $n!$ Bijektionen von M auf M'. Die Anzahl der $f : \{0, \dots, n\} \to \{0, \dots, n\}$ bestimmt man nun, indem man die Möglichkeiten für $f(n)$ zählt, dies sind $n+1$. Bei festem $f(n)$ gibt es $n!$ Bijektionen von $\{0, \dots, n-1\}$ auf $\{0, \dots, n\} \setminus \{f(n)\}$. Also ist

$$|S_{n+1}| = |S_n| \cdot (n+1) = (n+1)!.$$

□

Definition 1.13. Eine Menge heißt endlich, wenn es ein $n \in \mathbb{N}$ und eine Bijektion $f : \{0, \dots, n-1\} \to M$ gibt, also wenn $|M| \in \mathbb{N}$. Andernfalls heißt M unendlich.

Frage 1.14. *Sei M unendlich. Kann man dann $\mathbb{N}$ injektiv in M abbilden? Was meinen Sie dazu?*

Die axiomatischen Grundlagen, zu denen auch das Auswahlaxiom gehört, gestatten eine (inkonstruktive) positive Antwort. Erst etwa ab 1900 einigte sich die mathematische Mehrheit auf diese Auffassung. Die Festlegung der Axiome basiert auf kulturellen Auffassungen. Die der Mathematik zugrundeliegenden Axiome kann man nicht beweisen. Zum Glück besteht seit etwa 1930 Konsens für die Zermelo-Fraenkel'schen Axiome ZFC. Sie können diese im Anhang nachlesen. Die Gesetze für Gruppen, Körper usf. heißen auch manchmal auch Axiome.

1.2 Körper

Definition 1.15. Ein Körper ist eine Struktur $(K, +, \cdot, 0, 1)$ mit folgenden Gesetzen:

(K1) $(K, +, 0)$ ist eine abelsche Gruppe.

(K2) $(K \setminus \{0\}, \cdot, 1)$ ist eine abelsche Gruppe.

(K3) $(a+b)\cdot c = a\cdot c + b\cdot c$. (Hier gilt "Punkt vor Strich".)

Definition 1.16. Verzichtet man in (K2) auf die Kommutativität, so erhält man einen Schiefkörper. Verzichtet an in (K2) auf das Inverse, so erhält man, dass $(K \setminus \{0\}, \cdot, 1)$ nur noch eine sogenannte Halbgruppe mit 1 ist, und nennt dann $(K, +, \cdot, 0, 1)$ einen kommutativen Ring mit 1 . Verzichtet man in (K2) auf die Kommutativität und das Inverse, so erhält man einen Ring mit 1.

Beispiele: Körper $\mathbb{Q}$, $\mathbb{R}$, Gegenbeispiel: $\mathbb{Z}$ hingegen ist nur kommutativer Ring mit Eins.

Rechenregeln 1.17. *in Körpern.*

(1) $-(-a) = a$.

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

(3) $0 \cdot a = 0$.

(4) $a \cdot (-b) = -ab$.

(5) $(ab)c = a(bc)$.

Definition 1.18. Die Definition der komplexen Zahlen durch Cardano[2]. Die Menge der komplexen Zahlen ist $\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$. Wir definieren $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ und $\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ durch:
$(a+ib) + (c+id) = a + c + i(b+d)$
$(a+ib) \cdot (c+id) = ac - bd + i(bc+ad)$.

Definition 1.19. Die Gauß'sche Zahlenebene $\mathbb{C}_g = \{(a,b) : a, b \in \mathbb{R}\}$ Wir definieren $+ : \mathbb{C}_g \times \mathbb{C}_g \to \mathbb{C}_g$ und $\cdot : \mathbb{C}_g \times \mathbb{C}_g \to \mathbb{C}_g$ durch:

$$\begin{aligned} (a,b) + (c,d) &= (a+c, b+d), \\ (a,b) \cdot (c,d) &= (ac - bd, bc + ad). \end{aligned}$$

$f : \mathbb{C} \to C_g$, $f(a+ib) = (a,b)$ ist ein Isomorphismus von $(\mathbb{C}, +, \cdot)$ auf $(\mathbb{C}_g, +, \cdot)$.

Satz 1.20. *Der Fundamentalsatz der Algebra. Jedes nicht konstante Polynom über $\mathbb{C}$ hat eine Nullstelle in $\mathbb{C}$. D.h.: Für alle $n \in \mathbb{N} \setminus \{0\}$, für alle $a_0, \ldots, a_n \in \mathbb{C}$ mit $a_n \neq 0$ gibt es ein $x \in \mathbb{C}$ so dass*

$$\sum_{i=0}^{n} a_i x^i = 0.$$

[2] Gerolamo Cardano, 1501 – 1576

(In der Vorlesung nicht bewiesen!)

Lemma 1.21. *Sei $n \in \mathbb{N}$. Sei $f: \{0,\dots,n-1\} \to \{0,\dots,n-1\}$ injektiv. Dann ist f surjektiv.*

Beweis: Dies zeigt man durch vollständige Induktion über $\mathbb{N}$. Bei Induktionsschritt ändert man die beiden Punkte $(k = f^{-1}(n-1), n-1)$, $(n-1, f(n-1) = \ell)$ auf dem Graphen von f um in (k, ℓ), $(n-1, n-1)$ und erhält so f_{flip}. Man schneidet das Paar $(n-1, n-1)$ aus dem Graphen von f_{flip} weg und erhält so Funktion, auf die die Induktionsvoraussetzung passt, die also $\{0,\dots,n-2\}$ bijektiv auf $\{0,\dots,n-2\}$ abbildet. Dann fügt man $(n-1, n-1)$ hinzu, erhält also, dass f_{flip} bijektiv ist. Zuletzt macht man die Auswechslung wieder rückgängig, erhält also die bikektive Funktion f. □

Definition 1.22. Es seien $n \in \mathbb{N} \setminus \{0\}$, $x, y \in \mathbb{Z}$. Wir setzen $[x]_n \cdot_n [y]_n := [x \cdot y]_n$.

Dies ist wohldefiniert.

Lemma 1.23. *Sei p prim. Dann ist $(\mathbb{Z}_p, +_p, \cdot_p)$ ein Körper.*

Beweis: Sei $[x \cdot y_1]_p = [x \cdot y_2]_p$, für $x, y_i \in \{1,\dots,p-1\}$ und $y_1 < y_2$. Dann gibt es ein $m \geq 1$, so dass $m \cdot p = x \cdot (y_2 - y_1)$. Wir können aufgrund dieser Gleichung $x = x_1 \cdot x_2$ schreiben und $y_2 - y_1 = n_1 \cdot n_2$ mit natürlichen Zahlen aus $\{1,\dots,p-1\}$, so dass $m \cdot p = x_1 \cdot x_2 \cdot n_1 \cdot n_2$ und $m = x_1 \cdot n_1$ und $p = x_2 \cdot n_2$. Letzteres widerspricht der Tatsache, dass p prim ist. Die Annahme $y_1 < y_2$ ist daher falsch, und es folgt $y_1 = y_2$. Da p prim ist, ist die Abbildung $x \mapsto x \cdot_p y$ auf $\mathbb{Z}_p$ injektiv und nach dem Lemma von eben auch surjektiv. Da die 1 also im Bild liegt, findet man ein Linksinverses zu x. □

In der Algebra zeigt man: Es gibt zu jedem $n \geq 1$ Körper mit genau p^n Elementen, und jeder endliche Körper ist von Charakteristik p, p prim, und hat p^n Elemente für ein n.

1.3 Vektorräume

Zu den wichtigsten Strukturen der linearen Algebra gehören die Vektorräume.

Wir treffen die Konvention, dass wir nun griechische Buchstaben als Variablen für die Körperelemente schreiben und lateinische Buchstaben vorerst für Vektoren (und Elemente von Halbordnungen und vieles mehr) stehen.

Definition 1.24. $(V, K, +_K, \cdot_K, +_V, \cdot_s)$ ist ein K-Vektorraum oder Vektorraum über K, wenn folgendes gilt:

(1) $(K, +_K, \cdot_K)$ ist ein Körper. Sei 1 das neutrale Element bezüglich $\cdot_K$.

(2) $(V, +_V)$ ist eine abelsche Gruppe mit neutralem Element 0_V. V heißt die Menge der Vektoren, die Elemente aus V heißen Vektoren.

(3) $\cdot_s : K \times V \to V$ wird die skalare Multiplikation genannt. Sie hat folgende Eigenschaften:

   (a) $\forall \alpha, \beta \in K, \forall v \in V, (\alpha \cdot_K \beta) \cdot_s v = \alpha \cdot_s (\beta \cdot_s v)$.

   (b) $\forall \alpha, \beta \in K, \forall v \in V, (\alpha +_K \beta) \cdot_s v = \alpha \cdot_s v +_V \beta \cdot_s v$.

   (c) $\forall \alpha \in K, \forall v, w \in V, \alpha \cdot_s (v +_V w) = \alpha \cdot_s v +_V \alpha \cdot_s w$.

   (d) $\forall v \in V\ 1 \cdot_s v = v$.

Man schreibt oft nur V oder $(V, K, +_V \cdot_s)$ statt $(V, K, +_K, \cdot_K, +_V, \cdot_s)$, wenn man einige oder gar alle Strukturmerkmale aus dem Kontext ablesen kann. Außerdem schreiben wir nicht immer die Indizes unter + und $\cdot$. Man schreibt statt $\alpha \cdot v$ meistens αv.

Die leere Menge ist kein Vektorraum. Meistens schließt man Strukturen mit leerer Trägermenge aus, denn in diesen ist der unsinnige Satz $\forall x, x \neq x$ wahr, und in allen anderen Strukturen ist er falsch. Die Axiome für Gruppen, Körper, Vektorräume, Halbordnungen, usf. fordern entweder schon als Axiom die Existenz eines bestimmten Elements, oder man fordert, wie bei Halbordnungen, dass eine Halbordnung nicht leer sein soll.

Beispiel 1.25. Beispiele für Vektoräume:

(1) Der Nullraum $\{0\}$. Dies ist die Einermenge des Nullvektors. Man kann jeden Körper nehmen.

(2) Sei K ein Körper, $n \in \mathbb{N}$. K^n, $\mathbb{R}^n$. (Hierbei ist $K^0 = \{0\}$). Die Tupel sind mit komponentenweiser Addition und komponentenweiser skalarer Multiplikation verknüpfbar. So erhält man Vektorräume.

(3) $\mathbb{R}^{(\mathbb{N})} = \{f : \mathbb{N} \to \mathbb{R} : \exists k \forall n \geq k f(n) = 0\}$. Hier wird $f + g$ durch $(f + g)(n) = f(n) + g(n)$ definiert, und auch die skalare Multiplikation ist komponentenweise definiert.

   Der Raum $\mathbb{R}^{(\mathbb{N})}$ ist isomorph zum Vektorraum der $\mathbb{R}[x]$ der reellen Polynome in einer Variablen x. Wir setzen $\operatorname{supp}(f) = \{n : f(n) \neq 0\}$. Man identifiziert $f \in \mathbb{R}^{(\mathbb{N})}$ mit $\sum_{i \in \operatorname{supp}(f)} f(i) \cdot x^i$. Die Vektoraddition ist die Polynomaddition, und analog verfährt man mit der Multiplikation mit Skalaren. Finden Sie eine unendliche linear unabhängige Menge in diesem Raum? (S. Definition 1.36)

(4) $\mathbb{R}^{\mathbb{N}} = \{f : \mathbb{N} \to \mathbb{R}\}$. Wieder wird $f + g$ durch $(f + g)(n) = f(n) + g(n)$ definiert. Zeigen Sie, dass $\mathbb{R}^{\mathbb{N}} \neq \mathbb{R}^{(\mathbb{N})}$. Die Suche nach einer Basis (s. Definition 1.39) in $\mathbb{R}^{\mathbb{N}}$ ist eine Herausforderung.

Definition 1.26. Sei V ein Vektorraum. Eine Teilmenge $U \subseteq V$ heißt Unter(vektor)raum von V, wenn $U \neq \varnothing$ und

$$\forall \alpha, \beta \in K, \forall v, w \in U, \alpha v + \beta w \in U.$$

Lemma 1.27. *Unterräume sind Vektorräume.*

Rechenregeln 1.28. *in Vektorräumen*

(1) $\forall v \in V, 0_K \cdot_s v = 0_V$.

(2) $\forall \alpha \in K, \alpha \cdot_s 0_V = 0_V$.

(3) $\forall v \in V, \alpha \in K, \alpha \cdot_s v = 0 \to (\alpha = 0_K \vee v = 0_V)$.

(4) $\forall v \in V, \alpha \in K, (-\alpha) \cdot_s v = \alpha \cdot_s (-v) = -(\alpha \cdot_s v) = -\alpha v$.

Beispiele 1.29. für Unterräume:

(1) $\{0\}$, V sind Unterräume von V.

(2) Sei $m \leq n$, und sei K ein Körper.

$$\{(x_1, \ldots, x_n) \in K^n : x_{m+1} = \cdots = x_n = 0_K\}$$

ist ein Unterraum von K^n.

(3) $\mathbb{R}^{1000}$ ist ein Unterraum von $\mathbb{R}^{(\mathbb{N})}$, und letzteres ist wiederum ein Unterraum von $\mathbb{R}^{\mathbb{N}}$.

Ein Gegenbeispiel: $\{(x, y) \in \mathbb{R}^2 : x + y < 1\}$ ist kein Unterraum.

Lemma 1.30. *Der Schnitt beliebig vieler Unterräume ist ein Unterraum.*

Wir nehmen nun Skriptbuchstaben für Mengen von Teilmengen von V.
Sei $\mathcal{S}$ eine Menge von Unterräumen von V. Wir schreiben

$$\bigcap \mathcal{S} = \bigcap \{U : U \in \mathcal{S}\} = \bigcap_{U \in \mathcal{S}} U$$

für den Schnitt über $\mathcal{S}$. Falls $\mathcal{S} = \varnothing$, setzen wir $\bigcap \mathcal{S} = V$.

Definition 1.31. Sei V ein Vektorraum und $M \subseteq V$. Wir setzen

(1) $\mathcal{U}_M = \{U : U \text{ Unterraum von } V, U \supseteq M\}$.

(2) $\operatorname{span}(M) = \bigcap \mathcal{U}_M$.

(3) M heißt Erzeugendensystem für/von V wenn $\operatorname{span}(M) = V$.

Beobachtung 1.32. *(1)* *$M \subseteq M' \subseteq V$. Dann ist* $\operatorname{span}(M) \subseteq \operatorname{span}(M')$.

(2) Falls U Unterraum, so ist $\operatorname{span}(U) = U$.

(3) $\operatorname{span}(\operatorname{span}(M)) = \operatorname{span}(M)$.

(4) $\operatorname{span}(\varnothing) = \operatorname{span}(\{0_V\}) = \{0_V\}$.

Satz 1.33. *Sei V ein K-Vektorraum, $M \subseteq V$. Dann ist*

$$\operatorname{span}(M) = \{\sum_{i=0}^{n-1} \alpha_i v_i \,:\, n \in \mathbb{N}, \alpha_i \in K, v_i \in M\}.$$

Hierbei ist für $n = 0$ die leere Summe $\sum_{i=0}^{-1}$ vom Wert 0_V.

Beweis: Wir zeigen, dass die linke Seite als Teilmenge in der rechten enthalten ist: Die rechte Seite ist ein Unterraum, denn es gilt

$$\alpha(\sum_{i=0}^{n-1} \alpha_i v_i) + \beta(\sum_{i=0}^{m-1} \beta_i v_i) = \sum_{i=0}^{\max(m,n)-1} (\alpha\alpha_i + \beta\beta_i) v_i.$$

Daher ist die rechte Seite im Schnitt in der linken Seite auch aufgerufen und also eine Obermenge des Schnitts.

Wir zeigen, dass die rechte Seite in der linken als Teilmenge enthalten ist: Sei U ein beliebiger Unterraum, der M als Teilmenge enthält. Dann enthält U auch alle Summen der Form $\sum_{i=1}^{n} \alpha_i v_i$ mit $v_i \in M$. Da dies für jedes U gilt, ist jede dieser Summen auch in $\bigcap \mathcal{U}_M$. □

Definition 1.34. Ein Term des Typs $\sum_{i=1}^{n} \alpha_i v_i$ heißt Linearkombination. Eine Linearkombination heißt nicht trivial, wenn es ein i in $\{1, \ldots, n\}$ gibt mit $\alpha_i \neq 0$.

Beispiele 1.35. (1) Sei V ein K-Vektorraum. Dann ist in V:

$$\operatorname{span}(\{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}) = K^n.$$

(2) $\operatorname{span}(\{v_0, v_1\}) = \mathbb{R}^2$, falls v_0, $v_1 \in \mathbb{R}^2$ linear unabhängig sind (s.u.).

Definition 1.36. Sei V ein K-Vektorraum.

(1) Sei $n \in \mathbb{N} \setminus \{0\}$, seien $v_1, \ldots, v_n \in V$. Die Vektoren $v_1, \ldots, v_n$ heißen linear unabhängig, wenn

$$\forall \alpha_1 \ldots \forall \alpha_n (\sum_{i=1}^{n} \alpha_i v_i = 0_V \to \alpha_1 = \cdots = \alpha_n = 0).$$

Andernfalls heißen die $v_1, \ldots, v_n$ linear abhängig.

(2) Sei $L \subseteq V$. L heißt linear unabhängig, wenn

$$\forall n \forall v_1 \in L \ldots \forall v_n \in L(\bigwedge_{i \neq j} v_i \neq v_j \to v_1, \ldots, v_n \text{ ist linear unabhängig}).$$

Andernfalls heißt L linear abhängig.

Definition 1.37. Es seien I eine nicht leere Menge und $<$ eine zweistellige Relation auf I. Dann heißt $(I, <_I)$ lineare Ordnung (oder auch totale Ordnung), falls folgendes gilt:

(1) $<_I$ ist irreflexiv, d.h. $\forall i \in I, i \not<_I i$,

(2) $<_I$ ist total, d.h., $\forall i, j \in I, (i <_I j \vee i = j \vee i >_I j)$ und

(3) $<_I$ ist transitiv, d.h., $\forall i, j, k \in I, (i <_I j \wedge j <_I k \to i <_I k)$.

Beobachtung 1.38. *(1) Sei T linear unabhängig und sei $S \subseteq T$. Dann ist S linear unabhängig.*

(2) T ist linear unabhängig, wenn jede endliche Teilmenge $S \subseteq T$ linear unabhängig ist. Man sagt hierzu auch: Die lineare Unabhängigkeit ist „eine Eigenschaft von endlichem Charakter."

(3) Seien L_i, $i \in I$, linear unabhängig, sei $(I, <_I)$ linear geordnet und gelte für $i <_I j$ $L_i \subseteq L_j$. Dann ist $\bigcup_{i \in I} L_i$ linear unabhängig. Man sagt hierzu auch: Die (=jede) „aufsteigende Vereinigung" linear unabhängiger Mengen ist linear unabhängig.

Beweis: (1) und (2) folgen unmittelbar aus der Definition der linearen Unabhängigkeit. (3) Annahme, wie haben n, und paarweise verschiedene $v_i \in \bigcup\{L_j : j \in I\}$, die sich nicht trivial zum Nullvektor kombinieren. Dann gibt es eine Function $f: \{1, \ldots, n\} \to I$, so dass $v_i \in L_{f(i)}$. Die endliche Menge $\{f(i) : i = 1, \ldots, n\}$ hat in der linearen Ordnung $(I, <_I)$ ein größtes Element, d.h. ein $j_{\max} \in \{f(i) : i = 1, \ldots, n\}$, so dass für alle $i = 1, \ldots, n$, $f(i) \leq_I j_{\max}$. Nach der Voraussetzung über das Aufsteigen der L_j, $j \in J$, sind alle v_i, $i = 1, \ldots, n$, Elemente von $L_{j_{\max}}$. Dies widerspricht der linearen Unabhängigkeit von $L_{j_{\max}}$. □

Die Einschränkung auf aufsteigende Vereinigungen ist wichtig, denn die Vereinigung linear unabhängiger Mengen ist i.A. nicht linear unabhängig. Beispiel $M_1 = \{(0,1), (1,0)\}$, $M_2 = \{(1,1)\}$.

1.4 Basen

Definition 1.39. $B \subseteq V$ heißt Basis von V, wenn B ein linear unabhängiges Erzeugendensystem für V ist.

Für $V = \{0\}$ ist $\varnothing$ eine (die einzige) Basis. In jedem Vektorraum V ist die Menge $\{0_V\}$ linear abhängig.

Lemma 1.40. *Es sei $M \subseteq V$.*

(1) Wenn $v \in \operatorname{span}(M)$, so ist $\operatorname{span}(M \cup \{v\}) = \operatorname{span}(M)$.

(2) Wenn $v \in \operatorname{span}(M) \setminus M$, so ist $M \cup \{v\}$ linear abhängig.

(3) Wenn $v \notin M$ und $M \cup \{v\}$ linear abhängig ist mit einem nicht trivialen Vorfaktor für v, so ist $v \in \operatorname{span}(M)$.

Lemma 1.41 (Eine Vorstufe des Austauschlemmas von Steinitz[3])**.** *Sei $A \subseteq V$ linear abhängig und sei $L \subseteq A$ linear unabhängig. Dann gibt es $a \in A \setminus L$, das sich als Linearkombination von Vektoren aus $A \setminus \{a\}$ darstellen lässt.*

Beweis: Da A linear abhängig ist, gibt es eine nicht triviale Linearkombination

$$\sum_{a \in A_0} \alpha_a a = 0$$

mit einem endlichen $A_0 \subseteq A$. Nun ist mindestens ein $a \in A_0 \setminus L$ in dieser Summe mit $\alpha_a \neq 0$. Denn wäre für jedes $a \in A_0 \setminus L$ in dieser Summe $\alpha_a = 0$, dann wäre schon

$$\sum_{a \in L \cap A_0} \alpha_a a = 0$$

eine nicht triviale Linearkombination, im Gegensatz zur linearen Unabhängigkeit von L. Wir halten also ein $a \in A_0 \setminus L$ mit $\alpha_a \neq 0$ fest. Dann ist

$$\sum_{b \in A_0 \setminus \{a\}} \alpha_b b = -\alpha_a a.$$

Wir dividieren beide Seiten durch $-\alpha_a$ und erhalten somit

$$a = \sum_{b \in A_0 \setminus \{a\}} \alpha_b (-\alpha_a)^{-1} b,$$

wie gewünscht. □

Bemerkung 1.42. Für Moduln (d.h. Vektorräume über Ringen mit 1) ist das Lemma falsch. Im (Hauptideal-)Ring $\mathbb{Z}$ ist $2 \cdot 3 - 3 \cdot 2 = 0$, aber wir können die 3 nicht darstellen aus der 2.

Lemma 1.43. *Das Austauschlemma von Steinitz. Sei L_0 linear unahängig. Sei $b \in \operatorname{span}(L_0 \cup \{a\}) \setminus \operatorname{span}(L_0)$. Dann ist $a \in \operatorname{span}(L_0 \cup \{b\})$.*

[3] Ernst Steinitz, 1871–1928

Beweis: $A = L_0 \cup \{a, b\}$ ist linear abhängig und $L = L_0 \cup \{b\}$ ist linear unabhängig und $A \setminus L = \{a\}$. Nach dem vorigen Lemma ist also $a \in \operatorname{span}(A \setminus \{a\}) = \operatorname{span}(L)$. □

Satz 1.44. *Der Austauschsatz von Steinitz. Sei B eine Basis von V mit $n \geq 1$ Elementen. Sei L linear unabhängig. Dann gilt*

(a) $|L| \leq n$, *und*

(b) Es gibt $|L|$ Vektoren $v_1, \ldots v_{|L|}$ aus $B = \{b_1, \ldots, b_n\}$, so dass $L \cup B \setminus \{v_1, \ldots, v_{|L|}\}$ eine Basis ist.

Beweis: Sei $L = \{\ell_1, \ldots, \ell_{|L|}\}$. Nach dem Austauschlemma gibt es ein $b = v_1 \in B$, so dass $\{\ell_1\} \cup (B \setminus \{b\}) = B_1$ zumindest auch b erzeugt und damit ganz V erzeugt. Wir zeigen, dass B_1 eine Basis ist. Es fehlt also noch die lineare Unabhängigkeit.

Es seien $\gamma \in K$, $\beta_c \in K$, für $c \in B_1' = B \setminus \{b\}$, und es sei

$$\gamma \cdot v_1 + \sum_{c \in B_1'} \beta_c \cdot c = 0_V.$$

Wir müssen zeigen, dass $\gamma = 0$ und dass jedes $\beta_c = 0$ ist für $c \in B_1'$. Dies tun wir wie folgt:

Nach Wahl von b haben wir

$$v_1 = \alpha_b \cdot b + \sum_{c \in B_1'} \alpha_c \cdot c$$

mit $\alpha_b \neq 0$. Wir setzen die Gleichung für v_1 in die erste Gleichung ein und erhalten:

$$\gamma\alpha_b \cdot b + \sum_{c \in B_1'} (\gamma\alpha_c + \beta_c)c = 0_v.$$

Nach Voraussetzung ist B linear unabhängig. Daher ist $\gamma\alpha_b = 0$ und für $c \in B_1'$, $\gamma\alpha_c + \beta_c = 0$. Da $\alpha_b \neq 0$, is $\gamma = 0$, und daher sind alle $\beta_c = 0$ für $c \in B_1'$. Somit ist die lineare Unahängigkeit von B_1 gezeigt.

Im zweiten Schritt tauscht man ℓ_2 in B_1 hinein, d.h. man findet ein $v_2 \in B_1$, so dass $B_2 = \{\ell_2\} \cup B_1 \setminus \{v_2\} = \{\ell_1, \ell_2\} \cup B \setminus \{v_1, v_2\}$ eine Basis ist, usf. Nach weiteren $|L| - 2$ Schritten hat man eine Basis $B_{|L|} = L \cup B \setminus \{v_1, \ldots, v_{|L|}\}$. □

Bemerkung 1.45. Der Satz gilt auch für unendliche L und B, und kann mit transfiniter Induktion bewiesen werden. Diese lernt man erst später (oder nie).

Korollar 1.46. *Sei B eine endliche Basis von V. Dann hat jede Basis von V die Mächtigkeit von B.*

Beweis: Man tauscht B in jede andere Basis hinein. Beim Tausch fehlt kein Element und keines ist überzählig. □

Definition 1.47. Sei V ein Vektorraum. Die Dimension von V ist die Mächtigkeit einer (= jeder) Basis von V. Wir schreiben $\dim(V)$ dafür.

Bemerkung 1.48. Die Definition der Dimension ergibt auch für unendliche Basen Sinn. Dies wird zum Beispiel in der Logik-Vorlesung oder in der Mengenlehre-Vorlesung bewiesen.

Die endlichen Mächtigkeiten sind 0, 1, 2, ... Dann gibt es die abzählbar unendliche Mächtigkeit, die auch die Mächtigkeit von $\mathbb{N}$ ist. Danach kommen die Alephs $\aleph_0$, $\aleph_1$, ..., $\aleph_\alpha$, α Ordinalzahl, die wir noch nicht kennen. Die Mächtigkeit von $\mathbb{R}$, $|\mathbb{R}|$ oder $\mathrm{card}(\mathbb{R})$ geschrieben, wird von ZFC nicht auf der Aleph-Skala festgelegt. Die Dimension von $\mathbb{R}^{\mathbb{N}}$ ist $|\mathbb{R}|$. Es ist konsistent relativ zu ZFC, dass $|\mathbb{R}| = \aleph_1$.

Definition 1.49. Eine zweistellige Relation R auf einer Menge A heißt antisymmetrisch, falls für $x, y \in A$, xRy und yRx die Gleichheit $x = y$ impliziert.

Eine zweistellige Relation R auf einer Menge A heißt irreflexiv, falls für $x \in A$, nicht xRx.

Definition 1.50. (1) $(H, \leq_H)$ heißt Halbordnung, falls $\leq_H$ reflexiv, antisymmetrisch und transitiv ist. Man kann auch strikte Halbordnungen definieren: $(H, <_H)$ mir irreflexivem, transitiven $<_H$. Durch Hinzunehmen der Gleichheit gelangt man wieder zum ersten Strukturtypus, und umgekehrt kann man aus einem reflexiven, antisymmetrischen und transitiven $\leq_H$ die Gleichheit wegnehmen und erhält dann eine strikte Halbordnung.

(2) Eine Halbordnung $(H, \leq_H)$ heißt induktive Halbordnung , wenn jede durch $\leq_H$ linear geordnete Teilmenge $K \subseteq H$ (man sagt zu solchen Teilmengen auch Ketten) eine obere Schranke hat, d.h. es gibt $s \in H$, so dass für alle $k \in K$, $k \leq_H s$.

(3) Ein Element m in einer Halbordnung $(H, \leq_H)$ heißt maximales Element wenn $\forall n \in H(n \geq_H m \to n = m)$. (Verwechseln Sie dies nicht mit dem größten Element von weiter unten. In linearen Ordnungen fallen die beiden Begriffe zusammen, aber nicht in Halbordnungen.)

Jede lineare Ordnung $(I, <_I)$ ist eine strikte Halbordnung. Die Umkehrung gilt nicht.

Die lineare Ordnung $(\mathbb{N}, \leq)$ ist keine induktive Halbordnung, denn die Kette $\mathbb{N}$ hat keine obere Schranke.[4]

[4]Im Anhang wird beim Unendlichkeitsaxiom der Begriff „induktive Menge" definiert. $\mathbb{N}$ ist isomorph zu einer induktiven Menge . Die beiden Gebrauchsweisen von induktiv in „induktive Halbordnung" und „induktive Menge" haben nichts miteinander zu tun.

Vorspann: Der Basisergänzungssatz braucht zum Beweis das Lemma von Zorn, dessen Beweis nicht zum Stoff einer Anfängervorlesung gehört. Der Satz über die Existenz von Basen reicht tief in unsere Auffassung der Axiome hinein:

Satz 1.51 (Blass[5], 1984 [3])**.** ZF. *Wenn jeder Vektorraum eine Basis hat, dann gilt das Auswahlaxiom.*

Der Beweis dieses Satzes ist nicht Gegenstand einer Anfängervorlesung. Er könnte bei Interesse einmal in einem Seminar in etwa drei oder vier Sitzungen durchgeführt werden.

Lemma 1.52. *Das Lemma von Zorn[6]. Jede induktive Halbordnung hat ein maximales Element.*

Ich habe unten in den Exkurs einen Beweis aufgeschrieben, falls Sie Interesse haben. Er ist nicht Prüfungsstoff.

Wenn man das Lemma von Zorn akzeptiert, dann kann man die anderen Beweisschritte durchführen:

Satz 1.53. ZFC *Der Basisergänzungssatz von Steinitz (Im Original wohl in [14], aber wir geben einen späteren Beweis mit Hilfe des Lemmas von Zorn.). Sei V ein K-VR und sei A linear unabhängig, und sei E ein Erzeugendensystem. Dann gibt es $E' \subseteq E$, so dass $A \cup E'$ eine Basis bildet.*

Beweis: Wie schreiben

$$\mathcal{H} = \{L : A \subseteq L \subseteq A \cup E \text{ und } L \text{ ist linear unabhängig}\}.$$

Wir ordnen $\mathcal{H}$ mit $\subseteq$. Dann ist $(\mathcal{H}, \subseteq)$ eine induktive Halbordnung.

Wir zeigen dies: Die Relation $\subseteq$ ist antisymmetrisch, reflexiv und transitiv. Sei K eine Kette in $(\mathcal{H}, \subseteq)$. Wir bilden $\bigcup K$. Dann ist $\bigcup K \subseteq V$, $A \subseteq \bigcup K \subseteq A \cup E$. K ist linear unabhängig, denn jede Abhängigkeit bräuchte nur endlich viele Vektoren, und die kämen in einem einzigen Kettenglied schon vor, was der Voraussetzung $K \subseteq \mathcal{H}$ widerspräche. Also ist $\bigcup K$ eine obere Schranke von K.

Nun gibt es nach dem Lemma von Zorn ein maximales Element E' in $\mathcal{H}$. E' ist linear unabhängig, da $E' \in \mathcal{H}$. E' erzeugt V, denn andernfalls gibt es ein $v \in E \setminus \operatorname{span}(E')$. Dann ist $E' \cup \{v\}$ linear unabhängig, denn andernfalls könnte man nach der Vorstufe des Austauschlemmas v als Linearkombination von E' darstellen. $E' \cup \{v\} \in \mathcal{H}$ zeigt also, dass E' nicht maximal ist. □

[5] Andreas Blass, geb. 1947
[6] Max Zorn, 1906 – 1993

1.5 Exkurs: Beweis des Lemmas von Zorn

Dieser Abschnitt reicht meines Erachtens über eine Anfängervorlesung hinaus. Bei Interesse können Sie ihn studieren.

Definition 1.54. Sei A eine Menge. $(A, <)$ heißt Wohlordnung, wenn $(A, <)$ eine lineare Ordnung ist, in der jede nicht leere Teilmenge ein $<$-minimales Element hat.

Eine Menge A hat eine Wohlordnung, wenn es eine Relation $<$ aus A gibt, so dass $(A, <)$ eine Wohlordnung ist.

Definition 1.55. Das Auswahlaxiom sagt: Jede Menge M nicht leerer Mengen hat eine Auswahlfunktion, d.h. eine Funktion $f: M \to \bigcup_{m \in M} m$ mit $f(m) \in m$ für $m \in M$.

Wir beweisen gleich zwei sehr nützliche Äquivalente zum Auswahlaxiom:

Satz 1.56. *Der Wohlordnungssatz und das Lemma von Zorn. Folgende sind äquivalent auf der Basis der Zermelo-Fraenkel-Axiome* ZF *(die Liste kann man im Anhang nachlesen):*

(1) Das Auswahlaxiom.

(2) Jede induktive Halbordnung hat ein maximales Element.

(3) Jede Menge hat eine Wohlordnung.

Beweis

(1) impliziert (3) Dies ist er Zermelo'sche Wohlordnungssatz von 1904 [15]. Der Beweis braucht Ordinalzahlen und den Rekursionssatz für Ordinalzahlen, siehe z.B [6], [16, Kapitel 2]. Diese Gegenstände werden erst in späteren Vorlesungen gelehrt.

Sei A gegeben. Wir konstruieren eine Wohlordnung $\prec$ auf A wie folgt. Wir nehmen eine Auswahlfunktion h auf $\mathscr{P}(A) \setminus \{\varnothing\}$. Dann definieren wir rekursiv über die Klasse On aller Ordinalzahlen mit der $\in$-Ordnung $(\mathrm{On}, \in)$ eine Einbettung $F: \mathrm{On} \to A \cup \{A\}$ wie folgt:

$$F(\alpha) := h(A \setminus \{F(\beta) : \beta \in \alpha\}),$$

falls $\{F(\beta) : \beta \in \alpha\} \neq A$. Falls $\{F(\beta) : \beta \in \alpha\} = A$, dann ist $F(\alpha) := A$. Nach dem Rekursionssatz für On ist F eine wohldefinierte Operation. Da A eine Menge ist und da F injektiv ist, solange der Wert A nicht angenommen wird, gibt es nach dem Ersetzungsschema ein α mit $F(\alpha) = A$. Sei α minimal mit $F(\alpha) = A$. Dann ist $F \restriction \alpha =: f: \alpha \to A$ eine Bijektion (also insbesondere eine Funktion, ein Element des Mengenuniversums), die nun die Wohlordnung von $(\alpha, \in)$ auf A überträgt durch $a \prec b \leftrightarrow f^{-1}(a) \in f^{-1}(b)$. Von der klassengroßen Operation F wird also am Ende nur der mengengroße Anfangsabschnitt $F \restriction \alpha$ gebraucht.

(3) impliziert (1). Sei X eine Menge nicht leerer Mengen. Wir nehmen eine Wohlordnung $\prec$ auf $\bigcup X$. Dann setzen wir für $x \in X$, $f(x) =$ das $\prec$-minimale Element von x.

