Lineare Algebra 1

Stefan Schreieder

WS 2020/21

Inhaltsverzeichnis

Kapitel 0

Motivation

In diesem Kapitel diskutieren wir einige bekannte Konzepte aus der Schulmathematik. Hierbei liegt der Fokus auf der Anschauung und der Motivation. Im Gegensatz zum eigentlichen Inhalt dieser Vorlesung, wird in diesem motivierenden Kapitel auf mathematische Präzision verzichtet.

0.1 Der n-dimensionale reelle Standardvektorraum

Wir beginnen damit, die natürlichen Zahlen $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, die ganzen Zahlen $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$, sowie die reellen Zahlen $\mathbb{R}$ als bekannt anzunehmen. Geometrisch stellen wir uns die reellen Zahlen als Zahlengerade vor. Für eine natürliche Zahl $n \in \mathbb{N}$, sei $\mathbb{R}^n$ die Menge aller n-Tupel reeller Zahlen, d.h.

$$\mathbb{R}^n = \{(v_1, \dots, v_n) \mid v_1, \dots, v_n \in \mathbb{R}\}.$$

Hier gilt die Konvention dass $\mathbb{R}^0 := \{0\}$ die Menge mit nur einem Element ist. Ein n-Tupel $(v_1, \dots, v_n) \in \mathbb{R}^n$ nennen wir (Zeilen-)Vektor und die Menge aller solcher Vektoren, also $\mathbb{R}^n$, nennen wir den n-dimensionalen reellen Standardvektorraum. Manchmal nennen wir einen solchen Vektor auch einfach Punkt in $\mathbb{R}^n$.

Je nach Kontext ist es manchmal besser ein n-Tupel nicht als Zeilenvektor, sondern als Spaltenvektor

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

aufzufassen. Es macht in der Regel keinen echten Unterschied, ob man nun Zeilen- oder Spaltenvektoren verwendet. Da es aber später durchaus wichtig werden kann, ob man nun von Zeilen oder Spaltenvektoren spricht, sollte man sich zumindest in einer Argumentationslinie jeweils auf eine der beiden Konventionen festlegen. Hier in diesem Skript ist es vor allem eine Frage des zur Vefügung stehenden Platzes, der oft die Schreibweise als Zeilenvektoren gegenüber Spaltenvektoren als angenehmer erscheinen lässt.

Konkret ist also $\mathbb{R}^1 = \mathbb{R}$ die Zahlengerade. Die Menge $\mathbb{R}^2$ besteht aus allen Paaren von reellen Zahlen. Geometrisch entspricht das den Punkten in der reellen Ebene, da jeder solche Punkt (nach Einführung eines geeigneten Koordinatensystems) genau von zwei Koordinaten festgelegt wird. Die Menge $\mathbb{R}^3$ besteht aus allen Zeilenvektoren $(v_1, v_2, v_3) \in \mathbb{R}^3$ mit $v_i \in \mathbb{R}$ für $i = 1, 2, 3$. Also entspricht $\mathbb{R}^3$ den Punkten im drei-dimensionalen reellen Raum, denn ein jeder solcher Punkt kann (nach Einführung eines geeigneten Koordinatensystems) eindeutig durch drei Koordinaten festgelegt werden.

Für $n \geq 4$ gibt es keine direkte geometrische Anschauung von $\mathbb{R}^n$. Trotzdem kommt $\mathbb{R}^n$ für $n \geq 4$ völlig natürlich in unserem Leben vor. Zum Beispiel kann man $\mathbb{R}^4$ auffassen als den üblichen drei-dimensionalen Raum $\mathbb{R}^3$ der durch die ersten drei Koordinaten gegeben ist, zusammen mit einer zeitlichen Komponente, die als vierte Koordinate gegeben ist. Ähnlich kann man $\mathbb{R}^{3+n}$ als den Rgewöhnlichen drei-dimenisonalen Raum $\mathbb{R}^3$ auffassen, wobei man zusätzlich noch jedem Punkt einen Vektor der Länge n zuordnet, der bestimmte Eigenschaften des Punktes beschreibt (etwa physikalische Eigenschaften, wie Kräfte, Temperatur, Windstärke etc. die an jenem Punkt gelten).

Um noch ein Beispiel aus einem ganz anderen Bereich zu geben: Eine Bank die n Kunden hat kann zu jeder beliebigen Zeit die Kontostände aller Kunden messen und erhält damit einen Punkt in $\mathbb{R}^n$ der vom jeweiligen Zeitpunkt der Messung abhängt.

0.2 Rechenvorschriften

0.2.1 Skalarmultiplikation

Sei $v := (v_1, \ldots, v_n)$ ein Vektor im $\mathbb{R}^n$ und sei $\lambda \in \mathbb{R}$ eine reelle Zahl. Da man jeden Eintrag von v mit λ multiplizieren kann, können wir dann den Vektor $(\lambda \cdot v_1, \lambda \cdot v_2, \ldots, \lambda \cdot v_n)$ betrachten. Wir führen die Schreibweise $\lambda \cdot v$ für diesen Vektor ein. Wann immer wir in Zukunft eine solche Schreibweise einführen, werden wir folgendes Symbol verwenden:

$$\lambda \cdot v := (\lambda \cdot v_1, \lambda \cdot v_2, \ldots, \lambda \cdot v_n)$$

um zu verdeutlichen, dass die linke Seite der Gleichung per Definition der rechten Seite entspricht. Die obige Operation, welche der reellen Zahl λ und dem Vektor v den Vektor $\lambda \cdot v$ zuordnet nennen wir Skalarmultiplikation.

Wir fixieren nun $\lambda \in \mathbb{R}$ und betrachten die Abbildung

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n, v \mapsto \lambda \cdot v$$

die jedem Vektor v den Vektor $\lambda \cdot v$ zuordnet.

- Falls $\lambda = 0$, so kollabiert die obige Abbildung alle Vektoren auf den Nullpunkt $0 = (0, 0, \ldots, 0) \in \mathbb{R}^n$.
- Falls $\lambda = 1$, so entspricht obige Abbildung der Identität.
- Falls $\lambda > 1$, so entspricht obige Abbildung einer Streckung um den Faktor λ.
- Falls $0 < \lambda < 1$, so entspricht obige Abbildung einer Stauchung um den Faktor λ.
- Falls $\lambda = -1$, so entspricht obige Abbildung geometrisch der Spiegelung am Nullpunkt.
- Falls $\lambda < 0$, so entspricht obige Abbildung der Komposition einer Spiegelung am Nullpunkt zusammen mit einer Streckung bzw Stauchung um den Faktor $|\lambda|$.

Lemma 0.2.1. *Seien $\lambda, \lambda' \in \mathbb{R}$ reelle Zahlen und sei $v \in \mathbb{R}^n$. Dann gilt:*

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

In anderen Worten: es ist egal, ob man zuerst die Zahlen λ und λ' miteinander multipliziert und danach das Ergebnis wie oben Beschrieben auf den Vektor v anwendet, oder ob man zuerst λ' auf v anwendet und danach λ auf das so erhaltene Ergebnis anwendet.

Beweis. Per Definition gilt:

$$(\lambda \cdot \lambda') \cdot (v_1, \ldots, v_n) = ((\lambda \cdot \lambda') \cdot v_1, (\lambda \cdot \lambda') \cdot v_2, \ldots, (\lambda \cdot \lambda') \cdot v_n)$$

und

$$\begin{aligned} \lambda \cdot (\lambda' \cdot (v_1, \ldots, v_n)) &= \lambda \cdot (\lambda' \cdot v_1, \lambda' v_2, \ldots, \lambda' v_n) \\ &= (\lambda \cdot (\lambda' \cdot v_1), \lambda \cdot (\lambda' \cdot v_2), \ldots, \lambda \cdot (\lambda' \cdot v_n)). \end{aligned}$$

Da λ, λ' und v_i reelle Zahlen sind, und die Multiplikation in den reellen Zahlen das Assoziativgesetz, also

$$(\lambda \cdot \lambda') \cdot v_i = \lambda \cdot (\lambda' \cdot v_i)$$

erfüllt, sehen wir direkt dass

$$(\lambda \cdot \lambda') \cdot v = \lambda \cdot (\lambda' \cdot v)$$

gilt, weil alle Einträge dieser Vektoren übereinstimmen. □

0.2.2 Addition von Vektoren

Seien nun $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$ zwei Vektoren im $\mathbb{R}^n$. Wir können dann den Vektor

$$v + w := (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \in \mathbb{R}^n$$

betrachten und nenne ihn die Summe von v und w. Die Operation welche zwei Vektoren v und w ihre Summe $v + w$ zuordnet nennen wir Vektoraddition. Diese liefert eine Abbildung

$$\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad (v, w) \mapsto v + w$$

welche einem Paar (v, w) von Vektoren in $\mathbb{R}^n$ ihre Summe $v + w$ zuordnet.

Lemma 0.2.2. *Folgende Rechenregeln gelten für die oben definierte Vektoraddition in $\mathbb{R}^n$:*

(i) $v + w = w + v$ *für alle* $v, w \in \mathbb{R}^n$*;*

(ii) $v + 0 = v$ *für alle* $v \in \mathbb{R}^n$*, wobei* $0 = (0, 0, \ldots, 0) \in \mathbb{R}^n$ *den Nullvektor bezeichnet;*

(iii) $(\lambda \cdot v) + (\lambda \cdot w) = \lambda \cdot (v + v')$ *für alle* $\lambda \in \mathbb{R}$ *und* $v, v' \in \mathbb{R}^n$*.*

Beweis. Dies folgt genauso wie in Lemma 0.2.1 von den entsprechenden Rechenregeln in den reellen Zahlen. Wir geben die Details für (iii) und verwenden Spaltenvektoren in unserer Schreibweise. Unter Ausnutzung der Definitionen sowie der Rechenregeln für die reellen Zahlen erhalten wir

$$\lambda \cdot v + \lambda \cdot w = \lambda \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \lambda \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot v_1 \\ \lambda \cdot v_2 \\ \vdots \\ \lambda \cdot v_n \end{pmatrix} + \begin{pmatrix} \lambda \cdot w_1 \\ \lambda \cdot w_2 \\ \vdots \\ \lambda \cdot w_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot v_1 + \lambda \cdot w_1 \\ \lambda \cdot v_2 + \lambda \cdot w_2 \\ \vdots \\ \lambda \cdot v_n + \lambda \cdot w_n \end{pmatrix}$$

und

$$\lambda \cdot (v + w) = \lambda \cdot \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot (v_1 + w_1) \\ \lambda \cdot (v_2 + w_2) \\ \vdots \\ \lambda \cdot (v_n + w_n) \end{pmatrix} = \begin{pmatrix} \lambda \cdot v_1 + \lambda \cdot w_1 \\ \lambda \cdot v_2 + \lambda \cdot w_2 \\ \vdots \\ \lambda \cdot v_n + \lambda \cdot w_n \end{pmatrix}.$$

Also gilt $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$, wie behauptet. □

0.3 Geraden, Ebenen usw.

Die in den vorherigen Abschnitten eingeführten Operationen der Vektoraddition sowie der Skalarmultiplikation erlaubt es uns z.B. Geraden in $\mathbb{R}^2$ oder Ebenen in $\mathbb{R}^3$ zu beschreiben.

Definition 0.3.1. *Eine Teilmenge $L \subset \mathbb{R}^n$ heißt Gerade, falls es Vektoren $v, w \in \mathbb{R}^n$ mit $w \neq 0$ gibt, sodass*

$$L = \{v + \lambda \cdot w \mid \lambda \in \mathbb{R}\}$$

gilt. Wir verwenden hier auch die Schreibweise $v + \mathbb{R} \cdot w := \{v + \lambda \cdot w \mid \lambda \in \mathbb{R}\}$.

Per Definition ist eine Gerade also eine Teilmenge des $\mathbb{R}^n$ das als Bild der Abbildung

$$\mathbb{R} \longrightarrow \mathbb{R}^n, \quad \lambda \mapsto v + \lambda w$$

für Vektoren $v, w \in \mathbb{R}^n$ mit $w \neq 0$ beschrieben werden kann.

Satz 0.3.1. *Eine Teilmenge $L \subset \mathbb{R}^2$ ist genau dann eine Gerade, wenn es reelle Zahlen $a_1, a_2, b \in \mathbb{R}$ mit $a_1 \neq 0$ oder $a_2 \neq 0$ gibt, sodass L die Lösungsmenge der linearen Gleichung $a_1x_1 + a_2x_2 = b$ ist, d.h.*

$$L = \{(v_1, v_2) \in \mathbb{R}^2 \mid a_1v_1 + a_2v_2 = b\}.$$

Beweis. Der Beweis erfordert ein klein wenig Arbeit und soll in diesem motivierenden Kapitel übersprungen werden. □

Frage: Wie kann man eine Ebene in $\mathbb{R}^n$ (mit $n \geq 2$) beschreiben? Man könnte meinen, dass eine Ebene eine Teilmenge $E \subset \mathbb{R}^n$ ist, sodass es Vektoren $v, w_1, w_2 \in \mathbb{R}^n$ mit

$$E = \{v + \lambda_1 w_1 + \lambda_2 w_2 \mid (\lambda_1, \lambda_2) \in \mathbb{R}^2\}$$

gibt. Doch selbst wenn w_1 und w_2 beide ungleich Null sind, so wird obige Teilmenge des $\mathbb{R}^n$ sicher nicht immer eine Ebene beschreiben; dies klappt zum Beispiel nicht wenn $w_2 = w_1$ oder allgemeiner, wenn $w_2 \in w_1 \cdot \mathbb{R}$ gilt.

Definition 0.3.2. *Eine Teilmenge $E \subset \mathbb{R}^n$ heißt Ebene, falls es Vektoren $v, w_1, w_2 \in \mathbb{R}^n$ mit $w_1 \neq 0$ und $w_2 \notin \mathbb{R} \cdot w_1$ gibt, sodass*

$$E = \{v + \lambda_1 w_1 + \lambda_2 w_2 \mid (\lambda_1, \lambda_2) \in \mathbb{R}^2\}$$

gilt. Wir verwenden hier auch die Schreibweise

$$v + \mathbb{R} \cdot w_1 + \mathbb{R} w_2 := \{v + \lambda_1 w_1 + \lambda_2 w_2 \mid (\lambda_1, \lambda_2) \in \mathbb{R}^2\}.$$

Wenn man nun dieses 'Spiel' fortsetzt und sich fragt, wie man Unterräume des $\mathbb{R}^n$ beschreiben kann, die aussehen wie ein 'verschobener' $\mathbb{R}^m$ für ein $m \leq n$, so wird man unweigerlich auf das Konzept der linearen Unabhängigkeit von Vektoren stoßen, das wir später im Detail studieren werden.

0.4 Lineare Gleichungssysteme

Seien $n, m \geq 1$ positive natürliche Zahlen. Ein lineares Gleichungssystem ist ein System von Gleichungen der Form

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

wobei $a_{ij} \in \mathbb{R}$ und $b_i \in \mathbb{R}$ gegebene relle Zahlen sind und man Lösungen in den unbestimmten $x_1, \ldots, x_m$ sucht.

Die natürliche Frage bei einem solchen System an linearen Gleichungen lautet: Gibt es eine Lösung dieses Systems, d.h. gibt es einen Vektor $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ sodass alle oben Beschriebenen Gleichungen erfüllt sind? Falls eine solche Lösung gibt: Wie viele Lösungen gibt es und wie sieht der Raum aller Lösungen aus?

Lineare Gleichungssysteme tauchen überall in der Mathematik und anderen Wissenschaften auf. Der Versuch derartige Gleichungssysteme zu verstehen und zu lösen war einer der entscheidenden Triebfedern für die moderne Lineare Algebra.

Einem linearen Gleichungssystem wie oben kann man eine Abbildung

$$f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

mit

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} \mapsto \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1m}v_m \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2m}v_m \\ \vdots \\ a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nm}v_m \end{pmatrix}$$

zuordnen. Die Lösbarkeit von obigem System von linearen Gleichungen ist dann gleichbedeutend mit der Frage, ob es einen Punkt $v \in \mathbb{R}^n$ gibt, sodass

$$f(v) = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

gilt.

Die Abbildung f von oben hängt offensichtlich nur von den Koeffizienten a_{ij} ab und es ist oft nützlich diese in der Form einer Matrix anzuordnen:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}.$$

Um das Problem der Lösbarkeit von obigem Gleichungssystem genauer zu verstehen, kommen nun natürlicherweise folgende Fragen zum Vorschein:

- Wie genau hängt die Abbildung $f : \mathbb{R}^m \to \mathbb{R}^n$ von der Matrix A welche die Koeffizienten a_{ij} kodiert ab?

- Können wir grundlegende Eigenschaften der Abbildung $f : \mathbb{R}^m \to \mathbb{R}^n$ benennen die unabhängig von der konkreten Wahl der Koeffizienten a_{ij} immer gilt?
- Können wir die Familie aller solcher Abbildungen genauer studieren und Grundlegende Eigenschaften verstehen?

Um die Essenz einer mathematischen Aussage zu verstehen, ist es oft nötig das vorliegende Problem zu abstrahieren. Das ist oft aus verschiedenen Gründen eine gute Idee: Zum einen erweitert es den möglichen Anwendungsbereich, weil die Aussagen allgemeiner und weniger speziell werden. Zum anderen erleichtert es aber auch oft die Suche nach der Lösung, denn wenn man ersteinmal die "nicht-relevanteÏnformation ausgeblendet hat, so bleibt hoffentlich nur noch diejenige, die für das Problem wirklich relevant und wichtig ist, und man hofft, dass in dieser kondensierten Information auch die Lösung zu suchen ist, oder noch besser: von alleine ergibt. Im vorliegenden Fall stellen sich natürlicherweise zB folgende Fragen:

- Welche Struktur haben die Räume $\mathbb{R}^m$ und $\mathbb{R}^n$ sowie die Abbildung $f : \mathbb{R}^m \to \mathbb{R}^n$ genau?
- Können wir allgemein gültige Struktursätze über solche Räume und Abbildungen dazwischen zeigen?
- Welche Struktur von $\mathbb{R}$ braucht man wirklich um obige Konstruktion von linearen Gleichungssystemen durchführen zu können?
- Können wir in obiger Diskussion $\mathbb{R}$ durch andere Mengen mit vernünftigen Rechenvorschriften ersetzen, sodass der Formalismus im Wesentlichen immer noch so wie im Fall der reellen Zahlen funktioniert?

Kapitel 1

(Ein wenig) Mengenlehre

1.1 Mengen

Definition 1.1.1 (Cantor). *Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens (welche Elemente genannt werden) zu einem Ganzen.*

Obige Definition wurde von Cantor (1845-1918) aufgestellt; einige Begriffe benötigen eine Interpretation bzw. Erklärung:

- Objekte: mathematische Objekte;
- wohlunterschieden: Gleichheit muss klar sein , d.h. es muss jederzeit möglich sein die Gleichheit von zwei Objekten zu bestimmten;
- Zusammenfassung zu einem Ganzen: Damit meint Cantor, dass wir hiermit eine neues mathematisches Objekt erschaffen bzw. definieren.

Beispiele:

- Die Menge $\mathcal{Z} = \{0, 1, 2, 3, \ldots, 9\}$ aller Ziffern;
- Die Menge $\mathcal{A} := \{A, B, C, \ldots, Z\}$ aller Großbuchstaben;
- Die Menge $\mathfrak{a} := \{a, b, c, \ldots, z\}$ aller Kleinbuchstaben;
- Die Menge aller natürlicher Zahlen $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$;
- Die Menge aller ganzen Zahlen $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

Auch wenn unsere obige Schreibweise immer eine natürliche Ordnung der Elemente unserer Mengen enthält, sind Mengen grundsätzlich ungeordnet.

Eine Menge M definiert sich über ihre Elemente und wir schreiben $x \in M$ falls x ein Element der Menge M ist. Zwei Mengen sind genau dann gleich, wenn beide Mengen dieselben Elemente enthalten. Falls x kein Element der Menge M ist, so schreiben wir $x \notin M$.

Beispiele:

- $1 \in \mathcal{Z}$ aber $-1 \notin \mathcal{Z}$;
- $A \in \mathcal{A}$ aber $a \notin \mathcal{A}$;

- $a \in \mathfrak{a}$ aber $A \notin \mathfrak{a}$;
- $0 \in \mathbb{N}$ aber $A \notin \mathbb{N}$;
- $-2020 \in \mathbb{Z}$ aber $\infty \notin \mathbb{Z}$.

Definition 1.1.2. *Sei M eine Menge. Eine Menge N heißt Teilmenge von M, falls jedes Element aus N auch in M enthalten ist. Das heißt:*

$$x \in N \quad \Longrightarrow \quad x \in M.$$

Wenn dies erfüllt ist, so schreiben wir $N \subset M$.

Definition 1.1.3. *Die leere Menge $\emptyset$ ist die Menge welche kein Element enthält.*

Es gilt also für jede beliebige Menge M:

$$x \in M \quad \Longrightarrow \quad x \notin \emptyset.$$

Umgekehrt ist die leere Menge eine Teilmenge einer jeden beliebigen anderen Menge M:

$$\emptyset \subset M.$$

Definition 1.1.4. *Sei M eine Menge. Dann ist die Potenzmenge $\mathcal{P}(M)$ die Menge aller Teilmengen von M:*

$$N \in \mathcal{P}(M) \quad \Longleftrightarrow \quad N \subset M.$$

Beispiele:

- $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$ hat 4 (verschiedene) Elemente,
- $\mathcal{P}(\emptyset) = \{\emptyset\}$ hat ein Element,
- $\mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$ hat 2 (verschiedene) Elemente,
- $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$ hat 4 (verschiedene) Elemente.

Definition 1.1.5. *Seien M und N Mengen. So heißt*

- *$M \setminus N := \{x \mid x \in M \quad \text{und} \quad x \notin N\}$ das Komplement von N in M;*
- *$M \cap N := \{x \mid x \in M \quad \text{und} \quad x \in N\}$ der Schnitt von M und N;*
- *$M \cup N := \{x \mid x \in M \quad \text{oder} \quad x \in N\}$ die Vereinigung von M und N;*

Bemerkung 1.1.6. *In obiger Schreibweise wird das Wort 'und' auch oft durch das Symbol $\wedge$ ersetzt; ähnlich wird das Wort 'oder' durch das Symbol $\vee$ ersetzt wird.*

Allgemeiner können wir den Schnitt oder die Vereinigung von nicht nur zwei Mengen, sondern von einer beliebigen Anzahl an Mengen definieren. Formal bedeutet das, dass wir eine Index Menge I haben (z.B. $I = \{1,2\}$ oder $I = \mathbb{N}$ oder jede andere Menge die Ihnen in den Sinn kommt) und für jedes Element $i \in I$ eine Menge M_i gegeben haben. Wir definieren dann:

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

sowie

$$\bigcup_{i \in I} M_i := \{x \mid \exists i \in I \ \text{ mit } \ x \in M_i\}.$$

Bemerkung 1.1.7. *Das Symbol* $\forall$ *bedeutet 'für alle'; das Symbol* $\exists$ *bedeutet 'es existiert ein'.*

Lemma 1.1.8. *Seien* A, B *und* C *Mengen, so gilt:*

(i) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;

(ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

(iii) $A \cap B = A \iff A \subset B \iff A \cup B = B$.

Beweis. Zwie Mengen sind gleich genau dann wenn sie die selben Elemente besitzen. Um die erste Behauptung im Lemma zu zeigen, müssen wir also zeigen, dass jedes Element $x \in A \cap (B \cup C)$ auch in $(A \cap B) \cup (A \cap C)$ enthalten ist, und dass umgekehrt jedes Element $y \in (A \cap B) \cup (A \cap C)$ auch in $A \cap (B \cup C)$ enthalten ist. Das sind zwei unterschiedliche und voneinander unabhängige Argumentationsketten, die wir nun finden müssen.

Wir beginnen damit anzunehmen, dass wir ein Element $x \in A \cap (B \cup C)$ gegeben haben. Das bedeutet, dass $x \in A$ gilt und entweder $x \in B$ oder $x \in C$ gilt. Falls $x \in B$ gilt, so haben wir $x \in A \cap B$ und damit $x \in (A \cap B) \cup (A \cap C)$ wie gewünscht. Wenn andererseits $x \in C$ gilt, so haben wir $x \in A \cap C$ und damit ebenfalls $x \in (A \cap B) \cup (A \cap C)$ wie gewünscht. Insgesamt haben wir also

$$A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$$

gezeigt.

Sei nun umgekehrt ein Element $y \in (A \cap B) \cup (A \cap C)$ gegeben. Dann gilt $y \in A \cap B$ oder $y \in A \cap C$ (das oder hier bedeutet nicht 'entweder oder'; insbesondere könnten auch beide Aussagen wahr sein, auf jeden Fall ist aber mindestens eine der beiden Aussagen wahr). Falls $y \in A \cap B$ gilt, so gilt $y \in A$ und $y \in B \subset B \cup C$ und somit $y \in A \cap (B \cup C)$ wie gewünscht. Wenn andererseits $y \in A \cap C$ gilt, so gilt $y \in A$ und $y \in C \subset B \cup C$ und somit erhalten wir ebenfalls $y \in A \cap (B \cup C)$. Insgesamt haben wir damit

$$(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$$

gezeigt. Da wir die andere Inklusion oben schon gezeigt haben, schließen wir insgesamt dass $(A \cap B) \cup (A \cap C)$ und $A \cap (B \cup C)$ genau diesselben Elemente enthalten und somit übereinstimmen. Das beweist Aussage (i). Die anderen Aussagen werden ganz ähnlich bewiesen. $\square$

Definition 1.1.9. *Seien* M *und* N *Mengen. Falls* $M \cap N = \emptyset$ *gilt, so schreiben wir* $M \sqcup N$ *oder* $M \dot\cup N$ *anstelle von* $M \cup N$ *und nennen die Vereinigung von* M *und* N *auch disjunkte Vereinigung. Allgemeiner schreiben wir*

$$\dot\bigcup_{i \in I} M_i := \bigcup M_i$$

falls M_i *mit* $i \in I$ *Mengen sind, sodass* $M_i \cap M_j = \emptyset$ *für alle* $i \neq j$ *gilt.*

1.2 Russell's Paradoxon und die Axiome von Zermelo–Fraenkel

Die Mengenlehre die wir in vorherigem Abschnitt aufgesetzt haben wirkt auf den ersten Blick absolut natürlich und unproblematisch. Es war deshalb ein 'Schock' für die damalige Wissenschaft als Russell 1903 folgende Beobachtung publizierte.

Russell begann damit die Menge R aller Mengen M zu definieren, welche sich selbst nicht als Element enthalten. Also formal:

$$R := \{M \mid M \text{ ist eine Menge sodass } M \notin M \text{ gilt}\}.$$

Russell argumentierte dann, dass die Menge R gar nicht existieren kann. In der Tat, nehmen wir a die Menge R wie wir sie oben definiert haben und wie sie scheinbar der Definition von Cantor genügt existiert. Wir fragen uns dann, ob R ein Element von R ist und diese Frage sollte entweder eine positive oder negative Antwort haben. Es sollte also entweder $R \in R$ oder $R \notin R$ gelten. Tatsächlich kann aber keines von beidem gelten, wie man leicht sehen kann: Da R per definition nur Mengen enthält, die sich nicht selbst entahlten, kann sicherlich nicht $R \in R$ gelten. Falls aber nun $R \notin R$ gilt, so muss $R \in R$ nach Definition der Menge R gelten. Das ist ein Widerspruch der zeigt, dass R wie oben definiert keine Menge sein kann.

Diese Paradoxon kann man auf verschiedene Weisen lösen. In jedem Fall war nach Russell's Entdeckung klar, dass man den Mengenbegriff präziser und vorsichtiger definieren muss. Historisch hat sich dazu der Zugang von Zermelo und Fraenkel durchgesetzt, welche eine Reihe von grundlegenden Axiome aufgestellt haben, welche die Mengenlehre auf ein festes Fundament gestellt haben und insbesondere obiges Paradoxon ausschließt. Da die moderne Mathematik komplett auf den Axiomen von Remelo–Fraenkel basiert, wollen wir diese hier kurz angeben. Wir betonen aber gleichzeitig, dass ein weiterführendes Studium dieser Thematik uns zu weit von dem eigentlichen Ziel dieser Vorlesung (Lösen von linearen Gleichungssystemen) abbringen würde. Wir wollen hier lediglich darauf aufmerksam machen, dass man bei der Definition von Mengen ein klein wenig vorsichtig sein muss, um Probleme wie das von Russell oben auszuschließen.

Die Axiome von Zermelo–Fraenkel

(1) Mengen sind genau dann gleich, wenn sie dieselben Elemente enthalten.

(2) (Leerenmengenaxiom) Es gib die leere Menge $\emptyset$, also eine Menge ohne Elemente.

(3) (Paarmengenaxiom) Für zwei Objekte a und b gibt es immer eine Menge M die genau a und b enthält: $M = \{a, b\}$.

(4) (Vereinigungsaxiom) Für jede Menge A (von Mengen) gibt es eine Menge B die genau die Elemente der Elemente von A als Elemente hat. In anderen Worten, die folgende Menge B existiert:
$$B = \{b \mid \exists M \in A \text{ so dass } M \in A\}.$$
Zusammen mit dem obigen Axiom lässt sich damit insbesondere die Vereinigung von zwei Mengen definieren.

(5) (Unendlichkeitsaxiom) Es gibt eine Menge A welche die leere Menge enthält und zusätzlich mit jedem Element $x \in A$ auch die Menge $x \cup \{x\}$ enthält.

(6) (Potenzmengenaxiom) Für jede Menge A gibt es eine Menge $\mathcal{P}(A)$ deren Elemente genau die Teilmengen von A sind.

(7) (Regularitätssaxiom) Jede nichtleere Menge A enthält ein Element $B \in A$ sodass $A \cap B = \emptyset$, d.h. $x \in B \Rightarrow x \notin A$. (Dieses Axiom schließt das von Russell gefundene Paradoxon aus, denn die 'Menge' R von Russel die wir oben betrachtet hat erfüllt das Regularitätsaxiom nicht.)

(8) (Ersetzungsaxiom) Ist A eine Menge und wird jedes Element von A durch eine beliebige Menge ersetzt, so geht A in eine Menge über. (Beispiel: $\{\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}(\mathcal{P}(\mathbb{N})), \dots\}$ ist eine Menge.)

Zusätzlich zu den oben genannten Axiomen werden wir auch noch folgendes Axiom, das sogenannte Auswahlaxiom, fordern.

(9) Sei A eine Menge sodass

$$A = \dot{\bigcup}_{i \in I} A_i$$

die disjunkte Vereinigung von Teilmengen $A_i \subset A$ mit $i \in I$ ist. Dann gibt es eine Teilmenge $M \subset A$ sodass für jedes $i \in I$ die Menge M genau ein Element aus A_i enthält: $M \cap A_i = \{a_i\}$ für genau ein Element $a_i \in A_i$.

Das Auswahlaxiom sieht auf den ersten Blick harmlos und natürlich aus. Es folgt aber nicht aus den anderen Axiomen und hat in der Tat ebenso 'erstaunliche' wie 'offensichtliche' Konsequenzen. Zu den wohl erstaunlichsten Konsequenzen gehört das Banach–Tarski Paradoxon.

Satz 1.2.1. *(Banach–Tarski Paradoxon) Betrachte den drei-dimensionalen reellen Ball mit Durchmesser 1:*

$$B := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq 1\}.$$

Dann kann man B in sechs disjunkte Teilmengen zerlegen, sodass sich die einzelnen Teilmengen wieder so zusammenfügen lassen, dass zwei reelle Bälle von Radius 1 entstehen.

Die mathematische Erklärung von obigem Paradoxon liegt darin, dass die Teilmengen in die man B zerlegt so 'wild' sind, dass man ihnen kein Volumen zuordnen kann. (Offensichtlich kann man das nicht, denn sonst hätte der Ball mit Durchmesser 1 dasselbe Volumen wie zwei Bälle mit Durchmesser eins, also hätte er zwingend Volumen Null.) Obiger Satz wird in der Regel in der Maßtheorie (Analysis III) genauer beleuchtet und wir wollen uns hier nicht weiter damit beschäftigen.

