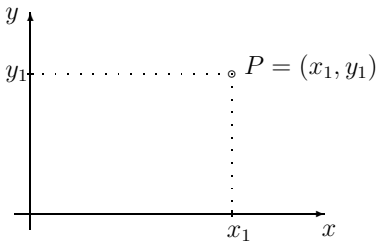
1. Vektorräume

Vorbemerkungen

Konkrete geometrische Fragestellungen in der Ebene oder im drei-dimensionalen Raum waren vielfach Ausgangspunkt bedeutender mathematischer Entwicklungen. Als Hilfsmittel zur Behandlung solcher Fragen wurden beispielsweise geometrische Konstruktionsverfahren mittels Zirkel und Lineal entwickelt. Eine andere Strategie besteht darin, geometrische Fragen in rechnerische Probleme umzusetzen, um durch "Ausrechnen" zu Lösungen zu gelangen. Dies ist das Vorgehen der *analytischen Geometrie*, die 1637 von René Descartes in seinem berühmten Werk "La Géométrie" begründet wurde. Ein Großteil der rechnerischen Methoden der analytischen Geometrie wiederum wird heute in erweiterter Form unter dem Begriff der *Linearen Algebra* zusammengefasst.

Wir wollen im Folgenden etwas näher auf die grundlegenden Ideen des Descartes'schen Ansatzes eingehen. Hierzu betrachten wir eine Ebene E (etwa in dem uns umgebenden drei-dimensionalen Raum), zeichnen einen Punkt von E als so genannten Nullpunkt 0 aus und wählen dann ein Koordinatensystem mit Koordinatenachsen x und y, die sich im Nullpunkt 0 schneiden. Identifizieren wir die Achsen x und y jeweils noch mit der Menge $\mathbb{R}$ der reellen Zahlen, so lassen sich die Punkte P von E als Paare reeller Zahlen interpretieren:



In der Tat, ist P ein Punkt in E, so konstruiere man die Parallele zu y durch P. Diese schneidet die Achse x in einem Punkt x_1. Entsprechend schneidet die Parallele zu x durch P die Achse y in einem Punkt y_1, so dass man aus P das Koordinatenpaar (x_1, y_1) erhält. Umgekehrt lässt sich P aus dem Paar (x_1, y_1) in einfacher Weise zurückgewinnen, und zwar als Schnittpunkt der Parallelen zu y durch x_1 und der Parallelen zu x durch y_1. Genauer stellt man fest, dass die Zuordnung $P \longmapsto (x_1, y_1)$ eine umkehrbar eindeutige Beziehung zwischen

den Punkten von E und den Paaren reeller Zahlen darstellt und man deshalb wie behauptet eine Identifizierung

$$E = \mathbb{R}^2 = \text{ Menge aller Paare reeller Zahlen}$$

vornehmen kann. Natürlich hängt diese Identifizierung von der Wahl des Nullpunktes 0 sowie der Koordinatenachsen x und y ab. Wir haben in obiger Abbildung ein rechtwinkliges Koordinatensystem angedeutet. Im Prinzip brauchen wir jedoch an dieser Stelle noch nichts über Winkel zu wissen. Es genügt, wenn wir als Koordinatenachsen zwei verschiedene Geraden x und y durch den Nullpunkt 0 verwenden. Genaueres hierzu werden wir noch in den Vorbemerkungen zu Kapitel 2 besprechen.

Es soll nun auch die Identifizierung der beiden Koordinatenachsen x und y mit der Menge $\mathbb{R}$ der reellen Zahlen noch etwas genauer beleuchtet werden. Durch Festlegen des Nullpunktes ist auf x und y jeweils die Streckungsabbildung mit Zentrum 0 und einer reellen Zahl als Streckungsfaktor definiert. Wählen wir etwa einen von 0 verschiedenen Punkt $1_x \in x$ aus und bezeichnen mit $\alpha \cdot 1_x$ das Bild von 1_x unter der Streckung mit Faktor α, so besteht x gerade aus allen Punkten $\alpha \cdot 1_x$, wobei α die reellen Zahlen durchläuft. Genauer können wir sagen, dass die Zuordnung $\alpha \longmapsto \alpha \cdot 1_x$ eine umkehrbar eindeutige Beziehung zwischen den reellen Zahlen und den Punkten von x erklärt. Nach Auswahl je eines von 0 verschiedenen Punktes $1_x \in x$ und entsprechend $1_y \in y$ sind daher x und y auf natürliche Weise mit der Menge $\mathbb{R}$ der reellen Zahlen zu identifizieren, wobei die Punkte $0, 1_x \in x$ bzw. $0, 1_y \in y$ den reellen Zahlen 0 und 1 entsprechen. Die Möglichkeit der freien Auswahl der Punkte $1_x \in x$ und $1_y \in y$ wie auch die Verwendung nicht notwendig rechtwinkliger Koordinatensysteme machen allerdings auf ein Problem aufmerksam: Der Abstand von Punkten in E wird unter der Identifizierung $E = \mathbb{R}^2$ nicht notwendig dem auf $\mathbb{R}^2$ üblichen euklidischen Abstand entsprechen, der für Punkte $P_1 = (x_1, y_1)$ und $P_2 = (x_2, y_2)$ durch

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

gegeben ist. Eine korrekte Reflektierung von Abständen auf E ist jedoch mit Hilfe der später noch zu diskutierenden *Skalarprodukte* möglich.

In der Mathematik ist man stets darum bestrebt, bei der Analyse von Phänomenen und Problemen, für die man sich interessiert, zu gewissen "einfachen Grundstrukturen" zu gelangen, die für das Bild, das sich dem Betrachter bietet, verantwortlich sind. Solchermaßen als wichtig erkannte Grundstrukturen untersucht man dann oftmals losgelöst von der eigentlichen Problematik, um herauszufinden, welche Auswirkungen diese haben; man spricht von einem *Modell*, das man untersucht. Modelle haben den Vorteil, dass sie in der Regel leichter zu überschauen sind, aber manchmal auch den Nachteil, dass sie den eigentlich zu untersuchenden Sachverhalt möglicherweise nur in Teilaspekten beschreiben können.

In unserem Falle liefert der Descartes'sche Ansatz die Erkenntnis, dass Punkte von Geraden, Ebenen oder des drei-dimensionalen Raums mittels Koordinaten zu beschreiben sind. Hierauf gestützt können wir, wie wir gesehen

haben, die Menge $\mathbb{R}^2$ aller Paare reeller Zahlen als Modell einer Ebene ansehen. Entsprechend bildet die Menge $\mathbb{R}^3$ aller Tripel reeller Zahlen ein Modell des drei-dimensionalen Raums, sowie natürlich $\mathbb{R} = \mathbb{R}^1$ ein Modell einer Geraden. Die Untersuchung solcher Modelle führt uns zum zentralen Thema dieses Kapitels, nämlich zu den *Vektorräumen*. Vektorräume beinhalten als fundamentale Struktur zwei Rechenoperationen, zum einen die Multiplikation von Skalaren (in unserem Falle reellen Zahlen) mit Vektoren, was man sich als einen Streckungsprozess vorstellen kann, und zum anderen die Addition von Vektoren. Wir wollen dies mit den zugehörigen geometrischen Konsequenzen einmal am Beispiel einer Ebene E und ihrem Modell $\mathbb{R}^2$ erläutern.

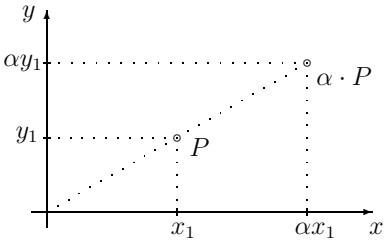
Wir beginnen mit der skalaren Multiplikation. Für

$$\alpha \in \mathbb{R}, \qquad P = (x_1, y_1) \in \mathbb{R}^2$$

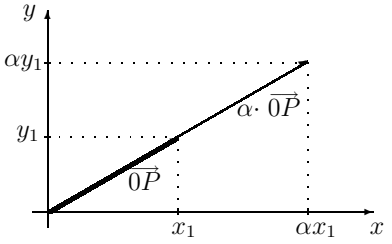
bezeichnet man mit

$$\alpha \cdot P = \alpha \cdot (x_1, y_1) := (\alpha x_1, \alpha y_1)$$

das Produkt von α und P, wobei sich in E folgendes Bild ergibt:



Die Multiplikation von Punkten $P \in E$ mit einem Skalar $\alpha \in \mathbb{R}$ ist folglich zu interpretieren als Streckungsabbildung mit Streckungszentrum 0 und Streckungsfaktor α. Besonders instruktiv lässt sich dies beschreiben, wenn man die Punkte $P \in E$ als "Vektoren" im Sinne gerichteter Strecken $\overrightarrow{0P}$ auffasst. Vektoren sind somit charakterisiert durch ihre Länge und ihre Richtung (außer für den Nullvektor $\overrightarrow{00}$, der keine bestimmte Richtung besitzt). Der Vektor $\alpha \cdot \overrightarrow{0P}$ geht dann aus $\overrightarrow{0P}$ hervor, indem man ihn mit α streckt, d. h. seine Länge mit α (oder, besser, mit dem Betrag $|\alpha|$) multipliziert und ansonsten die Richtung des Vektors beibehält bzw. invertiert, je nachdem ob $\alpha \geq 0$ oder $\alpha < 0$ gilt:



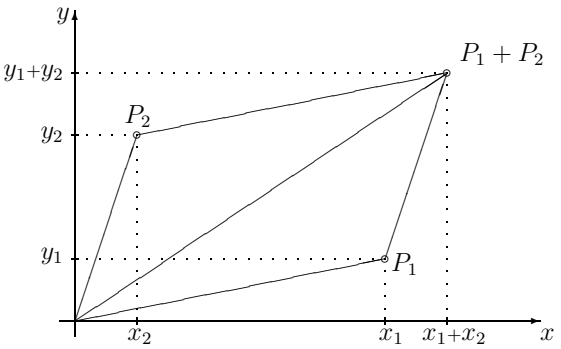
Als weitere Rechenoperation betrachten wir die Addition von Punkten in $\mathbb{R}^2$. Für

$$P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in \mathbb{R}^2$$

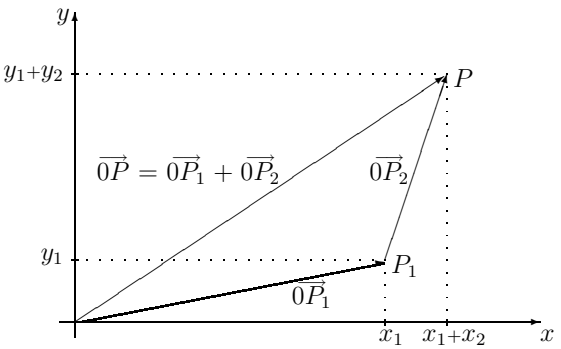
setzt man

$$P_1 + P_2 := (x_1 + x_2, y_1 + y_2),$$

was in E mittels folgender Skizze verdeutlicht werden möge:



Auch die Beschreibung der Addition in E gestaltet sich instruktiver, wenn man den Vektorstandpunkt im Sinne gerichteter Strecken zugrunde legt. Allerdings sollte man dabei zulassen, dass Vektoren als gerichtete Strecken parallel zu sich selbst verschoben und somit vom Koordinatenursprung als ihrem natürlichen Fußpunkt gelöst werden können. Die Summe der Vektoren $\overrightarrow{0P_1}$ und $\overrightarrow{0P_2}$ ergibt sich dann als Vektor $\overrightarrow{0P}$, wobei P derjenige Endpunkt ist, den man erhält, indem man beide Vektoren miteinander kombiniert, also den Vektor $\overrightarrow{0P_1}$ in 0 anlegt und den Vektor $\overrightarrow{0P_2}$ im Endpunkt P_1 von $\overrightarrow{0P_1}$, etwa wie folgt:



Dabei zeigt die obige Parallelogrammkonstruktion, dass sich das Ergebnis der Addition nicht ändert, wenn man alternativ den Vektor $\overrightarrow{0P_2}$ in 0 anlegt und anschließend den Vektor $\overrightarrow{0P_1}$ im Endpunkt von $\overrightarrow{0P_2}$. Die Addition von Vektoren hängt daher nicht von der Reihenfolge der Summanden ab, sie ist *kommutativ*.

Es mag etwas verwirrend wirken, wenn wir die Elemente des $\mathbb{R}^2$ einerseits als Punkte, sowie andererseits auch als (verschiebbare) Vektoren im Sinne gerichteter Strecken interpretieren. Im Prinzip könnte man eine begriffliche Trennung zwischen Punkten und Vektoren vornehmen, indem man den einem Punkt $P \in \mathbb{R}^2$ zugeordneten Vektor $\overrightarrow{0P}$ als *Translation* $Q \longmapsto P+Q$ interpretiert, d. h. als Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$, die einem Element $Q \in \mathbb{R}^2$ das Element $P + Q$ als Bild zuordnet. Wir wollen von dieser Möglichkeit allerdings keinen Gebrauch machen, da eine Trennung der Begriffe für unsere Zwecke keine Vorteile bringt und die Dinge lediglich komplizieren würde.

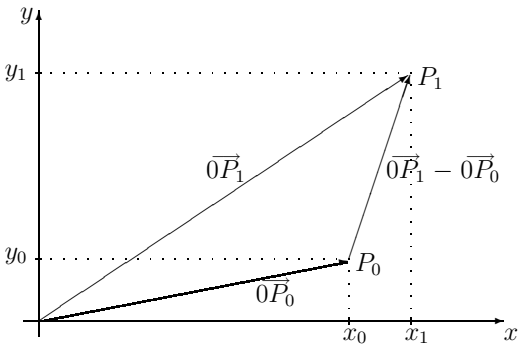
Als Nächstes wollen wir besprechen, dass die Addition von Punkten und Vektoren in $\mathbb{R}^2$ bzw. E auf natürliche Weise auch eine Subtraktion nach sich zieht. Für $P_0 = (x_0, y_0) \in \mathbb{R}^2$ setzt man

$$-P_0 = -(x_0, y_0) := (-1) \cdot (x_0, y_0) = (-x_0, -y_0)$$

und nennt dies das negative oder inverse Element zu P_0. Dieses ist in eindeutiger Weise charakterisiert als Element $Q \in \mathbb{R}^2$, welches der Gleichung $P_0 + Q = 0$ genügt. Die Subtraktion zweier Elemente $P_1 = (x_1, y_1)$ und $P_0 = (x_0, y_0)$ in $\mathbb{R}^2$ wird dann in nahe liegender Weise auf die Addition zurückgeführt, und zwar durch

$$P_1 - P_0 := P_1 + (-P_0) = (x_1 - x_0, y_1 - y_0).$$

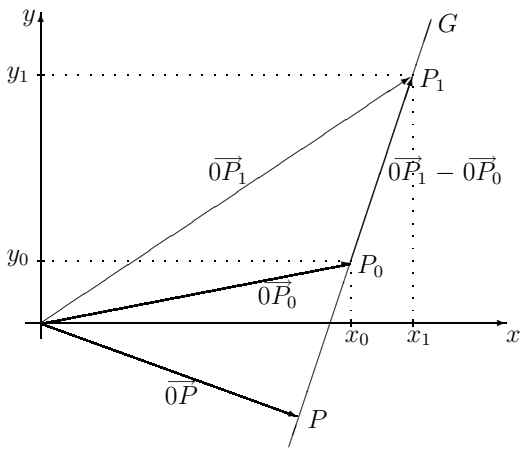
Legen wir wieder den Vektorstandpunkt in E zugrunde, so entsteht also $-\overrightarrow{0P_0}$ aus dem Vektor $\overrightarrow{0P_0}$ durch Invertieren seiner Richtung, wobei die Länge erhalten bleibt. Damit lässt sich die Differenz zweier Vektoren $\overrightarrow{0P_1}$ und $\overrightarrow{0P_0}$ wie folgt illustrieren:



Insbesondere erkennt man, dass die Summe der Vektoren $\overrightarrow{0P_0}$ und $\overrightarrow{0P_1} - \overrightarrow{0P_0}$ gerade den Vektor $\overrightarrow{0P_1}$ ergibt, was eine sinnvoll definierte Addition bzw. Subtraktion natürlich ohnehin leisten sollte. Allgemeiner kann man Summen des Typs

$$\overrightarrow{0P} = \overrightarrow{0P_0} + \alpha \cdot (\overrightarrow{0P_1} - \overrightarrow{0P_0})$$

mit unterschiedlichen Skalaren $\alpha \in \mathbb{R}$ bilden. Der Punkt P liegt dann für $P_0 \neq P_1$ stets auf der Geraden G, die durch P_0 und P_1 festgelegt ist, und zwar durchläuft P ganz G, wenn α ganz $\mathbb{R}$ durchläuft:



Die Gerade in E bzw. $\mathbb{R}^2$, welche die gegebenen Punkte P_0 und P_1 enthält, wird daher durch die Gleichung

$$G = \{P_0 + t \cdot (P_1 - P_0)\,;\, t \in \mathbb{R}\}$$

beschrieben. Sind zwei solche Geraden

$$G = \{P_0 + t \cdot (P_1 - P_0)\,;\, t \in \mathbb{R}\}, \qquad G' = \{P_0' + t \cdot (P_1' - P_0')\,;\, t \in \mathbb{R}\}$$

mit $P_0 \neq P_1$ und $P_0' \neq P_1'$ gegeben, so sind diese genau dann parallel, wenn $P_1 - P_0$ ein skalares Vielfaches von $P_1' - P_0'$ ist, bzw. umgekehrt, wenn $P_1' - P_0'$ ein skalares Vielfaches von $P_1 - P_0$ ist. Ist Letzteres nicht der Fall, so besitzen G und G' genau einen Schnittpunkt, wobei eine Berechnung dieses Schnittpunktes auf die Lösung eines so genannten linearen Gleichungssystems führt, welches aus 2 Gleichungen mit 2 Unbekannten, nämlich den Koordinaten des Schnittpunktes von G und G' besteht. Die Lösung von Gleichungssystemen dieses Typs wird uns noch ausführlich in Kapitel 3 beschäftigen.

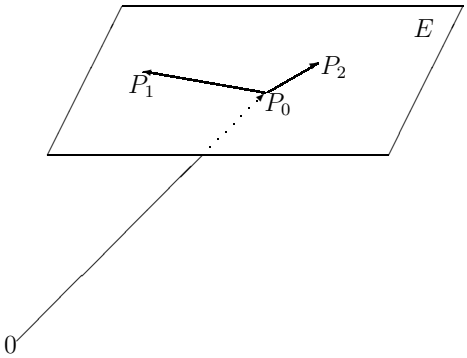
Die vorstehenden Überlegungen lassen sich ohne Probleme auf den dreidimensionalen Raum und sein Modell $\mathbb{R}^3$ verallgemeinern. Beispielsweise ist für zwei Punkte $P_0, P_1 \in \mathbb{R}^3$ wiederum

$$G = \{P_0 + t \cdot (P_1 - P_0)\,;\, t \in \mathbb{R}\}$$

die durch P_0 und P_1 bestimmte Gerade im $\mathbb{R}^3$. Für Punkte $P_0, P_1, P_2 \in \mathbb{R}^3$ kann man mit $P_1' := P_1 - P_0$ und $P_2' := P_2 - P_0$ entsprechend das Gebilde

$$E = \{P_0 + s \cdot P_1' + t \cdot P_2'\,;\, s, t \in \mathbb{R}\}$$

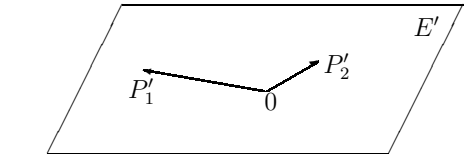
betrachten:



Wenn P_1' kein Vielfaches von P_2' und P_2' kein Vielfaches von P_1' ist, die Vektoren in 0 angetragen also nicht auf einer Geraden durch 0 liegen, so bezeichnet man P_1' und P_2' als *linear unabhängig*. In diesem Falle erkennt man E als Ebene, ansonsten als Gerade oder auch nur als Punkt. Da die Vektoren P_1' und P_2' hier eine entscheidende Rolle spielen, sollten wir auch das Gebilde

$$E' = \{s \cdot P_1' + t \cdot P_2' \,;\, s, t \in \mathbb{R}\}$$

betrachten, welches durch Verschieben von E um den Vektor $-\overrightarrow{0P}$ entsteht:



Im Rahmen der Vektorräume nennt man E' den von P_1' und P_2' *aufgespannten* oder *erzeugten linearen Unterraum* von $\mathbb{R}^3$. Allgemeiner kann man im $\mathbb{R}^3$ den von beliebig vielen Vektoren $Q_1, \ldots, Q_r$ erzeugten linearen Unterraum

$$U = \{t_1 Q_1 + \ldots + t_r Q_r \,;\, t_1, \ldots, t_r \in \mathbb{R}\}$$

betrachten. Für einen Vektor $Q \in \mathbb{R}^3$ sagt man, dass Q *linear von* $Q_1, \ldots, Q_r$ *abhängt*, falls $Q \in U$ gilt. Folgende Fälle sind möglich: Für $Q_1 = \ldots = Q_r = 0$ besteht U nur aus dem Nullpunkt 0. Ist aber einer der Vektoren $Q_1, \ldots, Q_r$ von 0 verschieden, etwa $Q_1 \neq 0$, so enthält U zumindest die durch Q_1 gegebene Gerade $G = \{tQ_1 \,;\, t \in \mathbb{R}\}$. Gehören auch $Q_2, \ldots, Q_r$ zu G, d. h. sind $Q_2, \ldots, Q_r$ linear abhängig von Q_1, so stimmt U mit G überein. Ist Letzteres nicht der Fall und gilt etwa $Q_2 \notin G$, so spannen Q_1 und Q_2 die Ebene $E = \{t_1 Q_1 + t_2 Q_2 \,;\, t_1, t_2 \in \mathbb{R}\}$ auf, so dass U zumindest diese Ebene enthält. Im Falle $Q_3, \ldots, Q_r \in E$, also wenn $Q_3, \ldots, Q_r$ linear von Q_1, Q_2 abhängen, stimmt U mit E überein. Ansonsten gibt es einen dieser Vektoren, etwa Q_3, der nicht zu E gehört. Die Vektoren Q_1, Q_2, Q_3 bilden dann sozusagen ein Koordinatensystem im $\mathbb{R}^3$, und man sieht dass U mit ganz $\mathbb{R}^3$ übereinstimmt, dass

also alle Vektoren im $\mathbb{R}^3$ linear von Q_1, Q_2, Q_3 abhängen. Insbesondere ergibt sich, dass ein linearer Unterraum im $\mathbb{R}^3$ entweder aus dem Nullpunkt, aus einer Geraden durch 0, aus einer Ebene durch 0 oder aus ganz $\mathbb{R}^3$ besteht.