(1) impliziert (2) [18]. Auch diese Implikation braucht Ordinalzahlen. Sei eine induktive Halbordnung $(H, \leq_H)$ gegeben. Wir konstruieren ein maximales Element von H wie folgt. Wir nehmen eine Auswahlfunktion h auf $\mathscr{P}(A) \setminus \{\varnothing\}$. Dann definieren wir rekursiv über die Klasse aller Ordinalzahlen mit der $\in$-Ordnung $(\mathrm{On}, \in)$ eine Einbettung $F\colon \mathrm{On} \to A \cup \{A\}$ wie folgt:

$$F(\alpha) := h\big((\text{Menge der oberen Schranken von } \{F(\beta) : \beta \in \alpha\}) \setminus \{F(\beta) : \beta \in \alpha\}\big),$$

falls (Menge der oberen Schranken von $\{F(\beta) : \beta \in \alpha\}) \setminus \{F(\beta) : \beta \in \alpha\} \neq \varnothing$, und $F(\alpha) = A$ sonst. Nach dem sogenannten Rekursionssatz für On ist F eine wohldefinierte Operation. Da A eine Menge ist und da F injektiv ist, solange der Wert A nicht angenommen wird, gibt es nach dem Ersetzungsschema ein α mit $F(\alpha) = A$. Sei α minimal mit $F(\alpha) = A$. Da $\{F(\beta) : \beta \in \alpha\} = K$ eine Kette ist, gibt es nur dann keine obere Schranke außerhalb K, wenn K ein größtes Element in K hat. Dann ist α eine Nachfolgerordinalzahl, $\alpha = \beta + 1$, und $F(\beta)$ ist ein maximales Element von H.

(2) impliziert (1). Sei Y eine Menge nicht leerer Mengen. Wir nehmen

$$H = \{(X, h) : X \subseteq Y, h \text{ Auswahlfunktion auf } X\},$$

und halbordnen mit $(X, h) \leq_H (X', h')$ wenn $X \subseteq X'$ und $h' \restriction X = h$. $(H, \leq_H)$ ist eine induktive Halbordnung, denn jede Kette hat als eine obere Schranke die Vereinigung der Kette. Nach dem Zorn'schen Lemma gibt es ein maximales Element (M, h). Man rechnet nach, dass wegen der Maximalität von M die Gleichheit $M = Y$ gilt. h ist also eine Auswahlfunktion auf Y. □

1.6 Der Verband der Unterräume

Definition 1.57. Sei $(H, \leq)$ eine Halbordnung.

(1) $m \in H$ heißt größtes Element, wenn $\forall h \in H h \leq m$. $k \in H$ heißt kleinstes Element, wenn $\forall h \in H h \geq k$.

(2) Sei $A \subseteq H$. Mit $\sup(A)$ (das nicht zu existieren braucht) bezeichnen wir die kleinste obere Schranke von A, also das Element s, so dass

$$\forall a \in A a \leq s \wedge \forall b(\forall a \in A b \geq a \to b \geq s).$$

Mit $\inf(A)$ (das nicht zu existieren braucht) bezeichnen wir die größte untere Schranke von A, also das Element i, so dass

$$\forall a \in A a \geq i \wedge \forall b(\forall a \in A b \leq a \to b \leq i).$$

(3) Wir schreiben $\sup(a,b)$ für $\sup(\{a,b\})$.

Definition 1.58. Eine Halbordnung mit größtem und mit kleinstem Element heißt Verband, wenn je zwei Elemente ein Supremum und ein Infimum haben.

Definition 1.59. Sei V ein Vektorraum, und seien U_1, U_2 Unterräume von V. Wir setzen

$$U_1 + U_2 = \{u_1 + u_2 : u_1 \in U_1, u_2 \in U_2\}.$$

Beobachtung 1.60. $U_1 + U_2 = \operatorname{span}(U_1 \cup U_2)$.

Satz 1.61. *Sei V ein Vektorraum. Dann ist die durch Inklusion geordnete Menge der Unteräume von V ein Verband. Dabei ist* $\sup(U_1,U_2) = U_1 + U_2$, $\inf(U_1,U_2) = U_1 \cap U_2$.

Bemerkung 1.62. Das Distributivgesetz $(U_1+U_2)\cap U_3 = U_1\cap U_3+U_2\cap U_3$ ist im Allgemeinen ab Dimension 2 falsch.

Lemma 1.63. *Sei V ein Vektorraum. Das Modularitätsgesetz: Sei $U_2 \subseteq U_3$. Dann ist*

$$(U_1 + U_2) \cap U_3 = U_1 \cap U_3 + U_2.$$

Man sagt hierzu: Das Modularitätsgesetz gilt im Verband der Unterräume.

Beweis: "$\subseteq$": Sei $x = u_1 + u_2$, $u_i \in U_i$ und $x \in U_3$. Da $u_2 \in U_3$ ist, ist somit $u_1 = x - u_2 \in U_3$. "$\supseteq$": Wenn $u_1 \in U_3$, so ist $x \in U_3$. □

Definition 1.64. (1) Sei V ein Vektorraum. U ein Unterraum von V. Ein Unterraum U' heißt zu U komplementärer Unterraum, wenn $U \cap U' = \{0\}$ und $U + U' = V$.

(2) Sei $(H, \leq_H)$ ein Verband, $h \in H$. k heißt Komplement von/zu h, falls $\sup(h,k)$ das größte Element ist und $\inf(h,k)$ das kleinste Element im Verband ist.

Satz 1.65. ZFC *Es sei V ein Vektorraum. Jeder Unterraum U hat einen komplementären Unterraum U' (der im Falle $V \neq U \neq \{0\}$ nicht eindeutig ist).*

Beweis: Man nimmt eine Basis B von U und ergänzt diese mit dem Basisergänzungssatz durch Hinzufügen von B' zu einer Basis $B \cup B'$ von V. Dann ist $U' = \operatorname{span}(B')$ ein komplementärer Unterraum. □

Lemma 1.66. *Es seien V ein endlichdimensionaler Vektorraum, U ein Unterraum und U' ein komplementärer Unterraum zu U. Dann ist*

$$\dim(V) = \dim(U) + \dim(U').$$

Beweis: Man addiert die Mächtigkeiten einer Basis B von U und einer Basisergänzung B'. □

Bemerkung 1.67. Das Lemma gilt auch für unendlichdimensionale Vektorräume, wird dann mit Kardinalzahlen und der kardinalen Addition formuliert.

Wir verlassen den Verband der Unterräume in der folgenden Definition, die ausgehend von zwei Vektorräumen deren direkte Summe (in der Strukturklasse der Vektorräume) definiert. Der neue Menge der Vektoren ist ein kartesisches Produkt. Dennoch ist diese Konstruktion näher an einer Summe als an einem Produkt, wie wir gleich begründen werden.

Definition 1.68. Seien V_1 und V_2 zwei K-Vektorräume. Dann definieren wir folgende Struktur:

$$V = V_1 \bigoplus V_2 = (V_1 \times V_2, K, +_K, \cdot_K, +_V, \cdot_s)$$

durch

$$\begin{aligned} (u_1, u_2) +_V (v_1, v_2) &:= (u_1 +_{V_1} v_1, u_2 +_{V_2} u_2), \\ \alpha \cdot_s (v_1, v_2) &:= (\alpha v_1, \alpha v_2). \end{aligned}$$

$V_1 \oplus V_2$ heißt die direkte Summe der Vektorräume V_1, V_2. Die direkte Summe ist nicht kommutativ.

Der Begriff in der folgenden Definition wird auch Gegenstand des zweiiten Kapitels sein.

Definition 1.69. Seien $V_1 = (V_1, K, +, \cdot, +_{V_1}, \cdot_{s,V_1})$ und $V_2 = (V_2, K, +, \cdot, +_{V_2}, \cdot_{s,V_2})$ zwei K-Vektorräume.

(1) Eine Abbildung $f: V_1 \to V_2$ heißt lineare Abbildung oder auch Vektorraumhomomorphismus, wenn

$$\forall \alpha, \beta \in K \forall v, w \in V_1, f(\alpha \cdot_{s,V_1} v +_{V_1} \beta \cdot_{s,V_1} w) = \alpha \cdot_{s,V_2} f(v) +_{V_2} \beta \cdot_{s,V_2} f(w).$$

(2) Eine Abbildung $f: V_1 \to V_2$ heißt Vektorraumisomorphismus, wenn f bijektiv und linear ist. Wir schreiben $V_1 \cong V_2$ oder $f: V_1 \xrightarrow{\cong} V_2$.

Lemma 1.70. *Sei V ein Vektorraum, seien U_1, U_2 Unterräume, und sei $U_1 \cap U_2 = \{0\}$. Dann gilt*

$$\begin{aligned} f \quad : \quad & U_1 \times U_2 \to U_1 + U_2, \\ & (u_1, u_2) \mapsto u_1 + u_2 \end{aligned}$$

ist bijektiv und ein Vektorraumisomorphismus von $U_1 \oplus U_2$ auf $U_1 + U_2$.

Die eben angegebene Abbildung f dient als bezeugender Isomorphismus im folgenden Korollar.

Korollar 1.71. *Wenn V ein Vektorraum ist und U ein Unterraum und U' ein zu U in V komplementärer Unterraum ist, dann ist*

$$V = U + U' \cong U \bigoplus U'.$$

Korollar 1.72. *Wenn V ein Vektorraum ist und U ein Unterraum. Dann sind je zwei zu U komplementäre Unterräume isomorph.*

Beweis: Seien U' und U'' zu U komplementäre Unterräume. Nach dem Lemma 1.70 gibt es $f: U \times U' \cong V$ und $f': V \cong U \times U''$. Die Hintereinanderschaltung ist von der Form $f' \circ f: U \times U' \cong U \times U''$. Wir beobachten, dass $(f' \circ f)(u, 0) = (u, 0)$ für $u \in U$. Nun schränken wir diese lineare Abbildung ein auf die Argumente der Form $(u, u') = (0_V, u')$ für $u' \in U'$. Dann haben wir $f' \circ f \restriction \{0\} \times U' : \{0\} \times U' \to \{0\} \times U''$. Die eingeschränkte lineare Abbildung $u' \mapsto (f' \circ f)(0, u')$ bildet daher U' isomorph auf U'' ab. □

Quotientenräume

Definition 1.73. Sei M eine Menge. Eine Menge $R \subseteq M \times M$ heißt Relation auf M.

1. R ist reflexiv, wenn für alle $m \in M$ mRm.
2. R ist symmetrisch, wenn für alle $m, n \in M$ mit mRn auch nRm gilt.
3. R heißt transitiv, wenn für alle $m, n, r \in M$ gilt. Wenn mRn und nRr, so mRr.
4. R heißt Äquivalenzrelation, wenn sie reflexiv, symmetrisch und transitiv ist.
5. Sei R eine Äquivalenzrelation auf M. $[m]_R = \{n \in M : nRm\}$ heißt die (R-)Äquivalenzklasse von m.
6. $M/R = \{[m]_R : m \in M\}$ heißt die Quotientenmenge (von R).
7. $S \subseteq M$ heißt Repräsentantensystem für M/R, falls

$$\forall m \in M, \exists^{=1} s \in S, sRm.$$

Lemma 1.74. *Wenn R eine Äquivalenzrelation ist, so gilt $[m]_R = [n]_R$ oder $[m]_R \cap [n]_R = \emptyset$.*

Satz 1.75. ZFC. *Jede Äquivalenzrelation hat ein Repräsentantensystem.*

Beweis: Man nimmt eine Auswahlfunktion $f: \{[m]_R : m \in M\} \to M$. Das Bild $\{f([m]_R) : m \in M\}$ ist ein Repräsentantensystem. Hierbei ist die Konvention wichtig, dass Mengen (nach dem Extensionalitätsaxiom, siehe Liste der Axiom im Anhang) keine Wiederholungen enthalten. □

Der folgende Typ von Äquivalenzrelationen ist zur Beschreibung von Vektorräumen nützlich.

Definition 1.76. Sei V ein Vektorraum und sei U ein Unterraum.

(1) Wir definieren für $v, w \in V$:
$$v \sim_U w :\Leftrightarrow v - w \in U.$$

(2) $v + U := \{v + u : u \in U\}$ nennt man Nebenklasse (von v bezüglich U).

(3) $V/U = \{[v]_{\sim_U} : v \in V\}$ heißt die Menge der Nebenklassen oder die Quotientenmenge.

(4) $\pi: V \to V/U$, $\pi(v) = [v]_{\sim_U}$ heißt die kanonische Projektion.

Beobachtung 1.77. *Sei V ein Vektorraum und sei U ein Unterraum, $v \in V$.*

(1) $[v]_{\sim_U} = v + U$.

(2) Die Relation $\sim_U$ ist eine Äquivalenzrelation. Daher sind je zwei Nebenklassen entweder gleich oder disjunkt.

Satz 1.78. *V/U trägt eine K-Vektorraumstruktur, die eindeutig dadurch bestimmt ist, dass die kanonische Projektion*
$$\pi: V \to V/U$$
linear ist.

Beweis: Wir definieren $0_{V/U} = [0_V]_{\sim_U} = U$ und $\alpha([v]_{\sim_U} + [v']_{\sim_U}) = [\alpha v + \alpha v']_{\sim_U}$ und rechnen nach, dass die so definierte skalare Multiplikation und Vektoraddition die Menge V/U zu einen Vektorraum machen, der via π ein homomorphes Bild von V ist. □

Das folgende Lemma greift vor und ist ein Spezialfall des Satzes 2.15 aus dem folgenden Kapitel 2.

Definition 1.79. Wenn $f: X \to Y$ eine Funktion ist, und $Z \subseteq X$, so schreibt man $f \restriction Z$ für die Funktion $g: Z \to Y$, die $g(z) = f(z)$ erfüllt. $f \restriction Z$ heißt die Einschränkung (oder Restriktion) von f auf Z.

Lemma 1.80. *Sei U' ein Komplement von U in V. Dann induziert die kanonische Projektion $\pi: V \to V/U$ via Einschränkung einen Isomorphismus*

$$\pi' = \pi \restriction U' : U' \xrightarrow{\cong} V/U.$$

Beweis: Da $U \cap U' = \{0\}$, ist $\pi \restriction U'$ injektiv. Da $U + U' = V$, ist $\pi \restriction U'$ surjektiv. □

Bemerkung 1.81. und *Denkaufgabe*: U' ist also ein Repräsentantensystem von V/U. Sei $\dim(V) \geq 2$. Dann gibt es einen Unterraum U von V, so dass V/U ein Repräsentantensystem hat, das kein Unterraum ist.

Korollar 1.82. *Wenn V endlichdimensional ist, dann ist* $\dim(V) = \dim(U) + \dim(V/U)$.

Beweis: Lemma 1.70 und Lemma 1.80. Isomorphe Vektorräume haben dieselbe Dimension. □

Das Korollar gilt auch für unendlichdimensionale Vektorräume.

Definition 1.83. $\operatorname{codim}_V(U) = \dim(V/U)$ heißt die Kodimension von U in V.

Satz 1.84. *Sei V endlichdimensional, und seien U_1 und U_2 Unterräume. Dann ist* $\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2)$.

Beweis: Es ist $U_2/U_1 = U_2/(U_1 \cap U_2)$. Nach Korollar 1.82 ist $\dim(U_2/U_1) + \dim(U_1 \cap U_2) = \dim(U_2)$ und $\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2/U_1)$. □

Der Satz gilt auch für unendlichdimensionale Vektorräume.

Ab hier bis zum Kapitelende den Stoff nicht vorgetragen. Stoff bis hierher bis auf den Exkurs über das Lemma von Zorn im Zeitraum vom 19.10. bis zum 15.11.2021 vorgetragen.

Definition 1.85. Sei V ein Vektorraum und sei H ein Unterraum. H heißt Hyperebene, wenn es einen eindimensionalen komplementären Unterraum H' gibt mit $V = H + H'$.

Bemerkung 1.86. Im Endlichdimensionalen ist die Definition äquivalent zu $\dim(H) = \dim(V) - 1$. Im Unendlichen gibt es nur die Addition von Kardinalzahlen, keine Subtraktion, daher ist die obige Definition in mehr Situationen geeignet.

Wie Satz 1.65 beweist man die folgende Verfeinerung des Satzes:

Satz 1.87. ZFC *Sei V ein Vektorraum. Seien U und U' Unterräume und $V = U + U'$. Dann hat U hat einen komplementären Unterraum U'', so dass $U'' \subseteq U'$.*

Im Spezialfall, dass U' endlichdimensional ist, kommt man ohne Auswahl aus:

Satz 1.88. ZF *Sei V ein Vektorraum. Sei H ein Unterraum und U' ein endlichdimensionaler Unterraum und $V = H + U'$. Dann hat H hat einen komplementären Unterraum U'', so dass $U'' \subseteq U'$.*

Beweis: Da $H + U' = V$, gilt: U'/H erzeugt V/H. Wir nehmen ein linear unabhängiges Erzeugendensystem B für $\subseteq U'/H$. Dieses ist endlich. Dann nehmen wir Vertreter für $b+H$ in U' für jedes $b+H \in B$. Die endliche Menge dieser Vertreter spannt einen Unterraum U'' mit den gewünschten Eigenschaften auf. □

Korollar 1.89. ZF *Sei V ein Vektorraum und H eine Hyperebene, und sei U ein Unterraum. Dann ist entweder $U \subseteq H$, oder $U \cap H$ ist eine Hyperebene in U.*

Beweis: Falls $U \subseteq H$, so ist $U \cap H = U$, und U ist keine Hyperebene in U. Falls $U \not\subseteq H$, so ist $U/H = H'/H$ für jedes Komplement H' von H. Wir nehmen nun mit dem vorigen Satz ein Komplement H' von H, so dass $H' \subseteq U$. Dann gilt nach dem Modularitätsgesetz, angewandt auf, $H + H' = V$, die Gleichung $(H + H') \cap U = H \cap U + H' = U$. Somit ist H' ein eindimensionales Komplement von $U \cap H$ in U. □

Beispiel 1.90. Sei $V = K^n$, $n \in \mathbb{N} \setminus \{0\}$, und seien $\alpha_i \in K$ und sei $\alpha_i \neq 0$ für ein i. Dann ist

$$H = \{(x_1, \ldots, x_n) : \sum_{i=1}^{n} \alpha_i x_i = 0\}$$

eine Hyperebene im Raum K^n.

1.7 Weitere Beispiele zu (Äquivalenz-)Relationen

Definition 1.91. Seien M, N Mengen. $R \subseteq M \times N$ heißt Relation auf (M, N). Falls $M = N$, so sagt man Relation auf M. Wir schreiben auch xRy für $(x, y) \in R$

Beispiel 1.92. Sei $f: M \to N$ eine Funktion. Dann ist $\mathrm{graph}(f) = \{(x, y) \in M \times N : f(x) = y\}$ eine Relation.

Beispiel im Beispiel: $f(x) = x^2$, definiert auf $\mathbb{R}$. Dann ist der Graph von f die folgende Relation: $R = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x = y^2\}$.

Beispiel 1.93. $C = \{f: \mathbb{N} \to \{0, 1\}\}$ heißt die Cantormenge. $R = \{(f, g) \in C \times C : \exists k \forall n \geq k f(n) = g(n)\}$ ist eine Äquivalenzrelation auf C. Jede Äquivalenzklasse ist abzählbar. C und C/R sind beide nicht abzählbar.

Wir wiederholen noch ein Beispiel aus der Zahlentheorie.

Definition 1.94. Seien p, $m \in \mathbb{Z}$.

(1) p teilt $m :\Leftrightarrow \exists z \in \mathbb{Z} p \cdot z = m$. Wir schreiben $p \mid m$ für p teilt m .

(2) $m \equiv n \mod (p) :\Leftrightarrow p \mid m - n$.

(3) $[x]_p = \{y : y \equiv x \mod (p)\}$.

(4) $\mathbb{Z}_p = \{[x]_p : x \in \mathbb{Z}\}$. Dieses ist eine Gruppe (mit geeignet definierter Addition), und im Falle primer p auch ein Körper.

Definition 1.95. $\mathcal{Z} \subseteq \mathscr{P}(M)$ heißt Zerlegung oder Partition von M wenn $\bigcup \mathcal{Z} = M$, $\varnothing \notin \mathcal{Z}$, und $\forall Z, Z' \in \mathcal{Z} Z \cap Z' = \varnothing$.

Satz 1.96. *Sei M eine Menge. Korrespondenz von Zerlegungen von M und Äquivalenzrelationen auf M. Es gibt*

$$F: \{\mathcal{Z} : \mathcal{Z} \text{ Zerlegung auf } M\} \to \{R : R \text{ Äquivalenzrelation auf } M\}$$

und

$$G: \{R : R \text{ Äquivalenzrelation auf } M\} \to \{\mathcal{Z} : \mathcal{Z} \text{ Zerlegung auf } M\},$$

so dass für alle R und $\mathcal{Z}$, $F(G(R)) = R$ und $G(F(\mathcal{Z})) = \mathcal{Z}$.

Beweis: $xF(\mathcal{Z})y \Leftrightarrow \exists Z \in \mathcal{Z} x, y \in Z$. $G(R) = M/R$. □

Kapitel 2

Lineare Abbildungen

2.1 Grundlegende Eigenschaften

Wir wiederholen Def. 1.69.

Definition 2.1. Seien V und W K-Vektorräume. $f: V \to W$ heißt lineare Abbildung oder Vektorraumhomomorphismus, wenn

$$\forall \alpha, \beta \in K \forall v_1, v_2 \in V f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2).$$

Im Fall $V = W$ nennt man solch ein f auch Vektorraumendomorphismus. Falls f bijektiv und linear ist, heißt f auch Vektorraumisomorphismus.

Definition 2.2. Sei $f: X \to Y$ eine Abbildung.

(1) Sei $X' \subseteq X$. Dann heißt $f \restriction X' : X' \to Y$ mit $(f \restriction X')(x') = f(x')$ für $x' \in X'$ die Einschränkung von f auf X'.

(2) Gelte $\forall y \in Y \exists^{=1} x \in X f(x) = y$. Dann schreiben wir f^{-1} für die Umkehrabbildung. $f^{-1}(y) = x$ gdw $f(x) = y$.

(3) Sei $Z \subseteq X$. $f[Z] := \{f(z) : z \in Z\} (= \mathrm{Im}(f \restriction Z))$ heißt das f-Bild von Z.

(4) Sei $Z \subseteq Y$. $f^{-1}[Z] := \{x : f(x) \in Z\}$ heißt die Urbildmenge von Z (unter f). Im Falle bijektiver f ist $f^{-1}[Z] = \{f^{-1}(z) : z \in Z\}$.

Die Menge $f^{-1}[\{z\}]$ heißt auch die z-Faser unter f.
Dann ist inbesondere für injektives $f: X \to Y$ und $z \in f[X]$, $f^{-1}[\{z\}] = \{f^{-1}(z)\}$.

Definition 2.3. Seien V, W K-Vektorräume. Sei $f: V \to W$ eine lineare Abbildung.

(1) $\mathrm{Im}(f) = \{f(v) : v \in V\}$ heißt der Bildraum von f.

(2) $\ker(f) = \{v \in V : f(v) = 0_W\}$ heißt der Kern von f.

Definition 2.4. Seien V, W K-Vektorräume.

(1) $\mathrm{Hom}(V,W) = \{f: V \to W : f \text{ linear}\}$.

(2) $\mathrm{End}(V) = \mathrm{Hom}(V,V)$ heißt die Menge (später Algebra mit $+$, $\cdot_s$ und $\circ$) der Endomorphismen von V.

(3) $V^* = \mathrm{Hom}(V,K)$ heißt der Dualraum zu V.

Beispiele 2.5. für lineare Abbildungen.

(1) Integration. Sei $V = \{f: [a,b] \to \mathbb{R} : f \text{ stetig}\}$. Wir definieren $L: V \to V$ via $L(f)(x) = \int_a^x f(y)dy$.

(2) Differentiation. Sei $V = \{f: [a,b] \to \mathbb{R} : f \text{ stetig differenzierbar}\}$, und sei $W = \{f: [a,b] \to \mathbb{R} : f \text{ stetig}\}$. Wir definieren $L: V \to W$ via $L(f) = f'$.

Lemma 2.6. *Sei $f: V \to W$ eine lineare Abbildung, und sei $f(a) = b$. Dann ist $f^{-1}[\{b\}] = a + \ker(f)$.*

Lemma 2.7. *Sei $f: V \to W$ eine lineare Abbildung.*

(1) $\ker(f)$ *ist ein Unterraum von V.*

(2) $\mathrm{Im}(f)$ *ist ein Unterraum von W.*

Lemma 2.8. *Sei $f: V \to W$ eine lineare Abbildung. Dann ist* $\ker(f) = \{0_V\}$*, gdw f injektiv ist.*

Lemma 2.9. *Sei $f: V \to W$ eine lineare Abbildung.*

(1) Sei Z ein Unterraum von W. $f^{-1}[Z]$ ist ein Unterraum von V.

(2) Sei Z ein Unterraum von V. $f[Z]$ ist ein Unterraum von W.

Lemma 2.10. *Sei $f: V \to W$ eine lineare Abbildung. Dann ist* $\ker(f) = \{0_V\}$ *gdw f linear unabhängige Mengen auf linear unabhängige Mengen abbildet.*

Beweis: Von der rechten Seite auf die linke Seite: Wir beweisen die Kontraposition. Falls $x \neq 0$ und $x \in \ker(f)$, dann ist $\{x\}$ linear unabhängig und $\{f(x)\} = f[\{x\}]$ linear abhängig. Also ist die rechte Seite nicht wahr.

Von der linken Seite auf die rechte Seite. Wieder arbeiten wir mit der Kontraposition. Falls f eine linear unabhängige Menge L auf eine linear abhängige Menge abbildet,

heißt dies, es gibt paarweise verschiedene $\ell_1, \dots, \ell_n \in L$, und ein nicht triviales n-Tupel $(\alpha_1, \dots, \alpha_n)$, so dass

$$\sum_i \alpha_i f(\ell_i) = 0 = f(\sum_i \alpha_i \ell_i).$$

Dann ist $\sum_i \alpha_i \ell_i \neq 0$, da die ℓ_i linear unabhängig sind und $(\alpha_1, \dots, \alpha_n)$ nicht trivial ist. Also haben wir $0 \neq \sum_i \alpha_i \ell_i \in \ker(f)$. □

Lemma 2.11. *Sei $f: V \to W$ eine lineare Abbildung, $M \subseteq V$. Dann ist* $\mathrm{span}(f[M]) = f[\mathrm{span}(M)]$.

Beweis: f kommutiert mit Linearkombinationen. □

Korollar 2.12. *Sei $f: V \to W$ eine injektive lineare Abbildung, B eine Basis von V. Dann ist $f[B]$ eine Basis von $f[V]$. Falls f bijektiv ist, ist $f[B]$ eine Basis von W.*

Korollar 2.13. *Sei $f: V \to W$ eine lineare Abbildung. Folgende sind äquivalent:*

(1) f ist bijektiv.

(2) Für jede Teilmenge B von V gilt: B ist Basis von V gdw $f[B]$ eine Basis von W ist.

(3) Es gibt eine Basis B von V, so dass $f[B]$ eine Basis von W ist.

Beweis: (1) impliziert (2). Da f injektiv ist, ist das Bild von B linear unabhängig. Da f surjektiv ist, erzeugt das Bild von f den Vektorraum W. Daher gilt $W = \mathrm{Im}(f) = f[V] = f[\mathrm{span}(B)] = \mathrm{span}(f[B])$.

(2) impliziert (3). Es gibt eine Basis B von V. Wir lesen (2) und erhalten dann, dass $f[B]$ eine Basis von W ist.

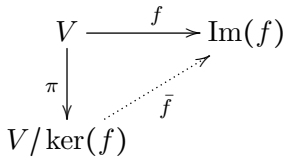
(3) impliziert (1). Da $f[B]$ linear unabhängig ist, ist f injektiv. Da $f[B]$ den Raum W erzeugt, ist f surjektiv. □

Definition 2.14. Seien V, W K-Vektorräume. V und W heißen isomorph wenn es eine bijektive lineare Abbildung $f: V \to W$ gibt. Solch eine Abbildung heißt Vektorraumisomorphismus.

Satz 2.15 (Der Noether'sche[1] Isomorphiesatz [13])**.** *Seien V, W K-Vektorräume. Sei $f: V \to W$ eine lineare Abbildung. Dann induziert f einen Isomorphismus $\bar{f}: V/\ker(f) \xrightarrow{\cong} \mathrm{Im}(f)$ via $\bar{f}(v + \ker(f)) = f(v)$.*

[1] Emmy Noether, 1882 – 1953

Man schreibt die Situation auch gerne in einem Diagramm



Hierbei ist $\pi(v) = v + \ker(f)$. Wir haben also $\bar{f} \circ \pi = f$. Man sagt hierzu auch: „f faktorisiert durch π" und „das Diagramm kommutiert."

Wir kombinieren den Isomorphiesatz mit dem Satz über die Dimension von Quotienten 1.82 und erhalten für jede lineare Abbildung eine nützliche Aufspaltung des Definitionsbereichsraumes:

Korollar 2.16. *Wenn V endlichdimensional ist und $f: V \to W$ linear ist, dann ist* $\dim(V) = \dim(\ker(f)) + \dim(\mathrm{Im}(f))$.

2.2 Lineare Abbildungen und Matrizen

Definition 2.17. Sei K ein Körper, und $m, n \in \mathbb{N} \setminus \{0\}$. Eine m-n-Matrix über K ist ein Schema der Form

$$A = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} \end{pmatrix}$$

mit $a_{i,j} \in K$. Man kann A als Funktion von $\{1, \ldots, m\} \times \{1, \ldots, n\} \to K$ auffassen mit $A(i,j) = \alpha_{i,j}$. Kürzere Schreibweisen für A sind auch

$$(\alpha_{i,j})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$$

oder nur $(\alpha_{i,j})_{i,j}$ oder nur $(\alpha_{i,j})$. Wichtig ist, dass es auf die Anordnung ankommt.

Wenn die $a_{i,j}$ paarweise verschieden sind, gibt es $(m \cdot n)!$ Anordnungen auf der Menge $\{\alpha_{i,j} : i = 1, \ldots, m, j = 1, \ldots, n\}$, die die Menge zu einer m-n-Matrix machen.

Definition 2.18. Sei K ein Körper, und $m, , n \in N \setminus \{0\}$.

(1) Die m-n-Nullmatrix ist gegeben durch $\alpha_{i,j} = 0$ für $i = 1, \ldots, m$, $j = 1, \ldots, n$.

(2) Sei nun $m = n$. Die m-m-Einheitsmatrix $1_{M_{m,m}(K)}$ ist gegeben durch $\alpha_{i,j} = \delta_{i,j}$ für $i = 1, \ldots, m$, $j = 1, \ldots, m$. Hierbei ist

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

das Kronecker-Delta[2]

Definition und Behauptung 2.19. Sei K ein Körper, und $m, , n \in \mathbb{N} \setminus \{0\}$. $M_{m,n}(K)$ bezeichnet die Menge der m-n-Matrizen über K. Wir definieren die komponentenweise Addition + und die komponentenweise Sklarenmultiplikation $\cdot_s$ auf $M_{m,n}(K)$ durch $(\alpha_{i,j}) + (\beta_{i,j}) = (\gamma_{i,j})$ mit $\gamma_{i,j} = \alpha_{i,j} +_K \beta_{i,j}$ und $\xi \cdot_s (\alpha_{i,j}) = (\beta_{i,j})$ mit $\beta_{i,j} = \xi \cdot_K \alpha_{i,j}$. So wird

$$(M_{m,n}(K), K, +_K, \cdot_K, +, \cdot_s)$$

zu einem Vektorraum über K, den wir ebenfalls mit $M_{m,n}(K)$ bezeichnen.

Definition und Behauptung 2.20. Sei K ein Körper, V, W K-Vektorräume. Dann ist $\mathrm{Hom}(V, W)$ mit der komponentenweise Addition $(f + g)(v) = f(v) + g(v)$ und der komponentenweise skalaren Multiplikation $(\alpha f)(v) = \alpha f(v)$ ein Vektorraum.

Satz 2.21. *Seien V, W Vektorräume der Dimension n bzw m. Dann ist* $\mathrm{Hom}(V, W)$ *isomorph zu $M_{m,n}(K)$. Für je zwei geordnete Basen $\vec{B} = (b_1, \ldots, b_n)$ von V und $\vec{C} = (c_1, \ldots, c_m)$ von W gilt folgendes: Jedes $f \in \mathrm{Hom}(V, W)$ bestimmt durch*

$$\text{für } j = 1, \ldots, n: \quad f(b_j) = \sum_{i=1}^{m} \alpha_{i,j} c_i \tag{2.1}$$

eine Matrix $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) = A_f = (\alpha_{i,j})$. *Bei festen $\vec{B}$, $\vec{C}$ ist die Abbildung*

$$\begin{aligned} i\colon \mathrm{Hom}(V, W) &\to M_{m,n}(K); \\ f &\mapsto i(f) = A_f \end{aligned}$$

ein Vektorraumisomorphismus.

Wir nennen die Umkehrung dieses Isomorphismus: Seien immer noch $\vec{B}$ und $\vec{C}$ geordnete Basen von V bzw. W. Jede Matrix $A \in M_{m,n}(K)$ bestimmt genau eine lineare Abbildung f_A mit

$$\text{für } j = 1, \ldots, n: \quad f_A(b_j) = \sum_{i=1}^{m} \alpha_{i,j} c_i.$$

Es gilt $A_{f_A} = A$ und $f_{A_f} = f$.

[2] Leopold Kronecker, 1823 – 1891

Definition 2.22. Seien $\ell, m, n \in \mathbb{N} \setminus \{0\}$, und sei $B = (\beta_{i,j}) \in M_{m,n}(K)$ $A = (\alpha_{k,i}) \in M_{\ell,m}(K)$. Wir definieren die Matrizenmultiplikation

$$\begin{aligned} \cdot : M_{\ell,m}(K) \times M_{m,n}(K) &\to M_{\ell,n}(K); \\ (A, B) &\mapsto A \cdot B = AB = C = (\gamma_{\ell,j}) \end{aligned}$$

durch die folgenden Gleichungen

$$\text{für } k = 1, \ldots, \ell, j = 1, \ldots, n \quad \gamma_{\ell,j} = \sum_{i=1}^{m} \alpha_{\ell,i}\beta_{i,j}.$$

Beobachtung 2.23. *1. Die Matrizenmultiplikation ist in beiden Argumenten linear. D.h., für alle $\alpha, \beta\gamma, \delta$, A, $A' \in M_{\ell,m}(K)$, B, $B' \in M_{m,n}(K)$ gilt: $(\alpha A + \beta A')(\gamma B + \delta B') = \alpha\gamma AB + \beta\gamma A'B + \beta\gamma AB' + \beta\delta A'B'$.*

2. Die Matrizenmultiplikation ist assoziativ.

3. Die Matrizenmultiplikation ist nicht kommutativ.

Beweis: Nachrechnen! Bei (3) nimmt man Dimension mindestens 2 und kann mit Scherungsmatrizen arbeiten. □

Satz 2.24. *Seien U, V, W Vektorräume der Dimension n bzw. m bzw. ℓ mit geordneten Basen $\vec{B} = (b_1, \ldots, b_n)$, $\vec{C} = (c_1, \ldots, c_m)$, $\vec{D} = (d_1, \ldots, d_\ell)$. Sei $g: U \to V$ linear und bzgl. $\vec{B}$ und $\vec{C}$ gegeben durch $g = f_B$, $B \in M_{m,n}(K)$. Sei $h: V \to W$ linear und bzg. $\vec{C}$ und $\vec{D}$ gegeben durch $h = f_A$, $A \in M_{\ell,m}(K)$. Dann ist $h \circ g: U \to W$ bzgl. $\vec{B}$ und $\vec{D}$ gegeben durch f_{AB}. Wir haben also $A_{h\circ g} = A_h A_g$, ausführlicher*

$$\mathrm{Mat}_{\vec{B}}^{\vec{D}}(h \circ g) = \mathrm{Mat}_{\vec{C}}^{\vec{D}}(h) \cdot \mathrm{Mat}_{\vec{B}}^{\vec{C}}(g)$$

und $f_{AB} = f_A \circ f_B$.