1.3 Relationen

Definition 1.3.1. *Seien M und N Mengen. Dann heißt*

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

Paarmenge oder kartesisches Produkt von M und N.

Definition 1.3.2. *Sei M eine Menge. Eine Relation auf M ist eine Teilmenge $R \subset M \times M$.*

Falls $R \subset M \times M$ eine Relation ist, so schreiben wir für $x, y \in M$:

$$x \sim_R y \iff (x, y) \in R.$$

Manchmal schreiben wir auch $\sim_R$ für eine Relation, die Teilmenge $R \subset M \times N$ is dann durch obige Äquivalenz charakterisiert. Wenn der Kontext klar ist, schreiben wir manchmal auch $x \sim y$ anstelle von $x \sim_R y$.

Definition 1.3.3. *Sei M eine Menge und R (bzw. $\sim_R$) eine Relation auf M. Wir nennen diese Relation*

- *reflexiv, falls für alle $x \in M$: $x \sim_R x$;*
- *symmetrisch, falls für alle $x, y \in M$: $x \sim_R y \Rightarrow y \sim_R x$;*
- *transitiv, falls für alle $x, y, z \in M$: falls $x \sim_R y$ und $y \sim_R z$, so gilt $x \sim_R z$.*

Definition 1.3.4. *Sei M eine Menge. Eine Relation R auf M heißt Äquivalenzrelation falls sie reflexiv, symmetrisch und transitiv ist.*

Beispiel 1. Sie M die Menge aller Menschen und betrachte folgende Relationen auf M:

- $x \sim_{R_1} y :\Longleftrightarrow x$ und y haben diesselbe Mutter.

  Diese Relation ist reflexiv, symmetrisch und transitiv, also ein Äquivalenzrelation.

- $x \sim_{R_2} y :\Longleftrightarrow x$ kennt y.

  Diese Relation ist reflexiv, aber sicherlich nicht transitiv, und auch nicht symmetrisch (wir kennen alle Donald Trump, aber er wird die meisten von uns nicht kennen). Dies ist also keine Äquivalenzrelation.

- $x \sim_{R_3} y :\Longleftrightarrow x$ wiegt höchstens so viel wie y.

  Diese Relation ist symmetrisch und transitiv, aber nicht reflexiv.

- $x \sim_{R_4} y :\Longleftrightarrow x$ und y haben diesselbe Staatsbürgerschaft.

  Diese Relation ist reflexiv und symmetrisch, aber nicht transitiv (es gibt Leute die zwei Staatsbürgerschaften haben), also ist es keine Äquivalenzrelation.

Beispiel 2. Sei $M = \mathbb{Z}$ die Menge der ganzen Zahlen und betrachte folgende Relationen auf $M = \mathbb{Z}$:

- $x \sim_{R_1} y :\Longleftrightarrow x$ und y sind beides gerade Zahlen.

  Diese Relation ist symmetrisch und transitiv, aber nicht reflexiv. Also ist es keine Äquivalenzrelation.

- $x \sim_{R_2} y :\Longleftrightarrow x - y$ ist eine gerade Zahl.

  Diese Relation ist reflexiv, symmetrisch und transitiv. Also ist es eine Äquivalenzrelation.

- $x \sim_{R_3} y :\Longleftrightarrow x - y$ ist durch 2020 teilbar.

  Diese Relation ist reflexiv, symmetrisch und transitiv. Also ist es eine Äquivalenzrelation.

Definition 1.3.5. *Sei M eine Menge und sei R eine Äquivalenzrelation auf M.*

- *Für jedes $x \in M$ nennen wir die Teilmenge $[x] := \{y \in M \mid x \sim_R y\}$ die Äquivalenzklasse von R.*
- *Jedes Element $y \in [x]$ heißt Vertreter von $[x]$ in M.*

Die Eigenschaften einer Äquivalenzrelation haben folgende wichtige Folgerung.

Lemma 1.3.6. *Sei M eine Menge und sei R eine Äquivalenzrelation auf M. Dann gilt:*

(i) $M = \bigcup_{x \in M} [x]$;

(ii) Falls $x, y \in M$ Elemente sind mit $x \sim_R y$, so gilt $[x] = [y]$.

(iii) Falls $x, y \in M$ Elemente sind, sodass $[x] \cap [y] \neq \emptyset$ gilt, so gilt $[x] = [y]$.

Beweis. Da $\sim_R$ reflexiv ist, gilt $x \sim_R x$ für alle $x \in M$ und somit gilt $x \in [x]$ für alle $x \in M$. Dies impliziert $M \subset \bigcup_{x \in M} [x]$. Andererseits ist für alle $x \in M$, $[x] \subset M$ eine Teilmenge von M, sodass $\bigcup_{x \in M} [x] \subset M$ folgt. Also insgesamt $M = \bigcup_{x \in M} [x]$, was zu beweisen war. Das zeigt (i).

Um (ii) zu beweisen, nehmen wir an, dass $x, y \in M$ Elemente sind mit $x \sim_R y$. Um $[x] = [y]$ zu zeigen müssen wir zeigen, dass beide Mengen übereinstimmen. Dazu muss man zeigen, dass beide Mengen diesselben Elemente besitzen und dazu wiederum genügt es zu zeigen, dass sowohl $[x] \subset [y]$ als auch $[y] \subset [x]$ gilt. Wir zeigen zuerst dass $[x] \subset [y]$ gilt. Sei dazu $z \in [x]$. Das bedeutet, dass $x \sim_R z$ gilt. Da unsere Relation symmetrisch ist, gilt also auch $z \sim_R x$. Nach Annahme gilt $x \sim_R y$ und somit impliziert $z \sim_R x$ per Transitivität unserer Relation dass $z \sim_R y$ gilt. Nutzen wir nun noch einmal die Symmetrie unserer Relation aus, so erhalten wir $y \sim_R z$ und damit $z \in [y]$. Wir haben also gerade gezeigt, dass jedes Element $z \in [x]$ auch in $[y]$ enthalten ist, und somit gilt also $[x] \subset [y]$. Die umgekehrte Richtung, dass also jedes Element in $[y]$ auch in $[x]$ enthalten ist, geht völlig analog (vertauschen Sie einfach die Symbole x und y in obiger Argumentation). Das beweist (ii).

Um (iii) zu beweisen, nehmen wir an, dass $[x] \cap [y] \neq \emptyset$ gilt. Das bedeutet, dass es ein Element $z \in M$ mit $z \in [x]$ und $z \in [y]$ gibt. Per Definition heißt das, dass $x \sim_R z$ und $y \sim_R z$ gilt. Da die Relation symmetrisch ist, gilt auch $z \sim_R y$. Da die Relation reflexiv ist, impliziert nun aber $x \sim_R z$ und $z \sim_R y$ dass $x \sim_R y$ gilt. Also folgt $[x] = [y]$ aus (ii). Das schließt den Beweis des Lemmas ab. □

Obiges Lemma zeigt, dass eine Äquivalenzrelation auf einer Menge M in natürliche Weise die Elemente aus M in verschiedene "Klassenäufteilt, also eine Partition der Menge M in disjunkte Teilmengen die ganz M überdecken liefert. Diese Teilmengen nenne wir Äquivalenzklassen.

Definition 1.3.7. *Sei M eine Menge und sei R eine Äquivalenzrelation auf M. Eine Teilmenge $N \subset M$ heißt Äquivalenzklasse der Relation R, falls ein $x \in M$ mit $N = [x]$ existiert. Die Menge aller Äquivalenzklassen*

$$M/\sim_R := \{N \in \mathcal{P}(M) \mid \exists x \in M \quad mit \quad N = [x]\} \subset \mathcal{P}(M)$$

heißt Quotientenmenge von M bezüglich der Äquivalenzrelation $\sim_R$. Manchmal schreiben wir auch $M/\sim$ oder M/R anstelle von $M/\sim_R$

Das Auswahlaxiom garantiert, dass wir für jede Äquivalenzklasse $N \in M/\sim_R$ ein Element $x \in M$ mit $x \in N$ whalen können. Insbesondere garantiert das Auswahlaxiom, dass wir jedes Element der Quotientenmenge $M/\sim_R$ in der Form $[x]$ für ein $x \in M$ schreiben können. (Diese Tatsache mag wie eine Tautologie erscheinen, aber man braucht das Auswahlaxiom dafür. Tatsächlich ist das Auswahlaxiom sogar äquivalent zu dieser Aussage, was ein weiteres Indiz dafür ist, dass es ein natürliches Axiom ist, wenn gleich wir daran erinnern, dass es durchaus erstaunliche und vielleicht etwas merkwürdige Konsequenzen, wie etwa das Banach-Tarski Paradoxon, hat.)

Beispiel 1.

Sei M die Menge aller EU-Bürger die genau eine Staatsbürgerschaft haben. Sei $\sim_R$ die Äquivalenzrelation auf M die gegeneb ist durch $x \sim_R y :\Longleftrightarrow x$ und y haben diesselbe Staatsbürgerschaft. Dann ist $M/\sim_R$ eine endliche Menge die genau für jedes EU-Mitgliedsland ein Element besitzt. Ein solches EU-Mitgliedsland wird in $M/\sim_R$ dem Element $[x]$ entsprechen, wobei x ein beliebiger EU-Bürger ist der die Staatsbürgerschaft von besagtem land besitzt und zusätzlich keine weitere Staatsbürgerschaft innehat. (Dieses beispiel nimmt an, dass es in jedem EU-Mitgliedsland mindestens einen solchen Bürger gibt. Wenn dem nicht so ist, würde jenes Land in der Quotientenmenge $M/\sim_R$ fehlen, da es keinen Vertreter $x \in M$ gäbe, der die Staatsbürgerschaft von jenem Land innehätte.)

Beispiel 2.

Sei $M = \mathbb{Z}$ die Menge aller ganzen Zahlen. Sei weiterhin R die Äquivalenzrelation auf $\mathbb{Z}$ die gegeben ist durch $x \sim_R y :\Longleftrightarrow x - y$ ist gerade, d.h. $x - y \in 2\mathbb{Z}$. Dann gilt

$$M/\sim_R = \{[0], [1]\}.$$

Diese Menge hat also genau 2 (verschiedene) Elemente.

Beispiel 3.
Sei $M = \mathbb{Z}$ die Menge aller ganzen Zahlen. Sei weiterhin R die Äquivalenzrelation auf $\mathbb{Z}$ die gegeben ist durch $x \sim_R y :\Longleftrightarrow x - y$ ist durch 2020 teilbar, d.h. $x - y \in 2020\mathbb{Z}$. Dann gilt

$$M/\sim_R = \{[0], [1], [2], [3], \ldots, [2018], [2019]\}.$$

Diese Menge hat also genau 2020 (verschiedene) Elemente.

1.4 Abbildungen

Definition 1.4.1. *Seien M und N Mengen. Eine Abbildung $f : M \to N$ ist eine Vorschrift, die jedem Element $x \in M$ genau ein Element $f(x) \in N$ zuordnet. Das Element $f(x) \in N$ heißt das Bild von x (unter der Abbildung f).*

Definition 1.4.2. *Seien M und N Mengen und sei $f : M \to N$ eine Abbildung.*

(i) Die Menge

$$\mathrm{Bild}(f) := \{f(x) \mid x \in M\}$$

heißt Bildmenge von f.

(ii) Für eine Teilmenge $N' \subset N$ heißt

$$f^{-1}(N') := \{x \in M \mid f(x) \in N'\}$$

das Urbild von N' unter der Abbildung f. In dem Spezialfall dass $N' = \{y\}$ eine Menge mit genau einem Element y ist, schreiben wir auch $f^{-1}(y)$ anstelle von $f^{-1}(N')$.

Definition 1.4.3. *Seien M und N Mengen und sei $f : M \to N$ eine Abbildung.*

(i) f heißt injektiv, falls für alle $x, y \in M$ gilt: $f(x) = f(y) \Rightarrow x = y$;

(ii) f heißt surjektiv, falls $\mathrm{Bild}(f) = N$ gilt, d.h. für jedes $y \in N$ gibt es ein $x \in M$ mit $f(x) = y$;

(iii) f heißt bijektiv, falls f injektiv und surjektiv ist.

Beispiel.

- Sei M eine Menge, so ist die Identität

$$\mathrm{id}_M : M \longrightarrow M, \quad x \mapsto \mathrm{id}_M(x) := x$$

  eine Abbildung von M nach M. Diese Abbildung ist bijektiv.

- Sei $N \subset M$ eine Teilmenge, dann ist die Abbildung

$$N \longrightarrow M, \quad x \mapsto x$$

  injektiv. Die Abbildung ist genau dann surjektiv, wenn $M = N$ gilt.

- Die Abbildung $f : \{1, 2, 3\} \to \{A, B\}$ mit $f(1) = f(2) = A$ und $f(3) = B$ ist surjektiv, aber nicht injektiv.

- Die Abbildung $f : \{1, 2, 3\} \to \{A, B, C\}$ mit $f(1) = f(2) = A$ und $f(3) = B$ ist weder surjektiv noch injektiv.

- $f : \{1, 2, 3\} \to \{A, B, C\}$ mit $f(1) = A$ und $f(2) = C$ und $f(3) = B$ ist bijektiv.

Definition 1.4.4. *Seien M, N und L Mengen und seien $f : M \to N$ und $g : N \to L$ Abbildungen. Dann ist die Komposition $g \circ f$ von f mit g gegeben durch die Abbildung*

$$g \circ f : M \longrightarrow N, \quad x \mapsto g(f(x)).$$

Lemma 1.4.5. *Falls $f : M \to N$, $g : N \to L$ und $h : L \to K$ Abbildungen zwischen Mengen sind, so gilt $(h \circ g) \circ f = h \circ (g \circ f)$.*

Beweis. Wir müssen zeigen, dass für alle $x \in M$,

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

gilt. Das rechnet man aber einfach nach:

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

und

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

Also,

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

wie behauptet. □

Lemma 1.4.6. *Sei $f : M \to N$ eine bijektive Abbildung zwischen zwei Mengen M und N. Dann gibt es genau eine Abbildung $f^{-1} : N \to M$ mit $f \circ f^{-1} = \mathrm{id}_N$ und $f^{-1} \circ f = \mathrm{id}_M$.*

Beweis. Es kann nur eine solche Abbildung geben, weil für $y \in N$,

$$f(f^{-1}(y)) = \mathrm{id}_N(y) = y$$

gelten muss, aber nach Injektivität von f höchstens ein $x \in M$ mit $f(x) = y$ existieren kann.
Wir zeigen nun die Existenz von f. Da f bijektiv ist, gibt es für beliebiges $y \in N$ genau ein $x_y \in X$ mit $f(x_y) = y$, sodass wir $f^{-1}(y) := x_y$ definieren können. Dies liefert eine Abbildung

$$f^{-1} : N \longrightarrow M$$

und die Definition der Abbildung impliziert $f^{-1}(f(x_y)) = f^{-1}(y) = x_y$ für alle $x \in M$.
Sei umgekehrt $y \in N$. Da f surjektiv ist, gilt $y = f(x)$ für ein $x \in M$. Dann haben wir aber

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

und nachdem was wir oben gesehen haben, wissen wir dass $f^{-1}(f(x)) = x$ gilt, sodass insgesamt

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

folgt. Dies gilt für beliebiges $y \in N$, sodass $f \circ f^{-1} = \mathrm{id}_N$ gilt. □

Lemma 1.4.7. *Seien M, N und L Mengen und seien $f : M \to N$ und $g : N \to L$ Abbildungen. Dann gilt:*

(i) Falls f und g injektiv sind, so ist $g \circ f$ injektiv;

(ii) Falls f und g surjektiv sind, so ist $g \circ f$ surjektiv;

(iii) Falls $g \circ f$ surjektiv ist, so ist g surjektiv;

(iv) Falls $g \circ f$ injektiv ist, so ist f injektiv.

Beweis. Wir beweisen (i) durch Widerspruch. Dazu nehmen wir an, dass $g \circ f$ injektiv nicht injektiv ist. Dann gibt es zwei verschiedene Elemente $x, y \in M$ mit $g(f(x)) = g(f(y))$. Da g injektiv ist, folgt daraus $f(x) = f(y)$ und die Injektivität von f impliziert $x = y$. Das ist ein Widerspruch. Also ist $g \circ f$ injektiv.

Aussage (ii) kann man ähnlich durch Widerspruch zeigen. Man kann aber ebenso wie in (i) auch direkt argumentieren. Wir führen hier das direkte Argument vor. Sie dazu $z \in L$. Da g surjektiv ist, gibt es ein $y \in N$ mit $g(y) = z$. Da f surjektiv ist, gibt es ein $x \in M$ mit $f(x) = y$. Dann gilt aber

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

Da $z \in L$ beliebig war, zeigt das, dass $g \circ f$ surjektiv ist, wie behauptet.

Um (iii) zu zeigen, nehmen wir nun an, dass $g \circ f$ surjektiv ist. Sei dann $z \in L$. Dann gibt es ein $x \in M$ mit $g(f(x)) = x$. Insbesondere erfüllt das Element $y = f(x) \in N$ die Bedingung $g(y) = z$ und somit ist g surjektiv.

Um (iv) zu zeigen, nehmen wir nun an, dass $g \circ f$ injektiv ist. Seien weiterhin $x, y \in M$ mit $f(x) = f(y)$ gegeben. Dann gilt sicherlich auch $g(f(x)) = g(f(y))$ und die Injektivität von $g \circ f$ liefert $x = y$. Das beweist (iv) und schließt insgesamt den Beweis des Lemmas ab. $\square$

Bemerkung 1.4.8. *Falls $g \circ f$ bijektiv ist, so ist nach obigem Lemma f injektiv und g surjektiv. Es ist aber nicht wahr, dass f oder g bijektiv sind, wie folgendes Beispiel zeigt. Sei $f : \{1, 2\} \to \{1, 2, 3\}$ die natürliche Inklusion, und sei $g : \{1, 2, 3\} \to \{1, 2\}$ gegeben durch $g(1) = 1$, $g(2) = g(3) = 2$. Dann ist $g \circ f = \mathrm{id}$ und somit bijektiv, aber f ist nicht surjektiv und g ist nicht injektiv.*

Kapitel 2

Gruppen, Ringe und Körper

2.1 Gruppen

Definition 2.1.1. *Sei G eine Menge. Eine Verknüpfung auf G ist eine Abbildung*

$$* : G \times G \longrightarrow G, \quad (g_1, g_2) \mapsto *(g_1, g_2) := g_1 * g_2.$$

Eine Verknüpfung auf G ist also eine Vorschrift, die jedem Paar (g_1, g_2) von Elementen aus G ein weiteres Element $g_1 \circ g_2 \in G$ zuordnet. Hier ist zu beachten, dass für $g_1 \neq g_2$ auch $(g_1, g_2) \neq (g_2, g_1)$ gilt und somit $g_1 * g_2$ nicht zwingend mit $g_2 * g_1$ übereinstimmen muss. Verknüpfungen für die $g_1 * g_2 = g_2 * g_1$ für alle $g_1, g_2 \in G$ gilt, heißen kommutativ oder abelsch.

Beispiele.

- Sei M eine beliebige Menge und sei $G := \mathrm{Abb}(M, M)$ die Menge aller Abbildungen von M nach M. So liefert die Verknüpfung von Abbildungen eine Verknüpfung auf der Menge G:
$$G \times G \longrightarrow G, \quad (f, g) \mapsto f \circ g.$$
Im Allgemeinen ist dies Verknüpfung nicht abelsch.

- Sei $G = \mathbb{N}, \mathbb{Z}, \mathbb{Q}$ oder $\mathbb{R}$, so ist
$$G \times G \longrightarrow G, \quad (x, y) \mapsto x + y$$
eine abelsche Verknüpfung. Ebenso ist,
$$G \times G \longrightarrow G, \quad (x, y) \mapsto x \cdot y$$
eine abelsche Verknüpfung.

- Sei $G = \mathbb{Q}$ oder $\mathbb{R}$, so ist
$$G \times G \longrightarrow G, \quad (x, y) \mapsto (x + y)/2$$
eine abelsche Verknüpfung.

Definition 2.1.2. *Eine Menge G zusammen mit einer Verknüpfung $*$ heißt Gruppe, wenn folgende Axiome erfüllt sind:*

1. *(Assoziativgesetz) Für alle $a, b, c \in G$ gilt $a * (b * c) = (a * b) * c$.*

2. *(Existenz eines neutralen Elements und von Inversen Elementen)*
   Es gibt ein Element $e \in G$, neutrales Element genannt, mit folgender Eigenschaft:
   (i) *für alle $a \in G$ gilt: $e * a = a$;*
   (ii) *zu jedem Element $a \in G$ gibt es ein Element a', inverses Element von a genannt, so dass $a' * a = e$ gilt.*

Die Gruppe G heißt abelsch oder kommutativ, falls die Verknüpfung abelsch/kommutativ ist (d.h. falls $a * b = b * a$ für alle $a, b \in G$ gilt).

Notation 1. *Sei $(G, *)$ eine Gruppe, d.h. eine Menge G zusammen mit einer Verknüpfung $*$ die obige Eigenschaften erfüllt. Falls die Verknüpfung aus dem Kontext klar ist, schreiben wir oft auch einfach G, anstelle von $(G, *)$. Weiterhin wird die Verknüpfung $*$ oft auch als $\cdot$ geschrieben. Dann schreibt man also $a * b = a \cdot b$. Das Inverse eines Elements $a \in G$ wird dann mit a^{-1} bezeichnet. Der Einfachheit halber schreibt man manchmal auch ab anstelle von $a \cdot b$.*

Beispiele.

- Die Mengen $G = \mathbb{Z}, \mathbb{Q}$ oder $\mathbb{R}$ mit $+$ als Verknüpfung sind Gruppen.
- $(\mathbb{N}, +)$ ist keine Gruppe, weil Inverse Elemente fehlen.
- $(\mathbb{R}, \cdot)$ ist keine Gruppe, weil $0 \in \mathbb{R}$ kein Inverses Element besitzt.
- Falls M eine Menge mit mehr als einem Element ist, so ist $G = \mathrm{Abb}(M, M)$ zusammen mit der Verknüpfung $\circ$ keine Gruppe, weil i.A. Inverse Elemente fehlen.
- Die Verknüpfung $(x, y) \mapsto (x + y)/2$ auf $\mathbb{Q}$ oder $\mathbb{R}$ ist keine Gruppe, weil das Assoziativgesetz nicht gilt.

Lemma 2.1.3. *Sei G eine Gruppe, so gilt:*

(1) *Das inverse Element a^{-1} zu einem Element $a \in G$ hat auch die Eigenschaft $a \cdot a^{-1} = e$;*

(2) *Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat auch die Eigenschaft $a \cdot e = a$ für alle $a \in G$;*

(3) *Das inverse Element a^{-1} zu einem Element $a \in G$ ist eindeutig bestimmt;*

(4) *Es gilt für $a, b \in G$:*
$$(a^{-1})^{-1} = a \quad \textit{und} \quad (ab)^{-1} = b^{-1}a^{-1}.$$

(5) *Es gelten folgende Kürzungsregeln:*
$$ab = ac \quad \Rightarrow b = c \quad \textit{und} \quad ba = ca \mid \Rightarrow b = c.$$

Beweis. Sei $a \in G$ mit inversem Element $a^{-1} \in G$, d.h. $a^{-1}a = e$. Sei weiterhin $(a^{-1})^{-1}$ ein inverses Element zu a^{-1}, d.h. $(a^{-1})^{-1}a^{-1} = e$. Dann gilt

$$aa^{-1} = eaa^{-1} = ((a^{-1})^{-1}a^{-1})aa^{-1} = (a^{-1})^{-1}(a^{-1}a)a^{-1} = (a^{-1})^{-1}ea^{-1} = (a^{-1})^{-1}a^{-1} = e.$$

Dies zeigt (1).

Angenommen $e' \in G$ ist ein zweites neutrales Element von G mit $e \neq e'$. Dann gilt nach (1) $e * e' = e'$ und $e * e' = e$, also $e = e'$. Das ist ein Widerspruch und somit ist das neutrale

Element $e \in G$ eindeutig durch die Eigenschaft "$e \cdot a = a$ für alle $a \in G$"bestimmt. Weiterhin gilt nach (1):

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

Das beweist (2).

Angenommen $a' \in G$ ist ein Element mit $a'a = e$. Dann gilt nach (1) auch $aa' = e$ und somit

$$a' = a'aa^{-1} = ea^{-1} = a^{-1}.$$

Das beweist (3).

Nach (1) ist a ein inverses Element zu a^{-1} und damit $(a^{-1})^{-1} = a$. Weiterhin gilt

$$b^{-1}a^{-1}ab = b^{-1}eb = b^{-1}b = e$$

und somit ist $b^{-1}a^{-1}$ ein inverses Element zu $ab \in G$. Das beweist (4).

Seien schließlich $a, b, c \in G$ Elemente. Falls $ab = ac$ gilt, dann gilt $a^{-1}(ab) = a^{-1}(ac)$ und damit auch $a^{-1}ab = a^{-1}ac$ sowie $b = c$, da $a^{-1}a = e$. Falls $ba = ca$ gilt, so gilt auch $(ba)a^{-1} = (ca)a^{-1}$ und damit auch $baa^{-1} = caa^{-1}$ und somit $b = c$, da $a^{-1}a = e$. Dies beweist (5) und somit ist der Beweis des Lemmas abgeschlossen. □

Definition 2.1.4. *Sei G eine Gruppe mit neutralem Element e. Eine Teilmenge $H \subset G$ heißt Untergruppe, falls $e \in H$ und falls für alle $a, b \in H$ auch $ab \in H$ und $a^{-1} \in H$ gilt.*

Die Bedingung $ab \in H$ für alle $a, b \in H$ impliziert dass die Einschränkug der Verknüpfung auf G eine Verknüpfung

$$H \times H \longrightarrow H, \quad (ab) \mapsto ab$$

liefert. Da $e \in H$ und $a^{-1} \in H$ für alle $a \in H$, ist H mit obiger Verknüpfung wieder eine Gruppe.

Beispiele.

- $(\mathbb{Z}, +) \subset (\mathbb{Q}, +)$ ist eine Untergruppe;
- $(\mathbb{Q}^*, \cdot) \subset (\mathbb{R}^*, \cdot)$ ist eine Untergruppe.

Definition 2.1.5. *Seien G und H Gruppen. Ein Gruppenhomomorphismus ist eine Abbildung $f : G \longrightarrow H$ sodass*

$$f(ab) = f(a)f(b) \in H$$

für alle $a, b \in G$ gilt.

Hier bezeichnet $ab \in G$ die Verknüpfung von a und b in G, wobei $f(a)f(b)$ die Verknüpfung von $f(a)$ und $f(b)$ in H bezeichne.

Lemma 2.1.6. *Sei $f : G \to H$ ein Gruppenhomomorphismus. Dann gilt:*

- *$f(e_G) = e_H$, wobei $e_G \in G$ und $e_H \in H$ die neutralen Elemente bezeichne.*
- *$f(a)^{-1} = f(a^{-1})$ für alle $a \in G$.*

Beweis. Wendet man die Gesetzmäßigkeit $f(ab) = f(a)f(b)$ auf $a = b = e_G$ an, so erhalten wir

$$f(e_G) = f(e_G)f(e_G) \in H.$$

Da H eine Gruppe ist, gilt nach Lemma 2.1.3 die Kürzungsregel, sodass die Identität $f(e_G) = f(e_G)f(e_G)$ schon $f(e_G) = e_H$ impliziert.

Sei nun $a \in G$. Dann gilt

$$f(a^{-1})f(a) = f(a^{-1}a) = f(e_G) = e_H.$$

Also,

$$f(a)^{-1} = f(a^{-1}),$$

was zu beweisen war. □

Beispiele.

- Für jede natürliche Zahl n ist $f : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto nx$ ein Gruppenhomomorphismus von $(\mathbb{Z}, +)$.
- Falls $H \subset G$ eine Untergruppe ist, so ist

  $$f : H \longrightarrow G, a \mapsto a$$

  ein injektiver Gruppenhomomorphismus.

Definition 2.1.7. *Seien G und H Gruppen. Ein Isomorphismus zwischen G und H ist eine Gruppenhomomorphismus $f : G \longrightarrow H$ welcher injektiv und surjektiv, also bijektiv ist.*

Lemma 2.1.8. *Sei $f : G \longrightarrow H$ ein Gruppenhomomorphismus. Dann ist f genau dann injektiv, falls $f^{-1}(e_H) = e_G$, wobei $e_H \in H$ und $e_G \in G$ die neutralen Elemente bezeichne.*

Beweis. Wir haben oben gesehen, dass $f(e_G) = e_H$ gelten muss. Falls f injektiv ist, so muss also $f^{-1}(e_H) = e_G$ gelten.
Nehmen wir umgekehrt $f^{-1}(e_H) = e_G$ and. Falls $a, b \in G$ mit $f(a) = f(b)$, so gilt

$$e_H = f(a)f(b)^{-1} = f(a)f(b^{-1}) = f(ab^{-1}).$$

Nach Annahme also $ab^{-1} = e_G$ und damit $a = b$. Dies beweist das Lemma. □

2.2 Ringe

Definition 2.2.1. *Eine Menge R zusammen mit zwei Verknüpfungen*

$$+ : R \times R \longrightarrow R, \quad (x, y) \mapsto x + y$$

und

$$\cdot : R \times R \longrightarrow R, \quad (x, y) \mapsto x \cdot y,$$

heißt Ring, falls folgendes gilt:

(1) $(R, +)$ ist eine abelsche Gruppe;

(2) Die Multiplikation ist assoziativ, d.h. für alle $x, y, z \in R$ gilt

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

(3) Das Distributivgesetz gilt. D.h. für alle $a, b, c \in R$ gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c$$

(wobei man wie üblich 'Punkt vor Strich' anwendet und in obigen Termen zuerst die Multiplikation und dann die Addition ausführt).

(4) Es gibt ein neutrales Element bzgl der Multiplikation (Eins Element genannt), d.h. ein Element $1 \in R$ mit $1 \cdot a = a$ für alle $a \in R$.

Beispiele

- Die Menge $\mathbb{Z}$ der ganzen Zahlen mit der üblichen Addition und Multiplikation ist ein Ring.
- Die Menge $\mathbb{Q}$ oder $\mathbb{R}$ mit der üblichen Addition und Multiplikation ist ein Ring.

Lemma 2.2.2. *Sei R ein Ring. Falls $0 \in R$ das neutrale Element der Addition bezeichnet, so gilt für alle $a \in R$:*

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

Beweis. Es gilt

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

Nach der Kürzungsregel in Gruppen gilt also

$$0 = a \cdot 0.$$

Ebenso haben wir

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

und somit $0 = 0 \cdot a$. Dies beweist das Lemma. □

2.2.1 Restklassenring

Sei $n \geq 1$ eine natürliche Zahl. Betrachte die Menge $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$ aller ganzen Zahlen die durch n teilbar sind. Wir betrachten dann die Relation $\sim$ auf $\mathbb{Z}$, die gegeben ist durch

$$x \sim y \quad :\Longleftrightarrow \quad x - y \in n\mathbb{Z}.$$

Auf Übungsblatt 3 zeigen Sie, dass dies eine Äquivalenzrelation ist. Die Menge der Äquivalenzklassen $\mathbb{Z}/\sim$ wird mit $\mathbb{Z}/n\mathbb{Z}$ bezeichnet. Die Äquivalenzklasse einer ganzen Zahl $a \in \mathbb{Z}$ wird mit $[a] \in \mathbb{Z}/n\mathbb{Z}$ bezeichnet. Konkret ist $[a] \subset \mathbb{Z}$ die Teilmenge aller ganzen Zahlen z mit $z - a \in n\mathbb{Z}$, also aller ganzen Zahlen z die von der Form $z = a + kn$ für ein $k \in \mathbb{Z}$ sind.

Lemma 2.2.3. *Seien $[a], [b] \in \mathbb{Z}/n\mathbb{Z}$. Die Vorschriften*

$$[a] + [b] := [a + b] \quad \text{und} \quad [a] \cdot [b] := [a \cdot b]$$

sind wohldefiniert, d.h. die Äquivalenzklassen $[a + b]$ und $[a \cdot b]$ hängen nur von den Klassen $[a]$ und $[b]$ ab, nicht aber von den Repräsentanten a. Die Menge $\mathbb{Z}/n\mathbb{Z}$ zusammen mit der oben definierten Addition und Multiplikation bildet einen kommutativen Ring mit genau n Elementen.

