Ralph-Hardo Schulz

Lineare Algebra / Analytische Geometrie I

Skriptum zur Vorlesung in der Lehrkräfteweiterbildung

Berlin 2019/2022
LaTeX–Erstellung unter Mitarbeit von Tscho Heringlehner

Inhaltsverzeichnis

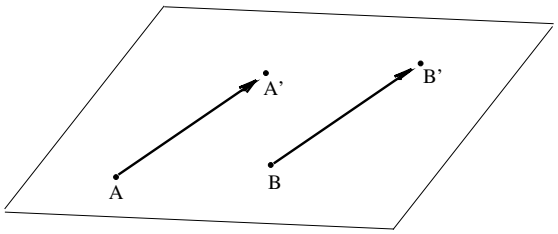
Kapitel I: Einleitung

1 Modelle von Vektorräumen

In diesem (der Motivation dienenden) Paragraphen behandeln wir Beispiele zur algebraischen Struktur 'Vektorraum'. Einige Begriffe und Gegebenheiten setzen wir dabei als bekannt voraus. Eine Präzisierung erfolgt später, wenn wir uns um eine mathematisch strenge Darstellung bemühen.

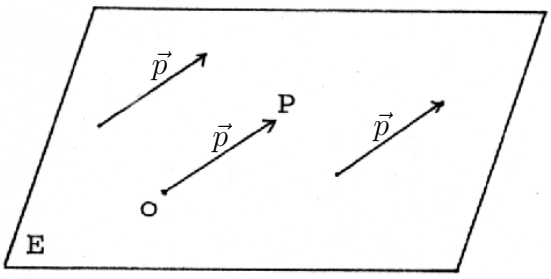
1.1 Vektoren der (Zeichen-) Ebene

(a) Was verstehen wir unter einem Vektor der Zeichenebene?
Wir betrachten eine 'Parallelverschiebung' (Translation) der 'Zeichenebene' E. Verbinden wir Urbild- und Bildpunkte (z.B. A und A', B und B', etc.) durch Pfeile,



Figur 1.1: Vektorgleiche Pfeile der Zeichenebene

so haben diese Pfeile $\overrightarrow{AA'}$, $\overrightarrow{BB'}$, ... die gleiche 'Länge' und (im Falle $A \neq A'$) die gleiche 'Richtung' (d.h. sie, sind parallel) und die gleiche 'Orientierung'; wir sagen: sie sind **vektorgleich**.
Die Menge $\vec{a}$ (zum Begriff 'Menge's.u. §2) aller zu $\overrightarrow{AA'}$ vektorgleichen Pfeile in E heißt **(freier) Vektor** der Zeichenebene. Ein Vektor in diesem Sinne ist also eine Menge von Pfeilen. Jeder der Pfeile von $\vec{a}$ (also u.a. $\overrightarrow{AA'}$) wird **Repräsentant** von $\vec{a}$ genannt; die Länge der Strecke $\overrightarrow{AA'}$ heißt **Länge** $|\vec{a}|$ von $\vec{a}$. Mit jedem Vektor $\vec{a}$ ist eine Parallelverschiebung $v_{\vec{a}}$ verbunden und umgekehrt.
Sei nun ein fester Punkt O (**Ursprung**) in E ausgewählt. Jetzt bestimmt jeder Punkt P einen Pfeil $\overrightarrow{OP}$; der dadurch repräsentierte Vektor $\vec{p}$ heißt **Ortsvektor** des Punktes P.



Figur 1.2: Ortsvektor und freier Vektor

Umgekehrt ist jeder Vektor $\vec{p}$ von E Ortsvektor eines Punktes (nämlich des Endpunktes des Repräsentanten von $\vec{p}$ mit Anfangspunkt O).

Bei festem O haben wir eine eindeutige Zuordnung zwischen

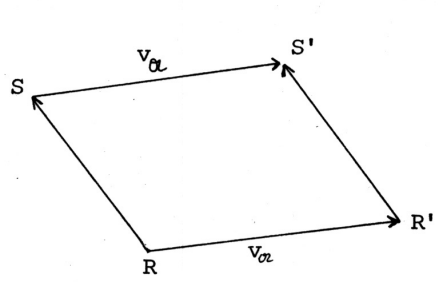
(i) dem Punkt P,

(ii) der Parallelverschiebung $v_{\vec{p}}$, die O in P überführt,

(iii') dem Pfeil $\vec{OP}$,

(iii) dem durch $\overrightarrow{OP}$ bestimmten Vektor $\vec{p}$.

Anmerkung. Eine Parallelverschiebung führt einen beliebigen Pfeil (z.B. $\overrightarrow{RS}$) in einen vektorgleichen Bildpfeil ($\overrightarrow{R'S'}$)**Vektoren bleiben also invariant unter Parallelverschiebungen**. (Umgekehrt gibt es zu zwei vektorgleichen Pfeilen eine Parallelverschiebung, die den einen Pfeil in den anderen überführt.) (Beweis mit den Eigenschaften der Parallelverschiebungen.)



Figur 1.3: Pfeile unter Parallelverschiebung [1]

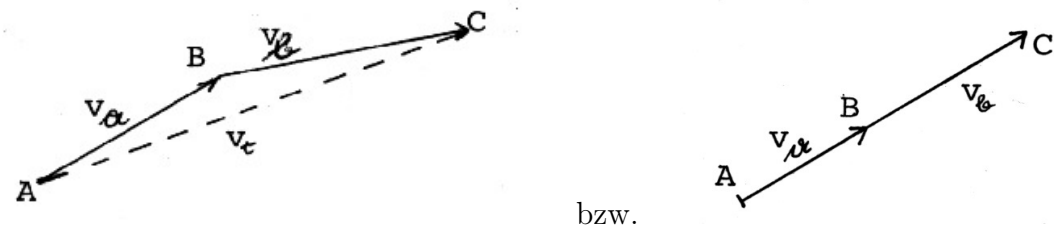
(b) Addition

Zunächst zeigen wir folgende

Bemerkung. Die Hintereinanderausführung zweier Parallelverschiebungen ergibt wieder eine Parallelverschiebung.

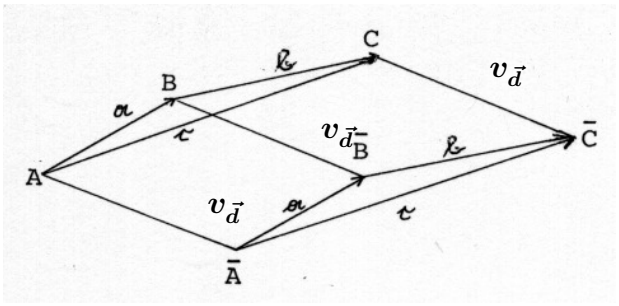
Um dies einzusehen, tragen wir in einem Punkt A den zur ersten Parallelverschiebung gehörenden Vektor $\vec{a}$ ab; den Endpunkt des Repräsentanten nennen wir B; in B tragen wir den zur zweiten Parallelverschiebung gehörenden Vektor $\vec{b}$ ab. Der Endpunkt dieses Pfeils sei C. (Wir definieren $\overrightarrow{AC} =: \overrightarrow{AB}+\overrightarrow{BC}$; der Endpunkt des 1. Pfeils muss gleich dem Anfangspunkt des 2. Pfeils sein.) Nun ist C das Bild von A bei Hintereinanderausführung beider Parallelverschiebungen. Der Pfeil $\overrightarrow{AC}$ definiert eine Parallelverschiebung $v_{\vec{c}}$.

[1] In den Figuren sind für Vektoren oft noch Frakturbuchstaben verwandt, also z.B. $\mathfrak{a}$ statt $\vec{a}$.



Figur 1.4: Hintereinanderausführung von Parallelverschiebungen

Wir müssen zeigen, dass $v_{\vec{c}}$ (bzw. der zugehörige Vektor $\vec{c}$) nicht vom Punkt A abhängt. Ist $\bar{A}$ nun irgend ein weiterer Punkt von E, so betrachten wir die zum Pfeil $\overrightarrow{A\bar{A}}$ des Vektors $\vec{d}$ gehörende Parallelverschiebung $v_{\vec{d}}$; sie führe die Punkte A, B, C in $\bar{A}$, $\bar{B}$, $\bar{C}$ über, s. Figur 1.5. Da Pfeile auf vektorgleiche Pfeile abgebildet werden, ist $\bar{C}$ der Bildpunkt von $\bar{A}$ bei Hintereinanderausführung von $v_{\vec{a}}$ und $v_{\vec{b}}$.

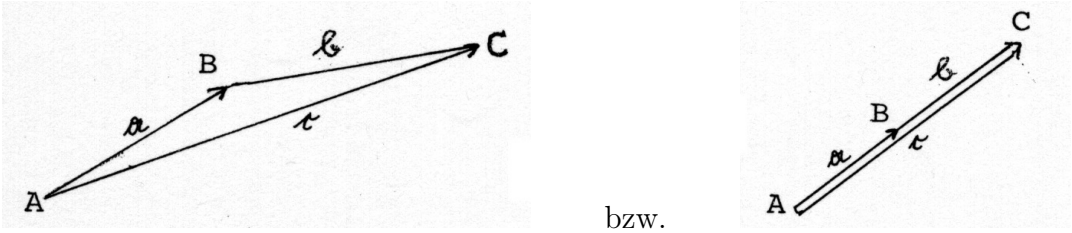


Figur 1.5: Unabhängigkeit der Addition von den Repräsentanten

Andererseits sind aber auch $\overrightarrow{AC}$ und $\overrightarrow{\bar{A}\bar{C}}$ vektorgleich. Jeder Bildpunkt entsteht also durch Abtragen des zu $\vec{c}$ gehörenden Repräsentanten, damit einheitlich durch die Parallelverschiebung, $v_{\vec{c}}$.
Aufgrund dieser Überlegungen und in Übereinstimmung mit den Erfordernissen einer Addition von Vektoren in der Physik (z.B. „Kräfteparallelogramm“ s.u.) definieren wir die Summe zweier Vektoren $\vec{a}$ und $\vec{b}$

$$\boxed{\vec{c} = \vec{a} + \vec{b} \qquad \text{mit Hilfe der Repräsentanten} \qquad \overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}}$$

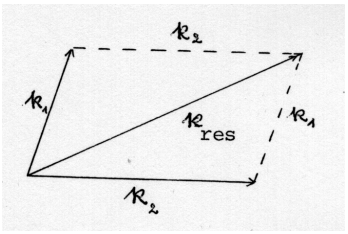
Hierbei ist A beliebiger Punkt von E und $\overrightarrow{AB}$, $\overrightarrow{BC}$, $\overrightarrow{AC}$ seien Repräsentanten von $\vec{a}$, $\vec{b}$ bzw. $\vec{c}$. Es ist wichtig, dass der Endpunkt (Spitze) des Repräsentanten von $\vec{a}$ mit dem Anfangspunkt (Fuß) des Repräsentanten von $\vec{b}$ übereinstimmt, sodass ein „Pfeilzug“ entsteht (**Spitze–Fuß–Regel**).