Das soeben eingeführte Konzept der *linearen Abhängigkeit* von Vektoren ist ein ganz zentraler Punkt, der in diesem Kapitel ausführlich im Rahmen der Vektorräume behandelt werden wird. Dabei nennt man ein System von Vektoren $Q_1, \ldots, Q_r$ *linear unabhängig*, wenn keiner dieser Vektoren von den restlichen linear abhängt. Die oben durchgeführte Überlegung zeigt beispielsweise, dass linear unabhängige Systeme im $\mathbb{R}^3$ aus höchstens 3 Elementen bestehen. Insbesondere werden uns linear unabhängige Systeme, so wie wir sie im obigen Beispiel für lineare Unterräume des $\mathbb{R}^3$ konstruiert haben, gestatten, den Begriff des Koordinatensystems oder der Dimension im Kontext der Vektorräume zu präzisieren. Als Verallgemeinerung linear unabhängiger Systeme von Vektoren werden wir schließlich noch so genannte *direkte Summen* von linearen Unterräumen eines Vektorraums studieren.

Wir haben bisher im Hinblick auf Vektorräume lediglich die Modelle $\mathbb{R}^n$ mit $n = 1, 2, 3$ betrachtet, wobei unser geometrisches Vorstellungsvermögen in erheblichem Maße bei unseren Argumentationen mit eingeflossen ist. Bei der Behandlung der Vektorräume in den nachfolgenden Abschnitten werden wir jedoch grundsätzlicher vorgehen, indem wir eine Reihe von Verallgemeinerungen zulassen und uns bei der Entwicklung der Theorie lediglich auf gewisse axiomatische Grundlagen stützen. Zunächst beschränken wir uns bei dem zugrunde liegenden Skalarenbereich nicht auf die reellen Zahlen $\mathbb{R}$, sondern lassen beliebige *Körper* zu. Körper sind zu sehen als Zahlsysteme mit gewissen Axiomen für die Addition und Multiplikation, die im Wesentlichen den Regeln für das Rechnen mit den reellen Zahlen entsprechen. So kennt man neben dem Körper $\mathbb{R}$ der reellen Zahlen beispielsweise den Körper $\mathbb{Q}$ der rationalen Zahlen wie auch den Körper $\mathbb{C}$ der komplexen Zahlen. Es gibt aber auch Körper, die nur aus endlich vielen Elementen bestehen.

Die Axiome eines Körpers bauen auf denen einer *Gruppe* auf, denn ein Körper bildet mit seiner Addition insbesondere auch eine Gruppe. So werden wir in diesem Kapitel nach gewissen Vorbereitungen über Mengen zunächst Gruppen studieren, ausgehend von den zugehörigen Gruppenaxiomen. Wir beschäftigen uns dann weiter mit Körpern und deren Rechenregeln und gelangen anschließend zu den Vektorräumen. Vektorräume sind immer in Verbindung mit einem entsprechenden Skalarenbereich zu sehen, dem zugehörigen Körper; man spricht von einem Vektorraum über einem Körper K oder von einem K-Vektorraum. Ein K-Vektorraum V ist ausgerüstet mit einer Addition und einer skalaren Multiplikation, d. h. für $a, b \in V$ und $\alpha \in K$ sind die Summe $a + b$ sowie das skalare Produkt $\alpha \cdot a$ als Elemente von V erklärt. Addition und skalare Multiplikation genügen dabei den so genannten Vektorraumaxiomen, welche bezüglich der Addition insbesondere die Gruppenaxiome enthalten. Prototyp eines K-Vektorraums ist für eine gegebene natürliche Zahl n die Menge

$$K^n = \{(a_1, \ldots, a_n)\,;\, a_1, \ldots, a_n \in K\}$$

aller n-Tupel mit Komponenten aus K, wobei Addition und skalare Multiplikation durch

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) := (a_1 + b_1, \ldots, a_n + b_n),$$
$$\alpha \cdot (a_1, \ldots, a_n) := (\alpha a_1, \ldots, \alpha a_n)$$

gegeben sind.

Insbesondere wird mit dieser Definition die oben angesprochene Reihe von Modellen $\mathbb{R}^n$ für $n = 1, 2, 3$ auf beliebige *Dimensionen* n verallgemeinert. Dies hat durchaus einen realen Hintergrund, denn um beispielsweise ein Teilchen im drei-dimensionalen Raum in zeitlicher Abhängigkeit zu beschreiben, benötigt man neben den 3 räumlichen Koordinaten noch eine zusätzliche zeitliche Koordinate, so dass man sich im Grunde genommen im Vektorraum $\mathbb{R}^4$ bewegt. In analoger Weise lassen sich Paare von Punkten im drei-dimensionalen Raum als Punkte des $\mathbb{R}^6$ charakterisieren.

1.1 Mengen und Abbildungen

Normalerweise müsste man hier mit einer streng axiomatischen Begründung der Mengenlehre beginnen. Da dies jedoch einen unverhältnismäßig großen Aufwand erfordern würde, wollen wir uns an dieser Stelle mit einem naiven Standpunkt begnügen und unter einer *Menge* lediglich eine Zusammenfassung gewisser Objekte verstehen, der so genannten *Elemente* dieser Menge. Eine Menge X ist somit in eindeutiger Weise durch ihre Elemente festgelegt, wobei wir $x \in X$ schreiben, wenn x ein Element von X ist, bzw. $x \notin X$, wenn dies nicht der Fall ist. Insbesondere werden wir folgende Mengen in natürlicher Weise als gegeben annehmen:

$$\begin{aligned}
\emptyset &= \text{ leere Menge,}\\
\mathbb{N} &= \{0, 1, 2, \ldots\} \text{ natürliche Zahlen,}\\
\mathbb{Z} &= \{0, \pm 1, \pm 2, \ldots\} \text{ ganze Zahlen,}\\
\mathbb{Q} &= \{p/q\,;\, p, q \in \mathbb{Z}, q \neq 0\} \text{ rationale Zahlen,}\\
\mathbb{R} &= \text{ reelle Zahlen.}
\end{aligned}$$

Es sei angemerkt, dass bei einer Menge, sofern wir sie in aufzählender Weise angeben, etwa $X = \{x_1, \ldots, x_n\}$, die Elemente $x_1, \ldots, x_n$ nicht notwendig paarweise verschieden sein müssen. Diese Konvention gilt auch für unendliche Mengen; man vergleiche hierzu etwa die obige Beschreibung von $\mathbb{Q}$.

Um Widersprüche zu vermeiden, sind die Mengenaxiome so ausgelegt, dass die Bildung von Mengen gewissen Restriktionen unterworfen ist. Beispielsweise darf eine Menge niemals sich selbst als Element enthalten, so dass insbesondere die Gesamtheit aller Mengen nicht als Menge angesehen werden kann, da sie sich selbst als Element enthalten würde. Einen Hinweis auf die hiermit verbundene Problematik liefert das folgende Paradoxon von Russel: Wir nehmen einmal in

naiver Weise an, dass man die Gesamtheit aller Mengen, die sich nicht selbst enthalten, also

$$X = \{\text{Mengen } A \text{ mit } A \notin A\},$$

als Menge betrachten kann. Fragt man sich dann, ob $X \in X$ oder $X \notin X$ gilt, so erhält man im Falle $X \in X$ nach Definition von X sofort $X \notin X$ und im Falle $X \notin X$ entsprechend $X \in X$. Es ergibt sich also $X \in X$ und $X \notin X$ zugleich, was keinen Sinn macht.

Wichtig für die Handhabung von Mengen sind die erlaubten Prozesse der Mengenbildung, auf die wir nachfolgend eingehen.

(1) Teilmengen. – Es sei X eine Menge und $P(x)$ eine Aussage, deren Gültigkeit (wahr oder falsch) man für Elemente $x \in X$ testen kann. Dann nennt man $Y = \{x \in X\,;\, P(x) \text{ ist wahr}\}$ eine *Teilmenge* von X und schreibt $Y \subset X$. Dabei ist auch $Y = X$ zugelassen. Gilt allerdings $Y \neq X$, so nennt man Y eine *echte* Teilmenge von X. Beispielsweise ist $\mathbb{R}_{>0} := \{x \in \mathbb{R}\,;\, x > 0\}$ eine (echte) Teilmenge von $\mathbb{R}$. Für eine gegebene Menge X bilden die Teilmengen von X wiederum eine Menge, die so genannte *Potenzmenge* $\mathfrak{P}(X)$.

(2) Vereinigung und Durchschnitt. – Es sei X eine Menge und I eine Indexmenge, d. h. eine Menge, deren Elemente wir als Indizes verwenden wollen. Ist dann für jedes $i \in I$ eine Teilmenge $X_i \subset X$ gegeben, so nennt man

$$\bigcup_{i\in I} X_i := \{x \in X\,;\ \text{es existiert ein } i \in I \text{ mit } x \in X_i\}$$

die *Vereinigung* der Mengen X_i, $i \in I$, sowie

$$\bigcap_{i\in I} X_i := \{x \in X\,;\, x \in X_i \text{ für alle } i \in I\}$$

den *Durchschnitt* dieser Mengen, wobei wir in beiden Fällen wiederum eine Teilmenge von X erhalten. Im Falle einer endlichen Indexmenge $I = \{1, \ldots, n\}$ schreibt man auch $X_1 \cup \ldots \cup X_n$ statt $\bigcup_{i\in I} X_i$ sowie $X_1 \cap \ldots \cap X_n$ statt $\bigcap_{i\in I} X_i$. Zwei Teilmengen $X', X'' \subset X$ werden als *disjunkt* bezeichnet, wenn ihr Durchschnitt leer ist, also $X' \cap X'' = \emptyset$ gilt. Als Variante zur Vereinigung von Mengen $X_i, i \in I$, kann man deren *disjunkte Vereinigung* $\coprod_{i\in I} X_i$ bilden. Hierunter versteht man die Gesamtheit aller Elemente, die in irgendeiner der Mengen X_i enthalten sind, wobei man allerdings für verschiedene Indizes $i, j \in I$ die Elemente von X_i als verschieden von allen Elementen aus X_j ansieht.

(3) Differenz von Mengen. – Sind X_1, X_2 Teilmengen einer Menge X, so heißt

$$X_1 - X_2 := \{x \in X_1\,;\, x \notin X_2\}$$

die *Differenz* von X_1 und X_2. Auch dies ist wieder eine Teilmenge von X, sogar von X_1.

(4) Kartesisches Produkt von Mengen. – Es seien $X_1, \ldots, X_n$ Mengen. Dann heißt

$$\prod_{i=1}^{n} X_i := \{(x_1, \ldots, x_n)\,;\, x_1 \in X_1, \ldots, x_n \in X_n\}$$

das *kartesische Produkt* der Mengen $X_1, \ldots, X_n$, wobei man für dieses Produkt auch die Notation $X_1 \times \ldots \times X_n$ verwendet bzw. X^n, falls $X_1 = \ldots = X_n = X$ gilt. Die Elemente $(x_1, \ldots, x_n)$ werden als *n-Tupel* mit Komponenten $x_i \in X_i$, $i = 1, \ldots, n$, bezeichnet. Es gilt genau dann $(x_1, \ldots, x_n) = (x'_1, \ldots, x'_n)$ für zwei n-Tupel, wenn man $x_i = x'_i$ für $i = 1, \ldots, n$ hat. In ähnlicher Weise lässt sich für eine Indexmenge I das kartesische Produkt $\prod_{i\in I} X_i$ von gegebenen Mengen X_i, $i \in I$, bilden. Man schreibt die Elemente eines solchen Produktes als Familien $(x_i)_{i\in I}$ von Elementen $x_i \in X_i$ und meint damit Tupel, deren Einträge mittels I indiziert werden. Sind die X_i Exemplare ein und derselben Menge X, so verwendet man statt $\prod_{i\in I} X_i$ auch die Notation X^I. Eine Familie $(x_i)_{i\in\emptyset}$, welche durch die leere Indexmenge $I = \emptyset$ indiziert ist, wird als *leer* bezeichnet. Demgemäß bestehen die kartesischen Produkte $\prod_{i\in I} X_i$ und X^I im Falle $I = \emptyset$ aus genau einem Element, nämlich der leeren Familie.

Als Nächstes kommen wir auf den Begriff der Abbildung zwischen Mengen zu sprechen.

Definition 1. *Eine* Abbildung $f\colon X \longrightarrow Y$ *zwischen zwei Mengen X und Y ist eine Vorschrift, welche jedem $x \in X$ ein wohlbestimmtes Element $y \in Y$ zuordnet, das dann mit $f(x)$ bezeichnet wird; man schreibt hierbei auch $x \longmapsto f(x)$. Dabei heißt X der* Definitionsbereich *und Y der* Bild- *oder* Wertebereich *der Abbildung f.*

Zu einer Menge X gibt es stets die *identische Abbildung* $\mathrm{id}_X\colon X \longrightarrow X$, $x \longmapsto x$. Im Übrigen kann man beispielsweise ein kartesisches Produkt des Typs X^I auch als Menge aller Abbildungen $I \longrightarrow X$ interpretieren.

Im Folgenden sei $f\colon X \longrightarrow Y$ wieder eine Abbildung zwischen zwei Mengen. Ist $g\colon Y \longrightarrow Z$ eine weitere Abbildung, so kann man f mit g komponieren; man erhält als Resultat die Abbildung

$$g \circ f\colon X \longrightarrow Z, \qquad x \longmapsto g(f(x)).$$

Für Teilmengen $M \subset X$ und $N \subset Y$ bezeichnet man

$$f(M) := \{y \in Y\,;\ \text{es existiert ein } x \in M \text{ mit } y = f(x)\}$$

als das *Bild* von M unter f sowie

$$f^{-1}(N) := \{x \in X\,;\, f(x) \in N\}$$

als das *Urbild* von N unter f; es handelt sich hierbei um Teilmengen von Y bzw. X. Besteht N aus nur einem einzigen Element y, also $N = \{y\}$, so schreibt man $f^{-1}(y)$ anstelle von $f^{-1}(\{y\})$. Weiter nennt man f *injektiv*, wenn aus $x, x' \in X$ mit $f(x) = f(x')$ stets $x = x'$ folgt, und *surjektiv*, wenn es zu jedem $y \in Y$

ein $x \in X$ mit $f(x) = y$ gibt. Schließlich heißt f *bijektiv*, wenn f injektiv und surjektiv zugleich ist.

Man kann sagen, dass f genau dann injektiv ist, wenn das Urbild $f^{-1}(y)$ eines jeden Punktes $y \in Y$ entweder leer ist oder aus genau einem Punkt $x \in X$ besteht. Weiter ist f genau dann surjektiv, wenn für jedes $y \in Y$ das Urbild $f^{-1}(y)$ nicht leer ist. Somit ist f genau dann bijektiv, wenn für jedes Element $y \in Y$ das Urbild $f^{-1}(y)$ aus genau einem Punkt x besteht. Man kann dann zu f die so genannte *Umkehrabbildung* $g\colon Y \longrightarrow X$ betrachten. Sie ordnet einem Punkt $y \in Y$ das eindeutig bestimmte Element $x \in f^{-1}(y)$ zu, und es gilt $g \circ f = \mathrm{id}_X$ sowie $f \circ g = \mathrm{id}_Y$. Zu einer Abbildung $f\colon X \longrightarrow Y$ bezeichnet man die Umkehrabbildung, sofern diese existiert, meist mit $f^{-1}\colon Y \longrightarrow X$.

Aufgaben

1. Es seien A, B, C Teilmengen einer Menge X. Man zeige:
   (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) $A - (B \cup C) = (A - B) \cap (A - C)$
   (iv) $A - (B \cap C) = (A - B) \cup (A - C)$
2. Es sei $f\colon X \longrightarrow Y$ eine Abbildung zwischen Mengen. Man zeige für Teilmengen $M_1, M_2 \subset X$ und $N_1, N_2 \subset Y$:
   (i) $f(M_1 \cup M_2) = f(M_1) \cup f(M_2)$
   (ii) $f(M_1 \cap M_2) \subset f(M_1) \cap f(M_2)$
   (iii) $f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2)$
   (iv) $f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2)$

   Gilt in (ii) sogar Gleichheit?
3. Es seien $X \xrightarrow{f} Y \xrightarrow{g} X$ Abbildungen von Mengen mit $g \circ f = \mathrm{id}$. Man zeige, dass f injektiv und g surjektiv ist.
4. (i) Gibt es eine bijektive Abbildung $\mathbb{N} \longrightarrow \mathbb{Z}$?
   (ii) Gibt es für $n \in \mathbb{N}$ eine bijektive Abbildung $\mathbb{N} \longrightarrow \mathbb{N} \times \{1, \ldots, n\}$?
   (iii) Gibt es eine bijektive Abbildung $\mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$?
   (iv) Gibt es eine bijektive Abbildung $\mathbb{N} \longrightarrow \mathbb{Q}$?
5. Es sei X eine Menge und $f\colon X \longrightarrow \mathfrak{P}(X)$ eine Abbildung von X in die zugehörige Potenzmenge. Man zeige, dass f nicht surjektiv sein kann.

1.2 Gruppen

Unter einer *inneren Verknüpfung* auf einer Menge M versteht man eine Abbildung $f\colon M \times M \longrightarrow M$. Sie ordnet jedem Paar (a, b) von Elementen aus M ein Element $f(a, b) \in M$ zu. Um den Charakter einer Verknüpfung auch in der Notation zum Ausdruck kommen zu lassen, werden wir anstelle von

$f(a,b)$ meist $a \cdot b$ schreiben. Bei kommutativen Verknüpfungen, also solchen, die $f(a,b) = f(b,a)$ für alle $a, b \in M$ erfüllen, verwenden wir auch die additive Schreibweise $a + b$.

Definition 1. *Eine Menge G mit einer inneren Verknüpfung $G \times G \longrightarrow G$, $(a,b) \longmapsto a \cdot b$, heißt eine* Gruppe, *wenn die folgenden Eigenschaften erfüllt sind*:

(i) *Die Verknüpfung ist* assoziativ, *d. h. es gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ für alle $a, b, c \in G$.*

(ii) *Es existiert ein* neutrales Element *e in G, d. h. ein Element $e \in G$ mit $e \cdot a = a \cdot e = a$ für alle $a \in G$.*[1]

(iii) *Zu jedem $a \in G$ gibt es ein* inverses Element, *d. h. ein Element $b \in G$ mit $a \cdot b = b \cdot a = e$. Dabei ist e das nach* (ii) *existierende* (*eindeutig bestimmte*) *neutrale Element von G.*

Die Gruppe heißt kommutativ *oder* abelsch, *falls die Verknüpfung kommutativ ist, d. h. falls zusätzlich gilt*:

(iv) *$a \cdot b = b \cdot a$ für alle $a, b \in G$.*

In der obigen Situation sagt man gewöhnlich einfach, G sei eine Gruppe, ohne die Verknüpfung "$\cdot$" explizit zu erwähnen. Beispiele für Gruppen sind:

(1) $\mathbb{Z}$ mit der Addition "+"

(2) $\mathbb{Q}$ mit der Addition "+" und $\mathbb{Q}^* := \mathbb{Q} - \{0\}$ mit der Multiplikation "$\cdot$"

(3) $\mathbb{R}$ mit der Addition "+" und $\mathbb{R}^* := \mathbb{R} - \{0\}$ mit der Multiplikation "$\cdot$"

(4) Für eine Menge X ist die Menge $\mathrm{Bij}(X,X)$ der bijektiven Selbstabbildungen $X \longrightarrow X$ eine Gruppe unter der Komposition von Abbildungen als Verknüpfung. Man prüft leicht nach, dass diese Gruppe nicht kommutativ ist, sofern X mindestens 3 paarweise verschiedene Elemente enthält.