2.2.2 Polynomring

Sei R ein Ring (zB. $R = \mathbb{Z}, \mathbb{Q}$ oder $R = \mathbb{R}$). Ein Polynom über R, ist eine Abbildung

$$\mathbb{N} = \{0, 1, 2, \ldots\} \longrightarrow R, \quad i \mapsto a_i$$

sodass ein $n \in \mathbb{N}$ mit $a_i = 0$ für $i \geq n + 1$ existiert. Üblicherweise schreibt man ein solches Polynom in der Gestalt

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

wobei x eine formale Variable bezeichnet. Der Grad eines Polynoms ist die kleinste natürliche Zahl $d \in \mathbb{N}$ sodass $a_i = 0$ für $i \geq d + 1$ gilt.

Definition 2.2.4. *Sei R ein Ring. Die Menge aller Polynome (von beliebigem Grad) über R bezeichnet man als $R[x]$:*

$$R[x] = \left\{ \sum_{i=0}^{n} a_i x^i \mid n \geq 0, \quad a_0, \ldots, a_n \in R \right\}.$$

Auf der Menge $R[x]$ aller Polynome über R definiert man eine Addition wie folgt:

$$+ : R[x] \times R[x] \longrightarrow R[x], \quad \left(\sum_{i=0}^{n} a_i x^i, \sum_{j=0}^{m} b_i x^i \right) \mapsto \sum_{i \geq 0} (a_i + b_i)\, x^i$$

und

$$\cdot : R[x] \times R[x] \longrightarrow R[x], \quad \left(\sum_{i=0}^{n} a_i x^i, \sum_{j=0}^{m} b_i x^i \right) \mapsto \sum_{i \geq 0} \left(\sum_{j=0}^{i} a_j b_{i-j} \right) x^i$$

wobei in obiger Notation, $a_i = 0$ für $i > n$ und $b_i = 0$ für $i > m$ gilt und somit die entsprechenden Summen endlich sind.

Lemma 2.2.5. *Sei R ein Ring, dann ist die Menge $R[x]$ aller Polynome über R wieder ein Ring. (Wir nennen diesen Ring den Polynomring über R.)*

Beweis. Ein Polynom heißt konstant, falls es Grad Null hat, d.h. falls es von der Form a_0 mit $a_0 \in R$ ist. Insbesondere liefern die konstanten Polynome also eine injektive Abbildung $R \to R[x]$. Das Bild des neutralen Elements $0 \in R$ der Addition ist dann auch ein neutrales Element der Addition in $R[x]$ und man überprüft leicht dass $(R[x], +)$ eine abelsche Gruppe ist. Entscheidend ist hier, dass $(R, +)$ eine abelsche Gruppe ist, da R ein Ring ist.

Sei $1 \in R$ das neutrale Element der Multiplikation auf R. Das konstante Polynom der Form 1 ist dann ein neutrales Element der Multiplikation und man überprüft leicht, dass die Multiplikation assoziativ ist und das Distributivgesetz gilt. Entscheidend ist jeweils, dass die entsprechenden Eigenschaften für R gelten. □

Definition 2.2.6. *Seien R und S Ringe. Ein Ringhomomorphismus ist eine Abbildung $f : R \to S$ sodass folgendes gilt:*

- *$f(0) = 0$, d.h. f bildet das neutrale Element bzgl der Addition von R auf das von S ab;*
- *$f(1) = 1$, d.h. f bildet das neutrale Element bzgl der Multiplikation von R auf das von S ab;*
- *für alle $a, b \in R$ gilt: $f(a + b) = f(a) + f(b)$ und $f(ab) = f(a)f(b)$.*

2.3 Körper

Definition 2.3.1. *Eine Menge K zusammen mit zwei Verknüpfungen*

$$+ : K \times K \longrightarrow K, \quad (a, b) \mapsto a + b$$

und

$$\cdot : K \times K \longrightarrow K, \quad (a, b) \mapsto a \cdot b$$

heißt Körper, falls K ein kommutativer Ring ist und zusätzlich folgendes gilt, wobei $K^* := K \setminus \{0\}$:

- *für alle* $a, b \in K^*$ *gilt* $ab \in K^*$;
- $(K^*, \cdot)$ *ist eine Gruppe mit neutralem Element* $1 \in K^*$ *(insbesondere gilt also* $1 \neq 0$*).*

Beispiele

- $\mathbb{Z}$ zusammen mit der üblichen Addition und Multiplikation ist ein Ring aber kein Körper, weil nicht jedes Element in $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ ein Inverses besitzt;
- $\mathbb{Q}$ ist ein Körper;
- $\mathbb{R}$ ist ein Körper.

Lemma 2.3.2. *Sei* K *ein Körper, so gilt für alle* $a, b, x, y \in K$:

(1) $0 \cdot a = a \cdot 0 = 0$;

(2) $a \cdot b = 0 \Rightarrow a = 0$ *oder* $b = 0$;

(3) $a(-b) = -(ab)$ *und* $(-a)(-b) = ab$;

(4) falls $x \cdot a = y \cdot a$ *und* $a \neq 0$ *gilt, so gilt schon* $x = y$.

Beweis. Die Behauptung $0 \cdot a = a \cdot 0 = 0$ in (1) haben wir bereits für beliebige Ringe nachgeprüft.

Sei nun $a \cdot b = 0$. Falls $a \neq 0$, so gilt $a^{-1}ab = a^{-1}0 = 0$ und damit $b = 0$. Also gilt $a = 0$ oder $b = 0$. Das zeigt (2).

Die erste Behauptung in (3) folgt aus

$$ab + (-a)b = (a - a)b = 0b = 0.$$

Die zweite folgt aus der ersten und der Tatsache, dass die Multiplikatoin kommutativ ist:

$$(-a)(-b) = -(-a)b = -b(-a) = -(-ba) = ba = ab.$$

Sei schließlich $x \cdot a = y \cdot a$ und $a \neq 0$. Dann gibt es ein multiplikatives Inverses $a^{-1} \in K^*$. Aus $x \cdot a = y \cdot a$ folgt dann

$$xaa^{-1} = yaa^{-1}$$

und wegen $aa^{-1} = 1$ folgt daraus $x = y$, wie behauptet. Dies beweist das Lemma. □

Definition 2.3.3. *Sei* K *ein Körper. Sei* $1 \in K^*$ *das neutrale Element der Multiplikation. Für eine natürliche zahl* $n \geq 1$*, betrachten wir das Element*

$$n \cdot 1 = 1 + 1 + 1 + \cdots + 1 \in K.$$

Falls $n \cdot 1 \neq 0 \in K$ *für alle* $n \geq 1$ *gilt, so sagen wir, dass* K *Charakterisitk* 0 *hat. Andernfalls sagen wir, dass* K *Charakteristik* n *hat, falls* n *die kleinste positive ganze Zahl mit* $n \cdot 1 = 0$ *ist.*

Lemma 2.3.4. *Sei* K *ein Körper. Die Charakteristik von* K *ist entweder* 0 *oder eine Primzahl.*

Beweis. Andenommen die Charakteristik von K ist $n > 0$. Dann gilt also $n \cdot 1 = 0$ und n ist die kleinste natürliche Zahl mit dieser Eigenschaft. Wir müssen zeigen, dass n eine Primzahl ist. Wir zeigen das per Widerspruch und nehmen an, dass n keine Primzahl ist. Dann gilt $n = ab$ für ganze Zahlen $1 < a, b < n$. Es gilt dann

$$0 = n \cdot 1 = (ab) \cdot 1.$$

Andererseits gilt

$$(a \cdot 1) \cdot (b \cdot 1) = \underbrace{(1 + 1 + \cdots + 1)}_{a\text{-mal}} \cdot \underbrace{(1 + 1 + \cdots + 1)}_{b\text{-mal}}$$

sodass das Distribiutivgesetz

$$(a \cdot 1) \cdot (b \cdot 1) = (ab) \cdot 1$$

impliziert. Also insgesamt

$$(a \cdot 1) \cdot (b \cdot 1) = (ab) \cdot 1 = 0.$$

Da K ein Körper ist, gilt nach Lemma 2.3.2 schon $a \cdot 1 = 0$ oder $b \cdot 1 = 0$. Dies widerspricht der Minimalität von n, womit das Lemma bewiesen ist. □

2.3.1 Körper mit endlich vielen Elementen

Proposition 2.3.5. *Sei $n \geq 1$ eine ganze Zahl. Der Restklassenring $\mathbb{Z}/n\mathbb{Z}$ ist ein Körper genau dann wenn $n = p$ eine Primzahl ist.*

Beweis. Wir wissen bereits (siehe Lemma 2.2.3), dass $\mathbb{Z}/n\mathbb{Z}$ ein kommutativer Ring ist. Es bleibt also zu untersuchen, wann die Multiplikation eine Verknüpfung auf $\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}$ induziert welche $\mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}$ zu einer abelschen Gruppe macht.

Falls $n = 1$ gilt, so hat $\mathbb{Z}/n\mathbb{Z}$ nur ein Element und kann damit kein Körper sein (da $0 \neq 1$ in jedem Körper gilt). Falls $n \geq 2$ keine Primzahl ist, so gilt $n = ab$ für positive ganze Zahlen a, b mit $1 < a, b < n$. Dann gilt aber $[a] \neq [0]$ und $[b] \neq [0]$ aber

$$[a][b] = [ab] = [0].$$

Da $[0]$ das neutrale Element der Addition von $\mathbb{Z}/n\mathbb{Z}$ ist und $[a] \neq [0]$ und $[b] \neq [0]$ gilt, widerspricht obige Schlussfolgerung vorherigem Lemma. Also kann $\mathbb{Z}/n\mathbb{Z}$ kein Körper sein. (Alternativ hätten wir auch so argumentieren können: Wäre $\mathbb{Z}/n\mathbb{Z}$ ein Körper, so ist seine Charakteristik gleich n und nach obigem Lemma muss dann n eine Primzahl sein.)

Sei umgekehrt $n = p$ eine Primzahl. Sei $[a] \in \mathbb{Z}/p\mathbb{Z}$ mit $[a] \neq [0]$. Das bedeutet, dass $a \in \mathbb{Z}$ eine ganze Zahl ist, die nicht durch p teilbar ist. Wir betrachten dann die Abbildung

$$\mathbb{Z}/p\mathbb{Z} \longrightarrow \mathbb{Z}/p\mathbb{Z}, \quad [x] \mapsto [a][x] = [ax].$$

Wir behaupten, dass diese Abbildung injektiv ist. In der Tat, falls $x, y \in \mathbb{Z}$ sodass

$$[ax] = [ay]$$

gilt, so ist $ax - ay = a(x - y)$ durch p teilbar. Da p eine Primzahl ist, und a nicht durch p teilbar ist, muss $x - y$ durch p teilbar sein. Also gilt $[x] = [y]$ und somit ist obige Abbildung injektiv. Dies impliziert insbesondere, dass für $[a], [b] \in \mathbb{Z}/n\mathbb{Z}$ mit $[a] \neq [0]$ und $[b] \neq [0]$

Da $\mathbb{Z}/p\mathbb{Z}$ eine endliche Menge ist, ist jede injektive Abbildung $\mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ schon surjektiv. Also ist obige Abbildung

$$\mathbb{Z}/p\mathbb{Z} \longrightarrow \mathbb{Z}/p\mathbb{Z}, \quad [x] \mapsto [a][x] = [ax]$$

surjektiv und somit muss es ein Element $a' \in \mathbb{Z}$ mit $[a][a'] = [1]$ geben. Das Element $[a'] \in \mathbb{Z}/p\mathbb{Z}$ ist also ein inverses Element zu $[a]$ bezüglich der Multiplikation auf

$$(\mathbb{Z}/p\mathbb{Z})^* = \mathbb{Z}/p\mathbb{Z} \setminus \{[0]\}.$$

Damit ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper. Dies schließt unseren Beweis ab. □

2.3.2 Der Körper der komplexen Zahlen

Betrachte die Menge $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ von paaren (a, b) reeller Zahlen zusammen mit den folgenden beiden Verknüpfungen:

$$+ : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}, \quad ((a, b), (a', b')) \mapsto (a + a', b + b')$$

und

$$\cdot : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}, \quad ((a, b), (a', b')) \mapsto \left(aa' - bb', ab' + ba'\right).$$

Um die Schreibweise zu vereinfachen, schreiben wir das Element $0, 1 \in \mathbb{C}$ auch al i. Mit dieser Schreibweise und mit Hilfe obiger Verknüpfung gilt dann

$$(a, b) = a + ib,$$

also lässt sich jede komplexe Zahl eindeutig als $a + ib$ mit reellen Zahlen a, b schreiben. Weiterhin gilt

$$(0, 1) \cdot (0, 1) = (-1, 0)$$

und somit

$$i \cdot i = -1.$$

Allgemeiner lassen sich die Verknüpfungen $+$ und $\cdot$ auf den komplexen Zahlen wie folgt schreiben:

$$(a + ib) + (a' + ib') = (a + a') + i(b + b')$$

und

$$(a + ib) \cdot (a' + ib') = aa' - bb' + i(ab' + a'b).$$

Lemma 2.3.6. *Die Menge $\mathbb{C}$ zusammen mit den oben eingeführten Verknüpfungen bildet einen Körper. Das neutrale Element der Addition ist $0 = (0, 0)$ und das neutrale Element der Multiplikation ist $1 = (1, 0)$.*

Beweis. Dies rechnet man einfach direkt nach, wobei man die üblichen Rechenregeln für die reellen Zahlen ausnutzt. $\square$

Die komplexen Zahlen enthalten die reellen Zahlen und somit auch die rationalen Zahlen als Unterkörper (also Teilmengen sodass die Einschränkung der Operationen auf $\mathbb{C}$ die üblichen Körperstrukturen auf $\mathbb{Q}$ und $\mathbb{R}$ liefert. Gauß hat die komplexen Zahlen im 18 Jahundert eingeführt. Seine Motivation bestand darin Nullstellen von Polynomen zu berchnen. Genauer ist aus der Schule bekannt, dass nicht jedes Polynom über $\mathbb{Q}$ oder $\mathbb{R}$ eine Nullstelle besitzt. Ein einfaches Beispiel ist das Polynom $x^2 + 1$, welches weder in $/Q$ noch in $\mathbb{R}$ eine Nullstelle besitzt. Ein komplizierteres Beispiel wäre $x^{2020} + x^2 + 1$. Auch dieses Polynom hat keine Nullstelle in $\mathbb{Q}$ oder $\mathbb{R}$, da der Ausdruck $x^{2020} + x^2 + 1$ für jedes $x \in \mathbb{R}$ immer positiv ist.

Das Polynom $x^2 + 1$ hat zwar keine Nullstelle in $\mathbb{R}$, aber es hat eine Nullstelle in $\mathbb{C}$, denn es gilt:

$$i^2 + 1 = -1 + 1 = 0.$$

Gauß hat bewiesen, dass tatsächlich jedes Polynom über $\mathbb{C}$ eine Nullstelle in $\mathbb{C}$ hat. Genauer gilt folgendes:

Satz 2.3.1 (Funamentalsatz der Algebra)**.** *Sei $f \in \mathbb{C}[x]$ und schreibe $f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ mit $a_n \neq 0$. Falls $n \geq 1$, so gibt es komplexe Zahlen $z_1, \ldots, z_n \in \mathbb{C}$ (Nullstellen von f genannt) mit*

$$f = a_n \cdot (x - z_1)(x - z_2) \cdots (x - z_n).$$

Der Beweis von obigem Satz sprengt den Rahmen dieser Vorlesung. Ein eleganter Beweis kann zum Beispiel mit Hilfe von Funktionentheorie (der Theorie von holomorphen Funktionen auf $\mathbb{C}$) gegeben werden.

Bemerkung 2.3.7. *In obigem Satz sind die Nullstellen $z_1, \ldots, z_n$ von f natürlich nicht notwendigerweise verschieden.*

Kapitel 3

Vektorräume

3.1 Grundlegende Definitionen und Eigenschaften

3.1.1 Definition und Beispiele von Vektorräumen

Definition 3.1.1. *Sei K ein Körper. Eine Menge V zusammen mit einer Verknüpfung*

$$+ : V \times V \longrightarrow V, \quad (v_1, v_2) \mapsto v_1 + v_2$$

(Addition genannt) und einer äußeren Verknüpfung[1]

$$\cdot : K \times V \longrightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v$$

(Skalarmultiplikation genannt) heißt Vektorraum über K (oder K-Vektorraum, oder einfach Vektorraum), wenn folgendes gilt:

(i) *$(V, +)$ ist eine abelsche Gruppe (das neutrale Element heißt Nullvektor und wird mit $0 \in V$ bezeichnet, das Inverse eines Elements $v \in V$ bzgl. $+$ wird mit $-v$ bezeichnet);*

(ii) *für alle $v, w \in V$ und $\lambda, \mu \in K$ gilt folgendes (wobei wir immer 'Punkt vor Strich' in unserer Notation anwenden):*

- $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$;
- $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$;
- $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$;
- $1 \cdot v = v$, *wobei $1 \in K^*$ das Einselement, also das neutrale Element von $(K^*, \cdot)$ bezeichne.*

Bemerkung 3.1.2. *Streng genommen müssten wir die Addition auf V in unserer Notation von der Addition K unterscheiden, z.B. indem wir für die Vektoraddition $\dot{+}$ und für die Addition in K einfach $+$ schreiben. Der Einfachheit verzichten wir auf diese Unterscheidung; die Frage ob ein gegebenes $+$-Zeichen die Addition in K oder die Addition in V bezeichne, lässt sich immer daran ablesen, ob die Elemente die damit Verknüpft werden in K oder in V liegen.*

Man beachte weiterhin, dass wir den Nullvektor, also das neutrale Element von $(V, +)$ mit demselben Symbol wie das neutrale Element $0 \in K$ von $(K, +)$ bezeichnen. Hier sollte wieder

[1] Streng genommen sind Verknüpfungen auf einer Menge Abbildungen $M \times M \to M$. Die vorliegende 'äußere Verknüpfung' ist eine Abbildung $K \times V \to V$ und da $K \neq V$ ist es keine Verknüpfung im strengen Sinn, sodass wir das Wort 'innere Verknüpfung' verwenden. In diesem Kontext wird die Verknüpfung $+$ auf V auch oft 'äußere Verknüpfung' genannt.

aus der jeweiligen Situation klar sein, welches der beiden Elemente gemeint ist. Da $0+0=0$ *sowohl in* V *als auch in* K *gilt, ist diese vereinfachte Schreibweise nicht in Konflikt mit oben genannter vereinfachter Schreibweise.*

Bemerkung 3.1.3. *Ein ähnliches Problem tritt bei der Skalarmultiplikation und der gewöhnlichen Multiplikation in* K *auf. Beide Verknüpfungen werden mit demselben Zeichen notiert. Wieder erschließt sich aus der Tatsache ob zwei Elemente in* K *miteinander oder ein Element aus* K *mit einem Element aus* V *verknüpft wird, welche der beiden Verknüpfungen gemeint ist.*

Beispiel. Sei K ein Körper und sei $n \geq 1$ eine positive natürliche Zahl. Betrachte dann die Menge

$$V := K^n := \underbrace{K \times K \times \cdots \times K}_{\text{n-Mal}}$$

von n-Tupel $v = (v_1, v_2, \ldots, v_n)$ von Elementen $v_i \in K$. Solche Tupel werden (Zeilen-) Vektoren genannt. Seien nun $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$ zwei Vektoren im K^n. Wie in Kapitel 0.1 können wir dann den Vektor

$$v + w := (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \in K^n$$

betrachten. Diese Verknüpfung liefert eine Vektoraddition

$$+ : K^n \times K^n \longrightarrow, \quad (v, w) \mapsto v + w.$$

Für ein Element $\lambda \in K$ und einen Vektor $v = (v_1, v_2, \ldots, v_n) \in K^n$, definieren wir

$$\lambda \cdot (v_1, v_2, \ldots, v_n) := (\lambda \cdot v_1, \lambda \cdot v_2, \ldots, \lambda \cdot v_n).$$

Dies liefert eine Skalarmultiplikation

$$K \times K^n \longrightarrow K^n.$$

Lemma 3.1.4. *Sei* K *ein Körper und sei* $n \geq 1$*. Die Menge* $V = K^n$ *zusammen mit der oben definierten Vektoraddition und Skalarmultiplikation ist ein* K*-Vektorraum.*

Beweis. Dies ist eine einfach Rechnung, die wir dem Leser überlassen. Der Nullevektor ist zum Beispiel gegeben durch $0 = (0, 0, 0, \ldots, 0)$ und für $v = (v_1, v_2, \ldots, v_n) \in K^n$ ist das Inverse bezüglich Vektoraddition gegeben durch $-v = (-v_1, -v_2, \ldots, -v_n)$. □

Bemerkung 3.1.5. *Sei* $V := \{0\}$ *die abelsche Gruppe mit nur einem Element. Dann ist* V *ein* K*-Vektorraum für jeden Körper* K*. Dieser Vektorraum heißt Nullraum.*

Per Konvention ist der Vektorraum K^0 *der Nullraum über* K*.*

Beispiel. Sei K ein Körper und sei M eine beliebige Menge. Sei $V := \mathrm{Abb}(M, K)$ die Menge aller Abbildungen $f : M \to K$. Wir definieren eine Vektoraddition auf V wie folgt:

$$+ : V \times V \longrightarrow V, \quad (f, g) \mapsto f + g,$$

wobei $f + g : M \to K$ definiert ist durch:

$$(f + g)(x) := f(x) + g(x) \ \text{ für alle } x \in M.$$

Weiterhin definieren wir eine Skalarmultiplikation

$$\cdot : K \times V \longrightarrow V, \quad (\lambda, f) \mapsto \lambda \cdot f,$$

wobei $\lambda \cdot f : M \to K$ definiert ist durch:

$$(\lambda \cdot f)(x) := \lambda \cdot f(x) \ \text{ für alle } x \in M.$$

Lemma 3.1.6. *Die Menge $V = \mathrm{Abb}(M, K)$ zusammen mit der oben definierten Vektoraddition und Skalarmultiplikation ist ein Vektorraum.*

Beweis. Die Vektorraum Axiome folgen durch eine einfache Rechnung aus der Tatsache dass K ein Körper ist. Wir lassen die Detail den LeserInnen zur Übung. □

Lemma 3.1.7. *Sei K ein Körper und sei V ein K-Vektorraum. Dann gilt:*

(i) $0 \cdot v = 0$*;*

(ii) $\lambda \cdot 0 = 0$ *für alle* $\lambda \in K$*;*

(iii) $\lambda \cdot v = 0 \Longrightarrow \lambda = 0$ *oder* $v = 0$*;*

(iv) $(-1) \cdot v = -v$*.*

Beweis. Es gilt:

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

und damit $0 \cdot v = 0$, weil $(V, +)$ eine abelsche Gruppe ist.. Das zeigt (i).
Weiterhin gilt

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

und somit $\lambda \cdot 0 = 0$, weil $(V, +)$ eine abelsche Gruppe ist. Dies zeigt (ii).
Angenommen $\lambda \cdot v = 0$ aber $\lambda \neq 0$. Da K ein Körper ist, gibt es dann ein multiplikatives Inverses $\lambda^{-1} \in K^*$ und somit

$$0 = \lambda^{-1} \cdot 0 = \lambda^{-1} \cdot (\lambda \cdot v) = (\lambda^{-1}\lambda) \cdot v = 1 \cdot v = v.$$

Das zeigt (iii).
Schließlich gilt

$$(-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = (-1 + 1) \cdot v = 0 \cdot v = 0$$

und damit $(-1) \cdot v = -v$, wie behauptet. Das beweist das Lemma. □

3.1.2 Untervektorräume

Definition 3.1.8. *Sei K ein Körper und sei V ein K-Vektorraum. Eine Teilmenge $U \subset V$ heißt Untervektorraum, falls folgendes gilt:*

(a) $U \neq \emptyset$*;*

(b) $v, w \in U \Rightarrow v + w \in U$*; (man sagt: U ist bezüglich Vektoraddition abgeschlossen);*

(c) $v \in U$, $\lambda \in K \Rightarrow \lambda v \in W$*; (man sagt: U ist bezüglich Skalarmultiplikation abgeschlossen).*

Beispiel 1. Sei $V = \mathbb{R}^2$.

- $U_1 = \{0\}$ ist ein Untervektorraum;
- $U_2 = \{(v_1, v_2) \in \mathbb{R}^2 \mid a_1 v_1 + a_2 v_2 = b\}$ ist ein Untervektorraum genau dann wenn $b = 0$;
- $U_3 = \{(v_1, v_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$ ist kein Untervektorraum, weil U_3 weder bezüglich Vektoraddition, noch bezüglich Skalarmultiplikation abgeschlossen ist;

- $U_4 = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 \geq 0 \quad \text{und} \quad v_2 \geq 0\}$ ist kein Untervektorraum: U_3 ist zwar bezüglich Vektoraddition abgeschlossen, aber nicht bezüglich Skalarmultiplikation.

Beispiel 2. Sei $V = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ der Vektorraum aller Abbildungen $f : \mathbb{R} \to \mathbb{R}$.

- Die Teilmenge $U_1 \subset V$ aller stetigen Abbildungen ist ein Untervektorraum, weil die Summe zweier stetigen Abbildungen sowie das Produkt einer stetigen Abbildung mit eine Skalar wieder stetig sind.
- Die Teilmenge $U_2 \subset V$ aller differenzierbaren Abbildungen ist ein Untervektorraum, weil die Summe zweier differenzierbarer Abbildungen sowie das Produkt einer stetigen Abbildung mit eine Skalar wieder differenzierbar sind.

Beispiel 3. Sei K ein Körper und sei $V := K^n$. Seien $a_{ij} \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$. Betrachte die Teilmenge $U \subset V$ die gegeben ist durch alle Vektoren $v = (v_1, \ldots, v_n) \in V$ sodass folgendes gilt:

$$\begin{aligned} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n &= 0 \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n &= 0 \\ &\vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n &= 0. \end{aligned}$$

Da $0 = (0, 0, \ldots, 0) \in U$ gilt, ist $U \neq \emptyset$. Weiterhin sieht man einfach, dass aus $v, w \in U$ auch $v + w \in U$ folgt. Und für $v \in U$ und $\lambda \in K$ gilt $\lambda v \in U$. Also ist $U \subset V$ ein Untervektorraum.

Lemma 3.1.9. *Sei V ein K-Vektorraum und sei $U \subset V$ ein Untervektorraum. Dann ist U mit der induzierten Vektoraddition und Skalarmultiplikation wieder ein K-Vektorraum.*

Beweis. Da $U \neq \emptyset$, gibt es ein Element $v \in U$. Dann ist auch $-v \in U$ uns damit auch $0 = v - v \in U$. Also gilt $0 \in U$.

Für jedes $v \in U$ gilt $-v \in U$ und für $v, w \in U$ gilt $v + w \in U$. Damit ist $U \subset V$ eine Untergruppe bzgl $+$ und damit selbst eine abelsche Gruppe, da $(V, +)$ eine abelsche Gruppe ist.

Bedingung (ii) in Definition 3.1.1 ist klar, weil diese Bedingung für alle $v, w \in V$ und $\lambda, \mu \in K$ gilt und damit insbesondere auch für alle $v, w \in U$ und $\lambda, \mu \in K$. Das beweist das Lemma. $\square$

3.1.3 Schnitt von Untervektorräumen

Lemma 3.1.10. *Sei K ein Körper und sei V ein K-Vektorraum. Für eine beliebige Indexmenge I und für jedes Element $i \in I$ sei ein Untervektorraum $U_i \subset V$ gegeben. Dann ist der Schnitt*

$$U := \bigcap_{i \in I} U_i$$

ein Untervektorraum von V.

Beweis. Sei $0 \in V$ der Nullvektor. Dann gilt $0 \in U_i$ für alle $i \in I$ und damit $0 \in U$. Also ist U nicht leer. Seien nun $v, w \in U$. Dann gilt $v, w \in U_i$ für alle $i \in I$. Da U_i ein Untervektorraum von V ist, gilt damit auch $v + w \in U_i$ für alle i. Also $v + w \in U$ und damit ist U bezüglich Vektoraddition abgeschlossen.

Sei nun $\lambda \in K$ und $v \in U$. Dann gilt $v \in U_i$ für alle i. Da U_i ein Untervektorraum von V ist, gilt dann auch $\lambda \cdot v \in U_i$ für alle i. Also $\lambda \cdot v \in U$ und damit ist U bezüglich Skalarmultiplikation abgeschlossen. Dies zeigt, dass U ein Untervektorraum von V ist. $\square$

Bemerkung 3.1.11. *Die Vereinigung von Unterverktorräumen ist im Allgemeinen kein Untervektorraum. Genauer gilt: falls $U_1, U_2 \subset V$ Untervektorräume eines K-Vektorraums V sind, so ist $U_1 \cup U_2$ genau dann ein Untervektorraum wenn $U_1 \subset U_2$ oder $U_2 \subset U_1$ gilt.*

Beweis. Angenommen $U := U_1 \cup U_2$ ist ein Untervektorraum von V. Nehmen wir weiterhin an, dass U_1 keine Teilmenge von U_2 ist. Dann gibt es ein $u_1 \in U_1$ mit $u_1 \notin U_2$. Sei nun $u_2 \in U_2$ beliebig. Da U ein Untervektorraum ist, muss dann $u_1 + u_2 \in U$ gelten. Also entweder $u_1 + u_2 \in U_1$ oder $u_1 + u_2 \in U_2$.

Fall 1. $u_1 + u_2 \in U_2$

Da $u_2 \in U_2$ gilt auch $-u_2 \in U_2$ und damit

$$u_1 = u_1 + u_2 - u_2 \in U_2.$$

Dies ist ein Widerspruch zu der Annahme, dass $u_1 \notin U_2$. Also kann Fall 1 nicht eintreten und es muss Fall 2 gelten:

Fall 2. $u_1 + u_2 \in U_1$

Da $u_1 \in U_1$, gilt dann $-u_1 \in U_1$ und somit auch

$$u_2 = -u_1 + u_1 + u_2 \in U_1.$$

Also gilt $u_2 \in U_1$ und da $u_2 \in U_2$ beliebig war, schließen wir daraus, dass $U_2 \subset U_1$ gilt. Dies schließt den Beweis ab. □

3.1.4 Linearer Span

Definition 3.1.12. *Sei K ein Körper und sei V ein K-Vektorraum. Sei $M \subset V$ eine beliebige Teilmenge.*

(1) Eine Linearkombination von Vektoren aus M ist ein Element $v \in V$ der Form

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots \lambda_r v_r$$

wobei $\lambda_1, \dots, \lambda_r \in K$ Skalare sind und $v_1, \dots, v_r \in M$ Vektoren in der Menge M sind.

(2) Falls $M \neq \emptyset$, so ist der lineare Span von M die Teilmenge

$$\operatorname{span}_K(M) \subset V$$

aller Linearkombinationen von Elementen aus M.

(3) Falls $M = \emptyset$, so definieren wir $\operatorname{span}_K(M) := \{0\}$.