Figur 1.6: Zur Addition von Vektoren

Diese Definition der Summe zweier Vektoren ist sinnvoll, da $\vec{c}$ dabei nicht von A abhängt. Gehen wir von den Vektoren zu den zugehörigen Parallelverschiebungen über! Wir haben gesehen, dass die Hintereinanderausrührung von $v_{\vec{a}}$ und $v_{\vec{b}}$ wieder eine Parallelverschiebung ist; deren eindeutig zugeordneter Vektor $\vec{c}$ hat u.a. den Repräsentanten $\overrightarrow{AC}$, ist aber nicht von A abhängig (s.o.), sondern nur von $\vec{a}$ und $\vec{b}$.

Anmerkung. Zwei in einem Punkt angreifende Kräfte addieren sich vektoriell (Kräfteparallelogramm):

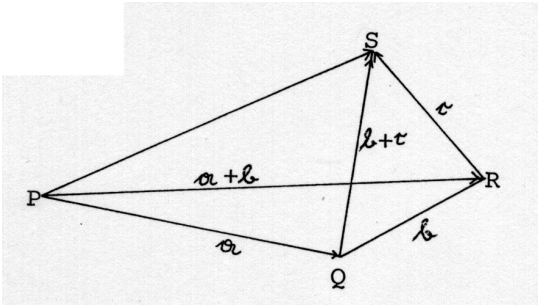


Figur 1.7: Kräfteparallelogramm: $\vec{k_{\text{res}}} = \vec{k_1} + \vec{k_2}$

Rechengesetze:

$$\boxed{(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}) \qquad \text{(Assoziativgesetz)}}$$

Beweisskizze: Siehe Figur 1.8 !

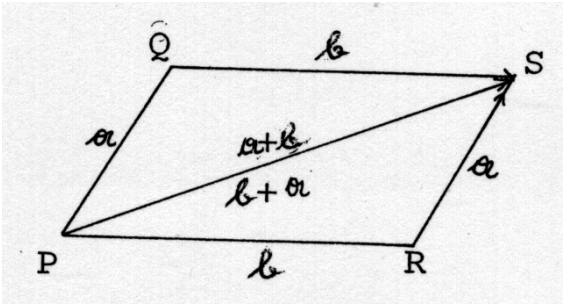


Figur 1.8: Zum Assoziativgesetz □

Gegenbeispiel aus der Umgangssprache: Fach(Werkstadt) ist nicht gleich (Fachwerk)Stadt.

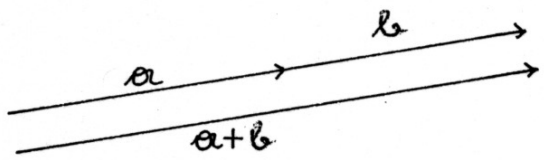
$$\boxed{\vec{a} + \vec{b} = \vec{b} + \vec{a} \qquad \text{(Kommutativgesetz)}}$$

Beweisskizze:
1. Fall: $\vec{a}$ und $\vec{b}$ haben verschiedene Richtung



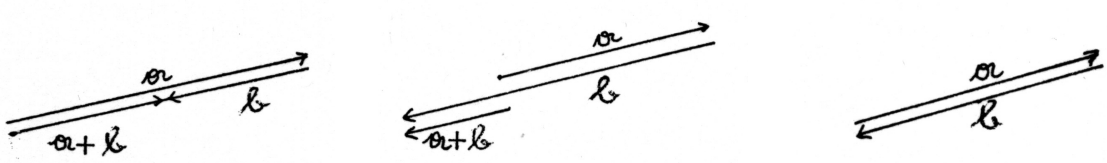
Figur 1.9 : Zum Kommutativgesetz

($\overrightarrow{RS}$ ist Repräsentant von $\vec{a}$, da $\overrightarrow{PQ}$ durch $v_{\vec{b}}$ in einen vektorgleichen Pfeil übergeht)
2. Fall: $\vec{a}$ und $\vec{b}$ haben gleiche Richtung. Diese Betrachtung überlassen wir dem Leser; als Hinweis diskutieren wir jedoch die verschieden Möglichkeiten für $\vec{a} + \vec{b}$.
Haben $\vec{a}$ und $\vec{b}$ die gleich Richtung ($\vec{a} \parallel \vec{b}$), hat auch $\vec{a} + \vec{b}$ diese Richtung. p Haben $\vec{a}$ und $\vec{b}$ zusätzlich die gleiche Orientierung, so auch $\vec{a} + \vec{b}$; der Summenvektor hat dann die Länge $|\vec{a}| + |\vec{b}|$.



Figur 1.10 a: Zur Orientierung des Summenvektors

Sind $\vec{a}$ und $\vec{b}$ gegensätzlich orientiert, so hat im Fall $|\vec{a}| > |\vec{b}|$ (bzw. $|\vec{b}| > |\vec{a}|$) die Summe $\vec{a} + \vec{b}$ und die Orientierung von $\vec{a}$ (bzw. $\vec{b}$) und die Länge $|\vec{a}| - |\vec{b}|$ (bzw. $|\vec{b}| - |\vec{a}|$).



für $|\vec{a}| > |\vec{b}|$ für $|\vec{b}| > |\vec{a}|$ für $|\vec{a}| = |\vec{b}|$

Figur 1.10 b,c,d: Zum Summenvektor – Beispiele für die weiteren Fälle

Es verbleibt der Fall, dass $\vec{a}$ und $\vec{b}$ parallel und gleich lang sind, aber gegensätzlich orientiert. (Anwendungsbeispiel: gleich große, aber entgegengesetzt gerichtete Kräfte). Es ist üblich, dann $-\vec{a}$ für $\vec{b}$ zu schreiben.

Ist $\overrightarrow{AB}$ Repräsentant von $\vec{a}$, so repräsentiert $\overrightarrow{BA}$ den Vektor $-\vec{a}$.
Die Summe $\vec{a} + (-\vec{a})$ wird repräsentiert durch die „Pfeile“ der Form $\overrightarrow{AA}$ ($= \overrightarrow{AB} + \overrightarrow{BA}$),

die wir ebenfalls vektorgleich nennen. Der Vektor, den sie bestimmen, heißt **Nullvektor** $\vec{o}$; er hat Länge 0 und undefinierte Richtung.

Eine (hier nicht beantwortete) Frage : Wie sieht die „Parallelverschiebung“ $v_{\vec{o}}$ aus, wie $v_{(-\vec{a})}$?

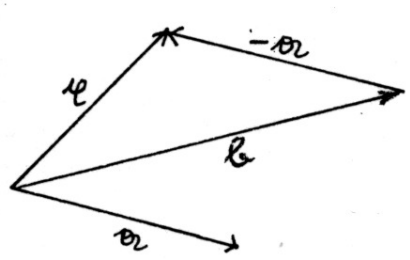
Wir halten fest: $\boxed{\vec{a}+(-\vec{a})=\vec{o}}$

Haben $-\vec{a}$ und $\vec{o}$ die Eigenschaften, die man aufgrund der Ähnlichkeit zur Bezeichnungsweise bei Zahlen erwartet? Tatsächlich gilt:

$$\boxed{\vec{c}+\vec{o}=\vec{c}} \qquad \text{für jeden Vektor } \vec{c} \text{ in } E\,;$$

$$\text{und } \boxed{\text{aus } \vec{x}+\vec{a}=\vec{b} \text{ folgt } \vec{x}=\vec{b}-\vec{a}};$$

denn aus $\vec{x}+\vec{a}=\vec{b}$ folgt durch Addition von $-\vec{a}$ die Gleichung $(\vec{x}+\vec{a})+(-\vec{a})=\vec{b}+(-\vec{a})$; mit der *Abkürzung* $\vec{b}-\vec{a}=\vec{b}+(-\vec{a})$ und wegen $(\vec{x}+\vec{a})+(-\vec{a})=\vec{x}+(\vec{a}+(-\vec{a}))=\vec{x}+\vec{o}=\vec{x}$ sieht man, dass höchstens diese Lösung möglich ist. Umgekehrt ist $\vec{x}=\vec{b}-\vec{a}$ tatsächlich Lösung, was man durch Einsetzen verifiziert. So lassen sich also solche Gleichungen lösen.(Vgl. Figur 1.11 !)



Figur 1.11: Zur Lösung der Gleichung $\vec{x}+\vec{a}=\vec{b}$

(c) S-Multiplikation (Multiplikation mit Skalaren)

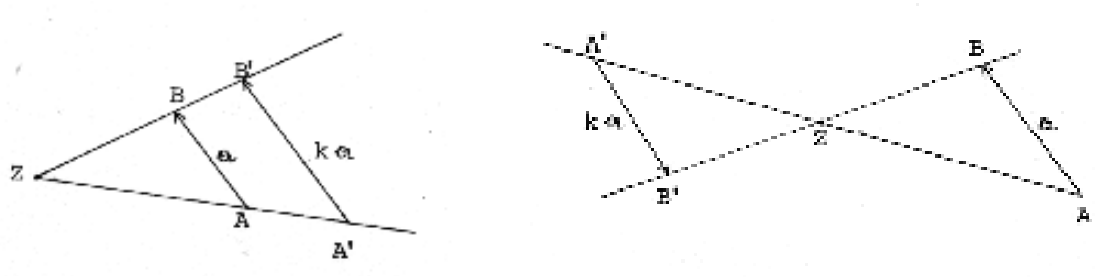
Eine zentrische Streckung mit dem (reellen, von 0 verschiedenen) Streckfaktor k und dem Zentrum Z bildet jeden Pfeil $\overrightarrow{AB}$ auf einen zu $\overrightarrow{AB}$ parallelen Pfeil $\overrightarrow{A'B'}$ der $|k|$-fachen Länge und der im Falle

$$\left.\begin{matrix} k>0 & \text{gleichen} \\ k<0 & \text{entgegengesetzten} \end{matrix}\right\}$$

Orientierung ab.

(Zum Zusammenhang mit den **Strahlensätzen** s. Figur 1.12 !).

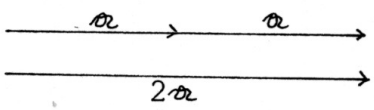
Ein Vektor wird also bei zentrischen Streckungen wieder auf einen (meist vom ursprünglichen verschiedenen) Vektor abgebildet.



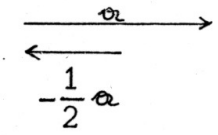
$k > 0$ $k < 0$

Figur 1.12: Pfeile unter zentrischen Streckungen mit Streckfaktor k : Es gilt ($\overrightarrow{ZA'} = k\overrightarrow{ZA}$).