Beweis: Für $j = 1, \ldots, n$ ist $g(b_j) = f_B(b_j) = \sum_{i=1}^{m} \beta_{i,j} c_i$. Für $i = 1, \ldots, m$ ist $h(c_i) = f_A(c_i) = \sum_{k=1}^{\ell} \alpha_{k,i} d_k$. Nun setzen wir ein und erhalten für $j = 1, \ldots, n$:

$$(h \circ g)(b_j) = h(\sum_{i=1}^{m} \beta_{i,j} c_i) = \sum_{i=1}^{m} \beta_{i,j} \sum_{k=1}^{\ell} \alpha_{k,i} d_k = \sum_{k=1}^{\ell} \sum_{i=1}^{m} \alpha_{k,i}\beta_{i,j} d_k = \sum_{k=1}^{\ell} \gamma_{k,j} d_k.$$

Hierbei haben wir sehr oft Distributivität und Kommutativität, also die Vektorraumgesetze, benutzt, um die Summationszeichen zu vertauschen. Wer skeptisch ist, kann die Korrektheit der vorgenommenen Vertauschung per Induktion über m und ℓ aus den Gesetzen, die über $m = 2$ und $\ell = 2$ sprechen, herleiten. □

Beispiel 2.25. Wir leiten aus Satz 2.21 und der Definition der Multiplikation eine wichtige Anwendung her: Nun nehmen wir den Spezialfall $V = K^n$, $W = K^m$ und die Standardbasen $\vec{B} = (e_1, \dots, e_n)$ mit $e_i = (0, \dots, 0, 1, 0 \dots, 0)$ mit einer 1 an i-ter Stelle, und als Spaltenvektor, also als ein Element von $M_{n,1}$ aufgefasst. $\vec{C} = (e_1, \dots, e_m)$. Sei $f: K^n \to K^m$ gegeben durch A, in anderen Worten $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) = A$. Dann stehen in der j-ten Spalte von A genau das f-Bild von e_j (als Spaltenvektor), d.h.,

$$f(\sum_{j=1}^{n} \xi_j e_j) = A \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}. \tag{2.2}$$

Konvention: Wir fassen die Multiplikation mit $A \in M_{n,m}(K)$ als Abbildung f_A von K^n in den K^m auf, eben als Multiplikation der Matrix von links aus an einen Spaltenvektor nach dem Muster (2.2). Wir schreiben auch φ_A hierfür, wenn $V = K^n$ mit der Standardbasis $\vec{B} = (e_1, \dots, e_n)$ und $W = K^m$ mit der Standardbasis $\vec{C} = (e_1, \dots, e_m)$.

2.3 Lineare Gleichungssysteme

Nun behandeln wir „praktische Rechenfertigkeiten".

Beispiel 2.26. Gesucht sind alle $(x_1, x_2, x_3) \in \mathbb{R}^3$, so dass

$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= 3 \text{ und} \\ x_1 + 5x_2 &= 7 \end{aligned} \tag{2.3}$$

Wir haben also $m = 2$ Gleichungen in $n = 3$ Variablen.

Definition 2.27. Sei K ein Körper. Seien $m, n \in \mathbb{N} \setminus \{0\}$, $\alpha_{i,j}, \beta_i \in K$ für $i = 1, \dots, m, j = 1, \dots, n$. Eine Konjunktion der Form

$$\text{Für } i = 1, \dots, m \quad \sum_{j=1}^{n} \alpha_{i,j} \cdot x_j = \beta_i \tag{2.4}$$

heißt lineares Gleichungssystem mit m Gleichungen in n Unbekannten.

Man kann das Gleichungssystem (2.4) auch in Matrizenmultiplikation schreiben:

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \tag{2.5}$$

mit $A = (\alpha_{i,j})$ und womöglich unterschiedlich "hohen" x und b (der Spaltenvektor aus $\beta_1, \dots, \beta_n$) oder noch ausführlicher

$$\begin{pmatrix} \sum_{j=1}^{n} \alpha_{1,j} x_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{m,j} x_j \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \tag{2.6}$$

Definition 2.28. (Zeilen-)Stufenform für ein lineares Gleichungssystem über einem Körper K mit n Unbekannten und m Zeilen. Eine Anordnung der Form

$$\begin{array}{cccccccccccccccc}
0 & \cdots & 0 & \alpha_{1,j_1} & * & \dots & * & * & * & \dots & * & * & * & \cdots & \alpha_{1,n} & \beta_1 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & a_{2,j_2} & * & \dots & * & * & * & \cdots & \alpha_{2,n} & \beta_2 \\
\vdots & & & & & & & & & & & & & & & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \dots & 0 & \alpha_{\varrho,j_\varrho} & * & \cdots & \alpha_{\varrho,n} & \beta_\varrho \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots & 0 & b_{\varrho+1} \\
\vdots & & & & & & & & & & & & & & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots & 0 & \beta_r \\
\vdots & & & & & & & & & & & & & & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \cdots & 0 & 0
\end{array}$$

heißt Stufenform oder genauer Zeilenstufenform eines linearen Gleichungssystems. Hierbei ist für $i = 1, \dots, \varrho$, $\alpha_{i,j_i} \neq 0$, für $i = \varrho + 1, \dots, r$, $\beta_i \neq 0$, für $i = r + 1, \dots, m$, $\beta_i = 0$ und die Sterne stehen für beliebige Elemente von K. Die Zahl $\varrho \leq m, n$ heißt der Rang des linearen Gleichungssystems. Die Zahl $r \in [\varrho, \min(m, n)]$ heißt der Rang der um die rechte Seite erweiterten Matrix des linearen Gleichungssystems.

Definition 2.29. Ein lineares Gleichungssystem heißt homogen, wenn $\beta_1 = \cdots = \beta_m = 0$.

Jedes homogene lineare Gleichungssystem hat den Nullvektor als Lösung, und es gibt womöglich noch weitere Lösungen.

Definition 2.30. Sei $r \in K$. Wir definieren $G_2 - rG_1$ für zwei lineare Gleichungen G_i der Form $\sum_{j=1}^{n} \alpha_{i,j} x_j = \beta_i$, $i = 1, 2$, als $\sum_{j=1}^{n} (\alpha_{2,j} - r\alpha_{1,i}) x_j = \beta_2 - r\beta_1$. Falls die Anzahl der Variablen nicht übereinstimmt, füllt man mit Vorfaktoren Null auf. Wir nehmen als Beipiel, dass die erste Gleichung nur $m < n$ Variablen hätte. Dann füllt kann man $\sum_{j=1}^{m} \alpha_{1,j} x_j = \beta_1$ durch $\sum_{j=1}^{m} \alpha_{1,j} x_j + \sum_{j=m+1}^{n} 0 \cdot x_j = \beta_1$ auf und kann danach die Gleichung $G_2 - rG_1$ bilden.

Definition 2.31. Sei I ein System von m Gleichungen in n Unbekannten. $L_I = \{(x_1, \dots, x_n) : (x_1, \dots, x_n)$ löst I$\}$

Das folgende Lemma ist der wichtigste Bestandteil des Gauß-Verfahrens:

Lemma 2.32. *Es seien G_1, G_2 lineare Gleichungen und es sei $r \in K$. Aus G_1 und G_2 folgen $G_2 - rG_1$ und G_1. Aus $G_2 - rG_1$ und G_1 folgen G_1 und G_2. Ebenso kann man jede Gleichung durch ein Vielfaches ersetzen mit Faktor ungleich Null.*

Ein Einzelschritt im Gauß-Verfahren: Sei ein LGS I mit den Gleichungen $G_1, \ldots, G_m$ gegeben. Sei $\alpha_{1,1} \neq 0$. Dann bilden wir das LGS I' mit $G_1' = G_1$, $G_i' = G_i - \frac{\alpha_{i,1}}{\alpha_{1,1}} G_1$. Dann gilt $L_I = L_{I'}$. Dies wird nun geschickt iteriert.

Lemma 2.33. *Eine Zeilenvertauschung kann man durch dreimaliges Addieren erreichen.*

Beweis: Seien $i_1 \neq i_2$, $1 \leq i_1, i_2 \leq m$ zwei Zeilenindizes. Man vertauscht Zeile i_1 mit Zeile i_2, indem man

(1) Zeile i_1 durch die Summe der beiden Zeilen ersetzt und Zeile i_2 stehen lässt.

(2) Im nächsten Schritt ändert man die Summenzeile nicht und ersetzt die i_2-te Zeile durch (-1)mal ihr Original plus einmal die jetzige Zeile i_1, die ja die Summenzeile ist. Dann steht jetzt die frühere i_1-Zeile an Stelle i_2, und in der Zeile i_1 steht immer noch die Summe.

(3) Im letzten Schritt zieht man die jetzige i_2-Zeile von der Summenzeile ab.

□

Satz 2.34. *Jedes lineare Gleichungssystem lässt sich durch Zeilenoperationen in ein Gleichungssystem in Stufenform umwandeln, das dieselbe Lösungsmenge hat (zwei solche lineare Gleichungssysteme nennt man äquivalent).*

Beweis: Einen recht formalen Beweis induktiv über $m \cdot n$ findet man im Beweis des Satzes 2.61.

Wir geben nun eine alternative Beschreibung des Gauß-Algorithmus[3] [4] :
Ein Hauptschritt besteht aus der Suche des nächsten sogenannten Pivots (Spitze) und einigen Zeilenoperationen, die alle Einträge in der Spalte unter dem Pivot zu Null machen. Man braucht ϱ (= Rang) von A Hauptschritte und erhält ϱ Pivots, die gerade den Rang bezeugen. Wir nummerieren die Hauptschritte mit Hauptschritt 1, ..., Hauptschritt k, ..., Hauptschritt ϱ. Im Hauptschritt Nummer k sind etwa $3 + (n - k)$ Unterschritte durchzuführen.

Erster Hauptschritt: Wir suchen das kleinste j, so dass in der Spalte j ein Eintrag nicht Null ist, sei der Wert α. Wir nennen den gefundenen Spaltenindex $j = j_1$. Sei $\alpha = \alpha_{i,j_1}$.

[3] Carl Friedrich Gauß, 1777 - 1855

[4] Abu Dscha'far Muhammad ibn Musa al Chwarizmi, 770 - (835-850).

Wir tauschen nun Zeile i mit der ersten Zeile. Danach ziehen wir von Zeile $i = 2, 3, \ldots, m$ jeweils $\alpha_{i,j_1}/\alpha$ mal die erste Zeile ab. Wir erhalten eine Matrix mit genau einem Eintrag ungleich 0 in der j_1-ten Spalte, an der ersten Stelle.

Sei nun $2 \leq k \leq \varrho$. Nach dem Hauptschritt Nummer $(k-1)$ haben wir ein Schema der Form

$$\begin{array}{cccccccccccccccc}
0 & \cdots & 0 & \alpha_{1,j_1} & * & \ldots & * & * & * & \ldots & * & * & * & \cdots & \alpha_{1,n} & \beta_1 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & a_{2,j_2} & * & \ldots & * & * & * & \cdots & \alpha_{2,n} & \beta_2 \\
\vdots & & & & & & & & & & & & & & & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \ldots & 0 & \alpha_{k-1,j_{k-1}} & * & \cdots & \alpha_{k-1,n} & \beta_{k-1} \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \ldots & 0 & 0 & * & \cdots & * & * \\
\vdots & & & & & & & & & & & & & & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \ldots & 0 & 0 & * & \cdots & * & * \\
\vdots & & & & & & & & & & & & & & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \ldots & 0 & 0 & * & \cdots & * & *
\end{array}$$

mit neuen Zahlen $\alpha_{i,j}$ und β_i erreicht.

Im k-ten Hauptschritt suchen wir die am weitestens links stehende Spalte, in der es einen Nichtnull Eintrag in Zeile $\geq k$ gibt. Sei dies Spalte j_k, und heiße der Eintrag wiederum α. Wir nennen ihn auch den k-Pivot. Falls es keinen solchen Eintrag mehr gibt, endet das Verfahren schon in $k-1$ Schritten und es ist $\varrho = k-1$. Wir nehmen daher nun an, dass es so einen Eintrag gibt. Wir tauschen die Zeile mit diesem Eintrag in die k-te Zeile und nennen die neue Zeile Z_k. Wiederum sorgen wir durch Abziehen von $\frac{\alpha_{k+1,j_k}}{\alpha}$ mal Z_k von der Zeile Z_{k+1}, $\frac{\alpha_{k+2,j_k}}{\alpha}$ mal Z_k von Z_{k+2}, usf, bis $\frac{\alpha_{m,j_k}}{\alpha}$ mal Z_k von Z_m. Unterhalb des k-Pivots in der Spalte j_k stehen jetzt nur Nullen. Daraus können weitere Spalten weiter rechts entstehen, die auch nur Nullen in Zeilen $k+1, \ldots, m$ haben. Das ist dann die Breite der Stufe k. □

Mit der Zeilenstufenform kann man nun die Lösungsmenge angeben: Genau dann, wenn es keine echte Stufe in der b-Spalte gibt, d.h., falls $\varrho = r$ und $\alpha_{1,j_1} \neq 0 \ldots \alpha_{r,j_r} \neq 0$ und die b-Spalte ab Höhe $\varrho + 1$ nur Nullen hat, gibt es (mindestens) eine Lösung. Man kann dies auch als Rangkriterium für die um die b-Spalte erweiterte Matrix

$$\big(A|b\big) = \begin{pmatrix} \alpha_{1,1} & \ldots & * & \ldots & \alpha_{1,n} & \beta_1 \\ \vdots & & & & \vdots & \vdots \\ \alpha_{n,1} & \ldots & * & \ldots & \alpha_{n,m} & \beta_m \end{pmatrix} \tag{2.7}$$

fassen.

Korollar 2.35. *Das Gauß-Verfahren liefert die Lösungsmenge in parametrisierter Form: Wir nehmen also an, dass es keine Stufe in der b-Spalte gibt, d.h. $\varrho = r$. Man wählt Werte*

für $x_n, \ldots, x_{j_r+1}$ frei und berechnet dann x_{j_r} mit der letzten relevanten Gleichung. Danach wählt man $x_{j_r-1}, \ldots x_{j_{r-1}+1}$ frei und rechnet $x_{j_{r-1}}$ aus. Man setzt das Verfahren fort, bis man bei x_{j_1} angelangt ist, das in Abhängigkeit von $x_n, \ldots, x_{j_1+1}$ berechnet wird. Die Variablen $x_1, \ldots, x_{j_1-1}$ kann man frei belegen. Die $n-r$ frei wählbaren Parameter sind also

$$x_1, \ldots, x_{j_1-1}, x_{j_1+1} \ldots, x_{j_2-1}, x_{j_2+1}, \ldots, x_{j_3-1}, \ldots, x_{j_r+1}, \ldots, x_n.$$

Die bestimmten Variablen sind

$$x_{j_1}, \ldots, x_{j_r}.$$

Alle so beschriebenen $x \in K^n$ bilden zusammen die Lösungsmenge

$$L = \{(x_1, \ldots, x_n) : \text{ oben genannte Parameter frei gewählt, } x_{j_k} \text{ für } k = 1, \ldots, r \text{ ist durch die } r \text{ Zeilen der Stufenform bestimmt}\}.$$

Immer, auch im Fall $\varrho < r$, ist der Kern der durch A bestimmten linearen Abbildung von Dimension $n - \varrho$.

Korollar 2.36. *Das lineare Gleichungssystem* (2.4) *ist genau dann lösbar, wenn der Rang der Matrix $(A|b)$ gleich dem der Matrix A ist.*

Ab hier Video 1, als Vorlesung für den 29.11.2021

Definition 2.37. Der Zeilenrang eine Matrix $A \in M_{m,n}(K)$ ist die maximale Anzahl an linear unbhängigen Zeilen. Der Spaltenrang von A ist die maximale Anzahl an linear unabhängigen Spalten.

Satz 2.38. *Der Zeilenrang jeder Matrix ist gleich dem Spaltenrang und heißt von nun an Rang.*

Beweis: In Satz 2.46 findet sich ein eleganter Beweis mit dem Noether'schen Isomorphiesatz 2.15.

Wir geben nun einen eigenständigen Beweis:

Der Zeilenrang in Zeilenstufenform ist die Anzahl der Stufen, bei uns oben ϱ. Man sieht an der Zeilenstufenform, dass die Spalten Nummer $j_1, \ldots j_\varrho$ linear unabhängig sind und dass es nicht mehr als ϱ linear unabhängige Spalten geben kann, da das Bild $\{Ax : x \in K^n\}$ durch $e_1, \ldots e_\varrho$ im K^m aufgespannt wird.

Zeilenoperationen ändern den Zeilenrang nicht:
Wie im Beweis des Austauschlemmas rechnet man durch Einsetzen in die Gleichungen für lineare Unabhängigkeit (für Vektoren, die aus Zeilen gebildet werden) nach: Der Zeilenrang ändert sich durch eine einzelne Zeilenoperation des Typs $(G_1, G_2) \mapsto G_1, G_2 - rG_1)$ oder des Typs $r \neq 0$, $G_1 \mapsto rG_1$ nicht.

Zeilenoperationen ändern den Spaltenrang nicht:
Seien $a_{j_1} = (\alpha_{i,j_1})_{1\le i\le m}, \dots, a_{j_k} = (\alpha_{i,j_k})_{1\le i\le m}$ Spalten von A. Die j_k sollen beliebig sein, nicht unbedingt die aus der Zeilenstufenform. Die Gleichung

$$\sum_{\kappa=1}^{k} \xi_\kappa a_{j_\kappa} = 0 \in K^m \tag{2.8}$$

ist ein Teil der Gleichung $Ax = 0$, die ξ_κ sind gerade die x_{j_κ}. Ebenso wie sich die Lösungsmenge von $Ax = 0$ durch Zeilenoperationen nicht ändert, ändert sich auch die Lösungsmenge von Gleichung (2.8) nicht. □

Bemerkung 2.39. Wenn man auch noch Spaltenoperationen zulässt, kann man die Zeilenstufenform noch verbessern zu Stufen der Breite eins. Dies entspricht jedoch einer Vertauschung der Variablen $x_1, \dots, x_n$. Unter Umständen sollte man also darüber Buch führen!

Definition 2.40. Sei $A \in M_{m,n}(K)$ und sei $\vec{B}$ die Standardbasis auf dem K^n und sei $\vec{C}$ die Standardbasis auf dem K^m. Wir definieren $\varphi_A: K^n \to K^m$ via $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(\varphi_A) = A$, also $\varphi_A(x) = Ax$ für $x \in K^n$ und $Ax \in K^m$.

Bemerkung 2.41.

Wir übertragen unsere Resultate über Nebenklassen, Umkehrbarkeit und Isomorphie auf $\varphi_A: K^n \to K^m$, $A \in M_{m,n}(K)$:

Korollar 2.42. *Sei $A \in M_{m,n}(K)$ und sei $H = \{x \in K^n : Ax = 0\}$. Dann gilt*

(1) Für jedes $b \in K^m$ ist $\{x \in K^n : Ax = b\}$ leer oder eine Nebenklasse von H.

(2) Falls $m < n$, so ist H nicht der Nullraum. Falls $m > n$, so hat nicht für jedes b die Gleichung $Ax = b$ eine Lösung.

(3) Falls $m = n$, so sind äquivalent:

- *(a) H ist der Nullraum.*
- *(b) $Ax = b$ ist lösbar für alle $b \in K^m$.*
- *(c) Für alle $b \in K^m$ gibt es höchstens ein $x \in K^m$ mit $Ax = b$.*

Definition und Behauptung 2.43. (1) Falls $A \in M_{m,m}(K)$ und der Fall (3)(a) vorliegt, sagt man A ist regulär oder A ist invertierbar. Wir schreiben A^{-1} für die inverse Matrix. Dann ist $A^{-1}A = AA^{-1} = 1_{M_{m,m}(K)}$. Nicht reguläre quadratische Matrizen heißen auch ausgeartete Matrizen.

(2) $\mathrm{GL}_n(K)$ bezeichnet die Menge (oder auch, mit der Matrizenmultiplikation, die Gruppe) der regulären n-n-Matrizen über K. Die Gruppe $(\mathrm{GL}_n(K), \cdot)$ (mit der Matrizenmultiplikation $\cdot$) wird die allgemeine lineare Gruppe über K genannt, GL steht für general linear.

Bemerkung 2.44. Die Gruppe $(\mathrm{GL}_n(K), \cdot)$ ist nicht abelsch für $n \geq 2$.

Beweis: Das folgende Beispiel lässt sich in jedem Körper finden, auch in $\mathbb{F}_2$.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \tag{2.9}$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \tag{2.10}$$

□

Definition 2.45. Sei V ein endlichdimensionaler Vektorraum. Sei $f \in \mathrm{Hom}(V, W)$. Wir definieren den Rang von f als

$$\mathrm{rang}(f) = \dim(\mathrm{Im}(f)).$$

Satz 2.46. *Seien V, W endlichdimensional, $f \in \mathrm{Hom}(V, W)$, gegeben durch $f = f_A$ (also $A = \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$ mit irgendwelchen geordneten Basen $\vec{B}$, $\vec{C}$). Dann ist* $\mathrm{rang}(f) = \mathrm{rang}(A)$.

Beweis: $\dim(\mathrm{Im}(f)) = \mathrm{rang}(A)$ für jedes A mit $A = \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$. Dies sieht man am besten am Spaltenrang, der ja die Dimension von $\mathrm{Im}(f)$ angibt. Nach dem Noether'schen Isomorphiesatz 2.15 ist dies gleich $\dim(V)/\dim(\ker(f)) = \dim(V) - \dim(\ker(f))$, also dem Zeilenrang. □

Man sieht also: Im Fall endlichdimensionaler Vektorräume V und W gilt: Für alle angeordneten Basen $\vec{B}$, $\vec{D}$ von V, $\vec{C}$, $\vec{E}$ von W, $f \in \mathrm{Hom}(V, W)$ ist

$$\mathrm{rang}(f) = \mathrm{rang}(\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)) = \mathrm{rang}(\mathrm{Mat}_{\vec{D}}^{\vec{E}}(f)).$$

2.4 Basiswechsel und Normalformen modulo Äquivalenz

Sei $\pi_{\vec{B}}(e_j) = b_j$ für $j = 1, \ldots, n$, $\pi_{\vec{C}}(e_i) = c_i$ für $i = 1, \ldots, m$, $\pi_{\vec{D}}(e_k) = d_k$ für $k = 1, \ldots, \ell$. Diese bestimmen Isomorphismen $\pi_{\vec{B}} \colon K^n \xrightarrow{\cong} U$, $\pi_{\vec{C}} \colon K^m \xrightarrow{\cong} V$, $\pi_{\vec{D}} \colon K^\ell \xrightarrow{\cong} W$.

Dann kommutiert folgendes Diagramm

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W \\ \uparrow{\scriptstyle \pi_{\vec{B}}} & & \uparrow{\scriptstyle \pi_{\vec{C}}} & & \uparrow{\scriptstyle \pi_{\vec{D}}} \\ K^n & \xrightarrow{\varphi_{A_f}} & K^m & \xrightarrow{\varphi_{A_g}} & K^\ell \end{array} \tag{2.11}$$

Somit ist nach Satz 2.21 und Beispiel 2.25

$$\begin{aligned} A_f &= \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) \quad \text{und} \quad \varphi_{A_h} = \pi_{\vec{C}}^{-1} \circ h \circ \pi_{\vec{B}} \\ A_g &= \mathrm{Mat}_{\vec{C}}^{\vec{D}}(g) \quad \text{und} \quad \varphi_{A_g} = \pi_{\vec{D}}^{-1} \circ g \circ \pi_{\vec{C}}. \end{aligned}$$

Nach Satz 2.24 ist $A_{g\circ f} = A_g A_f$.

Nun setzen wir $V = W$ und $i(c_i) = d_i$. Dadurch wird ein Isomorphismus $i\colon V \to V$ bestimmt, ein sogenannter Basiswechsel. Dann ist $\mathrm{Mat}_{\vec{D}}^{\vec{C}}(i)$ die Einheitsmatix. Wie setzen $g = \mathrm{id}$ und erhalten aus der Multiplikationsformel angewandt auf $f = \mathrm{id} \circ f$,

$$\mathrm{Mat}_{\vec{B}}^{\vec{D}}(f) = \mathrm{Mat}_{\vec{C}}^{\vec{D}}(\mathrm{id}) \cdot \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) \tag{2.12}$$

Im folgenden Lemma lernen wir, wie man $\mathrm{Mat}_{\vec{C}}^{\vec{D}}(\mathrm{id}) = M$ aus $\vec{D}$ und $\vec{C}$ berechnet. Nun setzen wir $U = V$ und $i(b_j) = c_j$ und $f = \mathrm{id}$ und erhalten aus der Multiplikationsformel, angewandt auf $g = g \circ \mathrm{id}$

$$\mathrm{Mat}_{\vec{B}}^{\vec{D}}(g) = \mathrm{Mat}_{\vec{C}}^{\vec{D}}(g) \cdot \mathrm{Mat}_{\vec{B}}^{\vec{C}}(\mathrm{id}) \tag{2.13}$$

Außerdem ist $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(\mathrm{id})$ wieder aus $\vec{C}$ und $\vec{B}$ zu bestimmen.

Ab hier Video 2, als Vorlesung für den 02.12.2021

Lemma 2.47 (Über die zu einer Basistransformation gehörende Koordinatentransformation)**.** *Seien V ein n-dimensionaler Vektorraum und $\vec{B}$ und $\vec{B}'$ Basen von V. Sei*

$$\text{für } j = 1, \dots, n, \quad b'_j = \sum_{i=1}^{n} \alpha_{i,j} b_i. \tag{2.14}$$

Sei zu $x' = (\xi'_1, \dots, \xi'_n)$ als Spaltenvektor der Spaltenvektor x gewählt, so dass

$$\sum_{j=1}^{n} \xi_j b_j = \sum_{j=1}^{n} \xi'_j b'_j. \tag{2.15}$$

Dann berechnet sich $x = (\xi_1, \dots, \xi_n)$ aus $x' = (\xi'_1, \dots, \xi'_n)$ wie folgt:

$$\xi_j = \sum_{i=1}^{n} \alpha_{j,i} \xi'_i. \tag{2.16}$$

Beweis: Der Koeffizient von b_j ist auf beiden Seiten von (2.15): $\xi_j = \sum_{i=1}^n \alpha_{i,j}\xi'_i$. Dies sieht man wie folgt:

$$\begin{aligned}\sum_{j=1}^n \xi_j b_j = \sum_{j=1}^n \xi'_j b'_j = \\ \sum_{k=1}^n \xi'_k b'_k = \sum_k \xi'_k \sum_{i=1}^n \alpha_{i,k} b_i = \\ \sum_k \xi'_k \sum_{j=1}^n \alpha_{j,k} b_j = \sum_{j=1}^n (\sum_{k=1}^n \xi'_k \alpha_{j,k}) b_j = \\ \sum_{j=1}^n (\sum_{i=1}^n \xi'_i \alpha_{j,i}) b_j.\end{aligned} \tag{2.17}$$

Hierbei haben wir beim ersten und beim dritten Gleichheutszeichen nur den Summationsindex umbenannt. Da die b_j linear unabhängig sind, ist für $j = 1, \ldots, n$,

$$\xi_j = \sum_{i=1}^n \alpha_{i,j}\xi'_i,$$

was zu zeigen war.

Beachten Sie, dass in (2.16) alte Koordinaten durch neue ausgedrückt werden und dass die Matrix aus der Gleichung (2.14) transponiert verwendet wird: Es wird über den Zeilenindex summiert. □

Nach diesen Vorbetrachtungen wechseln wir die Basen auf beiden Seiten und erhalten:

Satz 2.48 (Satz von der Transformation bei Basiswechsel)**.** *Sei f bezüglich der Basen $\vec{B}$ und $\vec{C}$ gegeben durch $A_f = \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$. Sei $h\colon \vec{B} \to \vec{B}'$ und sei $g\colon \vec{C} \to \vec{C}'$ Basiswechsel, d.h. $h(b_j) = b'_j$, $j = 1, \ldots, n$, und $g(c_j) = c'_j$, $i = 1, \ldots, m$. Hierbei schreiben wir die Matrizen zu den Basiswechseln als $\hat{B}$ mit*

$$\sum_{i=1}^m \xi_i b_i = \sum \xi'_i b'_i \text{ gdw } x' = \hat{B}x,$$

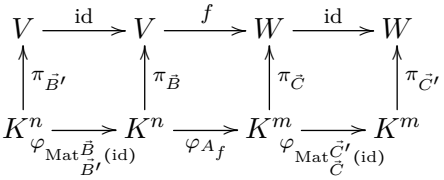
wobei x der Spaltenvektor aus $(\xi_1, \ldots, \xi_n)$ ist, und als $\hat{C}$ mit

$$\sum_{j=1}^n \zeta_j c_j = \sum_{j=1}^n \zeta'_j c'_j \text{ gdw } y' = \hat{C}y,$$

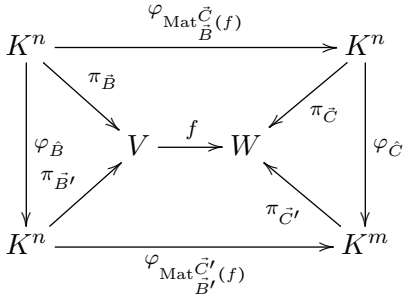
wobei y der Spaltenvektor aus $(\zeta_1, \ldots, \zeta_n)$ ist. Dann gilt

$$\begin{aligned}\mathrm{Mat}_{\vec{B}'}^{\vec{C}'}(f) &= \mathrm{Mat}_{\vec{C}}^{\vec{C}'}(\mathrm{id}) \cdot \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) \cdot \mathrm{Mat}_{\vec{B}'}^{\vec{B}}(\mathrm{id}) \\ &= \hat{C} \cdot \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) \cdot \hat{B}^{-1}.\end{aligned} \tag{2.18}$$

Beweis: Das Diagramm



oder auch das Diagramm



beschreibt die Voraussetzungen des Satzes. Nun folgt Gleichung (2.18) aus dem Zusammensetzen der Gleichungen (2.12) und (2.13).

Wir geben noch einen alternativen Beweis, der mit den Rechenregeln in den Gruppen $\mathrm{GL}_n(K)$ und $\mathrm{GL}_n(K)$ und der Assoziativität der Matrizenmultiplikation arbeitet (die ihrerseits wiederum auf Satz 2.24 bauen): Seien $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) = A$ und $\mathrm{Mat}_{\vec{B}'}^{\vec{C}'}(f) = A'$. Seien x die Koordinaten von v bzgl. $\vec{B}$ und x' die Koordinaten von v bzgl. $\vec{B}'$. Seien y die Koordinaten von $f(v)$ bzgl. $\vec{C}$ und y' die Koordinaten von $f(v)$ bzgl. $\vec{C}'$. Dann ist $y = Ax$ und $y' = A'x'$. Nun ist $y' = \hat{C}y$ und $x' = \hat{B}x$, also $x = \hat{B}^{-1}x$. Also

$$y' = \hat{C}Ax = \hat{C}A\hat{B}^{-1}x'$$

und

$$A' = \hat{C}A\hat{B}^{-1}$$

□

Definition 2.49. (1) Seien $A, B \in M_{m,n}(K)$. A und B heißen äquivalent, wenn es ein reguläres $C \in M_{n,n}(K)$ und ein reguläres $D \in M_{m,m}(K)$ gibt, so dass

$$B = DAC^{-1}.$$

(2) Seien $f, g \in \mathrm{Hom}(V, W)$. f und g heißen äquivalent, wenn es ein invertierbares $h \in \mathrm{Hom}(V, V)$ und ein invertierbares $i \in \mathrm{Hom}(W, W)$ gibt, so dass

$$g = i \circ f \circ h^{-1}.$$

(3) Seien $A, B \in M_{m,m}(K)$. A und B heißen ähnlich oder konjugiert, wenn es ein reguläres $C \in M_{m,m}(K)$ gibt, so dass

$$B = CAC^{-1}.$$

(4) Seien $f, g \in \mathrm{End}(V)$. f und g heißen ähnlich oder konjugiert, wenn es ein invertierbares $h \in \mathrm{End}(V)$ gibt, so dass

$$g = h \circ f \circ h^{-1}.$$

Beobachtung 2.50. *(1) Seien $\vec{B}$ und $\vec{C}$ endliche geordnete Basen von V bzw. W. Dann sind f und $g \in \mathrm{Hom}(V,W)$ äquivalent gdw $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$ und $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(g)$ äquivalent sind.*

(2) Sei $\vec{B}$ eine endliche geordnete Basis von V. Dann sind f und $g \in \mathrm{End}(V)$ ähnlich gdw $\mathrm{Mat}_{\vec{B}}^{\vec{B}}(f)$ und $\mathrm{Mat}_{\vec{B}}^{\vec{B}}(g)$ ähnlich sind.

(3) Die Äquivalenz von Matrizen, von Homomorphismen, und die Ähnlichkeit von quadratischen Matrizen, von Endomorphismen sind auf ihren jeweiligen Räumen Äquivalenzrelationen.

Satz 2.51 (Normalform modulo Äquivalenz, auch Normalform für lineare Abbildungen genannt)**.** *Sei $A \in M_{m,n}(K)$. Dann gibt es zu A eine äquivalente Matrix B der Form*

$$M_{m,n,k} = \begin{pmatrix} 1_{M_{k,k}(K)} & 0_{M_{k,n-k}(K)} \\ 0_{M_{m-k,k}(K)} & 0_{M_{m-k,n-k}(K)} \end{pmatrix} \tag{2.19}$$

Beweis: Wir nehmen eine Basis $b_{k+1}, \ldots, b_n$ von $\ker(f)$ und ergänzen sie um $b_1, \ldots, b_k$ zu einer Basis von V. Es sei $\vec{B} = (b_1, \ldots, b_n)$. Wir wählen $f(b_i) = c_i$ für $i = 1 \ldots, k$ und wählen einen komplementären Unterraum in W zu $\mathrm{Im}(f) = \mathrm{span}(\{c_1, \ldots, c_k\})$. Wir nehmen eine Basis $\{c_{k+1}, \ldots, c_m\}$ von W und setzen $\vec{C} = (c_1, \ldots, c_m)$. Dann ist $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$ in der Form (2.19). □

Korollar 2.52. *Eine m-n-Matrix A hat genau dann Rang k, wenn es eine reguläre n-n-Matrix B und eine eine reguläre m-m-Matrix C gibt, so dass CAB die Gestalt* (2.19) *hat.*

Definition 2.53 (Die n-n-Elementarmatrizen $E_{i,j}^{\lambda}$)**.** Für $i, j = 1, \ldots, n$, $i \neq j$ und $\lambda \in K$ definieren wir $E_{i,j}^{\lambda} = (\alpha_{i',j'})_{i',j'=1,\ldots,n}$ wie folgt:

$$\alpha_{i',j'} = \begin{cases} \delta_{i',j'}, & \text{für } (i',j') \neq (i,j); \\ \lambda, & \text{für } (i',j') = (i,j). \end{cases}$$

Für $i = 1, \dots, n$, $\lambda \in K \setminus \{0\}$ definieren wir $E_{i,i}^{\lambda} = E_i^{\lambda} = (\alpha_{i',j'})_{i',j'=1,\dots,n}$ wie folgt:

$$\alpha_{i',j'} = \begin{cases} \delta_{i',j'}, & \text{für } (i',j') \neq (i,i); \\ \lambda, & \text{für } (i',j') = (i,i). \end{cases}$$

Lemma 2.54. *Sei $A \in M_{m,n}(K)$. Die Rechtsmultiplikation mit einer Elementarmatrix aus $M_{n,n}(K)$ entspricht einer Spaltenumformung: Die Bildung von $AE_{i,j}^{\lambda}$ für $i \neq j$ addiert das λ-fache der i-ten Spalte von A zur j-ten Spalte. Die Bildung von AE_i^{λ} bildet das λ-fache der i-ten Spalte. Die Linksmultiplikation $E_{i,j}^{\lambda}A$ für $i \neq j$ addiert das λ-fache der j-ten Zeile zur i-ten Zeile von A. Die Bildung von $E_i^{\lambda}A$ bildet das λ-fache der i-ten Zeile.*

Beweis: Wir rechnen die Linksmultiplikation für $i \neq j$ nach: Seien $B = (\beta_{i',j'})$ eine m-n-Matrix, $E_{i,j}^{\lambda} = (\alpha_{k,i'})$ eine m-m-Matrix und

$$(\gamma_{k,j'})_{k=1,\dots,m,j'=1,\dots,n} = E_{i,j}^{\lambda} \cdot B.$$

Dann ist für $k = 1, \dots, m$ mit $k \neq i$, $j' = 1, \dots, n$, $\gamma_{k,j'} = \sum_{i'=1}^{m} \alpha_{k,i'}\beta_{i',j'} = \beta_{k,j'}$. Alle Zeilen außer eventuell der i-ten Zeile von B bleiben also unverändert durch die Multiplikation von links mit $E_{i,j}^{\lambda}$.

Sei nun $k = i$. Dann ist für $j' = 1, \dots, n$, $\gamma_{i,j'} = \sum_{i'=1}^{m} \alpha_{i,i'}\beta_{i',j'} = \beta_{i,j'} + \lambda\beta_{j,j'}$ Das λ-fache der $j$ten Zeile wurde zur i-ten Zeile von B addiert.