Lemma 3.1.13. *Sei K ein Körper und sei V ein K-Vektorraum. Sei $M \subset V$ eine beliebige Teilmenge. Dann ist der lineare Span $\operatorname{span}_K(M) \subset V$ ein Untervektorraum von V. Wir nennen $\operatorname{span}_K(M)$ deshalb auch den von M erzeugten (oder aufgespannten) Untervektorraum.*

Beweis. Falls $M = \emptyset$, so ist $\operatorname{span}_K(M) = \{0\}$ sicherlich ein Untervektorraum von V. Sei nun $M \neq \emptyset$. Dann ist $\operatorname{span}_K(M)$ sicherlich nicht-leer (weil z.B. $M \subset \operatorname{span}_K(M)$ gilt). Falls $v, w \in \operatorname{span}_K(M)$, so gilt

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r$$

für $\lambda_i \in K$ und $v_i \in M$, sowie

$$w = \mu_1 w_1 + \mu_2 w_2 + \cdots + \mu_s w_s$$

für $\mu_j \in K$ und $w_j \in M$. Dann ist aber auch

$$v + w = \lambda_1 v_1 + \lambda_2 v_2 + \dots \lambda_r v_r + \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_s w_s$$

eine Linearkombination von Elementen aus M. Ebenso ist für jedes $\lambda \in K$, der Vektor

$$\lambda \cdot v = (\lambda\lambda_1)v_1 + (\lambda\lambda_2)v_2 + \dots + (\lambda\lambda_r)v_r$$

eine Linearkombination von Elementen aus M. Also gilt

$$v + w \in \mathrm{span}_K(M) \quad \text{und} \quad \lambda \cdot v \in \mathrm{span}_K(M).$$

Dies beweist, dass $\mathrm{span}_K(M) \subset V$ ein Untervektorraum ist. □

Bemerkung 3.1.14. *Der Untervektorraum* $\mathrm{span}_K(M) \subset V$ *enthält* M*, weil* $1 \cdot v = v \in \mathrm{span}_K(M)$ *für alle* $v \in M$ *gilt. Umgekehrt muss jeder Untervektorraum* $U \subset V$ *der* M *enthält auch beliebige Linearkombinationen aus Elementen von* M *enthalten und somit muss* $\mathrm{span}_K(M) \subset U$ *gelten. Wir sehen also, dass* $\mathrm{span}_K(M)$ *der kleinste Untervektorraum von* V *ist, welcher* M *enthält. Formal heißt das folgendes: Falls* I *eine Indexmenge ist, sodass* $U_i \subset V$ *mit* $i \in I$ *alle Untervektorräume mit* $M \subset U_i$ *bezeichnet, so gilt*

$$\mathrm{span}_K(M) = \bigcap_{i \in I} U_i.$$

Bemerkung 3.1.15. *Sei* V *eine Menge. Sei* I *eine beliebige Indexmenge und sei für jedes* $i \in I$ *ein Element* $v_i \in V$ *gegeben. Diese 'Information' nennen wir eine Familie von Elementen aus* V *(mit Indexmenge* I*) und bezeichnen eine solche Familie mit* $(v_i)_{i \in I}$*. Der Unterschied zwischen einer Teilmenge von* V *und einer Familie von Elementen aus* V *ist, dass dasselbe Element in einer Familie öfters vorkommen kann, in einer Teilmenge aber nicht. Falls* $I = \{1, 2, \dots, n\}$*, so kann man außerdem vom* i*-ten Element einer Familie sprechen (mit* $1 \leq i \leq n$*), wohingegen dies nicht für Mengen funktioniert.*

Definition 3.1.16. *Sei* K *ein Körper und sei* V *ein* K*-Vektorraum. Für eine beliebige Indexmenge* I *und eine beliebige Familie* $(v_i)_{i \in I}$ *von Elementen* $v_i \in V$ *sei*

$$\mathrm{span}((v_i)_{i \in I}) := \mathrm{span}\left(\bigcup_{i \in I} \{v_i\}\right).$$

3.1.5 Homogene lineare Gleichungssysteme

Definition 3.1.17. *Sei* K *ein Körper und seien* $n, m \geq 1$ *natürliche Zahlen. Seien* $a_{ij} \in K$ *für* $1 \leq i \leq m$ *und* $1 \leq j \leq n$ *gegeben. Ein homogenes lineares Gleichungssystem von* m *Gleichungen in* n *Unbekannten besteht aus* m *Gleichungen*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= 0. \end{aligned}$$

Ein Vektor $v = (v_1, \dots, v_n) \in K^n$ *ist eine Lösung dieses Gleichungssystems, falls*

$$\begin{aligned} a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n &= 0 \\ a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n &= 0 \\ &\vdots \\ a_{m1}v_1 + a_{m2}v_2 + \dots + a_{mn}v_n &= 0. \end{aligned}$$

Wir haben schon gesehen, dass die Menge aller Lösungen eines gegebenen homogenen linearen Gleichungssystems ein Untervektorraum $U \subset K^n$ ist. Das folgende Lemma gibt eine einfache Regel, wann dieser Unterraum nicht der Nullraum ist. Wann es also von Null verschiedene Lösungen gibt.

Lemma 3.1.18 (Fundamentallemma)**.** *Ein homogenes lineares Gleichungssystem von m Gleichungen in n Unbekannten hat immer eine nichttriviale Lösung (also eine die verschieden vom Nullvektor ist), falls $n > m$ gilt.*

Beweis. Sei $n > m$. Der Beweis verwendet Induktion nach m. Falls $m = 1$, so ist $n \geq 2$ und damit ist die Aussage klar, denn die Gleichung

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$

hat für $n \geq 2$ immer eine nicht-triviale Lösung. (Falls $a_{11} = 0$ so ist $(1, 0, 0, \ldots, 0)$ eine nicht-triviale Lösung und falls $a_{11} \neq 0$, so ist $(-a_{12}, a_{11}, 0, 0, \ldots, 0)$ eine nicht-triviale Lösung.)

Sei nun $m > 1$. Falls $a_{j1} = 0$ für alle $j = 1, \ldots, m$, so ist $v = (1, 0, 0, \ldots, 0)$ eine nicht-triviale Lösung. Wir können also annehmen, dass mindestens ein a_{j1} ungleich Null ist. Nach umsortieren, können wir annehmen, dass $a_{11} \neq 0$ gilt. Die erste Gleichung ist damit äquivalent zu

$$x_1 = \frac{-a_{12}}{a_{11}}x_2 + \frac{-a_{13}}{a_{11}}x_3 + \cdots + \frac{-a_{1n}}{a_{11}}x_n.$$

Wenn wir das in die letzten $m - 1$ Gleichungen einsetzen, erhalten wir ein homogenes Gleichungssystem von $m - 1$ Gleichungen in den $n - 1$ Unbekannten $x_2, \ldots, x_n$. Nach Induktionsannahme gibt es davon eine nicht triviale Lösung $(v_2, \ldots, v_n) \in K^{n-1}$. Sei dann

$$v_1 := \frac{-a_{12}}{a_{11}}v_2 + \frac{-a_{13}}{a_{11}}v_3 + \cdots + \frac{-a_{1n}}{a_{11}}v_n.$$

Dann ist $(v_1, \ldots, v_n) \in K^n$ eine nicht-triviale Lösung des ursprünglichen Gleichungssystems. Das beweist das Lemma. $\square$

Das Beweisverfahren das im Lemma benutzt wurde liefert ein allgemeines Verfahren um beliebige lineare Gleichungssysteme zu lösen. Wir nennen dieses Verfahren den Gauß Algorithmus. Dazu starten wir mit einem homogenen linearen Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

Wir nennen die j-te Zeile dieses Systems mit L_j. Wir stellen dann fest, dass folgende Operationen die Lösungsmenge nicht ändern:

- Vertauschen der Zeilen.
- Ersetze für ein $j = \{1, 2, \ldots, m\}$ die j-te Zeile L_j furch λL_j mit $\lambda \in K^*$.
- Ersetze für ein $j = \{1, 2, \ldots, m\}$ die j-te Zeile L_j furch $L_j + \lambda L_h$ mit $\lambda \in K$ und $h \in \{1, 2, \ldots, m\} \setminus \{j\}$.

Nach hintereinander ausführen dieser Operationen, können wir jedes homogene linear Gleichungssystem in die Gestalt

$$\begin{aligned} 0x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ 0x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ 0x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

oder

$$\begin{aligned} 1x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ 0x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ 0x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

bringen.

Im ersten Fall wenden wir nun dasselbe Verfahren auf das Gleichungssystem

$$\begin{aligned} a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

von m Gleichungen und $n-1$ Unbekannten an. Im zweiten Fall wenden wir das Verfahren auf das Gleichungssystem

$$\begin{aligned} a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

von $m-1$ Gleichungen und $n-1$ Unbekannten an.

Wendet man das iterativ an, so kann man jedes lineare Gleichungssystem in sogenannte Zeilen-Stufenform bringen. Dieses Verfahren liefert einen sehr hilfreichen Algorithmus um lineare Gleichungssysteme zu lösen.

3.1.6 Lineare Unabhängigkeit

Viele verschiedene Teilmengen eines gegebenenen Vektorraums erzeugen denselben Untervektorraum. Bei der Suche nach einem möglichst kleinen (und somit effizienten) Erzeugendensystem eines gegebenen Untervektorraums stößt man automatisch auf folgende Definition.

Definition 3.1.19. *Sei K ein Körper und sei V ein K-Vektorraum.*

- *Eine endliche Familie $v_1, \ldots, v_r$ von Vektoren aus V heißt linear unabhängig, falls gilt: Falls $\lambda_1, \ldots, \lambda_r \in K$ mit*

  $$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = 0$$

  existieren, so gilt schon $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$.

  In anderen Worten: die Vektoren $v_1, \ldots, v_r$ sind linear unabhängig wenn sich der Nullvektor 0 nur trivial aus den gegebenen Vektoren linear kombinieren lässt.

- *Eine beliebige Familie $(v_i)_{i\in I}$ von Vektoren aus V heißt linear unabhängig falls jede endliche Teilfamilie (d.h. jede Familie der Form $(v_i)_{i\in J}$ wobei $J \subset I$ eine enfliche Teilmenge ist) linear unabhängig ist.*

- *Eine beliebige Familie $(v_i)_{i\in I}$ von Vektoren aus V heißt linear abhängig, falls die Familie nicht linear unabhängig ist.*

Lemma 3.1.20. *Für eine Familie $(v_i)_{i\in I}$ von Vektoren eines K-Vektorraums V sind die folgenden Bedingungen äquivalent:*

(1) $(v_i)_{i\in I}$ ist linear unabhängig;

(2) jeder Vektor $v \in \operatorname{span}_K((v_i)_{i\in I})$ lässt sich in eindeutiger Weise als Linearkombination der Vektoren v_i mit $i \in I$ schreiben.

Beweis. Die Implikation (2) $\Rightarrow$ (1) ist klar, weil nach (2) insbesondere der Nullvektor 0 eindeutig als Linear kombination der v_i geschrieben werden kann und somit aus

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = 0$$

schon $\lambda_i = 0$ für alle i gelten muss.

Nehmen wir umgekehrt an, dass (1) gilt. Wir wollen dann (2) zeigen. Sei dazu $v \in V$ und seien

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

und

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

zwei Linearkombinationen die beide v ergeben, wobei hier $\lambda_i, \mu_i \in K$ Skalare sind, sodass nur endlich viele der λ_i und μ_i ungleich Null sind. Dann gilt

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

Und nach (1) gilt dann schon $\lambda_i = \mu_i$ für alle i. Das beweist das Lemma. □

Beispiele.

- Ein einziger Vektor $v \in V$ ist linear abhängig genau dann wenn $v = 0$ der Nullvektor ist.

- Eine Familie von Vektoren $(v_i)_{i\in I}$ ist immer linear abhängig, wenn derselbe Vektor zweimal vorkommt.

Lemma 3.1.21 (Abhängigkeitslemma)**.** *Sei K ein Körper, V ein K-Vektorraum und seien $v_1, \ldots, v_n \in V$ linear unabhängig. Falls $w \in V$ ein Vektor ist, sodass $v_1, \ldots, v_n, w$ linear abhängig sind, so ist w eine Linearkombination der Vektoren $v_1, \ldots, v_n$.*

Beweis. Sei $w \in V$ ein Vektor ist, sodass $v_1, \ldots, v_n, w$ linear abhängig sind. Dann gibt es Skalare $\lambda_i \in K$ und $\mu \in K$ sodass

$$\sum_{i=1}^{n} \lambda_i v_i + \mu w = 0$$

gilt und gleichzeitig

$$(\lambda_1, \dots, \lambda_n, \mu) \neq (0, 0, 0 \dots, 0)$$

gilt. Falls $\mu = 0$ gilt, so wären $v_1, \dots, v_n$ linear abhängig, was unserer Voraussetzung widerspricht. Also gilt $\mu \neq 0$ und obige Gleichung lässt sich wie folgt umformen:

$$w = \sum_{i=1}^{n} \frac{-\lambda_i}{\mu} v_i.$$

Also ist w eine Linearkombination der Vektoren $v_1, \dots, v_n$, wie behauptet. □

Lemma 3.1.22 (Schrankenlemma)**.** *Sei K ein Körper und $n \geq 1$ eine natürliche Zahl. Sei V ein K-Vektorraum der von n- Vektoren erzeugt werden kann. Dann ist jede Familie von $n+1$-Vektoren linear abhängig.*

Beweis. Angenommen es gibt eine Familie von Vektoren $v_1, \dots, v_n$ mit

$$V = \operatorname{span}_K(v_1, \dots, v_n).$$

Seien dann $w_1, \dots, w_{n+1}$ beliebige Vektoren. Dann gibt es Skalare $a_{ij} \in K$ mit

$$w_j = \sum_{i=1}^{n} a_{ij} v_i.$$

Betrachten wir nun das homogene lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_{n+1} &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_{n+1} &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{n,n+1}x_{n+1} &= 0. \end{aligned}$$

Das ist ein Gleichungssystem von n Gleichungen in $n+1$ Unbekannten. Da $n+1 > n$ gilt, gibt es nach dem Fundamentallemma eine nichttriviale Lösung $(\lambda_1, \dots, \lambda_{n+1}) \in K^{n+1}$. Dann gilt aber

$$\sum_{j=1}^{n+1} \lambda_j w_j = 0$$

und damit sind $w_1, \dots, w_{n+1}$ linear abhängig. Das beweist das Schrankenlemma. □

3.2 Basis und Dimension

3.2.1 Definition und Beispiele

Definition 3.2.1.

- *Eine Familie $\mathcal{B} = (v_i)_{i \in I}$ von Vektoren in einem K-Vektorraum V heißt Erzeugendensystem, falls*

$$\operatorname{span}_K(\mathcal{B}) = V$$

gilt. V heißt endlich erzeugt, falls es ein Erzeugendensystem mit endlich vielen Elementen gibt.

- *Eine Familie $\mathcal{B} = (v_i)_{i\in I}$ von Vektoren in einem K-Vektorraum V heißt Basis, falls es ein linear unabhängiges Erzeugendensystem ist. Falls $\mathcal{B}$ eine endliche Familie ist, die eine Basis von V ist, so nennt man die Anzahl der Elemente in $\mathcal{B}$ (also die Mächtigkeit der Indexmenge I) die Länge der Basis.*

Beispiele.

- Die leere Familie ist per Definition eine Basis des Nullraums $V = \{0\}$.
- Sei $e_i \in K^n$ der Vektor der Form

$$e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$$

  wobei die 1 an i-ter Stelle steht. Dann ist $e_1, e_2, \dots, e_n$ eine Basis von K^n.

Lemma 3.2.2. *Sei V ein K-Vektorraum mit $V \neq \{0\}$. Falls $\mathcal{B} = (v_i)_{i\in I}$ eine Basis von V ist, so lässt sich jedes Element v von V eindeutig schreiben als*

$$v = \sum_{i\in I} \lambda_i v_i \quad \textit{nur endlich viele } \lambda_i \neq 0.$$

Beweis. Das folgt direkt aus Lemma 3.1.20. □

3.2.2 Basissatz

Satz 3.2.1 (Basissatz für endlich erzeugte Vektorräume)**.** *Sei $V \neq \{0\}$ ein endlich erzeugter K-Vektorraum. Dann gilt:*

(i) V besitzt eine endliche Basis.

(ii) Je zwei Basen von V haben gleich viele Elemente.

(iii) Sind $v_1, \dots, v_r$ linear unabhängige Vektoren von V, so gibt es eine natürliche Zahl $n \geq r$ und Vektoren $v_{r+1}, \dots, v_n \in V$, sodass

$$v_1, \dots, v_r, v_{r+1}, \dots, v_n$$

eine Basis von V ist.

Beweis. Da V endlich erzeugt ist, gibt es eine natürliche Zahl m, sodass V von m Vektoren erzeugt wird.

Wir zeigen zuerst (iii). Falls $V = \mathrm{span}_K(v_1, \dots, v_r)$ gilt, so sind wir fertig. Ansonsten wählen wir einen Vektor

$$v_{r+1} \in V \setminus \mathrm{span}_K(v_1, \dots, v_r).$$

Dann sind $v_1, \dots, v_{r+1}$ linear unabhängig, weil sonst aus dem Abhängigkeitslemma $v_{r+1} \in \mathrm{span}_K(v_1, \dots, v_r)$ folgen würde. Falls nun $V = \mathrm{span}_K(v_1, \dots, v_r, v_{r+1})$ gilt, so sind wir fertig. Wenn das nicht gilt wählen wir einen Vektor

$$v_{r+2} \in V \setminus \mathrm{span}_K(v_1, \dots, v_r, v_{r+1}).$$

und sehen wie zuvor dass dann $v_1, \dots, v_{r+1}, v_{r+2}$ linear unabhängig. Wir setzen dieses Verfahren fort, und wissen nach dem Schrankenlemma, dass das Verfahren nach endlich vielen

Schritten abbricht, weil mehr als m Vektoren in V immer linear abhängig sind. Wenn das Verfahren abbricht, haben wir also linear unabhängige Vektoren

$$v_1, \ldots, v_r, v_{r+1}, \ldots, v_n \in V$$

mit

$$V = \mathrm{span}_K(v_1, \ldots, v_n)$$

gefunden haben, wobei $m + n \geq r$ eine natürlich Zahl ist (nach dem Schrankenlemma muss $n \leq m$ gelten). Das beweist (iii).

Da $V \neq \{0\}$ gibt es einen Vektor $v_1 \in V$ mit $v_1 \neq 0$. Dann ist v_1 linear unabhängig und wir können (iii) auf v_1 anwenden. Das liefert eine endliche Basis $v_1, \ldots, v_n$ von V, womit (i) gezeigt ist.

Um (ii) zu zeigen, sei $\mathcal{B} = (v_1, \ldots, v_n)$ die oben konstruierte Basis von V und sei $\mathcal{B}'$ eine weitere beliebige Basis von V. Nach dem Schrankenlemma, sind beliebige $n+1$ Vektoren aus V linear abhängig. Also kann $\mathcal{B}'$ höchstens n verschiedene Elemente besitzen. Falls $\mathcal{B}'$ weniger als n Elemente besitzt, so könnte man V mit weniger als n Elementen erzeugen und damit wären nach dem Schrankenlemma je n Vektoren aus V linear abhängig, was der Tatsache, dass $v_1, \ldots, v_n$ eine Basis von V ist, widerspräche. Also kann $\mathcal{B}'$ nicht weniger Elemente haben als $\mathcal{B}$ und wir schließen insgesamt, dass $\mathcal{B}'$ genauso viele Elemente wie $\mathcal{B}$ (nämlich n Stück) haben muss. Das schließt den Beweis des Satzes ab. □

Allgemeiner gilt:

Theorem 3.2.3. *Jeder K-Vektorraum hat eine Basis.*

Beweis. Genauer zeigt man, dass für jede Familie von Vektoren $(v_i)_{i\in I}$ es eine Teilmenge $J \subset I$ gibt, sodass die Familie $(v_j)_{j\in J}$ linear unabhängig ist und

$$\mathrm{span}_K((v_i)_{i\in I}) = \mathrm{span}_K((v_j)_{j\in J})$$

gilt. Die Existenz von Basen folgt offensichtlich aus dieser Aussage, indem man mit einem beliebigen Erzeugendensystem von $(v_i)_{i\in I}$ von V startet. Der Beweis obiger Aussage verwendet das Auswahlaxiom (bzw. das Zornsche Lemma, welches zum Auswahlaxiom äquivalent ist), und soll in dieser Vorlesung weggelassen werden. □

3.2.3 Dimension

Definition 3.2.4. *Sei V ein endlich erzeugter K-Vektorraum. Dann ist die Dimension von V gegeben durch*

$$\dim_K(V) := \begin{cases} n, & \text{falls es eine endliche Basis mit } n \text{ Elementen gibt;} \\ \infty, & \text{falls es keine endliche Basis gibt.} \end{cases}$$

Obige Definition ist unabhängig von der Wahl der Basis nach dem vorherigen Satz. Insbesondere gilt: falls $\dim_K(V)$ endlich ist, so hat jede Basis von V genau $\dim_K(V)$ Elemente.

Korollar 3.2.5. *Sei V ein K-Vektorraum mit $\dim_K(V) = n < \infty$. Dann hat jede Familie von linear unabhängigen Vektoren in V Länge $\leq n$, wobei Gleichheit genau dann gilt, wenn es sich um eine Basis handelt.*

Beweis. Jede endliche Familie von linear unabhängigen Vektoren in V kann nach obigem Satz zu einer Basis ergänzt werden. Da die Länge einer Basis nach obigem Satz eindeutig ist, folgt die Behauptung. □

Satz 3.2.2. *Sei V ein K-Vektorraum. Dann ist V genau dann endlich erzeugt, wenn es eine natürliche Zahl n gibt, sodass jede Familie von mehr als n Vektoren aus V linear abhängig ist.*

Beweis. Wenn V endlich erzeugt ist, so ist $\dim_K(V)$ endlich und es gibt eine Basis mit $\dim_K(V)$ vielen Elementen. Nach obigem Korollar hat dann jede Familie von linear unabhängigen Vektoren in V Länge $\leq \dim_K(V)$.

Sei also umgekehrt eine natürliche Zahl n gegeben, sodass jede Familie von mehr als n Vektoren aus V linear abhängig ist. Sei dann m die größte natürliche Zahl, sodass es linear unabhängige Vektoren $v_1, \dots, v_m$ in V gibt. (Nach Annahme gilt $m \leq n$ und somit ist m insbesondere endlich, also wirklich eine natürliche Zahl.) Nach dem Abhängigkeitslemma muss $\operatorname{span}_K(v_1, \dots, v_m) = V$ gelten, weil man sonst $m+1$ linear unabhängige Vektoren in V finden könnte. Damit hat V eine endliche Basis und damit endliche Dimension. Das beweist den Satz. $\square$

Korollar 3.2.6. *Sei V ein K-Vektorraum und sei $U \subset V$ ein Untervektorraum. Falls $\dim_K(V) < \infty$ gilt, so ist U endlich-dimensional und es gilt $\dim_K(U) \leq \dim_K(V)$.*

Beweis. Da V endlich-dimensional ist, ist jede Familie von mehr als $\dim_K(V)$ vielen Vektoren in $U \subset V$ linear abhängig. Damit ist U endlich erzeugt nach obigem Satz. Insbesondere hat U eine Basis nach Satz 3.2.1 und jede Basis von U kann zu einer Basis von V ergänzt werden. Daraus folgt aber dass $\dim_K(U) \leq \dim_K(V)$ gelten muss. Das schließt den Beweis ab. $\square$

Bemerkung 3.2.7. *Der Leser sollte sich an dieser Stelle bewusst machen, dass es a priori überhaupt nicht klar ist, dass jeder Untervektorraum eines endlich erzeugten Vektorraums wieder endlich erzeugt ist. Insbesondere ist das a priori nicht klar für Untervektorräume von K^n.*

3.2.4 Basisauswahlsatz von Steinitz

Satz 3.2.3 (Basisauswahlsatz von Steinitz)**.** *Sei V ein K-Vektorraum. Sei weiterhin $v_1, \dots, v_r$ ein Erzeugendensystem von V. Dann gibt es eine natürliche Zahl $n \leq r$ und Indizes $1 \leq i_1 < i_2 < \dots < i_n \leq r$, sodass $v_{i_1}, \dots, v_{i_n}$ eine Basis von V sind.*

Beweis. Sei

$$n := \max\{k \mid \text{es gibt } k \text{ linear unabhängige Vektoren in } v_1, \dots, v_r\}.$$

Dann gibt es Indizes $1 \leq i_1 < i_2 < \dots < i_n \leq r$, sodass $v_{i_1}, \dots, v_{i_n}$ linear unabhängig sind. Für jeden Index $1 \leq h \leq r$ ist dann $v_{i_1}, \dots, v_{i_n}, v_h$ linear abhängig (wegen Maximalität von n). Nach dem Abhängigkeitslemma, gilt also

$$v_h \in \operatorname{span}_K(v_{i_1}, \dots, v_{i_n})$$

für alle $1 \leq h \leq r$. Damit gilt auch

$$\operatorname{span}_K(v_1, \dots, v_r) \subset \operatorname{span}_K(v_{i_1}, \dots, v_{i_n}).$$

Da $v_1, \dots, v_r$ ein Erzeugendensystem von V ist, gilt

$$V = \operatorname{span}_K(v_1, \dots, v_r).$$

Also insgesamt $V = \operatorname{span}_K(v_{i_1}, \dots, v_{i_n})$ sodass $v_{i_1}, \dots, v_{i_n}, v_h$ linear abhängiges Erzeugendensystem ist, also eine Basis. Das beweist den Satz. $\square$

3.3 Lineare Abbildung

3.3.1 Definition und einfache Eigenschaften

Definition 3.3.1. *Sei K ein Körper und sei V und W jeweils ein K-Vektorraum. Eine Abbildung $f : V \to W$ heißt Homomorphismus (oder auch lineare Abbildung), falls für alle $u, v \in V$ und $\lambda \in K$, folgendes gilt:*

- $f(u + v) = f(u) + f(v)$;
- $f(\lambda v) = \lambda f(v)$.

Lemma 3.3.2. *Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Dann gilt:*

(i) $f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i f(v_i)$ *für alle* $\lambda_1, \ldots, \lambda_n \in K$ *und* $v_1, \ldots, v_n \in V$;

(ii) $f(0) = 0$;

(iii) $f(-v) = -f(v)$ *für alle* $v \in V$;

(iv) falls U ein weiterer K-Vektorraum ist und $g : U \to V$ eine lineare Abbildung ist, so ist $f \circ g$ ebenfalls eine lineare Abbildung;

(v) falls f bijektiv ist, so ist $f^{-1} : W \to V$ wieder eine lineare Abbildung.

Beweis. (i) folgt direkt aus der Definition per vollständige Induktion nach n.

(ii) folgt aus der Tatsache, dass $f(0) = f(0 + 0) = f(0) + f(0)$ gilt und $(W, +)$ eine abelsche Gruppe ist.

Nach (ii) gilt $0 = f(0) = f(v - v) = f(v) + f(-v)$. Also $f(-v) = -f(v)$ für alle $v \in V$ wie in (iii) behauptet.

Um (iv) zu zeigen, seien $u, v \in U$ und $\lambda \in K$ beliebig. Dann gilt:;

$$(f \circ g)(u + v) = f(g(u + v)) = f(g(u) + g(v)) = f(g(u)) + f(g(v)) = (f \circ g)(u) + (f \circ g)(v)$$

und

$$(f \circ g)(\lambda v) = f(g(\lambda v)) = f(\lambda g(v)) = \lambda f(g(v)) = \lambda \cdot (f \circ g)(v).$$

Das beweist (iv).

Um (v) zu zeigen, nehmen wir nun an, dass f bijektiv ist mit Umkehrabbildung $f^{-1} : W \to V$. Dann gilt $f^{-1}(f(v)) = v$ und $f(f^{-1}(w)) = w$ für alle $v \in V$ und $w \in W$. Insbesondere gilt also für $w, w' \in W$:

$$f(f^{-1}(w) + f^{-1}(w')) = f(f^{-1}(w)) + f(f^{-1}(w')) = w + w'.$$

Andererseits gilt auch $f(f^{-1}(w + w')) = w + w'$. Da f injektiv ist, muss somit $f^{-1}(w + w')) = f^{-1}(w) + f^{-1}(w')$ gelten. Ähnlich gilt für alle $\lambda \in K$ und alle $w \in W$:

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

Andererseits gilt auch $f(f^{-1}(\lambda w)) = \lambda w$. Es folgt also wieder aus der Injektivität von f, dass $\lambda f^{-1}(w) = f^{-1}(\lambda w)$ gelten muss. Insgesamt sehen wir, dass f^{-1} eine lineare Abbildung ist, wie gewollt. □

Definition 3.3.3. *Eine lineare Abbildung $f : V \to W$ zwischen K-Vektorräumen heißt:*

- *Monomorphismus, falls f injektiv ist;*
- *Epimorphismus, falls f surjektiv ist;*
- *Isomorphismus, falls f bijektiv ist;*
- *Endomorphismus, falls $V = W$ gilt;*
- *Automorphismus, falls $V = W$ gilt und f ein Isomorphismus ist.*

Bemerkung 3.3.4. *Die Relation auf der Menge aller K-Vektorräumen die durch*

$$V \sim W \quad :\Longleftrightarrow \quad \exists \text{ Isomorphismus } f : V \to W$$

gegeben ist, ist eine Äquivalenzrelation.

3.3.2 Struktursatz über linearer Abbildungen deren Wertebereich endlich erzeugt ist

Der folgende Satz ist sehr nützlich und liefert eine Vielzahl von Beispielen linearer Abbildungen.

Satz 3.3.1. *Sei V ein endlich erzeugter K-Vektorraum mit Basis $v_1, \dots, v_n$. Sei W ein beliebiger K-Vektorraum und seien $w_1, \dots, w_n$ beliebige Vektoren in W. Dann gibt es genau eine lineare Abbildung $f : V \longrightarrow W$ mit $f(v_i) = w_i$ für alle $i = 1, \dots, n$.*

Beweis. Wir zeigen zuerst die Eindeutigkeit. Da $v_1, \dots, v_n$ eine Basis von V ist, hat jedes Element $v \in V$ eine eindeutige Darstellung

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

mit $\lambda_i \in \mathbb{R}$. Falls $f : V \to W$ eine beliebige lineare Abbildung ist, so gilt:

$$f(v) = f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i \cdot f(v_i).$$

Wenn zusätzlich $f(v_i) = w_i$ für alle $i = 1, \dots, n$ gilt, so muss also

$$f(v) = f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i \cdot w_i$$

gelten. Dies zeigt, dass f (falls es existiert) eindeutig ist. Gleichzeitig gibt obige Überlegung eine Strategie wie man eine lineare Abbildung f wie gewünscht definieren könnte. Wir verfolgen das im Folgenden.

Wir nutzen dieses mal aus, dass jedes Element $v \in V$ eine eindeutige (!) Darstellung

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

mit $\lambda_i \in \mathbb{R}$ hat. (Oben haben wir nur die Existenz irgendeiner solchen Darstellung benutzt.) Die Eindeutigkeit dieser Darstellung erlaubt es uns, eine Abbildung $f : V \to W$ via

$$f(\sum_{i=1}^{n} \lambda_i v_i) := \sum_{i=1}^{n} \lambda_i w_i$$

zu definieren. Es ist dann klar, dass $f(v_i) = w_i$ für alle $i = 1, \dots, n$ gilt. Man rechnet weiterhin einfach nach, dass f linear ist. Zum Beispiel gilt:

$$f(\sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^n \mu_i v_i) = f(\sum_{i=1}^n (\lambda_i + \mu_i) v_i) = \sum_{i=1}^n (\lambda_i + \mu_i) w_i = \sum_{i=1}^n \lambda_i w_i + \sum_{i=1}^n \mu_i w_i$$

und ebenso

$$f(\sum_{i=1}^n \lambda_i v_i) + f(\sum_{i=1}^n \mu_i v_i) = \sum_{i=1}^n \lambda_i w_i + \sum_{i=1}^n \mu_i w_i.$$

Als insgesamt:

$$f(\sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^n \mu_i v_i) = f(\sum_{i=1}^n \lambda_i v_i) + f(\sum_{i=1}^n \mu_i v_i)$$

für alle λ_i und μ_i und somit

$$f(u + v) = f(u) + f(v)$$

für alle $u, v \in V$. Ähnlich zeigt man

$$f(\lambda v) = \lambda f(v)$$

für alle $\lambda \in K$ und $v \in V$. Die schließt den Beweis ab. $\square$

3.3.3 Kern und Bild

Definition 3.3.5. *Sei K ein Körper und sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(1) Der Kern von f ist $\mathrm{Ker}(f) := \{v \in V \mid f(v) = 0\}$.

(2) Für $U \subset V$ schreiben wir $f(U) := \{f(u) \mid u \in U\}$ für das Bild von U unter der Abbildung f.

(3) Das Bild von f ist $\mathrm{Bild}(f) := f(V)$.