Mit $\boxed{k\vec{a}}$ bezeichnen wir den Vektor, der gleiche Richtung wie $\vec{a}$ hat, $|k|$-fache Länge und (für $k > 0$) gleiche bzw. (für $k < 0$) entgegengesetzte Orientierung; insbesondere ist $|k\vec{a}| = |k| \cdot |\vec{a}|$.
Beispiele: $1\vec{a} = \vec{a}$; $-\vec{a} = (-1)\vec{a}$



Figur 1.13 a: Beispiel zur S-Multiplikation:
$2\vec{a} = \vec{a} + \vec{a}$

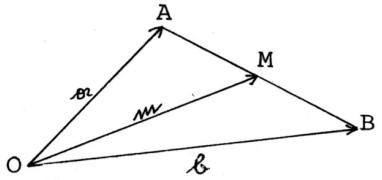


Figur 1.13 b: Beispiel zur S-Multiplikation:
$(-\frac{1}{2})\vec{a} = -(\frac{1}{2}\vec{a})$

Für $k = 0$ definieren wir: $0\vec{a} = \vec{o}$ für alle Vektoren $\vec{a}$ von E.

(Geometrische) Anwendung: **Bestimmung des Mittelpunktes einer Strecke**.

Seien $\vec{a}$ und $\vec{b}$ die Ortsvektoren (bzgl. des Ursprungs O) der Endpunkte A und B der gegebenen Strecke! Wie lässt sich dann der Ortsvektor $\vec{m}$ des Mittelpunktes M von AB angeben?



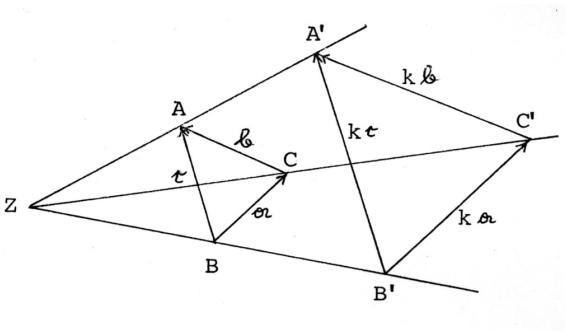
Figur 1.14 Mittelpunkt einer Strecke

$\overrightarrow{AB}$ ist Repräsentant von $\vec{b} - \vec{a}$ (nach der Spitze-Fuß-Regel)
$\overrightarrow{AM}$ ist Repräsentant von $\frac{1}{2}(\vec{b}-\vec{a})$ (nach Definition von Mittelpunkt und S-Multiplikation)
$\overrightarrow{OM}$ also Repäsentant von $\vec{a} + \frac{1}{2}(\vec{b} - \vec{a}) = \vec{a} + \frac{1}{2}\vec{b} - \frac{1}{2}\vec{a} = \frac{1}{2}(\vec{a} + \vec{b})$. Also $\boxed{\vec{m} = \frac{1}{2}(\vec{a} + \vec{b})}$.
Dabei haben wir mit Vektoren gerechnet, wie wir es von den reellen Zahlen gewohnt sind. Gerechtfertigt wird unser Vorgehen hier durch den Nachweis der Gültigkeit folgender **Rechenregeln:**

$\boxed{k(\vec{a} + \vec{b}) = k\vec{a} + k\vec{b}}$ (Distributivgesetz 1. Art)

Beweis:

Sind $\overrightarrow{BC}$ und $\overrightarrow{CA}$ Repräsentanten von $\vec{a}$ bzw. $\vec{b}$, und ist $\overrightarrow{BA}$ Pfeil von $\vec{c} = \vec{a} + \vec{b}$, so gilt für deren Bilder unter einer zentrischen Streckung mit Streckfaktor k: $\overrightarrow{B'C'}$, $\overrightarrow{C'A'}$, $\overrightarrow{B'A'}$ sind Repräsentanten von $k\vec{a}$, $k\vec{b}$ bzw. $k\vec{c}$, und es ist $k(\vec{a}+\vec{b}) = k\vec{c} = k\vec{a} + k\vec{b}$. □



Figur 1.15 Zur S-Multiplikation

$$\boxed{k_1(k_2\vec{a}) = (k_1 \cdot k_2)\vec{a}} \qquad \text{(gemischtes Assoziativgesetz)}$$

(Unmittelbar aus der Definition einzusehen).

$$\boxed{(k_1 + k_2)\vec{a} = k_1\vec{a} + k_2\vec{a}} \qquad \text{(Distributivgesetz 2. Art)}$$

↑ Addition reeller Zahlen ↑ Addition von Vektoren

Beweis – Andeutung.
$(k_1 + k_2)\vec{a}$ und $k_1\vec{a} + k_2\vec{a}$ haben beide die gleiche Orientierung (warum?) und die Länge $|k_1 + k_2| \cdot |\vec{a}|$
(Fallunterscheidungen nach den Vorzeichen von k_1 und k_2). □

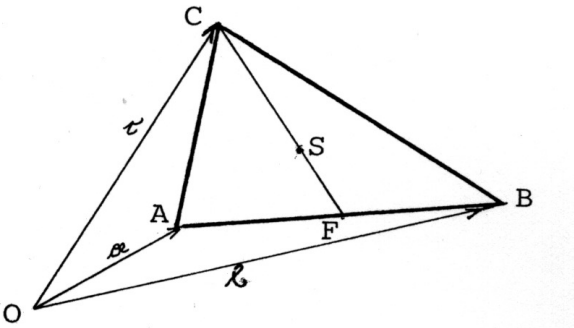
Weitere geometrische Anwendungen:

(i) Schnittpunkt der Seitenhalbierenden eines Dreiecks.

Wir wissen, dass die Seitenhalbierenden eines Dreiecks sich im Verhältnis 2 : 1 teilen. Wir wollen diesen Sachverhalt vektoriell herleiten. Das Dreieck sei gegeben durch die Punkte A, B, C mit den Ortsvektoren $\vec{a}$, $\vec{b}$ bzw. $\vec{c}$. Es sei F der Mittelpunkt der Strecke $\overline{AB}$ mit Endpunkten A und B. Wir suchen den Ortsvektor $\vec{s}$ des Punktes S, der die Strecke $\overline{CF}$ im Verhältnis 2 : 1 teilt.

$\vec{f} = \frac{1}{2}(\vec{a} + \vec{b})$ (s.o.)
$\vec{s} = \vec{f} + \frac{1}{3}(\vec{c} - \vec{f}) = \vec{f} + \frac{1}{3}\vec{c} - \frac{1}{3}\vec{f} = \frac{2}{3}\vec{f} + \frac{1}{3}\vec{c} =$
$\frac{1}{3}(\vec{a} + \vec{b} + \vec{c})$.

Da das Ergebnis symmetrisch in $\vec{a}$, $\vec{b}$, $\vec{c}$ ist, erhalten wir $\vec{s}$ auch als Ortsvektor der Punkte, die die anderen Seitenhalbierenden im Verhältngs 2 : 1 teilen; diese Punkte fallen also zusammen; S ist der Schnittpunkt aller drei Seitenhalbierenden.



Figur 1.16: Zum Seitenhalbierenden-Schnittpunkt

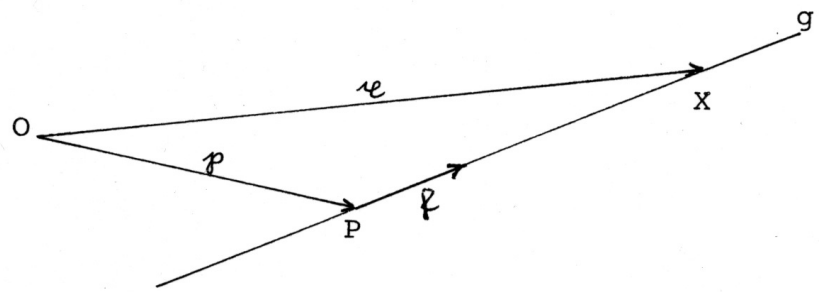
(ii) Geradengleichungen

Ist von einer Geraden g ein Punkt P mit Ortsvektor $\vec{p}$ und die Richtung durch einen Vektor $\vec{f}$ in Richtung g bekannt, dann hat jeder Punkt X von g einen Ortsvektor $\vec{x}$ der

Form

$$\boxed{\vec{x} = \vec{p} + k\vec{f}} \qquad \text{(vektorielle Punktrichtungsgleichung).}$$

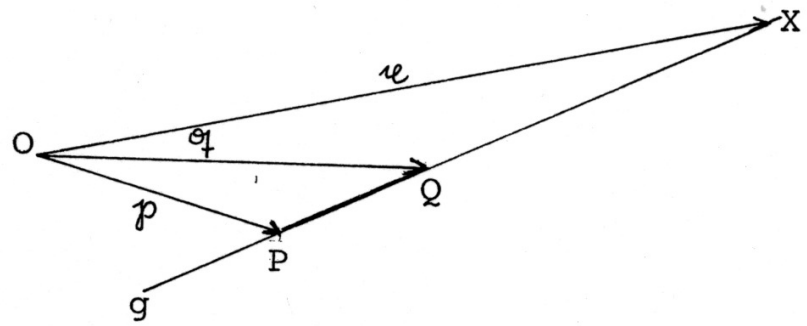
Umgekehrt liegt jeder solche Punkt auf g.



Figur 1.17: Zur Punkt-Richtungs-Gleichung einer Geraden

(Beim Beweis wird benutzt, dass parallele Vektoren skalare Vielfache von einander sind). Sind von g zwei Punkte P, Q mit Ortsvektoren $\vec{p}$ und $\vec{q}$ gegeben, so folgt mit $\vec{f} = \vec{q} - \vec{p}$

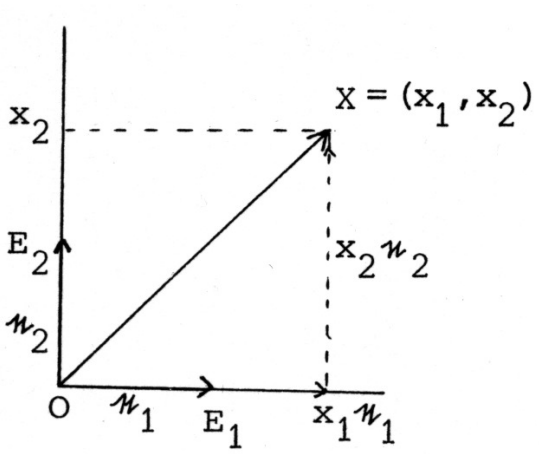
$$\boxed{\vec{x} = \vec{p} + k(\vec{q} - \vec{p})} \qquad \text{(vektorielle Zweipunktegleichung).}$$



Figur 1.18: Zur Zweipunktegleichung einer Geraden

(d) Komponentendarstellung

Von der Schule her sind wir gewohnt, Punkte durch Koordinatentupel darzustellen. Bis jetzt haben wir uns um eine koordinatenfreie Darstellung bemüht. Nun wollen wir die Verbindung zwischen dem (Orts-) Vektor und den Koordinaten eines Punktes herstellen. Dazu wählen wir in E ein **'kartesisches Koordinatensystem'** aus, also insbesondere Punkte O, E_1, E_2 derart, dass gilt: $|\overline{OE_1}| = 1 = |\overline{OE_2}|$ (Strecken der Länge 1) und $OE_1 \perp OE_2$ (Geraden, die aufeinander senkrecht stehen).