Wir rechnen nun die Rechtsmultiplikation für $i \neq j$ nach: Seien $B = (\beta_{i',j'})$ eine m-n-Matrix, $E_{i,j}^{\lambda} = (\alpha_{k,i'})$ eine n-n-Matrix und

$$(\gamma_{k,j'})_{k=1,\dots,m,j'=1,\dots,n} = B \cdot E_{i,j}^{\lambda}.$$

Dann ist für $k = 1, \dots, m$, $j' = 1, \dots, n$, $j' \neq j$, $\gamma_{k,j'} = \sum_{i'=1}^{m} \beta_{k,i'}\alpha_{i',j'} = \beta_{k,j'}$. Alle Spalten außer eventuell der j-ten Spalte von B bleiben also unverändert durch die Multiplikation von rechts mit $E_{i,j}^{\lambda}$.

Sei nun $j' = j$. Dann ist für $k = 1, \dots, m$, $\gamma_{k,j} = \sum_{i'=1}^{m} \beta_{k,i'}\alpha_{i',j} = \beta_{k,j} + \lambda\beta_{k,i}$ Das λ-fache der i-ten Spalte wurde zur j-ten Spalte von B addiert. □

bis hier in Video 2, für den 2.12.2021

Lemma 2.55. *$(E_{i,j}^{\lambda})^{-1} = E_{i,j}^{-\lambda}$ für $i \neq j$ und $(E_i^{\lambda})^{-1} = E_i^{\frac{1}{\lambda}}$. Die Umkehrungen von Elementarmatrizen sind also wieder Elementarmatrizen.*

Induktiv über die Dimension n zeigt man:

Satz 2.56. *Es seien $n \in \mathbb{N} \setminus \{0\}$, K ein Körper. Jede reguläre n-n-Matrix lässt sich als Produkt von endlich vielen Elementarmatrizen schreiben.*

Beweis: Für $n = 1$ ist $A = (\lambda)$ und $\lambda \neq 0$, also ist $A = E_1^\lambda$. Sei $A \in M_{n+1,n+1}(K)$ regulär. Die Vertauschung zweier Zeilen lässt sich durch Heranmultiplizieren geeigneter Elementarmatrizen von links durchführen. Wir bilden wir mit dem Gauß-Verfahren mit geeigneten Zeilenoperationen (i_k, j_k, ζ_k), $k = 1, \ldots, \frac{(n+4)(n+5)}{2}$,

$$A' = E_{i_1,j_1}^{\zeta_1} \cdot \cdots \cdot E_{i_k,j_k}^{\zeta_k} \cdot A,$$

so dass die letzte Zeile von A' wie $(0, \ldots, 0, 1)$ aussieht. Dies ist tatsächlich durch Zeilenoperationen, also Heranmultipizieren von Elementarmatrizen von links, erreichbar, denn im Beweis des Lemmas 3.28 wird ausgeführt, wie man Zeilenvertauschungen als Produkt von vier Elementarmatrizen schreiben kann. (Im zweiten Schritt des dortigen Beweises wird auch mit (-1) multipliziert, daher sind es dort zwei Multiplikationen mit Elementarmatrizen.) Im Hauptschritt k, $k = 1, \ldots, n+1$ braucht man maximal $n+3-k$ Elementarmatrizen-Multiplikationen und

$$\sum_{k=1}^{n+1} (k+3) \leq \sum_{k=1}^{n+4} k = \frac{(n+4)(n+5)}{2}.$$

Danach bildet man mit geeigneten (h_k, ℓ_k, ξ_k), $k = 1, \ldots, n$,

$$B = E_{h_1,\ell_1}^{\xi_1} \cdot \cdots \cdot E_{h_k,\ell_k}^{\xi_k} \cdot A',$$

so dass die letzte Spalte von B wie $(0, \ldots, 0, 1)$ als Spalte aussieht. Außerdem bleibt die letzte Zeile von B gleich wie die letzte Zeile von A'. Nun wenden wir die Induktionsvoraussetzung auf $B_{\text{kurz}} = B \restriction \{1, \ldots, n\} \times \{1, \ldots, n\}$ an. Diese liefert: B_{kurz} ist ein Produkt von (endlich vielen) n-n-Elementarmatizen. Wir fügen zu jeder dieser Elementarmatrizen als letzte Zeile $(0, \ldots, 0)$ der Länge n an und dann als $n+1$-te Spalte $(0, \ldots, 0, 1)$ der Länge $n+1$ als Spalte. So erhalten wir $(n+1)$-$(n+1)$-Elementarmatrizen, deren Produkt gerade B ist. Also hat B die Form $B = XA$, X Produkt von Elementarmatrizen. Daher ist $A = X^{-1}B$. B_{kurz} ist nach Induktionsvoraussetzung ein Produkt von Elementarmatrizen. Da das Hinzufügen der letzten Zeile und der letzten Spalte der Form $(0, \ldots, 0, 1)$ an jeden Faktor und an B_{kurz} an dieser Darstellung nichts ändert, ist auch B ein Produkt von Elementarmatrizen. Da das Inverse einer Elementarmatrix wieder eine Elementarmatrix ist, ist X^{-1} ein Produkt von Elementarmatrizen. Wir erhalten aus der Darstellung von B eine Darstellung von $A = X^{-1}B$. □

Korollar 2.57 (Aus Lemma 1.80 und dem Noether'schen Isomorphiesatz 2.15)**.** *Sei* $f \in \text{Hom}(V, W)$ *und sei* U' *ein in* V *zu* $\ker(f)$ *komplementärer Unterraum. Dann ist*

$$f \restriction U' : U' \to \text{Im}(f)$$

ein Vektorraumisomorphismus.

Satz 2.58. *Seien W, V K-Vektorräume der Dimension m und n und, sei $f: V \to W$ eine lineare Abbildung vom Rang k. Dann gibt es eine angeordnete Basis $\vec{B}$ von V und eine angeordnete Basis $\vec{C}$ von W, so dass f bezüglich der neuen angeordneten Basen durch eine Matrix in Form*

$$M_{m,n,k} = \begin{pmatrix} 1_{M_{k,k}(K)} & 0_{M_{k,n-k}(K)} \\ 0_{M_{m-k,k}(K)} & 0_{M_{m-k,n-k}(K)} \end{pmatrix} \tag{2.20}$$

dargestellt wird.

Beweis: Sei B gegeben. Wir nehmen eine Basis $b_{k+1}, \ldots, b_n$ von $\ker(f)$ und ergänzen sie um $b_1 \in B, \ldots, b_k \in B$ zu einer Basis von V. Dann ist $U' = \operatorname{span}(\{b_1, \ldots, b_k\})$ ein zu $\ker(f)$ komplementärer Unterraum und nach Korollar 2.57 $f \restriction U'$ ein Isomorphismus. Wir wählen $f(b_i) = c_i$ für $i = 1 \ldots, k$. Da $f \restriction U'$ injektiv ist, sind die $c_1, \ldots c_k$ linear unabhängig. Wir wählen einen komplementären Unterraum W' in W zu $\operatorname{Im}(f) = \operatorname{span}(\{c_1, \ldots, c_k\})$. Wir nehmen eine Basis $\{c_{k+1}, \ldots, c_m\}$ von W'. Dann ist $\operatorname{Mat}_{\vec{B}}^{\vec{C}}(f)$ in der Form (2.20). □

Korollar 2.59. *Jede m-n-Matrix lässt sich durch Zeilen- und Spaltenumformungen in die Gestalt* (2.20) *bringen. In anderen Worten: Jede m-n-Matrix lässt sich durch Heranmultiplizieren von n-n-Elementarmatrizen von rechts und von m-m-Elementarmatrizen von links in die Gestalt* (2.20) *bringen. Jede m-n-Matrix des Rangs k hat die Form*

$$P \cdot M_{m,n,k} \cdot P'$$

mit einem geeigneten Produkt P von m-m-Elementarmatrizen und einem geeigneten Produkt P' von n-n-Elementarmatrizen.

Beweis: Sei $A \in M_{m,n}(K)$. Man wählt Basen $\vec{B}$ von V und $\vec{C}$ von W, so dass $\operatorname{Mat}_{\vec{B}}^{\vec{C}}(\varphi_A) = M_{m,n,k}$ von der Form (2.19) ist. Dann stellt man die beiden Basiswechselmatrizen $\hat{B}$ (n-dimensional) und $\hat{C}$ (m-dimensional), die gemäß Gleichung (2.18) zwischen A und $M_{m,n,k}$ vermitteln, jeweils im $\mathrm{GL}_n(K)$ (von rechts) bzw. im $\mathrm{GL}_m(K)$ (von links) als Produkte von Elementarmatrizen dar nach Satz 2.56. □

Bemerkung 2.60. In der obigen Situation hat man $A = \operatorname{Mat}_{\vec{E}}^{\vec{E}}(\varphi_A)$ und $M_{m,n,k} = \operatorname{Mat}_{\tilde{B}}^{\tilde{C}}(\varphi_A) = \tilde{C}^{-1} A \tilde{B}$. Sei $b_j = (\beta_{k,j})_{k=1,\ldots,n}$, $c_i = (\gamma_{\ell,i})_{\ell=1,\ldots,m}$. Die Transformationsmatrix $\tilde{C} = (\gamma_{i,j})_{1 \le i,j \le m}$ ist gerade $(c_1|c_2|\cdots|c_m)$. Die Transformationsmatrix $\tilde{B} = (\beta_{i,j})_{1 \le i,j \le n}$ ist gerade $(b_1|b_2|\cdots|b_n)$.

Beweis: Es sei $M_{m,n,k} = A' = (\alpha'_{i,j})$, so dass $A' = \operatorname{Mat}_{\vec{B}}^{\vec{C}}(f)$ und mit $f = \varphi_A$, $f(e_j) =$

$\sum_{i=1}^{m} \alpha_{i,j} e_i$, also für $j = 1, \ldots, n$,

$$\begin{aligned} f(b_j) &= f(\sum_{r=1}^{n} \beta_{r,j} e_r) = \sum_{r=1}^{n} \beta_{r,j} f(e_r) \\ &= \sum_{r=1}^{n} \sum_{\ell=1}^{m} \beta_{r,j} \alpha_{\ell,r} e_\ell = \sum_{r=1}^{n} \sum_{\ell=1}^{m} \beta_{r,j} \alpha_{\ell,r} \sum_{i=1}^{m} (\tilde{C}^{-1})_{i,\ell} c_i \\ &= \sum_{i=1}^{m} (\sum_{\ell=1}^{m} (\tilde{C}^{-1})_{i,\ell} \sum_{r=1}^{n} \alpha_{\ell,r} \beta_{r,j}) c_i = \sum_{i=1}^{m} (\tilde{C}^{-1} A \tilde{B})_{i,j} c_i \\ &= \sum_{i=1}^{m} (M_{m,n,k})_{i,j} c_i = \sum_{i=1}^{k} \delta_{i,j} c_i. \end{aligned}$$

Da andererseits nach Definition von $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$ nach Gleichung 2.1 im Satz 2.21 gerade $f(b_j) = \sum_{i=1}^{m} \alpha'_{i,j} c_i$ ist, haben wir also $A' = \tilde{C}^{-1} A \tilde{B}$ verifiziert.

Die $\tilde{B}$, $\tilde{C}$ sind gerade die Umkehrmatrizen den $\hat{B}$, $\hat{C}$ von Satz 2.18, da wir ja in unserer Situation die neuen Basen $\vec{B}$, $\vec{C}$ (neu heißt für $A' = \mathrm{Mat}_{\vec{B}}^{\vec{C}}(f)$) haben. □

Verzichtet man auf Multiplikation von rechts, d.h. auf Spaltenumformungen, so erhält man wieder das Ergebnis des Gauß-Algorithmus: Die Stufenform vom Beginn des zweiten Kapitels:

Satz 2.61. *Seien W, V K-Vektorräume der Dimension m und n und, sei $f: V \to W$ eine lineare Abbildung vom Rang k. In V sei eine angeordnete Basis $\vec{B}$ fixiert. Dann kann man eine angeordnete Basis $\vec{C}$ von W wählen, so dass f bezüglich der neuen Basen durch eine Matrix in Stufenform hat.*

Beweis: Ändern der Basis $\vec{C}$ auf der Bildseite geschieht durch Heranmultiplizieren von links an $\mathrm{Mat}_{\vec{B}}^{\vec{C}}$. Heranmultiplizieren von links ist gleichwertig mit Zeilenumformungen. □

Je zwei ähnliche Matrizen sind äquivalent. Die Ähnlichkeit ist eine in der Regel eine echt feinere Äquivalenzrelation als die Äquivalenz von Matrizen, d.h. Ähnlichkeit impliziert Äquivalenz. Besonders einfache Vertreter der Ähnlichkeitsklassen, die Normalformen (modulo Ähnlichkeit) genannt werden, untersuchen wir im Kapitel über die Jordan'sche Normalform.

Definition 2.62. Sei K ein Körper. Eine K-Algebra [mit Eins] ist eine Struktur

$$(A, K, +_K, \cdot_K, +, \cdot, \cdot_s),$$

so dass

(1) $(A, K, +_K, \cdot_K, +, \cdot_s)$ ein K-Vektorraum ist, und

(2) $(A, +, \cdot)$ ein Ring [mit Eins] ist, und

(3) $\forall \lambda \in K,\ \forall a, b \in A\ (\lambda \cdot_s a) \cdot b = \lambda \cdot_s (a \cdot b) = a \cdot (\lambda \cdot_s b)$.

Lemma 2.63. *(1) Sei V ein K-Vektorraum. Dann ist $(\mathrm{End}(V), K, +_K, \cdot_K, +, \circ, \cdot_s)$ eine K-Algebra.*

(2) Sei $n \in \mathbb{N}$. Der Vektorraum $M_{n,n}(K)$ bildet mit der Matrizenmultiplikation als zusätzlichem Strukturmerkmal eine K-Algebra.

Definition 2.64. (Vektorraum-)Isomorphismen von V nach V heißen (Vektorraum-)Automorphismen. Mit $\mathrm{Aut}(V)$ bezeichnet man die Menge der Automorphismen oder auch die K-Algebra $(\mathrm{Aut}(V), K, +_k, \cdot_K, +, \circ, \cdot_s)$.

Satz 2.65. *Sei V n-dimensional, und sei $\vec{B}$ eine geordnete Basis von V. Dann ist*

$$\begin{aligned} \mathrm{Mat}_{\vec{B}}^{\vec{B}}\colon \mathrm{End}(V) &\to M_n(K), \\ f &\mapsto \mathrm{Mat}_{\vec{B}}^{\vec{B}}(f) \end{aligned}$$

ein K-Algebren-Isomorphismus, der $\mathrm{Aut}(V)$ auf $\mathrm{GL}_n(K)$ abbildet.

Beweis: Satz 2.21 und Satz 2.24 zusammen. □

Lemma 2.66. *Rechenregel. Sei $A \in M_{m,n}(K)$, $B \in M_{m,\ell}(K)$. Wir schreiben $(A \mid B) \in M_{m,n+\ell}(K)$ für die Hintereinanderschreibung von A und B. Sei $C \in M_{m,m}(K)$. Dann ist $C(A \mid B) = (CA \mid CB)$. Analoges gilt für die Untereinanderschreibung von Matrizen mit gleicher Spaltenzahl n und Heranmultiplizieren von n-n-Matrizen von rechts.*

Anwendung: $Ax = b$ gdw $(CA)x = Cb$. Man kann aber auch $A(x|y) = (b|c)$ suchen, usf. Das Gaußverfahren lässt sich mit beliebig vielen Spalten auf der rechten Seite durchführen.

Bemerkung 2.67. Ein wichtige Anwendung ist das Bestimmen der inversen Matrix: Gegeben seien $(A|E)$, $E = 1_{M_{n,n}(K)}$, $A \in \mathrm{GL}_n(K)$. Dann sucht man mit dem Gaußverfahren und einer Fortführung auf Diagonalform mit Einsen in der Hauptdigonalen ein C, so dass $C(A|E) = (E|CE) = (E|C)$. Man hat also gerade $C = A^{-1}$.

Kapitel 3

Determinanten

3.1 Die Signatur einer Permutation

Definition 3.1. Sei X eine endliche Menge. $X^2 = X \times X$. Eine Orientierung von X ist eine Abbildung $s\colon X^2 \setminus \{(x,x) : x \in X\} \to \{-1,1\}$, die dem Gesetz

$$s(x,y) = -s(y,x)$$

folgt.

Definition 3.2. $[X]^2 = \{\{x,y\} : x \neq y \in X\}$ ist eine übliche Schreibweise für die Menge der zweielementigen Teilmengen von X.

Übung 3.3. Es gibt auf X genau $2^{\frac{|X|(|X|-1)}{2}}$ Orientierungen. Es gibt auf X genau $|X|!$ lineare Ordungen. Jede lineare Ordnung erzeugt eine Orientierung. Für $|X| \geq 3$ gibt es auf X auch Orientierungen, die nicht von einer linearen Ordnung stammen.

Definition 3.4. und Behauptung und Beweis

(1) Sei $\pi \in \mathrm{Sym}(X)$ und sei s eine Orientierung auf X. Wir definieren das Vorzeichen oder die Signatur von π, $\mathrm{sign}(\pi)$, wie folgt:

$$\mathrm{sign}(\pi) := \prod_{\{x,y\} \in [X]^2} \frac{s(\pi(x),\pi(y))}{s(x,y)}.$$

Diese Funktion ist unabhängig von s: Denn sei t eine Orientierung auf X, die sich von s auf n ungeordneten Paaren unterscheidet, dann ist

$$\prod_{\{x,y\} \in [X]^2} \frac{s(\pi(x),\pi(y))}{t(\pi(x),\pi(y))} = \prod_{\{x,y\} \in [X]^2} \frac{s(x,y)}{t(x,y)} = (-1)^n.$$

(2) Sei $\pi \in \mathrm{Sym}(X)$, und sei s eine Orientierung auf X. Wir nennen $\{x, y\} \in [X]^2$ mit $s(\pi(x), \pi(y)) = -s(x, y)$ einen Fehlstand von π, s. Diese Definition hängt von s ab.

Beobachtung 3.5. *Aus Def. 3.4 (1)folgt:*

$$\mathrm{sign}(\pi) = (-1)^{\textit{Anzahl der Fehlstände von } \pi,\ s}.$$

Die Parität der Anzahl der Fehlstände ist somit auch invariant unter der Wahl der Orientierung s.

Übung 3.6. Ist die Anzahl der Fehlstände invariant unter der Wahl der Orientierung?

Lemma 3.7. *Die Signatur ist ein Gruppenhomomorphismus von* $\mathrm{Sym}(X)$ *nach* $(\{1, -1\}, \cdot)$.

Beweis: Übung. □

Definition 3.8. (1) Sei $k \geq 2$, $k \in \mathbb{N}$. Seien $x_1, \ldots, x_k$ paarweise verschiedene Elemente aus X. Mit

$$(x_1, \ldots, x_k)$$

bezeichnen wir die Permutation, die x_i auf x_{i+1} abbildet für $i = 1, \ldots, k-1$ und x_k auf x_1 abbildet und alle anderen Elemente festhält. Man nennt diese Permutation auch Zyklus/Zykel von $x_1, \ldots, x_k$ oder einen Zykel der Länge k. Die Zykel-Schreibweise ist mehrdeutig, da sie $X \setminus \{x_1, \ldots, x_k\}$ nicht notiert.

(2) Ein Zykel der Länge 2 heißt auch Transposition.

Beobachtung 3.9. *Ein Zykel der Länge k hat die Signatur* $(-1)^{k-1}$.

Beweis: $(x_1, \ldots, x_k)$ erzeugt bezüglich der Orientierung $s(x_i, x_j) = 1$ wenn $i < j$ genau die folgenden Fehlstände: $\{x_k, x_1\}$, …, $\{x_k, x_{k-1}\}$. □

Definition 3.10. Sei $\pi \in \mathrm{Sym}(X)$.

$$B_x = \{y : \exists z \in \mathbb{Z} : \pi^z(x) = y\}$$

heißt die π-Bahn von x. Hierbei ist π^z die z-malige Anwendung von π.

Definition 3.11. Sei $n \in \mathbb{N} \setminus \{0\}$. Sei $\mathrm{Sym}(\{1, \ldots, n\}) =: S_n$.

Lemma 3.12. *Sei X endlich. Jede Permutation von X lässt sich bis auf die Reihenfolge eindeutig als Hintereinanderausführung von disjunkten Zykeln schreiben.*

Beweis: Induktion über $|X|$. Falls $|X| = \varnothing$, ist die leere Permutation die einzige Bijektion. Sei nun $|X| = n+1$, $n \in \mathbb{N}$. Statt $\mathrm{Sym}(X)$ reicht es S_{n+1} zu betrachten. Sei $\pi \in S_{n+1}$.

1. Fall: Es gibt $\varnothing \neq Y \neq \{1, \dots, n+1\}$, so dass $\pi[Y] \subseteq Y$ (also $= Y$). Dann lassen sich nach Induktionsvoraussetzung $\pi \restriction Y$ und $\pi \restriction (X \setminus Y)$ jeweils bis auf die Reihenfolgen eindeutig als Produkt von Zykeln darstellen. Außerdem ist $\pi = (\pi \restriction Y \cup \mathrm{id}_{X \setminus Y}) \circ (\pi \restriction (X \setminus Y) \cup \mathrm{id}_Y)$.

2. Fall: Es gibt kein solches Y. Dann ist $\pi = (1, \pi(1), \pi^2(1), \dots, \pi^n(1))$ ein Zykel. □

Bemerkung: Bei unendlichem X hat man womöglich auch unendlich lange Bahnen, die wie $\mathbb{Z}$ angeordnet sind, also nicht mehr zyklisch sind. Auf jeder solchen Bahn sieht π wie die Shift-Abbildung $\pi^z(x) \mapsto \pi^{z+1}(x)$, $z \in \mathbb{Z}$, aus, für ein festes x aus der Bahn. Jede Permutation lässt sich als als (womöglich unendlich langes) Produkt endlicher Zykel und Shift-Abbildungen auf unendlichen Bahnen schreiben. Wir werden dies hier nicht beweisen.

Satz 3.13. *Jede Permutation auf einer endlichen Menge ist ein Produkt von Transpositionen.*

Beweis: Jeder Zykel ist Produkt endlich vieler Transpositionen: Sei $k \geq 2$. $(x_1, \dots, x_k) = (x_1, \dots, x_{k-1}) \circ (x_{k-1}, x_k)$. □

Definition 3.14. Ein Gruppenhomomorphismus, der alle Elemente auf das neutrale Element abbildet, heißt trivialer Homomorphismus.

Satz 3.15. *Sei $|X| \geq 2$ und X endlich.* sign *ist der einzige nicht triviale Gruppenhomomorphismus von* $\mathrm{Sym}(X)$ *auf* $\{-1, 1\}$.

Beweis: Nach Lemma 3.7 ist sign ein Gruppenhomomorphismus. von $\mathrm{Sym}(X)$ nach $\{-1, 1\}$, der im Falle $|X| \geq 2$ nicht trivial ist.

Sei nun φ ein beliebiger Gruppenhomomorphismus. von $\mathrm{Sym}(X)$ nach $\{-1, 1\}$, $|X| \geq 2$ und s nicht trivial.

Es gibt $\varepsilon \in \{1, -1\}$, so dass für jede Transposition (a, b), $\varphi(a, b) = \varepsilon$: Sei (a', b') eine Transposition und sei $\pi(a) = a'$ und $\pi(b) = b'$ und $\pi \in \mathrm{Sym}(X)$. Dann ist $(a, b) = \pi^{-1}(a', b')\pi$ und $\varphi(a, b) = \varphi(a', b')$.

Da φ nicht trivial sein soll, gibt es ein $(a, b) \in \mathrm{Sym}(X)$, so dass $\varphi(a, b) = -1$. Somit ist $\varphi(a, b) = -1$ für alle Transpositionen. Da φ und sign nun auf allen Transpositionen übereinstimmen und jede Permutation Produkt von Transpositionen ist, gilt $\varphi = \mathrm{sign}$. □

Eine andere Darstellung des Vorzeichens ist gegeben durch die folgende Formel.

Korollar 3.16. *Sei $\pi \in S_n$. Dann ist*

$$\mathrm{sign}(\pi) = \prod_{1 \leq i < j \leq n} \frac{\pi(j) - \pi(i)}{j - i}.$$

Definition 3.17. Sei $(G, *)$ eine Gruppe.

(1) $\varnothing \neq N \subseteq G$ heißt Normalteiler, wenn N eine Untergruppe von G ist und

$$\forall g \in G,\ \forall n \in N,\ g * n * g^{-1} \in N.$$

(2) G heißt einfach, wenn G keine Normalteiler außer $\{e\}$ und G (den sogenannten trivialen Normalteilern) hat.

Definition 3.18. Man nennt

$$A_n = \{\pi \in S_n : \operatorname{sign}(\pi) = 1\}$$

die alternierende Gruppe auf $\{1, \dots, n\}$.

Satz 3.19. *Sei X endlich, $|X| \geq 3$. $\operatorname{Sym}(X)$ hat den nicht trivialen Normalteiler*

$$A = \{\pi \in \operatorname{Sym}(X) : \operatorname{sign}(\pi) = 1\}.$$

Beweis: Man rechnet nach: A ist ein nicht trivialer Normalteiler. Wenn $\operatorname{sign}(\pi) = 1$, so ist auch $\operatorname{sign}(\sigma \circ \pi \circ \sigma^{-1}) = 1$. Da $|X| \geq 3$, gibt es das Element $\mathrm{id} \neq (x_1, x_2, x_3) \in A$, A ist also nicht die triviale Gruppe $(\{\mathrm{id}\}, \circ)$. Auch ist A nicht $\operatorname{Sym}(X)$, da $(x_1, x_2) \in \operatorname{Sym}(X) \setminus A$. □

Übung 3.20. Sei X unendlich. Geben Sie weitere Normalteiler in $\operatorname{Sym}(X)$.

Aus der Algebra zitieren wir den folgenden Satz, den wir hier nicht beweisen.

Satz 3.21. *A_n ist einfach für $n = 3$ und $n \geq 5$.*

3.2 k-Formen

Definition 3.22. Sei $k \in \mathbb{N} \setminus \{0\}$. Seien $V_1, \dots, V_k$ und W K-Vektorräume.

$$\mu : V_1 \times \cdots \times V_k \to W$$

heißt multilinear, wenn μ in allen Argumenten linear ist, d.h $\forall j = 1, \dots, k, \forall \alpha, \beta \in K, \forall v_i \in V_i, i = 1, \dots, \ell, \forall v_j' \in V_j$

$$\mu(v_1, \dots, v_{j-1}, \alpha v_j + \beta v_j', v_{j+1}, \dots, v_k) = \alpha\mu(v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_k) + \beta\mu(v_1, \dots, v_{j-1}, v_j', v_{j+1}, \dots, v_k).$$

V^k bezeichnet das k-fache kartesische Produkt von V.

Definition 3.23. Eine multilineare Abbildung $\mu: V^k \to K$ heißt alternierend oder k-Form, wenn für alle $a_1, \ldots, a_k \in V$ gilt: Falls ein Vektor in $a_1, \ldots, a_k$ zweimal vorkommt, gilt $\mu(a_1, \ldots, a_k) = 0$.

Beobachtung 3.24. *Eine multilineare Abbildung ist genau dann alternierend, wenn sie scherungsinvariant ist, d.h. für $i \neq j$, $\lambda \in K$,*

$$\mu(v_1, \ldots, v_{i-1}, v_i \quad + \quad \lambda\beta v_j, v_{i+1}, \ldots, v_k) \qquad = \qquad \mu(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k).$$

Lemma 3.25. *Eine Multilinearform ist genau dann alternierend, wenn für alle linear abhängigen Tupel $a_1, \ldots, a_k$ gilt $\mu(a_1, \ldots, a_k) = 0$.*

Beweis: Sei μ alternierend, und sei o.B.d.A. a_1 eine Linearkombination von $a_2, \ldots, a_k$. Dann ist $\mu(a_1, \ldots, a_k) = 0$. Zur Rückrichtung: Sei $a_j = a_i$ für $i \neq j$. Dann ist $a_1, \ldots, a_n$ linear anhängig. □

Gilt die Vorwärtsrichtung auch in Moduln (Vektorräumen über Ringen statt Körpern)? Ja, bei Nullteilerfreiheit.

Lemma 3.26. *(1) Jede Linearkombination von (alternierenden) k-Formen ist eine (alternierende) k-Form.*

(2) Sie $\varphi: V \to W$ eine lineare Abbildung, und sei μ eine (alternierende) k-Form auf W. Dann ist μ^φ, definiert durch

$$\mu^\varphi(x_1, \ldots, x_k) = \mu(\varphi(x_1), \ldots, \varphi(x_k))$$

eine (alternierende) k-Form auf V.

Beweis: : Nachrechnen (im Kopf oder aufschreiben). □

Definition 3.27. Man bezeichnet mit $A^k V$ die Menge (und auch den K-Vektorraum) der alternierenden k-Formen auf V.

Lemma 3.28. *Sei μ eine k-Form auf V. Dann ist für alle $a_1, \ldots, a_k \in V$*

$$\mu(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_k) = -\mu(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_k).$$

Beweis: Wieder hilft der schon im Beweis von Lemma 3.28 benutzte Trick: „Vertauschen entspricht dreimal Scheren und einmal Multiplizieren mit (-1)":

$$\begin{aligned}
&\mu(a_1,\dots,a_{i-1},a_i,a_{i+1}\dots,a_{j-1},a_j,a_{j+1},\dots,a_k) =\\
&\mu(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_{j-1},a_i+a_j,a_{j+1},\dots,a_k) =\\
&\mu(a_1,\dots,a_{i-1},-a_j,a_{i+1},\dots,a_{j-1},a_i+a_j,a_{j+1},\dots,a_k) =\\
&\mu(a_1,\dots,a_{i-1},-a_j,a_{i+1},\dots,a_{j-1},a_i,a_{j+1},\dots,a_k) =\\
&-\mu(a_1,\dots,a_{i-1},a_j,a_{i+1},\dots,a_{j-1},a_i,a_{j+1},\dots,a_k).
\end{aligned}$$

Bei den ersten drei Gleichheitszeichen haben wir geschert, beim vierten (-1) herausgezogen. □

Bemerkung 3.29. Wenn K nicht die Charakteristik 2 hat (d.h., dass $1_k +_K 1_K \neq 0_k$), folgt aus der gegebenen Eigenschaft wieder das Alternieren.

Definition 3.30. Für beliebige Abbildungen $\tau\colon\{1,\dots,k\}\to\{1,\dots,k\}$ sei $\operatorname{sign}(\tau)=0$, falls τ nicht bijektiv ist.

Beobachtung 3.31. *Seien* $\sigma,\tau\colon\{1,\dots,k\}\to\{1,\dots,k\}$. *Dann ist* $\operatorname{sign}(\sigma\circ\tau)=\operatorname{sign}(\sigma)\cdot\operatorname{sign}(\tau)$.

Korollar 3.32. *Sei μ eine alternierende k-Form und sei $\sigma\colon\{1,\dots,k\}\to\{1,\dots,k\}$ und seien $a_1,\dots,a_k\in V$. Dann gilt*

$$\mu(a_{\sigma(1)},\dots,a_{\sigma(k)})=\operatorname{sign}(\sigma)\mu(a_1,\dots,a_k).$$

Beweis: 1. $\sigma\notin S_k$. Dann sind beide Seiten Null.

2. Fall $\sigma\in S_k$. Dann stellen wir σ als Produkt von Transpositionen dar. Für jede Transposition τ gilt $\mu(a_{\tau(1)},\dots,a_{\tau(k)})=(-1)\mu(a_1,\dots,a_k)$ und auch für jedes $\pi\in S_k$ $\mu(a_{\tau(\pi(1))},\dots,a_{\tau(\pi(k))})=(-1)\mu(a_{\pi(1)},\dots,a_{\pi(k)})$. Die Behauptung folgt also durch Induktion über die Anzahl der Transpositionen, die als Produkt σ ergeben. □

Korollar 3.33. *Sei μ eine alternierende k-Form und $(a_1,\dots,a_n)$ sei eine angeordnete Basis von V. Dann ist μ bestimmt durch*

$$\{(i_1,\dots,i_k,\mu(a_{i_1},\dots,a_{i_k}))\,:\,1\le i_1<\dots<i_k\le n\}.$$

Lemma 3.34. *Es gibt* $\binom{n}{k}:=\frac{n!}{k!(n-k)!}$ *Wahlen von* $\{i_1,\dots,i_k\}$ *aus* $\{1,\dots,n\}$.

Korollar 3.35. *Wenn* $\dim(V)<k$, *dann ist die Null-Funktion, d.h. die Funktion, die jedes k-Tupel auf* 0 *abbildet, die einzige alternierende k-Form auf V.*

3.3 Determinanten

Satz 3.36. *Sei $b_1, \dots, b_n$ eine angeordnete Basis von V, $\beta \in K$. Dann gibt es genau eine alternierende n-Form μ auf V mit $\mu(b_1, \dots, b_n) = \beta$.*

Beweis: Die Eindeutigkeit folgt aus 3.33. Wir nehmen eine Multilinearform μ, indem wir für $\tau \in \mathrm{Sym}(\{1, \dots, n\})$ setzen

$$\mu(b_{\tau(1)}, \dots, b_{\tau(n)}) = \mathrm{sign}(\tau)\beta.$$

Dann ist für alle τ

$$\mu(b_{\tau(1)}, \dots, b_{\tau(n)}) = \mathrm{sign}(\tau)\mu(b_1, \dots, b_n).$$

Falls die Charakteristik von K nicht 2 ist sind wir nach Bemwerkung 3.29 fertig. Nun zeigen wir für allgemeines K, dass μ alternierend ist. Seien nun zwei Argumente gleich, der Einfachheit halber seien das erste und das zweite Argument $a = \sum_{i=1}^n \xi_i b_i$. Dann ist für alle j_ℓ, $\mu(a, a, b_{j_3}, \dots, b_{j_n}) = \sum_{i=1}^n \xi_i^2 \mu(b_i, b_i, \dots) + \sum_{1 \le i < j \le n} \xi_i \xi_j (\mu(b_i, b_j, \dots) + \mu(b_j, b_i, \dots)) = 0$. μ ist also alternierend. □

Definition 3.37. Eine k-Form mit Bildmenge $\{0\}$ heißt triviale k-Form.

Korollar 3.38. *Sei V n-dimensional und sei μ eine nicht triviale alternierende n-Form auf V. Dann ist jede andere alternierende n-Form ein Vielfaches von μ.*

Definition 3.39 (Die Standardform der alternierenden n-Form)**.** Es sei μ_0 die alternierende n-Form auf K^n, so dass μ_0 die kanonische angeordnete Basis $(e_1, \dots, e_n)$ abbildet auf $\mu_0(e_1, \dots, e_n) = 1$.

Definition 3.40. Sei $A \in M_{n,n}(K)$, und sei für $j = 1, \dots, n$ sein a_j die j-te Spalte von A. Es sei μ_0 die standardisierte alternierende n-Form. Dann ist die Determinante von A,

$$\det(A) := \mu_0(a_1, \dots, a_n).$$

Wir haben also $\det(1_{M_{n,n}(K)}) = 1$. Man schreibt auch $|A|$ für $\det(A)$, besonders wenn man die Einträge von A ausgeschrieben hat.