Lemma 3.3.6. *Sei K ein Körper und sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(i) $\mathrm{Ker}(f) \subset V$ *ist ein linearer Unterraum.*

(ii) $\mathrm{Ker}(f) = \{0\} \iff f$ *ist injektiv (d.h. ein Monomorphismus)*

Beweis. Es gilt $f(0) = 0$, sodass $0 \in \mathrm{Ker}(f)$ (also ist $\mathrm{Ker}(f) \neq \emptyset$). Falls $u, v \in \mathrm{Ker}(f)$, so gilt

$$0 = 0 + 0 = f(u) + f(v) = f(u + v)$$

und somit $u + v \in \mathrm{Ker}(f)$. Weiterhin gilt für alle $\lambda \in K$:

$$f(\lambda v) = \lambda \cdot f(v) = \lambda \cdot 0 = 0.$$

Also $\lambda v \in \mathrm{Ker}(f)$. Das zeigt, dass $\mathrm{Ker}(f) \subset V$ ein linearer Unterraum ist.

Die zweite Aussage folgt von Lemma 2.1.8. Wir wiederholen das Argument zur Übung. Da $\mathrm{Ker}(f) = f^{-1}(0)$ das Urbild von 0 ist, muss $\mathrm{Ker}(f) = \{0\}$ gelten, wenn f injektiv ist. Sei umgekehrt $\mathrm{Ker}(f) = \{0\}$ und seien $u, v \in V$ mit $f(u) = f(v)$ gegeben. Dann gilt

$$0 = f(u) - f(v) = f(u) + f(-v) = f(u - v)$$

und damit $u - v \in \mathrm{Ker}(f) = \{0\}$. Also, $u - v = 0$ und somit $u = v$. Das zeigt, dass f injektiv ist, wie behauptet. Das schließt den Beweis ab. $\square$

Lemma 3.3.7. *Sei K ein Körper und sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen.*

(i) *$f(U) \subset W$ ist ein linearer Unterraum falls $U \subset V$ ein linearer Unterraum ist.*

(ii) $\mathrm{Bild}(f) \subset W$ *ist ein linearer Unterraum.*

(iii) *Falls $(v_i)_{i\in I}$ ein Erzeugendensystem von V ist, so ist $(f(v_i))_{i\in I}$ ein Erzeugendensystem von* $\mathrm{Bild}(f)$.

(iv) *Falls $v_1, \ldots, v_r$ linear abhängig in V sind, so sind $f(v_1), \ldots, f(v_n)$ linear abhängig in W.*

(v) *Falls $f(v_1), \ldots, f(v_n)$ linear unabhängig in W sind, so sind $v_1, \ldots, v_r$ linear unabhängig in V.*

(vi) *Falls V endlich erzeugt ist, so ist auch* $\mathrm{Bild}(f)$ *endlich erzeugt und es gilt* $\dim_K(\mathrm{Bild}(f)) < \dim_K(V)$.

Beweis. Aussage (i) ist eine leichte Übung.

(ii) folgt aus (i).

(iii): jedes Element $v \in V$ ist von der Form $v = \sum \lambda_i v_i$, wobei hier nur endlich viele λ_i ungleich Null sind. Dann gilt aber

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

Also $f(v) \in \mathrm{span}_K((f(v_i))_{i\in I})$. Also: $\mathrm{Bild}(f) \subset \mathrm{span}_K((f(v_i))_{i\in I})$ Die Umkehrung ist klar, sodass

$$\mathrm{Bild}(f) = \mathrm{span}_K((f(v_i))_{i\in I})$$

folgt. Das beweist (iii).

(iv)+(v): Falls $\sum \lambda_i v_i = 0$, so gilt $0 = f(\sum \lambda_i v_i) = \sum \lambda_i f(v_i)$. Das beweist (iv) und (v).

(vi) folgt aus (iii), denn wenn $v_1, \ldots, v_n$ eine Basis von V ist (die nach dem Basissatz existiert), so ist $f(v_1), \ldots, f(v_n)$ nach (iii) ein Erzeugendensystem von $\mathrm{Bild}(f)$. Also ist $\mathrm{Bild}(f)$ endlich erzeugt und $\dim_K(\mathrm{Bild}(f)) \leq n$ nach dem Steinitz'schen Auswahlsatz. $\square$

Beispiel.

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad (x_1, x_2, x_3) \mapsto (x_1 - x_2, x_1 - x_3, 0).$$

Hier überprüft man einfach, dass

$$\mathrm{Bild}(f) = \mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$$

und

$$\mathrm{Ker}(f) = \{(x, x, x) \mid x \in \mathbb{R}\} \cong \mathbb{R}.$$

Insbesondere gilt $\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Bild}(f)) = 3 = \dim(\mathbb{R}^3)$. Wir werden im nächsten Abschnitt sehen, dass diese Formel viel allgemeiner gilt.

Wir wollen noch folgende wichtige Folgerung aus obigen Überlegungen festhalten.

Satz 3.3.2. *Zwei endlich erzeugte K-Vektorräume V und W sind genau dann isomorph, wenn* $\dim_K V = \dim_K W$ *gilt.*

Beweis. Falls $f : V \to W$ ein Isomorphismus ist, und $v_1, \ldots, v_n$ eine Basis von V ist, so ist $f(v_1), \ldots, f(v_n)$ ein Erzeugendensystem von W. Also gilt $\dim_K W \leq \dim_K V$. Wenn wir dasselbe Argument auf f^{-1} anwenden, finden wir $\dim_K V \leq \dim_K W$. Also muss insgesamt $\dim_K V = \dim_K W$ gelten, wie behauptet.

Sei umgekehrt $n := \dim_K(V) = \dim_K(W)$ und seien $v_1, \ldots, v_n \in V$ und $w_1, \ldots, w_n \in W$ Basen von V, bzw. W. Dann gibt es eine lineare Abbildung

$$f : V \longrightarrow W$$

mit $f(v_i) = w_i$ für alle $i = 1, \ldots, n$. Da $w_1, \ldots, w_n \in W$ eine Basis von W ist, gilt $\text{Bild}(f) = W$, also ist f surjektiv. Weiterhin ist f injektiv, denn

$$f(\sum \lambda_i v_i) = \sum \lambda_i f(v_i) = \sum \lambda_i w_i$$

kann nur Null sein, wenn $\lambda_i = 0$ für alle $i = 1, \ldots, n$ gilt. Das beweist den Satz. □

Korollar 3.3.8. *Falls V ein endlich dimensinoaler K-Vektorraum ist, so gilt $V \cong K^n$ mit $n = \dim_K V$.*

3.3.4 Dimensionssatz für Homomorphismen

Satz 3.3.3. *Sei $f : V \longrightarrow W$ eine lineare Abbildung zwischen endlich erzeugten K-Vektorräumen. Dann gilt*

$$\dim(V) = \dim(\text{Ker}(f)) + \dim(\text{Bild}(f)).$$

Beweis. Wähle eine Basis $v_1, \ldots, v_r$ von $\text{Ker}(f)$ und ergänze diese zu einer Basis $v_1, \ldots, v_n$ von V. Dann ist $\text{Bild}(f)$ von $f(v_{r+1}), \ldots, f(v_n)$ erzeugt, und somit gilt $\dim(\text{Bild}(f)) \leq n - r$. Andererseits sind $f(v_{r+1}), \ldots, f(v_n)$ linear unabhängig, denn falls

$$\sum_{i=r+1}^{n} \lambda_i f(v_i) = 0$$

gilt, so gilt auch

$$f(\sum_{i=r+1}^{n} \lambda_i v_i) = 0,$$

also

$$\sum_{i=r+1}^{n} \lambda_i v_i \in \text{Ker}(f).$$

Da $v_1, \ldots, v_r$ eine Basis von $\text{Ker}(f)$ ist, folgt daraus dass entweder $\lambda_i = 0$ für $i = r+1, \ldots, n$, oder dass $v_1, \ldots, v_n$ linear abhängig sind. Letzteres ist falsch, sodass die Aussage die wir zeigen wollen folgt.

Insgesamt sehen wir also, dass $f(v_{r+1}), \ldots, f(v_n)$ eine Basis von $\text{Bild}(f)$ ist, und somit gilt

$$\dim(\text{Ker}(f)) + \dim(\text{Bild}(f)) = r + (n - r) = n = \dim_K V.$$

Das schließt den Beweis ab. □

Bemerkung 3.3.9. *Für eine lineare Abbildung $f : V \to W$ zwischen endlich erzeugten K-Vektorräumen heißt $\dim_K(\text{Bild}(f))$ auch der Rang von f:*

$$\text{Rang}(f) := \dim_K(\text{Bild}(f)).$$

Nach obigem Satz gilt also:

$$\dim_K V = \dim_K(\text{Ker}(f)) + \text{Rang}(f).$$

Korollar 3.3.10. *Sei $f : V \to W$ eine lineare Abbildung zwischen zwei endlich dimensionalen K-Vektorräumen derselben Dimension: $\dim_K V = \dim_K W$. Dann ist f genau dann ein Isomorphismus, wenn f entweder injektiv oder surjektiv ist.*

Beweis. Übungsblatt 8! $\square$

3.4 Summen und Komplemente

Definition 3.4.1. *Sei K ein Körper und sei V ein K-Vektorraum. Seien I eine beliebige Indexmenge und sei für jedes $i \in I$ ein Untervektorraum $U_i \subset V$ gegeben.*

(1) Die Summe der U_i's ist gegeben durch

$$\sum_{i\in I} U_i := \operatorname{span}_K\left(\bigcup_{i\in I} U_i\right).$$

(2) Falls für alle $i_0 \in I$

$$U_{i_0} \cap \sum_{i\in I\setminus\{i_0\}} U_i = \{0\}$$

gilt, so sagen wir, dass die U_i's eine direkte Summe bilden, oder in direkter Summe stehen. Wir schreiben dann

$$\bigoplus_{i\in I} U_i := \sum_{i\in I} U_i.$$

Der Einfachheit halber konzentrieren wir uns im folgenden auf den Fall dass die Indexmenge I endlich ist. Falls I nichtleer ist, können wir dann annehmen, dass es eine natürliche Zahl r mit $I = \{1, 2, \ldots, r\}$ gibt.

Lemma 3.4.2. *Endlich viele Untervektorräume $U_1, \ldots, U_r$ eines K-Vektorraums V sind in direkter Summe genau dann wenn die Abbildung*

$$f : U_1 \times U_2 \times \cdots \times U_r \longrightarrow V, \quad (u_1, \ldots, u_r) \mapsto \sum_{i=1}^{r} u_i$$

injektiv ist.

Beweis. Das endliche Produkt $U_1 \times U_2 \times \cdots \times U_r$ ist in natürlicherweise ein K-Vektorraum und die angegebene Abbildung ist K-linear. Also ist f genau dann injektiv, wenn $\operatorname{Ker}(f) = \{0\}$ gilt. Letzteres gilt per Definition genau dann wenn die U_i's in direkter Summe sind. $\square$

Korollar 3.4.3. *Falls endlich viele Untervektorräume $U_1, \ldots, U_r$ eines K-Vektorraums V eine direkte Summe bilden, so gilt*

$$U_1 \times U_2 \times \cdots \times U_r \cong \bigoplus_{i=1}^{r} U_i.$$

Beweis. Die Abbildung f aus obigem Lemma ist linear und injektiv, falls $U_1, \ldots, U_r$ eine direkte Summe bilden. Also gilt

$$U_1 \times U_2 \times \cdots \times U_r \cong \operatorname{Bild}(f).$$

Aber das Bild von f ist per Definition genau $\bigoplus_{i=1}^{r} U_i$. Das beweist das Korollar. $\square$

Definition 3.4.4. *Sei V ein K-Vektorraum und sei $U \subset V$ ein Untervektorraum. Ein Komplement zu U ist ein Untervektorraum $W \subset V$ mit*

$$U \oplus W = V.$$

Das heißt, $U \cap W = \{0\}$ und $V = \mathrm{span}_K(U \cup W)$.

Theorem 3.4.5. *Sei V ein K-Vektorraum und sei $U \subset V$ ein Untervektorraum. Dann existiert immer ein Komplement $W \subset V$ zu U.*

Beweis. Wir beweisen nur den Fall in dem V endlich erzeugt ist. Falls $\dim_K V = \infty$, so benutzt das Theorem das Zornsche Lemma, also das Auswahlaxiom.

Falls V endlich erzeugt ist, so ist auch U endlich erzeugt (nach Korollar 3.2.6). Nach dem Basissatz hat U also eine endliche Basis $v_1, \ldots, v_r$. Nach dem Basisergänzuungssatz kann diese zu einer Basis $v_1, \ldots, v_n$ von V ergänzt werden. Falls $r = n$ gilt, so ist $U = V$ und wir setzen $W = \{0\}$. Falls $r < n$ gilt, so setzen wir $W := \mathrm{span}_K(v_{r+1}, \ldots, v_n)$ und erhalten

$$U + W = V \quad \text{und} \quad U \cap W = \{0\}.$$

(Beide Aussagen folgen einfach aus der Tatsache, dass $v_1, \ldots, v_n$ eine Basis von V ist.) Das beweist das Theorem. □

Bemerkung 3.4.6. *Komplemente sind im Allgemeinen nicht eindeutig bestimmt. Das sieht man aus obigem Beweis, denn die linear unabhängigen Vektoren $v_1, \ldots, v_r$ können auf verschiedene Art und Weisen zu einer Basis $v_1, \ldots, v_n$ ergänzt werden (siehe Beweis des Basis Ergänzungssatzes).*

Kapitel 4

Matrizen

4.1 Matrizen – Definition und grundlegende Eigenschaften

4.1.1 Der Vektorraum aller linearer Abbildungen

Lemma 4.1.1. *Seien V und W K-Vektorräume. Dann hat die Menge* $\mathrm{Hom}_K(V,W)$ *aller linearer Abbildungen $f : V \to W$ in natürlicherweise die Struktur eines K-Vektorraums, wobei Vektoraddition und Skalarmultiplikation, wie folgt definiert sind:*

$$(f+g) : V \to W, \quad v \mapsto f(v) + g(v) \quad \textit{und} \quad \lambda \cdot f : V \to W, \quad v \mapsto \lambda \cdot f(v)$$

für alle $f, g \in \mathrm{Hom}_K(V,W)$ und $\lambda \in K$.

Beweis. Das ist eine leichte Übung. □

Lemma 4.1.2. *Seien V und W endlich erzeugte K-Vektorräume. Dann gilt*

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

Beweis. Sei $v_1, \ldots, v_n$ eine Basis von V und sei $w_1, \ldots, w_m$ eine Basis von W. Für i, j mit $1 \leq i \leq n$ und $1 \leq j \leq m$ betrachten wir dann die lineare Abbildung

$$f_{ij} : V \longrightarrow W$$

die eindeutig dadurch festgelegt ist, dass

$$f_{ij}(v_k) = \begin{cases} w_j & \text{falls } k = i; \\ 0 & \text{sonst.} \end{cases}$$

Es ist eine einfache Übung nachzuweisen, dass die linearen Abbildungen f_{ij} mit $1 \leq i \leq n$ und $1 \leq j \leq m$ eine Basis von $\mathrm{Hom}_K(V,W)$ bilden. Also gilt

$$\dim_K \mathrm{Hom}_K(V,W) = m \cdot n = \dim_K V \cdot \dim_K W$$

wie behauptet. □

4.1.2 Matrizen als darstellendes Schema einer linearen Abbildung

Erinnerung: Jeder endlich-erzeugte K-Vektorraum V ist isomorph zu K^n für ein eindeutig bestimmtes $n \in \mathbb{N}$. Genauer gilt $n = \dim V$, d.h. n ist die Länge einer Basis von V. Um lineare Abbildungen $f : V \to W$ zwischen endlich erzeugten K-Vektorräumen zu untersuchen, können/wollen wir uns deshalb erst einmal auf den Fall $V = K^n$ und $W = K^m$ konzentrieren.

Wir wissen, dass eine lineare Abbildung

$$f : K^n \longrightarrow K^m$$

eindeutig dadurch festgelegt ist, welche Werte f auf einer Basis von K^n annimmt. Umgekehrt gibt es für jede Wahl von solchen Werten eine eindeutige lineare Abbildung. Um lineare Abbildungen $f : K^n \to K^m$ zu studieren, sollten wir also eine Basis von K^n festlegen und es erscheint vernünftig die Standardbasis $e_1, \ldots, e_n$ zu wählen, wobei hier

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \cdots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

ist.

Eine lineare Abbildung $f : K^n \to K^m$ ist eindeutig durch die Werte $f(e_1), \ldots, f(e_n)$ auf obiger Standardbasis festgelegt. Jeder der Werte $f(e_i)$ ist ein Vektor in K^m. Wir können also schreiben:

$$f(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{mi} \end{pmatrix} \in K^m$$

für Elemente $a_{ji} \in K$ wobei $i = 1, \ldots, n$ und $j = 1, \ldots, m$.

Die lineare Abbildung f ist damit eindeutig durch die Elemente $a_{ji} \in K$ mit $i = 1, \ldots, n$ und $j = 1, \ldots, m$ festgelegt. Es ist nützlich diese Elemente in Form einer Matrix anzuordnen.

Definition 4.1.3. *Sei K ein Körper und seien $n, m \geq 1$ positive natürliche Zahlen. Eine $m \times n$ Matrix ist ein rechteckiges Schema bestehend aus m Zeilen und n Spalten von Elementen aus K:*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

Die Menge aller $m \times n$-Matrizen bezeichnen wir mit $K^{m \times n}$.

Lemma 4.1.4. *Die Menge $K^{m \times n}$ aller $m \times n$-Matrizen ist in natürlicher weise ein K-Vektorraum der zu K^{mn} isomorph ist.*

Beweis. Offensichtlich ist $K^{m \times n}$ natürlicherweise zu K^{mn} in bijektion, indem man die Einträge einer Matrix in geeigneter Weise untereinander schreibt und dadurch eine $m \times n$ Matrix eindeutig einem Vektor mit nm Einträgen zuordnet. □

Bemerkung 4.1.5. *Wir haben oben gesehen, dass die Menge* $\mathrm{Hom}_K(K^n, K^m)$ *aller linearer Abbildungen* $f : K^n \to K^m$ *in natürlicherweise mit der Menge aller* $m \times n$*-Matrizen identifiziert werden kann (d.h. es gibt eine natürliche bijektive Abbildung zwischen den beiden Mengen), wobei* f *zu der Matrix*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \in K^{m \times n} \quad \textit{mit} \quad f(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{mi} \end{pmatrix} \in K^m \ \ \forall\ i = 1, \ldots, n$$

gehört. Sowohl $\mathrm{Hom}_K(K^n, K^m)$ *als auch* $K^{m \times n}$ *ist ein* K*-Vektorraum und obige bijektive Abbildung ist ein Isomorphismus von* K*-Vektorräumen:*

$$\mathrm{Hom}_K(K^n, K^m) \cong K^{m \times n}.$$

4.1.3 Produkt von Matrizen

Die Menge aller Matrizen ist nicht nur ein Vektorraum, hat also eine Vektoraddition und Skalarmultiplikation, sondern man kann auch eine Multiplikation definieren. Der Grund hierfür ist in obiger Bemerkung zu finden: $K^{m \times n}$ kann natürlicherweise mit $\mathrm{Hom}_K(K^n, K^m)$ identifiziert werden. (Beachten Sie hier, dass sich in dieser Identifikation die Reihenfolge von m und n vertauscht.) Aber für positive natürlicher Zahlen ℓ, m, n gibt es eine K-lineare Abbildung

$$\mathrm{Hom}_K(K^m, K^\ell) \times \mathrm{Hom}_K(K^n, K^m) \longrightarrow \mathrm{Hom}_K(K^n, K^\ell), \quad (f, g) \mapsto f \circ g.$$

Auf der Matrizenseite, entspricht das einer Abbildung

$$K^{\ell \times m} \times K^{m \times n} \longrightarrow K^{\ell \times n}, \quad (A, B) \mapsto A \cdot B.$$

Man rechnet nach, dass die so definierte Verknüpfung $(A, B) \mapsto A \cdot B$ wie folgt explizit beschrieben werden kann.

Definition 4.1.6. *Sei* K *ein Körper und seien* $A = (a_{ij}) \in K^{\ell \times m}$ *und* $B = (b_{jk}) \in K^{m \times n}$ *zwei Matrizen, so heißt*

$$A \cdot B := \begin{pmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2n} \\ c_{31} & c_{32} & c_{33} & \cdots & c_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_{\ell 1} & c_{\ell 2} & c_{\ell 3} & \cdots & c_{\ell n} \end{pmatrix} \in K^{\ell \times n} \qquad \textit{mit} \quad c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}$$

das Matrixprodukt von A *und* B*.*

Lemma 4.1.7. *Seien* $f : K^m \to K^\ell$ *und* $g : K^n \to K^m$ *lineare Abbildungen mit zugehörigen Matrizen (bezüglich der jeweiligen Standardbasen)*

$$A_f \in K^{\ell \times m} \quad \textit{und} \quad B_g \in K^{m \times n}.$$

Dann ist die lineare Abbildung $f \circ g : K^n \to K^\ell$ *(bezüglich den jeweiligen Standardbasen) durch die Matrix*

$$A_f \cdot B_g \in K^{\ell \times n}$$

gegeben.

Beweis. Sei $A_f = (a_{kj})$ und $B_g = (b_{ji})$. Sei weiterhin $e_1, \ldots, e_n$ die Standardbasis von K^n und sei $e'_1 \ldots, e'_m$ die Standardbasis von K^m und $e''_1, \ldots, e''_\ell$ die Standardbasis von K^ℓ. Dann gilt:

$$g(e_i) = \sum_{j=1}^{m} b_{ji} e'_j \quad \text{und} \quad f(e'_j) = \sum_{k=1}^{\ell} a_{kj} e''_k.$$

Nach Linearität also:

$$f \circ g(e_i) = f(\sum_{j=1}^{m} b_{ji} e'_j) = \sum_{j=1}^{m} b_{ji} f(e'_j) = \sum_{j=1}^{m} b_{ji} \sum_{k=1}^{\ell} a_{kj} e''_k = \sum_{j=1}^{m} \sum_{k=1}^{\ell} b_{ji} a_{kj} e''_k = \sum_{k=1}^{\ell} \sum_{j=1}^{m} a_{kj} b_{ji} e''_k.$$

Also gehört $f \circ g : K^n \to K^\ell$ bezüglich der oben gewählten Standardbasen zu der $\ell \times n$ Matrix $C_{f \circ g} := (c_{ki})$ mit

$$c_{ki} = \sum_{j=1}^{m} a_{kj} b_{ji}.$$

Nach Definition des Matrixprodukts gilt also:

$$A_f \cdot B_g = C_{f \circ g}.$$

Das beweist das Lemma. $\square$

Lemma 4.1.8. *Sei $f : K^n \to K^m$ eine lineare Abbildung und sei $A_f = (a_{ji}) \in K^{m \times n}$ die $m \times n$ Matrix welche f bezüglich den Standardbasen beschreibt. Dann gilt für jedes $v \in K^n$,*

$$f(v) = A_f \cdot v \in K^{m \times 1} = K^m$$

wobei wir hier $v \in K^n = K^{n \times 1}$ als Spaltenvektor schreiben und wir das Matrixprodukt

$$K^{m \times n} \times K^{n \times 1} \longrightarrow K^{m \times 1}$$

ausnutzen.

Beweis. Wir geben zwei Beweise des Lemmas.

Beweis 1: Das Lemma folgt direkt aus Lemma 4.1.7. In der Tat, falls wir v fixieren, so gibt es eine lineare Abbildung

$$g : K \to K^n, \quad \lambda \mapsto \lambda \cdot v.$$

Wir schreiben

$$v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}.$$

Die lineare Abbildung g wird dann bezüglich der Standardbasen von einer $n \times 1$ Matrix B_g dargestellt, die genau die Form

$$B_g = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

hat. Nach Lemma 4.1.7 wird dann die Komposition

$$K \xrightarrow{g} K^n \xrightarrow{f} K^m$$

von der Matrix

$$A_f \cdot B_g = A_f \cdot v \in K^{m\times 1}$$

dargestellt. Dies ist also die Matrix welche die lineare Abbildung

$$K \longrightarrow K^m, \quad \lambda \mapsto \lambda \cdot f(v)$$

bezüglich den Standardbasen darstellt. Das bedeutet aber genau dass

$$f(v) = A_f \cdot v$$

gilt, wie behauptet.

Beweis 2: Wir schreiben

$$v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}.$$

Dann gilt:

$$A_f \cdot v = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_i a_{1i}x_i \\ \sum_i a_{2i}x_i \\ \sum_i a_{3i}x_i \\ \vdots \\ \sum_i a_{mi}x_i \end{pmatrix} \in K^{m\times 1} = K^m.$$

Falls $e_1, \dots, e_n$ und $e'_1, \dots, e'_m$ die Standardbasen von K^n und K^m sind, so gilt andererseits

$$f(e_i) = \sum_j a_{ji} e'_j$$

und damit

$$f(v) = f(\sum_i x_i e_i) = \sum_i x_i f(e_i) = \sum_i x_i \sum_j a_{ji} e'_j = \sum_i \sum_j a_{ji} x_i e'_j.$$

In Vektorschreibweise liest sich das wie folgt:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} \sum_i a_{1i}x_i \\ \sum_i a_{2i}x_i \\ \sum_i a_{3i}x_i \\ \vdots \\ \sum_i a_{ni}x_i \end{pmatrix} \in K^m.$$

Vergleicht man beide Terme, erhält man

$$f(v) = A_f \cdot v$$

wie behauptet. Das beweist das Lemma. □

Proposition 4.1.9. *Sei K ein Körper und sei n eine positive natürliche Zahl. Dann gibt es einen K-linearen Isomorphismus*

$$\mathrm{Hom}_K(K^n, K^m) \xrightarrow{\cong} K^{m\times n}, \quad f \mapsto A_f$$

der einer linearen Abbildung $f : K^n \to K^m$ die Darstellende matrix bezüglich der Einheitsbasis zuordnet. Die Umkehrabbildung von obigem Isomorphismus ist der Isomorphismus

$$K^{m\times n} \xrightarrow{\cong} \mathrm{Hom}_K(K^n, K^m), \quad A \mapsto h_A$$

der einer $m \times n$ Matrix die lineare Abbildung

$$h_A : K^n \longrightarrow K^m, \quad v \mapsto A \cdot v$$

zuordnet.

Beweis. Das folgt direkt aus den obigen beiden Lemmata. □

Bemerkung 4.1.10. *Lemma 4.1.7 besagt, dass in obiger Proposition die Formel $A_{f\circ g} = A_f \cdot A_g$ gilt. Da nach obiger Proposition $h_{A_f} = f$ gilt, impliziert das auch folgende wichtige Relation: $h_{AB} = h_A \circ h_B$.*

Lemma 4.1.11. *Für Matrizen A, B, C und $\lambda \in K$ gelten folgende Rechenregeln, wobei immer angenommen wird, dass die Matrizen jeweils so gewählt sind, dass die jeweiligen Matrixprodukte definiert sind:*

(i) $(\lambda \cdot A) \cdot B = A \cdot (\lambda \cdot B)$;

(ii) $A(B + C) = A \cdot B + A \cdot C$;

(iii) $(A + B) \cdot C = AC + BC$;

(iv) $(AB)C = A(BC)$.

Beweis. Das folgt aus den entsprechenden Eigenschaften für die Komposition linearer Abbildungen. Alternativ können die Regeln auch explizit nachgerechnet werden. □

Beispiel. Im Allgemeinen ist $AB \neq BA$, selbst wenn beide Seiten definiert sind. Zum Beispiel gilt:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

4.1.4 Allgemeine lineare Gruppe

Definition 4.1.12. *Sei n eine positive natürliche Zahl. Die Matrix*

$$\mathbb{1}_n := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \in K^{n\times n}$$

heißt Einheitsmatrix, oder auch $n \times n$-Einheitsmatrix.

Lemma 4.1.13. *Für jede $n \times n$ Matrix $A \in K^{n\times n}$ gilt: $A \cdot \mathbb{1}_n = \mathbb{1}_n \cdot A = A$.*

Beweis. Das folgt direkt aus der Definition des Matrizenprodukts. □

Definition 4.1.14. *Sei K ein Körper und sei n eine positive natürliche Zahl. Eine $n \times n$ Matrix $A \in K^{n\times n}$ heißt invertierbar, falls es eine Matrix $A' \in K^{n\times n}$ mit*

$$A'A = AA' = \mathbb{1}_n$$

gibt. Die Menge aller invertierbaren $n \times n$ Matrizen wird mit $\mathrm{GL}(n,K)$ bezeichnet.

Bemerkung 4.1.15. *Eine Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn die zugehörige lineare Abbildung*

$$h_A : K^n \to K^n, \quad v \mapsto A \cdot v$$

ein Automorphismus ist, d.h. bijektiv ist. Das inverse zu A ist dann durch die darstellende Matrix von h_A^{-1} bezüglich der Einheitsbasen gegeben.

Proposition 4.1.16. $\mathrm{GL}(n,K)$ *ist eine Gruppe bezüglich der Matrixmultiplikation. Das neutrale Element ist durch $\mathbb{1}_n$ gegeben.*

Beweis. Wir wissen schon, dass die Matrixmultiplikation

$$K^{n\times n} \times K^{n\times n} \longrightarrow K^{n\times n}$$

assoziativ ist. Als nächstes müssen wir zeigen, dass die Matrixmultiplikation die Untermenge der invertierbaren Matrizen respektiert, d.h. dass das Produkt von zwei invertierbaren Matrizen wieder invertierbar ist. Seien dazu $A, B \in \mathrm{GL}(n,K)$. Dann gibt es Matrizen A', B' mit $A'A = \mathbb{1}_n$ und $B'B = \mathbb{1}_n$. Also:

$$B'A'AB = B'\mathbb{1}_n B = B'B = \mathbb{1}_n \quad \text{und} \quad ABB'A' = A\mathbb{1}_n A' = \mathbb{1}_n.$$

Somit gilt also $AB \in \mathrm{GL}(n,K)$. Dies zeigt, dass die Matrixmultiplikation eine Verknüpfung

$$\mathrm{GL}(n,K) \times \mathrm{GL}(n,K) \longrightarrow \mathrm{GL}(n,K)$$

liefert. Diese Verknüpfung ist assoziativ, weil die Matrixmultiplikation im Allgemeinen diese Eigenschaft hat. Weiterhin ist $\mathbb{1}_n$ ein neutrales Element und es gibt per Definition ein links ud rechts inverses zu jedem Element $A \in \mathrm{GL}(n,K)$. (Beachten Sie hier, dass per Definition das inverse A' zu A wieder in $\mathrm{GL}(n,K)$ enthalten ist.) Das zeigt, dass $\mathrm{GL}(n,K)$ eine Gruppe ist. □

Bemerkung 4.1.17. *Für $n \geq 2$ ist $\mathrm{GL}(n,K)$ nicht abelsch. Das kann man zum Beispiel für $n = 2$ einfach anhand der folgenden Beispiele sehen:*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Man rechnet einfach nach, dass die zugehörigen linearen Abbildungen $h_A, h_B : K^2 \to K^2$ invertierbar sind, sodass wirklich $A, B \in \mathrm{GL}(n,K)$ gilt. Andererseits gilt:

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \textit{und} \quad BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Also ist $\mathrm{GL}(2,K)$ nicht abelsch. Der Fall $n \geq 2$ folgt dann daraus, dass es einen injektiven Gruppenhomomorphismus $\mathrm{GL}(2,K) \to \mathrm{GL}(n,K)$ gibt, der dadurch gegeben ist, dass man $K^n \cong K^2 \oplus K^{n-2}$ ausnutzt und einen Automorphismus $f : K^2 \to K^2$ auf den Automorphismus

$$(f, \mathrm{id}_{K^{n-2}}) : K^2 \oplus K^{n-2} \longrightarrow K^2 \oplus K^{n-2}$$

abbildet. Auf Seite der Matrizen bedeutet dies, dass man aus einer invertierbaren 2×2 Matrix A eine invertierbare $n \times n$ Matrix bastelt in dem man rechts unten eine $n-2$ Einheitsmatrix $\mathbb{1}_{n-2}$ anfügt und den Rest mit Nullen auffüllt. Schematisch sieht diese Matrix dann so aus:

$$\begin{pmatrix} A & 0 \\ 0 & \mathbb{1}_{n-2} \end{pmatrix},$$

wobei die Einträge hier selbst wieder Matrizen sind.