Figur 1.19: Kartesische Koordinaten

Nun seien $\vec{e}_1$ und $\vec{e}_2$ Ortsvektoren von E_1 bzw. E_2.
Der Punkt X mit den kartesischen Koordinaten (x_1, x_2) hat dann den Ortsvektor $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$ (**Komponentendarstellung**), wofür wir auch $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2}$ schreiben. (Siehe Figur 1.19 !)

Wir haben also folgende Entsprechung: $\boxed{X = (x_1, x_2) \longleftrightarrow \vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2}}$.

Addition in Komponentendarstellung:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} = (x_1\vec{e}_1 + x_2\vec{e}_2) + (y_1\vec{e}_1 + y_2\vec{e}_2) = (x_1 + y_1)\vec{e}_1 + (x_2 + y_2)\vec{e}_2 = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2}$$

S-Multiplikation in Komponentendarstellung:

$$k\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} = k(x_1\vec{e}_1 + x_2\vec{e}_2) = (kx_1)\vec{e}_1 + (kx_2)\vec{e}_2 = \begin{pmatrix} kx_1 \\ kx_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2}.$$

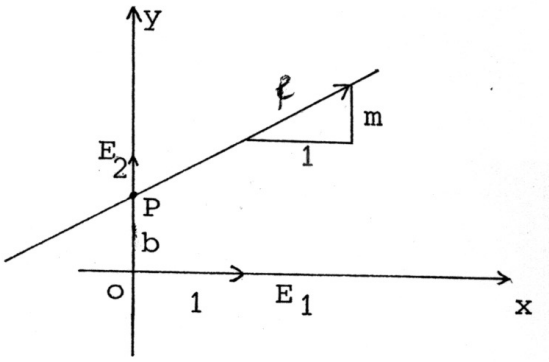
(Schreiben Sie die beiden Rechnungen ausführlicher hin! Welche Rechengesetze wurden benutzt?)

Beispiel.
Sei g Gerade mit Achsenabschnitt b und Steigung $m \neq \infty$. Es geht dann g durch den Punkt P mit Ortsvektor $\vec{p} = \begin{pmatrix} 0 \\ b \end{pmatrix}_{\vec{e}_1,\vec{e}_2}$ und hat $\begin{pmatrix} 1 \\ m \end{pmatrix}_{\vec{e}_1,\vec{e}_2}$ als Richtungsvektor $\vec{f}$. Die (vektorielle) Punktrichtungsgleichung $\vec{x} = \vec{p} + k\vec{f}$ wird zu

$$\begin{pmatrix} x \\ y \end{pmatrix}_{\vec{e}_1,\vec{e}_2} = \begin{pmatrix} 0 \\ b \end{pmatrix}_{\vec{e}_1,\vec{e}_2} + k\begin{pmatrix} 1 \\ m \end{pmatrix}_{\vec{e}_1,\vec{e}_2} = \begin{pmatrix} k \\ b + km \end{pmatrix}$$

woraus $x = k$ und $y = b + km$, also die gewohnte Gleichung für den einen der beiden Geradentypen

$\boxed{y = mx + b}$ folgt.



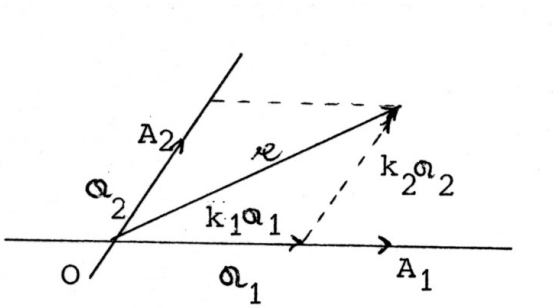
Figur 1.20: Zur Geradengleichung

Anmerkung. Wir haben uns eben eines kartesischen Koordinatensystems bedient. Will man den Begriff des Senkrechtstehens noch vermeiden, so lassen sich ähnliche Betrachtungen auch für ein (sogenanntes affines) Koordinatensystem $(O; A_1, A_2)$ bzw. für $\vec{a}_1$, $\vec{a}_2$ statt $\vec{e}_1$, $\vec{e}_2$ anstellen, wobei von den Punkten O, A_1, A_2 nur gefordert wird, dass sie nicht auf einer Geraden liegen, bzw. von $\vec{a}_1$, $\vec{a}_2$, dass sie nicht parallel sind; s. Figur 1.21 ! Auch dann lässt sich jeder Vektor (eindeutig) in der Form

$$\vec{x} = k_1\vec{a}_1 + k_2\vec{a}_2 = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}_{\vec{a}_1,\vec{a}_2}$$

als **Linearkombination** von $\vec{a}_1$ und $\vec{a}_2$ schreiben.

Wir sagen $\vec{a}_1$ und $\vec{a}_2$ spannen die Ebene E auf. Die Bedingung „$\vec{a}_1$ und $\vec{a}_2$ sind nicht parallel“ bedeutet, dass keine Beziehung der Form $\vec{a}_1 = k\vec{a}_2$ oder $\vec{a}_2 = k\vec{a}_1$ bestehen darf. Fassen wir die beiden letzten Gleichungen zusammen, dann fordern wir für $\vec{a}_1$ und $\vec{a}_2$, dass die Gleichung



Figur 1.21: Affine Koordinaten eines Vektors

$$x_1\vec{a}_1 + x_2\vec{a}_2 = \vec{o}$$

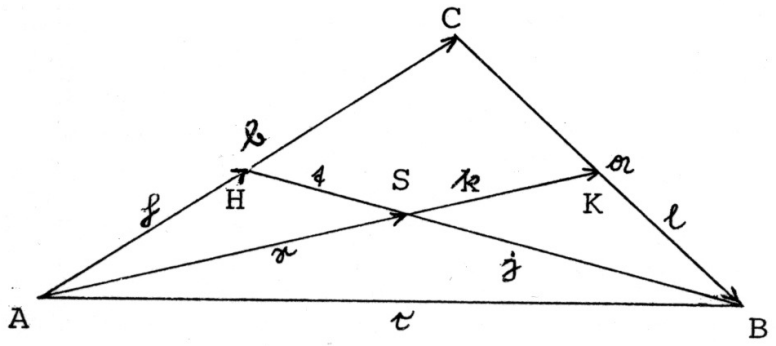
nur die Lösung $x_1 = 0 = x_2$ hat. In diesem Fall nennen wir $\vec{a}_1$ und $\vec{a}_2$ **linear unabhängig**.

Insbesondere sind $\vec{e}_1$ und $\vec{e}_2$ linear unabhängig.

(e) Anhang: Koordinatenfreie Beweise geometrischer Sätze mit Hilfe der Vektorrechnung

Wir gehen hier kurz auf Beweisverfahren für Figuren ein, die sich durch Zeichnen von Geraden, Strecken, Parallelen und Teilen von Strecken konstruieren lassen. Als Beispiel wählen wir nochmals den Satz über die Seitenhalbierenden eines Dreiecks, ohne jedoch das Teilverhältnis als bekannt vorauszusetzen.

(i) In den Figuren auftretende Strecken werden mit einer geeigneten Orientierung versehen, sodass zu Vektoren übergegangen werden kann. (S.z.B. Figur 1.22 !)



Figur 1.22: Zuordnung von Vektoren

Beispiel.

$\vec{s}$	sei	repräsentiert	durch	$\overrightarrow{AS}$		$\vec{t}$	sei repräsentiert .	durch	$\overrightarrow{HS}$
$\vec{k}$	"	"	"	$\overrightarrow{AK}$		$\vec{l}$	"	"	$\overrightarrow{KB}$
$\vec{h}$	"	"	"	$\overrightarrow{AH}$		$\vec{j}$	"	"	$\overrightarrow{HB}$
$\vec{a}$	"	"	"	$\overrightarrow{CB}$		$\vec{b}$	"	"	$\overrightarrow{AC}$
$\vec{c}$	"	"	"	$\overrightarrow{AB}$					

(ii) Zwei geeignete linear unabhängige Vektoren $\vec{a}$, $\vec{b}$ werden ausgewählt. (Im Beispiel ist dies schon geschehen, $\vec{a}$ und $\vec{b}$ sind linear unabhängig.) Jeder Vektor der Ebene lässt sich dann als Linearkombination dieser Vektoren schreiben, zunächst meist nur mit uns unbekannten Koeffizienten.

(iii) Die Voraussetzungen sowie konstruktionsbedingten Beziehungen werden als Gleichungen geschrieben.

Beispiel.
$\vec{h} = \frac{1}{2}\vec{b}$, $\vec{l} = \frac{1}{2}\vec{a}$, $\vec{s} = x\vec{k}$, $\vec{t} = y\vec{j}$ (mit noch zu bestimmenden Zahlen x, y).

(iv) Geeignete geschlossene Pfeilzüge werden durch Gleichungen ausgedrückt.

Beispiel.
$\vec{s} = \vec{h} + \vec{t}$, $\vec{h} + \vec{j} = \vec{c}$, $\vec{k} + \vec{l} = \vec{c}$, $\vec{c} = \vec{a} + \vec{b}$.

(v) Die Gleichungen werden auf Beziehungen zwischen $\vec{a}$ und $\vec{b}$ reduziert.

Beispiel.
$\vec{s} = \vec{h} + \vec{t} = \frac{1}{2}\vec{b} + y\vec{j} = \frac{1}{2}\vec{b} + y(\vec{c} - \vec{h}) = \frac{1}{2}\vec{b} + y(\vec{a} + \vec{b} - \frac{1}{2}\vec{b}) = y\vec{a} + \frac{1}{2}(y+1)\vec{b}$;

andererseits:

$\vec{s} = x\vec{k} = x(\vec{c} - \vec{l}) = x(\vec{a} + \vec{b} - \frac{1}{2}\vec{a}) = \frac{1}{2}x\vec{a} + x\vec{b}$.

Hieraus erhalten wir (nach Koeffizienten von $\vec{a}$ und $\vec{b}$) sortiert:

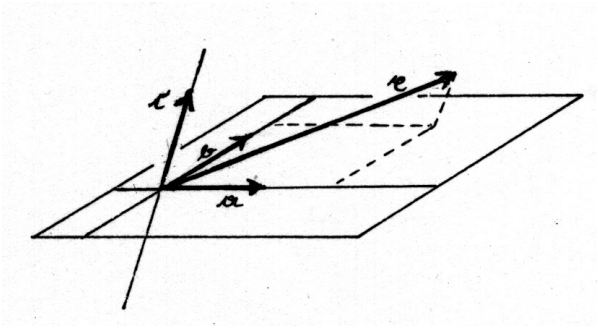
$$(\frac{1}{2}x - y)\vec{a} = [\frac{1}{2}(y+1) - x]\vec{b}.$$

(vi) Da $\vec{a}$ und $\vec{b}$ linear unabhängig (also nicht parallel) sind, folgt jeweils aus einer Gleichung der Form $k\vec{a} + m\vec{b} = \vec{o}$, dass $k = m = 0$ gilt.