Lemma 3.41 (Die Leibnizformel)**.** *Sei $A = (\alpha_{i,j})_{i,j=1,\dots,n}$. Dann ist*

$$\det(A) = \sum_{\pi \in S_n} \mathrm{sign}(\pi) \prod_{j=1}^n \alpha_{\pi(j),j}. \tag{3.1}$$

Man nennt (3.1) *die Leibniz-Formel*[1]*.*

[1] Gottfried Wilhelm von Leibniz, 1646 – 1716

Beweis: Nach Definition, zweimaliger Anwendung der Linearität, Korollar 3.32 und Def. 3.30 erhalten wir

$$\begin{aligned}\det(A) =& \mu_0(\sum_i \alpha_{i,1}e_i, \dots, \sum_i \alpha_{i,n}e_i)\\ =& \sum_{\tau:\{1,\dots,n\}\to\{1,\dots,n\}} \mu_0(\alpha_{\tau(1),1}e_{\tau(1)}, \dots, \alpha_{\tau(n),n}e_{\tau(n)})\\ =& \sum_{\tau:\{1,\dots,n\}\to\{1,\dots,n\}} (\prod_{j=1}^{n} \alpha_{\tau(j),j})\mu_0(e_{\tau(1)}, \dots, e_{\tau(n)})\\ =& \sum_{\tau:\{1,\dots,n\}\to\{1,\dots,n\}} \operatorname{sign}(\tau)\prod_{j=1}^{n} \alpha_{\tau(j),j}\\ =& \sum_{\pi\in S_n} \operatorname{sign}(\pi)\prod_{j=1}^{n} \alpha_{\pi(j),j}.\end{aligned}$$

Nun verifiziert man das diese einzige Möglichkeit für die Determinante auch tatsächlich die Standard-n-Form ist. □

Lemma 3.42.

$$\det((a)) = a,$$

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - b_1a_2,$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_2b_1c_3 - a_1b_3c_2 - a_3b_2c_1.$$

Für obere Dreiecksmatrizen ist

$$\begin{vmatrix} \lambda_1 & & & \\ 0 & \lambda_2 & * & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{vmatrix} = \prod_{j=1}^{n} \lambda_j,$$

und schließlich ist $\det(E_i^\lambda) = \lambda$, $\det(E_{i,j}^\lambda) = 1$.

Lemma 3.43. *Sei* $A \in M_{m,m}(K)$, $B \in M_{m,n}(K)$, $C \in M_{n,n}(K)$. *Dann ist* $\begin{vmatrix} A & B \\ 0_{M_{n,m}(K)} & C \end{vmatrix} = \det(A)\cdot\det(C)$.

Beweis: Das ist mit (3.1) leicht nachzurechnen. Man kann sich auf $\pi \in S_{m+n}$ beschränken, die $\{1,\dots,m\}$ und $\{m+1,\dots,n\}$ invariant lassen. Jede solche Permutation ist das Produkt ihrer Einschränkungen auf $\{1,\dots,m\}$ und $\{m+1,\dots,n\}$. Die Signatur der Hintereinanderausführung ist das Produkt der Signaturen der einzelnen Permutationen. □

Satz 3.44. *(1)* $\det(AB) = \det(A)\det(B)$.

(2) $\det(A) = 0$ *gdw* A *singulär.*

Beweis: (1) Seien $b_1,\dots,b_n$ die Spalten von B. Dann ist für jedes feste A die Funktion

$$B \mapsto \mu(b_1,\dots,b_n) = |Ab_1, Ab_2, \dots, Ab_n| = \det(AB)$$

eine n-Form auf dem K^n. $\mu(e_1,\dots,e_n) = \det(A)$ und der Eindeutigkeitssatz zeigen, dass $\mu = \det(A)\cdot\mu_0$.

(2) Wenn A invertierbar ist, ist $\det(A)\cdot\det(A^{-1}) = \det(1_{M_{n,n}(K)}) = 1$, also $\det(A) \neq 0$. Wenn A singulär ist, sind die Spalten von A linear abhängig, und daher ist $\det(A) = 0$. □

Korollar 3.45. $\det\colon \mathrm{GL}_n(K) \to (K\setminus\{0\},\cdot)$ *ist ein Gruppenhomomorphismus.*

Der Kern $\varphi^{-1}[\{e_2\}]$ jedes Gruppenhomomorphismus $\varphi\colon (G_1,*_1,e_1) \to (G_2,*_2,e_2)$ ist eine Untergruppe (sogar ein Normalteiler) von G_1. Dies zeigt man wie in Lemma 2.7. Für $n > 1$ ist det nicht injektiv.

Definition 3.46. Die spezielle lineare Gruppe ist

$$\mathrm{SL}_n(K) = \{A \in M_{n,n}(K) : \det(A) = 1\}.$$

Definition 3.47. Die Transponierte einer m-n-Matrix $A = (\alpha_{i,j})_{1=1,\dots,m,j=1,\dots,n}$ ist die n-m-Matrix

$$A^\top = (\alpha_{j,i})_{j=1,\dots,n,i=1,\dots,m}$$

Die Spalten von $A^\top$ sind also die Zeilen von A.

Lemma 3.48. $(AB)^\top = B^\top A^\top$.

Beweis: Seien $a_1,\dots,a_k$ die Zeilen von A und $b_1,\dots,b_n$ die Spalten von B. Aus

$$AB = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix}$$

folgt

$$(AB)^\top = \begin{pmatrix} b_1^\top \\ \vdots \\ b_n^\top \end{pmatrix} \begin{pmatrix} a_1^\top & \dots & a_k^\top \end{pmatrix} = B^\top A^\top$$

□

Satz 3.49. *(1)* $\det(A^\top) = \det(A)$.

(2) det *ist eine alternierende Multilinearform der Zeilen von A.*

Beweis: (1)

$$\begin{aligned}\det(A^\top) &= \sum_{\pi\in S_n} \operatorname{sign}(\pi) \prod_{i=1}^{n} \alpha_{i,\pi(i)} \\ &= \sum_{\pi\in S_n} \operatorname{sign}(\pi) \prod_{i=1}^{n} \alpha_{\pi^{-1}(i),i} \\ &= \sum_{\pi\in S_n} \operatorname{sign}(\pi) \prod_{i=1}^{n} \alpha_{\pi(i),i} \\ &= \det(A)\end{aligned}$$

Hierbei nutzen wir $\operatorname{sign}(\pi) = \operatorname{sign}(\pi^{-1})$. (2) folgt aus (1). □

3.4 Der Laplace'sche Entwicklungssatz

Satz 3.50. *Die Cramer'sche*[2] *Regel . Sei $Ax = b$ ein lineareres Gleichungssystem, und $a_1, \ldots, a_n$ seien die Spalten von A. Wenn A regulär ist, dann errechnet sich der Lösungsvektor $x = (\xi_1, \ldots, \xi_n)$ (als Spalte) wie folgt: Für $j = 1, \ldots, n$ ist*

$$\xi_j = \frac{\det(a_1, \ldots, a_{j-1}, b, a_{j+1}, a_n)}{\det(A)}.$$

Beweis: $Ax = b$ heißt ausführlich $\xi_1 a_1 + \cdots + \xi_n a_n = b$. Die Gleichung ist äquivalent zur Aussage: Für $j = 1, \ldots, n$ ist

$$\begin{aligned}\det(a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_n) &= \det(a_1, \ldots, \sum_i \xi_i a_i, \ldots, a_n) \\ &= \sum_i \xi_i \det(a, a_{j-1}, a_i, a_{j+1}, \ldots, a_n) \\ &= \xi_j \det(a_1, \ldots, a_n) = \xi_j \det(A),\end{aligned}$$

□

Definition 3.51. Sei $A \in M_{n,n}(K)$ und seien $i, j \in \{1, \ldots, n\}$. Dann bezeichnet $A_{i,j}$ die Matrix, die aus A durch Streichung der i-ten Zeile und der j-ten Spalte entsteht.

[2] Gabriel Cramer, 1704 – 1752

Satz 3.52. *Der Laplace'sche Entwicklungssatz*[3] *Sei $A = (\alpha_{i,j})$ eine n-n-Matrix und sei j_0 ein Spaltenindex. Dann ist*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j_0} \alpha_{i,j_0} \det(A_{i,j_0}).$$

Man nennt die rechte Seite auch die Entwicklung der Determinante nach der j_0-ten Spalte.

Beweis: Seien a_j die Spalten von A. $a_{j_0} = \sum_i \alpha_{i,j_0} e_i$. Daher liefert die Linearität in der j_0-ten Spalte $\det(A) = \sum_i \alpha_{i,j_0} |a_1, \ldots, a_{j_0-1}, e_i, a_{j_0+1}, \ldots, a_n|$. Wir verschieben in der Matrix $(a_1, \ldots, \ldots, a_{j_0-1}, e_i, a_{j_0+1}, \ldots, a_n)$ die j_0-te Spalte nach vorn und die i-te Zeile nach oben und erhalten

$$\hat{A_{i,j_0}} := \begin{pmatrix} 1 & * \\ 0_{K^{n-1}} & A_{i,j_0} \end{pmatrix}$$

Die bei den Verschiebungen verwendeten zyklischen Permutationen $(1 \cdots j_0)$ und $(1 \cdots i)$ haben die Signaturen $(-1)^{j_0-1}$ und $(-1)^{i-1}$. Aus Satz 3.49 folgt

$$|a_1, \ldots, a_{j_0-1}, e_i, a_{j_0+1}, \ldots, a_n| = (-1)^{(i-1)+(j_0-1)} \det(\hat{A_{i,j_0}}) = (-1)^{i+j_0} \det(A_{i,j_0}).$$

□

Definition 3.53. Wir definieren die Adjunkte $\mathrm{adj}(A) = (\gamma_{i,j})$ einer n-n-Matrix A durch

$$\gamma_{i,j} = (-1)^{i+j} \det(A_{j,i}).$$

Satz 3.54. $\mathrm{adj}(A)A = \det(A) 1_{M_{n,n}(K)}$

Beweis: Sei c_j die j-te Zeile von $\mathrm{adj}(A)$ und sei $b = (\beta_1, \ldots, \beta_n)$ als Spaltenvektor. Nach dem Entwicklungssatz ist $c_j b = \sum_{k=1}^{n} (-1)^{j+k} \beta_k \det(A_{k,j}) = |a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_n|$. Nun setzen wir $b = a_i$ ein und erhalten

$$c_j a_i = \begin{cases} \det(A), & \text{wenn } j = i; \\ 0 & \text{sonst.} \end{cases}$$

□

Bemerkung 3.55. Dann ist auch $A\,\mathrm{adj}(A) = \det(A) 1_{M_{n,n}(K)}$. Dies folgt im Fall $\det(A) \neq 0$ aus Lemma 1.3 über Inverse in Gruppen: $ab = e \to ba = e$. Wir rechnen hier in der Gruppe $\mathrm{GL}_n(K)$. Im Fall $\det(A) = 0$ transponiert man den Beweis von eben: Sei b ein Zeilenvektor und sei c_j die j-te Spalte von $\mathrm{adj}(A)$, berechnen bc_j, ...

[3] Pierre-Simon Marquis de Laplace 1749 – 1827

Definition 3.56. Seien $\beta_1, \ldots, \beta_n \in K$, $n \in \mathbb{N} \setminus \{0\}$. Eine Matrix $(\alpha_{i,j}) \in M_{n,n}(K)$ mit

$$\alpha_{i,j} = \beta_j^{i-1}$$

heißt Vandermonde-Matrix[4] zu $\beta_1, \ldots, \beta_n$. Auch die Transponierte einer Matrix der angegebenen Form heißt Vandermonde-Matrix.

Satz 3.57. *Seien $n \in \mathbb{N} \setminus \{0\}$ und $\beta_1, \ldots, \beta_n \in K$, und sei A ein Vandermonde-Matrix zu $\beta_1, \ldots, \beta_n \in K$. Dann ist*

$$\det(A) = \prod_{1 \le i < j \le n} ((\beta_j - \beta_i).$$

Beweis: Wir starten mit

$$A = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \beta_1 & \beta_2 & \ldots & \beta_n \\ \beta_1^2 & \beta_2^2 & \ldots & \beta_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{n-1} & \beta_2^{n-1} & \ldots & \beta_n^{n-1} \end{pmatrix}$$

Man subtrahiert für $k = n, \ldots 2$ das β_1-fache der $(k-1)$-ten Zeile von der k-ten Zeile und erhält eine Matrix mit $(1, 0, \ldots, 0)^\top$ als erster Spalte und $(1, \beta_i - \beta_1, \ldots, \beta_i^{n-1} - \beta_1 \beta_i^{n-2})^\top$ als i-ter Spalte für $i = 2, \ldots, n$. Dann entwickelt man nach der ersten Spalte und zieht aus der j-ten Spalte für $j = 2, \ldots, n$ jeweils den Faktor $(\beta_j - \beta_1)$ heraus. So ist $\det(A)$ gleich dem Produkt $(\beta_n - \beta_1) \cdot \cdots \cdot (\beta_2 - \beta_1)$ und der folgenden $n-1$-$n-1$-Determinante

$$d = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \beta_2 & \beta_3 & \ldots & \beta_n \\ \beta_2^2 & \beta_3^2 & \ldots & \beta_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_2^{n-2} & \beta_3^{n-2} & \ldots & \beta_n^{n-2} \end{vmatrix}.$$

Nun ist $d = \prod_{2 \le i < j \le n} (\beta_j - \beta_i)$, nach Induktionsvoraussetzung. □

Beweis von Vandermonde im Fall $n = 3$ im Jahre 1771, von Cauchy 1815 im allgemeinen Fall.

Nun geben wir eine Anwendung: Die Approximation durch Polynome.

Satz 3.58. *Seien β_i, $1 \le i \le n$ paarweise verschiedene Körperelemente und sei γ_i, $1 \le i \le n$ ein Bild-Tupel. Dann gibt es genau ein Polynom des Grades $n-1$ der Form $\sum_{i=0}^{n-1} \alpha_{i+1} x^i$, so dass*

$$(I) \qquad \text{für } i = 1, \ldots, n: \qquad \sum_{j=0}^{n-1} \alpha_{i+1} \beta_i^j = \gamma_i.$$

[4] Alexandre-Théophile Vandermonde, 1735 – 1796

Beweis: Die Suche nach den Unbekannten $(\alpha_1, \ldots, \alpha_n)$ ist gegeben durch das lineares Gleichungssystem (I) $A\vec{\alpha} = \vec{\gamma}$. Die transponierte Matrix ist die Vandermonde-Matrix

$$A^T = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \beta_1 & \beta_2 & \ldots & \beta_n \\ \beta_1^2 & \beta_2^2 & \ldots & \beta_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{n-1} & \beta_2^{n-1} & \ldots & \beta_n^{n-1} \end{pmatrix}.$$

Da nach Voraussetzung die Argumentstellen β_i, $1 \leq i \leq n$, paarweise verschieden sind, erhalten wir aus Satz 3.49 und aus Satz 3.57 die Gleichungen $\det(A) = \det(A^T) \neq 0$. □

3.5 Geometrische Bedeutung der Determinanten

Orientierung

Definition 3.59. Eine angeordnete Basis $(b_1, \ldots, , b_n)$ heißt positiv orientiert, wenn $\mu_0(a_1, \ldots, a_n) > 0$.

Die folgende Betrachtung benutzt Begriffe aus der Analysis und der Topologie. Wir können $(\mathbb{R}^n)^n$ auf kanonische Weise mit einer Topologie versehen (die wie $\mathbb{R}^{n^2}$ mit der von $\mathbb{R}$ geerbten Produkttopologie aussieht). Sei

$$B\colon [0,1] \to \{\text{geordnete Basen von } \mathbb{R}^n\}$$

eine stetige Funktion, und sei $B(0) = \vec{B}_0$ und $B(1) = \vec{B}_1$. Die Abbildung $t \mapsto \mu_0(B(t))$ ist eine stetige Funktion von t, die nicht Null wird. Also haben $\vec{B}_0$ und $\vec{B}_1$ dasselbe Vorzeichen. Man kann eine positiv orientierte Matrix nicht auf stetige Weise im Raum der angeordneten Basen in eine negativ orientierte Matrix überführen.

Jede nicht triviale n-Form μ gestattet eine Definition von Orientierung nach folgendem Muster: $\vec{B}$ ist positiv orientiert bzgl. μ wenn $\mu(\vec{B}) > 0$.

Zwei n-Formen μ und μ' definieren dieselbe Orientierung gdw $\mu = \lambda\mu'$ für ein positives λ.

$\mathbb{R}^n$ hat also genau zwei Orientierungen.

Anregung: Man informiere sich über den Begriff eines angeordneten Körpers. Man überlege sich, wie dies in anderen angeordneten Körpern aussieht. $\mathbb{C}$ hat keine Anordnung.

Volumen

Definition 3.60. Für Vektoren $a_1, \ldots, a_n \in \mathbb{R}^n$ sei $\mathrm{vol}_n(a_1, \ldots, a_n)$ das in der Analysis definierte n-dimensionale Volumen des Parallelepipeds

$$\mathrm{PE}(a_1, \ldots, a_n) = \{\lambda_1 a_1 + \cdots + \lambda_n a_n \,:\, 0 \leq \lambda_i \leq 1 \text{ für alle } i\}$$

Satz 3.61. $\mathrm{vol}_n(a_1, \ldots, a_n) = |\det(a_1, \ldots, a_n)|$.

Beweis: Induktiv über n. Sei a_1' der Normalenvektor von a_1 auf die Hyperebene $\mathrm{span}(a_2, \ldots, a_n)$. Dann ist

$$\mathrm{vol}_n(a_1, \ldots, a_n) = |a_1'| \,\mathrm{vol}_{n-1}(a_2, \ldots, a_n) \tag{3.2}$$

nach dem Cavalieri-Prinzip[5]. Wir nehmen o.B.d.A. an, dass $a_1', e_2, \ldots, e_n$ linear unabhängig sind. (Nach dem Basisergänzungssatz gibt es $n-1$ Vektoren unter den e_i, die zusammen mit a_1' eine Basis bilden.) Wir nehmen als neue geordnete Basis $(a_1', e_2, \ldots, e_n)$ und schreiben a_i in den neuen Koordinaten als a_i'. Dann ist nach der Scherungsinvarianz $|\det(a_1, a_2, \ldots, a_n)| = |\det(a_1', a_2, \ldots, a_n)|$. Nach dem Satz über die Basistransformation ist $\det(a_1', a_2, \ldots, a_n) = \det(a_1', a_2', \ldots, a_n')$. Wir schreiben $a_i' = (\alpha_{i,1}', \ldots, \alpha_{i,n}')$ als Spaltenvektor für $i = 2, \ldots, n$. Dann ist $\alpha_{i,1}' = 0$, da a_1' ein Normalenvektor ist. Außerdem sind die Koordinaten von a_1' in der neuen Basis gerade $(|a_1'|, *, \ldots, *)$ Wir schreiben $a_i'' = (\alpha_{i,2}', \ldots \alpha_{i,n}')$ als Spaltenvektor. Nach Lemma 3.43 folgt nun

$$|\det(a_1', a_2', \ldots, a_n')| = |a_1'| |\det_{n-1}(a_2'', \ldots, a_n'')|. \tag{3.3}$$

Nun ist nach Induktionsvoraussetzung $\mathrm{vol}_{n-1}(a_2, \ldots, a_n) = |\det_{n-1}(a_2'', \ldots, a_n'')|$. Nun setzen wir die Invarianz unter Basiswechsel, die Gleichung (3.3) und die Induktionsvoraussetzung in die Gleichung (3.2) ein und erhalten die Induktionsbehauptung. □

Endomorphismen

Definition 3.62. Sei $\varphi \in \mathrm{End}(V)$, V n-dimensional, und sei A eine Matrixdarstellung von φ (bezüglich irgendeiner geordneten Basis $\vec{B}$). Dann definieren wir

$$\det(\varphi) = \det(A).$$

Dies ist nach Satz 2.48 und dem Korollar 3.45 wohldefiniert.

Satz 3.63. *Seien* $\varphi, \psi \in \mathrm{End}(V)$. $\det(\varphi \circ \psi) = \det(\varphi) \cdot \det(\psi)$.

Beweis: Dies folgt aus dem Multiplikationssatz 2.24.

[5] Bonaventura Francesco Cavalieri, 1598 – 1647

Kapitel 4

Dualräume

4.1 Der Dualraum

Wir wiederholen

Definition 4.1. Sei V ein K-Vektorraum. $V^* = \mathrm{Hom}(V, K)$ heißt der der Dualraum von V. Die Elemente von V^* heißen auch Linearformen von V.

Identifiziert man den K^n mit dem Raum $M_{n,1}(K)$ der n-dimensionalen Spaltenvektoren, $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ so kann man lineare Abbildungen $\lambda : K^n \to K$ als Zeilenvektoren $\lambda = (\alpha_1, \ldots, \alpha_n)$ schreiben. Die Berechung von $\lambda(x) = \sum \alpha_i \xi_i$ entspricht dann gerade der Matrizenmultiplikation

$$\lambda(x) = (\alpha_1, \ldots, \alpha_n) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Lemma 4.2. *Sei V ein n-dimensionaler K-Vektorraum, und sei $\vec{B} = (b_1, \ldots, b_n)$ eine angeordnete Basis von V. Die Koordinatenfunktionale λ_i, $i = 1, \ldots, n$ mit*

$$\lambda_i(\sum_j \xi_j b_j) = \xi_i$$

bilden eine Basis von V^*. Man schreibt auch $\lambda_i = b_i^*$. Vorsicht: λ_i hängt nicht nur b_i ab, sondern von ganz $\vec{B}$.

Beweis: $\lambda_1, \ldots, \lambda_n$ ist linear unabhängig in V^*. Sei $\sum \gamma_i \lambda_i = 0_{V^*}$. Zwei Funktionen sind gleich, gdw sie auf allen Argumenten übereinstimmen. Wir haben also $\forall v \in V \ \sum \gamma_i \lambda_i(v) =$

$0_{V^*}(v) = 0_K$. Wir spezialisieren nun, indem wir $v = b_j$ einsetzen für $j = 1, \ldots, n$. Dann ist $\sum \gamma_i \lambda_i(b_j) = \gamma_j = 0$ für jedes j.

Die λ_i erzeugen V^*. Sei $\lambda \in V^*$. Dann ist λ bestimmt durch seine Werte $\lambda(b_i)$, $i = 1, \ldots, n$. Durch Einsetzen von b_j, $j = 1, \ldots, n$ auf beiden Seiten rechnet man $\lambda = \sum_{i=1}^{n} \lambda(b_i)\lambda_i$ nach. □

Beobachtung 4.3. *$(K^n)^*$ ist isomorph zum Vektorraum der Zeilenvektoren der Länge n.*

Beweis: Wir definieren $e_i^* \in (K^n)^*$ durch $e_i^*(e_j) = \delta_{i,j}$. Dann ist $\{e_1^*, \ldots, e_n^*\}$ eine Basis von $(K^n)^*$. Die Abbildung $^*: K^n \to (K^n)^*$, die e_i auf e_i^* abbildet, bestimmt einen Vektorraumisomorphismus. e_i^* kann man sich als Zeile $(0, \ldots 0, 1, 0 \ldots, 0)$ mit dem Eintrag 1 an i-ter Stelle vorstellen. Dann ist der Wert $e_i^*(e_j)$ gerade das Matrizenprodukt der Zeile e_i und der Spalte e_j. □

Korollar 4.4. *Wenn V endlichdimensional ist, so ist V^* isomorph zu V.*

Beweis: Je zwei K-Vektorräume der gleichen Dimension sind isomoprh (siehe 4.13). Sei $n = \dim(V)$ und sei $\pi: V \overset{cong}{\to} K^n$. Seien e_i^* wie in 4.3 definiert. Wir definieren $\pi^*: (K^n)^* \to V^*$ durch

$$\pi^*(e_i^*) = e_i^* \circ \pi.$$

π^* heißt die zu π duale Abbildung . Man sagt auch π induziert die duale Abbildung π^*.

Dann ist

$$\begin{array}{ccc} V & \xrightarrow{h} & V^* \\ \cong\downarrow \pi & & \pi^*\uparrow\cong \\ K^n & \xrightarrow[\cong]{*} & (K^n)^* \end{array}$$

kommutativ und $h = \pi^* \circ {}^* \circ \pi$ ein Isomorphismus der gewünschten Art (der keineswegs eindeutig ist, sondern von *, π und π^* abhängt). □

Definition 4.5. Sei $\vec{B} = (b_1, \ldots, b_n)$ eine angeordnete Basis von V $b_i^*(b_j) = \delta_{i,j}$. Dann heißt $(b_1^*, \ldots, b_n^*)$ die zu $\vec{B}$ duale angeordnete Basis von V.

Bemerkung: Der bestimmte Artikel ist gerechtfertigt: Es gibt genau eine duale angeordnete Basis zur angeordneten Basis $\vec{B}$.

Definition 4.6. $V^{**} = (V^*)^*$ heißt der Bidualraum zu V.

Definition 4.7.

$$\begin{array}{rcl} \Phi: V & \to & V^{**} \\ x & \mapsto & (\lambda \mapsto \lambda(x), \lambda \in V^*) \end{array}$$

heißt die kanonische Abbildung von V nach V^{**}.

Lemma 4.8. *(AC, falls V unendlichdimensional.) Die kanonische Abbildung $\Phi: V \to V^{**}$ ist injektiv und linear.*

Beweis: Φ ist linear. Wenn $x \neq 0_V$, gibt es eine Linearform λ mit $\lambda(x) = \Phi(x)(\lambda) \neq 0$. Also ist Φ injektiv. □

Satz 4.9. *(ZFC) Die kanonische Abbildung $\Phi: V \to V^{**}$ ist bijektiv gdw V endlichdimensional ist.*

Beweis: Wenn V endlichdimensional ist, ist $\dim(V) = \dim(V^*) = \dim(V^{**})$ und somit ist jede injektive lineare Abbildung von V nach V^{**} surjektiv. □

4.2 Der Dualraum eines unendlichdimenisonalen Vektorraums

Wenn V unendlichdimensional ist, dann gibt es eine angeordnete Basis C von V, die o.B.d.A. mit $\{(i, c_i) : i \in \mathbb{N}\}$ beginnt. Wir nehmen $\lambda \in V^*$, so dass $\lambda(c_i) = 1$ für $i \in \mathbb{N}$ und $\lambda(c) = 0$ für $c \in C \setminus \{c_i : i \in \mathbb{N}\}$. Dann setzen wir die linear unabhängige Menge $\{\lambda\} \cup \{c_i^* : i \in \mathbb{N}\}$ zu einer Basis B von V^* fort. Wir nehmen $\mu \in V^{**}$ mit $\mu(\lambda) = 1$ und $\mu(b) = 0$ für $b \in B \setminus \{\lambda\}$. Dann ist $\mu \notin \text{bild}(\Phi)$: Wir nehmen das Gegenteil an: Sei $v \in V$ und $\Phi(v) = \mu$, d.h.

$$\Phi(v)(\lambda') = \lambda'(v) = \mu(\lambda') \text{ für alle } \lambda' \in V^*. \qquad (*)$$

Nun ist $v = \sum_{j \in J_0} \alpha_j c_j + \sum_{c \in C_0} \alpha_c c$ für ein endliches $J_0 \subseteq \mathbb{N}$ und ein endliches $C_0 \subseteq C$. Wir setzen $\lambda' = \lambda$ ein in Gleichung $(*)$ und erhalten $\lambda(v) = \sum_{j \in J_0} \alpha_j = \mu(\lambda) = 1$. Wir setzen für $i \in J_0$, $\lambda' = c_i^*$ ein in Gleichung $(*)$ und erhalten $c_i^*(v) = \alpha_i = \mu(c_i^*) = 0$. Dies ist ein Widerspruch. □

Wenn V endlichdimensional ist, sind nach Obigem V und V^{**} isomorph. Für unendlichdimensionale V liegt hingegen ein stark konträrer Sachverhalt vor: Nicht nur die kanonische Einbettung, sondern jede Einbettung ist nicht surjektiv. Der Dualraum wird exponentiell groß.

Definition 4.10. Seien K, C Mengen. K^C ist die Menge aller Funktionen von C nach K.

Der Beweis des folgenden Satzes reicht über die Anfängervorlesung hinaus.Er benutzt wiederum den Vandermonde-Trick und daneben noch leichte Abschätzungen aus der Kardinalzahlenarithmetik. Von der Linearen Algebra-Vorlesung 2012/13 wurde die Frage, ob V^{**} isomorph zu V sein kann im Unendlichdimensionalen, an uns herausgetragen, und hier ist eine Antwort. Es wird nicht erwartet, dass man diesen Stoff in einer Prüfung über Lineare Algebra beherrscht.

Satz 4.11. *(ZFC) Wenn V unendlichdimensional ist, C eine Basis von V ist und B eine Basis von V^* ist, gibt es*

(1) eine Bijektion zwischen B und V^* und

(2) eine Bijektion zwischen V^* und K^C und

(3) keine Surjektion von C auf B.

Hierbei ist K^C die Menge aller Funktionen von C nach K.

Beweis: Wir zeigen (1), nach einem Beweis von Martin Ziegler. Sei B eine Basis von V^*, also $|B| = \dim(V^*)$, wenn man schon unendliche Mächtigkeiten kennt. (Hier wurde AC benutzt.)

Nun gilt, sobald B oder K unendlich sind: Es gibt eine Bijektion von $K \cup B$ auf V^*. Hierzu braucht man das Auswahlaxiom dafür, dass B und K Mächtigkeiten haben. Wir zeigen zuerst, dass es eine Injektion von V^* in $B \cup K$ gibt. Dieser Beweis verwendet Zitate, die erst in der Vorlesung Mathematische Logik bewiesen werden, nämlich den Satz von Gerhard Hessenberg[1] $\kappa \times \kappa \cong \kappa$ für unendliche Mengen κ. Wir machen diese Schritte anschaulich plausibel: $\lambda = \sum_{b \in B_0} \alpha_b b$ ist ein beliebiges Element von V^*. $B_0 \subseteq B$ ist endlich. Wir kodieren (d.h. bilden λ injektiv ab) λ durch $\{(b, \alpha_b) : b \in B_0\}$. Nun schätzen wir die Menge der Codes ab. Diese ist $\{K^{B_0} : B_0 \subseteq B, B_0 \text{ endlich}\}$, und letzteres kann man injektiv in $K \cup B$ abbilden, denn $B \times B \cong B$, und daher $\bigcup_{n \in \mathbb{N}} B^n \cong B$ und $\{B_0 : B_0 \subseteq B, B_0 \text{ endlich}\} \cong B$. Aus den gleichen Gründen gilt $K^{B_0} \cong K$, und $K \times B \cong K \cup B$.

Wir benutzen zur genauen Konstruktion der behaupteten Bijektionen den erwähnten Satz von Hessenberg und einige leichte Kardinalzahlarithmetik, z.B. [16] oder ein Mengenlehrebuch. (Diese Konstruktionen von Bijektionen und Injektionen geht wieder in ZF.) Außerdem hilft folgender Kunstgriff: Aus zwei Injektionen $A \to B$ und $B \to A$ kann man eine Bijektion $A \cong B$ bauen nach dem Satz von Cantor, Schröder und Bernstein, [6] oder sehr gut auf der englischen oder der deutschen Wikipedia.

Nun zeigen wir: Es gibt eine Injektion von K nach B: Sei C eine Basis von V, die o.B.d.A. mit paarweise verschiedenen c_i, $i \in \mathbb{N}$, beginnt. (Hier haben wir wieder AC benutzt, nun geht es in ZF weiter.) Für paarweise verschiedene $\beta_1, \dots, \beta_n \in K$ ist

$$(1, \beta_1, \beta_1^2, \beta_1^3, \dots, \beta_1^{n-1}), (1, \beta_2, \dots, \beta_2^{n-1}), \dots, (1, \beta_n, \beta_n^2, \beta_n^3, \dots, \beta_n^{n-1})$$

linear unabhängig in V^* nach Satz 3.57. Hierbei ist $(\alpha_0, \dots, \alpha_{n-1}) \in V^*$ dasselbe wie $\sum_{i=0}^{n-1} \alpha_i c_i^*$ und $c_i^*(c_j) = \delta_{i,j}$ und $c_i^*(c) = 0$ für $c \in C \setminus \{c_j : j \in \mathbb{N}\}$. Nun ist auch $(1, \beta, \beta^2, \dots) = \sum_{i \in \mathbb{N}} \beta^i c_i^* \in V^*$. Nach der Rechnung mit der Vandermonde'schen Determinante ist jede endliche Teilmenge von $\{(1, \beta, \beta^2, \dots) : \beta \in K\}$ linear unabhängig.

[1] Gerhard Hessenberg, 1874 – 1925

Ausführlich zeigt man dies so: Sei $n \in \mathbb{N} \setminus \{0\}$ beliebig, und seien $\beta_1, \dots \beta_n$ paarweise verschiedene Element aus K. $\sum_{i=1}^n \alpha_i(1, b_i^1, \dots) = 0_{V^*}$ mit paarweise verschiedenen β_i. Dann ist für $c_0, \dots, c_{n-1}$ $\sum_{i=1}^n \alpha_i(1, b_i^1, \dots)(c_j) = \sum_{i=1}^n \alpha_i \beta_i^{j-1} = 0$. Wenn wir $x = (\alpha_1, \dots, \alpha_n)$ als Spaltenvektor sehen, haben wir also ein lineares Gleichungssystem $Ax = 0_{K^n}$ mit eine Vandermondematrix A, deren j-te Zeile gerade $(\beta_1^{j-1}, \beta_2^{j-1}, \dots, \beta_n^{j-1})$ ist. Da die Vandermondematrix regulär ist, ist $\alpha_1 = \dots = \alpha_n = 0$ die einzige Lösung.

Also ist

$$B' := \{(1, \beta, \beta^2, \dots) : \beta \in K\}$$

eine linear unabhängige Menge und die Abbildung

$$i: \beta \mapsto (1, \beta, \beta^2, \dots)$$

ist eine Injektion von K in eine linear unabhängige Menge B'. Nun arbeiten wir in ZFC und ergänzen B' zu einer Basis B'' von V und nehmen dann eine Bijektion b zwischen B'' und B mit dem unbewiesenen Satz 4.13. Dann ist $b \circ i: K \to B$ wie gewünscht.

Aus der Bijektion zwischen V^* und $B \cup K$ (in der o.B.d.A. eine Hälfte von B ausgelassen wird AC) und der Injektion von K in (eine Hälfte von, hier wieder AC) B erhalten wir: Es gibt eine Injektion von V^* in B. Umgekehrt ist die Identität eine Injektion von B in V^*. Nun haben wir also eine Bijektion zwischen B und V^*, und (1) ist gezeigt.

(2) Es gibt folgende Bijektion zwischen V^* und K^C: Jedes $f: C \to K$ bestimmt durch lineare Fortsetzung auf V ein Element $\lambda \in V^*$ mit $\lambda \restriction C = f$. Falls $f, g: C \to K$, $f \neq g$, so bestimmen sie zwei verschiedene Formen in V^*. Jedes $\lambda \in V^*$ ist durch seine Einschränkung auf C bestimmt.

(3) Da K mindestens zwei Elemente hat, gibt es eine Injektion von $\{0,1\}^C$ nach K^C (Übung 3(iii) auf Blatt 3). Nach dem Satz von Cantor (s.u.) gibt es keine Surjektion von C nach $\{0,1\}^C \cong \mathscr{P}(C)$. Daher gibt es auch keine Surjektion von C auf K^C und keine Surjektion von C auf B. □

Wir zeigen hier noch den eben benutzten Satz von Cantor:

Satz 4.12. *(Cantor) Sei C eine Menge. Sei $f: C \to \mathscr{P}(C)$. Dann ist f nicht surjektiv.*

Beweis: Nach dem Aussonderungsschema gibt es die Menge

$$D = \{b \in C : b \notin f(b)\},$$

die sogenannte Diagonalmenge. Dann ist für $b \in C$, $f(b) \neq D$. Denn, angenommen $f(b) = D$. Dann ist $b \in D$ gdw $b \notin f(b)$, also $b \in D$ gdw $b \notin D$. Widerspruch. □

Ohne Beweis schreiben wir folgenden wichtigen Satz:

Satz 4.13. ZFC *Zwischen je zwei Basen eines Vektorraums gibt es eine Bijektion. Anders gesagt: Je zwei Basen desselben Vektorraums sind gleichmächtig. Die Dimension von V ist definiert als die Mächtigkeit einer Basis.*

Man braucht zum Beweis Kenntnisse über Mächtigkeiten, auch Kardinalzahlen genannt, die zum Beispiel in der Logik-Vorlesung vermittelt werden. Sehr elegant und kurz kann man dies zum Beispiel in Kapitel 2 des Ziegler Skripts lesen `http://home.mathematik.uni-freiburg.de/ziegler/skripte/logik.pdf`. An linearer Algebra braucht man nur die Techniken aus Kapitel 2 dieses Skripts. Bei Interesse können Sie also den Beweis später selbst führen.

Bemerkung 4.14. Hier ist noch ohne Beweis eine Eigenschaft von Mächtigkeiten, die aus ZFC folgt und manchmal stillschweigend (in Satz 4.11 und in diesem Skript aber außer in dieser Bemerkung gerade nicht) verwendet wird: Seien A, B nicht leere Mengen. Dann sind äquivalent

(1) $|A| \leq |B|$.

(2) Es gibt eine Injektion von A nach B.

(3) Es gibt eine Surjektion von B auf A.