Wie bereits behauptet, ist in einer Gruppe G das neutrale Element e eindeutig bestimmt. Ist nämlich $e' \in G$ ein weiteres neutrales Element, so folgt $e = e' \cdot e = e'$. Auf ähnliche Weise zeigt man, dass das zu einem Element $a \in G$ gehörige inverse Element $b \in G$ eindeutig bestimmt ist. Hat man nämlich ein weiteres inverses Element $b' \in G$ zu a, so folgt

$$b = e \cdot b = (b' \cdot a) \cdot b = b' \cdot (a \cdot b) = b' \cdot e = b'.$$

Die gerade durchgeführten Schlüsse benötigen (neben den Eigenschaften von e und b) lediglich, dass e' links-neutral ist, d. h. die Eigenschaft $e' \cdot a = a$ für alle $a \in G$ besitzt, sowie dass b' links-invers zu a ist, d. h. die Gleichung $b' \cdot a = e$ erfüllt. Entsprechend kann man für rechts-neutrale bzw. rechts-inverse Elemente schließen. In einer Gruppe stimmt daher jedes links- (bzw. rechts-) neutrale Element mit dem eindeutigen neutralen Element $e \in G$ überein, ist

[1] Das neutrale Element e ist, wie wir sogleich sehen werden, durch seine definierende Eigenschaft eindeutig bestimmt.

also insbesondere auch rechts- (bzw. links-) neutral. In ähnlicher Weise sieht man, dass links-inverse Elemente auch rechts-invers bzw. rechts-inverse Elemente auch links-invers sind. Wir können sogar noch einen Schritt weitergehen und die definierenden Bedingungen einer Gruppe in diesem Sinne abschwächen:

Bemerkung 2. *Es genügt, in Definition* 1 *anstelle von* (ii) *und* (iii) *lediglich die Existenz eines Elementes $e \in G$ mit folgenden Eigenschaften zu fordern*:
(ii′) *e ist links-neutral in G, d. h. es gilt $e \cdot a = a$ für alle $a \in G$.*
(iii′) *Zu jedem $a \in G$ existiert ein bezüglich e links-inverses Element in G, d. h. ein Element $b \in G$ mit $b \cdot a = e$.*

Beweis. Es sei G eine Menge mit einer multiplikativ geschriebenen Verknüpfung und einem Element $e \in G$, so dass die Bedingungen (i), (ii′) und (iii′) erfüllt sind. Um zu sehen, dass G eine Gruppe ist, haben wir zu zeigen, dass die Bedingungen (ii) und (iii) von Definition 1 gelten. Wir zeigen zunächst für Elemente $a \in G$, dass jedes Element $b \in G$, welches links-invers zu a bezüglich e ist, auch rechts-invers zu a bezüglich e ist. Gelte also $b \cdot a = e$, und sei c ein links-inverses Element zu b, so dass also $c \cdot b = e$ gilt. Hieraus folgt

$$\begin{aligned} a \cdot b = (e \cdot a) \cdot b = ((c \cdot b) \cdot a) \cdot b = (c \cdot (b \cdot a)) \cdot b \\ = (c \cdot e) \cdot b = c \cdot (e \cdot b) = c \cdot b = e, \end{aligned}$$

so dass b rechts-invers zu a bezüglich e ist. Es bleibt noch zu zeigen, dass das links-neutrale Element e auch rechts-neutral ist. Sei also $a \in G$. Ist dann $b \in G$ links-invers zu a bezüglich e, so ist b, wie wir gesehen haben, auch rechts-invers zu a bezüglich e, und es folgt

$$a \cdot e = a \cdot (b \cdot a) = (a \cdot b) \cdot a = e \cdot a = a,$$

also ist e rechts-neutral. □

Gewöhnlich wird das neutrale Element e einer Gruppe G bei multiplikativer Schreibweise der Verknüpfung als *Einselement* bezeichnet, und man schreibt 1 anstelle von e. Für das inverse Element zu $a \in G$ benutzt man die Schreibweise a^{-1}. Im Übrigen ist es bei multiplikativ geschriebenen Gruppenverknüpfungen üblich, das Verknüpfungszeichen "$\cdot$" zu unterdrücken, sofern dies nicht zu Verwechslungen führt. Für endlich viele Elemente $a_1, \ldots, a_n \in G$ definiert man das Produkt dieser Elemente durch

$$\prod_{i=1}^{n} a_i := a_1 \cdot \ldots \cdot a_n.$$

Eine spezielle Klammerung ist hierbei aufgrund des Assoziativgesetzes nicht notwendig; auf einen detaillierten Beweis dieser "offensichtlichen" Tatsache verzichten wir jedoch an dieser Stelle. Wir werden im Folgenden endliche Folgen $a_1, \ldots, a_n \in G$ meist für Indizes $n \in \mathbb{N}$ betrachten, so dass hier insbesondere

auch der Fall $n = 0$ zugelassen ist. Es handelt sich dann um die *leere* Folge, und man erklärt das zugehörige leere Produkt durch

$$\prod_{i=1}^{0} a_i := 1.$$

Wie schon gesagt, verwendet man bei kommutativen Verknüpfungen meist die additive Schreibweise. Das neutrale Element einer kommutativen Gruppe wird dann als *Nullelement* 0 geschrieben und das Inverse zu einem Element $a \in G$ als $-a$. Statt $a + (-a')$ verwendet man üblicherweise die Notation $a - a'$. Endliche Summen von Elementen $a_i \in G$, $i = 1, \ldots, n$, schreibt man in der Form $\sum_{i=1}^{n} a_i$, wobei die leere Summe durch $\sum_{i=1}^{0} a_i := 0$ definiert ist.

Definition 3. *Es sei G eine Gruppe. Eine Teilmenge $H \subset G$ heißt* Untergruppe *von G, wenn gilt*[2]:

(i) $a, b \in H \Longrightarrow ab \in H$,
(ii) $1 \in H$,
(iii) $a \in H \Longrightarrow a^{-1} \in H$.

Ist also $H \subset G$ eine Untergruppe, so induziert die Gruppenverküpfung $G \times G \longrightarrow G$ eine Verknüpfung $H \times H \longrightarrow H$, und H ist mit dieser Verknüpfung selbst wieder eine Gruppe. Umgekehrt, ist Letzteres der Fall, so kann man leicht zeigen, dass H eine Untergruppe von G ist. Im Übrigen sieht man sofort ein, dass eine nicht-leere Teilmenge $H \subset G$ bereits dann eine Untergruppe von G ist, wenn die Bedingung $a, b \in H \Longrightarrow ab^{-1} \in H$ erfüllt ist. Eine Gruppe G enthält stets die trivialen Untergruppen $\{1\}$ und G.

Als Nächstes wollen wir einige elementare Rechenregeln für das Rechnen in Gruppen behandeln. Für Elemente $a, b, c \in G$ gilt:

(1) $\begin{array}{l} ab = ac \Longrightarrow b = c \\ ac = bc \Longrightarrow a = b \end{array}$ (Kürzungsregeln)

(2) $(a^{-1})^{-1} = a$

(3) $(ab)^{-1} = b^{-1}a^{-1}$

Zum Nachweis von (1) multipliziere man von links mit a^{-1} bzw. von rechts mit c^{-1}. Im Falle (2) schließe man wie folgt. $(a^{-1})^{-1}$ ist, wie wir gesehen haben, dasjenige eindeutig bestimmte Element in G, welches (von links oder rechts) mit a^{-1} multipliziert 1 ergibt. Wegen $a^{-1}a = 1$ ergibt sich $(a^{-1})^{-1} = a$. Entsprechend erhält man $(ab)^{-1} = b^{-1}a^{-1}$, da $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}b = 1$ gilt.

Abschließend wollen wir noch eine spezielle Charakterisierung von Gruppen geben.

[2] Nachfolgend steht $\Longrightarrow$ für die so genannte *Implikation*. Für Aussagen A und B schreibt man $A \Longrightarrow B$ oder $B \Longleftarrow A$, wenn B aus A folgt. Entsprechend bedeutet $A \Longleftrightarrow B$, dass A und B äquivalent sind.

Satz 4. *Eine nicht-leere Menge G mit einer Verknüpfung $(a, b) \longmapsto a \cdot b$ ist genau dann eine Gruppe, wenn gilt*:

(i) *Die Verknüpfung ist assoziativ.*

(ii) *Zu $a, b \in G$ gibt es stets Elemente $x, y \in G$ mit $x \cdot a = b$ und $a \cdot y = b$.*

Sind diese Bedingungen erfüllt, so sind die Elemente x, y in (ii) *eindeutig durch a, b bestimmt.*

Beweis. Ist G eine Gruppe, so multipliziere man die Gleichungen in (ii) von links bzw. rechts mit a^{-1}. Es folgt, dass $x = ba^{-1}$ bzw. $y = a^{-1}b$ die eindeutig bestimmten Lösungen sind. Seien nun umgekehrt die Bedingungen des Satzes erfüllt, und sei $a \in G$. Dann existiert nach (ii) ein Element $e \in G$ mit $ea = a$. Zu $b \in G$ existiert weiter ein $y \in G$ mit $ay = b$, und es folgt

$$eb = eay = ay = b,$$

also ist e links-neutral. Weiter folgt die Existenz links-inverser Elemente nach (ii). Somit ist G eine Gruppe nach Bemerkung 2. □

Aufgaben

1. Für eine Menge X betrachte man die Menge $\mathrm{Bij}(X, X)$ der bijektiven Selbstabbildungen. Man prüfe nach, dass $\mathrm{Bij}(X, X)$ unter der Komposition von Abbildungen eine Gruppe bildet und zeige, dass diese nicht kommutativ ist, sofern X mindestens 3 verschiedene Elemente besitzt.

2. Es sei G eine Gruppe und $H \subset G$ eine Teilmenge. Man zeige, dass H genau dann eine Untergruppe von G ist, wenn die Gruppenverknüpfung von G eine Verknüpfung auf H induziert (d. h. wenn für $a, b \in H$ stets $ab \in H$ gilt) und wenn H mit dieser Verknüpfung selbst wieder eine Gruppe ist.

3. Es sei G eine Gruppe und $H \subset G$ eine Teilmenge. Man zeige, dass H genau dann eine Untergruppe von G ist, wenn gilt:
   (i) $H \neq \emptyset$
   (ii) $a, b \in H \Longrightarrow ab^{-1} \in H$

4. Es sei G eine Gruppe mit Untergruppen $H_1, H_2 \subset G$. Man zeige, dass $H_1 \cup H_2$ genau dann eine Untergruppe von G ist, wenn $H_1 \subset H_2$ oder $H_2 \subset H_1$ gilt.

5. Für eine Gruppe G betrachte man die Abbildung $i\colon G \longrightarrow G$, $g \longmapsto g^{-1}$. Man zeige:
   (i) i ist bijektiv.
   (ii) Ist $A \subset G$ eine Teilmenge mit $i(A) \subset A$, so gilt bereits $i(A) = A$; man nennt A dann *symmetrisch*.
   (iii) Für jede Teilmenge $A \subset G$ sind $A \cup i(A)$ und $A \cap i(A)$ symmetrisch.

6. Es sei G eine Gruppe mit $a^2 = 1$ für alle $a \in G$. Man zeige, dass G abelsch ist.

7. Es sei G eine endliche abelsche Gruppe. Dann gilt $\prod_{g \in G} g^2 = 1$.

8. Für ein $n \in \mathbb{N} - \{0\}$ betrachte man die Teilmenge $R_n = \{0, 1, \ldots, n-1\} \subset \mathbb{N}$. Es sei $\pi\colon \mathbb{Z} \longrightarrow R_n$ die Abbildung, welche einer ganzen Zahl aus $\mathbb{Z}$ jeweils deren nicht-negativen Rest bei Division durch n zuordnet. Man zeige:
   (i) Es existiert eine eindeutig bestimmte Verknüpfung $(a, b) \longmapsto a + b$ auf R_n, so dass für $x, y \in \mathbb{Z}$ stets $\pi(x + y) = \pi(x) + \pi(y)$ gilt.
   (ii) R_n ist mit dieser Verknüpfung eine abelsche Gruppe.
9. Es sei G eine Gruppe. Auf der Potenzmenge $\mathfrak{P}(G)$ betrachte man die durch

$$(A, B) \longmapsto A \cdot B = \{a \cdot b \in G\,;\, a \in A, b \in B\}$$

gegebene Verknüpfung. Man zeige, dass diese Verknüpfung assoziativ ist und ein neutrales Element besitzt. Ist $\mathfrak{P}(G)$ mit dieser Verknüpfung sogar eine Gruppe? Falls nein, zu welchen Elementen $A \in \mathfrak{P}(G)$ gibt es inverse Elemente?

1.3 Körper

Ein Körper ist eine additiv geschriebene abelsche Gruppe, auf der zusätzlich eine Multiplikation mit gewissen Eigenschaften definiert ist, nach dem Vorbild der rationalen oder der reellen Zahlen. Genauer:

Definition 1. *Ein* Körper *ist eine Menge K mit zwei inneren Verknüpfungen, geschrieben als Addition "+" und Multiplikation "$\cdot$", so dass folgende Bedingungen erfüllt sind*:
(i) $(a + b) + c = a + (b + c)$ *für* $a, b, c \in K$ (Assoziativgesetz der Addition).
(ii) *Es existiert ein Element* $0 \in K$ *mit* $0 + a = a$ *für alle* $a \in K$ (neutrales Element der Addition).
(iii) *Zu* $a \in K$ *existiert ein Element* $b \in K$ *mit* $b + a = 0$ (inverses Element der Addition).
(iv) $a + b = b + a$ *für* $a, b \in K$ (Kommutativgesetz der Addition).
(v) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ *für* $a, b, c \in K$ (Assoziativgesetz der Multiplikation).
(vi) *Es existiert ein Element* $1 \in K$ *mit* $1 \cdot a = a$ *für alle* $a \in K$ (neutrales Element der Multiplikation).
(vii) *Zu* $a \in K - \{0\}$ *existiert ein Element* $b \in K$ *mit* $b \cdot a = 1$ (inverses Element der Multiplikation).
(viii) $a \cdot b = b \cdot a$ *für* $a, b \in K$ (Kommutativgesetz der Multiplikation).
(ix) $a \cdot (b + c) = a \cdot b + a \cdot c$ *und* $(a + b) \cdot c = a \cdot c + b \cdot c$ *für* $a, b, c \in K$ (Distributivgesetze).
(x) $1 \neq 0$.

Bei den Distributivgesetzen (ix) hätten wir eigentlich auf der rechten Seite die Terme $a \cdot b$, $a \cdot c$, $b \cdot c$ jeweils in Klammern setzen müssen. Man vereinbart jedoch, dass die Multiplikation "$\cdot$" Vorrang vor der Addition "+" hat, so dass Klammerungen dann entbehrlich sind. Auch sei darauf hingewiesen, dass das Multiplikationszeichen "$\cdot$", ähnlich wie im Falle von Gruppen, vielfach nicht

ausgeschrieben wird. Schließlich nennt man 0 das *Nullelement* und 1 das *Einselement* von K.

Als Nächstes wollen wir einige simple Rechenregeln für das Rechnen in Körpern K behandeln.

(1) $0a = a0 = 0$ für $a \in K$, denn es gilt

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

(2) $(-1)a = -a$ für $a \in K$, denn

$$a + (-1)a = 1a + (-1)a = (1-1)a = 0a = 0.$$

(3) $(-a)b = a(-b) = -ab$, $(-a)(-b) = ab$ für $a, b \in K$; dies ergibt sich unter Benutzung von (2).

(4) Für $a, b \in K$ folgt aus $ab = 0$ bereits $a = 0$ oder $b = 0$. Denn aus $ab = 0$ mit $a \neq 0 \neq b$ würde sich sonst als Widerspruch

$$1 = abb^{-1}a^{-1} = 0b^{-1}a^{-1} = 0$$

ergeben.

Man kann also in Körpern in etwa so rechnen, wie man dies von den rationalen oder reellen Zahlen her gewohnt ist. Doch sei schon an dieser Stelle auf Unterschiede zum Vorbild vertrauter Zahlbereiche hingewiesen. Für eine natürliche Zahl $n \in \mathbb{N}$ und ein Element $a \in K$ ist es üblich, die n-fache Summe von a mit sich selbst als $n \cdot a$ zu bezeichnen, wobei dann insbesondere $n \cdot a = 0$ für $n = 0$ oder $a = 0$ gilt. Weiter setzt man $n \cdot a = (-n) \cdot (-a)$ für negative ganze Zahlen n. Es folgt jedoch aus $n \cdot a = 0$ nicht notwendig $n = 0$ oder $a = 0$, wie wir an konkreten Beispielen noch feststellen werden.

Unter Verwendung des Gruppenbegriffs lassen sich Körper in übersichtlicher Weise wie folgt charakterisieren:

Bemerkung 2. *Die Bedingungen* (i) - (x) *in Definition* 1 *sind äquivalent zu den folgenden Bedingungen*:

(i) *K ist eine abelsche Gruppe bezüglich der Addition.*
(ii) *$K^* = K - \{0\}$ ist eine abelsche Gruppe bezüglich der Multiplikation.*
(iii) *Es gelten die Distributivgesetze* (ix) *aus Definition* 1.

Beweis. Zunächst ist klar, dass die Bedingungen (i) - (iv) aus Definition 1 diejenigen einer kommutativen additiven Gruppe sind. Weiter folgt aus obiger Regel (4), dass für einen Körper K die Teilmenge $K^* = K - \{0\}$ abgeschlossen unter der Multiplikation ist und dass mit einem Element $a \in K^*$ wegen $a \cdot a^{-1} = 1$ auch dessen inverses a^{-1} zu K^* gehört. Somit sieht man, dass K^* eine abelsche Gruppe bezüglich der Multiplikation ist, und es implizieren die Bedingungen aus Definition 1 die Bedingungen von Bemerkung 2.

Seien nun umgekehrt die Bedingungen aus Bemerkung 2 erfüllt. Um hieraus die Bedingungen von Definition 1 abzuleiten, braucht man lediglich zu wissen,

dass in der Situation von Bemerkung 2 die Beziehung $0a = 0 = a0$ für alle $a \in K$ gilt. Diese kann man jedoch mit Hilfe der Distributivgesetze auf gleiche Weise herleiten, wie wir dies bereits oben bei den Rechenregeln getan haben. □

Ähnlich wie bei Gruppen hat man auch bei Körpern den Begriff des Unter- oder Teilkörpers.

Definition 3. *Es sei K ein Körper. Eine Teilmenge $L \subset K$ heißt ein* Teilkörper *von K, wenn gilt*:

(i) $a, b \in L \Longrightarrow a + b, a \cdot b \in L$.
(ii) $0, 1 \in L$.
(iii) $a \in L \Longrightarrow -a \in L$.
(iv) $a \in L$, $a \neq 0 \Longrightarrow a^{-1} \in L$.

Es ist klar, dass eine Teilmenge $L \subset K$ genau dann ein Teilkörper von K ist, wenn Addition und Multiplikation auf K sich zu Verknüpfungen $L \times L \longrightarrow L$ einschränken und wenn L unter diesen Verknüpfungen selbst ein Körper ist. Bekannte Beispiele für Körper sind die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$, wobei $\mathbb{Q}$ ein Teilkörper von $\mathbb{R}$ ist. Ein Körper enthält mindestens 2 verschiedene Elemente, nämlich das neutrale Element der Addition und das neutrale Element der Multiplikation, also 0 und 1. Andererseits gibt es aber auch einen Körper K, der aus genau 2 Elementen besteht. Man betrachte nämlich die Teilmenge $\{0, 1\} \subset \mathbb{Z}$ und setze:

$$\begin{aligned} 0 + 0 = 0, \qquad & 0 + 1 = 1 + 0 = 1, \qquad & 1 + 1 = 0, \\ 0 \cdot 0 = 0, \qquad & 0 \cdot 1 = 1 \cdot 0 = 0, \qquad & 1 \cdot 1 = 1. \end{aligned}$$

Eine Verifikation der Körperaxiome zeigt, dass diese Verknüpfungen auf $\{0, 1\}$ in der Tat die Struktur eines Körpers definieren; man bezeichnet diesen meist mit $\mathbb{F}_2$. Natürlich ist $\mathbb{F}_2$ kein Teilkörper von $\mathbb{Q}$ oder $\mathbb{R}$, denn es gilt $2 \cdot 1 = 1 + 1 = 0$, wobei 2 als natürliche Zahl, nicht aber als Element von $\mathbb{F}_2$ aufzufassen ist.