Im Beispiel:
$\frac{1}{2}x - y = 0$ und $\frac{1}{2}(y+1) - x = 0$, also $y = \frac{1}{2}x$ und $x = \frac{2}{3}$. Das Teilverhältnis ist als $1 : 2$, wie erwartet.

1.2 Vektoren des Anschaungsraumes

Die Ausführungen von (1.1) a) – c) gelten fast wörtlich, wenn man jeweils „Zeichenebene E “ durch „Anschauungsraum“ ersetzt. Ein Vektor ist dann eine Menge vektorgleicher Pfeile im Raum. Die höhere Dimension schlägt sich lediglich in den Teilen (d) und (e) explizit nieder, wo eine 3. Koordinate und entsprechend eine 3. Komponente (mit einem 3. Koordinatensystemvektor) hinzugezogen werden muss. Der Raum lässt sich nicht durch 2 Vektoren, sondern erst durch drei Vektoren $\vec{a}$, $\vec{b}$, $\vec{c}$ aufspannen. Dabei darf $\vec{a}$ kein Vielfaches von $\vec{b}$ sein und $\vec{c}$ keine Linearkombination von $\vec{a}$ und $\vec{b}$.



Figur 1.23: Affines Koordinatensystem im Raum

1.3 Lösungen von linearen Gleichungssystemen

Die Behandlung linearer Gleichungen ist eine sowohl inner-mathematisch als auch für praktische Anwendungen bedeutsame Problemstellung. Wie von der Analytischen Geometrie gingen von ihr wesentliche Anstöße zur Entwicklung der Theorie der Vektorräume aus.

Diese Zusammenhänge werden plausibel, wenn man bedenkt, dass sich lineare Gleichungen oft geometrisch interpretieren lassen bzw. umgekehrt sich geometrische Gebilde oft durch Gleichungen beschreiben lassen.

Beispiel.

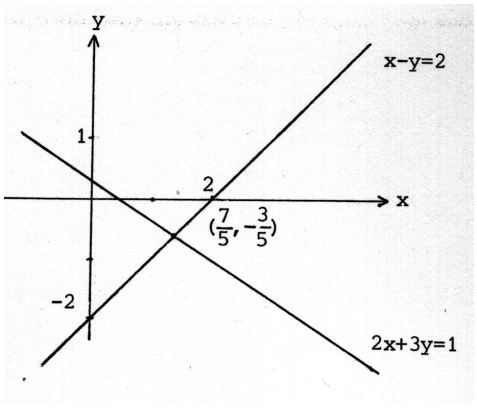
(a) Den beiden Gleichungen

$$\begin{cases} 2x & +3y & = 1 \\ x & -y & = 2 \end{cases}$$

entsprechen im üblichen x-y-Koordinatensystem der Ebene zwei Geraden; deren Schnittpunkt hat, da er auf beiden Geraden liegt, beide Gleichungen erfüllende Koordinaten

$$x = \tfrac{7}{5} \quad \text{und} \quad y = -\tfrac{3}{5}.$$

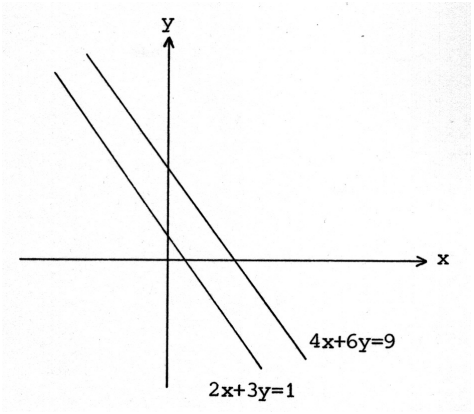
Dies ist die einzige gemeinsame Lösung beider Gleichungen.



Figur 1.24: Graphische Lösung zweier Gleichungen

(b) Zu $\begin{cases} 2x & +3y & = 1 \\ 4x & +6y & = 9 \end{cases}$

gehören in geometrischer Interpretation zwei parallele Geraden. (Beide Geraden haben die Steigung $-\frac{2}{3}$.) Das Fehlen eines Schnittpunktes bedeutet, dass das Gleichungssystem keine Lösung besitzt.



Figur 1.25: Veranschaulichung der Unlösbarkeit eines Gleichungssystems

(c) Betrachten wir die Lösungen der (linearen) Gleichung

$$-2x + 5y + z = 0, \qquad x, y, z \text{ reell.} \qquad (*)$$

Erfüllen die drei reellen Zahlen x_1, y_1, z_1 diese Gleichung, so sprechen wir von einer **Lösung** von $(*)$ und schreiben das „Lösungstripel“ in der Form $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$. Die Gleichung $(*)$ hat u.a. die Lösungen $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$, $\begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$, allgemeiner $\begin{bmatrix} k \cdot 1 \\ 0 \\ k \cdot 2 \end{bmatrix}$ für k reell.

In Anlehnung an die Koordinatenschreibweise der S-Multiplikation bei Vektoren definieren wir zunächst eine „**S-Multiplikation**“ zwischen reellen Zahlen und Lösungstripeln:

$$k \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} := \begin{bmatrix} kx_1 \\ ky_1 \\ kz_1 \end{bmatrix}$$

Diese Definition wird durch die obigen Beispiele von Lösungen nahegelegt, ist aber sonst völlig willkürlich. Wir erhalten damit: Mit dem Tripel $\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ ist auch $k\alpha_1$ Lösungstripel für jede reelle Zahl k.
Analysieren wir den Beweis hierfür („Multiplikation der Gleichung“ mit k), so zeigt sich allgemeiner:

$$\textbf{Ist } \boldsymbol{\xi} \textbf{ Lösung von } (*)\textbf{, so auch } \boldsymbol{k\xi} \ . \qquad \text{(i)}$$

(Beweis ?)
Aber noch haben wir nicht alle Lösungen angegeben, so z.B. nicht solche, für die $x = 0$, $y \neq 0$ ist, wie z.B. $\beta_1 = \begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix}$ (und alle skalaren Vielfachen $k\beta_1$ mit k reell, $k \neq 0$).

Durch Multiplikation der Gleichung $(*)$ mit einer reellen Zahl k kamen wir eben von einer Lösung zu (sogar mehr als endlich) vielen. Es liegt nahe auch andere Operationen mit Gleichungen zu betrachten.

Seien also $\xi_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ und $\xi_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ Lösungen von $(*)$;

es gilt dann	$-2x_1$	$+ 5y_1$	$+ z_1$	$= 0$
und	$-2x_2$	$+ 5y_2$	$+ z_2$	$= 0$.
Wir „addieren“ und erhalten	$-2(x_1 + x_2)$	$+ 5(y_1 + y_2)$	$+ (z_1 + z_2)$	$= 0$.

Es ist also auch $\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix}$ Lösung von $(*)$.

Dies veranlasst uns, in Analogie zur Koordinatendarstellung der Vektoraddition eine **Addition zwischen Lösungstripeln** zu definieren:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} := \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix}.$$

In dieser Schreibweise formuliert, haben wir also folgendes gezeigt:

Sind $\boldsymbol{\xi_1}$ und $\boldsymbol{\xi_2}$ Lösungen von $(*)$, so ist es auch $\boldsymbol{\xi_1 + \xi_2}$. (ii)

(Diese Tatsache hängt wie (i) wesentlich davon ab, dass $(*)$ linear ist und das „konstante Glied“ (rechte Seite) in $(*)$ Null ist; wir sprechen von einer **homogen linearen Gleichung** .)

Mit den Lösungen α_1 und β_1 von $(*)$ folgt aus (i) und (ii) nun: Jede „Linearkombination„ $h\alpha_1 + k\beta_1$, h, k reell, ist Lösung von $(*)$.

Sind dies alle Lösungen? Für eine beliebige Lösung $\xi_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ gilt $z_1 = 2x_1 - 5y_1$, also

$$\xi_1 = \begin{bmatrix} x_1 \\ y_1 \\ 2x_1 - 5y_1 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 2x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ y_1 \\ -5y_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + y_1 \begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix} = x_1\alpha_1 + y_1\beta_1.$$

Damit lässt sich jede Lösung von $(*)$ als Linearkombination von $\boldsymbol{\alpha_1}$ und $\boldsymbol{\beta_1}$ darstellen; umgekehrt ist jede Linearkombination dieser Elemente Lösung von $(*)$.

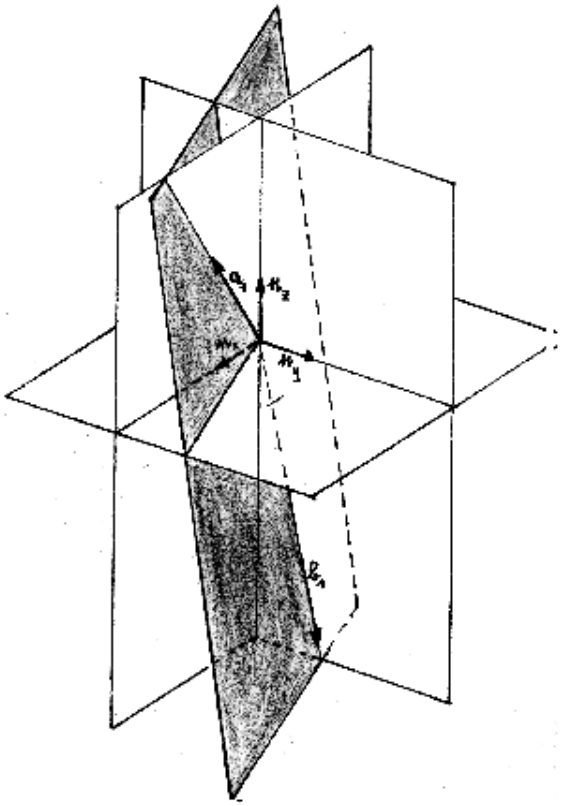
Interpretieren wir x, y, z als die kartesischen Koordinaten von Punkten des Anschauungsraumes (wir schreiben sie als Ortsvektoren), so entspricht einem Lösungstripel $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$

nun $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}_{\vec{e}_x,\vec{e}_y,\vec{e}_z}$. Die durch die Gleichung $(*)$ bestimmten Vektoren sind dann Linearkombinationen der Vektoren $\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}_{\vec{e}_x,\vec{e}_y,\vec{e}_z}$ (entspricht α_1) und $\vec{b_1} = \begin{pmatrix} 0 \\ 1 \\ -5 \end{pmatrix}_{\vec{e}_x,\vec{e}_y,\vec{e}_z}$ (entspricht β_1).

Geometrisch gesehen ist also

$$-2x + 5y + z = 0$$

die Gleichung einer Ebene durch den Ursprung.



Figur 1.26: Eine Ebene als Lösung einer linearen Gleichung.