4.2 Äquivalenz von Matrizen

Definition 4.2.1. *Zwei Matrizen $A, B \in K^{m \times n}$ heißen äquivalent, geschrieben $A \sim B$, falls es invertierbare Matrizen $S \in \mathrm{GL}(n, K)$ und $T \in \mathrm{GL}(m, K)$ gibt, sodass folgendes gilt:*

$$A = TBS.$$

4.2.1 Spaltenrang

Wir erinnern daran, dass für eine lineare Abbildung $f : V \to W$ zwischen K-Vektorräumen, der Rang von f als

$$\mathrm{Rang}(f) := \dim_K(\mathrm{Bild}(f))$$

definiert wird.

Definition 4.2.2. *Sei $A \in K^{m \times n}$ Matrix mit zugehöriger linearer Abbildung $h_A : K^n \to K^m$. Dann heißt*

$$\mathrm{Rang}(A) := \mathrm{Rang}(h_A)$$

der Rang (oder Spaltenrang) von A.

Lemma 4.2.3. *Der Rang/Spaltenrang einer Matrix $A \in K^{m \times n}$ ist die Dimension des Unterraums $U \subset K^m$ der von den Spalten der Matrix A erzeugt wird.*

Beweis. Der Rang von A ist explizit durch die Dimension des Unterraums von K^m gegeben, der durch

$$h_A(e_1), \ldots, h_A(e_n)$$

aufgespannt wird. Da $h_A(e_i)$ genau die i-te Spalte der Matrix A ist, folgt das Lemma. $\square$

Frage: Wie berechnet man den Rang einer Matrix explizit?

Lemma 4.2.4. *Falls zwei Matrizen $A, B \in K^{m \times n}$ äquivalent sind, so haben A und B denselben Rang.*

Beweis. Das ist eine direkte Folgerung aus der Definition. In der Tat, falls $A \sim B$, so gilt $A = TBS$ für invertierbare Matrizen T und S. Dann sind aber die Abbildungen h_T und h_S invertierbar und damit ist das Bild von

$$h_A = h_T \circ h_B \circ h_S$$

isomorph zu dem Bild von h_B. In der Tat,

$$\mathrm{Bild}(h_A) = h_T(h_B(h_S(K^n))) = h_T(h_B(K^n)) = h_T(\mathrm{Bild}(h_B))$$

und damit induziert der Isomorphismus

$$h_T : K^m \longrightarrow K^m$$

einen Isomorphismus $\mathrm{Bild}(h_B) \cong \mathrm{Bild}(h_A)$. Insbesondere haben also die Bilder von h_A und h_B dieselbe Dimension. Also gilt $\mathrm{Rang}(A) = \mathrm{Rang}(B)$. $\square$

Satz 4.2.1. *Sei $A \in K^{m\times n}$ eine $m \times n$ Matrix. Dann ist A äquivalent zu einer Matrix B von der Blockgestalt*

$$B = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

wobei $\mathbb{1}_r$ die $r \times r$-Einheitsmatrix ist und $r = \mathrm{Rang}(A)$ der Rang von A ist.

Beweis. Betrachte die lineare Abbildung

$$h_A : K^n \longrightarrow K^m.$$

Per Definition gilt $r := \mathrm{Rang}(A) = \dim \mathrm{Bild}(h_A)$. Nach der Dimensionsformel gilt

$$\dim(\mathrm{Ker}(h_A)) = n - r.$$

Wir wählen dann eine Basis $v_{r+1}, \dots v_n$ von $\mathrm{Ker}(h_A)$ und ergänzen dies zu einer Basis $v_1, \dots, v_n$ von V. Dann ist $w_1 := f(v_1), \dots, w_r := f(v_r)$ eine Basis des Bildes $\mathrm{Bild}(h_A)$, denn es erzeugt das Bild und das Bild hat Dimension r. Wir ergänzen $w_1, \dots, w_r$ zu einer Basis $w_1, \dots, w_m$ von K^m. Betrachte dann die linearen Abbildungen

$$\phi : K^n \longrightarrow K^n, \quad \sum_i \lambda_i e_i \mapsto \sum_i \lambda_i v_i$$

wobei $e_1, \dots, e_n$ die Standardbasis von K^n ist und

$$\psi : K^m \longrightarrow K^m, \quad \sum_j \mu_j w_i \mapsto \sum_j \mu_j e'_j$$

wobei $e'_1, \dots, e'_m$ die Standardbasis von K^m ist. Dies sind jeweils Isomorphismen, weil Basen auf Basen abgebildet werden. Also gibt es invertierbare Matrizen $S \in \mathrm{GL}(n, K)$ und $T \in \mathrm{GL}(m, K)$ mit $\phi = h_S$ und $\psi = h_T$. Das Matrizenprodukt $B := TAS$ ist dann äquivalent zu A. Die zugehörige lineare Abbildung h_B hat die Eigenschaft $h_B = h_T \circ h_A \circ h_S : K^n \to K^m$. Die Matrix B ist nun genau die zu h_B gehörige Matrix, und damit gilt

$$h_B(e_i) = \sum_{j=1}^{m} b_{ji} e'_j$$

Andererseits gilt:

$$h_B(e_i) = h_T \circ h_A \circ h_S(e_j) = \psi(h_A(\phi(e_i))) = \psi(h_A(v_i))$$

Falls $i > r$, so gilt also $h_B(e_i) = 0$. Falls $i \le r$, so gilt

$$h_B(e_i) = \psi(h_A(v_i)) = \psi(w_i) = e'_i.$$

Also gilt:

$$B = (b_{ji}) = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

wie behauptet. Das beweist den Satz. □

Korollar 4.2.5. *Zwei Matrizen $A, B \in K^{m\times n}$ sind genau dann äquivalent zu einander, wenn sie denselben Rang haben.*

Beweis. Dies folgt direkt aus obigem Satz und der Tatsache, dass die Relation $A \sim B$ eine Äquivalenzrelation ist. □

4.2.2 Zeilenrang und die transponierte Matrix

Definition 4.2.6. *Sei $A = (a_{ij}) \in K^{m\times n}$ eine $m \times n$ Matrix. Dann ist die zu A transponierte Matrix gegeben durch*

$$A^t := (a_{ji}) \in K^{n\times m}.$$

Manchmal schreibt man auch A^T anstelle von A^t.

Beispiel.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

Lemma 4.2.7. *Seien $A \in K^{\ell\times m}$ und $B\imath K^{m\times n}$ Matrizen. Dann gilt*

$$(AB)^t = B^t A^t.$$

Beweis. Sei $A = (a_{kj})$ und $B = (b_{ji})$. Sei weiterhin $C := AB$. Dann gilt $C = (c_{ki})$ mit

$$c_{ki} = \sum_j a_{kj} b_{ji}.$$

Weiterhin gilt: $A^t = (a^t_{jk})$ und $B^t = (b^t_{ij})$ wobei $a^t_{jk} = a_{kj}$ und $b^t_{ij} = b_{ji}$ gilt. Sei dann $D := B^t A^t$. Dann gilt $D = (d_{ik})$ mit

$$d_{ik} = \sum_j b^t_{ij} a^t_{jk} = a_{kj} b_{ji} = c_{ki}.$$

Also: $D^t = C$ und damit $(AB)^t = B^t A^t$. Das beweist das Lemma. □

Korollar 4.2.8. $A \in \mathrm{GL}(n, K) \Rightarrow A^t \in \mathrm{GL}(n, K)$.

Beweis. Das folgt direkt aus obigem Lemma. Denn falls $A'A = AA' = \mathbb{1}_n$ gilt, so gilt nach obigem Lemma auch

$$A^t (A')^t = (A')^t A^t = \mathbb{1}_n.$$

Also ist auch A^t invertierbar. □

Satz 4.2.2. *Sei $A \in K^{m\times n}$. Dann haben A und $A^t \in K^{n\times m}$ denselben Rang. In anderen Worten: der von den Spalten von A aufgespannte Unterraum von K^m hat dieselbe Dimension wie der von den Zeilen von A aufgespannte Unterraum von K^n. Wir sagen auch: der Spaltenrang und der Zeilenrang von A stimmen überein.*

Beweis. Nach Satz 4.2.1 gibt es invertierbare Matrizen S, T, sodass $B := SAT$ die Gestalt

$$B = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

hat, wobei $r = \mathrm{Rang}(A)$ gilt. Damit gilt aber:

$$B = B^t = (SAT)^t = T^t A^t S^t.$$

Da das transponierte einer invertierbaren Matrix wieder invertierbar ist, folgt dass A^t äquivalent zu B ist und damit denselben Rang wie B hat. Also $\mathrm{Rang}(A) = \mathrm{Rang}(B) = \mathrm{Rang}(A^t)$. □

Bemerkung 4.2.9. *Wir haben gesehen, dass A zu einer linearen Abbildung $h_A : K^n \to K^m$ gehört. Die transponierte Matrix gehört demnach zu einer Abbildung $h_{A^t} : K^m \to K^n$.*

Diese Operation, also einer linearen Abbildung $V \to W$ eine lineare Abbildung in die entgegengesetzte Richtung zuzuordnen ist von der Perspektive von linearen Abbildungen nicht offensichtlich. Dies ist ein allgemeines Phänomen: wenn man ein bestimmtes mathematisches Objekt auf verschiedene Arten beschreiben kann, ist es oft so, dass eine Sichtweise Dinge einfach/offensichtlich erscheinen lässt, die in der anderen Sichtweise überraschend sind.

In dem uns vorliegenden Fall kann man trotzdem erklären, wo die Abbildung 'in die andere Richtung' herkommt. Wir tun dies in dem folgenden Unterkapitel, das sich in erster Linie an interessierte Leser richtet und nicht klausurrelevant ist.

Die transponierte Matrix als lineare Abbildung

(Dieser Abschnitt ist nicht klausurrelevant und richtet sich in erster Linie an interessierte Leser.)

Falls V und W K-Vektorräume sind, und $f : V \to W$ eine lineare Abbildung ist, so gibt es eine natürliche lineare Abbildung

$$f^\vee : \mathrm{Hom}_K(W, K) \longrightarrow \mathrm{Hom}_K(V, K), \quad \phi \mapsto \phi \circ f,$$

die wir duale Abbildung zu f nennen.

Sei nun $V = K^n$ und $W = K^m$. Dann induzieren nach Aufgabe 2 auf Blatt09 die Standardbasis von V und W natürliche Basen (duale Basen genannt) $\phi_1, \ldots, \phi_n$ von $\mathrm{Hom}_K(V, K)$ und $\psi_1, \ldots, \psi_m$ von $\mathrm{Hom}_K(W, K)$ die durch

$$\phi_i(e_j) = \delta_{ij} \quad \text{und} \quad \psi_l(e'_k) = \delta_{lk}$$

gegeben sind, wobei δ_{ij} und δ_{lk} das Kronecker Delta bezeichnet (d.h. $\delta_{ij} \in \{0, 1\}$ ist genau dann 1 wenn $i = j$ und analog für δ_{lk}). Die Wahl der duale Basen liefert Vektorraum Isomorphismen

$$\Phi : \mathrm{Hom}_K(V, K) \xrightarrow{\cong} K^n, \quad \sum_{i=1}^{n} \lambda_i \phi_i \mapsto \sum_{i=1}^{n} \lambda_i e_i \tag{4.1}$$

und

$$\Psi : K^m \xrightarrow{\cong} \mathrm{Hom}_K(W, K), \quad \sum_{j=1}^{m} \mu_j e'_j \mapsto \sum_{j=1}^{m} \mu_j \psi_j, \tag{4.2}$$

wobei $e_1, \ldots, e_n$ die Standardbasis von K^n und $e'_1, \ldots, e'_m$ die Standardbasis von K^m ist.

Lemma 4.2.10. *Sei $A = (a_{ji}) \in K^{m \times n}$ eine $m \times n$ Matrix mit zugehöriger linearer Abbildung $h_A : V \to W$, wobei $V := K^n$ und $W := K^m$. Betrachte die Komposition*

$$K^m \xrightarrow{\Psi} \mathrm{Hom}_K(W, K) \xrightarrow{h_A^\vee} \mathrm{Hom}_K(V, K) \xrightarrow{\Phi} K^n$$

wobei $h_A^\vee$ die zu h_A duale Abbildung ist und Φ bzw. Ψ die Isomorphismen in (4.1)bzw. (4.2) sind. Dann ist die obige Komposition eine lineare Abbildung $h^\vee : K^m \to K^n$ deren darstellende Matrix bezüglich der Standardbasen die zu A transponierte Matrix $A^t \in K^{n \times m}$ ist.

Beweis. Sicherlich ist $h^\vee = \Phi \circ h_A^\vee \circ \Psi$ als Komposition linearer Abbildungen wieder linear. Weiterhin gilt:
$$h^\vee(e'_j) = \Phi(h_A^\vee(\psi_j)).$$
Hier ist $h_A^\vee(\psi_j) \in \mathrm{Hom}(V, K)$ die lineare Abbildung
$$V \longrightarrow K, \quad v \mapsto \psi_j(h_A(v)) = \psi_j(A \cdot v)$$
wobei wir daran erinnern, dass $V = K^n$ gilt und $h_A(v) = A \cdot v$ gilt. Da $\phi_1, \dots, \phi_n$ eine Basis von $\mathrm{Hom}_K(V, K)$ ist, muss es Elemente $\lambda_1, \dots \lambda_n \in K$ geben, sodass
$$h_A^\vee(\psi_j) = \sum_{l=1}^n \lambda_l \phi_l$$
gilt. Der Koeffizient λ_i kann nun dadurch gefunden werden, indem man in obige Gleichung den Vektor e_i einsetzt. In der Tat erhält man dann
$$h_A^\vee(\psi_j)(e_i) = \sum_{l=1}^n \lambda_l \phi_l(e_i) = \sum_{l=1}^n \lambda_l \delta_{li} = \lambda_i.$$
Andererseits gilt
$$h_A^\vee(\psi_j)(e_i) = \psi_j(A \cdot e_i) = \psi_j(\sum_{k=1}^m a_{ki} e'_k) = \sum_{k=1}^m a_{ki} \delta_{j,k} = a_{ji}.$$
Also: $\lambda_i = a_{ji}$ und damit
$$h_A^\vee(\psi_j) = \sum_{i=1}^n a_{ji} \phi_i.$$
Setzt man das oben ein, erhält man
$$h^\vee(e'_j) = \Phi(h_A^\vee(\psi_j)) = \Phi(\sum_{i=1}^n a_{ji} \phi_i) = \sum_{i=1}^n a_{ji} e_i.$$
Das zeigt, dass die lineare Abbildung $h^\vee : K^m \to K^n$ bezüglich der Standardbasen von der Matrix
$$A^T = (a_{ij})_{\substack{i \in \{1, \dots, n\} \\ j \in \{1, \dots, m\}}}$$
gegeben ist. Das schließt den Beweis des Lemmas ab. □

4.3 Elementarmatrizen und Gaußalgorithmus

4.3.1 Elementare Zeilen- und Spaltenumformungen

Lemma 4.3.1. *Sei $A \in K^{m \times n}$ eine $m \times n$-Matrix. Der Rang von A ändert sich nicht durch folgende Umformungen:*

(S1) Zwei Spalten vertauschen;

(S2) Zu einer Spalte ein skalares Vielfaches einer anderen Spalte dazu addieren;

(S3) Eine Spalte mit $0 \neq \lambda \in K$ multiplizieren.

Beweis. Es ist offensichtlich, dass sich der von den Spalten aufgespannte Unterraum durch keine der drei Operationen ändert. Also ändert sich auch die Dimension von diesem Raum nicht, und damit also der Rang. □

Lemma 4.3.2. *Sei $A \in K^{m\times n}$ eine $m \times n$-Matrix. Der Rang von A ändert sich nicht durch folgende Umformungen:*

(Z1) Zwei Zeilen vertauschen;

(Z2) Zu einer Zeile ein skalares Vielfaches einer anderen Zeile dazu addieren;

(Z3) Eine Zeile mit $0 \neq \lambda \in K$ multiplizieren.

Beweis. Es ist offensichtlich, dass sich der von den Zeilen aufgespannte Unterraum durch keine der drei Operationen ändert. Also ändert sich auch die Dimension von diesem Raum nicht, und damit also der Zeilenrang. Da wir schon gezeigt haben, dass der Zeilen und Spalten-rang übereinstimmt, folgt die Aussage. □

Obige beiden Lemmata sind rechnerisch äußerst nützlich um den Rang einer Matrix zu bestimmen.

Beispiel. Elementare Zeilenumformungen liefert

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -12 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Wendet man nun noch elementare Spaltenumformungen an, erhalten wir:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Also insgesamt:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und somit hat obige Matrix A den Rang 2.

Satz 4.3.1. *Sei $A \in K^{m\times n}$. Durch hintereinander ausführen von elementaren Zeilen- und Spaltenumformungen, kann man A in die Form*

$$\begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

wobei $r = \operatorname{Rang}(A)$ ist, wie in Satz 4.2.1 bringen.

Beweis. Nach dem Gaußalgorithmus, den wir schon gesehen haben, kann man A durch elementare Zeilenumformungen in sogenannte Zeilen-Stufenform bringen. Wendet man auf das Ergebnis nun geeignete Spaltenumformungen an, so erhält man die gewünschte Form. □

4.3.2 Elementarmatrizen

Definition 4.3.3. *Sie n eine positive natürliche Zahl und sei K ein Körper. Für $i, j \in \{1, 2, \ldots, n\}$ sei $E_{ij} \in K^{n\times n}$ die Matrix deren Einträge alle Null sind, bis auf den Eintrag an der Stelle (i, j), welcher 1 sei.*

Mit obiger Definition gilt zum Beispiel $\mathbb{1}_n = \sum_{i=1}^n E_{ii}$.

Lemma 4.3.4. *Sei n eine positive natürliche Zahl und sei K ein Körper. Die Zeilenumformungen aus Lemma 4.3.2 (bzw Spaltenumformungen aus Lemma 4.3.1) sind durch Multiplikation von links (bzw rechts) mit folgenden Matrizen gegeben, die wir Elementarmatrizen nennen:*

(1) Vertauschung der i-ten und j-ten Zeile/Spalte:

$$V_{ij} := \mathbb{1}_n + E_{ij} + E_{ji} - E_{ii} - E_{jj};$$

(2) Zur i-ten Zeile/Spalte λ-Mal j-te Zeile /Spalte addieren (wobei $i \neq j$):

$$A_{ij}(\lambda) := \mathbb{1}_n + \lambda E_{ij};$$

(3) Multiplikation der i-ten Zeile/Spalte mit $0 \neq \lambda \in K$:

$$M_i(\lambda) := \mathbb{1}_n + (\lambda - 1)E_{ii}.$$

Beweis. Das rechnet man direkt nach. □

Lemma 4.3.5. *Elementarmatrizen sind invertierbar und die Inversen sind wieder Elementarmatrizen.*

Beweis. Das folgt daraus, dass man diejeweilige Zeilen oder Spaltenumformung rückgängig machen kann. Genauer gilt:

$$M_i(\lambda) \cdot M_i(1/\lambda) = M_i(1/\lambda) \cdot M_i(\lambda) = \mathbb{1}_n, \quad V_{ij} \cdot V_{ij} = \mathbb{1}_n$$

und

$$A_{ij}(\lambda) \cdot A_{ij}(-\lambda) = A_{ij}(-\lambda) \cdot A_{ij}(\lambda) = \mathbb{1}_n.$$

□

Korollar 4.3.6. *Sei $A \in K^{m\times n}$. Dann gibt es Matrizen $T \in \mathrm{GL}(n, K)$ und $S \in \mathrm{GL}(m, K)$ die jeweils explizit als Produkte von Elementarmatrizen gegeben sind, und so dass*

$$TAS = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $r = \mathrm{Rang}(A)$ ist.

Beweis. Das folgt direkt aus Lemma 4.3.4 und Satz 4.3.1. □

Man beachte dass eine Matrix $A \in K^{n\times n}$ genau dann invertierbar ist, wenn $\mathrm{Rang}(A) = n$ gilt. In diesem Fall verwandeln nach Satz 4.3.1 geeignete Zeilen und Spaltenumformungen A in die Einheitsmatrix.

Satz 4.3.2. *Sei $A \in K^{n\times n}$ invertierbar. Wendet man nun auf die Einheitsmatrix $\mathbb{1}_n$ genau die Zeilenumformungen an die nötig sind um A in die Einheitsmatrix zu verwandeln, so verwandelt sich $\mathbb{1}_n$ in das Inverse A^{-1} zu A.*

Beweis. Das Verfahren verwandelt A in TA und $\mathbb{1}_n$ in T, wobei T ein Produkt von Elementarmatrizen ist. Da

$$TA = \mathbb{1}_n$$

gilt, haben wir

$$A = T^{-1}.$$

Das beweist die Behauptung. □

Bemerkung 4.3.7. *Obiger Satz ist in der Praxis sehr nützlich um das Inverse zu einer invertierbaren Matrix $A \in K^{n\times n}$ zu finden. Dazu muss man a priori noch nicht einmal wissen, ob A invertierbar ist. Man beginnt mit A und wendet Zeilenumformungen an, bis A eine obere Dreiecksmatrix wird, also eine matrix deren Einträge unter der Diagonalen Null sind. Wenn dann die Einträge auf der Diagonalen nicht all von Null verschieden sind, so ist A nicht invertierbar. Wenn abber die Diagonaleinträge alle von Null verschieden sind, so kann man weitere Zeilenumformungen anwenden bis man die Einheitsmatrix erhält.*

Beispiel. Betrachte die Matrix $A \in \mathbb{R}^{2\times 2}$ mit

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}.$$

Wir führen dann wie folgt Zeilenumformungen aus, um A in die Einheitsmatrix zu verwandeln:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -3 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & -1 & -3 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & 1 & 3 & -1 \end{array}\right).$$

Also gilt:

$$A^{-1} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}.$$

Bemerkung 4.3.8. *In obigem Verfahren kann man anstatt Zeilen- auch Spaltenumformungen durchführen. Es ist aber wichtig, dass man nicht Zeilen- mit Spaltenumformungen vermischt. In der Tat, man kann zwar Zeilen und Spaltenumformungen anwenden um aus A die Matrix $TAS = \mathbb{1}_n$ zu erhalten und gleichzeitig $\mathbb{1}_n$ in TS zu verwandeln. Dann gilt aber $A = T^{-1}S^{-1}$ und damit $A^{-1} = ST$, was im Allgemeinen von TS verschieden ist.*

Korollar 4.3.9. *Jedes Element in* $\mathrm{GL}(n, K)$ *ist ein Produkt von Elementarmatrizen.*

Beweis. Sei $A \in K^{n\times n}$ eine invertierbare Matrix. Dann gilt: $\mathrm{Rang}(A) = n$ und damit nach obigem Satz

$$TA = \mathbb{1}_n,$$

wobei T ein Produkt von Elementarmatrizen ist. Dann ist auch T^{-1} ein Produkte von Elementarmatrizen und es gilt:

$$A = T^{-1}.$$

Das beweist das Korollar. □

4.4 Lineare Gleichungssysteme

Definition 4.4.1. *Sei K ein Körper, $A = (a_{ji})_{\substack{j=1,\dots,m\\ i=1,\dots,n}} \in K^{m\times n}$ eine $m \times n$-Matrix und*

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in K^m$$

ein Spaltenvektor der Länge m. Das Gleichungssystem

$$\sum_{i=1}^{n} a_{ji}x_i = b_j \qquad j = 1,\dots,m$$

heißt lineares Gleichungssystem über K. Wir schreiben dafür auch $Ax = b$, wobei

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

der Vektor der Unbekannten ist. Die Lösungsmenge (oder der Lösungsraum) L des linearen Gleichungssystems $Ax = b$ besteht aus allen Vektoren $v \in K^n$ mit $Av = b$.

Bemerkung 4.4.2. *Ein lineares Gleichungssystem $Ax = b$ heißt homogen, falls $b = 0$ gilt.*

Die Matrix $A \in K^{m\times n}$ hat eine zugehörige lineare Abbildung

$$h_A : K^n \longrightarrow K^m, \quad v \mapsto A \cdot v.$$

Die Lösungsmenge L des Gleichungssystems $Ax = b$ ist also genau das Urbild des Vektors $b \in K^m$:

$$L = h_A^{-1}(b).$$

Satz 4.4.1. *Sei K ein Körper und seien $A \in K^{m\times n}$ und $b \in K^m$ gegeben. Sei L die Lösungsmenge des linearen Gleichungssystems $Ax = b$. Sei weiterhin $r = \mathrm{Rang}(A)$. Dann gilt:*

(a) $L \neq \emptyset \Leftrightarrow$ die Matrix $(A|b) \in K^{m\times(n+1)}$ hat Rang r.

(b) Falls $v_0 \in L$ (also $L \neq \emptyset$), so gibt es einen linearen Unterraum $U \subset K^n$ der Dimension $m - r$, sodass

$$L = U + v_0 = \{u + v_0 \mid u \in U\}.$$

(c) Der lineare Unterraum $U \subset K^n$ aus (b) ist die Lösungsmenge des homogenen Gleichungssystems $Ax = 0$. Insbesondere hängt U also nicht weder von der Wahl der Lösung $v_0 \in L$, noch von $b \in K^m$, sondern nur von A ab.

Beweis. Betrachte die zu A zugehörige Abbildung $h_A : K^n \to K^m$. Wir haben dann gesehen, dass $L = h_A^{-1}(b)$ gilt. Also ist L genau dann nicht-leer, wenn $b \in \mathrm{Bild}(h_A)$ gilt. Letzteres ist dazu Äquivalent, dass b eine Linearkombination der Spalten von A ist. Und somit ist die Bedingung dazu äquivalent, dass

$$\mathrm{Rang}(A|b) = \mathrm{Rang}(A)$$

gilt. Da nach Voraussetzung $r = \mathrm{Rang}(A)$ gilt, folgt Teil (a).

Um (b) und (c) zu zeigen, nehmen wir nun an, dass es eine Lösung $v_0 \in L$ gibt. Dann gilt also $h_A(v_0) = b$. Falls $v \in h_A^{-1}(0)$, so gilt dann wegen Linearität:

$$h_A(v + v_0) = h_V(v) + h_A(v_0) = b.$$

Umgekehrt gilt für jedes $v_1 \in L$:

$$h_A(v_1 - v_0) = h_A(v_1) - h_A(v_0) = b - b = 0.$$

Diese beiden Eigenschaften beweisen folgende Identität:

$$h_A^{-1}(b) = h_A^{-1}(0) + v_0.$$

Also,

$$L = U + v_0,$$

wobei $U = h_A^{-1}(0) = \mathrm{Ker}(h_A)$ der Kern von h_A ist. Dies ist ein linearer Unterraum. Nach der Dimensionsformel gilt:

$$\dim(\mathrm{Ker}(h_A)) + \dim(\mathrm{Bild}(h_A)) = m.$$

Da $r = \mathrm{Rang}(A)$ per Definition die Dimension von $\mathrm{Bild}(h_A)$ ist, gilt also:

$$\dim(U) = m - r,$$

wie behauptet. □

Wie berechnet man die Lösungen von $Ax = b$ explizit?

- Strategie 1: Rate eine Lösung und berechne alle Lösungen von $Ax = 0$ via dem Gaußalgorithmus.
- Strategie 2: Wende den Gaußalgorithmus direkt auf $Ax = b$, wie folgt an.

Lemma 4.4.3. *Sei K ein Körper und sei $Ax = b$ mit $A \in K^{m\times n}$ und $b \in K^m$ eine lineares Gleichungssystem. Betrachte die Matrix $(A|b) \in K^{m\times(n+1)}$ und wende eine elementare Zeilenumformung an:*

$$(A|b) \rightsquigarrow (A'|b').$$

Dann haben die Gleichungssysteme $Ax = b$ und $A'x = b'$ dieselben Lösungsmengen.

Beweis. Übung. □

Obiges Lemma zeigt, dass man den Gaußalgorithmus anwenden kann, um die Lösungsmenge eines Gleichungssystems $Ax = b$ zu finden.

Beispiel. Betrachte das lineare Gelcihungssystem $Ax = b$, wobei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathbb{R}^{3\times 3} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3.$$

Wir wenden nun den Gaußalgorithmus auf $(A|b) \in \mathbb{R}^{3\times 4}$ an:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & -6 & -12 & -4 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 2/3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Nach einer weiteren elementaren Zeilenumformung erhalten wir:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & -1/3 \\ 0 & 1 & 2 & 2/3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Also hat das lineare Gleichungssystem $Ax = b$ dieselbe Lösungsmenge wie das Gleichungssystem

$$x_1 - x_3 = -1/3$$

$$x_2 + 2x_3 = 2/3.$$

Also ist die Lösungsmenge L nicht-leer und wir haben

$$L = \begin{pmatrix} -1/3 \\ 2/3 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \cdot \mathbb{R}.$$

Kapitel 5

Determinanten

5.1 Determinante von 2×2 Matrizen

Wir beginnen mit der folgenden provisorischen Definition einer Determinantenfunktion auf der Menge aller 2×2-Matrizen

$$\det : K^{2\times 2} \longrightarrow K, \qquad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \det A := ad - bc$$

Die oben definierte Determinante hat folgende wichtige Eigenschaften, für mehr Details, siehe [Fi14, §3.1].

Eigenschaft 1.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0 \quad \Longleftrightarrow \quad \operatorname{Rang} \begin{pmatrix} a & b \\ c & d \end{pmatrix} < 2.$$

Also: $\det A = 0$ genau dann wenn die Spaltenvektoren linear abhängig sind.

Beweis. Falls $\det A = 0$ gilt, so haben wir $ad = bc$ und damit

$$d \cdot \begin{pmatrix} a \\ c \end{pmatrix} = c \cdot \begin{pmatrix} b \\ d \end{pmatrix}$$

Also sind die Spalten von A linear abhängig, es sei denn $d = c = 0$. Aber falls $d = c = 0$, so sind die Spalten von A offensichtlich ebenfalls linear abhängig. Dies zeigt, die Richtung $\Rightarrow$ in Eigenschaft 1.

Für die Umkehrung, nehmen wir an, dass die Spalten linear abhängig sind. Aber dann ist eine Spalte ein Vielfaches der anderen Spalte und das zeigt direkt, dass $\det A = 0$ gilt. Dies beweist Eigenschaft 1. □

Eigenschaft 2. Sei nun $K = \mathbb{R}$ und sei $A = (v_1|v_2)$ eine reelle 2×2-Matrix mit den Spaltenvektoren v_1 und v_2:

$$v_1 = \begin{pmatrix} a \\ c \end{pmatrix}, \quad v_2 = \begin{pmatrix} b \\ d \end{pmatrix} \quad \text{und} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Betrachte das von v_1 und v_2 aufgespannte Parallelogramm $P = \{tv_1 + sv_2 \mid t, s \in [0, 1]\}$. Wir nehmen an, dass Die Ecken von P, also die Vektoren 0, v_1, $v_1 + v_2$ und v_2 in dieser Reihenfolge gegen den Urzeigersinn durchlaufen werden. Dann ist der Flächeninhalt $|P|$ von P gegeben durch

$$|P| = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Beweis. Sowohl der Flächeninhalt von P, als auch der Ausdruck $\det A$ verhalten sich linear unter Skalierung der Vektoren v_1 und v_2. Wir können deshalb annehmen, dass v_1 und v_2 die Länge 1 haben, d.h. $a^2 + c^2 = 1$ und $b^2 + d^2 = 1$. Wir nehmen nun der Einfachheit halber an, dass v_1 und v_2 beide im ersten Quadranten liegen. Sei dann α der Winkel der von der x-Achse und dem Strahl $\mathbb{R}_{\geq 0} v_1$ aufgespannt wird. Sei weiterhin β der Winkel der von der x-Achse und dem Strahl $\mathbb{R}_{\geq 0} v_2$ aufgespannt wird. Dann gilt: Das Parallelogramm P hat eine Seite der Länge 1 und die zugehörige Höhe ist durch $h = \sin(\beta - \alpha)$ gegeben. Der Flächeninhalt ist durch Länge mal Höhe gegeben, sodass wir folgendes erhalten:

$$|P| = \sin(\beta - \alpha).$$

Nach dem Additionssatz für den Sinus, den wir hier als bekannt annehmen wollen, gilt aber dann:

$$|P| = \sin(\beta - \alpha) = \cos\alpha \cdot \sin\beta - \cos\beta \cdot \sin\alpha.$$

Da v_1 und v_2 Länge 1 haben, gilt andererseits:

$$v_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}.$$

Also:

$$\det A = \cos\alpha \cdot \sin\beta - \sin\alpha \cdot \cos\beta.$$

Das beweist $|P| = \det A$, wie gewünscht. □

Bemerkung 5.1.1. *Eigenschaft 2 erklärt Eigenschaft 1 geometrisch: das von v_1 und v_2 aufgespannte Parallelogramm hat genau dann den Flächeninhalt 0, wenn v_1 und v_2 linear abhängig sind, sodass das Parallelogramm entartet ist.*

5.2 Determinanten: axiomatischer Zugang

Wir wollen die Determinante von 2×2-Matrizen die wir in vorherigem Abschnitt besprochen haben auf $n \times n$-Matrizen für beliebiges $n \geq 1$ verallgemeinern. Direkt eine passende Formel anzuschreiben ist gar nicht so einfach. Die Idee ist stattdessen wichtige Eigenschaften der Determinante im Fall von 2×2-Matrizen als Axiome zu verwenden und davon ausgehend so viele Eigenschaften abzuleiten, dass man automatisch auf die richtige Formel stößt.