Als Nächstes wollen wir den kleinsten Teilkörper von $\mathbb{R}$ konstruieren, der $\sqrt{2}$ enthält, also diejenige positive reelle Zahl, die mit sich selbst multipliziert 2 ergibt. Dieser Körper wird üblicherweise mit $\mathbb{Q}(\sqrt{2})$ bezeichnet. Zunächst zeigen wir:

Lemma 4. $\sqrt{2} \notin \mathbb{Q}$.

Beweis. Wir führen den Beweis indirekt, also durch Widerspruch, und nehmen $\sqrt{2} \in \mathbb{Q}$ an, etwa $\sqrt{2} = p/q$ mit $p, q \in \mathbb{Z} - \{0\}$. Den Bruch p/q können wir als gekürzt annehmen. Insbesondere sind dann p und q nicht beide durch 2 teilbar. Aus der Gleichung $p^2/q^2 = 2$ ergibt sich $p^2 = 2q^2$ und damit, dass p^2 gerade ist. Da das Quadrat einer ungeraden Zahl stets ungerade ist, muss auch p gerade sein, etwa $p = 2\tilde{p}$ mit einem Element $\tilde{p} \in \mathbb{Z}$. Es folgt $2q^2 = 4\tilde{p}^2$ bzw. $q^2 = 2\tilde{p}^2$ und damit wie soeben, dass 2 ein Teiler von q ist. Damit ist 2 sowohl

ein Teiler von p wie auch von q. Dies hatten wir jedoch zuvor ausgeschlossen. Die Annahme $\sqrt{2} \in \mathbb{Q}$ führt daher zu einem Widerspruch, ist folglich nicht haltbar, und es gilt $\sqrt{2} \notin \mathbb{Q}$. □

Als Folgerung erhalten wir:

Lemma 5. *Für $a, b \in \mathbb{Q}$ gilt*

$$a + b\sqrt{2} \neq 0 \iff a \neq 0 \text{ oder } b \neq 0.$$

Beweis. Die Implikation "$\Longrightarrow$" ist trivial. Um die Umkehrung "$\Longleftarrow$" zu zeigen, gehen wir wieder indirekt vor und nehmen an, es gäbe Zahlen $a, b \in \mathbb{Q}$ mit $a + b\sqrt{2} = 0$, wobei a und b nicht beide verschwinden mögen. Dann folgt notwendig $a \neq 0 \neq b$ und somit $\sqrt{2} = -ab^{-1} \in \mathbb{Q}$ im Widerspruch zu Lemma 4. □

Wir definieren nun $\mathbb{Q}(\sqrt{2})$ als Teilmenge von $\mathbb{R}$ durch

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}\,;\, a, b \in \mathbb{Q}\}.$$

Satz 6. *$\mathbb{Q}(\sqrt{2})$ ist ein echter Teilkörper von $\mathbb{R}$, der wiederum $\mathbb{Q}$ als echten Teilkörper enthält. Es ist $\mathbb{Q}(\sqrt{2})$ der kleinste Teilkörper von $\mathbb{R}$, der $\sqrt{2}$ enthält.*

Beweis. Zunächst soll gezeigt werden, dass $\mathbb{Q}(\sqrt{2})$ ein Teilkörper von $\mathbb{R}$ ist. Um die Abgeschlossenheit von $\mathbb{Q}(\sqrt{2})$ unter der Addition und Multiplikation zu zeigen, betrachte man Elemente $a + b\sqrt{2}$, $a' + b'\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ mit $a, b, a', b' \in \mathbb{Z}$. Dann folgt

$$\begin{aligned}(a + b\sqrt{2}) + (a' + b'\sqrt{2}) &= (a + a') + (b + b')\sqrt{2} && \in \mathbb{Q}(\sqrt{2}),\\ (a + b\sqrt{2}) \cdot (a' + b'\sqrt{2}) &= (aa' + 2bb') + (ab' + a'b)\sqrt{2} && \in \mathbb{Q}(\sqrt{2}),\end{aligned}$$

d. h. Bedingung (i) aus Definition 3 ist erfüllt. Dasselbe gilt für Bedingung (ii), denn $0 = 0 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ und $1 = 1 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Weiter ist mit $a + b\sqrt{2}$ auch $-(a + b\sqrt{2}) = (-a) + (-b)\sqrt{2}$ als inverses Element bezüglich der Addition in $\mathbb{Q}(\sqrt{2})$ enthalten, so dass auch Bedingung (iii) aus Definition 3 erfüllt ist.

Etwas schwieriger ist Bedingung (iv) aus Definition 3 nachzuweisen. Sei $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ von Null verschieden, also $a \neq 0$ oder $b \neq 0$ nach Lemma 5. Dann gilt $a - b\sqrt{2} \neq 0$, ebenfalls nach Lemma 5, und wir können schreiben:

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} \quad \in \mathbb{Q}(\sqrt{2}).$$

Insgesamt ergibt sich, dass $\mathbb{Q}(\sqrt{2})$ ein Teilkörper von $\mathbb{R}$ ist, und zwar ein echter Teilkörper, da beispielsweise $\sqrt{3}$ nicht zu $\mathbb{Q}(\sqrt{2})$ gehört. Letzteres zeigt man, indem man ähnlich argumentiert wie im Beweis zu Lemma 4. Im Übrigen enthält $\mathbb{Q}(\sqrt{2})$ den Körper der rationalen Zahlen als echten Teilkörper wegen $\sqrt{2} \notin \mathbb{Q}$.

Es bleibt noch zu zeigen, dass $\mathbb{Q}(\sqrt{2})$ der kleinste Teilkörper von $\mathbb{R}$ ist, der $\sqrt{2}$ enthält. Ist zunächst K ein beliebiger Teilkörper von $\mathbb{R}$, so enthält K notwendig alle Elemente der Form $n \cdot 1$ mit $n \in \mathbb{Z}$, es gilt also $\mathbb{Z} \subset K$. Dann muss K aber auch alle Brüche der Form p/q mit $p, q \in \mathbb{Z}$, $q \neq 0$, und damit $\mathbb{Q}$ enthalten. Folglich ist $\mathbb{Q}$ der kleinste Teilkörper von $\mathbb{R}$. Gilt nun $\sqrt{2} \in K$, so enthält K notwendig auch alle Ausdrücke der Form $a + b\sqrt{2}$ mit $a, b \in \mathbb{Q}$ und damit $\mathbb{Q}(\sqrt{2})$. Also ist $\mathbb{Q}(\sqrt{2})$ der (eindeutig bestimmte) kleinste Teilkörper von $\mathbb{R}$, der $\sqrt{2}$ enthält. □

Als Nächstes wollen wir von dem Körper $\mathbb{R}$ der reellen Zahlen ausgehen und diesen zum Körper $\mathbb{C}$ der komplexen Zahlen erweitern. Man setze

$$\mathbb{C} := \mathbb{R} \times \mathbb{R} = \{(a, a')\,;\, a, a' \in \mathbb{R}\}$$

und definiere Addition bzw. Multiplikation auf $\mathbb{C}$ durch

$$\begin{aligned}(a, a') + (b, b') &:= (a + b, a' + b'),\\ (a, a') \cdot (b, b') &:= (ab - a'b', ab' + a'b).\end{aligned}$$

Man prüft leicht nach, dass $\mathbb{C}$ mit diesen Verknüpfungen einen Körper bildet. Dabei ist $0_{\mathbb{C}} = (0, 0)$ das Nullelement sowie $-(a, a') = (-a, -a')$ das inverse Element bezüglich der Addition zu $(a, a') \in \mathbb{C}$. Weiter ist $1_{\mathbb{C}} = (1, 0)$ das Einselement von $\mathbb{C}$, und das inverse Element bezüglich der Multiplikation zu einem Element $(a, a') \neq 0_{\mathbb{C}}$ wird gegeben durch

$$(a, a')^{-1} = \left(\frac{a}{a^2 + a'^2}, -\frac{a'}{a^2 + a'^2}\right).$$

Exemplarisch wollen wir das Assoziativgesetz der Multiplikation nachweisen. Für $(a, a'), (b, b'), (c, c') \in \mathbb{C}$ rechnet man

$$\begin{aligned}&\big((a, a')(b, b')\big)(c, c') = (ab - a'b', ab' + a'b)(c, c')\\ &= (abc - a'b'c - ab'c' - a'bc', abc' - a'b'c' + ab'c + a'bc)\end{aligned}$$

sowie

$$\begin{aligned}&(a, a')\big((b, b')(c, c')\big) = (a, a')(bc - b'c', bc' + b'c)\\ &= (abc - ab'c' - a'bc' - a'b'c, abc' + ab'c + a'bc - a'b'c'),\end{aligned}$$

d. h. es gilt

$$\big((a, a')(b, b')\big)(c, c') = (a, a')\big((b, b')(c, c')\big).$$

Man stellt weiter fest, dass die Elemente der Form $(a, 0)$ einen Teilkörper $K \subset \mathbb{C}$ bilden. Es gilt nämlich $0_{\mathbb{C}}, 1_{\mathbb{C}} \in K$ sowie für $(a, 0), (b, 0) \in K$

$$\begin{aligned}(a, 0) + (b, 0) &= (a + b, 0) \in K,\\ (a, 0) \cdot (b, 0) &= (a \cdot b, 0) \in K,\\ -(a, 0) &= (-a, 0) \in K,\\ (a, 0)^{-1} &= (a^{-1}, 0) \in K, \text{ falls } a \neq 0.\end{aligned}$$

Man kann nun durch $a \longmapsto (a,0)$ eine natürliche Identifikation zwischen den Elementen von $\mathbb{R}$ und denen von K erklären. Da diese Identifikation auch die Körperstrukturen von $\mathbb{R}$ bzw. K respektiert, lässt sich $\mathbb{R}$ sogar als Körper mit dem Teilkörper $K \subset \mathbb{C}$ identifizieren. Somit können wir nun $\mathbb{R}$ als Teilkörper von $\mathbb{C}$ auffassen und brauchen nicht mehr zwischen dem Null- bzw. Einselement in $\mathbb{R}$ und $\mathbb{C}$ zu unterscheiden.

Üblicherweise bezeichnet man das Element $(0,1) \in \mathbb{C}$ als komplexe Zahl i; diese besitzt die Eigenschaft $i^2 = -1$, ist also zu interpretieren als Quadratwurzel aus -1. Komplexe Zahlen $z = (a, a')$ lassen sich sodann in der Form

$$z = (a,0) + (0,a') = (a,0) + (a',0) \cdot (0,1) = a + a'i$$

schreiben. Dabei wird a als *Realteil* und a' als *Imaginärteil* von z bezeichnet. Es gelten die Formeln

$$\begin{aligned}
(a + a'i) + (b + b'i) &= (a+b) + (a'+b')i, \\
(a + a'i) \cdot (b + b'i) &= (ab - a'b') + (ab' + a'b)i, \\
-(a + a'i) &= -a - a'i, \\
(a + a'i)^{-1} &= \frac{a}{a^2 + a'^2} - \frac{a'}{a^2 + a'^2} i,
\end{aligned}$$

letztere unter der Voraussetzung $a + a'i \neq 0$, also $a \neq 0$ oder $a' \neq 0$.

Als Beispiel für das Rechnen in Körpern wollen wir schließlich noch die binomische Formel herleiten. Sei also K ein beliebiger Körper. Für $a \in K$ und $n \in \mathbb{N}$ definiert man üblicherweise a^n als das n-fache Produkt von a mit sich selbst. Dabei ist a^0 das leere Produkt, also $a^0 = 1$. Außerdem kann man a^{-n} für $a \neq 0$ durch $(a^{-1})^n$ erklären, so dass dann a^n für ganzzahlige Exponenten n definiert ist. Für das Rechnen mit solchen Potenzen gelten die gewöhnlichen Potenzgesetze.

Seien $a, b \in K$, und sei $n \in \mathbb{N}$ eine natürliche Zahl. Zur Berechnung von $(a+b)^n$ wählen wir zunächst eine kombinatorische Methode. Hierzu stellen wir uns $(a+b)^n$ als n-faches Produkt vor:

$$(a+b)^n = (a+b) \cdot \ldots \cdot (a+b)$$

Die rechte Seite kann man unter sukzessiver Benutzung der Distributivgesetze ausrechnen, indem man aus jeder Klammer einen Summanden auswählt (also jeweils a oder b), das Produkt über die ausgewählten Elemente bildet und schließlich alle Produkte dieses Typs zu verschiedenen Wahlen summiert. Somit folgt

$$(a+b)^n = \sum_{i=0}^{n} \alpha(i) a^{n-i} b^i,$$

wobei $\alpha(i)$ gleich der Anzahl der Möglichkeiten ist, den Summanden b genau i-mal aus den n Klammern $(a+b)$ auszuwählen, mit anderen Worten, gleich der Anzahl der i-elementigen Teilmengen in $\{1, \ldots, n\}$. Will man i Elemente in $\{1, \ldots, n\}$ auswählen, so gibt es für das erste Element n Wahlmöglichkeiten, für

das zweite $n-1$ und so weiter, schließlich für das i-te Element noch $n-i+1$ Möglichkeiten. Insgesamt haben wir daher

$$n(n-1)\dots(n-i+1)$$

Möglichkeiten für diesen Auswahlprozess. Nun ist aber zu berücksichtigen, dass eine i-elementige Teilmenge $\{t_1,\dots,t_i\}$ von $\{1,\dots,n\}$, die in einem solchen Prozess konstruiert wird, nicht davon abhängt, in welcher Reihenfolge die Elemente $t_1,\dots,t_i$ ausgewählt werden. Wir müssen daher die obige Anzahl noch durch die Anzahl der Möglichkeiten dividieren, die Elemente $t_1,\dots,t_i$ in ihrer Reihenfolge zu vertauschen, also durch die Anzahl der bijektiven Selbstabbildungen $\pi\colon \{1,\dots,i\} \longrightarrow \{1,\dots,i\}$. Will man eine solche Abbildung π definieren, so hat man zur Festsetzung von $\pi(1)$ zunächst i Möglichkeiten, für $\pi(2)$ noch $i-1$ Möglichkeiten usw. Die Anzahl der bijektiven Selbstabbildungen von $\{1,\dots,i\}$ ist deshalb $i! = 1\cdot\ldots\cdot i$, und es ergibt sich

$$\alpha(i) = \frac{n(n-1)\dots(n-i+1)}{1\cdot 2\cdot\ldots\cdot i},$$

wobei man hierfür auch $\binom{n}{i}$ schreibt, also

$$\binom{n}{i} = \frac{n(n-1)\dots(n-i+1)}{1\cdot 2\cdot\ldots\cdot i} = \frac{n!}{i!(n-i)!}, \qquad 0\le i\le n.$$

In den Extremfällen $i=0$ bzw. $i=n$ erweist sich unsere Konvention bezüglich leerer Produkte als sinnvoll, es gilt $0!=1$ sowie $\binom{n}{0}=1=\binom{n}{n}$ und insbesondere $\binom{0}{0}=1$. Insgesamt folgt die bekannte *binomische Formel*

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i}b^i.$$

Die Koeffizienten $\binom{n}{i}\in\mathbb{N}$ werden als *Binomialkoeffizienten* bezeichnet.

Wir wollen noch einen präziseren Beweis für diese Formel geben, wobei wir die Gelegenheit nutzen, um das Prinzip der *vollständigen Induktion* zu erklären. Wenn man zeigen will, dass eine Aussage $A(n)$ für alle natürlichen Zahlen $n\in\mathbb{N}$ gültig ist, so genügt es nach diesem Prinzip, Folgendes zu zeigen:

(1) Es gilt $A(0)$ (Induktionsanfang).

(2) Für beliebiges $n\in\mathbb{N}$ kann man aus der Gültigkeit von $A(n)$ (Induktionsvoraussetzung) auf die Gültigkeit von $A(n+1)$ schließen (Induktionsschluss).

Natürlich kann man die vollständige Induktion statt bei $n=0$ auch bei einer anderen Zahl $n=n_0\in\mathbb{N}$ oder sogar bei einer Zahl $n=n_0\in\mathbb{Z}$ beginnen. Führt man den Induktionsschluss dann für ganze Zahlen $n\ge n_0$ durch, so ergibt sich die Gültigkeit von $A(n)$ für alle ganzen Zahlen $n\ge n_0$. Als Variante dieses Prinzips darf man beim Induktionsschluss zum Nachweis von $A(n+1)$ zusätzlich benutzen, dass die Aussage $A(m)$ bereits für alle m mit $n_0\le m\le n$ gilt, wobei

der Induktionsanfang wiederum bei $n = n_0$ liegen möge. In unserem Fall soll die Aussage $A(n)$ aus zwei Teilen bestehen und für $n \in \mathbb{N}$ wie folgt lauten:

$$\binom{n}{i} \in \mathbb{N} \quad \text{für} \quad 0 \le i \le n,$$
$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i;$$

die Binomialkoeffizienten $\binom{n}{i}$ sind dabei wie oben durch

$$\binom{n}{i} = \frac{n(n-1)\dots(n-i+1)}{1 \cdot 2 \cdot \dots \cdot i} = \frac{n!}{i!(n-i)!}$$

gegeben. Der Induktionsanfang bei $n = 0$ ist leicht durchzuführen; denn man hat $\binom{0}{0} = 1 \in \mathbb{N}$ und $(a+b)^0 = 1 = \binom{0}{0} a^0 b^0$, d. h. $A(0)$ ist richtig. Zum Induktionsschluss betrachten wir ein beliebiges $n \in \mathbb{N}$ und nehmen an, dass $A(n)$ richtig ist. Dann können wir wie folgt rechnen:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n = (a+b) \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i \\
&= \sum_{i=0}^{n} \binom{n}{i} a^{n+1-i} b^i + \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^{i+1} \\
&= a^{n+1} + \sum_{i=1}^{n} \binom{n}{i} a^{n+1-i} b^i + \sum_{i=0}^{n-1} \binom{n}{i} a^{n-i} b^{i+1} + b^{n+1} \\
&= a^{n+1} + \sum_{i=1}^{n} \binom{n}{i} a^{n+1-i} b^i + \sum_{i=1}^{n} \binom{n}{i-1} a^{n+1-i} b^i + b^{n+1} \\
&= a^{n+1} + \sum_{i=1}^{n} \left[\binom{n}{i} + \binom{n}{i-1} \right] a^{n+1-i} b^i + b^{n+1}
\end{aligned}$$

Nun hat man aber für $1 \le i \le n$

$$\begin{aligned}
\binom{n}{i} + \binom{n}{i-1} &= \frac{n!}{i!(n-i)!} + \frac{n!}{(i-1)!(n-i+1)!} \\
&= \frac{n!(n-i+1) + n!i}{i!(n-i+1)!} = \frac{n!(n+1)}{i!(n-i+1)!} \\
&= \frac{(n+1)!}{i!(n+1-i)!} = \binom{n+1}{i},
\end{aligned}$$

so dass sich wie gewünscht

$$(a+b)^{n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} a^{n+1-i} b^i$$

ergibt. Außerdem folgt aus $\binom{n}{i}, \binom{n}{i-1} \in \mathbb{N}$, dass auch $\binom{n+1}{i}$ eine natürliche Zahl ist. Die binomische Formel ist daher per Induktion bewiesen.

Aufgaben

1. Es sei K eine endliche Menge mit zwei Verknüpfungen "+" und "$\cdot$", welche den Bedingungen (i) – (x) von Definition 1 genügen, wobei jedoch die Bedingung (vii) ersetzt sei durch

   (vii′) Für $a, b \in K - \{0\}$ gilt $ab \in K - \{0\}$.

   Man zeige, dass K ein Körper ist.

2. Es sei K ein endlicher Körper. Für $n \in \mathbb{N}$ und $a \in K$ bezeichne $na = a + \ldots + a$ die n-fache Summe von a mit sich selber.
   (i) Es existiert ein $n \in \mathbb{N} - \{0\}$, so dass $na = 0$ für alle $a \in K$ gilt.
   (ii) Wählt man n wie vorstehend minimal, so ist n eine Primzahl, die so genannte *Charakteristik* von K.

3. Man betrachte für $n \in \mathbb{N} - \{0\}$ die Menge R_n aus Abschnitt 1.2, Aufgabe 8 mit der dort erklärten Addition, welche auf R_n die Struktur einer additiven abelschen Gruppe definiert. Man zeige:
   (i) Auf R_n lässt sich in eindeutiger Weise eine Multiplikation erklären, so dass alle Bedingungen von Definition 1, mit eventueller Ausnahme von (vii) erfüllt sind.
   (ii) Ist p eine Primzahl, so ist R_p sogar ein Körper; dieser wird auch mit $\mathbb{F}_p$ bezeichnet.