Es hat sich hier gezeigt, dass die „Anleihe “ bei der Vektorrechnung (Addition, S-Multiplikation, geometrische Interpretation) wesentlich dazu beigetragen hat, die Lösung der Gleichung $(*)$ „in den Griff zu bekommen“. Dies hängt damit zusammen, dass in der Menge der Lösungen sich jedes Tripel wie ein geometrischer Vektor verhält. *Formal* kommt dies durch die Möglichkeit einer analogen Schreibweise, einer analogen Definition von Addition und S-Multiplikation sowie durch die **Gültigkeit analoger Rechengesetze** zum Ausdruck:

Es gilt nämlich wieder (für alle Lösungstripel α, β, γ)

- Assoziativität der Addition: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
- Existenz eines Nullelements: $O := \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (triviale Lösung) mit $\alpha + O = \alpha$.
- Möglichkeit zur Subtraktion: Zu jedem α existiert ein $-\alpha$ mit $\alpha + (-\alpha) = O$.
- Kommutativgesetz der Addition: $\alpha + \beta = \beta + \alpha$

sowie

- Assoziativgesetz der S-Multiplikation: $k(h\alpha) = (k \cdot h)\alpha$.

- 1. Distributivgesetz: $(k+h)\alpha = k\alpha + h\alpha$.
- 2. Distributivgesetz: $k(\alpha+\beta) = k\alpha + k\beta$.
- Neutralität der reellen 1: $1\alpha = \alpha$.

Unter (i) und (ii) haben wir außerdem bewiesen, dass Summenbildung und S-Multiplikation nicht aus der Menge der Lösungen heraus führt. Alle diese Tatsachen sowie die Darstellbarkeit der Lösungen als reelle Linearkombinationen von α_1 und β_1 finden ihren Niederschlag in der Möglichkeit, die Lösungen geometrisch als Ebene im Raum darzustellen.

Ergänzung zu (1.2) & (1.3)

(i) Skalarprodukt

Seien $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ Einheitsvektoren eines kartesischen Koordinatensystems des Raumes; seien

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\vec{e}_1,\vec{e}_2,\vec{e}_3} \quad \text{und} \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{\vec{e}_1,\vec{e}_2,\vec{e}_3}$$

beliebige Vektoren des Raumes. Dann definiert man das **(kanonische) Skalarprodukt** von $\vec{x}$ und $\vec{y}$ durch[2]

$$\vec{x} \cdot \vec{y} := \sum_{i=1}^{3} x_i y_i.$$

Zu den Rechenregeln beachte man Aufgabe 1.2a !
Auf dem Skalarprodukt lassen wir auch folgende Begriffe „aufbauen“:

Betrag (Länge) eines Vektors: $|\vec{x}| = \sqrt{\vec{x}\cdot\vec{x}} = \sqrt{x_1^2+x_2^2+x_3^2}$ (in Übereinstimmung mit $|\vec{e}_i| = 1$ und der geometrischen Länge der Pfeile des Vektors $\vec{x}$). [3]
Orthogonalität zweier Vektoren: $\vec{x}\perp\vec{y}$: $\iff \vec{x}\cdot\vec{y} = 0$ (insbesondere $\vec{e}_i\perp\vec{e}_j$ für $i \neq j$).
Anmerkung: Man kann zeigen, dass $\vec{x}\cdot\vec{y} = |\vec{x}||\vec{y}|\cos\sphericalangle(\vec{x},\vec{y})$ gilt.

(ii) Zur Ebenengleichung

Sei nun die Gleichung

$$a_1x_1 + a_2x_2 + a_3x_3 = b \tag{1}$$

gegeben, wobei nicht alle a_i Null seinen.
Jeder Lösung $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ ordnen wir den Punkt des Raumes mit Ortsvektor $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\vec{e}_1,\vec{e}_2,\vec{e}_3}$ zu. Welches Gebilde beschreibt dann Gleichung (1)?

[2] Dabei bezeichnet $\sum_{i=1}^{k} a_i$ die Summe $a_1 + a_2 + \ldots a_k$.
[3] Beweis durch zweimalige Anwendung des Satzes von Pythagoras!

Setzt man $\vec{m} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}_{\vec{e_1},\vec{e_2},\vec{e_3}}$, so lautet (1) nun $\vec{m} \cdot \vec{x} = b$.

Sei $\vec{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}_{\vec{e_1},\vec{e_2},\vec{e_3}}$ Ortsvektor eines Punktes, dessen Koordinaten (1) erfüllen; solch ein Punkt existiert (warum?).
Wegen $b = \vec{m}\vec{p}$ (denn $\vec{p}$ erfüllt definitionsgemäß (1)) ist auch folgende Gleichung zu (1) äquivalent:

$$\vec{m}(\vec{x} - \vec{p}) = 0. \tag{2}$$

Mit $\vec{\hat{x}} := \vec{x} - \vec{p}$ erhalten wir $\vec{m}\vec{\hat{x}} = 0$. Ähnlich wie im Beispiel (1.3) erhählt man allgemein: es existieren zwei (linear unabhängige) Vektoren $\vec{\hat{y}_1}$, $\vec{\hat{y}_2}$ derart, dass $\vec{m}\vec{\hat{y}} = 0$ ist genau dann, wenn $\vec{\hat{y}} = \vec{y} - \vec{p} \in \mathbb{R}\vec{\hat{y}_1} + \mathbb{R}\vec{\hat{y}_2}$ $(:= \{k_1\vec{\hat{y}_1} + k_2\vec{\hat{y}_2} | k_1, k_2 \in \mathbb{R}\})$. Als Lösungsmenge von (1) ergibt sich dann $\vec{p} + \mathbb{R}\vec{\hat{y}_1} + \mathbb{R}\vec{\hat{y}_2}$, also eine **Ebene** E, die wegen $\vec{m} \perp (\vec{x} - \vec{p})$ für jedes $\vec{x}$, das Ortsvektor eines Punktes von E ist, orthogonal zu $\vec{m}$ ist. Normiert man $\vec{m}$ zu $\vec{n} = \frac{\vec{m}}{|\vec{m}|}$ (Länge 1), so gilt auch $\vec{n}(\vec{x} - \vec{p}) = 0$.
Durch die Gleichung $\vec{n}(\vec{x} - \vec{p}) = 0$ mit $|\vec{n}| = 1$ ist eine Ebene definiert. Hierbei ist $\vec{p}$ Ortsvektor eines Punktes der Ebene, sowie $\vec{n}$ ein normierter Richtungsvektor einer auf E „senkrecht stehenden" Geraden.
Die Gleichung $\vec{n}\vec{x} = d$ mit $d = \vec{n}\vec{p}$ (bzw. $\vec{n}\vec{x} - d = 0$ oder $\vec{n}(\vec{x} - \vec{p}) = 0$) mit $|\vec{n}| = 1$ heißt **Hessesche Normalform** der Gleichung der Ebene E. (Siehe auch Figur 1.27 !)

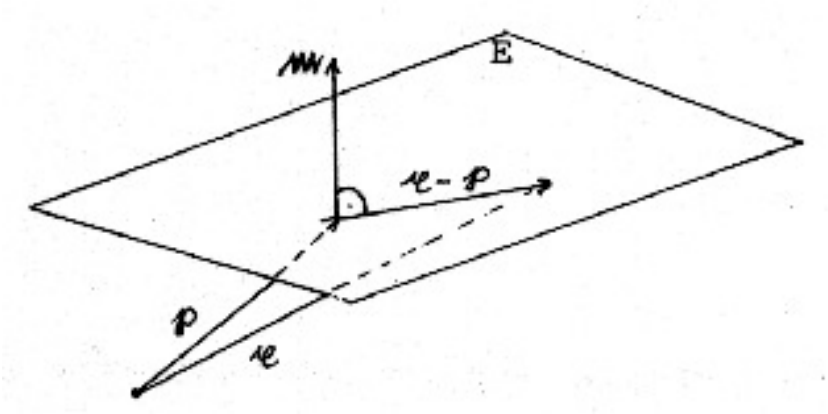
Beispiel. $-3x_1 + 2x_2 - 6x_3 = -27$
Es ist $|\vec{m}| = \left| \begin{pmatrix} -3 \\ 2 \\ -6 \end{pmatrix} \right| = \sqrt{9 + 4 + 36} = 7$, und die Hessesche Normalform der zugehörigen Ebenen-Gleichung lautet: $-\frac{3}{7}x_1 + \frac{2}{7}x_2 - \frac{6}{7}x_3 + \frac{27}{7} = 0$. Die Ebene hat als Normalenvektor: $\vec{n} = \frac{1}{7}\begin{pmatrix} -3 \\ 2 \\ -6 \end{pmatrix}_{\vec{e_1},\vec{e_2},\vec{e_3}}$.

(iii) Vektorprodukt (Kreuzprodukt)

Neben dem Skalarprodukt ist noch eine weitere Produktbildung bei Vektoren des Raumes von Interesse, das **Vektorprodukt**. Bezüglich eines kartesischen Koordinatensystems definiert man

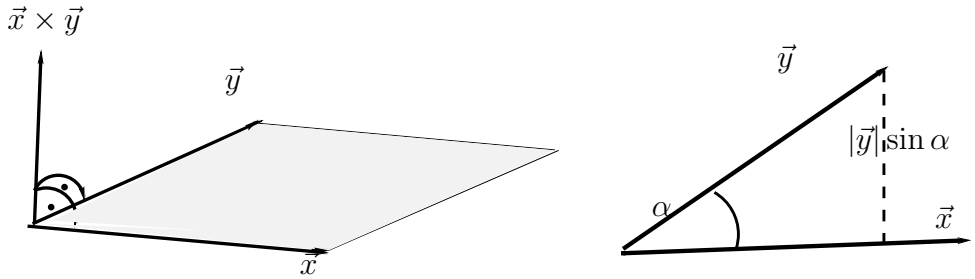
$$\vec{x} \times \vec{y} := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{\vec{e_1},\vec{e_2},\vec{e_3}} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{\vec{e_1},\vec{e_2},\vec{e_3}} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}_{\vec{e_1},\vec{e_2},\vec{e_3}}$$

Man kann zeigen, dass



Figur 1.27: Zur Normalform einer Ebenengleichung : $\vec{m}(\vec{x} - \vec{p}) = 0$

- $\vec{x} \times \vec{y} = 0$ genau im Falle der linearen Abhängigkeit von $\vec{x}$ und $\vec{y}$ gilt;
- im anderen Falle $\vec{x} \times \vec{y}$ auf der von $\vec{x}$ und $\vec{y}$ aufgespannten Nullpunktsebene senkrecht steht;



a) b)

Figur 1.28: a) $\vec{x} \times \vec{y}$ steht senkrecht auf der von $\vec{x}$ und $\vec{y}$ aufgespannten Ebene.
b) Zum Flächeninhalt des von $\vec{x}$ und $\vec{y}$ aufgespannten Parallelogramms.

- die Länge $|\vec{x} \times \vec{y}|$ des Vektorprodukts gleich dem Flächeninhalt des von $\vec{x}$ und $\vec{y}$ aufgespannten Parallelogramms ist, nämlich $|\vec{x} \times \vec{y}| = |\vec{x}| \cdot |\vec{y}| \cdot \sin \sphericalangle(\vec{x}, \vec{y})$. (Siehe Figur 1.28 !)