Definition 5.2.1. *Sei K ein Körper und n eine von Null verschiedene natürliche Zahl. Eine Abbildung*

$$\det : K^{n \times n} \longrightarrow K, \quad A \mapsto \det A$$

heißt Determinante oder Determinantenfunktion, falls folgende Eigenschaften gelten:

(D1) det *ist multi-linear, d.h. linear in jeder Zeile;*

(D2) det *ist alternierend, d.h.* $\det A = 0$ *falls A zwei gleiche Zeilen hat;*

(D3) det *ist normiert, d.h.* $\det \mathbb{1}_n = 1$.

Bemerkung 5.2.2. *Offensichtlich erfüllt die Determinante von 2×2-Matrizen, die wir in vorherigem Abschnitt betrachtet haben obige Axiome. Allgemein werden wir sehen, dass es für beliebiges n genau eine Determinante mit obigen Eigenschaften gibt. Dies erfordert allerdings ein wenig Arbeit.*

Bemerkung 5.2.3. *Die Eigenschaft, dass* det *linear in den Zeilen ist impliziert nicht, dass* $\det(A + A') = \det A + \det A'$ *gilt. In der Tat ist das bereits im Fall von 2×2-Matrizen im Allgemeinen falsch.*

Satz 5.2.1. *Sei* $\det : K^{n\times n} \to K$ *eine Determinantenfunktion. Dann gelten folgende weitere Eigenschaften:*

(D4) *Für $\lambda \in K$ gilt* $\det(\lambda \cdot A) = \lambda^n \det A$;

(D5) *Ist eine Zeile von A gleich Null, so gilt* $\det A = 0$;

(D6) *Sei $\lambda \in K$ und $i \neq j$. Entsteht B aus A durch Addition der λ-fachen j-ten Zeile zur i-ten Zeile, so ist* $\det B = \det A$;

(D7) *Entsteht B aus A durch eine Zeilenverstauschung, so gilt* $\det A = -\det B$;

(D8) *Ist A eine obere Dreiecksmatrix, also*

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * & * \\ 0 & \lambda_2 & \cdots & * & * \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

so gilt: $\det A = \lambda_1\lambda_2 \cdots \lambda_n$;

(D9) *Falls A von der Gestalt*

$$A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$$

für quadratische Matrizen A_1 und A_2 ist, so gilt $\det A = \det A_1 \cdot \det A_2$;

(D10) $\det A = 0 \iff \mathrm{Rang}(A) < n$;

(D11) $\det(A \cdot B) = \det A \cdot \det B$.

Beweis. (D4) und (D5) folgen direkt aus der Multilinearität (D1).

(D6): Seien $z_1, \ldots, z_n$ die Zeilen von A. Dann gilt:

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix} + \lambda \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_j \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix} + \det \begin{pmatrix} z_1 \\ \vdots \\ \lambda \cdot z_j \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ z_i + \lambda z_j \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix},$$

wobei wir im ersten Gleichheitszeichen (D2) und in den beiden letzten Gleichheitszeichen (D1) ausnutzen. Das beweist (D6).

(D7): Multilinearität, d.h. (D1), impliziert:

$$\det\begin{pmatrix} z_1 \\ \vdots \\ z_i + z_j \\ \vdots \\ z_i + z_j \\ \vdots \\ v_n \end{pmatrix} = \det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix} + \det\begin{pmatrix} z_1 \\ \vdots \\ z_j \\ \vdots \\ z_i \\ \vdots \\ v_n \end{pmatrix} + \det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_i \\ \vdots \\ v_n \end{pmatrix} + \det\begin{pmatrix} z_1 \\ \vdots \\ z_j \\ \vdots \\ z_j \\ \vdots \\ v_n \end{pmatrix}$$

Sowohl der Term auf der linken Seite, als auch die letzten beiden Terme auf der rechten Seite verschwinden nach (D2), sodass (D7) folgt.

(D8): Angenommen $\lambda_i \neq 0$ für alle i. Aufgrund von Multilinearität gilt dann

$$\det\begin{pmatrix} \lambda_1 & * & \cdots & * & * \\ 0 & \lambda_2 & \cdots & * & * \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & * \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \cdot \det\begin{pmatrix} 1 & *' & \cdots & *' & *' \\ 0 & 1 & \cdots & *' & *' \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & *' \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

wobei $*'$ nicht näher spezifizierte Einträge sind, die i.A. nicht mit $*$ übereinstimmen. Nach (D6) und dem Gauß-Algorithmus gilt:

$$\det\begin{pmatrix} 1 & *' & \cdots & *' & *' \\ 0 & 1 & \cdots & *' & *' \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & *' \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = \det \mathbb{1}_n.$$

Die Behauptung folgt nun aus (D3) (d.h. aus $\det \mathbb{1}_n = 1$).

Falls aber $\lambda_i = 0$ für ein i, so wählen wir i maximal mit dieser Eigenschaft. Dann gilt $\lambda_{i+1}, \ldots, \lambda_n \neq 0$ und wir können Zeilenumformungen der Art (Z2) anwenden um jeden Eintrag der i-ten Zeile zu einer Null zu verwandeln. Nach (D7) ändert sich dadurch die Determinante nicht. Wir können also annehmen, dass die i-te Zeile von A verschwindet und damit gilt $\det A = 0 \cdot \det A = 0$ nach (D1), angewendet auf die i-te Zeile.

(D9): Der Beweis ist ähnlich wie (D10) und (D11) unten und wir skizzieren an dieser Stelle das Argument nur: Nach Anwenden der elementaren Zeilenumformungen (Z1) und (Z2) (d.h. vertauschen und ein Vielfaches einer Zeile zu einer anderen addieren), können wir mit Hilfe des Gaußalgorithmus und wegen (D6) und (D7) annehmen, dass A_1 und A_2 obere Dreiecksmatrizen sind. Die Aussage folgt dann aus (D8).

(D10): Nach Anwenden des Gaußalgorithmus existiert ein Produkt $T \in \mathrm{GL}(n, K)$ von Elemntarmatrizen V_{ij} und $A_{ij}(\lambda)$, sodass

$$T \cdot A = \begin{pmatrix} \mathbb{1}_r & * \\ 0 & 0 \end{pmatrix}$$

gilt. Sei m die Anzahl der Vertauschungsmatrizen, die in dem Produkt T vorkommen. Dann gilt nach (D6) und (D7):

$$\det(A) = (-1)^m \det(TA).$$

Da $\mathrm{Rang}(A) = \mathrm{Rang}(TA)$ gilt, genügt es die Aussage für die Matrix TA zu zeigen. In diesem Fall folgt (D10) aber direkt aus (D8).

(D11): Falls A nicht invertierbar ist, so gilt $\operatorname{Rang} A < n$ und damit auch $\operatorname{Rang}(A \cdot B) < n$ (denn $h_{A \cdot B} = h_A \circ h_B$ und damit $\operatorname{Bild}(h_{AB}) \subset \operatorname{Bild}(h_A)$). Also stimmt die Formel, falls A nicht invertierbar ist. Falls A aber invertierbar ist, so gibt es nach dem Gauß-Algorithmus eine Matrix $T \in \mathrm{GL}(n, K)$ welche ein Produkt von Elementarmatrizen V_{ij} und $A_{ij}(\lambda)$ ist, sodass TA eine Diagonalmatrix ist, d.h. dass $\lambda_1, \ldots, \lambda_n \in K$ existieren, sodass

$$T \cdot A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Sei m die Anzahl der Vertauschungen, die in T vorkommt. Dann gilt nach (D6) und (D7):

$$\det(AB) = (-1)^m \det(TAB).$$

Da TA eine Diagonalmatrix mit den Diagonal-Einträgen $\lambda_1, \ldots, \lambda_n$ ist, ist das Produkt TAB dadurch gegeben, indem man mit B startet und für alle $i = 1, \ldots, n$, die i-te Zeile von B mit λ_i multipliziert. Aufgrund der Multilinearität (siehe (D1)) gilt dann aber:

$$\det(AB) = (-1)^m \det(TAB) = (-1)^m \lambda_1 \cdots \lambda_n \cdot \det B.$$

Andererseits gilt:

$$\det(A) = (-1)^m \det(TA) = (-1)^m \lambda_1 \cdots \lambda_n$$

nach (D8). Das beweist

$$\det(AB) = \det(A) \cdot \det(B),$$

wie gewünscht. □

Korollar 5.2.4. *Sei* $\det : K^{n \times n} \to K$ *eine Determinantenfunktion. Dann gilt* $\det A^t = \det A$.

Beweis. Wir zeigen das Korollar zuerst in dem Fall dass A eine Elementarmatrix ist. Falls $A = V_{ij}$, so gilt $V_{ij}^t = V_{ij}$ und damit ist die Aussage klar. Falls $A = A_{ij}(\lambda)$ für $i \neq j$, dann $A^t = A_{ji}(\lambda)$ und damit folgt die Aussage aus (D6), was $\det(A_{ij}(\lambda)) = 1 = \det(A_{ji}(\lambda))$ impliziert. Falls schließlich $A = M_i(\lambda)$ gilt, so ist $A = A^t$ und die Aussage ist wieder klar.

Sei nun A eine beliebige Matrix. Da $rk(A) = rk(A^t)$ gilt, genügt es den Fall zu behandeln in dem $rkA = n$, also A invertierbar ist. Dann ist $A = E_1 \cdots E_m$ ein Produkt von Elementarmatrizen $E_1, \ldots, E_m$, siehe Korollar 4.3.9. Also gilt $A^t = E_m^t E_{m-1}^t \cdots E_1^t$ (beachte dass sich die Reihenfolge in Produkten umdreht, wenn man die Matrizen transponiert) und damit

$$\det A = \prod_{i=1}^m \det(E_i) \quad \text{und} \quad \det A^t = \prod_{i=1}^m \det(E_i^t)$$

wegen (D11). Das Korollar folgt damit aus der Tatsache, dass es im Fall von Elementarmatrizen stimmt, d.h. $\det(E_i) = \det(E_i^t)$ gilt für alle i, wie wir oben nachgerechnet haben. □

Korollar 5.2.5. *Falls eine Determinantenfunktion* $\det : K^{n \times n} \to K$ *existiert, so ist diese eindeutig bestimmt. Es kann also höchstens eine solche Funktion geben.*

Beweis. Nach (D10) genügt es zu zeigen, dass $\det(A)$ eindeutig bestimmt ist in dem Fall dass A invertierbar ist. Aber dann gibt es eine Matrix $T \in \mathrm{GL}(n, K)$ welche ein Produkt von Elementarmatrizen V_{ij} und $A_{ij}(\lambda)$ ist, sodass TA eine obere Dreicksmatrix mit Diagonaleinträgen $\lambda_1, \ldots, \lambda_n \in K$ ist. Dann gilt aber nach (D6), (D7) und (D8):

$$\det A = (-1)^m \lambda_1 \ldots \lambda_n, \tag{5.1}$$

also ist $\det A$ eindeutig bestimmt. □

Bemerkung 5.2.6. *Die Existenz einer Determinantenfunktion haben wir (außer im Falll $n \leq 2$) noch nicht nachgewiesen. Es ist verlockend obige Beobachtung zu verwenden, und die Formel (5.1) als Definition zu verwenden. Das Problem dabei ist, dass das Produkt T nicht eindeutig bestimmt ist und somit a priori unklar ist, weshalb eine Definition via (5.1) wohldefiniert sein sollte. Ein besonders anschauliches Beisiel ist gegeben wenn man eine Matrix A betrachtet die aus $\mathbb{1}_n$ durch vertauschen von Zeilen entsteht. Verschiedene Kompositionen von solchen Vertauschungen können zum selben Ergebnis führen, sodass die in T vorkommenden Elementarmatrizen nicht eindeutig bestimmt sind und insbesondere die Anzahl der Vertauschungen, die in der Formel (5.1) eingeht von Wahlen abhängt und somit nicht wohldefiniert ist. In diesem Spezialfall reduziert sich dann die Wohldefiniertheit auf das Problem, ob die Parität der Anzahl der Vertauschungen wohldefiniert ist und wir werden im nächsten Abschnitt sehen, dass dies der Fall ist.*

Sobald wir die Existenz von Determinanten nachgeprüft haben, erlaubt (5.1) aber immerhin eine gute Methode wie man Determinanten berechnen kann.

5.3 Symmetrische Gruppe

Definition 5.3.1. *Sei $n \geq 1$ eine natürliche Zahl. Die symmetrische Gruppe S_n ist die Menge aller bijektiven Abbildungen*

$$\sigma : \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}.$$

Lemma 5.3.2. *Die symmetrische Gruppe S_n ist eine Gruppe bezüglich Verknüpfung von Abbildungen. Das neutrale Element ist die Identitätsabbildung $\mathrm{id} \in S_n$. Das Inverse zu $\sigma \in S_n$ ist die Umkehrabbildung σ^{-1}.*

Beweis. Das ist klar. □

Bemerkung 5.3.3. *Im Allgemeinen ist S_n nicht abelsch. Präziser gilt: S_n ist genau dann abelsch wenn $n \leq 2$ gilt.*

Bemerkung 5.3.4. *S_n hat genau $n!$ viele Elemente, da es genau $n!$ viele bijektive Abbildungen $\sigma : \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}$ gibt. In der Tat, für $\sigma(1)$ hat man n, Möglichkeiten, für $\sigma(2)$ bleiben dann noch $n-1$, für $\sigma(3)$ noch $n-2$ und so weiter.*

Schreibweise: Zweizeilenform

Sei $\sigma \in S_n$, so schreiben wir

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(n) \end{pmatrix}.$$

In dieser Schreibweise gilt zum Beispiel folgende Identität in S_4:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}.$$

Definition 5.3.5. *Eine Transposition ist ein Elemente $\tau \in S_n$, sodass es Indices $1 \leq i < j \leq n$ mit folgender Eigenschaft gibt:*

$$\tau(j) = i, \quad \tau(i) = j \quad \textit{und} \quad \tau(k) = k \quad \forall k \notin \{i, j\}.$$

Wir bezeichnen eine solche Transposition mit dem Symbol $\tau = (i\ j)$.

Folgendes Element in S_4 ist zum Beispiel eine Transposition:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}.$$

Definition 5.3.6. *Sei $\sigma \in S_n$ und sei $I := \{1, 2, \ldots, n\}$. Die Anzahl der Fehlstände von σ ist die Mächtigkeit der folgenden Menge:*

$$\{(i,j) \in I^2 \mid i < j \quad und \quad \sigma(i) > \sigma(j)\}.$$

Im folgenden bezeichnet $\{1, -1\} \subset \mathbb{Q}^*$ die multiplikative Untergruppe mit den beiden Elementen 1 und -1.

Satz 5.3.1. *Sei $n \geq 1$. Für $\sigma \in S_n$ definieren wir*

$$\epsilon(\sigma) := \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Dann gilt:

(1) $\epsilon(\sigma) = (-1)^k$, wobei k die Anzahl der Fehlstände von σ ist.

(2) $\epsilon : S_n \longrightarrow \{1, -1\}$ ist ein Gruppenhomomorphismus, d.h. $\epsilon(\sigma \circ \tau) = \epsilon(\sigma) \cdot \epsilon(\tau)$.

(3) Falls τ eine Transposition ist, so gilt $\epsilon(\tau) = -1$.

Beweis. Falls i, j durch alle natürlichen Zahlen mit $1 \leq i < j \leq n$ laufen, so durchlaufen (i, j) und $\{\sigma(i), \sigma(j)\}$ genau die zweielementigen Teilmengen von $I = \{1, \ldots, n\}$. Also gilt

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

ist, so folgt direkt aus der Definition die Ungleichung $(-1)^k \cdot \epsilon(\sigma) > 0$, und damit $(-1)^k \cdot \epsilon(\sigma) = 1$, also $\epsilon(\sigma) = (-1)^k$. Das beweist (1).
Wir zeigen schließlich (2):

$$\epsilon(\sigma \circ \tau) = \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} = \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{1 \leq i < j \leq n} \frac{\tau(j) - \tau(i)}{j - i}.$$

Die Behauptung folgt damit aus

$$\epsilon(\tau) = \prod_{1 \leq i < j \leq n} \frac{\tau(j) - \tau(i)}{j - i}$$

und

$$\epsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)},$$

wobei die letzte Gleichheit folgendes ausnutzt: Falls i, j durch alle natürlichen Zahlen mit $1 \leq i < j \leq n$ laufen, so durchlaufen $\{i, j\}$ und $\{\tau(i), \tau(j)\}$ genau die zweielementigen Teilmengen von $I = \{1, \ldots, n\}$. Also können wir $\tau(i)$ durch i und $\tau(j)$ durch j substituieren. Hier entsteht a priori das Problem, dass für $i < j$, nicht unbedingt $\tau(i) < \tau(j)$ gelten muss. Dies führt allerdings nicht zu einem Vorzeichenfehler, da Folgendes gilt:

$$\frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} = \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)}.$$

Das schließt den Beweis von (2) ab.

Schließlich zeigen wir (3). Sei dazu $\tau = (i\ j)$ mit $1 \leq i < j \leq n$ eine beliebige Transposition. Um $\sigma(\tau)$ zu berechnen müssen wir die Anzahl der Fehlstände von τ berechnen. Das heißt wir müssen alle Paare (a, b) mit $1 \leq a < b \leq n$ mit der Eigenschaft $\tau(b) < \tau(a)$ zählen. Wir führen im folgenden eine Liste aller solcher Paare, also aller Fehlstände auf:

- Das Paar (i, j) ist ein Fehlstand;
- Für alle $i < h < j$ ist sowohl (i, h) als auch (h, j) ein Fehlstand.

Obige Liste ist vollständig, weil für $1 \leq a < b \leq n$ mit $a < i$ oder $b > j$ das Paar (a, b) kein Fehlstand für τ ist (denn in jenem Fall gilt $\tau(a) < \tau(b)$). In obiger Liste stehen genau $1 + 2(j - i - 1)$ viele Paare, also gilt nach (1):

$$\epsilon(\tau) = (-1)^{1+2(j-i-1)} = -1.$$

Das beweist (3). □

Definition 5.3.7. *Der Gruppenhomomorphismus ϵ aus obigem Satz heißt Signum:*

$$\mathrm{sign} : S_n \longrightarrow \{1, -1\}, \quad \sigma \mapsto \mathrm{sign}(\sigma) := \epsilon(\sigma).$$

Korollar 5.3.8. *Sei $\sigma \in S_n$. Falls $\sigma = \tau_1 \circ \cdots \circ \tau_m$ ein Produkt von m Transpositionen τ_i ist, so gilt: $(-1)^m$ hängt nur von σ ab.*

Beweis. Das folgt direkt aus

$$\mathrm{sign}(\sigma) = \prod_{i=1}^{m} \mathrm{sign}(\tau_i) = (-1)^m.$$

□

Bemerkung 5.3.9. *Man kann zeigen, dass jedes Element in S_n eine Komposition von Transpositionen ist. Wir skizzieren den Beweis hier. Sei dazu $\sigma \in S_n$. Falls $\sigma(i) = i$ für alle $i \in I = \{1, 2, \ldots, n\}$ gilt, so ist die Aussage trivial. Ansonsten sei $i_0 \in I$ minimal mit der Eigenschaft dass $\sigma(i_0) \neq i_0$ und $\sigma(i) = i$ für $i < i_0$ gilt. Sei dann $\tau_1 = (i_0\ \sigma(i_0))$ die Transposition welche i_0 und $\sigma(i_0)$ vertauscht. Die Permutation $\sigma_1 := \tau_1 \circ \sigma$ hat dann die Eigenschaft dass $\sigma_1(i) = i$ für alle $i \leq i_0$ gilt. Per Induktion nach i_0 können wir Annehmen dass $\sigma_1 = \tau_2 \circ \cdots \circ \tau_m$ für Transpositionen τ_i gilt. Da aber $\tau_1 = \tau_1^{-1}$ gilt, folgt dann*

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_m.$$

Das beweist die Behauptung, dass σ als Komposition von Transpositionen geschrieben werden kann. Man beachte aber, dass diese Darstellung nicht eindeutig ist (da $\tau \circ \tau = \mathrm{id}$ für jede Transposition gilt, kann man zB solche Kopositionen beliebig dazwischen schalten ohne das Ergebnis zu ändern).

Definition 5.3.10. *Der Kern des Gruppenhomomorphismus*

$$\mathrm{sign} : S_n \longrightarrow \{-1, 1\}$$

aus Satz 5.3.1 bezeichnen wir mit $A_n := \mathrm{Ker}(\mathrm{sign}) \subset S_n$, d.h.

$$A_n := \{\sigma \in S_n \mid \mathrm{sign}(\sigma) = 1\}.$$

Wir nennen A_n die alternierende Gruppe von Grad n.

Bemerkung 5.3.11. *Man rechnet direkt nach, dass der Kern eines beliebigen Gruppenhomomorphismuses eine Untergruppe ist. Insbesondere ist $A_n \subset S_n$ eine Untergruppe und damit selbst eine Gruppe.*

Korollar 5.3.12. *Sei $\tau \in S_n$ ein beliebiges Element mit $\operatorname{sign}(\tau) = -1$. Dann gilt:*

$$S_n = A_n \sqcup A_n\tau,$$

wobei $A_n\tau := \{\sigma \circ \tau \mid \sigma \in A_n\}$.

Beweis. Es gilt sicherlich

$$S_n = \operatorname{sign}^{-1}(1) \sqcup \operatorname{sign}^{-1}(-1).$$

Da $A_n = \operatorname{sign}^{-1}(1)$, genügt es $A_n\tau = \operatorname{sign}^{-1}(-1)$ zu zeigen. Einerseits ist aber $A_n\tau \subset \operatorname{sign}^{-1}(-1)$ klar, weil $\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau) = 1 \cdot (-1) = -1$ für alle $\sigma \in A_n$ gilt. Ist andererseits $\sigma' \in \operatorname{sign}^{-1}(-1)$, so gilt:

$$\operatorname{sign}(\sigma' \circ \tau^{-1}) = \operatorname{sign}(\sigma')\operatorname{sign}(\tau^{-1}) = \operatorname{sign}(\sigma')\operatorname{sign}(\tau)^{-1} = \operatorname{sign}(\sigma') \cdot (-1) = -1 \cdot (-1) = 1.$$

Also, $\sigma := \sigma' \circ \tau^{-1} \in A_n$ und damit $\sigma' = \sigma \circ \tau \in A_n\tau$. Das beweist $A_n\tau \supset \operatorname{sign}^{-1}(-1)$, und somit insgesamt $A_n\tau = \operatorname{sign}^{-1}(-1)$. Das schließt den Beweis des Korollars ab. □

Bemerkung 5.3.13. *Wir haben auf einem Übungszettel gesehen, dass für ein festes Element $\tau \in S_n$, die Abbildung*

$$S_n \longrightarrow S_n, \qquad \sigma \mapsto \sigma\tau$$

die durch Rechtstranslation gegeben ist bijektiv ist. Insbesondere haben A_n und $A_n\tau$ gleich viele Elemente. Da S_n genau $n!$ verschiedene Elemente hat, gilt also:

$$|A_n| = \frac{n!}{2}.$$

5.4 Leibnizsche Summenformel

Satz 5.4.1. *Sei K ein Körper und sei n eine positive natürliche Zahl. Dann ist eine Determinantenfunktion auf $K^{n\times n}$ durch die folgende Leibnizsche Formel gegeben:*

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

wobei $A = (a_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}}$.

Bemerkung 5.4.1. *Die Leibnizsche Summenformel hat für jedes Element aus S_n einen Summanden, also insgesamt $n!$ Summanden. Für $n = 2$ erhalten wir z.B.:*

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Das stimmt mit der bereits bekannten Formel für 2×2-Matrizen überein.

Beweis von Satz 5.4.1. Wir müssen nachweisen, dass die Leibnizsche Summenformel die Bedingungen (D1), (D2) und (D3) aus Definition 5.2.1 erfüllt.

(D1): Wir bezeichnen die i-te Zeile von A mit z_i. Also:

$$A = \begin{pmatrix} \vdots \\ z_{i-1} \\ z_i \\ z_{i+1} \\ \vdots \end{pmatrix}.$$

Sei nun $z_i = \lambda z_i' + \mu z_i''$ mit $\lambda, \mu \in K$. Dann gilt:

$$\begin{aligned}
\det \begin{pmatrix} \vdots \\ z_{i-1} \\ \lambda z_i' + \mu z_i'' \\ z_{i+1} \\ \vdots \end{pmatrix} &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} \cdots (\lambda a'_{i\sigma(i)} + \mu a''_{i\sigma(i)}) \cdots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot \lambda \cdot a_{1\sigma(1)} \cdots a'_{i\sigma(i)} \cdots a_{n\sigma(n)} + \\
&\quad \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot \mu \cdot a_{1\sigma(1)} \cdots a''_{i\sigma(i)} \cdots a_{n\sigma(n)} \\
&= \lambda \det \begin{pmatrix} \vdots \\ z_{i-1} \\ z_i' \\ z_{i+1} \\ \vdots \end{pmatrix} + \mu \det \begin{pmatrix} \vdots \\ z_{i-1} \\ z_i'' \\ z_{i+1} \\ \vdots \end{pmatrix}.
\end{aligned}$$

Das zeigt, dass det multilinear ist, also (D1) erfüllt.

(D2): Angenommen es gibt Indices $1 \leq i < j \leq n$ sodass die i-te und j-te Zeile von A übereinstimmen. Sei dann $\tau \in S_n$ die Transposition, die i und j vertauscht. Nach Korollar 5.3.12 gilt

$$S_n = A_n \sqcup A_n \tau.$$

Also:

$$\det A = \sum_{\sigma \in A_n} \operatorname{sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)} + \sum_{\sigma \in A_n} \operatorname{sign}(\sigma \circ \tau) \cdot a_{1,\sigma(\tau(1))} \cdot \dots \cdot a_{n,\sigma(\tau(n))}. \tag{5.2}$$

Da sign nach Satz 5.3.1 ein Gruppenhomomorphismus ist, gilt

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau) = -\operatorname{sign}(\sigma). \tag{5.3}$$

Da außerdem die i-te und j-te Zeile von A identisch ist, gilt $a_{ik} = a_{jk}$ für alle $1 \leq k \leq n$. Da τ die Transposition ist, welche i und j vertauscht, gilt $\tau(i) = j$, $\tau(j) = i$ und $\tau(k) = k$ sonst. Also

$$a_{k,\sigma(\tau(k))} = \begin{cases} a_{k,\sigma(k)}, & \text{falls } k \notin \{i, .j\}, \\ a_{i,\sigma(j)} = a_{j,\sigma(j)}, & \text{falls } k = i, \\ a_{j,\sigma(i)} = a_{i,\sigma(i)}, & \text{falls } k = j. \end{cases}$$

Insgesamt zeigt das:

$$a_{1,\sigma(\tau(1))} \cdot \dots \cdot a_{n,\sigma(\tau(n))} = a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}.$$

Zusammen mit (5.3) folgt damit, dass der Ausdruck in (5.2) verschwindet. Das beweist (D2), und schließt unseren Beweis des Satzes ab. $\square$

Kapitel 6

Diagonalisierbarkeit

6.1 Ähnliche Matrizen

Definition 6.1.1. *Sei K ein Körper und sei $n \geq 1$ eine natürlich Zahl. Zwei Matrizen $A, B \in K^{n\times n}$ heißen ähnlich, falls es eine invertierbare Matrix $S \in \mathrm{GL}(n, K)$ gibt, sodass*

$$B = S^{-1}AS$$

gilt.

Bemerkung 6.1.2. *Ähnlichkeit von Matrizen liefert eine Äquivalenczrelation auf der Menge aller $n \times n$-Matrizen. Diese Äquivalenzrelation darf aber nicht mit der Relation aus Definition 4.2.1 verwechselt werden. Zwei ähnliche Matrizen sind zwar äquivalent zu einander wie in Definition 4.2.1 aber die Umkehrung ist im Allgemeinen nicht richtig.*

Definition 6.1.3. *Sei K ein Körper. Eine Matrix $D \in K^{n\times n}$ heißt Diagonalmatrix, oder einfach diagonal, falls es Elemente $\lambda_1, \ldots, \lambda_n \in K$ gibt, sodass folgendes gilt:*

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

In diesem Fall schreiben wir auch $D = D(\lambda_1, \ldots, \lambda_n)$ oder $D = diag(\lambda_1, \ldots, \lambda_n)$.

Definition 6.1.4. *Eine Matrix $A \in K^{n\times n}$ heißt diagonalisierbar, falls A ähnlich zu einer Diagonalmatrix D ist:*

$$D = S^{-1}AS$$

für ein $S \in \mathrm{GL}(n, K)$.

Definition 6.1.5. *Sei $A \in K^{n\times n}$. Ein Vektor $v \in K^n$ mit $v \neq 0$ heißt Eigenvektor von A, falls $A \cdot v = \lambda v$ für ein $\lambda \in K$ gilt. Der Wert λ heißt dann Eigenwert von A.*

Satz 6.1.1. *Eine Matrix $A \in K^{n\times n}$ ist genau dann diagonalisierbar, wenn es eine Basis $v_1, \ldots, v_n$ von Eigenvektoren gibt.*

Beweis. Wir nehmen zuerst an, dass A diagonalisierbar ist, d.h. es gibt eine invertierbare Matrix S sodass $D = S^{-1}AS$ eine Diagonalmatrix ist: $D = diag(\lambda_1, \ldots, \lambda_n)$. Seien dann

$v_1, \ldots, v_n$ die Spalten von S, d.h. $S = (v_1|v_2|\ldots|v_n)$, und sei $e_1, \ldots, e_n$ die Standardbasis von K^n. Dann gilt

$$SDe_i = S\lambda_i e_i = \lambda_i v_i.$$

Andererseits gilt

$$ASe_i = A(Se_i) = Av_i.$$

Da $D = S^{-1}AS$ zu $SD = AS$ äquivalent ist, gilt also $Av_i = \lambda_i v_i$ und damit ist v_i ein Eigenvektor zum Eigenwert λ_i. (Beachten Sie hier, dass $v_i \neq 0$, da $v_1, \ldots, v_n$ die Spalten von S sind und S vollen Rang hat.) Dies zeigt eine Richtung im Satz.