4. Man konstruiere einen Körper mit 4 Elementen.

5. Man weise nach, dass $\sqrt{3}$ nicht zu $\mathbb{Q}(\sqrt{2})$ gehört.

6. Man bestimme den kleinsten Teilkörper von $\mathbb{C}$, welcher die komplexe Zahl i enthält.

7. Für eine Aussage $A(n)$, die für $n \in \mathbb{N}$ definiert ist, betrachte man folgende Bedingungen:
   (i) $A(0)$ ist wahr.
   (ii) Für alle $n \in \mathbb{N}$ gilt: Ist $A(n)$ wahr, so auch $A(n+1)$.
   (iii) Für alle $n \in \mathbb{N}$ gilt: Ist $A(i)$ für alle $i \in \mathbb{N}$ mit $i \leq n$ wahr, so auch $A(n+1)$.

   Man zeige mittels eines formalen Schlusses, dass das Induktionsprinzip, welches die Bedingungen (i) und (ii) umfasst, äquivalent zu demjenigen ist, das die Bedingungen (i) und (iii) umfasst.

8. Es sei $A(m, n)$ eine Aussage, die für $m, n \in \mathbb{N}$ erklärt sei. Die folgenden Aussagen seien wahr:
   (i) $A(0, 0)$
   (ii) $A(i, j) \Longrightarrow A(i+1, j)$ für $i, j \in \mathbb{N}$.
   (iii) $A(i, j) \Longrightarrow A(i, j+1)$ für $i, j \in \mathbb{N}$.

Man zeige, dass dann $A(i,j)$ für alle $i,j \in \mathbb{N}$ wahr ist (*Prinzip der Doppelinduktion*). Lassen sich die Bedingungen (ii) bzw. (iii) noch abschwächen?

9. Für $n \in \mathbb{N}$ und Elemente $q \neq 1$ eines Körpers K leite man die Formel für die *geometrische Reihe* her:
$$\sum_{i=0}^{n} q^i = \frac{1-q^{n+1}}{1-q}$$

10. Man beweise für $k,n \in \mathbb{N}$ mit $n \geq k \geq 1$:
$$\sum_{i=k-1}^{n-1} \binom{i}{k-1} = \binom{n}{k}$$

11. Für $k,n \in \mathbb{N}$ zeige man, dass die Menge $\{(a_1,\ldots,a_n) \in \mathbb{N}^n\,;\, a_1+\ldots+a_n = k\}$ genau
$$\binom{k+n-1}{n-1}$$
Elemente besitzt.

1.4 Vektorräume

Wir wollen nun die eingangs angedeutete Vektorrechnung auf eine axiomatische Grundlage stellen, indem wir Vektorräume über Körpern betrachten. Vektoren werden wir im Folgenden stets mit lateinischen Buchstaben $a,b,c,\ldots$ bezeichnen, Skalare aus dem zugehörigen Körper dagegen mit griechischen Buchstaben $\alpha,\beta,\gamma,\ldots$

Definition 1. *Es sei K ein Körper. Ein* K-Vektorraum *ist eine Menge V mit einer inneren Verknüpfung $V \times V \longrightarrow V$, $(a,b) \longmapsto a+b$, genannt* Addition, *und einer äußeren Verknüpfung $K \times V \longrightarrow V$, genannt* skalare Multiplikation, *so dass gilt*:

(i) *V ist eine abelsche Gruppe bezüglich der Addition "+".*

(ii) *$(\alpha+\beta)\cdot a = \alpha\cdot a + \beta\cdot a$, $\alpha\cdot(a+b) = \alpha\cdot a + \alpha\cdot b$ für alle $\alpha,\beta \in K$, $a,b \in V$, d. h. Addition und Multiplikation verhalten sich distributiv.*

(iii) *$(\alpha\cdot\beta)\cdot a = \alpha\cdot(\beta\cdot a)$ für alle $\alpha,\beta \in K$, $a \in V$, d. h. die skalare Multiplikation ist assoziativ.*

(iv) *$1\cdot a = a$ für das Einselement $1 \in K$ und alle $a \in V$.*

Elemente eines Vektorraums werden auch als Vektoren bezeichnet. Wie jede Gruppe enthält ein K-Vektorraum mindestens ein Element, nämlich den Nullvektor 0 als neutrales Element. Andererseits kann man eine einelementige Menge $V = \{0\}$ stets zu einem K-Vektorraum machen, indem man $0+0=0$ und $\alpha\cdot 0 = 0$ für $\alpha \in K$ definiert. Man nennt V dann den *Nullraum* und schreibt in suggestiver Weise $V = 0$, wobei man streng genommen zwischen 0 als Nullelement und 0 als Nullraum zu unterscheiden hat. Ist L ein Körper

und K ein Teilkörper, so kann man L stets als K-Vektorraum auffassen. Als Vektorraumaddition auf L nehme man die gegebene Körperaddition und als skalare Multiplikation $K \times L \longrightarrow L$ die Einschränkung der Körpermultiplikation $L \times L \longrightarrow L$. Insbesondere ist $\mathbb{C}$ auf diese Weise ein Vektorraum über $\mathbb{Q}, \mathbb{Q}(\sqrt{2})$ oder $\mathbb{R}$. Im Übrigen ist jeder Körper K ein Vektorraum über sich selbst.

Für das Rechnen mit Vektoren gelten die gewöhnlichen Rechenregeln, die wir im Folgenden auflisten. Dabei haben wir an dieser Stelle der Deutlichkeit halber 0_K für das Nullelement von K und 0_V für den Nullvektor in V geschrieben, eine Unterscheidung, die wir im Weiteren allerdings nicht mehr machen werden.

(1) $\alpha \cdot 0_V = 0_V$ für alle $\alpha \in K$.

(2) $0_K \cdot a = 0_V$ für alle $a \in V$.

(3) $(-\alpha) \cdot a = \alpha \cdot (-a) = -\alpha \cdot a$ für alle $\alpha \in K, a \in V$.

(4) Aus $\alpha \cdot a = 0_V$ für $\alpha \in K$ und $a \in V$ folgt bereits $\alpha = 0_K$ oder $a = 0_V$.

Die Regeln (1) - (3) beweist man genauso wie die entsprechenden Regeln für das Rechnen in Körpern. Gleiches gilt für (4), wobei wir hier die Argumentation noch einmal ausführen wollen. Gilt nämlich $\alpha \cdot a = 0$ mit $\alpha \neq 0$, so ergibt sich

$$a = (\alpha^{-1} \cdot \alpha) \cdot a = \alpha^{-1} \cdot (\alpha \cdot a) = \alpha^{-1} \cdot 0_V = 0_V.$$

Als weitere Regeln führen wir noch die allgemeinen Distributivgesetze auf; es seien $\alpha, \alpha_i, \beta_i \in K$ sowie $a, a_i \in V$ für $i = 1, \ldots, n$.

$$\alpha \cdot \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \alpha a_i$$

$$(\sum_{i=1}^{n} \alpha_i) \cdot a = \sum_{i=1}^{n} \alpha_i a$$

$$\sum_{i=1}^{n} \alpha_i a_i + \sum_{i=1}^{n} \beta_i a_i = \sum_{i=1}^{n} (\alpha_i + \beta_i) a_i$$

Definition 2. *Es sei V ein K-Vektorraum. Eine Teilmenge $U \subset V$ heißt ein* K-Untervektorraum *oder* linearer Unterraum *von V, wenn gilt*:

(i) $U \neq \emptyset$

(ii) $a, b \in U \Longrightarrow a + b \in U$

(iii) $\alpha \in K$, $a \in U \Longrightarrow \alpha a \in U$

Für einen Vektor $a \in V$ ist

$$K \cdot a := \{\alpha a \,;\, \alpha \in K\}$$

stets ein linearer Unterraum von V. In Falle $a \neq 0$ kann man hier von einer "Geraden" sprechen, für $a = 0$ ist $K \cdot a$ der Nullraum. Jeder Vektorraum enthält folglich den Nullraum und sich selbst als lineare Unterräume. Fassen wir weiter

etwa $\mathbb{C}$ als $\mathbb{Q}$-Vektorraum auf, so erkennt man $\mathbb{R}$ und $\mathbb{Q}(\sqrt{2})$ als lineare Unterräume. Im Übrigen ist die Bezeichnung K-Untervektorraum in Definition 2 gerechtfertigt, denn es gilt:

Bemerkung 3. *Eine Teilmenge U eines K-Vektorraumes V ist genau dann ein K-Untervektorraum, wenn U abgeschlossen unter der Addition und der skalaren Multiplikation mit Elementen aus K ist, und wenn U mit diesen Verknüpfungen selbst ein K-Vektorraum ist.*

Beweis. Die behauptete Äquivalenz ist in einfacher Weise zu verifizieren. Wir wollen hier nur zeigen, dass jeder lineare Unterraum $U \subset V$ die in Bemerkung 3 genannten Bedingungen erfüllt. Sei also $U \subset V$ wie in Definition 2. Zunächst besagen die Bedingungen (ii) und (iii), dass U abgeschlossen unter der Addition und der skalaren Multiplikation ist. Weiter übertragen sich allgemeine Eigenschaften der Verknüpfungen wie Assoziativität, Kommutativität, Distributivität usw. in direkter Weise von V auf U. Nach Voraussetzung gilt $U \neq \emptyset$. Es enthält U daher ein Element a. Dann gehört auch $-a = (-1)a$ zu U und damit der Nullvektor $0 = a - a$. Also ist klar, dass U eine additive Untergruppe von V und insgesamt mit den von V induzierten Verknüpfungen ein K-Vektorraum ist. □

Als wichtigstes Beispiel eines Vektorraums über einem Körper K wollen wir das n-fache kartesische Produkt

$$K^n = \{(\alpha_1, \ldots, \alpha_n)\,;\, \alpha_i \in K \text{ für } i = 1, \ldots, n\}$$

betrachten, wobei $n \in \mathbb{N}$ sei. Die Addition $K^n \times K^n \longrightarrow K^n$ werde erklärt durch

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n),$$

sowie die skalare Multiplikation $K \times K^n \longrightarrow K^n$ durch

$$\alpha \cdot (\alpha_1, \ldots, \alpha_n) = (\alpha \cdot \alpha_1, \ldots, \alpha \cdot \alpha_n).$$

Das n-Tupel $(0, \ldots, 0) \in K^n$ definiert dann den Nullvektor in K^n, den wir üblicherweise wieder mit 0 bezeichnen, und es ist $(-\alpha_1, \ldots, -\alpha_n)$ das inverse Element bezüglich der Addition zu einem Element $(\alpha_1, \ldots, \alpha_n) \in K^n$. Im Falle $n = 0$ ist K^n als einelementige Menge anzusehen, welche nur aus dem leeren Tupel besteht; K^0 ist somit der Nullraum. Weiter lässt sich K^m für $m \leq n$ in kanonischer[3] Weise als linearer Unterraum von K^n auffassen, indem man die Elemente $(\alpha_1, \ldots, \alpha_m) \in K^m$ mit denen des Typs $(\alpha_1, \ldots, \alpha_m, 0, \ldots, 0) \in K^n$ identifiziert.

Anschaulich können wir den Vektorraum K^n für $K = \mathbb{R}$ und $n = 2$ als Modell einer Ebene und für $n = 3$ als Modell des gewöhnlichen dreidimensionalen

[3] Die Bezeichnung "kanonisch" werden wir im Folgenden noch häufiger verwenden. Wir meinen hiermit eine Möglichkeit, die sich in nahe liegender Weise als die einfachste Lösung anbietet.

Raumes ansehen. Als Untervektorräume der Ebene $\mathbb{R}^2$ gibt es, wie wir noch sehen werden, außer den trivialen linearen Unterräumen 0 und $\mathbb{R}^2$ lediglich die Geraden des Typs $\mathbb{R}a$ zu von Null verschiedenen Vektoren $a \in \mathbb{R}^2$.

Die obige Konstruktion des Vektorraums K^n lässt sich allgemeiner für einen K-Vektorraum W anstelle von K durchführen. Man erhält dann das n-fache kartesische Produkt W^n von W als K-Vektorraum mit komponentenweiser Addition und skalarer Multiplikation. Darüber hinaus kann man für eine beliebige Familie von K-Vektorräumen $(V_i)_{i\in I}$ das kartesische Produkt $V = \prod_{i\in I} V_i$ als K-Vektorraum auffassen, wiederum mit komponentenweisen Verknüpfungen, indem man also für $\alpha \in K$ und $(v_i)_{i\in I}, (v')_{i\in I} \in V$ setzt:

$$(v_i)_{i\in I} + (v'_i)_{i\in I} = (v_i + v'_i)_{i\in I}, \qquad \alpha \cdot (v_i)_{i\in I} = (\alpha \cdot v_i)_{i\in I}.$$

Viele interessante Vektorräume sind als Räume von Abbildungen oder Funktionen zu sehen. Sei etwa K ein Körper und X eine Menge. Dann bildet die Menge $V = \mathrm{Abb}(X, K)$ aller Abbildungen von X nach K auf natürliche Weise einen K-Vektorraum. Man erkläre nämlich die Summe zweier Elemente $f, g \in V$ als Abbildung

$$f + g\colon X \longrightarrow K, \qquad x \longmapsto f(x) + g(x),$$

sowie das skalare Produkt eines Elementes $\alpha \in K$ mit einem Element $f \in V$ durch

$$\alpha f\colon X \longrightarrow K, \qquad x \longmapsto \alpha f(x).$$

Es ist leicht nachzurechnen, dass V mit diesen Verknüpfungen einen K-Vektorraum bildet, den so genannten *Vektorraum der K-wertigen Funktionen auf X* (der im Übrigen mit dem kartesischen Produkt K^X übereinstimmt, dessen Faktoren K durch die Elemente der Menge X parametrisiert werden). Die Nullabbildung

$$0\colon X \longrightarrow K, \qquad x \longmapsto 0$$

ist das Nullelement, und das negative Element zu einem $f \in V$ wird gegeben durch

$$-f\colon X \longrightarrow K, \qquad x \longmapsto -(f(x)).$$

Setzt man beispielsweise $K = \mathbb{R}$ und $X = \{\alpha \in \mathbb{R}\,;\, 0 \leq \alpha \leq 1\}$, so ist $V = \mathrm{Abb}(X, \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller reellwertigen Funktionen auf dem Einheitsintervall in $\mathbb{R}$. Lineare Unterräume werden gebildet von den stetigen Funktionen, den differenzierbaren Funktionen bzw. von den Polynomen.

Im Folgenden sei K stets ein Körper. Wir wollen uns etwas genauer mit dem Problem der Konstruktion von linearen Unterräumen in einem K-Vektorraum V beschäftigen.

Lemma 4. *Es sei V ein K-Vektorraum und $(U_i)_{i\in I}$ eine Familie von linearen Unterräumen. Dann ist $U = \bigcap_{i\in I} U_i$ ebenfalls ein linearer Unterraum von V.*

Beweis. Um zu sehen, dass U ein linearer Unterraum von V ist, verifizieren wir die Bedingungen von Definition 2. Aus $0 \in U_i$ für alle i folgt $0 \in U$. Seien nun

$\alpha \in K$ und $a, b \in U$. Dann ergibt sich $a, b \in U_i$ für alle i, also $a + b, \alpha a \in U_i$ und somit $a+b, \alpha a \in U$. Folglich erfüllt U die definierenden Eigenschaften eines linearen Unterraums von V. □

Satz und Definition 5. *Es sei V ein K-Vektorraum und $A \subset V$ eine Teilmenge. Dann ist*

$$\langle A\rangle := \{\sum_{i=1}^{r} \alpha_i a_i \,;\, r \in \mathbb{N}, \alpha_i \in K, a_i \in A \text{ für } i = 1, \ldots, r\}$$

ein linearer Unterraum von V, und dieser stimmt überein mit dem linearen Unterraum

$$\bigcap_{A \subset U} U \subset V,$$

den man gemäß Lemma 4 erhält, wenn man den Durchschnitt über alle linearen Unterräume U in V bildet, die A enthalten.

Folglich ist $\langle A\rangle$ der kleinste lineare Unterraum in V, der A enthält, was bedeutet, dass jeder lineare Unterraum $U \subset V$, der A enthält, auch bereits $\langle A\rangle$ enthalten muss. Man nennt $\langle A\rangle$ den von A in V erzeugten linearen Unterraum *oder auch die* lineare Hülle *von A in V.*

In ähnlicher Weise definiert man für eine Familie $\mathfrak{A} = (a_i)_{i\in I}$ von Elementen aus V den von $\mathfrak{A}$ erzeugten linearen Unterraum $\langle\mathfrak{A}\rangle \subset V$ durch $\langle\mathfrak{A}\rangle = \langle A\rangle$ mit $A = \{a_i; i \in I\}$. Aus der Definition und obigem Satz ergeben sich in direkter Weise die folgenden elementaren Eigenschaften für erzeugte lineare Unterräume in einem Vektorraum V:

(1) $\langle\emptyset\rangle = 0$

(2) $A \subset \langle A\rangle$ für eine Teilmenge $A \subset V$.

(3) $\langle U\rangle = U$ für einen linearen Unterraum $U \subset V$.

(4) $A \subset B \Longrightarrow \langle A\rangle \subset \langle B\rangle$ und $A \subset \langle B\rangle \Longrightarrow \langle A\rangle \subset \langle B\rangle$ für Teilmengen $A, B \subset V$.

Nun zum *Beweis* von Satz 5. Wir zeigen zunächst, dass $\langle A\rangle$ ein linearer Unterraum von V ist. Es gilt $\langle A\rangle \neq \emptyset$, denn der Nullvektor 0 lässt sich als leere Summe $\sum_{i=1}^{0} \alpha_i a_i$ schreiben (oder für $A \neq \emptyset$ auch als entsprechende echte Summe mit Koeffizienten $\alpha_i = 0$), gehört also zu $\langle A\rangle$. Seien weiter $\alpha \in K$ sowie

$$a = \sum_{i=1}^{r} \alpha_i a_i, \qquad b = \sum_{j=1}^{s} \beta_j b_j$$

Elemente von $\langle A\rangle$. Dann folgt

$$\alpha a = \sum_{i=1}^{r} (\alpha\alpha_i) a_i \in \langle A\rangle$$

sowie

$$a + b = \sum_{i=1}^{r} \alpha_i a_i + \sum_{j=1}^{s} \beta_j b_j = \sum_{i=1}^{r+s} \alpha_i a_i \in \langle A \rangle,$$

wenn wir $\alpha_{r+j} = \beta_j$ und $a_{r+j} = b_j$ für $j = 1, \ldots, s$ setzen. Somit ist $\langle A \rangle$ ein linearer Unterraum von V.

Ist U ein beliebiger linearer Unterraum von V, der A enthält, so muss U aufgrund der definierenden Eigenschaften eines linearen Unterraums auch alle Linearkombinationen $\sum_{i=1}^{r} \alpha_i a_i$ mit Elementen $a_1, \ldots, a_r \in A$ und Koeffizienten $\alpha_1, \ldots, \alpha_r \in K$ enthalten. Somit ergibt sich $\langle A \rangle \subset U$ und damit $\langle A \rangle \subset \bigcap_{A \subset U} U$. Andererseits schließt man aus der Gleichung $a = 1 \cdot a$ für $a \in A$ natürlich $A \subset \langle A \rangle$, so dass auch $\langle A \rangle$ zu der Menge aller linearen Unterräume $U \subset V$ gehört, die A enthalten. Insbesondere ergibt sich $\langle A \rangle = \bigcap_{A \subset \mathfrak{U}} U$, und man erkennt $\langle A \rangle$ als kleinsten linearen Unterraum von V, der A enthält. □

Definition 6. *Es sei V ein K-Vektorraum. Eine Familie $\mathfrak{A} = (a_i)_{i \in I}$ von Elementen aus V heißt ein* Erzeugendensystem *von V, wenn jedes $a \in V$ eine Darstellung $a = \sum_{i \in I} \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$ besitzt, wobei $\alpha_i = 0$ für fast alle $i \in I$ gilt, d. h. für alle $i \in I$, bis auf endlich viele Ausnahmen. Mit anderen Worten, $\mathfrak{A}$ ist ein Erzeugendensystem von V, wenn $V = \langle \mathfrak{A} \rangle$ gilt. Weiter nennt man V* endlich erzeugt, *wenn V ein endliches Erzeugendensystem $a_1, \ldots, a_n$ besitzt.*

Jeder K-Vektorraum V besitzt ein Erzeugendensystem, denn es gilt beispielsweise $\langle V \rangle = V$. Weiter gilt:

$V = \langle 1 \rangle$ für $V = \mathbb{Q}$ als $\mathbb{Q}$-Vektorraum,

$V = \langle 1, \sqrt{2} \rangle$ für $V = \mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$-Vektorraum,

$V = \langle 1, i \rangle$ für $V = \mathbb{C}$ als $\mathbb{R}$-Vektorraum,

$V = \langle e_1, \ldots, e_n \rangle$ für $V = K^n$ als K-Vektorraum.