Wenn man mit Satz 4.13 die Definition von Dimension auch für unendliche Dimensionen hat, kann man 4.11 (3) zusammen mit „non (3) impliziert non (1)" umformulieren zu: Die Dimension von V ist echt kleiner als die Dimension von V^*.

Wenn man Mächtigkeiten kennt, kann man sagen: In Satz 4.11(1) und (2) wird

$$|V^*| = \dim(V^*) \text{ und } \dim(V^*) = |K|^{\dim V}$$

bewiesen. Damit endet die Bemerkung.

Korollar 4.15. ZFC *Sei V ein unendlichdimensionaler K-Vektorraum, und sei C eine Basis von V. Sei D eine Basis von V^{**}. Dann gibt es keine Surjektion von C auf D.*

Beweis: Sei B eine Basis von V^*, $b_0 \in B$. Sei D' eine Basis von V^{**}, die B^* als Teilmenge enthält (Basisergänzungssatz) und bijektiv zu D ist. Annahme: Es gibt eine Surjektion $s: C \to D'$. Wir zeigen: Wir können s umbauen zu einer Surjektion von C auf B. Wir setzen

$$s'(c) = \begin{cases} b, & \text{wenn } s(c) = b^* \in D', \\ b_0, & \text{wenn } s(c) \in D' \setminus \{b^* : b \in B\}. \end{cases}$$

Dann ist $s': C \to B$ surjektiv. Die Annahme, dass es eine Surjektion $s: C \to D'$ gibt, ist also falsch. Da $D \cong D'$, gibt es auch keine Surjektion $s: C \to D$. □

Ab hier gehört der Stoff wieder zur Vorlesung des ersten Jahres.

Definition 4.16. Ein Raum V heißt reflexiv, wenn $V \cong V^{**}$.

Es gibt also nach Korollar 4.15 keinen unendlichdimensionalen reflexiven Vektorraum. Bei anderen Räumen in der Funktionanalysis und Definitionen von V^*, die nur mit stetigen Formen und Funktionen arbeiten, ist Reflexivität hingegen ein wichtiges, vorkommendes und erwünschtes Phänomen.

4.3 Duale Basen

Identifiziert man wie oben die Elemente von K^m und K^n mit Spaltenvektoren und die Elemente der Dualräume mit Zeilenvektoren, so wird für eine m-n-Matrix A die Abbildung $f = f_A: K^n \to K^m$ gegeben durch

$$f: x \mapsto Ax.$$

und f^* durch

$$f^*: \lambda \mapsto \lambda A$$

weil

$$(\lambda A)x = \lambda(Ax).$$

Dieser an sich einfache Sachverhalt wird nun komplizierter, wenn man auch die dualen Abbildungen immer durch Spalten beschreiben möchte:

Lemma 4.17. *Seien $\vec{B} = (b_1, \dots, b_n)$ und $\vec{B}' = (b'_1, \dots, b'_n)$ zwei angeordnete Basen von V, und seien $(b_1^*, \dots, b_n^*)$ und $((b'_1)^*, \dots, (b'_n)^*)$ die entsprechenden dualen angeordneten Basen. Wir nehmen an, dass $b'_j = \sum_{i=1}^n \alpha_{i,j} b_i$, also $(b'_1 \dots, b'_n) = (b_1, \dots, b_n)A$. Dann ist $\begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix} = A \begin{pmatrix} (b'_1)^* \\ \vdots \\ (b'_n)^* \end{pmatrix}$, oder, häufiger so geschrieben $(\begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix})^\top = (\begin{pmatrix} (b'_1)^* \\ \vdots \\ (b'_n)^* \end{pmatrix})^\top A^\top$ und formuliert: Der Übergang der dualen angeordneten Basen wird vermittelt durch das Inverse der transponierten Matrix.*

Beweis:

$$\begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix} \cdot \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} = 1_{M_{n,n}(K)}$$

und ebenso für die gestrichenen Basen. Also ist

$$1_{M_{n,n}(K)} = \begin{pmatrix} (b_1')^* \\ \vdots \\ (b_n')^* \end{pmatrix} \cdot \begin{pmatrix} b_1' & \dots & b_n' \end{pmatrix} = X^\top \begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix} \cdot \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} A = X^\top A$$

und $X^\top = A^{-1}$. □

4.4 Duale Abbildungen

Definition 4.18. Sei $\pi \in \mathrm{Hom}(V, W)$. Dann definieren wir $\pi^* \in \mathrm{Hom}(W^*, V^*)$ durch

$$\pi^*(\lambda) = \lambda \circ \pi \text{ für } \lambda \in W^*.$$

π^* heißt die zu π duale Abbildung.

Lemma 4.19. *(A) Die Dualisierung* $^*\colon \mathrm{Hom}(V,W) \to \mathrm{Hom}(W^*, V^*)$, $\pi \mapsto \pi^*$ *ist ein kontravarianter linearer Funktor: Das heißt, es gelten die Punkte (1) bis (4):*

(1) $\mathrm{id}_V^* = \mathrm{id}_{V^*}$.

(2) $f \in \mathrm{Hom}(U,V)$, $g \in \mathrm{Hom}(V,W)$. *Dann ist* $(g \circ f)^* = f^* \circ g^*$.

(3) Wenn f *und* $g \in \mathrm{Hom}(V,W)$. *Dann ist* $(f+g)^* = f^* + g^*$.

(4) Für $\alpha \in K$ *ist* $(\alpha f)^* = \alpha f^*$.

*(B) (*ZFC*, falls* W *unendlichdimensional.) Außerdem ist der Funktor treu, d.h.: Aus* $f^* = 0$ *folgt* $f = 0$.

*(C) (*ZFC*, falls* V *unendlichdimensional.) Für endlichdimensionales* W *ist der Funktor voll, d.h.: Für jedes* $g \in \mathrm{Hom}(W^*, V^*)$ *gibt es* $f\colon V \to W$ *mit* $f^* = g$.

*(*ZFC*.) Für unendlichdimensionales* W *und* $V \neq \{0\}$ *ist der Funktor nicht voll.*

Beweis: (A) (1) $(\mathrm{id}_V)^*(\lambda) = \lambda(\mathrm{id}_V) = \lambda = \mathrm{id}_{V^*}(\lambda)$.

(2) $(g \circ f)^*(\lambda) = \lambda(g \circ f) = g^*(\lambda) \circ f f^*(g^*(\lambda)) = (f^* \circ g^*)(\lambda)$.

(3) $(f+g)^*(\lambda) = \lambda(f+g) = f^*(\lambda) + g^*(\lambda) f^*(g^*(\lambda)) = (f^* + g^*)(\lambda)$.

(4) $(\alpha f)^*(\lambda) = \lambda(\alpha f) = \alpha f^*(\lambda) f^*(g^*(\lambda)) = (\alpha f^*)(\lambda)$.

(B) Wenn $f^* = 0 \in \mathrm{Hom}(W^*, V^*)$, und $x \in V$, so ist für $\lambda \in W^*$, $\lambda(f(x)) = f^*(\lambda)(x) = 0$. Da dies für alle $\lambda \in W^*$ und alle $x \in V$ gilt, ist $f = 0 \in \mathrm{Hom}(V,W)$.

(C) Sei $(b_1^*, \dots, b_n^*)$ eine angeordnete Basis von W^*, die zur angeordneten Basis $(b_1, \dots, b_n)$ von W dual ist. Sei $C_0 = \{\lambda_1, \dots, \lambda_m\} \subseteq V^*$ eine Basis von $\mathrm{span}(\{g(b_i^*) : i = 1, \dots, n\}$, also $|C_0| =: m \leq n$.

Sei $g \in \mathrm{Hom}(W^*, V^*)$ gegen durch

$$\text{für } k = 1, \dots, n \quad g(b_k^*) = \sum_{i=1}^{m} \alpha_{i,k} \lambda_i. \tag{$\bullet$}$$

Gesucht ist ein $\varphi \in \mathrm{Hom}(V, W)$, so dass $\varphi^* = g$.

Nach Satz 1.88 ist $\dim(V/\bigcap_{j=1,\dots,m} \ker(\lambda_j)) = m$.

Man sieht leicht (in einigen Beweisschritten, die hier ausgelassen sind): Zur angeordneten linear unabhängigen Menge Es gibt linear unabhängige

$$[d_1], \dots, [d_m] \in V/(\bigcap_{j=1,\dots m} \ker(\lambda_j)),$$
$$\text{so dass für alle } i, j \in \{1, \dots, m\}, \lambda_i(d_j) = \delta_{i,j}.$$

Sei $D_0 = \{d_i : 1 \leq i \leq m\} \subseteq V$. Die Vektoren $d_1, \dots d_m$ sind auch in V linear unabhängig.

Sei $D \supseteq D_0$ eine Basis von V, so dass $(D \setminus D_0) \subseteq \bigcap_{i=1}^{m} \ker(\lambda_j)$. Nach Satz 1.88 gibt es so ein D.

Sei $\varphi \in \mathrm{Hom}(V, W)$ gegeben durch:

$$\begin{gathered} \text{Für } i = 1, \dots, m \text{ sei } \varphi(d_i) = \sum_{j=1}^{n} \alpha_{i,j} b_j. \\ \text{Für } d \in D \setminus D_0 \text{ sei } \varphi(d) = 0_W. \end{gathered} \tag{$\bullet\bullet$}$$

Wir zeigen, dass $\varphi^* = g$:

Wir sehen für $k = 1, \dots, n$, $j = 1, \dots, m$,

$$\varphi^*(b_k^*)(d_j) = b_k^*(\varphi(d_j)) = \alpha_{j,k} = g(b_k^*)(d_j).$$

Wir sehen für $k = 1, \dots, n$, $d \in D \setminus D_0$,

$$\varphi^*(b_k^*)(d) = b_k^*(\varphi(d)) = b_k^*(0) = 0_K = g(b_k^*)(d).$$

Wir geben noch einen einfacheren Beweis für (C) im Fall, dass auch V endlichdimensional ist. Dann haben $\mathrm{Hom}(V, W)$ und $\mathrm{Hom}(W^*, V^*)$ beide die Dimension $\dim(V) \cdot \dim(W)$, und die Abbildung

$$^*\colon \mathrm{Hom}(V, W) \to \mathrm{Hom}(W^*, V^*)$$

ist injektiv nach dem vorigen Punkt, also surjektiv.

(ZFC). Dies geht nun wie der Beweis des Satzes 4.9. Sei nun W unendlichdimensional. Sei C eine angeordnete Basis von W, die mit $\{(i, c_i) : i \in \mathbb{N}\}$ beginnt. Sei $b \in V^*$ nicht der Nullvektor. Wir definieren $\lambda \in \mathrm{Hom}(W^*)$, $\lambda = \sum_{i \in \mathbb{N}} c_i^*$. Hierbei ist $c_i^*(c_j) = \delta_{i,j}$ und $c_i^*(c) = 0$

für $c \in C \setminus \{c_i : i \in \mathbb{N}\}$. Wie in Satzes 4.9 zeigt man, dass $L = \{\lambda\} \cup \{c_i : i \in \mathbb{N}\}$ linear unabhängig ist. Wir setzen L zu einer Basis D von W^* fort. Sei $D' = D \setminus (\{c_i^* : i \in \omega\} \dot\cup \{\lambda\})$ eine Basis von W^*. Wir definieren $\mu \in \mathrm{Hom}(W^*, V^*)$ durch $\mu(\lambda) = b$ und $\mu(d) = 0_{V^*}$ für $d \in \{c_i^* : i \in \mathbb{N}\} \cup D'$. Dann ist für $f \in \mathrm{Hom}(V, W)$, $f^* \neq \mu$. Wir nehmen das Gegenteil an: Sei

$$(\lambda' \circ f)(v) = \mu(\lambda')(v) \text{ für alle } \lambda' \in W^*,\ v \in V. \qquad (\diamond)$$

Wir nehmen ein v mit $b(v) \neq 0$. Nun ist $f(v) = \sum_{j \in J_0} \alpha_j c_j + \sum_{c \in C_0} \alpha_c c \in W$ für ein endliches $J_0 \subseteq \mathbb{N}$ und ein endliches $C_0 \subseteq C$ und geeignete $\alpha_j \in K$, $\alpha_c \in K$. Wir setzen $\lambda' = \lambda$ ein in Gleichung $(\diamond)$ und erhalten $\lambda(f(v)) = \sum_{j \in J_0} \alpha_j = \mu(\lambda)(v) = b(v) \neq 0$. Wir setzen für $i \in J_0$, $\lambda' = c_i^*$ ein in Gleichung $(\diamond)$ und erhalten $c_i^*(f(v)) = \alpha_i = \mu(c_i^*)(v) = 0_{V^*}(v) = 0$. Dies ist ein Widerspruch. □

Wir erwähnen noch einmal: Identifiziert man wie oben die Elemente von K^m und K^n mit Spaltenvektoren und die Elemente der Dualräume mit Zeilenvektoren, so wird für eine m-n- Matrix A die Abbildung $f = f_A : K^n \to K^m$ gegeben durch

$$f : x \mapsto Ax.$$

und f^* durch

$$f^* : \lambda \mapsto \lambda A$$

weil

$$(\lambda A)x = \lambda(Ax).$$

Nun folgt wieder ein Übergang in die Spaltenschreibweise, ein Pendant to Lemma 4.17:

Satz 4.20. *Seien V und W endlichdimensionale K-Vektorräume, und sei $f : V \to W$ linear. Sei $\vec{B}$ eine angeordnete Basis von V, sei $\vec{C}$ eine angeordnete Basis von W. Wenn* $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) = A$, *so ist* $\mathrm{Mat}_{\vec{C}^*}^{\vec{B}^*}(f^*) = A^\top$.

Beweis: $\vec{B}^* \vec{B} = 1_{M_{n,n}(K)}$, $\vec{C}^* \vec{C} = 1_{M_{m,m}(K)}$.

Nach Konvention heißt $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) = A$ dasselbe wie: Für alle b_j ist $f(b_j) = \sum_i c_i \alpha_{i,j}$. Letzteres ist auch äquivalent zu: $f(\vec{B}) = \vec{C}A$. Nun ist $X = \mathrm{Mat}_{\vec{C}^*}^{\vec{B}^*}(f^*)$ gesucht, d.h., $f^*((\vec{C}^*)^\top) = (\vec{B}^*)^\top X$, denn Mat ist ja in Spaltenschreibweise. In Zeilen heißt dies $f^*(\vec{C}^*) = X^\top \vec{B}^*$. Dann ist

$$X^\top = X^\top \vec{B}^* \vec{B} = f^*(\vec{C}^*)\vec{B} = \vec{C}^* f(\vec{B}) = \vec{C}^* \vec{C} A = A.$$

Also ist $X = A^\top$. □

Bemerkung: Falls nur W endlichdimensional ist, hat man in $(\bullet)$ und $(\bullet\bullet)$ auch gerade die Transposition, nur mit dem Unterschied, dass ein behauptetes f^* gegeben ist und f gesucht ist.

Korollar 4.21. *Sei V endlichdimensional und sei $\varphi \in \mathrm{End}(V)$. Dann ist* $\det(\varphi^*) = \det(\varphi)$.

Lemma 4.22. *(1) Der Dualraum einer direkten Summe lässt sich auf natürliche Weise mit der direkten Summe der Dualräume identifizieren.* $(V_1 \oplus V_2)^* = V_1^* \oplus V_2^*$.

(2) Das Duale der Einbettung $V_1 \to V_1 \oplus V_2$ ist die Projektion $V_1^* \oplus V_2^* \to V_1^*$. Das Duale der Projektion $V_1 \oplus V_2 \to V_1$ ist die Einbettung $V_1^* \to V_1^* \oplus V_2^*$.

Beweis: (1) Jede lineare Abbildung $\lambda \in (V_1 \oplus V_2)^*$ ist bestimmt durch ihrer beiden Einschränkungen $\lambda_i = \lambda \restriction V_i$. Wir identifizieren hier V_1 mit seiner Einbettung $V_1 \times \{0_{V_2}\}$ in $V_1 \oplus V_2$ und verfahren analog mit v_2. Die Abbildung $\lambda \mapsto \lambda_1 \oplus \lambda_2$ mit $(\lambda_1 \oplus \lambda_2)(v) = \lambda_1(v_1) + \lambda_2(v_2)$ für $v = (v_1, v_2) \in V_1 \oplus V_2$ ist der gewünschte Isomorphismus.

(2) Sei $\varphi(v_1) = (v_1, 0) \in V_1 \times V_2$ die Einbettung. Dann ist für $\lambda \in V_1^* \oplus V_2^*$ $\varphi^*(\lambda)(v_1) = \lambda \circ \varphi(v_1) = \lambda(v_1, 0) \in K$.

Sei $\pi(v_1, v_2) = v_1$ die Projektion. Dann ist für $\lambda_1 \in V_1^*$ $\pi^*(\lambda_1)(v_1, v_2) = \lambda_1(\pi(v_1, v_2)) = \lambda_1(v_1) = \lambda_1(v_1) + 0_{V_2^*}(v_2) \in K$. □

Satz 4.23. *(*AC*, falls W unendlichdimensional). Sei $f \in \mathrm{Hom}(V, W)$. Dann gilt*

(1) f ist genau dann injektiv, wenn f^* surjektiv ist.

(2) f ist genau dann surjektiv, wenn f^* injektiv ist.

(3) Wenn W endlichdimensional ist, haben f und f^* den gleichen Rang. Für unendlichdimensionales $\mathrm{bild}(f)$ *gilt die Aussage nicht. Dies folgt aus Satz 4.11.*

Beweis: Wir geben einen Beweis für endlichdimensionale V, W. Seien $\vec{B} = (b_1, \ldots, b_n)$ und $\vec{C} = (c_1, \ldots, c_m)$ angeordnete Basen von V bzw. W, sei $\mathrm{Mat}_{\vec{B}}^{\vec{C}}(f) = A \in M_{m,n}(K)$. Dann gilt

$$f \text{ surj.} \Leftrightarrow \mathrm{rang}(A) = \dim(W) = m \Leftrightarrow \mathrm{rang}(A^\top) = m = \dim(W^*) \Leftrightarrow f^* \text{ inj.} \tag{1}$$

$$f \text{ inj.} \Leftrightarrow \mathrm{rang}(A) = \dim(V) = n \Leftrightarrow \mathrm{rang}(A^\top) = n = \dim(V^*) \Leftrightarrow f^* \text{ surj.} \tag{2}$$

$$\mathrm{rang}(f) = \mathrm{rang}(A) = \mathrm{rang}(A^\top) = \mathrm{rang}(f^*). \tag{3}$$

Wir geben noch einen Beweis, der auch für unendlichdimensionale V, W funktioniert.

1. Sei f injektiv und $\lambda \in V^*$. Wir sollen ein $\mu \in W^*$ finden mit $f^*(\mu) = \lambda$. Wir nehmen $\mu(w) = \lambda(f^{-1}(w))$ für $w \in \mathrm{bild}(f)$ und setzen μ auf ganz W fort (hier AC, falls W unendlichdim.). Dann ist für jedes $v \in V$, $f^*(\mu)(v) = \mu(f(v)) = \lambda(v)$.

Sei f^* surjektiv und sei $f(v) = 0$. Wir sollen $v = 0$ zeigen. Wir zeigen hierfür: $\lambda(v) = 0$ für jedes $\lambda \in V^*$. Da f^* surjektiv ist, ist $\lambda = f^*(\mu)$ für ein $\mu \in W^*$. $\lambda(v) = f^*(\mu)(v) = \mu(f(v)) = \mu(0) = 0$.

2. Sei f nicht surjektiv. Dann gibt es $w \in W \setminus \mathrm{Bild}(f)$. AC Wir ergänzen w zu einer Basis C von W und setzen $w^*(w) = 1$ und $w^*(c) = 0$ für $c \in C \setminus \{w\}$. Dann ist $w^* \in \ker(f^*)$, denn für jedes $v \in V$ ist $f^*(w^*)(v) = w^*(f(v)) = 0$.

Sei f surjektiv und sei $f^*(\lambda) = 0$. Wir sollen $\lambda = 0$ zeigen. Wir zeigen hierfür: $\lambda(w) = 0$ für jedes $w \in W$. Da f surjektiv ist, ist $w = f(v)$ für ein $v \in V$. $\lambda(w) = \lambda(f(v)) = f^*(\lambda)(v) = 0$.

3. Der Rang von f ist die Dimension von $\mathrm{bild}(f)$. Wir nehmen an: $w_1, \ldots, w_m \in W$ bilden eine Basis von $\mathrm{bild}(f)$. Wir ergänzen $\{w_1, \ldots, w_m\}$ zu einer Basis C von W. Dann sei $w_i^* \in W^*$ definiert durch $w_i^*(w_j) = \delta_{i,j}$ und $w_i^*(c) = 0$ für $c \in C \setminus \{w_1, \ldots, w_m\}$. Wir behaupten: $f^*(w_1^*), \ldots, f^*(w_m^*) \in W^*$ ist eine Basis von $\mathrm{bild}(f^*)$.

$f^*(w_1^*), \ldots, f^*(w_m^*)$ sind linear unabhängig: Sei $\sum_{i=1}^m \alpha_i f^*(w_i^*) = 0_{V^*}$, d.h., $\forall v \in V$, $\sum_{i=1}^m \alpha_i f^*(w_i^*)(v) = 0_K$. Dann ist insbesondere für v_k mit $f(v_k) = w_k$, $k = 1, \ldots, m$, $\sum_{i=1}^m \alpha_i f^*(w_i^*)(v_k) = \sum_{i=1}^m \alpha_i (w_i^*) \cdot f(v_k) = \sum_{i=1}^m \alpha_i (w_i^*)(w_k) = \alpha_k = 0_K$.

$f^*(w_1^*), \ldots, f^*(w_m^*)$ sind ein Erzeugendensystem für $\mathrm{bild}(f^*)$. Sei $\lambda \in \mathrm{bild}(f^*) \subseteq V^*$. Dann gibt es ein $\mu \in W^*$, so dass $\lambda = f^*(\mu) = \mu \circ f$. Wir bilden $\mu(w_i)$, $i = 1, \ldots m$. Dann ist $\mu = \sum_i \mu(w_i) w_i^*$ und $\lambda = \sum_{i=1}^m \mu(w_i) f^*(w_i^*)$. Denn für jedes $v \in V$ ist $\sum_{i=1}^m \mu(w_i) f^*(w_i^*)(v) = (\sum_i \mu(w_i) w_i^*)(f(v)) = \mu(f(v)) = \lambda(v)$, wie gewünscht. Die Basis $\{f^*(w_1^*), \ldots, f^*(w_m^*)\}$ bezeugt also $\dim(\mathrm{bild}(f)) = \dim(\mathrm{bild}(f^*))$.

Hier ist noch ein Beweis mit einem Diagramm. Es gibt einen zu $\ker(f)$ in V komplementären Unterraum X und einen zu $\mathrm{Im}(f)$ komplementären Unterraum Y in W. Nun rechnet man nach, dass folgendes Diagramm kommutiert:

$$\begin{array}{ccccccc}
\ker(f) + X & \xrightarrow{\pi} & X & \xrightarrow[f]{\cong} & \mathrm{Im}(f) & \xrightarrow[\mathrm{id}]{} & \mathrm{Im}(f) + Y \\
\downarrow{\scriptstyle *} & & \downarrow{\scriptstyle *\restriction X} & & \downarrow{\scriptstyle *\restriction \mathrm{Im}(f)} & & \downarrow{\scriptstyle *} \\
X^* + (\ker(f))^* & \xleftarrow{\mathrm{id}} & X^* & \xleftarrow[\cong]{f^*} & (\mathrm{Im}(f))^* & \xleftarrow{\pi} & (\mathrm{Im}(f))^* + Y^*
\end{array}$$

□

4.5 Duale Paare

Definition 4.24. (1) Seien V, W und U K-Vektorräume. Eine Abbildung

$$(,): V \times W \to U$$

heißt bilinear , wenn $(,)$ in beiden Argumenten linear ist, d.h. $(\alpha v + \beta v', w) = \alpha(v, w) + \beta(v', w)$ und $(v, \alpha w + \beta w') = \alpha(v, w) + \beta(v, w')$.

(2) Seien V und W, K-Vektorräume. Eine bilineare Abbildung

$$(,):V\times W\to K$$

heißt Bilinearform .

(3) Sei V K-Vektorraum. Eine Abbildung $(,):V\times V\to K$ heißt symmetrisch falls für alle $x,y\in V$ $(x,y)=(y,x)$.

(4) Sei V K-Vektorraum. Eine symmetrische Abbildung $(,):V\times V\to K$ heißt positiv definit falls für alle $x\in V$ $(x,x)=0\to x=0_V$.

(5) Sei V K-Vektorraum. Eine symmetrische positiv definite Bilinearform $(,):V\times V\to K$ heißt Skalarprodukt .

(6) Ein $\mathbb{R}$-Vektorraum mit einem Skalarprodukt heißt euklidischer Vektorraum..

Seien V und W endlichdimensionale K-Vektorräume, und sei

$$(,):V\times W\to K$$

eine Bilinearform. Seien $\vec{B}=(b_1,\dots,b_m)$ eine geordnete Basis von V und $\vec{C}=(c_1,\dots,c_n)$ eine geordnete Basis von W. $(,)$ ist durch eine m-n-Matrix

$$B=(\beta_{i,j})_{i=1,\dots,m;j=1,\dots,n}=(b_i,c_j)_{i=1,\dots,m;j=1,\dots,n}$$

eindeutig beschrieben. Wenn wir die Koordinaten von $v\in V$ und $w\in W$ bezüglich der gegebenen Basen als Spaltenvektoren x und y schreiben ergibt sich $(v,w)=(x^\top\vec{B},\vec{C}y)=x^\top(\vec{B}^\top,\vec{C})y=x^\top By$.

Die auf $K^m\times K^n$ bezüglich der kanonischen Basen gegebene Bilinearform bezeichnen wir mit $(,)_B$. Fasst man die Elemente von K^m und von K^n als Spaltenvektoren auf, so ist $(x,y)_B=x^\top By$.

Lemma 4.25. *Wenn $\vec{C}'=\vec{C}E$ und $\vec{B}'=\vec{B}D$, und $(,)$ durch die Basen $\vec{B}$ auf V und $\vec{C}$ auf W beschrieben wird, so wird $(,)$ in den Basen $\vec{B}'$ und $\vec{C}'$ durch $B'=D^\top BE$ beschrieben.*

Beweis $(\vec{B}'^\top,\vec{C}')=((\vec{B}D)^\top,\vec{C}E)=(D^\top\vec{B}^\top,\vec{C}E)=D^\top BE$.

Definition 4.26. Der Rang von $(,)$ ist der Rang einer Matrixdarstellung von $(,)$. Dieser Rang hängt nicht von der Wahl der Basen ab.

Definition 4.27. Das Paar V,W heißt duales Paar, wenn es eine Bilinearform $(,):V\times W\to K$ gibt, so dass $\dim(V)$ endlich ist und $\dim(V)=\dim(W)=\mathrm{Rang}((,))$.

Beispiele 4.28. (1) Wenn (V,W) ein duales Paar sind mit bezeugender Bilinearform $(,)$, dann sind auch (W,V) ein duales Paar, mit $(y,x)'=(x,y)$.

(2) Wenn W endlichdimensional ist, und $V = W^*$, dann sind (W^*, W) ein duales Paar mit erzeugender Bilinearform $(\lambda, x) = \lambda(x)$.

Satz 4.29. *Seien V, W K-Vektorräume, Dann entspricht jeder Bilinearform $(,)$ eine lineare Abbildung $\Phi_{(,)}: V \to W^*$ vermöge*

$$\Phi_{(,)}(v)(w) = (v, w).$$

$(,) \mapsto \Phi_{(,)}$ ist eine Bijektion zwischen allen Bilinearformen und allen linearen Abbildungen $\Phi: V \to W^*$. Die Linearität im ersten Argument bewirkt, dass Φ linear ist, die Linearität im zweiten Argument bewirkt, dass $\text{bild}(\Phi) \subseteq W^*$. □

Satz 4.30. *Sei W endlichdimensional. V, W bilden ein duales Paar mit Zeuge $(,)$ gdw $\Phi_{(,)}$ ein Isomorphismus ist.*

Beweis: Sei $n = \dim(V) = \dim(W) = \text{rang}((,))$. Seien $\vec{B}$ und $\vec{C}$ angeordnete Basen, B eine $(,)$ beschreibende Matrix bzgl. der angegebenen angeordneten Basen. Sei v in den Koordinaten $(\xi_1, \ldots, \xi_n)$ als Spalte bzgl. $\vec{B}$ gegeben. Dann ist $\Phi(v) \in W^*$ gegeben durch die Matrix $x^\top B^\top$. Die n-n-Matrix B hat vollen Rang, gdw $\{x^\top B^\top : v = \sum_i \xi_i b_i \in V\} = W^*$. Wir haben also gezeigt: B hat vollen Rang, gdw. Φ surjektiv ist. Eine lineare Abbildungen zwischen Räumen gleicher Dimension ist surjektiv gdw sie ein Isomorphismus ist.

Sei nun umgekehrt $\Phi: V \to W^*$ ein Isomorphismus. Sei $\vec{B}$ eine Basis von V, $\vec{C}$ eine Basis von W. Dann ist $\dim(V) = \dim(W^*) = \dim(W) = n$. Seien $(\xi_1, \ldots, \xi_n)$ die Koordinaten von v bzgl $\vec{B}$. Sei $w \in W$ durch seine Koordinaten $y = (\varepsilon_1, \ldots, \varepsilon_n)$ als Spalte bzgl. $\vec{C}$ gegeben. Da $\Phi(v) = (y \mapsto x^\top B y) = 0_{(K^n)^*}$ genau für $v = 0$ (also $x = 0$), hat B Rang n. □

Korollar 4.31. *Seien V, W endlichdimensional. $(,): V \times W \to K$ sei eine Bilinearform. Dann sind äquivalent:*

(1) V, W bilden ein duales Paar mit bezeugender Form $(,)$.

(2) $\forall v(\forall w(v, w) = 0 \to v = 0)$ und $\forall w(\forall v(v, w) = 0 \to w = 0)$. Man sagt hierzu: $(,)$ ist nicht ausgeartet.

Beweis: Die erste Klausel von (2) sagt: Φ ist injektiv, die zweite sagt, Φ ist surjektiv. □

Definition 4.32. Seien nun V, W ein duales Paar mit Zeugen $(,)$. Wir nennen zwei angeordnete Basen $\vec{B}$ und $\vec{C}$ dual wenn

$$(b_i, c_j) = \delta_{i,j}.$$

Wenn $\vec{C}$ vorgegeben ist und $\lambda_i = c_i^*$, dann gewinnt man $\vec{B}$ durch $b_i = \Phi^{-1}(\lambda_i)$. Ein duales Basenpaar gibt es genau dann, wenn V und W ein duales Paar bilden.

Definition 4.33. Sei V, W, $(,)$ ein duales Paar und $\Phi = \Phi_{(,)}$. Sei nun $f \in \mathrm{End}(W)$. Dann überträgt sich f zu einem Endomorphismus f^t von V durch $f^t = \Phi^{-1} \circ f^* \circ \Phi$. Man nennt f^t den zu f adjungierten Endomorphismus.

Wie in Satz 4.20 wird f^t bzgl. der dualen Basis durch die transponierte Matrix dargestellt.

Lemma 4.34. *f^t ist bestimmt durch*

$$(f^t(v), w) = (v, f(w)).$$

Beweis: $(f^t(v), w) = (\Phi^{-1} f^* \Phi(v), w) = (f^* \Phi(v))(w) = \Phi(v)(f(w)) = (v, f(w))$. □

Wenn (V, W) ein duales Paar ist, ist auch (W, V) ein duales Paar, und daher kann man auch für $f \in \mathrm{End}(V)$ einen dualen Endomorphismus $f^t \in \mathrm{End}(W)$ definieren, mit dem vertauschten $(,)$ und einem entsprechenden Ψ. Dann erhält man $f^{tt} = f$. Denn $(f^{tt}(v), w) = (v, f^t(w)) = (f(v), w)$.

Definition 4.35. Für ein duales Paar (V, W) mit Zeugen $(,)$ und einen Unterraum U von V definieren wir den zu U (bzgl. $(,)$) orthogonalen Unterraum von W

$$U^\perp := \{w \in W \,:\, \forall u \in U (u, w) = 0\}.$$

Lemma 4.36. *Sei V, W mit $(,)$ ein duales Paar.*

(1) $(U^\perp)^\perp = U$.

(2) Mit der Bilinearform $[v + U, w] = (v, w)$ wird V/U, $U^\perp$ zu einem dualen Paar. Es gilt $\dim(U) = \dim(U^\perp) = \dim(V)$.

Beweis: (1) Sei $x \in U$. Dann ist $x \in (U^\perp)^\perp$. Wenn $x \notin U$, dann gibt es eine Linearform λ in V^*, die auf U verschwindet und auf x den Wert 1 annimmt. Dann gibt es auch ein $w \in W$ mit $(U, w) = 0$ und $(x, w) = 1$. $w \in U^\perp$, und daher $x \notin (U^\perp)^\perp$.

(2) $[,]$ ist wohldefiniert, denn wenn v und v' zur selben Nebenklasse von U gehören, ist für alle $w \in U^\perp$, $(v', w) = (v, w) + (v - v', w) = (v, w)$. Wir zeigen, dass $[,]$ nicht ausgeartet ist. Wenn $(v' + U, w) = 0$ für alle $v' \in V$, ist $w = 0$. Wenn $(v + U, w') = 0$ für alle $w' \in U^\perp$, gehört v zu $(U^\perp)^\perp = U$. Also in $v + U = 0$. □

Kapitel 5

Symmetrische Matrizen

Definition 5.1. Eine Abbildung $\varphi: K^n \to K$ heißt quadratische Form, wenn

$$\varphi(x) = \sum_{i,j=1}^{n} \alpha_{i,j}\xi_i\xi_j.$$

Wir nehmen nun an, dass die Charakteristik von K nicht 2 ist, d.h., dass $1_k +_K 1_K \neq 0_k$. Dann kann man $\sigma_{i,j} = \frac{\alpha_{i,j}+\alpha_{j,i}}{2}$ setzen und $\varphi(x) = \sum_{i,j=1}^{n} \sigma_{i,j}\xi_i\xi_j = x^\top Sx$ mit einer symmetrischen Matrix S schreiben.

Definition 5.2. Eine Matrix $S \in M_{n,n}(K)$ heißt symmetrisch, wenn $\sigma_{i,j} = \sigma_{j,i}$.

Wenn die quadratische Form $\varphi(x)$ in Koordinaten $x = (\xi_1, \ldots, \xi_n)^\top$ bzgl. der angeordneten Basis $\vec{B}$ auf V durch $\varphi(x) = x^\top Sx$ beschrieben wird, $\vec{B} = \vec{B}'W^\top$ mit einer regulären Matrix W, so wird φ in Koordinaten y bzgl. der angeordneten Basis $\vec{B}'$ durch $\varphi(y) = y^\top W^\top SWy$ beschrieben. Dies beweist man wie Satz 2.48 oder Lemma 4.17. Wir suchen nun Basiswechsel W, so dass $W^\top SW = S'$ besonders einfach aussieht.

Lemma 5.3. *Die symmetrischen n-n-Matrizen über K bilden einen Vektorraum* $\mathrm{Sym}(n; K)$ *der Dimension* $n(n+1)/2$.

Lemma 5.4. *Sei $W \in M_{n,n}(K)$. Dann ist $S \mapsto W^\top SW$ ein Endomorphismus von* $\mathrm{Sym}(n; k)$, *der bijektiv ist, gdw W invertierbar ist.*

Definition 5.5. Sei $S \in \mathrm{Sym}(n, K)$, $r \in \{1, \ldots, n\}$ Unter dem r-ten Hauptminor $\delta_r(S)$ von S versteht man die Determinante der Untermatrix von S, die durch Streichung der letzten $n - r$ Zeilen und $n - r$ Spalten aus S entsteht.

Definition 5.6. $A \in M_{n,n}(K)$ heißt unipotent, wenn A die Gestalt

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

hat.