  Damit ist das Vektorprodukt u.a. zur Bestimmung von Flächeninhalten, Normalenvektoren und Winkelgrößen benutzbar.

Achtung: Assoziativgesetz und Kommutativgesetz sind verletzt!

Beispiel: $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3} \times \begin{pmatrix} 0 \\ 1 \\ -5 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3} = \begin{pmatrix} 0 \cdot (-5) - 2 \cdot 1 \\ 2 \cdot 0 - 1(-5) \\ 1 \cdot 1 - 0 \cdot 0 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3} = \begin{pmatrix} -2 \\ 5 \\ 1 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3}$ steht in Übereinstimmung damit, dass $\begin{pmatrix} -2 \\ 5 \\ 1 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3}$ ein Normalenvektor der von $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3}$ und

$\begin{pmatrix} 0 \\ 1 \\ -5 \end{pmatrix}_{\vec{e}_1\vec{e}_2\vec{e}_3}$ aufgespannten, in (1.3)(c) durch die Gleichung $-2x + 5y + z = 0$ beschriebenen Ebene ist.

1.4 Lösung einer homogenen linearen Differentialgleichung

Wir untersuchen die Differentialgleichung

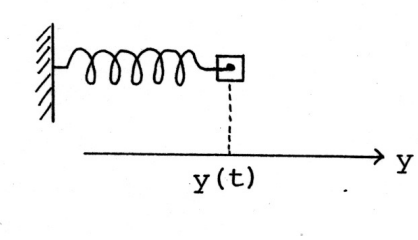
$$y''(t) + 4y(t) = 0 \qquad (**)$$

($y = y(t)$ bezeichen dabei eine reelle Funktion und y'' die zweimalige Ableitung von y nach t, die oft auch als $\ddot{y}$ geschrieben wird).

Anwendungsbeispiel. Ungedämpfter harmonischer Oszillator: Ein Körper der Masse $m = 1$ kg sei an einer Feder der Federkonstanten $D = 4\text{N/Meter}$ befestigt. Feder- und Trägheitskraft seien die einzigen auf den Körper wirkenden Kräfte.
Bei der Auslenkung $y(t)$ aus der Ruhelage gilt dann zu jedem Zeitpunkt t für die rücktreibende Kraft $K = -Dy$ (s. Figur 1.29 !) und damit

$$m \cdot y''(t) = -D \cdot y(t).$$

Geben wir t in sec, y in Meter an, so gilt $(**)$ für die Maßzahlen. Von Anfangsbedingungen sehen wir zunächst ab. Die Differentialgleichung $(**)$ besitzt Lösungen, d.h. zweimal differenzierbare reelle Funktionen, die die Gleichung erfüllen. Unter Anwendung der Kettenregel verifiziert man, dass z.B. die Funktion y_1 mit



Figur 1.29:
Kraft bei Auslenkung aus der Ruhelage

$$y_1(t) = \sin 2t$$

eine solche Lösung ist.
Um einen Überblick über die Lösungen zu erhalten, überlegen wir uns (in Analogie zu (1.3)), wie wir gegebenenfalls aus bekannten Lösungen neue konstruieren können.
Sei etwa y eine Lösung; dann folgt aus $y''(t) + 4y(t) = 0$ durch Multiplikation mit reellem k

$$ky''(t) + 4ky(t) = 0.$$

Bezeichnen wir diejenige Funktion, deren Werte für jedes t gerade das k-fache der entsprechenden Werte von y sind, mit ky, sodass also

$$(ky)(t) := k \cdot y(t)$$

und $(ky)''(t) = ky''(t)$ für alle t gilt, **dann ist auch ky Lösung von** $(**)$. Insbesondere ist k_1y_1 mit $(k_1y_1)(t) = k_1 \cdot \sin 2t$ eine Lösung für jedes reelle k_1. Ähnlich ist y_2 mit $y_2(t) = \cos 2t$ Lösung und damit für jedes reelle k auch k_2y_2 mit $(k_2y_2)(t) = k_2 \cos 2t$.
Die Analogie zu (1.3) legt nahe, auch hier die „Summe“ y_1 und y_2 zweier Lösungen y_1 und y_2 zu untersuchen. Dazu definieren wir die Summe von Funktionen y_1 und y_2 durch

$$(y_1 + y_2)(t) := y_1(t) + y_2(t).$$

Durch Addition der Gleichungen $y_1''(t) + 4y_1(t) = 0$ und $y_2''(t) + 4y_2(t) = 0$ und Anwendung der Differentiationsregel für Summen ergibt sich: **Sind y_1, y_2 Lösungen von** $(**)$**, so ist auch $y_1 + y_2$ Lösung.**

Die Menge der Lösungen von $(**)$ ist also abgeschlossen bzgl. Multiplikation mit einem Skalar („S-Multiplikation“) und bzgl. Addition.

In Anwendung auf die bereits gefundenen Lösungen ($y_1(t) = \sin 2t$ und $y_2(t) = \cos 2t$) ergibt sich: Jede Linearkombination y mit

$$y(t) = k_1 \sin 2t + k_2 \cos 2t \ , \quad k_1, k_2 \text{ reell},$$

('Superposition') ist Lösung. Es ist möglich zu zeigen dass jede Lösung von $(**)$ diese Form hat.

Anmerkung 1. Vorgegebene **Anfangsbedingungen** sondern dann Funktionen mit bestimmten k_1, k_1, aus der Menge der Lösungen aus. Wegen $y(0) = k_2$ und $y'(0) = 2k_1$ wird z.B. die Bedingung

(i) $y(0) = 0$ und $y'(0) = 0$ nur durch die Funktion $\vec{o}$ mit $\vec{o}(t) = 0$ für alle t (**Nullfunktion**) (Körper in Ruhelage) bzw.

(ii) $y(0) = 2$ und $y'(0) = 1$ nur durch die Funktion y mit
$$y(t) = \tfrac{1}{2} \sin 2t + 2 \cos 2t = \sqrt{\tfrac{1}{4} + 4} \sin(2t + \varphi) \text{ (für } \tan \varphi = 4)$$
(ungedämpfte harmonische Schwingung) erfüllt.

Anmerkung 2. Wir vermerken, dass für Addition und S-Multiplikation bei Funktionen, insbesondere bei den Lösungen von $(**)$, entsprechende Rechengesetze wie bei den vorherigen Beispielen gelten, also:

Assoziativität der Addition Kommutativität der Addition Existenz einer Null (hier Nullfunktion $\vec{o}$) mit $f + \vec{o} = f$ Möglichkeit zur Subtraktion (Umkehrung der Addition): Zu f definieren wir $-f$ durch $(-f)(t) = -f(t)$; dann ist $f + (-f) = \vec{o}$ und $(g + f) + (-f) = g$. Assoziativgesetz der S-Multiplikation Beide Distributivgesetze der S-Multiplikation

Beweis. Übungsaufgabe

Anmerkung 3. Wie in den vorangegangenen Beispielen fällt auf, dass sich alle vorkommenden Vektoren als „Linearkombinationen“ von einigen wenigen darstellen lassen. Die Angabe solcher Erzeugenden ist ein wesentliches Beschreibungsprinzip der (linearen) Algebra.

1.5 Ein Code (als binäres Modell)

(a) Dualcode

Beim Dualcode werden Zahlen im **Binärsystem** (Dualsystem) dargestellt. Dieses Zahlsystem verwendet statt der Zahlen 0 bis 9 des Dezimalsystems nur 0 und 1. Bei beiden Systemen gibt die Stellung einer Ziffer in der Darstellung einer Zahl ihre Wertigkeit an (Positionssysteme), beim Binärsystem aber nicht als Potenzen zur Basis 10, sondern zur Basis 2.

Beispiel. dezimal 1001 = dezimal $(1 \cdot 10^3 + 0 \cdot 10^2 + 0 \cdot 10^1 + 1 \cdot 10^0)$
binär 1001 = dezimal $(1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0)$ = dezimal 9.
Die folgende Tabelle zeigt die ersten 10 Zahlen:

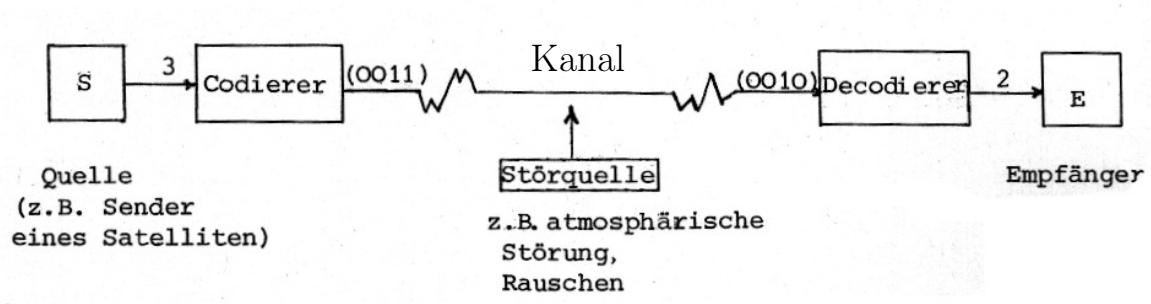
dezimal		dual	„Auffüllen“ zu „Wörtern“ der Länge 4 (Tetraden)
0	$0 \cdot 2^0$	0	0 0 0 0
1	$1 \cdot 2^0$	1	0 0 0 1
2	$1 \cdot 2^1 + 0 \cdot 2^0$	1 0	0 0 1 0
3	$1 \cdot 2^1 + 1 \cdot 2^0$	1 1	0 0 1 1
4	$1 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0$	1 0 0	0 1 0 0
5	$1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$	1 0 1	0 1 0 1
6	$1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0$	1 1 0	0 1 1 0
7	$1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0$	1 1 1	0 1 1 1
8	$1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0$	1 0 0 0	1 0 0 0
9	$1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$	1 0 0 1	1 0 0 1

Tabelle 1.1: Binäre Darstellung der ersten 10 Zahlen

Bei größeren Zahlen kann man entsprechend codieren (geschlossene Codierung) oder jede Dezimalstelle einzeln codieren. *Anmerkung:* Es handelt sich hier nur um eine von vielen Möglichkeiten der numerischen Codierung; s. auch “ ASCII”-Code, “Unicode”, “UTF-8” etc.).

(b) Datenübertragung über gestörte Kanäle und binäre Addition

Ein (schematisches) Beispiel einer Übertragung über einen gestörten Kanal (Funkstrecke/Atmosphä
oder Kabel, Magnetband, Glasfaser etc.) zeigt Figur 1.30.



Figur 1.30: Schema eines Übertragungssytems und Übertragungsbeispiel

	die gesendete Nachricht	$N_1 =$	O	O	1	1
	das empfangenes Signal	$E_1 =$	O	O	1	O
Dabei sind:	der „Fehlervektor“	$F_1 =$	O	O	O	1.