Dabei sei $e_i \in K^n$ für $i = 1, \ldots, n$ der i-te Einheitsvektor, also

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

wobei die 1 genau an der i-ten Stelle steht. Auf präzisere Weise können wir

$$e_i = (\delta_{1i}, \ldots, \delta_{ni})$$

schreiben mit

$$\delta_{hi} = \begin{cases} 1 & \text{für } h = i \\ 0 & \text{sonst} \end{cases};$$

δ_{hi} ist das so genannte *Kronecker-Symbol.*

Aufgaben

K sei stets ein Körper.

1. Es sei V ein K-Vektorraum und $U \subset V$ ein linearer Unterraum. Für welche Elemente $a \in V$ ist $a + U := \{a + u\,;\, u \in U\}$ wiederum ein linearer Unterraum von V?

2. Es sei V ein K-Vektorraum und $\mathfrak{A} = (A_i)_{i\in I}$ eine Familie von Teilmengen von V. Die Familie $\mathfrak{A}$ möge folgende Bedingung erfüllen: Zu je zwei Indizes $i, j \in I$ existiert stets ein Index $k \in I$ mit $A_i \cup A_j \subset A_k$. Man zeige

$$\langle \bigcup_{i\in I} A_i \rangle = \bigcup_{i\in I} \langle A_i \rangle.$$

   Gilt diese Beziehung auch ohne die Voraussetzung an die Familie $\mathfrak{A}$?

3. Es sei V ein endlich erzeugter K-Vektorraum. Dann lässt sich jedes beliebige Erzeugendensystem von V zu einem endlichen Erzeugendensystem verkleinern.

4. Es sei K Teilkörper eines Körpers L und V ein L-Vektorraum. Ist dann $x_1, \ldots, x_n$ ein Erzeugendensystem von V als L-Vektorraum und $\alpha_1, \ldots, \alpha_m$ ein Erzeugendensystem von L, aufgefasst als K-Vektorraum, so bilden die Produkte $\alpha_i x_j$ mit $i = 1, \ldots, m$ und $j = 1, \ldots, n$ ein Erzeugendensystem von V als K-Vektorraum.

5. Es seien $x, y \in \mathbb{R}^2$ Punkte, die nicht gemeinsam auf einer Geraden durch den Nullpunkt $0 \in \mathbb{R}^2$ liegen, d. h. es gelte $x \neq 0 \neq y$ sowie $\alpha x \neq \beta y$ für alle $\alpha, \beta \in \mathbb{R}^*$. Man zeige, dass x, y bereits ein Erzeugendensystem von $\mathbb{R}^2$ bilden. Gilt eine entsprechende Aussage auch, wenn man $\mathbb{R}$ durch einen beliebigen Körper K ersetzt?

6. Man betrachte das kartesische Produkt $\mathbb{Q}^{\mathbb{N}} = \prod_{i\in\mathbb{N}} \mathbb{Q}$ als $\mathbb{Q}$-Vektorraum. Kann dieser Vektorraum ein abzählbares Erzeugendensystem besitzen, d. h. ein Erzeugendensystem des Typs $(x_i)_{i\in\mathbb{N}}$?

1.5 Linear unabhängige Systeme und Basen von Vektorräumen

Sind $a_1, \ldots, a_n$ Vektoren eines K-Vektorraums V, so sagt man, wie bereits in den Vorbemerkungen erwähnt, a_n *hänge linear von* $a_1, \ldots, a_{n-1}$ *ab*, wenn es Koeffizienten $\alpha_1, \ldots, \alpha_{n-1} \in K$ mit $a_n = \sum_{i=1}^{n-1} \alpha_i a_i$ gibt, wenn also $a_n \in \langle a_1, \ldots, a_{n-1} \rangle$ gilt. Man sagt in diesem Falle auch, a_n lasse sich aus den Vektoren $a_1, \ldots, a_{n-1}$ *linear kombinieren* oder a_n sei eine *Linearkombination* von $a_1, \ldots, a_{n-1}$. Wenn man für ein System von Vektoren $a_1, \ldots, a_n$ weiß, dass irgendeiner dieser Vektoren von den übrigen linear abhängt, so bezeichnet man das System gemeinhin als *linear abhängig*. (System ist hier im Sinne von Familie gemeint; das System der $a_1, \ldots, a_n$ wäre präziser als Familie $(a_i)_{i=1\ldots n}$ zu notieren.) Andererseits heißt das System der $a_1, \ldots, a_n$ *linear unabhängig*, wenn keiner dieser Vektoren von den übrigen linear abhängt. Der Begriff der linearen Abhängigkeit bzw. Unabhängigkeit von Vektoren ist in der Linearen Algebra von fundamentaler

Wichtigkeit. Für eine formelmäßige Handhabung dieses Begriffes ist folgende (äquivalente) Definition besonders geeignet, auf die wir uns im Weiteren stets stützen werden.

Definition 1. *Ein System von Vektoren $a_1, \ldots, a_n$ eines K-Vektorraums V heißt* linear unabhängig, *wenn aus einer Gleichung $\sum_{i=1}^{n} \alpha_i a_i = 0$ mit Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$ notwendig $\alpha_1 = \ldots = \alpha_n = 0$ folgt, wenn sich also der Nullvektor $0 \in V$ nur in trivialer Weise als Linearkombination der Vektoren $a_1, \ldots, a_n$ darstellen lässt. Ist diese Bedingung nicht gegeben, so bezeichnet man das System $a_1, \ldots, a_n$ als* linear abhängig.

Ein System von Vektoren $a_1, \ldots, a_n$ ist also genau dann linear abhängig, wenn es Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum_{i=1}^{n} \alpha_i a_i = 0$ gibt, wobei die α_i nicht sämtlich verschwinden. Dies ist äquivalent zu der bereits oben erwähnten Bedingung, dass einer der Vektoren $a_1, \ldots, a_n$ eine Linearkombination der restlichen ist, denn die Gleichung $\sum_{i=1}^{n} \alpha_i a_i = 0$ ist für $\alpha_{i_0} \neq 0$ äquivalent zu $a_{i_0} = -\sum_{i \neq i_0} \alpha_{i_0}^{-1} \alpha_i a_i$. Beispielsweise bildet der Nullvektor $0 \in V$ ein linear abhängiges System, aber auch jedes System von Vektoren, in dem einer der Vektoren mehrfach vorkommt, ist linear abhängig. Dagegen ist ein System, welches aus genau einem Vektor $a \neq 0$ besteht, stets linear unabhängig. Ähnlich wie bei der Konvention der leeren Summe betrachtet man Systeme von Vektoren $a_1, \ldots, a_n$ auch im Falle $n = 0$ und meint damit dann das leere System. Auch das leere System erkennt man in nahe liegender Weise als linear unabhängig.

Um die Sprache zu vereinfachen, erwähnt man in der Situation von Definition 1 meist nur die zu betrachtenden Vektoren $a_1, \ldots, a_n$, ohne besonders darauf hinzuweisen, dass das System dieser Vektoren gemeint ist. So sagt man etwa in unpräziser Ausdrucksweise, die Vektoren $a_1, \ldots, a_n$ seien linear unabhängig, womit man natürlich nicht meint, dass jeder der Vektoren a_i für sich genommen ein linear unabhängiges System bildet (was lediglich $a_i \neq 0$ bedeuten würde), sondern dass das System $(a_i)_{i=1\ldots n}$ linear unabhängig ist.

Mit 1.3/5 sehen wir beispielsweise, dass die Elemente $1, \sqrt{2} \in \mathbb{Q}(\sqrt{2})$ ein linear unabhängiges System bilden, wenn wir $\mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$-Vektorraum auffassen. Entsprechendes gilt für die Elemente $1, i$ in $\mathbb{C}$ als $\mathbb{R}$-Vektorraum. Wichtig ist auch, dass für $n \in \mathbb{N}$ die "Einheitsvektoren" $e_1, \ldots, e_n \in K^n$ ein linear unabhängiges System bilden. Denn für $\alpha_1, \ldots, \alpha_n \in K$ gilt

$$\sum_{i=1}^{r} \alpha_i e_i = (\alpha_1, \ldots, \alpha_n),$$

also verschwindet diese Summe genau dann, wenn das Element $(\alpha_1, \ldots, \alpha_n)$ verschwindet, d. h. wenn $\alpha_i = 0$ für $i = 1, \ldots, n$ gilt.

Wir haben die lineare Abhängigkeit bzw. Unabhängigkeit in Definition 1 der Einfachheit halber nur für endliche Systeme von Vektoren formuliert. Die Begriffe übertragen sich aber in nahe liegender Weise auf beliebige Systeme $(a_i)_{i \in I}$, wenn man vereinbart, dass eine Linearkombination der a_i ein Ausdruck der Form $\sum_{i \in I} \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$ ist, wobei die α_i für fast alle

$i \in I$ verschwinden, d. h. für alle $i \in I$ bis auf endlich viele Ausnahmen. Eine solche Linearkombination ist daher in Wahrheit eine *endliche* Linearkombination, stellt also ein Element in V dar. Man bezeichnet ein System $(a_i)_{i\in I}$ von Vektoren aus V als *linear unabhängig*, wenn aus dem Verschwinden einer Linearkombination der a_i, also einer Gleichung $\sum_{i\in I} \alpha_i a_i = 0$, notwendig $\alpha_i = 0$ für alle $i \in I$ folgt. Das System $(a_i)_{i\in I}$ ist daher genau dann linear unabhängig, wenn jedes endliche Teilsystem von $(a_i)_{i\in I}$ linear unabhängig im Sinne von Definition 1 ist. Entsprechend ist $(a_i)_{i\in I}$ genau dann linear abhängig, wenn es ein endliches Teilsystem gibt, welches linear abhängig im Sinne von Definition 1 ist.

Satz 2. *Es seien $a_1, \ldots, a_n$ Vektoren eines K-Vektorraums V. Dann ist äquivalent*:

(i) *Die Vektoren $a_1, \ldots, a_n$ sind linear unabhängig.*

(ii) *Ist $a = \sum_{i=1}^n \alpha_i a_i$ eine Darstellung eines Elementes $a \in \langle a_1, \ldots, a_n\rangle$ mit Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$, so sind diese eindeutig durch a bestimmt.*

Beweis. Wir nehmen zunächst Bedingung (i) als gegeben an. Sind dann

$$a = \sum_{i=1}^{n} \alpha_i a_i = \sum_{i=1}^{n} \alpha_i' a_i$$

zwei Darstellungen von a als Linearkombination der a_i, so ist $\sum_{i=1}^n (\alpha_i - \alpha_i')a_i$ eine Linearkombination, die den Nullvektor 0 darstellt. Mit (i) folgt $\alpha_i - \alpha_i' = 0$, also $\alpha_i = \alpha_i'$ für alle i, d. h. die Darstellung von a als Linearkombination der a_i ist eindeutig.

Sei nun umgekehrt Bedingung (ii) gegeben. Um die lineare Unabhängigkeit des Systems der a_i zu zeigen, betrachten wir eine Gleichung $\sum_{i=1}^n \alpha_i a_i = 0$ mit Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$. Da trivialerweise $\sum_{i=1}^n 0 \cdot a_i = 0$ gilt, ergibt sich $\alpha_i = 0$ für alle i, wenn man (ii) benutzt. □

Sind die Bedingungen des Satzes erfüllt, so nennt man das System der a_i eine *Basis* des linearen Unterraumes $\langle a_1, \ldots, a_n\rangle$ von V. Man vereinbart nämlich:

Definition 3. *Ein System von Vektoren $a_1, \ldots, a_n$ eines K-Vektorraums V wird als* (endliche) Basis *von V bezeichnet, wenn gilt*:

(i) *Die Vektoren $a_1, \ldots, a_n$ bilden ein Erzeugendensystem von V; d. h. man hat $V = \langle a_1, \ldots, a_n\rangle$.*

(ii) *Das System der Vektoren $a_1, \ldots, a_n$ ist linear unabhängig.*

Allgemeiner heißt ein (nicht notwendig endliches) System von Vektoren eines Vektorraums V eine Basis, *wenn es sich um ein Erzeugendensystem handelt, welches linear unabhängig ist.*

Mit Satz 2 ergibt sich sofort:

Bemerkung 4. *Vektoren $a_1, \ldots, a_n$ eines K-Vektorraumes V bilden genau dann eine Basis, wenn gilt: Jedes $a \in V$ besitzt eine Darstellung $a = \sum_{i=1}^{n} \alpha_i a_i$ mit eindeutig bestimmten Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$.*

Fassen wir die bisher betrachteten Beispiele von Erzeugendensystemen und linear unabhängigen Systemen zusammen, so ergibt sich:

(1) Das leere System bildet eine Basis des Nullraums über einem gegebenen Körper K, also des K-Vektorraums $V = 0$.

(2) Die Elemente $1, \sqrt{2}$ bilden eine Basis von $\mathbb{Q}(\sqrt{2})$ als $\mathbb{Q}$-Vektorraum.

(3) Die Elemente $1, i$ bilden eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum.

(4) Für einen Körper K und $n \in \mathbb{N}$ bilden die Einheitsvektoren $e_1, \ldots, e_n$ eine Basis des K-Vektorraums K^n.

Die Kenntnis von Basen in Vektorräumen ist verantwortlich dafür, dass man etwa Fragen zur linearen Unabhängigkeit von Vektoren auf das Lösen linearer Gleichungssysteme zurückführen kann. Wir wollen dies am Beispiel des K-Vektorraums K^n und der Basis $e_1, \ldots, e_n$ einmal demonstrieren. Gegeben seien Vektoren $a_1, \ldots, a_r \in K^n$, etwa

$$a_j = (\alpha_{1j}, \ldots, \alpha_{nj}) = \sum_{i=1}^{n} \alpha_{ij} e_i, \qquad j = 1, \ldots, r.$$

Die Frage, ob $a_1, \ldots, a_r$ linear abhängig sind oder nicht, ist dann äquivalent zu der Frage, ob es ein nicht-triviales r-Tupel $(\xi_1, \ldots, \xi_r) \in K^r$ gibt mit $\sum_{j=1}^{r} \xi_j a_j = 0$, d. h. ob das lineare Gleichungssystem

$$\begin{aligned} \xi_1 \alpha_{11} + \ldots + \xi_r \alpha_{1r} &= 0 \\ &\ldots \\ \xi_1 \alpha_{n1} + \ldots + \xi_r \alpha_{nr} &= 0 \end{aligned}$$

eine nicht-triviale Lösung $(\xi_1, \ldots, \xi_r) \in K^r$ besitzt. Techniken zur Lösung solcher Gleichungssysteme werden wir im Abschnitt 3.5 kennen lernen.

Als Nächstes wollen wir ein technisches Lemma beweisen, welches insbesondere für die Handhabung und Charakterisierung von Vektorraumbasen von großem Nutzen ist.

Lemma 5. *Für Vektoren $a_1, \ldots, a_n$ eines K-Vektorraums V ist äquivalent*:

(i) *$a_1, \ldots, a_n$ sind linear abhängig.*

(ii) *Einer der Vektoren $a_1, \ldots, a_n$ ist eine Linearkombination der restlichen, d. h. es existiert ein $p \in \{1, \ldots, n\}$ mit $a_p \in \langle a_1, \ldots, a_{p-1}, a_{p+1}, \ldots, a_n \rangle$.*

(iii) *Es existiert ein $p \in \{1, \ldots, n\}$ mit*

$$\langle a_1, \ldots, a_n \rangle = \langle a_1, \ldots, a_{p-1}, a_{p+1}, \ldots, a_n \rangle.$$

Sind die Vektoren $a_1, \ldots, a_r$ für ein $r < n$ linear unabhängig, so folgen aus (i) *die Bedingungen* (ii) *und* (iii) *bereits für ein $p \in \{r+1, \ldots, n\}$.*

Beweis. Wir beginnen mit der Implikation von (i) nach (ii). Seien also $a_1, \dots, a_n$ linear abhängig. Man wähle dann $r \in \{0, \dots, n\}$ maximal mit der Eigenschaft, dass das System der Vektoren $a_1, \dots, a_r$ linear unabhängig ist; im Falle $r = 0$ sei hiermit das leere System gemeint, welches stets linear unabhängig ist. Insbesondere gilt $r < n$ aufgrund der Voraussetzung in (i), und $a_1, \dots, a_{r+1}$ sind linear abhängig. Es existiert folglich eine Gleichung $\sum_{i=1}^{r+1} \alpha_i a_i = 0$ mit Koeffizienten $\alpha_i \in K$, die nicht sämtlich verschwinden. Dabei gilt notwendigerweise $\alpha_{r+1} \neq 0$, denn anderenfalls hätte man die Gleichung $\sum_{i=1}^{r} \alpha_i a_i = 0$, wobei die Koeffizienten nicht sämtlich verschwinden würden, die $a_1, \dots, a_r$ also linear abhängig wären. Die erstere Gleichung lässt sich daher nach a_{r+1} auflösen, man erhält $a_{r+1} = -\sum_{i=1}^{r} \alpha_{r+1}^{-1} \alpha_i a_i$ und damit $a_{r+1} \in \langle a_1, \dots, a_r \rangle$, wie in (ii) und der Zusatzaussage behauptet.

Sei nun Bedingung (ii) erfüllt, d. h. es gelte für ein $p \in \{1, \dots, n\}$ die Beziehung $a_p \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$. Man hat dann

$$a_1, \dots, a_n \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$$

und somit

$$(*) \qquad \langle a_1, \dots, a_n \rangle \subset \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle,$$

denn $\langle a_1, \dots, a_n \rangle$ ist der kleinste lineare Unterraum von V, der $a_1, \dots, a_n$ enthält. Da die umgekehrte Inklusion trivialerweise erfüllt ist, ergibt sich Bedingung (iii).

Der Vollständigkeit halber wollen wir hier auch noch darauf hinweisen, dass sich die Inklusion $(*)$ leicht durch direktes Nachrechnen herleiten lässt. Es gelte etwa $a_p = \sum_{i \neq p} \alpha_i a_i$ mit Koeffizienten $\alpha_i \in K$. Für jedes $b \in \langle a_1, \dots, a_n \rangle$ mit einer Darstellung $b = \sum_{i=1}^{n} \beta_i a_i$ und Koeffizienten $\beta_i \in K$ ergibt sich dann

$$b = \sum_{i \neq p} \beta_i a_i + \beta_p \sum_{i \neq p} \alpha_i a_i = \sum_{i \neq p} (\beta_i + \beta_p \alpha_i) a_i,$$

also $b \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$, und somit

$$\langle a_1, \dots, a_n \rangle \subset \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle.$$

Sei schließlich Bedingung (iii) gegeben, für ein $p \in \{1, \dots, n\}$ gelte also

$$\langle a_1, \dots, a_n \rangle = \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$$

Dann folgt insbesondere $a_p \in \langle a_1, \dots, a_{p-1}, a_{p+1}, \dots, a_n \rangle$, etwa $a_p = \sum_{i \neq p} \alpha_i a_i$ mit gewissen Koeffizienten $\alpha_i \in K$, und die Gleichung $(-1)a_p + \sum_{i \neq p} \alpha_i a_i = 0$ zeigt, dass $a_1, \dots, a_n$ linear abhängig sind. Damit ist gezeigt, dass die Bedingungen (i), (ii) und (iii) äquivalent sind.

Sind nun die Vektoren $a_1, \dots, a_r$ für ein gegebenes $r < n$ linear unabhängig, $a_1, \dots, a_n$ aber insgesamt linear abhängig, so gilt, wie wir gesehen haben, Bedingung (ii) für ein $p \in \{r+1, \dots, n\}$. Für dieses p ist dann auch Bedingung (iii) erfüllt, so dass die zusätzliche Behauptung ebenfalls bewiesen ist. □

Das gerade bewiesene Lemma lässt einige interessante Schlussfolgerungen zu.

Satz 6. *Jeder endlich erzeugte K-Vektorraum besitzt eine Basis, und jede solche Basis ist endlich.*

Beweis. Es sei $a_1, \ldots, a_n$ ein Erzeugendensystem des betrachteten K-Vektorraums V, d. h. es gelte $V = \langle a_1, \ldots, a_n \rangle$. Indem wir dieses System verkleinern, können wir $a_1, \ldots, a_n$ als minimales Erzeugendensystem voraussetzen. Die Äquivalenz der Bedingungen (i) und (iii) in Lemma 5 zeigt dann, dass die Vektoren $a_1, \ldots, a_n$ linear unabhängig sind, also eine Basis bilden.