Satz 5.7. *Satz von Jacobi*[1] *Sei S symmetrisch und sind alle Hauptminoren $\delta_r = \delta_r(S)$ von S nicht Null, dann gibt es eine unipotente n-n-Matrix W mit $S = W^\top DW$ und D ist eine Diagonalmatrix mit Diagonalelementen δ_1, δ_2/δ_1, ..., δ_n/δ_{n-1}.*

Beweis:

$$S = \begin{pmatrix} T & w \\ w^\top & \alpha \end{pmatrix}$$

mit $\alpha \in K$, $w \in K^{n-1}$, $T \in \mathrm{Sym}(n-1, K)$. Ist T invertierbar, also der Minor nicht Null, so ist

$$S = A^\top \begin{pmatrix} T & 0 \\ 0 & \omega \end{pmatrix} A, \text{ mit } A = \begin{pmatrix} 1_{M_{n-1,n-1}(K)} & T^{-1}w \\ 0 & 1 \end{pmatrix},$$

und $\omega = \alpha - w^\top T^{-1} w$. Nun macht man per Induktion weiter mit der Bearbeitung von T. Man verifiziert, dass $\omega = \det(S)/\det(T)$ und dass das Produkt zweier unipotenter Matrizen und das Inverse einer unipotenten Matrix wieder unipotent sind. □

Satz 5.8. *Normalformen für symmetrische Matrizen. Zu jedem $S \in \mathrm{Sym}(n, K)$ ein $W \in \mathrm{GL}(n, K)$ und eine Diagonalmatrix S mit $S = W^\top DW$.*

Beweis: Induktion über n. Wir nehmen $S \neq 0_{M_{n,n}(K)}$ an, denn sonst ist nicht zu beweisen. Dann gibt es ein $v \in K^n$, so dass $v^\top Sv \neq 0$. Sonst wäre für alle x, $y \in K^n$, $2x^\top Sy = (x+y)^\top S(x+y) - x^\top Sx - y^\top Sy = 0$ und daher $e_i^\top Se_j = \sigma_{j,i} = 0$ für alle i, j.

Wir nehmen an, dass $v = v_n$ ist und $\vec{B} = (v_1, \dots, v_n)$ eine angeordnete Basis von K^n ist. Andernfalls geht man zu einer anderen Basis $\vec{B}'$ über mit einer (nicht notwendig unipotenten) Übergangsmatrix W, und S geht über in $S_1 = W^\top SW$. Nun haben wir

$$S_1 = W^\top SW = \begin{pmatrix} * & * \\ * & \alpha \end{pmatrix}.$$

[1] Carl Gustav Jacob Jacobi, 1804 – 1851

Nach Wahl von v_n ist $\alpha = v_n^\top S v_n \neq 0$. Nun gibt es eine reguläre Matrix U wie im vorigen Beweis (mit vertauschten Rollen von α und von T von dort), so dass

$$U^\top S_1 U = \begin{pmatrix} T & 0 \\ 0 & \alpha \end{pmatrix}.$$

Nach Induktionsvoraussetzung gibt es W_1, so dass $D = W_1^\top T W_1$ diagonal ist. Insgesamt haben wie also

$$\begin{pmatrix} W_1^\top & 0 \\ 0 & 1 \end{pmatrix} U^\top W^\top S W U \begin{pmatrix} W_1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} D & 0 \\ 0 & \alpha \end{pmatrix}.$$

□

Satz 5.9. *Der Trägheitssatz von Sylvester[2] Zu jeder reellen symmetrischen n-n-Matrix S gibt es eindeutig bestimmte Zahlen p und $q \in \mathbb{N}$ und $W \in \mathrm{GL}(n, K)$, so dass*

$$S = W^\top \begin{pmatrix} 1_{M_{p,p}(K)} & 0 & 0 \\ 0 & -1_{M_{q,q}(K)} & 0 \\ 0 & 0 & 0 \end{pmatrix} W.$$

p wird der positive Trägheitsindex, q der negative Trägheitsindex, $p-q$ wird manchmal der Trägheitsindex genannt.

Beweis: Wir nehmen an, dass D_1 und D_2 zwei unterschiedliche diagonale Darstellungen für S sind: $D_i = W_i^\top S W_i$ beginne mit p_i Einsen auf der Diagonalen, und dann q_i minus Einsen, und $p_1 < p_2$. Der Rang von D_i ist $p_i + q_i = r$, unabhängig von i, da beide S darstellen. Sei $\varphi_i(x) = x^\top D_i x$. Nun seien $y = (\eta_1, \ldots, \eta_n)^\top$ die Koordinaten von $v \in K^n$ nach dem Basiswechsel mit W_1, $x = (\xi_1, \ldots, \xi_n)^\top$ nach dem Basiswechsel mit W_2. Sei $U = W_2(W_1^{-1}) \restriction \{1, \ldots, r\}^2$ eine reguläre r-r-Matrix. Dann ist

$$\begin{aligned} \varphi_1(y) = \varphi_2(Uy) = \varphi_2(x) =& \eta_1^2 + \cdots + \eta_{p_1}^2 - \eta_{p_1+1}^2 - \cdots - \eta_r^2 \\ =& \xi_1 + \cdots + \xi_{p_2}^2 - \xi_{p_2+1}^2 - \cdots - \xi_r^2. \end{aligned} \tag{5.1}$$

Nun hat das lineare Gleichungssystem $\eta_1 = \cdots = \eta_{p_1} = 0 = \xi_{p_2+1} = \cdots = \xi_r$, in dem die Werte von $(\eta_1, \ldots, \eta_r)$ gesucht sind, eine nicht triviale Lösung $(\eta_1, \ldots, \eta_r)$, denn sein Rang ist $r - (p_2 - p_1) < r$. Für diese $(\eta_1, \ldots, \eta_r)$ wäre aber nach (5.1) $-\eta_{p_1+1}^2 - \cdots - \eta_r^2 = \xi_1 + \cdots + \xi_{p_2}^2$, also doch $\eta_i = 0$ für alle i. □

[2] James Joseph Sylvester, 1814 – 1897

Kapitel 6

Anhang: Mengen und Existenzbeweise

Dieses Kapitel gehört nicht zum Stoffkanon des ersten Jahres. Es ist für Interessierte geschrieben. In diesem Kapitel nennen wir die Axiome und beweisen auf der Basis der Axiome von Zermelo und Fraenkel, dass es einen vollständigen archimedisch angeordneten Körper $(\mathbb{R}, +, \cdot, 0, 1, <)$ gibt. Dieser ist bis auf Isomorphie eindeutig.

6.1 Mengen, Axiome und das Axiomensystem von Zermelo und Fraenkel

Die meisten Definitionen enthalten den Ausdruck: „Sei X eine Menge." Daher stellen wir uns die Frage:

Was ist eine Menge?

Die folgende Antwort wird von den allermeisten mathematischen Schulen akzeptiert:

Definition 6.1. Jedes Objekt ist eine Menge (in jedem Modell von ZFC), wenn es Element eines Modells des Axiomensystems ZFC ist, das heißt, wenn man seine Existenz aus den Axiomen von Zermelo und Fraenkel ZFC (mit den klassischen Beweisregeln) herleiten kann.

Definition 6.2. Jedes Objekt ist eine Menge (in einem Modell von ZFC), wenn es Element eines Modells des Axiomensystems ZFC ist, das heißt, wenn man seine Existenz aus den Axiomen von Zermelo und Fraenkel ZFC (mit den klassischen Beweisregeln) nicht widerlegen kann.

Zu dieser weiter gefassten Definition kann man Mengenlehre und insbesondere Unabhängigkeitsbeweise studieren. Wir fahren nun mit der ersten Definition fort.

Wir betrachten in dieser Vorlesung Mengen, die in *allen* Modellen von ZFC vorkommen, das heißt Mengen, deren Existenz man in ZFC beweisen kann. Wir werden einige solcher Existenzbeweise führen, darunter Existenzbeweise für $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

Vielleicht werden Sie im Laufe Ihres Studiums Beispiele für Mengen kennenlernen, die es nur in manchen ZFC-Modellen gibt. Wenn Sie sich für solche Mengen interessieren, können Sie später einmal eine Vorlesung über Unabhängigkeitsbeweise hören. ZFC hat viele verschiedene Modelle.

Nun folgen erste Erklärungsschritte der Begriffe, die in der ersten Definition vorkommen.

Beweise führen wir nach den klassischen Beweisregeln, die wir üben, aber hier nicht auflisten (Es sind etwa zehn Regeln.). Indirekte Beweise sind gestattet. Eine Aussage ist bewiesen, (genau dann) wenn ihr Negat widerlegt ist.

Ein ZFC-Modell besteht aus einem Universum aller Mengen und einer $\in$-Relation, $x \in y$ wird interpretiert als: „x ist eine Menge und y ist eine Menge und x ist Element von y." In einem ZFC-Modell soll alle ZFC-Axiome gelten.

Wir schauen uns im Duden den Eintrag „Axiom" an:

„1. als absolut richtig erkannter Grundsatz; gültige Wahrheit, die keines Beweises bedarf

2. nicht abgeleitete Aussage eines Wissenschaftsbereichs, aus der andere Aussagen deduziert werden"

Bei den Axiomen ZFC Zermelo[1], Fraenkel[2] und Choice (steht für Axiom of Choice, das Auswahlaxiom), handelt es sich eher um Axiome nach der ersten Definitionsmöglichkeit. Man kann sie nicht beweisen. Die meisten ZFC-Axiome und auch den Aufbau der Sprache der ersten Stufe kann man aus der Alltagserfahrung im Umgang mit endlich langen Zeichenreihen nachvollziehen, sozusagen in gewissem Maß experimentell verifizieren. Hingegen beim Unendlichkeitsaxiom und beim Auswahlaxiom hat man in der Natur kaum Vorbilder und kann außer Gedankenexperimenten keine Experimente machen. Das Wort „. . . keines Beweises bedarf" im Duden-Eintrag ist daher für unser Anliegen, die Mathematik auf sicheren Grundlagen möglichst rein deduktiv aufzubauen, eher eine beschönigende Notlösung. Mathematik heißt übrigens auf Deutsch: Kunst des Lernens. Wir schauen uns nun die Axiome, die ab 1930 etwa als Axiome der gesamten Mathematik gelten, an:

[1] Ernst Zermelo, 1871 – 1953
[2] Abraham Halevi Fraenkel, 1891 – 1965

ZFC: Die Zermelo–Fraenkel–Axiome mit Auswahlaxiom

Abgeschrieben und zusammengestellt aus [12].

(0) Existenzaxiom.
Es gibt eine Menge x. Diese Axiom gehört streng genommen nicht zu ZFC sondern zu den logischen Axiomen. Da wir die (sogenannten klassischen) Beweisregeln nicht auflisten werden, haben wir dieses Axiom hier aufgeschrieben.

(1) Extensionalitätsaxiom.
Je zwei Mengen, die dieselben Elemente enthalten, sind gleich. Beispiel: $\{1,1,2\} = \{1,2\} = \{2,1\}$.

(2) Paarmengenaxiom.
Zu je zwei Mengen x, y gibt es die Paarmenge $\{x,y\}$.

(3) Aussonderungsschema.
Zu jeder Menge x und zu jeder in der mengentheoretischen Sprache der ersten Stufe formulierbaren Eigenschaft φ und zu jeder Parametermenge p gibt es die Menge

$$\{u \in x : \varphi(u,p)\}.$$

Bsp: $\{n \in \mathbb{N} : n \text{ prim}\}$, $\{(n,m) \in \mathbb{N}\times\mathbb{N} : n,m \text{ Primzahlpaar}\}$, $\varnothing = \{u \in x : u \neq u\}$, hierbei sei x aus dem Existenzaxiom genommen.

(4) Vereinigungmengenaxiom.
Zu jeder Menge x gibt es die Vereinigungsmenge von x, das ist die Menge, die die Elemente jeder Menge, die Element von x ist, als Elemente enthält. Wir schreiben $\bigcup x$ dafür.

$$\bigcup x = \{z : \exists y(y \in x \wedge z \in y)\}.$$

Bsp. $\bigcup\{x,y\} = x \cup y$

(5) Potenzmengenaxiom.
Zu jeder Menge gibt es die Potenzmenge, $\mathscr{P}(x) = \{y : y \subseteq x\}$. Wir schreiben $y \subseteq x$ als Abkürzung für $\forall z(z \in y \to z \in x)$.

(6) Unendlichkeitsaxiom.
Es gibt eine unendliche Menge. Formalisiert: Es gibt eine Menge x, so dass x induktiv ist. Letzteres heißt:

$$\varnothing \in x \wedge \forall y(y \in x \to y \cup \{y\} \in x).$$

(Vorsicht: Dieser Gebrauch von "induktiv" hat nichts mit der induktiven Halbordnung zu tun, ist eher gegenteilig!)

(7) Fundierungsaxiom.
Jede nicht leere Menge x hat ein $\in$-minimales Element u, d.h. $u \in x \wedge \forall z \in x \neg z \in u$.
Konsequenzen: $x \neq \{x\}$. Im Mengenuniversum gibt es keine unendlich lang ansteigendenden Ketten $x_0 \ni x_1 \ni x_2 \dots$ und keine $\in$-Schleifen $x_0 \in x_1 \in \dots \in x_n \in x_0$.

(8) Ersetzungsschema.
Sei φ eine erststufige Eigenschaft, die auf x einen funktionalen Zusammenhang beschreibt. Dann gibt es die Bildmenge von x unter dieser Funktion. Formal: $\forall u \in x \exists^{=1} w \varphi(u, w, p) \to \exists y \forall u \in x \exists w \in y \varphi(u, w, p)$

(9) Auswahlaxiom.
Zu jeder Menge nicht leerer Mengen gibt es eine Auswahlfunktion. $\forall y((y \in x \to y \neq \varnothing) \to \exists f \forall y \in x f(y) \in y))$

Axiome (0) bis (8) zusammen heißen ZF. Axiome (0) bis (9) zusammen heissen ZFC. Beweise, bei denen die axiomatischen Grundlagen nicht genannt sind, werden auf der Basis von ZFC geführt. Dies gilt wahrscheinlich so für alle Mathematik, die Sie hier in den ersten Jahren oder jemals hören.

Manche Mathematiker(innen) deklarieren bei Benutzung das Auswahlaxiom, da es wegen seinen inkonstruktiven Charakters vielleicht weniger plausibel aussieht. Sie haben in der Schule sicherlich schon das Auswahlaxiom benutzt in der Analysis, womöglich, ohne es bemerkt zu haben. Vom Anliegen der Widerspruchsfreiheit her ist das Auswahlaxiom hingegen eine ungefährliche Zugabe:

Satz 6.3. *(1) Gödel[3] 1931 (Drei Anwendungen des zweite Gödelsche Unvollständigkeitssatzes). Wenn* ZF *widerspruchsfrei ist, dann beweist* ZF *nicht, dass es widerspruchsfrei ist. Wenn* ZFC *widerspruchsfrei ist, dann beweist* ZFC *nicht, dass es widerspruchsfrei ist. Die Zahlentheorie beweist nicht, dass sie widerspruchsfrei ist.*

(2) Gödel 1938. Wenn ZF *widerspruchsfrei ist, so auch* ZFC*.*

Beweise zum ersten dieser Sätze können Sie in der Vorlesung „Mathematische Logik" hören. Der zweite Satz wird manchmal in einer Vorlesung über axiomatische Mengenlehre bewiesen. Wir werden ZFC zu den Sätzen hinzuschreiben, die nicht in ZF alleine geführt werden können.

„Speziellere" Axiomensysteme: Gruppenaxiome, vollständige archimedisch angeordnete Körper, Ringaxiome, Ordnungsaxiome, Axiome für Boole'sche Algebren und viele mehr.

Bei vielen spezielleren Axiomensystemen beweist ZFC jeweils, dass das speziellere Axiomensystem widerspruchsfrei ist. Sie werden im Laufe Ihres Studiums eine Fülle solcher Axiomensysteme sehen.

[3] Kurt Gödel, 1906–1976

Nun schauen wir uns noch an, was im Aussonderungsschema und im Ersetzungsschema mit „erststufiger mengentheoretischer Eigenschaft" gemeint ist:

Atomare Eigenschaften sind: $x = y$, $x \in y$ (und nur diese). Wenn φ und ψ Eigenschaften sind und x eine Variable ist, so sind $(\varphi \wedge \psi)$, $(\varphi \to \psi)$, $(\varphi \vee \psi)$, $(\varphi \leftrightarrow \psi)$, $\neg\varphi$, $\exists x\varphi$, $\forall x\varphi$ Eigenschaften. Jede Eigenschaft lässt sich in endlich vielen Schritten aus den atomaren Eigenschaften aufbauen. Die Junktoren und die Quantoren tragen hierbei ihre übliche Bedeutung $\exists$ steht für „es gibt eine Menge" und $\forall$ steht für „für alle Mengen". Unsere Klammernkonventionen gestatten eindeutige Lesbarkeit. $\forall x \in y\varphi$ steht als Abkürzung für $\forall x(x \in y \to \varphi)$ und $\exists x \in y\varphi$ steht als Abkürzung für $\exists x(x \in y \wedge \varphi)$. „Es gibt genau ein x mit φ" schreibt man als $\exists^{=1}x\varphi$. Dies steht als Abkürzung für $\exists x(\varphi(x) \wedge \forall y(\varphi(y) \to x = y))$. Definierte Terme und Eigenschaften können, sozusagen als Abkürzungen, als Formeln in den Schemata und im Aufbau neuer Begriffe verwendet werden: Beispiele: $\varnothing$ (mit seinem Existenzbeweis von oben), $\{x\}$, $\bigcup x$, $\mathscr{P}(x)$, $x \subseteq y$ für $\forall z(z \in x \to z = y)$, und fast alles, was Sie im Lauf der Vorlesungen sehen werden, kann zum Aufbau weiterer Mengen verwendet werden.

Wir brauchen in den Schemata diese klare Festlegung der sprachlichen Möglichkeiten, um Paradoxa der Art

Die kleinste Zahl, die nicht mir vierzig Buchstaben definiert werden kann.

in den Definitionen in der Mathematik zu verbieten.

Definition 6.4. Wir nennen beliebige Zusammenfassungen von Mengen Klassen. Die Allklasse ist die Klasse aller Mengen.

Satz 6.5. *Die Allklasse ist keine Menge.*

Beweis: Annahme: Es gäbe die Menge x aller Mengen. Dann bilden wir die Russell'sche Menge [4]

$$y = y_{\text{Russell}} = \{u \in x : u \notin u\}.$$

Wir haben $y \in y$ impliziert $y \notin y$, und $y \notin y$ impliziert $y \in y$. □

Wir kennzeichnen das Ende eines Beweises mit einem Kästchen.

6.2 Die natürlichen Zahlen

Definition 6.6. Die Terme der Art $\underline{0} = \varnothing$,
$\underline{n+1} = \underline{n} \cup \{\underline{n}\}$, heißen die von Neumann'schen[5] natürlichen Zahlen.

[4] Bertrand Russell, 1872 – 1970

[5] John von Neumann, 1903 – 1957

Definition 6.7.

$$N = \{\underline{0}, \underline{1}, \dots\}.$$

heißt die von Neumann'sche Menge der natürlichen Zahlen.

Gibt es N?

Satz 6.8. *N ist eine Menge.*

Beweis: Wir nehmen ein x, wie es im Unendlichkeitsaxiom gegeben ist. Dann schreiben wir

$$N = \{z \in x : \forall y((y \subseteq x \wedge y \text{ induktiv}) \to z \in y)\}. \qquad (*)$$

Dies ist die offizielle Definition von N in ZFC. □

Die natürlichen Zahlen $(\mathbb{N}, 0, +1)$ sind nun irgendeine isomorphe Kopie der von Neumann'schen Menge mit ihrer Struktur. Die von Neumann'sche Struktur selbst ist gut, aber man darf sich auch etwas anderes aussuchen.

Bemerkung 6.9. x ist in y enthalten, ist im Deutschen zweideutig: Es kann sowohl $x \in y$ and auch $x \subseteq y$ gemeint sein.

Definition und Behauptung 6.10. Seien x, y Mengen. Dann gibt es folgende Mengen

(1) $x \cap y = \{z \in x : z \in y\}$ heißt die Schnittmenge von x und y.

(2) Sei $y \neq \varnothing$. $\bigcap y = \{z \in \bigcup y : \forall x \in y z \in x\}$ heißt der Schnitt über x. Als Randfall kann man definieren $\bigcap \varnothing$ = Allklasse.

(3) $x \cup y = \bigcup\{x, y\}$ heißt die Vereinigung von x und y.

(4) $x \setminus y = \{z \in x : z \notin y\}$ heißt x ohne y.

(5) $x \times y = \{(u, v) : u \in x, v \in y\}$ heißt das kartesische[6] Produkt von x und y.

Bemerkung: Für den Beweis der Existenz des kartesischen Produkts wird man zunächst geordnete Tupel (u, v) als geeigenete Mengen modellieren, z.B. als $\{u, \{u, v\}\}$. Man braucht ja bei Tupeln immer nur folgende Eigenschaft

$$(x, y) = (u, z) \to (x = u \wedge y = z).$$

Dann zeigt man die Existenz des kartesischen Produkts am einfachsten mit dem Paarmengenaxiom, dem Vereinigungsmengenaxiom, dem Potenzmengenaxiom und dem Aussonderungsaxiom. Sportliche Leute schaffen es auch ohne das Potenzmengenaxiom, nehmen dafür das Ersetzungsschema mit geeigneten einfachen funktionalen Zuordnungen. Wenn Sie die

[6] René Descartes, 1596 – 1650

Sache sehr interessiert, schauen Sie in einem Mengenlehrebuch nach. Man schaut am besten zuerst unter Kuratowski[7] pairs oder ordered pairs.

Relationen sind Teilmengen von kartesischen Produkten von Mengen, zum Beispiel $<$ auf $\mathbb{N}$: Diese Relation kann man mit $\{(m,n) \in \mathbb{N}\times\mathbb{N} : m < n\}$ identifizieren. Für klassengroße Pendants zu Relationen gibt es meines Wissens kein eigenes Wort. Beispiel: Die $\in$-Relation.

Definition 6.11. (und Bemerkungen) Seien x, y Mengen. Sei $f: x \to y$ eine Funktion, das heißt, $f \subseteq x \times y$ und für jedes $u \in x$ gibt es genau ein $v \in y$, so dass $(u,v) \in f$. Man schreibt für letzteres eher $f(u) = v$. Die Menge x heißt der Definitionsbereich von f, die Menge y heißt Zielbereich. Man kann eine Funktion f mit ihren Graphen

$$\{(u,v) \in x \times y \,:\, f(u) = v\}$$

identifizieren, verliert dabei jedoch die Information über den Zielbereich y. Er wird ersetzt durch den minimalen Zielbereich

$$\text{bild}(f) = f[x] = \{f(u) \,:\, u \in x\} \subseteq y$$

oder irgendeine Obermenge der Bildmenge. $\text{bild}(f)$ heißt die Bildmenge von f. Zu einer klassen-großen eindeutigen Zuordnung sagt man auch Operation oder Funktional.

Bem.: Wir schreiben nicht $f(x)$ für $f[x]$ (obwohl dieses in manchen Gebieten Usus ist), da wir uns der Möglichkeit, dass $f(x) \neq f[x]$, nicht begeben wollen. Denken Sie nur an die von Neumann'schen natürlichen Zahlen. Es gibt nur eine Funktion auf N, die $f(x) = f[x]$ für alle $x \in N$ erfüllt, nämlich die Identität. (Beweis: Übung). Natürlich wollen wir viele verschiedene Funktionen betrachten.

Wichtige Eigenschaften von Funktionen sind:

Definition 6.12. Sei $f: X \to Y$ eine Funktion, und sei $g: Y \to Z$ eine Funktion.

(1) $f: X \to Y$ heißt injektiv, wenn für je zwei $x, y \in X$, aus $f(x) = f(y)$ immer $x = y$ folgt.

(2) $f: X \to Y$ heißt surjektiv, wenn $\text{bild}(f) = Y$.

(3) f heißt bijektiv, wenn f injektiv und surjektiv ist.

(4) $g \circ f$ ist die Funktion zuerst f, dann g. Also für $x \in X$, $(g \circ f)(x) := g(f(x))$. $g \circ f: X \to Z$.

Nun wollen wir unter den Funktionen solche beschreiben, die gewisse Merkmale treu kopieren:

[7] Casimir Kuratowski, 1896–1980

Definition 6.13. (1) Eine Struktur ist eine nicht leere Menge M zusammen mit Funktionen auf M^n für ein n, Konstanten aus der Trägermenge und Relationen, das sind Teilmengen von M^n. Zum Beispiel $(\mathbb{N}, +, 0, 1, \cdot, <)$. Eine Struktur kann auch von einer oder mehreren Sorten keine Bestandteile haben, zum Beispiel gehört zu einer Gruppe zunächst einmal keine Relation.

(2) Zwei Strukturen heißen isomorph, wenn es eine Isomorphismus zwischen ihnen gibt. Ein Isomorphismus ist eine Bijektion von der einen Trägermenge auf die andere, die alle Strukturmerkmale in beide Richtungen treu abbildet.

(3) Eine Einbettung ist eine Injektion, die alle Strukturmerkmale treu abbildet und die Relationen in beide Richtungen treu erhält.

(4) Ein Homomorphismus $h: (A, f_A, R_A) \to (B, f_B, R_B)$ ist eine Funktion, die mit den Funktionen auf den jeweiligen Strukturen kommutiert, und die die Relationen in die Vorwärtsrichtung erhält: Zum Beispiel soll eine dreistellige Funktion homomorph abgebildet werden: $h(f_A(x, y, z)) = f_B(h(x), h(y), h(z))$ Für eine Relation soll gelten $xR_Ay \to f(x)R_Bf(y)$. Beachten Sie, dass hier wirklich nur die Implikation steht.

Nun kehren wir zu unserem Anliegen, weitere Eigenschaften und Beschreibungen von N und $\mathbb{N}$ zu finden, zurück:

In unserem Fall gilt also für jede Möglichkeit von $\mathbb{N}$: Es gibt eine bijektive Abbildung $f: N \to \mathbb{N}$, nämlich $f(\underline{n}) = n$. f ist strukturerhaltend, denn $f(\underline{0}) = 0$ und $f(\underline{n+1}) = n + 1 = f(\underline{n}) + 1$.

Nun kommen wir zu einem viel kleineren und schwächeren Axiomensystem als ZFC: den Peano-Axiomen.

Definition 6.14. Sei X eine Menge, $0_X \in X$ und $\sigma_X: X \to X$ eine Funktion. Wir sagen $(X, 0_X, \sigma_X)$ erfüllt die Peano-Axiome[8], wenn folgendes gilt

(1) σ_X ist eine Bijektion von X auf $X \setminus \{0_X\}$.

(2) Für alle $M \subseteq X$ gilt: $\big(0_X \in M \wedge \forall y(y \in M \to \sigma_X(y) \in M)\big) \to M = X$.

Satz 6.15. *Die von Neumann'schen natürlichen Zahlen N mit der leeren Menge und der Nachfolgerfunktion $x \mapsto x \cup \{x\}$ erfüllen die Peano-Axiome.*

Beweis: Wir zeigen das zweite Axiom. Annahme nicht. Sei M ein Gegenbeispiel. Dann ist M induktiv und eine echte Teilmenge von N, im Widerspruch zu $(*)$. □

Bemerkung 6.16. Das zweite Peano-Axiom ist gerade das Prinzip der vollständigen Induktion. d.h., für jedes Modell $(X, 0_X, \sigma_X)$ und jede (definierbare oder nicht definierbare)

[8] Giuseppe Peano 1858–1932

Eigenschaft gilt: Wenn 0_X die Eigenschaft hat und wenn sich die Eigenschaft auf Nachfolger vererbt, dann haben alle Elemente aus X die Eigenschaft.

Korollar 6.17. *Für die natürlichen Zahlen* $(\mathbb{N}, 0, +1)$ *gilt das Prinzip der vollständigen Induktion.*

Beweis: $(N, \varnothing,$ Nachfolgerfunktion) ist isomorph to $(\mathbb{N}, 0, +1)$. Isomorphe Strukturen erfüllen dieselben Wahrheiten. Nun folgt das Korollar aus Satz 6.15. □

Wir haben also in den Peano-Axiomen ein Beispiel für ein Axiomensystem, für das man in ZFC ein Modell bauen kann und für das man somit in ZFC einen Widerspruchsfreiheitsbeweis hat.

Frage 6.18. *Gibt es Strukturen, die nicht isomorph zu* $(\mathbb{N}, 0, +1)$ *sind und trotzdem vollständige Induktion gestatten?*

Ja, die gibt es. Es gibt Induktionen über längere lineare Ordnungen gewisser Art, sogenannte Wohlordnungen. Wir werden im Lauf dieses Semesters ein Beispiel kennenlernen. Der allgemeinste Induktionstyp ist die Induktion über fundierte mengenähnliche Relationen. Dieses können Sie bei Interesse später einmal studieren.

Nun fahren wir weiter mit den natürlichen Zahlen fort:

Satz 6.19. *Satz von der Definition durch Rekursion über* $\mathbb{N}$. *Sei* φ *eine Formel oder Rechenvorschrift, die* 0 *eine Menge* $\varphi(0)$ *zuordnet und für jedes* $n \in \mathbb{N}$ *der Ausgangsmenge* $\{(0, \varphi(0)), \ldots, (n, \varphi(n))\}$ *eine Fortsetzung* $\{(0, \varphi(0)), \ldots, (n, \varphi(n)),$ $(n+1, \varphi(n+1))\}$ *zuordnet. Dann gibt es genau eine Funktion* $f: \mathbb{N} \to V$, *die der rekursiven Rechenvorschrift genügt und* $f(n) = \varphi(n)$ *erfüllt.*

Wir erinnern hier an Mathematikgeschichte: Zermelo hat in seiner Arbeit an den Axiomen (etwa von 1900 bis 1908) übersehen, dass man das Ersetzungsschema braucht. Dies fand Fraenkel in den 1920er Jahren (und unabhängig von ihm noch Skolem[9] und Mirimanoff[10]. Zur Veranschaulichung schildern wir ohne Beweise ein Beispiel: Ohne Ersetzungsaxiom kann man das Potenzmengenaxiom nur endlich oft iterieren, und erreicht so nur Mengen gewisser Schachtelungstiefe. Wenn man nun für jedes $n \in \mathbb{N}$ die Potenzmengenbildung (von einer festen Startmenge ausgehend) n Mal iteriert hat, dann möchte man gerne alle n-fachen Gebilde vereinigen. Aber sie sind nicht unbedingt Elemente einer gemeinsamen Menge, es sei denn, man hat das Ersetzungsaxiom.

(Beweisskizze, für später gedacht, wenn Sie schon etwas Erfahrung gesammelt haben.) Im einfacheren Fall, dass man schon eine Zielmenge y kennt, in der alle rekursiv definierten

[9] Thoralf Albert Skolem, 1887 – 1963
[10] Dmitry Semionovitch Mirimanoff, 1861 – 1945

$\varphi(n)$ liegen, ist der Graph von f folgende Menge: $\{(n,u) \in \mathbb{N}\times y : \forall g\Big(\big(g(0) = \varphi(0) \wedge \forall m(g \restriction (m+1) = \{(0,\varphi(0)),\dots,(m,\varphi(n))\} \to g(m+1) = \varphi(m+1))\big) \to g(n) = u\Big)\}$. In diesem Fall braucht man das Ersetzungsaxiom nicht.

Im allgemeinen Fall wendet man das Ersetzungsschema auf $\mathbb{N}$ und ψ an, wobei $\psi(n,w)$ sagt: $w = \{(0,\varphi(0)),\dots,(n,\varphi(n))\}$. Danach ist die Vereinigungsmenge der vom Schema gelieferten Bildmenge eine Obermenge der gesuchten rekursiv definierte Funktion. Aus dieser Obermenge kann man wieder wie oben aussondern. □

Wir schauen uns Beispiele für rekursiv definierte Funktionen an:

Proposition 6.20. *Je zwei Modelle der Peano-Axiome sind isomorph. Wir sagen dazu auch: Die Peano-Axiome haben modulo Isomorphie genau ein Modell. Die Peano-Axiome sind kategorisch.*

Beweis: Seien $(\mathbb{N},0,+1)$ und $(Y,0_Y,\sigma_Y)$ zwei Modelle. Wir setzen $f(0) = 0_Y$. $f(n+1) = \sigma_Y(f(n))$. f ist injektiv, da für $m \neq n$, $f(n) \neq f(n)$, wie induktiv über n folgt: Für $n = 0$ ist nichts zu zeigen. Wir zeigen den Schritt von n auf $n+1$: Annahme $f(n+1) = f(m)$ für ein $m \neq m+1$. Dann ist $m > 0$, da $f(0) = 0_Y \neq \sigma_Y(f(n))$. Dann ist $\sigma_Y(f(n)) = \sigma_Y(f(n-1))$. Da σ_Y nach dem ersten Peano-Axiom injektiv ist, folgt $f(n) = f(m-1)$, im Widerspruch zur Induktionsvoraussetzung. Die Funktion f ist wohldefiniert und auf ganz $\mathbb{N}$ definiert nach dem Rekursionsprinzip. f ist surjektiv, da auch $\mathrm{bild}(f) \subseteq Y$ induktiv ist. f ist ein Isomorphismus von $(\mathbb{N},0,+1)$ auf $(Y,0_Y,\sigma_Y)$. □

Definition 6.21. Die rekursive Definition der Addition, der Multiplikation und der Exponentiation auf $\mathbb{N}$. Sei $m \in \mathbb{N}$. m wird als sogenannter Parameter festgehalten. Die Rekursion läuft über n: Wir definieren

$m + 0 = m$,
$m + (n+1) = (m+n) + 1$.
$m \cdot 0 = 0$,
$m \cdot (n+1) = m \cdot n + m$.
$m^0 = 1$,
$m^{n+1} = m^n \cdot m$.

Definition 6.22. Eine Menge M heißt endlich, wenn es eine natürliche Zahl n und eine Bijektion f gibt, so dass

$$f\colon \{0,1,\dots,n-1\} \to M,$$

Falls $n = 0$, steht auf der linken Seite gerade die leere Menge. n heißt die Mächtigkeit von M, man schreibt $|M| = n$ oder auch $\#(M) = n$.

Definition 6.23. M und X heißen gleichmächtig, wenn es eine Bijektion $f\colon M \to X$ gibt.

Die Gleichmächtigkeit ist eine Äquivalenzrelation (siehe Def. 6.29) auf dem Mengenuniversum mit klassen-vielen klassen-großen Äquivalenzklassen. In einer späteren Vorlesung werden Sie vielleicht lernen: Ein Repräsentantensystem für die Gleichmächtigkeit sind die Kardinalzahlen. Die unendlichen Kardinalzahlen werden auch die Alephs oder ℵ's genannt nach dem ersten Buchstaben des hebräischen Alphabets.

Übung: Sind Teilmengen endlicher Mengen endlich? Gibt es eine „größte" endliche Menge?

Definition 6.24. (Schwache Definition, Dedekind-unendlich[11]) Eine Menge, die nicht endlich ist, heißt unendlich.

Sie denken vielleicht an alternative, „positivere" Definitionen. Die gibt es in der Tat:

Satz 6.25. ZFC *Die folgenden Aussagen sind äquivalent:*

(1) M ist nicht endlich.

(2) Es gibt eine Injektion $f\colon \mathbb{N} \to M$.

(3) Es gibt eine Surjektion $f\colon M \to \mathbb{N}$.

Beweis: (1) impliziert (2). Sei h eine Auswahlfunktion auf $\mathscr{P}(M) \setminus \{\varnothing\}$. (Man kann auch beweisen, dass zu diesem Beweis wirklich ein Teil des Auswahlaxioms gebraucht wird.) Wir definieren rekursiv: Für $n \in \mathbb{N}$: $f(n) = h(M \setminus \{f(i) : i < n\})$. f ist injektiv, da für $n > m$, $f(n) \in M \setminus \{f(i) : i < n\}$, und insbesondere $f(m)$ in der Menge vorkommt, die abgezogen wird. Wir haben, dass für jedes n die Menge $M \setminus \{f(i) : i < n\} \neq \varnothing$ ist, denn andernfalls gäbe es ein n, so dass $f\colon \{0, \dots, n-1\} \to M$ bijektiv ist. Nach dem Satz über die rekursive Definition $f\colon \mathbb{N} \to M$ wohldefiniert.

(2) impliziert (3): Sei $f\colon \mathbb{N} \to M$ injektiv. Wir definieren $g\colon \mathrm{bild(f)} \to \mathbb{N}$ durch $g(f(n)) = n$. Für $y \in M \setminus \mathrm{bild}(f)$ setzen wir $g(y) = 0$. Dann ist $g\colon M \to \mathbb{N}$ surjektiv. Dies gilt allgemein:

Jede Umkehrfunktion ist surjektiv.