Für die Umkehrung sei $v_1, \ldots, v_n$ eine Basis von Eigenvektoren von A. Das bedeutet, dass es Elemente $\lambda_1, \ldots, \lambda_n \in K$ mit $Av_i = \lambda_i v_i$ für alle $i = 1, \ldots, n$ gibt. Dann ist $S := (v_1|v_2|\ldots|v_n) \in K^{n\times n}$ eine Matrix von Rang n und damit invertierbar. Weiterhin gilt:

$$ASe_i = A(Se_i) = Av_i = \lambda_i v_i$$

und falls $D := D(\lambda_1, \ldots, \lambda_n)$, so gilt

$$DSe_i = Dv_i = \lambda_i v_i.$$

Also insgesamt

$$ASe_i = DSe_i$$

für alle i. Da für eine $n \times n$ Matrix M der Vektor $M \cdot e_i$ genau der i-te Spalten Vektor ist, folgt also, dass AS und DS dieselben Spalten haben, und damit übereinstimmen. Also, $AS = SD$ bzw $D = S^{-1}AS$, wie gewünscht. □

Bemerkung 6.1.6. *Der Beweis zeigt: die Spalten von S sind die Eigenvektoren $v_1, \ldots, v_n$ und die Einträge auf der Diagonalen von D sind die zugehörigen Eigenwerte von A.*

Frage: Warum ist das nützlich?

Falls A diagonalisierbar ist, also

$$A = S \cdot D(\lambda_1, \ldots, \lambda_n) \cdot S^{-1}$$

gilt, so gilt

$$A^m = S \cdot D(\lambda_1^m, \ldots, \lambda_n^m) \cdot S^{-1}$$

für alle $m \geq 1$. Man kann also Potenzen von A sehr einfach und schnell ausrechnen. Das ist nützlich wenn man ein dynamisches System verstehen will, also verstehen will, was im Limes $n \to \infty$ passiert wenn man eine lineare Abbildung n-Mal hintereinander ausführt.

Neben der Tatsache, dass man mit diagonalisierbaren Matrizen einfacher rechnen kann, ist auch die geometrische Interpretation besonders anschaulich: Falls A diagonalisierbar ist, so gibt es nach obigem Satz eine Basis von Eigenvektoren und das bedeutet, dass die lineare Abbildung $h_A : K^n \to K^n$ einfach dadurch bestimmt ist, dass man eine Basis $v_1, \ldots, v_n$ von K^n hat und h_A jeden dieser Basisvektoren um ein bestimmtes Skalar skaliert, also $h_A(v_i) = \lambda_i v_i$.

6.2 Das charakteristische Polynom

Frage: Wie überprüft man in der Praxis, ob eine Matrix $A \in K^{n\times n}$ diagonalisierbar ist.

Definition 6.2.1. *Sei $A \in K^{n\times n}$. Die Funktion*

$$\chi_A : K \longrightarrow K, \qquad \chi_A(x) := \det(x \cdot \mathbb{1}_n - A)$$

heißt charakteristische Polynomfunktion.

Bemerkung 6.2.2. *Es gilt*

$$\chi_A(x) = \det \begin{pmatrix} x - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{31} & -a_{32} & x - a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ -a_{n1} & -a_{n2} & x - a_{n3} & \cdots & x - a_{nn} \end{pmatrix}.$$

Nach der Leibnizschen Summenformel ist obige Determinante eine Summe, wobei jeder Summand ein Produkt von genau n Einträgen der Matrix $x\mathbb{1}_n - A$ ist. Daraus folgt, dass $\chi_A(x)$ ein Polynom von Grad n über K ist. Der höchste Potenz x^n von x tritt auf, wenn man die Diagonaleinträge von $x\mathbb{1}_n - A$ miteinander multipliziert. Das entspricht in der Leibnizformel dem Summanden mit $\sigma = \mathrm{id} \in S_n$. Da $\mathrm{sign}(\mathrm{id}) = 1$ gilt, folgt, dass der Leitkoeffizient des Polynoms $\chi_A(x)$ x^n ist. Das heißt:

$$\chi_A(x) = x^n + c_1 x^{n-1} + \cdots + c_n$$

für Konstanten $c_1, \ldots, c_n \in K$. Da $\chi_A(0) = \det(-A) = (-1)^n \det A$ gilt, haben wir hier: $c_n = (-1)^n \det A$. Ähnlich rechnet man nach, dass

$$c_1 = a_{11} + a_{22} + a_{33} + \cdots + a_{nn}$$

gilt. Diesen Ausdruck, also die Summe aller Diagonaleinträge, nennt man die Spur von A:

$$\mathrm{Spur}(A) := \sum_i a_{ii}.$$

Lemma 6.2.3. *Sind $A, B \in K^{n\times n}$ zueinander ähnlich, so gilt $\chi_A(x) = \chi_B(x)$.*

Beweis. Nach Annahme gibt es $S \in \mathrm{GL}(n, K)$ mit $B = S^{-1}AS$. Für $x \in K$ rechnet man nach:

$$S^{-1}(x \cdot \mathbb{1}_n - A)S = x \cdot (S^{-1}\mathbb{1}_n S) - S^{-1}AS = x \cdot \mathbb{1}_n - B.$$

Also:

$$\chi_B(x) = \det(x \cdot \mathbb{1}_n - B) = \det(S^{-1}(x \cdot \mathbb{1}_n - A)S) = \det(S^{-1}) \det(x \cdot \mathbb{1}_n - A) \det(S) = \chi_A(x)$$

da $\det(S^{-1}) = \det(S)^{-1}$. □

Lemma 6.2.4. *Sei $A \in K^{n\times n}$. Ein Element $\lambda \in K$ ist genau dann Eigenwert von A, falls $\lambda \cdot \mathbb{1}_n - A$ nicht-trivialen Kern hat, oder äquivalent: nicht invertierbar ist.*

Beweis. Falls λ ein Eigenwert ist, so gibt es einen Vektor $v \in K^n$ mit $0 \neq v$ sodass $Av = \lambda v$. Damit hat die Matrix

$$\lambda \cdot \mathbb{1}_n - A$$

einen nichttrivialen Kern und somit kann der Rang obiger Matrix nicht maximal sein (nach der Dimensionsformel, siehe Satz 3.3.3).

Falls umgekehrt $\lambda \cdot \mathbb{1}_n - A$ nicht-trivialen Kern hat, so gibt es einen Vektor $v \in V$ mit $0 \neq v$ und

$$0 = (\lambda \cdot \mathbb{1}_n - A)v = \lambda v - Av.$$

Also ist $v \neq 0$ mit $Av = \lambda v$. Das beweist das Lemma. □

Korollar 6.2.5. *Falls $A \in K^{n\times n}$ diagonalisierbar ist, so zerfällt $\chi_A(x)$ in Linearfaktoren, d.h.*

$$\chi_A(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$$

für bestimmte Elemente $\lambda_1,\dots,\lambda_n \in K$ (die nicht unbedingt paarweise verschieden sein müssen).

Beweis. Falls A diagonalisierbar ist, so ist A zu einer Diagonalmatrix $D = D(\lambda_1,\dots,\lambda_n)$ ähnlich. Nach obigem Lemma gilt also:

$$\chi_A(x) = \chi_D(x) = \det(x\mathbb{1}_n - D) = \det()D(x-\lambda_1, x-\lambda_2,\dots,x-\lambda_n)) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n).$$

Das beweist das Korollar. □

Bemerkung 6.2.6. *Obiges Korollar ist nützlich um zu zeigen, dass bestimmte Matrizen nicht invertierbar sein können. Dazu betrachten wir die Matrix $A \in \mathbb{R}^{2\times 2}$ die einer Drehung um den Winkel α gegen den Uhrzeigersinn entspricht. Es gilt dann*

$$Ae_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \quad \textit{und} \quad Ae_2 = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}.$$

Damit gilt also:

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

und somit

$$\chi_A(x) = (x-\cos\alpha)^2 + (\sin\alpha)^2 = x^2 - 2\cos\alpha x + 1.$$

Dieses Polynom hat eine reelle Nullstelle genau dann wenn $\cos\alpha = \pm 1$, d.h. genau dann wenn $\alpha = 0$ oder $\alpha = \pi$. Insbesondere sehen wir, dass die Drehung um einen Winkel α mit $0 < \alpha < \pi$ nicht diagonalisierbar ist. Genauer folgt aus dem nachfolgenden Satz, dass $\chi_A(x)$ keinen einzigen Eigenvektor hat, was auch unserer Anschauung entspricht.

Satz 6.2.1. *Für $A \in K^{n\times n}$. Ein Element $\lambda \in K$ ist genau dann Eigenwert von A, falls $\chi_A(\lambda) = 0$ gilt.*

Beweis. Nach Lemma 6.2.4 ist λ genau dann ein Eigenwert, falls $\lambda\cdot\mathbb{1}_n - A$ nicht-trivialen Kern hat. Nach der Dimensionsformel (siehe Satz 3.3.3) ist das dazu äquivalent, dass $\mathrm{Rang}(\lambda\cdot\mathbb{1}_n - A) < n$ gilt. Nach (D10) in Satz 5.2.1 ist das wiederum zu $\det(\lambda\cdot\mathbb{1}_n - A) = 0$ äquivalent und somit zu $\chi_A(\lambda) = 0$. Das schließt den Beweis des Satzes ab. □

6.3 Eigenräume

Lemma 6.2.4 motiviert folgende Definition.

Definition 6.3.1. *Sei $A \in K^{n\times n}$ und sei $\lambda \in K$ ein Eigenwert von A. Dann nennen wir*

$$\mathrm{Eig}(A,\lambda) := \mathrm{Ker}(\lambda\cdot\mathbb{1}_n - A) := \{v \in K^n \mid (\lambda\mathbb{1}_n - A)v = 0\}$$

den Eigenraum von A zu λ.

Da der Kern einer linearen Abbildung immer ein linearer Unterraum ist, ist $\mathrm{Eig}(A,\lambda) \subset K^n$ ein linearer Unterraum.

Lemma 6.3.2. *Sei $A \in K^{n\times n}$ und seien $\lambda_1, \dots, \lambda_r$ verschiedene Eigenwerte von A. Dann sind die Eigenräume $\mathrm{Eig}(A, \lambda_1), \dots, \mathrm{Eig}(A, \lambda_r) \subset K^n$ in direkter Summe, d.h. für jeden Eigenwert λ_{i_0} gilt:*

$$\mathrm{span}_K(\mathrm{Eig}(A, \lambda_i) \mid i \neq i_0) \cap \mathrm{Eig}(A, \lambda_{i_0}) = 0.$$

Beweis. Wir beweisen die Aussage per Induktion nach r. Für $r = 1$ ist die Aussage offensichtlich richtig. Sei also $r \geq 2$ und $v_{i_0} \in \mathrm{Eig}(A, \lambda_{i_0})$ ein nichttrivialer Eigenvektor zum Eigenwert λ_{i_0}. Wir wollen zeigen, dass

$$v_{i_0} \notin \mathrm{span}_K(\mathrm{Eig}(A, \lambda_i) \mid i \neq i_0)$$

gilt. Bis auf Umbenennung der Indizes, können wir $i_0 = r$ annehmen. Wir wollen weiterhin einen Beweis durch Widerspruch führen und nehmen an, dass

$$v_r \in \mathrm{span}_K(\mathrm{Eig}(A, \lambda_i) \mid i \in \{1, 2, \dots, r-1\})$$

gilt. Da $\mathrm{Eig}(A, \lambda_i) \subset K^n$ ein linearer Unterraum ist für alle i, muss dann

$$v_r = \sum_{i=1}^{r-1} v_i$$

für gewisse $v_i \in \mathrm{Eig}(A, \lambda_i)$ gelten. Wenden wir darauf A an, so erhalten wir

$$\lambda_r v_r = A v_r = \sum_{i=1}^{r-1} A v_i = \sum_{i=1}^{r-1} \lambda_i v_i.$$

Andererseits folgt aus $v_r = \sum_{i=1}^{r-1} v_i$ bereits $\lambda_r v_r = \sum_{i=1}^{r-1} \lambda_r v_i$ und damit

$$\sum_{i=1}^{r-1} (\lambda_i - \lambda_r) v_i = 0.$$

Da die Eigenwerte $\lambda_1, \dots, \lambda_r$ alle verschieden waren, gilt $\lambda_i - \lambda_r \neq 0$ für alle $i = 1, \dots, r-1$. Also folgt, dass die Vektoren $v_1, \dots, v_{r-1}$ linear abhängig sind. Nach Induktionsvoraussetzung ist das nur möglich, falls $v_1 = v_2 = \dots = v_{r-1} = 0$ gilt. Dann ist aber auch $v_r = \sum_{i=1}^{r-1} v_i = 0$, was unserer Annahme $v_r \neq 0$ widerspricht. Das beweist das Lemma. □

Satz 6.3.1. *Eine Matrix $A \in K^{n\times n}$ ist genau dann diagonalisierbar, falls*

$$K^n = \bigoplus_{\lambda \in K} \mathrm{Eig}(A, \lambda)$$

gilt.

Beweis. Nach Lemma 6.3.2 stehen die Eigenräume zu den Eigenwerten von A immer in direkter Summe. Es genügt also zu zeigen, dass

$$K^n = \mathrm{span}_K(\mathrm{Eig}(A, \lambda) \mid \lambda \in K)$$

genau dann gilt, wenn A diagonalisierbar ist. Nun gilt aber nach dem Steinitzschen Auswahlsatz (siehe Satz 3.2.3) dass

$$K^n = \mathrm{span}_K(\mathrm{Eig}(A, \lambda) \mid \lambda \in K)$$

genau dann gilt, wenn K^n eine Basis von Eigenvektoren von A besitzt. Letzteres ist nach Satz 6.1.1 äquivalent dazu, dass A diagonalisierbar ist. Das schließt den Beweis des Satzes ab. □

Beispiel. Wir haben gesehen, dass das charakteristische Polynom einer diagonalisierbaren Matrix in Linearfaktoren zerfällt. Das folgende Beispiel zeigt, dass die Umkehrung nicht gilt. Sei dazu
$$A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$
Dann gilt
$$\chi_A(x) = (x-2)^2$$
und damit zerfällt das charakteristische Polynom in Linearfaktoren über $\mathbb{R}$. Weiterhin ist $\lambda = 2$ die einzige Nullstelle von $\chi_A(x)$ und damit ist $\lambda = 2$ der einzige Eigenwert von A. Andererseits gilt
$$\mathrm{Eig}(A,2) = \mathrm{Ker}(2\mathbb{1}_n - A) = \mathrm{Ker}\begin{pmatrix} 0 & -3 \\ 0 & 0 \end{pmatrix} \cong \mathbb{R}.$$
Also gilt $\mathbb{R}^2 \neq \bigoplus_\lambda \mathrm{Eig}(A,\lambda)$ und damit ist A nach obigem Satz nicht diagonalisierbar. Andererseits haben wir folgendes einfaches positives Kriterium.

Korollar 6.3.3. *Sei $A \in K^{n\times n}$. Falls A n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$ hat, so ist A diagonalisierbar.*

Beweis. Da λ_i ein Eigenwert von A ist, gibt es einen Eigenvektor v_i mit $v_i \neq 0$ und $Av_i = \lambda_i v_i$. Da die Eigenwerte $\lambda_1, \dots, \lambda_n$ paarweise verschieden sind, folgt aus Lemma 6.3.2 dass $v_1, \dots, v_n$ linear unabhängig, also eine Basis von K^n sind. Das Korollar folgt damit aus Satz 6.1.1. □

Definition 6.3.4. *Sei $A \in K^{n\times n}$ und sei λ ein Eigenwert von A.*

(1) Die geometrische Vielfachheit von λ ist die Dimension des Eigenraums $\mathrm{Eig}(A,\lambda)$.

(2) Falls λ eine Nullstelle von $\chi_A(x)$ ist, so ist die algebraische Vielfachheit von λ die größte natürliche Zahl m, sodass wir $\chi_A(x) = (x-\lambda)^m \cdot f(x)$ für ein Polynom f schreiben können.

Obiges Korollar hat folgende Verallgemeinerung.

Satz 6.3.2. *Eine Matrix $A \in K^{n\times n}$ ist genau dann diagonalisierbar, wenn das charakteristische Polynom $\chi_A(x)$ in Linearfaktoren zerfällt und für jede Nullstelle λ von $\chi_A(x)$ die algebraische mit der geometrischen Vielfachheit übereinstimmen.*

Beweis. Falls A diagonalisierbar ist, so gilt $D = S^{-1}AS$ für eine invertierbare Matrix $S \in \mathrm{GL}(n,K)$ und eine Diagonalmatrix $D = D(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Dann gilt $\chi_A(x) = \chi_D(x) = \prod_i (x-\lambda_i)$. Also ist die algebraische Vielfachheit einer jeden Nullstelle λ von $\chi_A(x)$ genau die Anzahl der Indizes i sodass $\lambda = \lambda_i$ gilt. Andererseits ist
$$\mathrm{Eig}(A,\lambda) = \mathrm{Ker}(\lambda\mathbb{1}_n - A) \cong S^{-1}\,\mathrm{Ker}((\lambda\mathbb{1}_n - A))S = \mathrm{Ker}(S^{-1}(\lambda\mathbb{1}_n - A)S) = \mathrm{Ker}(\lambda\mathbb{1}_n - S^{-1}AS.)$$
Da $D = S^{-1}AS$ erhalten wir also:
$$\mathrm{Eig}(A,\lambda) \cong \mathrm{Ker}(\lambda\mathbb{1}_n - D).$$
Somit ist klar, dass die geometrische Vielfachheit von λ ebenfalls der Anzahl entspricht wie oft λ in $\lambda_1, \dots, \lambda_n$ vorkommt. Dies beweist eine Richtung im Satz.

Für die Umkehrung nehmen wir an, dass $\chi_A(x) = \prod_i (x-\lambda_i)$ in Linearfaktoren zerfällt. Genauer können wir dann
$$\chi_A(x) = (x-\lambda_{i_1})^{m_1} \cdot (x-\lambda_{i_2})^{m_2} \cdots (x-\lambda_{i_r})^{m_r}$$

für $\lambda_{i_j} \neq \lambda_{i_h}$ für alle $j \neq h$ schreiben. Insbesondere ist dann die algebraische Multiplizität von λ_{i_j} genau m_j. Da $\chi_A(x)$ Grad n hat, gilt $n = \sum_{j=1}^r m_j$. Nach Annahme ist die algebraische und geometrische Vielfachheit einer jeden Nullstelle von $\chi_A(x)$ identisch. Das bedeutet, dass $\mathrm{Eig}(A, \lambda_{i_j})$ Dimension m_j hat. Nach Lemma 6.3.2 sind die Eigenräume zu verschiedenen Eigenwerten in direkter Summe, sodass

$$\bigoplus_{j=1}^n \mathrm{Eig}(A, \lambda_{i_j}) \subset K^n$$

gilt. Andererseits hat aber obiger Unterraum die Dimension $n = \sum_{j=1}^r m_j$ und somit muss

$$\bigoplus_{j=1}^n \mathrm{Eig}(A, \lambda_{i_j}) = K^n$$

gelten. Also hat K^n eine Basis von Eigenvektoren und damit ist A diagonalisierbar nach Satz 6.1.1. □

6.4 Darstellende Matrix bezüglich einer Basis

Sei K ein Körper und sei $f : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorräumen. Seien $\mathcal{B}_V := (v_1, \ldots, v_n)$ und $\mathcal{B}_W := (w_1, \ldots, w_m)$ geordnete Basen von V bzw. W. Dann gilt

$$f(v_i) := \sum_{j=1}^m a_{ji} w_j$$

für eindeutig bestimmte $a_{ji} \in K$.

Im folgenden wird der Begriff "Basisïmmer "geordnete Basis"meinen.

Definition 6.4.1. *Die Matrix $A := (a_{ji})_{1 \leq j \leq m, 1 \leq i \leq n} \in K^{m \times n}$ heißt die darstellende Matrix von f bezüglich der Basen $\mathcal{B}_V$ und $\mathcal{B}_W$. Wir schreiben:*

$$M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) := A.$$

Bemerkung 6.4.2. *Falls $V = K^n$, $W = K^m$ und $\mathcal{B}_V$ und $\mathcal{B}_W$ die jeweiligen Standardbasen sind, so folgt direkt aus der Definition die Beziehung $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) = A_f$, wobei A_f die darstellende Matrix von f bzgl der Standardbasis aus Proposition 4.1.9 ist.*

Satz 6.4.1. *Sei $f : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorräumen. Sei $r := \mathrm{Rang}(f) := \dim_K(\mathrm{Bild}(f))$. Dann gibt es Basen $\mathcal{B}_V$ von V und $\mathcal{B}_W$ von W mit*

$$M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Beweis. Das folgt genauso wie in Satz 4.2.1. Sei dazu $n = \dim V$ und $m = \dim W$. Nach der Dimensionformel hat $\mathrm{Ker}(f)$ die Dimension $n - r$, wir können also eine Basis $v_{r+1}, \ldots, v_n$ von $\mathrm{Ker}(f)$ wählen und diese zu einer Basis $\mathcal{B}_V = (v_1, \ldots, v_n)$ von V ergänzen. Dann ist $w_1 = f(v_1), \ldots, w_r = f(v_r)$ eine Basis von $\mathrm{Bild}(f) \subset W$ und wir ergänzen diese Basis zu einer Basis $\mathcal{B}_W = (w_1, \ldots, w_m)$ von W. Es folgt nun direkt aus der Definition dass

$$M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

gilt, wie behauptet. □

Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W. Obiger Satz zeigt, dass man Basen $\mathcal{B}_V$ und $\mathcal{B}_W$ von V und W finden kann, sodass die zugehörige Matrix möglichst einfach ist. Hat man nun eine lineare Abbildung $f : V \to V$, also einen Endomorphismus von V, so ist es natürlich, dass man nur eine Basis $\mathcal{B}_V$ wählen möchte und dann an der Matrix $M_{\mathcal{B}_V}^{\mathcal{B}_V}(f)$ interessiert ist. Da man also für den Start und Zielbereich dieselbe Basis wählt, ist es natürlich schwieriger wieder eine möglichst einfache Gestalt der Matrix $M_{\mathcal{B}_V}^{\mathcal{B}_V}(f)$ zu finden.

Problem 6.4.3. *Sei V ein endlich-dimensionaler K-Vektorraum. Gibt es eine Basis $\mathcal{B}_V$ von V, sodass die Matrix $M_{\mathcal{B}_V}^{\mathcal{B}_V}(f)$ möglichst einfach aussieht?*

Ziel dieses Kapitels ist es zu sehen, dass obige Frage automatisch auf das Konzept von ähnlichen Matrizen und Diagonalisierbarkeit von Matrizen hinausläuft. Wir beginnen mit folgendem technisch wichtigen Lemma.

Lemma 6.4.4. *Seien $g : U \to V$ und $f : V \to W$ lineare Abbildung zwischen endlich-dimensionalen K-Vektorräumen mit n Basen $\mathcal{B}_U$, $\mathcal{B}_V, \mathcal{B}_W$. Dann gilt*

$$M_{\mathcal{B}_W}^{\mathcal{B}_U}(f \circ g) = M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \cdot M_{\mathcal{B}_V}^{\mathcal{B}_U}(g).$$

Beweis. Sei $\mathcal{B}_U = (u_1, \ldots, u_h)$, $\mathcal{B}_V = (v_1, \ldots, v_n)$ und $\mathcal{B}_W = (w_1, \ldots, w_m)$. Sei weiterhin

$$g(u_l) = \sum_{i=1}^{n} b_{il} v_i \quad \text{und} \quad f(v_i) = \sum_{j=1}^{m} a_{ji} w_j.$$

Dann gilt $M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) = A = (a_{ji})$, $M_{\mathcal{B}_V}^{\mathcal{B}_U}(g) = B = (b_{il})$. Falls weiterhin $M_{\mathcal{B}_W}^{\mathcal{B}_U}(f \circ g) = C = (c_{jl})$ sei, so gilt

$$\sum_{j=1}^{m} c_{jl} w_l = (f \circ g)(u_l).$$

Andererseits gilt

$$(f \circ g)(u_l) = f(g(u_l)) = \sum_{i=1}^{n} b_{il} f(v_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_{il} a_{ji} w_j = \sum_{j=1}^{m} \sum_{i=1}^{n} b_{il} a_{ji} w_j.$$

Insgesamt folgt also $c_{jl} = \sum_{i=1}^{n} a_{ji} b_{il}$. Das heißt, $C = AB$, was zu beweisen war. □

Korollar 6.4.5. *Sie V ein endlich-dimensionaler K-Vektorraum und seien $\mathcal{B}_V$ und $\mathcal{B}'_V$ zwei Basen von V. Dann gilt*

$$M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id}) \in \mathrm{GL}(n, K) \quad \textit{mit} \quad M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id})^{-1} = M_{\mathcal{B}_V}^{\mathcal{B}'_V}(\mathrm{id})$$

Wir nennen $M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id})$ eine Basiswechsel-Matrix.

Beweis. Nach obigem Lemma gilt

$$M_{\mathcal{B}_V}^{\mathcal{B}'_V}(\mathrm{id}) \cdot M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id}) = M_{\mathcal{B}_V}^{\mathcal{B}_V}(\mathrm{id}) \;\; \text{und} \;\; M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id}) \cdot M_{\mathcal{B}_V}^{\mathcal{B}'_V}(\mathrm{id}) = M_{\mathcal{B}'_V}^{\mathcal{B}'_V}(\mathrm{id}).$$

Direkt aus der Definition folgt aber $M_{\mathcal{B}_V}^{\mathcal{B}_V}(\mathrm{id}) = \mathbb{1}_n$ und $M_{\mathcal{B}'_V}^{\mathcal{B}'_V}(\mathrm{id}) = \mathbb{1}_n$. Also ist $M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id})$ invertierbar mit $M_{\mathcal{B}'_V}^{\mathcal{B}_V}(\mathrm{id})^{-1} = M_{\mathcal{B}_V}^{\mathcal{B}'_V}(\mathrm{id})$, wie behauptet. □

Korollar 6.4.6. *Sie V ein endlich-dimensionaler K-Vektorraum. Seien $\mathcal{B}_V$ und $\mathcal{B}'_V$ zwei Basen von V. Dann gilt für jede lineare Abbildung $f : V \to V$ folgendes:*

$$M^{\mathcal{B}'_V}_{\mathcal{B}'_V}(f) = S^{-1} \cdot M^{\mathcal{B}_V}_{\mathcal{B}_V}(f) \cdot S,$$

wobei $S := M^{\mathcal{B}'_V}_{\mathcal{B}_V}(\mathrm{id})$. Das heißt, $M^{\mathcal{B}'_V}_{\mathcal{B}'_V}(f)$ und $M^{\mathcal{B}_V}_{\mathcal{B}_V}(f)$ sind zu einander ähnliche Matrizen.

Beweis. Nach obigem Korollar gilt $S^{-1} = M^{\mathcal{B}_V}_{\mathcal{B}'_V}(\mathrm{id})$. Also,

$$S^{-1} \cdot M^{\mathcal{B}_V}_{\mathcal{B}_V}(f) \cdot S = M^{\mathcal{B}_V}_{\mathcal{B}'_V}(\mathrm{id}) \circ M^{\mathcal{B}_V}_{\mathcal{B}_V}(f) \cdot M^{\mathcal{B}'_V}_{\mathcal{B}_V}(\mathrm{id}).$$

Nach obigem Lemma gilt dann:

$$S \cdot M^{\mathcal{B}_V}_{\mathcal{B}_V}(f) \cdot S^{-1} = M^{\mathcal{B}'_V}_{\mathcal{B}'_V}(\mathrm{id} \circ f \circ \mathrm{id})$$

und somit folgt die Behauptung aus $f = \mathrm{id} \circ f \circ \mathrm{id}$. □

Bemerkung 6.4.7. *Sei $V = K^n$ und sei $\mathcal{B}$ die Standardbasis von V. Falls $\mathcal{B}' = (v_1, \dots, v_n)$ eine beliebige Basis von $V = K^n$ ist, so ist die Basiswechselmatrix*

$$M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}) = (v_1|v_2|\cdots|v_n)$$

gegeben durch die Matrix mit deren i-te Spalte v_i ist.

Sei umgekehrt $S \in \mathrm{GL}(n, K)$ beliebig. Dann hat S vollen Rang und damit sind die Spalten $v_1, \dots, v_n$ von S eine Basis $\mathcal{B}'$ von K^n und es gilt, wie wir oben gesehen haben, die Beziehung

$$S = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}).$$

Satz 6.4.2. *Sei V ein n-dimensionaler K-Vektorraum (z.B. $V = K^n$). Zwei Matrizen $A, B \in K^{n\times n}$ sind genau dann ähnlich falls sie dieselbe lineare Abbildung $f : V \to V$ bezüglich verschiedener Basen von V darstellen, d.h. falls es Basen $\mathcal{B}$ und $\mathcal{B}'$ von V gibt mit $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ und $B = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$.*

Beweis. Falls $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ und $B = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$ gilt, so sind A und B ähnlich nach Korollar 6.4.6.

Falls umgekehrt $A = SBS^{-1}$ gilt, so wähle eine beliebige Basis von V und identifiziere damit V mit K^n. Wir betrachten dann $f : V \to V$ $v \mapsto Bv$. Falls $\mathcal{B}$ die Standardbasis von $V = K^n$ ist, so gilt also $B = M^{\mathcal{B}}_{\mathcal{B}}(f)$. Da S invertierbar ist, hat S^{-1} vollen Rang und somit sind die Spalten von S^{-1} eine Basis $\mathcal{B}'$ von V. Es gilt dann nach obiger Bemerkung:

$$S^{-1} = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}).$$

Nach Korollar 6.4.6 gilt demnach:

$$A = SBS^{-1} = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}) = M^{\mathcal{B}'}_{\mathcal{B}'}(f).$$

Also stellen A und B den Endomorphismus f bezüglich verschiedener Basen da. Das schließt den Beweis ab. □

Satz 6.4.3. *Zwei Matrizen $A, B \in K^{n\times n}$ sind genau dann ähnlich falls es eine Basis $\mathcal{B}$ von K^n gibt, sodass die lineare Abbildung*

$$f := h_A : K^n \longrightarrow K^n, \quad v \mapsto Av$$

bezüglich der Basis $\mathcal{B}'$ die Gestalt $B = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$ hat.

Beweis. Falls A und B ähnlich sind, so gibt es eine invertierbare Matrix S mit $B = S^{-1}AS$. Sei dann $\mathcal{B} = (e_1, \ldots, e_n)$ die Einheitsbasis von K^n und sie $\mathcal{B}' = (v_1, \ldots, v_n)$ die Basis von K^n die durch die Spalten von S gegeben ist, d.h. $S = (v_1|v_2|\ldots|v_n)$. Dann gilt nach Bemerkung 6.4.7:

$$S = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}).$$

Sei weiterhin $f := h_A : K^n \to K^n$ die lineare Abbildung $v \mapsto Av$. Dann gilt nach Korollar 6.4.6:

$$M_{\mathcal{B}'}^{\mathcal{B}'}(f) = S^{-1}M_{\mathcal{B}}^{\mathcal{B}}(f)S = S^{-1}AS = B.$$

Dies zeigt eine Richtung.

Für die Umkehrung, sei $f = h_A : K^n \to K^n$ die lineare Abbildung $v \mapsto Av$. Sei weiterhin $\mathcal{B}$ die Standardbasis von K^n. Angenommen es gibt eine weitere Basis $\mathcal{B}'$ mit $B = M_{\mathcal{B}'}^{\mathcal{B}'}(f)$. Dann gilt nach Korollar 6.4.6:

$$B = S^{-1}M_{\mathcal{B}}^{\mathcal{B}}(f)S = S^{-1}AS$$

und damit sind A und B ähnlich. Das schließt den Beweis ab. □

Korollar 6.4.8. *Eine Matrix $A \in K^{n\times n}$ ist genau dann diagonalisierbar, wenn es eine Basis $\mathcal{B}'$ von K^n gibt, sodass die lineare Abbildung $h_A : K^n \to K^n$ bezüglich der Basis $\mathcal{B}'$ eine Diagonalmatrix D ist, also $D = M_{\mathcal{B}'}^{\mathcal{B}'}(h_A)$.*

Beweis. Das folgt direkt aus obigem Satz und der Definition von Diagonalisierbarkeit. □

Literaturverzeichnis

[Fi14] G. Fischer, *Lineare Algebra*, 18. Auflage, Springer, 2014.