Ist nun $(b_j)_{j \in J}$ eine weitere Basis von V, so lässt sich jeder der Vektoren $a_1, \ldots, a_n$ als Linearkombination von endlich vielen der Vektoren b_j, $j \in J$, darstellen. Es existiert deshalb eine endliche Teilmenge $J' \subset J$ mit

$$V = \langle a_1, \ldots, a_n \rangle \subset \langle b_j \,;\, j \in J' \rangle \subset V.$$

Das System $(b_j)_{j \in J'}$ bildet somit ein Erzeugendensystem von V. Dieses ist als Teilsystem von $(b_j)_{j \in J}$ sogar linear unabhängig und stellt deshalb, ebenso wie $(b_j)_{j \in J}$, eine Basis dar. Dann folgt aber notwendig $J = J'$, und man erkennt J als endlich. □

Satz 7. *Es sei V ein K-Vektorraum und $a_1, \ldots, a_n$ ein System von Vektoren aus V. Dann ist äquivalent*:

(i) *$a_1, \ldots, a_n$ bilden eine Basis von V.*
(ii) *$a_1, \ldots, a_n$ ist ein maximales linear unabhängiges System in V.*
(iii) *$a_1, \ldots, a_n$ ist ein minimales Erzeugendensystem von V.*

Beweis. Sei zunächst Bedingung (i) als gegeben angenommen, sei also $a_1, \ldots, a_n$ eine Basis von V. Für beliebiges $a \in V$ gilt dann

$$V = \langle a_1, \ldots, a_n \rangle = \langle a, a_1, \ldots, a_n \rangle,$$

und man schließt aus der Äquivalenz (i) $\Longleftrightarrow$ (iii) von Lemma 5, dass das System $a, a_1, \ldots, a_n$ linear abhängig ist. Also ist $a_1, \ldots, a_n$ ein maximales linear unabhängiges System in V.

Als Nächstes gehen wir von Bedingung (ii) aus, sei also $a_1, \ldots, a_n$ ein maximales linear unabhängiges System in V. Ist dann $a \in V$ beliebig, so ist das System $a_1, \ldots, a_n, a$ linear abhängig, und es existiert eine nicht-triviale Linearkombination mit Koeffizienten aus K

$$\alpha a + \sum_{i=1}^{n} \alpha_i a_i = 0,$$

welche die Null darstellt. Aus der linearen Unabhängigkeit der $a_1, \ldots, a_n$ ergibt sich mittels Lemma 5 (man vergleiche den Beweis der Implikation (i) $\Longrightarrow$ (ii) in Lemma 5), dass zumindest der Koeffizient α nicht verschwindet. Folglich lässt

sich vorstehende Gleichung nach a auflösen, und man erhält $a \in \langle a_1, \ldots, a_n \rangle$, d. h. $a_1, \ldots, a_n$ ist ein Erzeugendensystem von V. Weiter folgt aus der linearen Unabhängigkeit der $a_1, \ldots, a_n$, indem man die Äquivalenz (i) $\Longleftrightarrow$ (iii) aus Lemma 5 benutzt, dass $a_1, \ldots, a_n$ ein minimales Erzeugendensystem von V ist.

Nehmen wir schließlich $a_1, \ldots, a_n$ wie in Bedingung (iii) als minimales Erzeugendensystem an, so zeigt die Äquivalenz (i) $\Longleftrightarrow$ (iii) aus Lemma 5, dass $a_1, \ldots, a_n$ dann notwendig ein linear unabhängiges System ist, also eine Basis, da es bereits ein Erzeugendensystem ist. □

Satz 8 (Basisergänzungssatz). *In einem K-Vektorraum V betrachte man ein linear unabhängiges System $a_1, \ldots, a_r$ sowie ein Erzeugendensystem $b_1, \ldots, b_m$. Dann lässt sich das System der a_i durch Elemente des Systems der b_j zu einer Basis von V ergänzen, d. h. es existieren paarweise verschiedene Indizes $i(r+1), \ldots, i(n) \in \{1, \ldots, m\}$ mit der Eigenschaft, dass die Vektoren*

$$a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)}$$

eine Basis von V bilden.

Beweis. Für $n \geq r$ betrachte man paarweise verschiedene Indizes

$$i(r+1), \ldots, i(n) \in \{1, \ldots, m\},$$

so dass

$$(*) \qquad V = \langle a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)} \rangle$$

gilt. Die Gleichung ist beispielsweise für $n = r + m$ erfüllt, wenn man $i_{r+j} = j$ für $j = 1, \ldots, m$ setzt. Man betrachte nun eine Gleichung $(*)$, wobei $n \geq r$ minimal gewählt sei. Dann ist $a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)}$ ein linear unabhängiges Erzeugendensystem, stellt also eine Basis von V dar. Anderenfalls wäre dieses System nämlich linear abhängig, und man könnte es aufgrund der Äquivalenz (i) $\Longleftrightarrow$ (iii) aus Lemma 5 zu einem echt kleineren Erzeugendensystem verkürzen. Da die Vektoren $a_1, \ldots, a_r$ jedoch linear unabhängig sind, ergibt sich mit Lemma 5 in dieser Situation, dass man einen der Vektoren $b_{i(r+1)}, \ldots, b_{i(n)}$ fortlassen kann, was aber wegen der Minimalität von n ausgeschlossen ist. Das Erzeugendensystem $a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)}$ ist daher linear unabhängig und folglich eine Basis.

Wir wollen noch auf einen zweiten Beweis eingehen, der den Vorteil hat, dass er im Hinblick auf nicht-endliche Basen verallgemeinerungsfähig ist. Hierzu betrachten wir Indizes

$$i(r+1), \ldots, i(n) \in \{1, \ldots, m\},$$

nunmehr aber mit der Bedingung, dass die Vektoren

$$a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)}$$

linear unabhängig sind. Wir dürfen n als maximal gewählt annehmen. Mit Lemma 5 ergibt sich dann

$$b_1, \ldots, b_m \in \langle a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)} \rangle$$

und folglich

$$V = \langle b_1, \ldots, b_m \rangle \subset \langle a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)} \rangle,$$

so dass $a_1, \ldots, a_r, b_{i(r+1)}, \ldots, b_{i(n)}$ ein Erzeugendensystem und damit eine Basis von V bilden. □

Theorem 9. *In einem K-Vektorraum V mögen die Elemente $a_1, \ldots, a_n$ eine Basis sowie $b_1, \ldots, b_m$ ein Erzeugendensystem bilden. Dann gilt $n \leq m$. Weiter ist $b_1, \ldots, b_m$ genau dann eine Basis, wenn $n = m$ gilt. Je zwei Basen eines endlich erzeugten K-Vektorraums V bestehen folglich aus gleichviel Elementen.*

Beweis. Aufgrund des Basisergänzungssatzes 8 lässt sich das System $a_2, \ldots, a_n$ durch Elemente des Systems $b_1, \ldots, b_m$ zu einer Basis $b_{i(1)}, \ldots, b_{i(r_1)}, a_2, \ldots, a_n$ ergänzen, wobei natürlich $r_1 \geq 1$ gelten muss; vgl. Lemma 5. Lässt man bei dieser Basis das Element a_2 fort, so kann man das entstehende System wiederum durch Elemente des Systems $b_1, \ldots, b_m$ zu einer Basis von V ergänzen, etwa zu

$$b_{i(1)}, \ldots, b_{i(r_1)}, b_{i(r_1+1)}, \ldots, b_{i(r_1+r_2)}, a_3, \ldots, a_n.$$

Fährt man auf diese Weise fort, so gelangt man nach n Schritten zu einer Basis $b_{i(1)}, \ldots, b_{i(r_1+\ldots+r_n)}$, wobei die Indizes $i(1), \ldots, i(r_1 + \ldots + r_n) \in \{1, \ldots, m\}$ notwendig paarweise verschieden sind. Es folgt $r_1 + \ldots + r_n \leq m$ und wegen $r_i \geq 1$ insbesondere $n \leq m$, wie behauptet.

Ist nun $b_1, \ldots, b_m$ bereits eine Basis, so kann man die Rolle der a_i und b_j vertauschen und erhält auf diese Weise $m \leq n$, also insbesondere $m = n$. Bildet andererseits $b_1, \ldots, b_m$ mit $m = n$ ein Erzeugendensystem von V, so kann man dieses System zu einem minimalen Erzeugendensystem von V verkleinern, also zu einer Basis; vgl. Satz 7. Da wir aber schon wissen, dass Basen in V aus genau n Elementen bestehen, folgt, dass $b_1, \ldots, b_m$ notwendig eine Basis von V ist.

Da endlich erzeugte K-Vektorräume gemäß Satz 6 lediglich endliche Basen besitzen, ergibt sich insbesondere, dass je zwei Basen eines solchen Vektorraums aus gleichviel Elementen bestehen. □

Für ein System $a_1, \ldots, a_n$ von Elementen bezeichnet man die natürliche Zahl n als die *Länge* dieses Systems. Gelegentlich werden wir auch unendlichen Systemen $(a_i)_{i \in I}$, also Systemen mit unendlicher Indexmenge I, eine Länge zuordnen, nämlich die Länge ∞. Wir werden dabei nicht zwischen verschiedenen Graden der Unendlichkeit unterscheiden, etwa abzählbar unendlich (z. B. $I = \mathbb{N}$) oder überabzählbar unendlich (z. B. $I = \mathbb{R}$).

Definition 10. *Es sei V ein K-Vektorraum. Besitzt dann V eine Basis endlicher Länge n, so bezeichnet man n als die* Dimension *von V, in Zeichen*

$\dim_K V = n$. *Gibt es andererseits in* V *keine Basis endlicher Länge, also kein endliches maximales linear unabhängiges System, so sagen wir, die Dimension von* V *sei* unendlich, $\dim_K V = \infty$.

Aufgrund von Theorem 9 ist die Dimension eines Vektorraums wohldefiniert. Der Nullraum $V = 0$ hat die Dimension 0, jeder K-Vektorraum $V \neq 0$ eine Dimension > 0. Wir wollen noch einige weitere Eigenschaften der Dimension eines Vektorraums zusammenstellen, die sich auf einfache Weise aus den bisher gewonnenen Ergebnissen folgern lassen.

Korollar 11. *Es sei* V *ein* K*-Vektorraum und* $n \in \mathbb{N}$. *Dann ist äquivalent*:

(i) $\dim_K V = n$.

(ii) *Es existiert in* V *ein linear unabhängiges System von* n *Vektoren, und jeweils* $n+1$ *Vektoren sind linear abhängig.*

Beweis. Sei zunächst Bedingung (i) gegeben. Jede Basis von V bildet dann ein linear unabhängiges System bestehend ans n Vektoren. Ist andererseits $y_1, \ldots, y_{n+1}$ ein System von $n+1$ Vektoren aus V und nehmen wir an, dass dieses linear unabhängig ist, so können wir das System gemäß Satz 8 zu einer Basis von V ergänzen. Man hätte dann $\dim_K V \geq n+1$ im Widerspruch zu unserer Voraussetzung. Aus (i) ergibt sich folglich (ii).

Ist umgekehrt Bedingung (ii) gegeben, so gibt es in V ein maximales linear unabhängiges System bestehend aus n Vektoren. Dieses bildet eine Basis, und es folgt $\dim_K V = n$. □

Korollar 12. *Es sei* V *ein* K*-Vektorraum und* $n \in \mathbb{N}$. *Dann ist äquivalent*:

(i) $\dim_K V \geq n$.

(ii) *Es existiert in* V *ein linear unabhängiges System von* n *Vektoren.*

Beweis. Bedingung (i) impliziert trivialerweise Bedingung (ii), auch im Falle unendlicher Dimension, da dann keine endlichen Basen, also keine endlichen maximalen linear unabhängigen Systeme in V existieren können. Gehen wir umgekehrt von (ii) aus, so ist nur im Falle $\dim_K V < \infty$ etwas zu zeigen. Jedes linear unabhängige System von Vektoren $a_1, \ldots, a_n \in V$ lässt sich dann gemäß Satz 8 zu einer Basis von V ergänzen, und es folgt wie gewünscht $\dim_K V \geq n$. □

Korollar 13. *Für einen* K*-Vektorraum* V *ist äquivalent*:

(i) $\dim_K V = \infty$.

(ii) *Es existiert eine Folge von Vektoren* $a_1, a_2, \ldots \in V$, *so dass für jedes* $n \in \mathbb{N}$ *das System* $a_1, \ldots, a_n$ *linear unabhängig ist.*

(iii) *Es existiert eine Folge von Vektoren* $a_1, a_2, \ldots \in V$, *so dass das System* $(a_i)_{i \in \mathbb{N}}$ *linear unabhängig ist.*

(iv) *Zu jedem* $n \in \mathbb{N}$ *gibt es ein linear unabhängiges System, bestehend aus* n *Vektoren von* V.

Beweis. Wir gehen aus von Bedingung (i). Sei also $\dim_K V = \infty$. Dann gibt es in V keine endlichen Basen und somit keine endlichen maximalen linear unabhängigen Systeme. Als Konsequenz ist es möglich, eine Folge von Vektoren $a_1, a_2, \ldots \in V$ wie in (ii) gewünscht zu konstruieren. Weiter folgt aus (ii) unmittelbar Bedingung (iii), da zu jeder endlichen Teilmenge $I \subset \mathbb{N}$ ein $n \in \mathbb{N}$ existiert mit $I \subset \{1, \ldots, n\}$. Die Implikation (iii) $\Longrightarrow$ (iv) ist trivial, und (iv) $\Longrightarrow$ (i) schließlich ergibt sich mit Korollar 12. □

Korollar 14. *Es sei V ein K-Vektorraum und $U \subset V$ ein Teilraum. Dann gilt*:
(i) $\dim_K U \leq \dim_K V$
(ii) *Aus* $\dim_K U = \dim_K V < \infty$ *folgt bereits* $U = V$.

Beweis. Die erste Behauptung folgt mittels Korollar 12 aus der Tatsache, dass ein linear unabhängiges System von Vektoren aus U auch in V linear unabhängig ist. Die zweite Behauptung gilt, da man in einem endlich-dimensionalen K-Vektorraum V ein linear unabhängiges System, beispielsweise eine Basis von U, stets zu einer Basis von V ergänzen kann. □

Wir wollen nun noch einige Beispiele betrachten.

(1) Ist K ein Körper, $n \in \mathbb{N}$, so folgt $\dim_K K^n = n$.

(2) $\dim_{\mathbb{R}} \mathbb{C} = 2$

(3) $\dim_{\mathbb{Q}} \mathbb{Q}(\sqrt{2}) = 2$

(4) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$. Dies zeigt man am einfachsten mit Hilfe eines Abzählbarkeitsarguments. Jeder endlich-dimensionale $\mathbb{Q}$-Vektorraum wäre, ebenso wie $\mathbb{Q}$, abzählbar, jedoch ist $\mathbb{R}$ nicht abzählbar.

(5) Sei K ein Körper, X eine Menge und $V = \mathrm{Abb}(X, K)$ der K-Vektorraum der K-wertigen Funktionen auf X. Besteht X dann aus $n < \infty$ Elementen, so gilt $\dim_K V = n$, wohingegen man für unendliches X die Gleichung $\dim_K V = \infty$ hat. Wir wollen dies im Folgenden begründen. Für $x \in X$ bezeichne $f_x \colon X \longrightarrow K$ diejenige Funktion, die durch $f_x(x) = 1$ und $f_x(y) = 0$ für $y \neq x$ gegeben ist. Dann ist für jeweils endlich viele paarweise verschiedene Elemente $x_1, \ldots, x_n \in X$ das System $f_{x_1}, \ldots, f_{x_n}$ linear unabhängig in V, denn aus einer Gleichung $\sum_{i=1}^{n} \alpha_i f_{x_i} = 0$ mit Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$ folgt

$$0 = (\sum_{i=1}^{n} \alpha_i f_{x_i})(x_j) = \alpha_j$$

für $j = 1, \ldots, n$. Hieraus ergibt sich bereits $\dim_K V = \infty$, wenn X unendlich viele Elemente besitzt. Da wir andererseits für endliches X jedes $f \in V$ in der Form

$$f = \sum_{x \in X} f(x) f_x$$

schreiben können, ist das System $(f_x)_{x\in X}$ in diesem Falle ein Erzeugendensystem und somit eine Basis von V, so dass man $\dim_K V = n$ hat, wenn X aus $n < \infty$ Elementen besteht.

Abschließend soll noch angedeutet werden, wie die Theorie dieses Abschnitts aussieht, wenn man sich nicht auf endlich erzeugte K-Vektorräume beschränkt. Man muss dann auch unendliche Basen zulassen, wie sie in Definition 3 mit eingeschlossen sind. Man prüft leicht nach, dass die in Satz 2 und Lemma 5 gegebenen Charakterisierungen linearer Abhängigkeit bzw. Unabhängigkeit sinngemäß auch für beliebige Systeme von Vektoren gelten. Als Folgerung übertragen sich die Resultate von Bemerkung 4 und Satz 7 auf den Fall nicht notwendig endlicher Basen.

Etwas problematischer ist der Beweis des Analogons zu Satz 6, dass nämlich jeder K-Vektorraum V eine Basis oder, in äquivalenter Sprechweise, ein maximales linear unabhängiges System besitzt. Die Existenz eines solchen Systems zeigt man am einfachsten mit Hilfe des so genannten Zornschen Lemmas, welches dem Gebiet der Mengenlehre zuzuordnen ist. Das Lemma geht von einer *teilweise geordneten* Menge M aus, wobei teilweise geordnet bedeutet, dass zwischen gewissen Elementen von M eine Relation "$\leq$" besteht, und zwar mit den folgenden Eigenschaften:

$$\begin{gathered} x \leq x \text{ für alle } x \in M \\ x \leq y, y \leq z \Longrightarrow x \leq z \\ x \leq y, y \leq x \Longrightarrow x = y \end{gathered}$$

Man nennt eine Teilmenge $N \subset M$ *streng geordnet*, wenn für je zwei Elemente $x, y \in N$ stets $x \leq y$ oder $y \leq x$ gilt. Weiter heißt ein Element $z \in M$ eine *obere Schranke* von N, wenn $x \leq z$ für alle $x \in N$ gilt. Das Lemma von Zorn lautet nun wie folgt:

Ist M eine teilweise geordnete Menge und besitzt jede streng geordnete Teilmenge von M eine obere Schranke in M, so existiert in M ein maximales Element.

Dabei heißt ein Element $z \in M$ *maximal*, wenn aus $z \leq x$ mit $x \in M$ stets $x = z$ folgt. In unserer konkreten Situation definiere man M als die Menge aller Teilmengen von V, deren Elemente ein linear unabhängiges System von Vektoren in V bilden. Für zwei solche Mengen $A, B \subset V$ setze man $A \leq B$, falls $A \subset B$ gilt. Die Voraussetzungen des Lemmas von Zorn sind dann für M erfüllt, als obere Schranke einer streng geordneten Teilmenge $N \subset M$ dient beispielsweise die Vereinigung aller Teilmengen $A \in N$, also

$$\bigcup_{A\in N} A \subset V.$$

Man erhält somit aus dem Zornschen Lemma die Existenz eines maximalen Elementes in V, d. h. eines maximalen linear unabhängigen Systems von Vektoren in V und damit gemäß Satz 7 einer Basis von V.

Wie wir gesehen haben, lässt sich die Existenz maximaler linear unabhängiger Systeme problemlos mit Hilfe des Zornschen Lemmas beweisen. Ähnliches kann man für minimale Erzeugendensysteme nicht behaupten, und dies ist der Grund dafür, dass die im Beweis zu Satz 6 benutzte Idee, Basen durch Minimieren von Erzeugendensystemen zu konstruieren, im Allgemeinfall nicht zum Ziel führt. Auch der Basisergänzungssatz 8 lässt sich mit Hilfe des Zornschen Lemmas auf den Fall unendlicher Systeme verallgemeinern, wenn man die im Beweis zu Satz 8 gegebene Argumentation im Sinne maximaler linear unabhängiger Systeme mit dem Zornschen Lemma kombiniert. Man kann sogar die Aussage von Theorem 9, dass nämlich je zwei Basen $(a_i)_{i\in I}$ und $(b_j)_{j\in J}$ eines K-Vektorraums V aus "gleichvielen" Elementen bestehen, auf unendlich-dimensionale Vektorräume verallgemeinern. Dabei ist "gleichviel" in dem Sinne zu präzisieren, dass es eine bijektive Abbildung $I \longrightarrow J$ gibt. Man nennt I und J bzw. die Basen $(a_i)_{i\in I}$ und $(b_j)_{j\in J}$ dann auch *gleichmächtig*[4]. Die Mächtigkeitsklasse einer solchen Basis könnten wir als Dimension von V bezeichnen, jedoch wollen wir im Sinne von Definition 10 nicht zwischen verschiedenen unendlichen Dimensionen unterscheiden.