(3) $\Rightarrow$ (1): Annahme M ist endlich, und g sei ist eine Bijektion $g\colon \{0, 1, \dots, n-1\} \to M$. Dann ist $f \circ g\colon \{0, \dots, n-1\} \to \mathbb{N}$ surjektiv. Jedoch ist $\max\{(f \circ g)(i) : i < n\} + 1 \notin \mathrm{bild}(f \circ g)$. Widerspruch. □

Auf den natürlichen Zahlen kann man die Addition die Multiplikation und die Exponentiation statt durch obige Rekursion auch (äquivalent) durch das Ausrechnen von Mächtigkeiten endlicher Mengen definieren:

Proposition 6.26. *Seien M und N endliche Mengen. $|m| = m$.*
$|M \cup N| = |M| + |N|$, falls $M \cap N = \varnothing$.

[11] Richard Dedekind, 1831–1916

$|M \times N| = |M| \cdot |N|$.
$|M|^{|N|} = |\{f \,:\, f\colon N \to M\}|$.

Beweis: Induktiv über $|N|$.

Proposition 6.27. *Für $(\mathbb{N}, +, \cdot, 0, 1)$ gelten folgende Rechengesetze:*

(1) Assoziativgesetze
Für ℓ, m, $n \in \mathbb{N}$ gelten
$(l + m) + n = (l + m) + n$ *und*
$(lm)n = l(mn)$.

(2) Neutrale Elemente. Für $n \in \mathbb{N}$ ist
$n + 0 = 0$
$n \cdot 1 = n$

(3) Kommutativgesetze
$m + n = n + m$,
$m \cdot n = n \cdot m$

(4) Distributivgesetz
$l(m + n) = lm + ln$

(5) Kürzungsregeln
$l + m = l + n \to m = n$
$l \cdot n = l \cdot m \to l = 0 \vee n = m$

Beweis: Man zeigt beim Kommutativgesetz erst durch Induktion $n + 1 = 1 + n$. Danach steigt man in die rekursive Definition von + ein. Bei den Kürzungsregeln nimmt man $m \leq n$ an und beweist die Aussage $\forall m \leq n(l + m = l + n \to n = m)$ induktiv über n. Diese Technik ist bei Aussagen mit mehreren Variablen sehr nützlich. □

6.3 Ganze, rationale Zahlen und reelle Zahlen

Wir konstruieren nun $\mathbb{Z}$, so dass wir eine isomorphe Kopie von $(\mathbb{N}, 0, +)$ als Teilstruktur (Substruktur, Unterstruktur) von $(\mathbb{Z}, 0, +)$ wiederfinden. Unser Wunsch ist: $(\mathbb{Z}, 0, +)$ soll eine kommutative Gruppe bilden.

Definition 6.28. (a) Eine Struktur $(G, \circ)$ heißt Gruppe, wenn sie die Gruppenaxiome erfüllt. Diese sind:

(G1) das Assoziativgesetz:
Für all $x, y, z \in G$ gilt $(x \circ y) \circ z = x \circ (y \circ z)$.

(G2) das Gesetz vom neutralen Element e:
Es gibt ein e e ist linksneutral, d.h. Für alle $x \in G$ gilt $e \circ x = x$.

(G3) die Existenz vom Inversen:
Es gibt ein linksneutrales Element e so dass
jedes Element ein Linksinverses bezüglich e hat, d.h. zu jedem $x \in G$ gibt es $y \in G$, so dass $y \circ x = e$.

(b) Eine Struktur $(G, \circ)$ heißt abelsche Gruppe, wenn sie die Gruppenaxiome erfüllt und zusätzlich kommutativ (auch abelsch genannt) ist. D.h., das Kommutativgesetz gilt: Für alle $x, y \in G$ ist $x \circ y = y \circ x$.

Das Gesetz von inversen Element soll also gelten:

$$\forall z \in \mathbb{Z}\ \exists y \in \mathbb{Z}\ z + y = 0.$$

Die Idee ist, $\mathbb{Z}$ als geeigneten Teil von $\mathbb{N} \times \mathbb{N}$ wiederzufinden. Hierzu führen wir das Dividieren durch Äquivalenzrelationen ein, das im Jargon auch „Rechnen modulo ..." oder „faktorisieren" genannt wird.

Definition 6.29. Sei M eine Menge.

(1) R heißt (zweistellige) Relation auf M wenn $R \subseteq M \times M$.

(2) R heißt reflexiv wenn für alle $x \in M$, xRx.

(3) R heißt symmetrisch: Für alle $x, y \in M$ gilt $xRy \to yRx$.

(4) R heißt transitiv für alle $x, y, z \in M$ gilt: $(xRy \wedge yRz \to xRz)$.

(5) R heißt Äquivalenzrelation auf M wenn R reflexiv, symmetrisch und transitiv ist.

Definition 6.30. Sei R eine Äquivalenzrelation auf M und $x \in M$.

(1) Die folgende Teilmenge von M

$$[x]_R := \{y \in M\ :\ xRy\}$$

heißt die Äquivalenzklasse von x.

(2) Wir nennen M/R die Quotientenmenge.

$$M/R := \{[x]_R\ :\ x \in M\}$$

Bem.: $M/R \subseteq \mathscr{P}(M)$.

(3) $p\colon M \to M/R$ mit $p(x) = [x]_R$ heißt die Quotientenabbildung.

Frage 6.31. *Hat die Quotientenabbildung eine Umkehrung?*

Wir suchen eine Teilmenge $S \subseteq M$, so dass die Einschränkung von p auf S, geschrieben $p \restriction S$, injektiv ist. Solch eine Teilmenge heißt Repräsentantensystem für M und R. Wir geben eine äquivalente Definition:

Definition 6.32. $S \subseteq M$ heißt Repräsentantensystem für R auf M wenn

$$\forall y \in M/R\ \exists^{=1} x \in S\ x \in y$$

S ist als gerade die Bildmenge der Auswahlfunktion

$$\{(y, S(y)) : y \in M/R\}.$$

Nach dem Auswahlaxiom existiert eine Auswahlfunktion und somit ein Repräsentantensystem. In der Tat gilt:

Satz 6.33. *In* ZF *gilt: Das Auswahlaxiom ist wahr genau dann, wenn jede Äquivalenzrelation ein Repräsenantensystem hat.*

Beweis: Die eine Richtung haben wir gerade eben gezeigt. Habe nun also jede Äquivalenzrelation ein Repräsenantensystem. Sei eine Menge X nicht leerer Mengen gegeben. Wir betrachten nun

$$\hat{X} = \{\{x\} \times x : x \in X\}.$$

Dann ist $\hat{X}$ eine Menge disjunkter nicht leerer Mengen. Wir definieren eine Äquivalenzrelation R auf $\bigcup \hat{X}$ durch

$$(x, u) R (y, v) \text{ wenn } x = y.$$

Sei S ein Repräsentantensystem für R auf $\bigcup \hat{X}$. Dann ist S gerade eine Auswahlfunktion für X. □

Übung 6.34. Lösen Sie nun die Schlumpfaufgabe auf Seite 31 des Ersti-Heftes „Perfekter Körper" vom Oktober 2021. Jetzt können Sie die Schlümpfe sicherlich beraten, für welche Äquivalenzrelation sie sich bei ihrer Strategiebesprechung ein Repräsentantensystem aussuchen sollten.

Der folgende Satz beschreibt einige für die Algebra wichtige Eigenschaften von Quotienten.

Proposition 6.35. *es sei R eine Äquivalenzrelation auf M.*

(1) $y \in [x]_R$ *gdw* yRx.

(2) Die Quotientenabbbildung p *is surjektiv. Es gilt* $p(x) = p(y)$ *gdw* xRy.

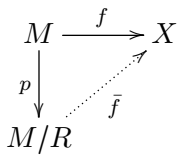
(3) Es sei X *eine Menge. Sei* $f\colon M \to X$ *eine Funktion. Dann sind äquivalent*

(a) $\forall x, y \in M(xRy \to f(x) = f(y))$. *(Man sagt dazu: „*$f$ *hängt nur von der Äquivalenzklasse ab" oder* f *ist unabhängig von den Repräsentanten)*

(b) Es gibt die von f *und* R *induzierte Abbildung* $\bar{f}\colon M/R \to X$, *so dass*

$$\forall x \in M\, \bar{f}(p(x)) = \bar{f}([x]_R) = f(x).$$

Man schreibt die Eigenschaft (3) (b) auch gerne in einem Diagramm



Beweis: (1) und (2) sind klar. (3): Die Implikation von (b) nach (a) ist einfach.

Wir zeigen die Implikation von (a) nach (b): Zuerst geben wir einen dummen Beweis, der mit Kanonen auf Spatzen schießt: Wir definieren $\bar{f}$. Sei hierzu S eine Auswahlfunktion für M und R. Wir setzen für $y \in M/R$,

$$\bar{f}(y) = f(S(y)).$$

Falls $y = p(x)$, so haben wir $\bar{f}(p(x)) = f(S(p(x))) = f(x)$, obwohl im Allgemeinen $S(p(x)) \neq x$ und nur $S(p(x))Rx$. Aber letzteres reicht ja aus, da f unabhängig von den Repräsentanten ist.

Nun geben wir einen Beweis, der das Auswahlaxiom nicht braucht: Wir setzen für $y \in M/R$,

$$\bar{f}(y) = \bigcup\{f(z) : z \in y\}.$$

Die Menge $\{f(z) : z \in y\}$ hat nur ein Element. Es gilt immer $\bigcup\{u\} = u$. Sei nun $y = p(x) = x/R$. Dann ist $x \in y$, und daher kommt $f(x)$ in der Menge $\{f(z) : z \in y\}$ als Element vor. Da diese Menge eine Einermenge ist, haben wir $\{f(z) : z \in y\} = \{f(x)\}$. Dann ist $\bigcup\{f(z) : z \in y\} = f(x)$, wie gewünscht. □

Man sagt „$\bar{f}$ is wohldefiniert", wenn Eigenschaft (3)(a) vorliegt.

Bei der Definition von + und · auf $\mathbb{Z}$ werden wir Wohldefiniertheit auch für Funktionen $\bar{f}\colon M/R \times M/R \to X$ und für Relationen $<\subseteq M/R \times M/R$ benutzen. Proposition 6.35 gilt sinngemäß auch für diese: Im Kriterium (3)(a) wird man nun für beide Argumente unabhängig voneinander Unabhängigkeit von den Repräsentanten voraussetzen.

Definition 6.36. und Behauptung. (Die ganzen Zahlen)

(1) Wir definieren eine Relation R auf $\mathbb{N}\times\mathbb{N}$, d.h. $R\subseteq(\mathbb{N}\times\mathbb{N})\times(\mathbb{N}\times\mathbb{N})$ wie folgt

$$(m,n)R(p,q) \text{ wenn } m+q=p+n.$$

(2) Wir setzen

$$\mathbb{Z}=(\mathbb{N}\times\mathbb{N})/R$$

(3) Wir definieren die Addition auf $\mathbb{Z}$:

$$[(m,n)]_R+[(p,q)]_R=[(m+p,n+q)]_R.$$

Überlegen Sie sich, ob dies wohldefiniert ist.

(4) Wir definieren die Multiplikation auf $\mathbb{Z}$ durch

$$[(m,n)]_R\cdot[(p,q)]_R=[(mp+nq,mq+np)]_R.$$

(5) Wir definieren das additive Inverse $-[(m,n)]_R$ als

$$-[(m,n)]_R=[(n,m)]_R.$$

(6) Wir definieren die lineare Ordnung $<$ auf $\mathbb{Z}$ durch

$$[(m,n)]_R<[(p,q)]_R \text{ wenn } m+p<_{\mathbb{N}} n+q.$$

Auch dies ist wohldefiniert: Wieder weist man nach, dass die Definition nicht von den Repräsentanten abhängt.

Überlegen Sie sich: In Übereinstimmung mit Proposition 6.35 müssten wir in (3) bis (6) nun $\bar{+}$, $\bar{\cdot}$, $\bar{<}$ auf der linken Seite schreiben. Wir schenken uns die Oberstriche.

Beweis: Wir beweisen (4), da es am kompliziertesten aussieht. Sei

$$(m',n')R(m,n).$$

Das heißt

$$m'+n=m+n'.$$

Wir sollen zeigen

$$(mp+nq,mq+np)R(m'p+n'q,m'q+n'p),$$

d.h.

$$mp+nq+m'q+n'p=m'p+n'q+mq+np.$$

Nach dem Distributivgesetz und das Kommutativgesetz ist die linke Seite $(m + n')p + (n + m')q$. Nun setzt man hierin $m + n' = m' + n$ ein, tauscht also beide Vorfaktoren aus. Dann hat man nach einer weiteren Anwendung des Distributivgesetzes gerade die rechte Seite. Danach ersetzt man (p, q) durch ein äquivalentes (p', q') und zeigt mit der analogen Technik, dass

$$(m'p + n'q, m'q + n'p)R(m'p' + n'q', m'q' + n'p').$$

Führte man beide Ersetzungen geichzeitig durch, geriete man in längere Arbeit mit dem Umordnen von Termen und Hinzufügen geeigneter Summanden. □

Definition 6.37. Eine Gruppe $(G, 0, \circ, <)$ heißt angeordnete Gruppe, wenn die Kürzungsregel gilt: $a < b \leftrightarrow a \circ c < b \circ c$.

Definition 6.38. Ein Ring $(R, +, \cdot, 0, 1, <)$ heißt angeordneter Ring, wenn $(R, +, \cdot, 0, 1)$ ein Ring ist und wenn $(R, +, <)$ eine angeordnete Gruppe ist und wenn auch für $\cdot$ die Kürzungsregel gilt: $c > 0 \to (a < b \leftrightarrow ac < bc)$.

Definition 6.39. Ein angeordneter Ring $(R, +, \cdot, 0, 1, <)$ heißt archimedisch oder archimedisch angeordnet, wenn es eine Einbettung $i\colon (\mathbb{N}, +, \cdot, 0, 1, <) \to (R, +, \cdot, 0, 1, <)$ gibt, so dass $i[\mathbb{N}]$ konfinal in $(R, <)$ liegt, d.h. $\forall r \in R \exists n \in \mathbb{N} r \leq i(n)$.

Nun rechnet man nach:

Proposition 6.40. *(1)* $i(n) = [(n, 0)]_R$ *ist eine Einbettung von* $(\mathbb{N}, +, 0)$ *in* $(\mathbb{Z}, +, [(0, 0)]_R)$.

(2) $(\mathbb{Z}, +, [(0, 0)]_R)$ *ist eine abelsche Gruppe.*

(3) $(\mathbb{Z}, +, \cdot, (20, 20)/R, [(21, 20)]_R, <)$ *ist ein kommutativer archimedisch angeordneter Ring mit Eins.*

Definition 6.41. und Behauptung. (Die rationalen Zahlen)

(1) Wir definieren eine Relation E auf $\mathbb{Z}\times(\mathbb{N}\setminus\{0\})$, d.h. $R \subseteq (\mathbb{Z}\times(\mathbb{N}\setminus\{0\}))\times(\mathbb{Z}\times(\mathbb{N}\setminus\{0\}))$ wie folgt

$$(m, n)E(p, q) \text{ wenn } m \cdot q = p \cdot n.$$

(2) Wir setzen

$$\mathbb{Q} = (\mathbb{Z} \times (\mathbb{N} \setminus \{0\}))/E$$

(3) Wir definieren die Addition auf $\mathbb{Q}$:

$$[(m, n)]_E + [(p, q)]_E = [(mq + pn, nq)]_E.$$

Überlegen Sie sich, ob dies wohldefiniert ist.

(4) Wir definieren die Multiplikation auf $\mathbb{Q}$ durch

$$[(m,n)]_E \cdot [(p,q)]_E = [(mp,qn)]_E.$$

(5) Wir definieren für $m \neq 0$ das multiplikative Inverse $([(m,n)]_E)^{-1}$ als

$$([(m,n)]_E)^{-1} := \begin{cases} [(n,m)]_E, & \text{wenn } m > 0, \\ [(-n,-m)]_E, & \text{wenn } m < 0. \end{cases}$$

(6) Wir setzen $0_{\mathbb{Q}} = [(0,1000)]_E$, $1_{\mathbb{Q}} = [(17,17)]_E$.

(7) Wir definieren die lineare Ordnung $<$ auf $\mathbb{Q}$ durch

$$[(m,n)]_E < [(p,q)]_E \text{ wenn } mq <_{\mathbb{N}} pn.$$

Auch dies ist wohldefiniert: Wieder weist man nach, dass die Definition nicht von den Repräsentanten abhängt.

Beweis: Wir beweisen (3), da es am kompliziertesten aussieht. Sei

$$(m',n')E(m,n).$$

Das heißt

$$m' \cdot n = m \cdot n'.$$

Wir sollen zeigen

$$(mq+np,nq)E(m'q+n'p,n'q),$$

d.h.

$$(mq+np)n'q = (m'q+n'p)nq.$$

Nach dem Distributivgesetz und das Kommutativgesetz ist die linke Seite $mqn'p+npn'q$. Nun setzt man hierin $m \cdot n' = m' \cdot n$ ein, tauscht also den ersten Vorfaktor aus. Dann ist die linke Seite also $m'qnp+npn'q$. Die rechte Seite ist $m'qnp+n'pnq$ nach dem Distributivgesetz, stimmt also mit der linken Seite überein. Danach ersetzt man (p,q) durch ein äquivalentes (p',q') und zeigt mit der analogen Technik, dass

$$(m'q+n'p,n'q)E(m'q'+n'p',n'q').$$

□

Definition 6.42. Eine lineare Ordnung heißt *dicht*, wenn $\forall x < z \exists y(x < y < z)$.

Nun rechnet man (Skeptiker müssen viel rechnen) nach:

Proposition 6.43. *(1)* $i(z) = (z,1)/E$ *ist eine Einbettung von* $(\mathbb{Z},+,\cdot,0,1,<)$ *in* $(\mathbb{Q},+\cdot,0_\mathbb{Q},1_\mathbb{Q},<$ $)$.

(2) $(\mathbb{Q},+,\cdot,0_\mathbb{Q},1_\mathbb{Q},<)$ *ist ein archimedisch angeordneter Körper mit einer dichten linearen Ordnung.*

Definition 6.44. Wir definieren die Betragsfunktion auf $\mathbb{Q}$, durch

$$|q| = \begin{cases} q, & \text{wenn } q \geq 0, \\ -q, & \text{wenn } q < 0. \end{cases}$$

Erinnern Sie sich an die Cauchyfolgen.

Definition 6.45. Eine Funktion $f\colon \mathbb{N} \to \mathbb{Q}$ heißt Cauchyfolge (mit rationalen Einträgen)[12] oder Cauchyfolge in $\mathbb{Q}$, wenn

$$\forall \varepsilon > 0\ \exists n_0\ \forall m, m' \geq n_0\ |f(m) - f(m')| < \varepsilon. \tag{6.1}$$

Mit „Cauchyfolge" meint man meistens eine reellwertige Cauchyfolge, d.h. eine Funktion $f\colon \mathbb{N} \to \mathbb{R}$ (siehe unten), für die wiederum (6.1) gilt.

Definition 6.46. und Behauptung. (Die reellen Zahlen)

(1) Wir definieren eine Relation eq auf $\{f : f\colon \mathbb{N} \to \mathbb{Q}, f \text{ Cauchyfolge}\} \subseteq \mathscr{P}(\mathbb{N} \times \mathbb{Q})$ wie folgt

$$f \text{ eq } g \text{ wenn } \lim_{n\to\infty} |f(n) - g(n)| = 0.$$

Dies ist eine Äquivalenzrelation.

(2) Wir setzen

$$\mathbb{R} = \{f : f\colon \mathbb{N} \to \mathbb{Q}, f \text{ Cauchyfolge}\}/\text{eq}$$

(3) Wir definieren die Addition auf $\mathbb{R}$:

$$[f]_{\text{eq}} + [g]_{\text{eq}} = [(f+g)]_{\text{eq}}.$$

Hier ist für $n \in \mathbb{N}$, $(f+g)(n) = f(n) + g(n)$. Überlegen Sie sich, ob die Addition wohldefiniert ist.

(4) Wir definieren die Multiplikation auf $\mathbb{R}$ durch

$$[f]_{\text{eq}} \cdot [g]_{\text{eq}} = [(f \cdot g)]_{\text{eq}}.$$

Hier ist für $n \in \mathbb{N}$, $(f \cdot g)(n) = f(n) \cdot g(n)$. Überlegen Sie sich, ob die Multiplikation wohldefiniert ist.

[12] Augustin-Louis Cauchy, 1789–1857

(5) $0_{\mathbb{R}}$ sei die eq-Klasse einer Funktion mit $\lim_{n\to\infty} f(n) = 0_{\mathbb{Q}}$. $1_{\mathbb{R}}$ sei die eq-Klasse irgendeiner Funktion mit $\lim_{n\to\infty} f(n) = 1_{\mathbb{Q}}$.

(6) Wir definieren die lineare Ordnung $<$ auf $\mathbb{R}$ durch

$$[f]_{\mathrm{eq}} \leq [g]_{\mathrm{eq}} \text{ wenn } \lim_{n\to\infty} g(n) - f(n) \geq 0.$$

Auch dies ist wohldefiniert: Wieder weist man nach, dass die Definition nicht von den Repräsentanten abhängt.

(7) Wir definieren $|[f]_{\mathrm{eq}}| := [(n \mapsto |f(n)|)]_{\mathrm{eq}}$. Durch diese Setzung wird die Betragsfunktion von $i[\mathbb{Q}](= i(\mathbb{Q}))$ auf $\mathbb{R}$ erweitert.

Nun kommt ein sehr schöner Satz, im Original von Cantor[13].

Satz 6.47. *(1)* $i(q) = [c_q]_{\mathrm{eq}}$ *mit* $c_q(n) = q$ *für* $n \in \mathbb{N}$ *ist eine Einbettung von* $(\mathbb{Q}, +, \cdot, 0_{\mathbb{Q}}, 1_{\mathbb{Q}}, <)$ *in* $(\mathbb{R}, +, \cdot, 0_{\mathbb{R}}, 1_{\mathbb{R}}, <)$.

(2) $(\mathbb{R}, +, \cdot, 0_{\mathbb{R}}, 1_{\mathbb{R}}, <)$ *ist ein vollständiger archimedisch geordneter Körper.*

Beweis: Wir beweisen die meiner Meinung nach schwierigste Aussage:

Lemma 6.48. *In* $\mathbb{R}$ *hat jede Cauchyfolge einen Grenzwert.*

Beweis: Sei $(r_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, und sei $r_n = [f_n]_{\mathrm{eq}}$, f_n eine Cauchyfolge mit rationalen Einträgen.

Dann gilt, da $(r_n)_n$ eine Cauchyfolge ist:

$$\forall \varepsilon > 0 \exists n_0 = n_0(\varepsilon) \forall m, m' \geq n_0(\varepsilon) \lim_{k\to\infty} |f_m(k) - f_{m'}(k)| < \varepsilon. \tag{6.2}$$

Dies heißt ausgeschrieben

$$\forall \varepsilon > 0 \exists n_0 = n_0(\varepsilon) \forall m, m' \geq n_0 \forall \varepsilon_1 > 0$$
$$\exists k_{\mathrm{riesig}}(\varepsilon, m, m', \varepsilon_1) \forall k \geq k_{\mathrm{riesig}}(\varepsilon, m, m', \varepsilon_1) |f_m(k) - f_{m'}(k)| < \varepsilon + \varepsilon_1.$$

Außerdem ist jedes f_u eine Cauchyfolge, daher gilt für jedes u:

$$\forall \varepsilon > 0 \exists k_0 = k_0(\varepsilon, u) \forall m, m' \geq k_0(\varepsilon, u) |f_u(m') - f_u(m)| < \varepsilon. \tag{6.3}$$

Wir definieren nun eine (schnell wachsende) Funktion $g\colon \mathbb{N} \to \mathbb{N}$ wie folgt.

Gegeben n nehmen wir $g(n)$, so dass

$$\forall m, m' \geq g(n)\ \forall k \leq n\, |f_k(m) - f_k(m')| \leq \frac{1}{n}. \tag{6.4}$$

[13] Georg Cantor, 1845 – 1918

Überlegen Sie sich, dass es so ein g tatsächlich gibt: Wir nutzen zum Finden von $g(n)$ endlich oft die Voraussetzung (6.3). Wir setzen nun

$$f_{\mathrm{diag}}(n) \coloneqq f_n(g(n)).$$

Man nennt solche Definitionen, die aus einer Folge von Folgen durch Einsetzung eines Arguments einmal als Index und einmal als Argument eine quer darüberlaufende Folge bilden, Diagonalisierungen. Diagonalisierung ist ein wichtiger Kunstgriff in vielen Gebieten der Mathematik. Auch in (6.4) haben wir n schon an zwei verschiedenen Stellen eingesetzt.

Wir behaupten

(a) f_{diag} ist eine Cauchyfolge und

(b)
$$\lim_{n\to\infty} r_n = [f_{\mathrm{diag}}]_{\mathrm{eq}}.$$

Wir zeigen zuerst (a). Sei $\varepsilon > 0$ gegeben. Wir nehmen n_0, so dass $\varepsilon < \frac{1}{3\cdot n_0}$ und für $m, m' \geq n_0, |r_m - r_{m'}| < \frac{\varepsilon}{6}$. Dann ist für $m, m' \geq n_0$,

$$\begin{aligned}
&f_{\mathrm{diag}}(m) - f_{\mathrm{diag}}(m') = \\
&f_m(g(m)) - g_{m'}(g(m')) = \\
&(f_m(g(m)) - f_m(k)) + (f_m(k) - f_{m'}(k)) + (f_{m'}(k) - f_{m'}(g(m'))).
\end{aligned}$$

Jeder Summand hat Betrag $< \frac{\varepsilon}{3}$, wenn wir $k \geq \max(g(m), g(m'))$ so groß wählen, dass $|f_m(k) - f_{m'}(k)| < |r_m - r_{m'}| + \frac{\varepsilon}{6} < \frac{\varepsilon}{6} + \frac{\varepsilon}{6}$. Solch ein k gibt es, da $r_m = [f_m]_{\mathrm{eq}}$ und $r_{m'} = [f_{m'}]_{\mathrm{eq}}$.

Nun zeigen wir (b). Wir haben zu zeigen:

$$\forall \varepsilon > 0 \exists n_0 \forall m \geq n_0 \lim_{k\to\infty} |f_m(k) - f_{\mathrm{diag}}(k)| < \varepsilon.$$

Dies bedeutet ausgeschrieben:

$$\begin{aligned}
&\forall \varepsilon > 0 \exists n_0 = n_0(\varepsilon) \forall m \geq n_0 \forall \varepsilon_1 > 0 \exists k_0(\varepsilon, m, \varepsilon_1) = k_0 \\
&\forall k \geq k_0 |f_m(k) - f_{\mathrm{diag}}(k)| < \varepsilon + \varepsilon_1.
\end{aligned} \tag{6.5}$$

Gegeben ε, wir nehmen n_0 wie $n_0(\varepsilon)$ in (6.2).

Damit erhalten wir

$$\forall m, m' \geq n_0 \lim_{k\to\infty} |f_m(k) - f_{m'}(k)| < \varepsilon.$$

Letzteres heißt ausgeschrieben:

$$\forall m, m' \geq n_0 \forall \varepsilon_1 > 0 \exists k_0 \forall k \geq k_0 |f_m(k) - f_{m'}(k)| < \varepsilon + \varepsilon_1. \tag{6.6}$$

Gegeben m und ε_1 nehmen wir nun $k_0(\varepsilon, m, \frac{\varepsilon_1}{3}) \geq \max(n_0, \frac{3}{\varepsilon_1})$ als k_0 zur Cauchyfolge $(f_m(u))_{u\in\mathbb{N}}$ wie in (6.3). Damit erhalten wir

$$\forall k, k' \geq k_0 |f_m(k') - f_m(k)| < \frac{\varepsilon_1}{3} \tag{6.7}$$

Dann haben wir für jedes $k \geq k_0 \geq n_0$ eine Zahl $k_{\text{riesig}}(\varepsilon, k, m, \varepsilon_1) \geq \max(k, g(k))$ zur Cauchyfolge $|f_m - f_k|$ wie in (6.6) zu $\varepsilon + \frac{\varepsilon_1}{3}$.

Nun haben wir nach unseren Vorgehen
ε gegeben,
$n_0 = n_0(\varepsilon)$ gewählt,
$m \geq n_0$, ε_1 gegeben,
$k_0 = k_0(\varepsilon, m, \frac{\varepsilon_1}{3})$ gewählt,
$k \geq k_0$ gegeben,
$k_{\text{riesig}} = k_{\text{riesig}}(\varepsilon, k, m, \frac{\varepsilon_1}{3})$ gewählt. Wir rechnen nun:

$$f_m(k) - f_{\text{diag}}(k) = f_m(k) - f_m(k_{\text{riesig}}) + f_m(k_{\text{riesig}}) - f_k(k_{\text{riesig}}) + f_k(k_{\text{riesig}}) - f_k(g(k)).$$

Nach der Dreieicksungleichung und nach nach (6.7), (6.6) und (6.4) (jeweils zur Abschätzung der drei Summanden) ergibt sich

$$|f_m(k) - f_{\text{diag}}(k)| \leq$$
$$|f_m(k) - f_m(k_{\text{riesig}})| + |f_m(k_{\text{riesig}}) - f_k(k_{\text{riesig}})| + |f_k(k_{\text{riesig}}) - f_k(g(k))| <$$
$$\frac{\varepsilon_1}{3} + \varepsilon + \frac{\varepsilon_1}{3} + \frac{\varepsilon_1}{3} \leq \varepsilon + \varepsilon_1.$$

Damit haben wir die Behauptung (6.5) mit ihrer recht komplexen Schachtelung der Quantoren gezeigt. In (6.5) kommt k_{riesig} nicht vor, die Vorgabe k hingegen schon. Die Zahl k_{riesig} haben wir nur als Hilfsmittel benutzt. □

Proposition 6.49. *In $\mathbb{R}$ gilt für Cauchyfolgen mit rationalen Einträgen: $[f]_{\text{eq}} = \lim_{n\to\infty} i(f(n))$. Hierbei ist $i\colon \mathbb{Q} \to \mathbb{R}$ die in Aussage (1) des Satzes genannte Einbettung.*

Beweis: Übung.

Übung 6.50. 1. Basteln Sie für $n \in \mathbb{N}$ jeweils eine rationalwertige Cauchyfolge $(f_n(k))_{k\in\mathbb{N}}$, so dass $([f_n]_{\text{eq}})$ eine Cauchyfolge in $\mathbb{R}$ ist und so dass $(f_n(n))_{n\in\mathbb{N}}$ keine Cauchyfolge ist.

2. Vielleicht könnte man sich im Beweis gerade eben auf Cauchyfolgen mit gewisser schneller Konvergenzrate beschränken und den Beweis so hoffentlich vereinfachen. Wir dürfen uns ja beliebige Repräsentanten für die f_n mit $r_n = [f_n]_{\mathrm{eq}}$ auswählen. So kann man, nach Vorbereitung, zu $g(n) = n$ kommen. Überlegen Sie sich, wie man das genau durchführen kann.

3. Jemand möchte gerne noch mehr reelle Zahlen haben und faktorisiert daher bei der Konstruktion von $\mathbb{R}$ nicht nach eq. Kann man dann noch rechnen? Was passiert mit der Archimedizität? Wenn Sie diese Frage interessiert, können Sie sich über Nichtstandard-Analysis informieren.

Korollar 6.51. *(*ZF *beweist:) Das Axiomensystem* KAV *für die vollständigen archimedisch geordneten Körper ist hat ein Modell, ist also widerspruchsfrei.*

Satz 6.52. KAV *hat bis auf in beide Richtungen stetige Isomorphismen genau ein Modell. Man sagt dazu: Das Axiomensystem* KAV *ist kategorisch.*

Beweis: Seien $(\mathbb{R}, +, \cdot, 0, 1, <)$ und $(S, +_s, \cdot_S, 0_S, 1_S, <_S)$ zwei vollständige archimedisch angeordnete Körper. Wir geben einen Isomorphismus $f\colon \mathbb{R} \to S$ an. Wir vereinfachen die Schreibweise für $\mathbb{R}$, und nehmen an, dass $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ nicht nur in $\mathbb{R}$ eingebettet sind, sondern Teilmengen sind. Wir setzen $f(0) = 0_S$, $f(n+1) = f(n) +_S 1_S$, $f(-n) = -_S f(n)$, $f(\frac{p}{q}) = \frac{f(p)}{f(q)}$. Hier ist auf der rechten Seite der Bruch in S und $\cdot_S$ ausgerechnet gemeint. Dann haben wir f schon auf $\mathbb{Q}$ definiert. Nun setzen wir f auf eindeutige Weise stetig fort, indem wir zu $r \in \mathbb{R}$ eine Cauchyfolge mit rationalen Einträgen wählen (wir brauchen nach dem Beweis von Prop. 6.35 das Auswahlaxiom nicht), sagen wir $(q_n)_n$. Dann setzen wir $f(r) = \lim_n f(q_n)$, hier wird der Limes in S gebildet. f ist surjektiv, da auch S ein vollständig archimedisch angeordneter Körper ist. Auch die Umkehrabbildung von f ist stetig.

f kommutiert mit $+$ und $\cdot$ nach Konstruktion: $f(r + r') = f(r) +_S f(r')$. Dies zeigt man wieder von den rationalen Zahlen ausgehend. Ebenso zeigt man, dass f die lineare Ordnung von $\mathbb{R}$ treu auf die lineare Ordnung von S abbildet. □

Literaturverzeichnis

[1] E. Artin. *Geometric algebra*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. Reprint of the 1957 original, A Wiley-Interscience Publication.

[2] Victor Bangert. *Lineare Algebra 2006/2007, Vorlesungsskript*. http://home.mathematik.uni-freiburg.de/geometrie/bangert. Universität Freiburg, 2007.

[3] Andreas Blass. Existence of bases implies the axiom of choice. In *Axiomatic set theory (Boulder, Colo., 1983)*, volume 31 of *Contemp. Math.*, pages 31–33. Amer. Math. Soc., Providence, RI, 1984.

[4] Egbert Brieskorn. *Lineare Algebra und analytische Geometrie. I*. Friedr. Vieweg & Sohn, Braunschweig, 1983. With historical notes by Erhard Scholz.

[5] Theodor Bröcker. *Lineare Algebra und analytische Geometrie*. Grundstudium Mathematik. [Basic Study of Mathematics]. Birkhäuser Verlag, Basel, 2003. Ein Lehrbuch für Physiker und Mathematiker. [A textbook for physicists and mathematicians].

[6] H.-D. Ebbinghaus. *Einführung in die Mengenlehre*. Hochschultaschenbuch, 4 edition, 2003.

[7] Sebastian Goette. *Lineare Algebra 2012/2013, Vorlesungsskript*. http://home.mathematik.uni-freiburg.de/frank/index.de.html. Universität Freiburg, 2013.

[8] Werner Greub. *Linear algebra*. Springer-Verlag, New York, fourth edition, 1975. Graduate Texts in Mathematics, No. 23.

[9] Klaus Jänich. *Linear algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1994.

[10] Max Koecher. *Lineare Algebra und analytische Geometrie*, volume 2 of *Grundwissen Mathematik [Basic Knowledge in Mathematics]*. Springer-Verlag, Berlin, 1983.

[11] Hans-Joachim Kowalsky. *Lineare Algebra*. Walter de Gruyter, Berlin-New York, 1977. Achte Auflage, de Gruyter Lehrbuch.

[12] Kenneth Kunen. *Set theory*, volume 34 of *Studies in Logic (London)*. College Publications, London, 2nd edition, 2013.

[13] Emmy Noether. Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern. *Math. Ann.*, 96(1):26–61, 1927.

[14] Ernst Steinitz. Zur Theorie der Moduln. *Math. Ann.*, 52(1):1–57, 1899.

[15] Ernst Zermelo. Beweis, daß jede Menge wohlgeordnet werden kann. *Math. Ann.*, 59:514–516, 1904.

[16] Martin Ziegler. *Mathematische Logik*. Mathematik kompakt. Birkhäuser, 2010.

[17] Martin Ziegler. *Lineare Algebra, Vorlesungsskript*. http://home.mathematik.uni-freiburg.de/ziegler. Universität Freiburg, 2012.

[18] Max Zorn. A remark on a method in transfinite algebra. *Bull. Amer. Math. Soc. N.S.*, 41:667–670, 1935.

Symbol- und Stichwortverzeichnis