Schließlich sei noch angemerkt, dass man in Korollar 14 (ii) nicht auf die Bedingung $\dim_K V < \infty$ verzichten kann. Um dies einzusehen, betrachte man einen K-Vektorraum V von unendlicher Dimension und ein abzählbar unendliches linear unabhängiges System $(a_i)_{i\in\mathbb{N}}$ von Vektoren in V. Dann ist einerseits $(a_{2\cdot i})_{i\in\mathbb{N}}$ gleichmächtig zu $(a_i)_{i\in\mathbb{N}}$, andererseits aber $\langle a_0, a_2, a_4, \ldots\rangle$ ein echter linearer Unterraum von $\langle a_0, a_1, a_2, \ldots\rangle$.

Aufgaben

1. Man betrachte $\mathbb{R}^3$ als $\mathbb{R}$-Vektorraum und überprüfe folgende Systeme von Vektoren auf lineare Abhängigkeit bzw. lineare Unabhängigkeit:
   (i) $(1,0,-1),(1,2,1),(0,-3,2)$
   (ii) $(1,1,1),(1,1,0),(1,0,0)$
   (iii) $(9,1,5),(17,11,14),(9,1,5)$
   (iv) $(1,2,3),(4,5,6),(6,9,12)$
   (v) $(1,9,7),(2,3,4),(9,7,6),(6,6,6)$
   (vi) $(1,\alpha,0),(\alpha,1,0),(0,\alpha,1)$, wobei α eine reelle Zahl sei.
2. Es seien U, U' lineare Unterräume eines K-Vektorraums V mit $U\cap U' = 0$. Bilden $x_1,\ldots,x_r \in U$ und $y_1,\ldots,y_s \in U'$ linear unabhängige Systeme, so auch die Vektoren $x_1,\ldots,x_r,y_1,\ldots,y_s$ in V.
3. Für welche natürlichen Zahlen $n \in \mathbb{N}$ gibt es in $\mathbb{R}^n$ eine unendliche Folge von Vektoren $a_1, a_2, \ldots$ mit der Eigenschaft, dass je zwei Vektoren dieser Folge linear unabhängig über $\mathbb{R}$ sind, also ein linear unabhängiges System im $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ bilden?

[4] Dass je zwei Basen eines K-Vektorraums gleichmächtig sind, beweist man wie in "Bosch, Algebra", Abschnitt 7.1. Die dortige Argumentation im Sinne von Transzendenzbasen und algebraischer Unabhängigkeit überträgt sich in direkter Weise auf den Fall von Vektorraumbasen und linearer Unabhängigkeit.

4. Man betrachte den $\mathbb{R}$-Vektorraum aller Funktionen $p\colon \mathbb{R} \longrightarrow \mathbb{R}$, die durch polynomiale Ausdrücke der Form $p(x) = \sum_{i=1}^{r} \alpha_i x^i$ mit Koeffizienten $\alpha_i \in \mathbb{R}$ und variablem $r \in \mathbb{N}$ gegeben sind. Man gebe eine Basis dieses Vektorraums an. (Hinweis: Man darf benutzen, dass nicht-triviale reelle Polynome höchstens endlich viele Nullstellen haben.)

5. Es sei $x_1, \ldots, x_n$ eine Basis eines K-Vektorraums V. Für gegebene Koeffizienten $\alpha_{ij} \in K$, $1 \leq i < j \leq n$ setze man $y_j = x_j + \sum_{i<j} \alpha_{ij} x_i$, $j = 1, \ldots, n$, und zeige, dass dann auch $y_1, \ldots, y_n$ eine Basis von V bilden.

6. Es sei $\mathbb{F}$ ein endlicher Körper mit q Elementen und V ein $\mathbb{F}$-Vektorraum der Dimension n.
   (i) Man bestimme die Anzahl der Elemente von V.
   (ii) Man bestimme die Anzahl der Teilmengen in V, deren Elemente jeweils eine Basis von V bilden.

7. Es sei $X = X_1 \amalg \ldots \amalg X_n$ eine Menge mit einer endlichen Zerlegung in nichtleere paarweise disjunkte Teilmengen. Für einen Körper K betrachte man den K-Vektorraum V aller K-wertigen Funktionen $X \longrightarrow K$, sowie den linearen Unterraum U derjenigen Funktionen, die auf jedem der X_i konstant sind. Man berechne $\dim_K U$.

   Welche Dimension erhält man, wenn die X_i nicht notwendig paarweise disjunkt sind und lediglich $X = \bigcup_{i=1}^{n} X_i$ gilt?

8. Es sei K ein Körper. Für gegebene Elemente $\gamma_1, \ldots, \gamma_n \in K$ betrachte man die Teilmenge
$$U = \{(\alpha_1, \ldots, \alpha_n) \in K^n \,;\, \sum_{i=1}^{n} \alpha_i \gamma_i = 0\}.$$
   Man zeige, dass U ein linearer Unterraum von K^n ist und berechne $\dim_K U$.

1.6 Direkte Summen

Zu Vektoren $a_1, \ldots, a_r$ eines K-Vektorraums V kann man die linearen Unterräume $Ka_i = \langle a_i \rangle$, $i = 1, \ldots, r$, betrachten. Bilden die a_i ein Erzeugendensystem von V, so lässt sich jedes Element $b \in V$ in der Form $b = \sum_{i=1}^{r} b_i$ schreiben mit Vektoren $b_i \in Ka_i$; wir werden sagen, dass V die *Summe* der linearen Unterräume Ka_i ist. Bilden $a_1, \ldots, a_r$ sogar eine Basis von V, so überlegt man leicht mit 1.5/4, dass in einer solchen Darstellung die Vektoren $b_i \in Ka_i$ eindeutig durch b bestimmt sind. Wir werden sagen, dass V die *direkte Summe* der Ka_i ist. Im Folgenden sollen Summe und direkte Summe für den Fall beliebiger linearer Unterräume von V erklärt werden.

Definition 1. *Es seien $U_1, \ldots, U_r$ lineare Unterräume eines K-Vektorraums V. Dann wird die* Summe *dieser Unterräume erklärt durch*
$$\sum_{i=1}^{r} U_i = \{\sum_{i=1}^{r} b_i \,;\, b_i \in U_i \text{ für } i = 1, \ldots, r\}.$$

Es ist unmittelbar klar, dass $\sum_{i=1}^{r} U_i$ wieder ein linearer Unterraum von V ist, nämlich der von $U_1, \ldots, U_r$ erzeugte lineare Unterraum $\langle U_1 \cup \ldots \cup U_r \rangle \subset V$. Insbesondere sieht man, dass die Summe von linearen Unterräumen in V assoziativ ist.

Satz 2. *Für eine Summe $U = \sum_{i=1}^{r} U_i$ von linearen Unterräumen $U_1, \ldots, U_r$ eines K-Vektorraums V sind folgende Aussagen äquivalent:*

(i) *Jedes $b \in U$ hat eine Darstellung $b = \sum_{i=1}^{r} b_i$ mit eindeutig bestimmten Vektoren $b_i \in U_i$, $i = 1, \ldots, r$.*

(ii) *Aus einer Gleichung $\sum_{i=1}^{r} b_i = 0$ mit Vektoren $b_i \in U_i$ folgt $b_i = 0$ für $i = 1, \ldots, r$.*

(iii) *Für $p = 1, \ldots, r$ gilt $U_p \cap \sum_{i \neq p} U_i = 0$.*

Beweis. Bedingung (i) impliziert trivialerweise (ii). Um (iii) aus (ii) herzuleiten, betrachte man einen Index $p \in \{1, \ldots, r\}$ und einen Vektor $b \in U_p \cap \sum_{i \neq p} U_i$, etwa $b = \sum_{i \neq p} b_i$ mit Summanden $b_i \in U_i$. Dann gilt $-b + \sum_{i \neq p} b_i = 0$, und es folgt aus (ii) insbesondere $b = 0$. Somit hat man $U_p \cap \sum_{i \neq p} U_i = 0$.

Sei nun Bedingung (iii) gegeben, und sei $b \in V$ auf zwei Weisen als Summe von Vektoren $b_i, b_i' \in U_i$ dargestellt, also

$$b = \sum_{i=1}^{r} b_i = \sum_{i=1}^{r} b_i'.$$

Dann folgt $\sum_{i=1}^{r} b_i - b_i' = 0$ und somit

$$b_p - b_p' = -\sum_{i \neq p} b_i - b_i' \in U_p \cap \sum_{i \neq p} U_i = 0, \qquad p = 1, \ldots, n.$$

Insbesondere ergibt sich $b_i = b_i'$ für $i = 1, \ldots, r$, d. h. (i) ist erfüllt. □

Definition 3. *Es sei $U = \sum_{i=1}^{r} U_i$ eine Summe von linearen Unterräumen $U_1, \ldots, U_r$ eines K-Vektorraums V. Dann heißt U die* direkte Summe *der U_i, in Zeichen $U = \bigoplus_{i=1}^{r} U_i$, wenn die äquivalenten Bedingungen von Satz 2 erfüllt sind.*

Die Summe $U = \sum_{i=1}^{r} U_i$ von linearen Unterräumen $U_1, \ldots, U_r \subset V$ wird auch mit $U_1 + \ldots + U_r$ bezeichnet, und man schreibt $U_1 \oplus \ldots \oplus U_r$, falls diese Summe direkt ist. Eine Summe $U_1 + U_2$ zweier linearer Unterräume $U_1, U_2 \subset V$ ist genau dann direkt, wenn $U_1 \cap U_2 = 0$ gilt. Wie schon angedeutet, ist die Direktheit einer Summe von linearen Unterräumen anzusehen als Verallgemeinerung der linearen Unabhängigkeit eines Systems von Vektoren. Für Vektoren $a_1, \ldots, a_r \in V$ ist nämlich

$$\langle a_1, \ldots, a_r \rangle = \bigoplus_{i=1}^{r} K a_i, \qquad a_i \neq 0 \text{ für } i = 1, \ldots, r$$

äquivalent zur linearen Unabhängigkeit von $a_1, \ldots, a_r$.

Eine leichte Variante der gerade definierten direkten Summe stellt die so genannte *konstruierte direkte Summe* dar. Man geht hierbei von gegebenen K-Vektorräumen $V_1, \ldots, V_r$ aus und konstruiert einen K-Vektorraum V, in dem sich die V_i als lineare Unterräume auffassen lassen, und zwar derart, dass V die direkte Summe der V_i ist, also $V = \bigoplus_{i=1}^r V_i$ im Sinne von Definition 3 gilt. Dabei wird V als das kartesische Produkt der V_i definiert, $V = \prod_{i=1}^r V_i$, und man fasst dieses Produkt als K-Vektorraum mit komponentenweiser Addition und skalarer Multiplikation auf. In V bilden dann die Teilmengen

$$V_i' = \{(v_1, \ldots, v_r) \in V \,;\, v_j = 0 \text{ für } j \neq i\}, \qquad i = 1, \ldots, r,$$

lineare Unterräume, und man stellt unschwer fest, dass V die direkte Summe der V_i' ist. Da die natürliche Abbildung

$$\iota_i \colon V_i \longrightarrow V_i', \qquad v \longmapsto (v_1, \ldots, v_r) \text{ mit } v_j = \begin{cases} v & \text{für } j = i \\ 0 & \text{für } j \neq i \end{cases},$$

für $i = 1, \ldots, r$ jeweils bijektiv ist und zudem die Vektorraumstrukturen auf V_i und V_i' respektiert, können wir jeweils V_i mit V_i' unter ι_i identifizieren und auf diese Weise V als direkte Summe der linearen Unterräume $V_1, \ldots, V_r$ auffassen, in Zeichen $V = \bigoplus_{i=1}^r V_i$. Wir nennen V die konstruierte direkte Summe der V_i. Hierbei ist jedoch ein wenig Vorsicht geboten. Ist beispielsweise ein K-Vektorraum V die (nicht notwendig direkte) Summe gewisser linearer Unterräume $U_1, \ldots, U_r \subset V$, gilt also $V = \sum_{i=1}^r U_i$, so kann man zusätzlich die direkte Summe $U = \bigoplus_{i=1}^r U_i$ konstruieren. Mit Hilfe der nachfolgenden Dimensionsformeln kann man dann $\dim_K U \geq \dim_K V$ zeigen, wobei Gleichheit nur dann gilt, wenn V bereits die direkte Summe der U_i ist. Im Allgemeinen wird daher U wesentlich verschieden von V sein.

Satz 4. *Es sei V ein K-Vektorraum und $U \subset V$ ein linearer Unterraum. Dann existiert ein linearer Unterraum $U' \subset V$ mit $V = U \oplus U'$. Für jedes solche U' gilt*

$$\dim_K V = \dim_K U + \dim_K U'.$$

Man nennt U' in dieser Situation ein *Komplement* zu U. Komplemente von linearen Unterräumen sind abgesehen von den trivialen Fällen $U = 0$ und $U = V$ nicht eindeutig bestimmt.

Beweis zu Satz 4. Ist V von endlicher Dimension, so wähle man eine Basis $a_1, \ldots, a_r$ von U und ergänze diese durch Vektoren $a_{r+1}, \ldots, a_n$ gemäß 1.5/8 zu einer Basis von V. Dann gilt

$$V = \bigoplus_{i=1}^n Ka_i, \qquad U = \bigoplus_{i=1}^r Ka_i,$$

und es ergibt sich $V = U \oplus U'$ mit $U' = \bigoplus_{i=r+1}^n Ka_i$.

Ist V nicht notwendig von endlicher Dimension, so können wir im Prinzip genauso schließen, indem wir nicht-endliche Basen zulassen und die Ausführungen am Schluss von Abschnitt 1.5 beachten. Wir können dann eine Basis $(a_i)_{i\in I}$ von U wählen und diese mittels des Basisergänzungssatzes durch ein System $(a'_j)_{j\in J}$ von Vektoren zu einer Basis von V ergänzen. Es folgt $V = U \oplus U'$ mit $U' = \langle (a'_j)_{j\in J}\rangle$.

Zum Beweis der Dimensionsformel betrachte man ein Komplement U' zu U. Sind U und U' von endlicher Dimension, so wähle man eine Basis $b_1, \ldots, b_r$ von U bzw. $b_{r+1}, \ldots, b_n$ von U'. Dann bilden $b_1, \ldots, b_n$ eine Basis von V, und es folgt

$$\dim_K V = n = r + (n - r) = \dim_K U + \dim_K U',$$

wie gewünscht. Ist mindestens einer der beiden Räume U und U' von unendlicher Dimension, so gilt dies erst recht für V, und die Dimensionsformel ist trivialerweise erfüllt. □

Als Anwendung wollen wir noch eine Dimensionsformel für Untervektorräume beweisen.

Satz 5. *Es seien U, U' lineare Unterräume eines K-Vektorraumes V. Dann gilt*

$$\dim_K U + \dim_K U' = \dim_K(U + U') + \dim_K(U \cap U').$$

Beweis. Zunächst seien U und U' von endlicher Dimension. Dann ist $U + U'$ endlich erzeugt und damit nach 1.5/6 ebenfalls von endlicher Dimension. Man wähle nun gemäß Satz 4 ein Komplement W von $U \cap U'$ in U, sowie ein Komplement W' von $U \cap U'$ in U', also mit

$$U = (U \cap U') \oplus W, \qquad U' = (U \cap U') \oplus W'.$$

Dann gilt

$$U + U' = (U \cap U') + W + W',$$

und wir behaupten, dass diese Summe direkt ist. In der Tat, gilt $a + b + b' = 0$ für gewisse Elemente $a \in U \cap U'$, $b \in W$, $b' \in W'$, so ergibt sich

$$b = -(a + b') \in (U \cap U') + W' = U'$$

und wegen $b \in W \subset U$ sogar

$$b \in (U \cap U') \cap W = 0,$$

da U die direkte Summe der linearen Unterräume $U \cap U'$ und W ist. Man erhält also $b = 0$ und auf entsprechende Weise auch $b' = 0$. Damit folgt aber auch $a = 0$, was bedeutet, dass $U + U'$ die direkte Summe der linearen Unterräume $U \cap U'$, W und W' ist. Die behauptete Dimensionsformel ergibt sich dann mittels Satz 4 aus der Rechnung

$$\begin{aligned}&\dim(U+U')\\&\quad=\dim(U\cap U')+\dim W+\dim W'\\&\quad=\big[\dim(U\cap U')+\dim W\big]+\big[\dim(U\cap U')+\dim W'\big]-\dim(U\cap U')\\&\quad=\dim U+\dim U'-\dim(U\cap U').\end{aligned}$$

Abschließend bleibt noch der Fall zu betrachten, wo (mindestens) einer der linearen Unterräume U, U' nicht von endlicher Dimension ist. Gilt etwa $\dim_K U = \infty$, so folgt hieraus insbesondere $\dim_K(U+U') = \infty$, und wir haben $\dim_K U + \dim_K U'$ wie auch $\dim_K(U+U') + \dim_K(U\cap U')$ als ∞ zu interpretieren. Die behauptete Dimensionsformel ist also auch in diesem Fall gültig. □

Korollar 6. *Es seien U, U' endlich-dimensionale lineare Unterräume eines K-Vektorraums V. Dann ist äquivalent*:

(i) $U + U' = U \oplus U'$.

(ii) $\dim_K(U+U') = \dim_K U + \dim_K U'$.

Beweis. Bedingung (ii) ist aufgrund von Satz 5 äquivalent zu $\dim_K(U\cap U') = 0$, also zu $U \cap U' = 0$ und damit zu Bedingung (i). □

Aufgaben

1. Man bestimme Komplemente zu folgenden linearen Unterräumen des $\mathbb{R}^3$ bzw. $\mathbb{R}^4$:
   (i) $U = \langle (1,2,3), (-2,3,1), (4,1,5)\rangle$
   (ii) $U = \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 \,;\, 3x_1 - 2x_2 + x_3 + 2x_4 = 0\}$
2. Man betrachte eine Zerlegung $V = \sum_{i=1}^n U_i$ eines endlich-dimensionalen K-Vektorraums V in Untervektorräume $U_i \subset V$ und zeige, dass die Summe der U_i genau dann direkt ist, wenn $\dim_K V = \sum_{i=1}^n \dim_K U_i$ gilt.
3. Man betrachte eine direkte Summenzerlegung $V = \bigoplus_{i=1}^n U_i$ eines K-Vektorraums V in lineare Unterräume $U_i \subset V$, sowie für jedes $i = 1,\dots,n$ eine Familie $(x_{ij})_{j\in J_i}$ von Elementen $x_{ij} \in U_i$. Man zeige:
   (i) Die Elemente x_{ij}, $i = 1,\dots,n$, $j \in J_i$, bilden genau dann ein Erzeugendensystem von V, wenn für jedes $i = 1,\dots,n$ die Elemente x_{ij}, $j \in J_i$, ein Erzeugendensystem von U_i bilden.
   (ii) Die Elemente x_{ij}, $i = 1,\dots,n$, $j \in J_i$, sind genau dann linear unabhängig in V, wenn für jedes $i = 1,\dots,n$ die Elemente x_{ij}, $j \in J_i$, linear unabhängig in U_i sind.
   (iii) Die Elemente x_{ij}, $i = 1,\dots,n$, $j \in J_i$, bilden genau dann eine Basis von V, wenn für jedes $i = 1,\dots,n$ die Elemente x_{ij}, $j \in J_i$, eine Basis von U_i bilden.
4. Es seien U_1, U_2, U_3 lineare Unterräume eines K-Vektorraums V. Man zeige:

$$\begin{aligned}&\dim_K U_1 + \dim_K U_2 + \dim_K U_3\\&\qquad=\dim_K(U_1+U_2+U_3)+\dim_K((U_1+U_2)\cap U_3)+\dim_K(U_1\cap U_2)\end{aligned}$$

5. Es sei $U_1 \subset U_2 \subset \ldots \subset U_n$ eine Folge linearer Unterräume eines K-Vektorraums V mit $\dim_K V < \infty$. Man zeige, dass es jeweils ein Komplement U_i' zu U_i gibt, $i = 1, \ldots, n$, mit $U_1' \supset U_2' \supset \ldots \supset U_n'$.

6. Es sei U ein linearer Unterraum eines endlich-dimensionalen K-Vektorraums V. Unter welcher Dimensionsbedingung gibt es lineare Unterräume U_1, U_2 in V mit $U \subsetneq U_i \subsetneq V$, $i = 1, 2$, und $U = U_1 \cap U_2$?

7. Es seien U, U' zwei lineare Unterräume eines endlich-dimensionalen K-Vektorraums V. Unter welcher Dimensionsbedingung besitzen U und U' ein gemeinsames Komplement in V?

8. Es sei U ein linearer Unterraum eines endlich-dimensionalen K-Vektorraums V. Unter welcher Dimensionsbedingung kann man Komplemente $U_1, \ldots, U_r$ zu U in V finden, derart dass $\sum_{i=1}^r U_i = \bigoplus_{i=1}^r U_i$ gilt?