Zur mathematischen Beschreibung des Sachverhalts setzen wir $N = E \oplus F$, wobei „komponentenweise“ folgendermaßen zu „addieren“ ist: [4]

$0 \oplus 0 = 0$	(kein Fehler)
$1 \oplus 0 = 1$	(kein Fehler)
$0 \oplus 1 = 1$	(Fehler)
$1 \oplus 1 = 0$	(Fehler).

Die letzten 4 Vereinbarungen schreiben wir auch in der Form einer **„Verknüpfungstafel“**:

$\oplus$	0	1
0	0	1
1	1	0 .

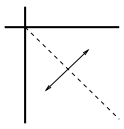
Anmerkung: Wie man sieht und wie wir weiter unten noch vertiefen werden, hängt die Addition eng mit der Anzahl der Summanden „1“ in einer Summe[5] zusammen. Ist diese Anzahl (die sogenannte **Parität**) gerade, so ist die Summe 0, andernfalls 1. Ersetzt man daher 0 durch g (wie „gerade“) und 1 durch u (wie „ungerade“) , so ergibt sich als Verknüpfungstafel:

$\oplus$	g	u
g	g	u
u	u	g .

Diese entspricht den bekannten Additionsregeln:

gerade	plus	gerade	=	gerade
gerade	plus	ungerade	=	ungerade
ungerade	plus	gerade	=	ungerade
ungerade	plus	ungerade	=	gerade .

Wir kommen zu weiteren Eigenschaften dieser Addition auf $\{0, 1\}$ (bzw. $\{g, u\}$):
Ist die *Abgeschlossenheit* der Addition schon vorher ersichtlich, so erkennt man an der Symmetrie



der Verknüpfungstafel sofort die *Kommutativität der Addition*; auch die *Neutralität* der Null ist leicht zu erkennen. Jedes Element ist sein eigenes *Inverses*. Schließlich gilt noch das *Assoziativgesetz*, wie sich durch Prüfen der einzelnen Fälle zeigt.

[4] also nicht, wie bei der gewöhnlichen Addition im Dualsystem, $1 + 1 = 10$, d.h. 0 mit Übertrag 1. Beachten Sie auch die untenstehende Anmerkung zu einer Addition auf $\{u, g\}$!

[5] auch mit mehreren Summanden

(c) Codierung zur Fehlererkennung

Mit dem Ziel, **einen** Übertragungsfehler pro Wort erkennen zu können, fügen wir den bisherigen Codewörtern noch je ein Kontrollsymbol (ähnlich den Prüfziffern bei Kontonummern) hinzu. Dabei benutzen wir die eben definierte Addition auf $\{0,1\}$ und bilden die „Quersumme“ der Informationssymbole (**„Paritätskontrolle“**): Anstelle von $a_1a_2a_3a_4$ senden wir $\underbrace{a_1a_2a_3a_4}_{\text{Informationsteil}}\ \underbrace{a_1 \oplus a_2 \oplus a_3 \oplus a_4}_{\text{Kontrollziffer}}$.

Beispiel. Für 0 0 1 1 senden wir 0 0 1 1 0.

Anmerkung. Auch hier haben wir wieder nur eine von vielen Möglichkeiten zur Codierung herausgegriffen; der entsprechende Code heißt **Paritätscode**; bei ihm ist die Anzahl der Einsen im Codewort gerade, d.h. die binäre Quersumme der Symbole gleich 0:

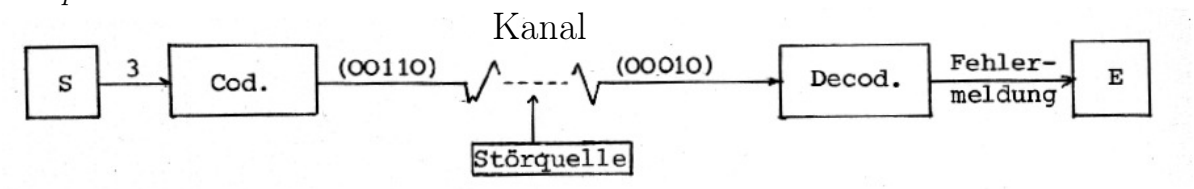
$$\begin{aligned} a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus (a_1 \oplus a_2 \oplus a_3 \oplus a_4) & \quad \text{Assoziativgesetz, Kommutativgesetz} \\ &= (a_1 \oplus a_1) \oplus (a_2 \oplus a_2) \oplus (a_3 \oplus a_3) \oplus (a_4 \oplus a_4) \quad \text{Selbstinverse} \\ &= 0 \oplus 0 \oplus 0 \oplus 0 \quad \text{Neutralität der 0} \\ &= 0 \end{aligned}$$

Codewörter sind daher Wörter der Form $a_1a_2a_3a_4a_5$, die die (Paritätskontroll-) Gleichung

$$a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5 = 0$$

erfüllen (und umgekehrt). Wörter, die durch Abänderung einer Komponente aus einem Codewort entstehen, haben Quersumme 1 und sind damit als fehlerhaft zu erkennen:

Beispiel.



Figur 1.31: Beispiel einer Übertragung mit Fehlererkennung

(d) Weiter Strukturierung

Wir vermerken, dass wir auch zwischen den Wörtern der Länge 5 aus Nullen und Einsen eine Addition komponentenweise definieren können:

Beispiel.

$$\begin{array}{rcccccc} E_2 & = & 0 & 0 & 0 & 1 & 0 \\ F_2 & = & 0 & 0 & 1 & 0 & 0 \\ \hline N_2 = E_2 \oplus F_2 & = & 0 & 0 & 1 & 1 & 0\ . \end{array}$$

Dabei übertragen sich viele der Rechengesetze, insbesondere die von uns herausgestellten: Abgeschlossenheit der Addition, Assoziativgesetz, Existenz eines neutralen Elements (hier 00000), Existenz eines Inversen zu jedem Wort (hier das Wort selbst) und das Kommutativgesetz.

Ferner ist es möglich, eine (banal erscheinende) **S-Multiplikation** zu definieren durch $0 \cdot W = 00000$ bzw. $1 \cdot W = W$ für jedes der betrachteten Wörter W.

Definiert man, motiviert auch durch die Multiplikationsregeln

$$\begin{array}{ccccc} \text{gerade} & \text{mal} & \text{gerade} & = & \text{gerade} \\ \text{gerade} & \text{mal} & \text{ungerade} & = & \text{gerade} \\ \text{ungerade} & \text{mal} & \text{gerade} & = & \text{gerade} \\ \text{ungerade} & \text{mal} & \text{ungerade} & = & \text{ungerade} \end{array},$$

eine Multiplikation zwischen den Elementen $0\,(\hat{=}\,g)$ und $1\,(\hat{=}\,u)$ durch folgende Verknüpfungstafel:

$\odot$	0	1
0	0	0
1	0	1

,

so kann man die S-Multiplikation wieder komponentenweise ausführen. Für sie gilt neben der Abgeschlossenheit und den beiden Distributivgesetzen auch das Assoziativgesetz

$$k \cdot (l \cdot W) = (k \odot l) \cdot W.$$

Damit ist, zumindest was die Rechengesetze betrifft, eine Analogie zwischen der eben behandelten Struktur und den vorigen hergestellt; als neuer „Tatbestand“ ist nun hinzugekommen, dass die **Skalare nicht mehr reelle Zahlen** sind (jedoch noch den meisten der für reelle Zahlen gültigen algebraischen Rechenregeln genügen.)
Ähnlich wie in (1.3) lassen sich die Lösungen der linearen Gleichung $a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5 = 0$, die ja Codewörter $a_1a_2a_3a_4a_5$ ergeben, als Linearkombinationen einiger weniger darstellen. „Erzeugende“ sind z.B. die Codewörter $\alpha_1 = 10001$, $\alpha_2 = 01001$, $\alpha_3 = 00101$ und $\alpha_4 = 00011$. Jedes Codewort ist dann von der Form

$$k_1\alpha_1 \oplus k_2\alpha_2 \oplus k_3\alpha_3 \oplus k_4\alpha_4 \quad \text{mit } k_1, k_2, k_3, k_4 \in \{0, 1\}.$$

Es wird sich herausstellen, dass die bei jedem der obigen Beispiele herausgestellten Rechengesetze deren (algebraische) Struktur schon wesentlich festlegen, sodass eine **gemeinsame Theorie** möglich und sinnvoll wird. Diese Theorie ist die des (abstrakten) Vektorraums (auch linearer Raum genannt), des Hauptgegenstandes unserer Vorlesung. Parallel zu diesem Thema versuchen wir, unsere Sprache zu präzisieren und zu normieren. Dies wird, nach evtl. anfänglichen Schwierigkeiten, sowohl die Kommunikation erleichtern als auch die Argumentation sicherer und nachprüfbarer gestalten.

Übungsaufgaben:
Aufgabe 1.1 Bestimmen Sie die Koeffizienten a_1, a_0 der Differentialgleichung

$$y'' + a_1y' + a_0y = 0$$

so, dass y_1 und y_{-1} mit $y_1(x) = e^x$ und $y_{-1}(x) = e^{-x}$ Lösungen dieser Gleichung sind. Sind dann auch die (Hyperbel-)Funktionen sinh und cosh mit

$$\sinh(x) := \frac{e^x - e^{-x}}{2} \quad \text{und} \quad \cosh(x) = \frac{e^x + e^{-x}}{2}$$

Lösungen dieser Differentialgleichung?

Anmerkung: Ohne Beweis dürfen Sie hier Eigenschaften der Exponentialfunktion verwenden, u.a. die Tatsache, dass $e^x \neq 0$ für alle $x \in \mathbb{R}$ gilt und $(e^x)' = e^x$ ist.

Aufgabe 1.2 a) Zeigen Sie für das kanonische Skalarprodukt des Raumes:

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

für alle Vektoren $\vec{a}, \vec{b}, \vec{c}$ des Raumes!
b) Benutzen Sie Teil a) zum Nachweis der Formel

$$|\vec{a} + \vec{b}|^2 + |\vec{a} - \vec{b}|^2 = 2(|\vec{a}|^2 + |\vec{b}|^2) \quad !$$

c) Warum heißt die Gleichung aus b) „**Parallelogrammgleichung**“? Interpretieren Sie sie elementargeometrisch mit Hilfe einer Skizze!

Aufgabe 1.3 [6]:
Wir definieren eine Abbildung von binären Tetraden auf binäre 8-Tupel

$$\mathfrak{c} : a_1a_2a_3a_4 \mapsto a_1a_2a_3a_4a_5a_6a_7a_8$$

durch

$$\begin{aligned} a_5 &:= a_1 \oplus a_3; \\ a_6 &:= a_2 \oplus a_4; \\ a_7 &:= a_1 \oplus a_2; \\ a_8 &:= a_3 \oplus a_4. \end{aligned}$$

Zeigen Sie:
1. $\mathfrak{c}$ ist additiv, d.h. es gilt : $\mathfrak{c}(a_1a_2a_3a_4 \oplus b_1b_2b_3b_4) = \mathfrak{c}(a_1a_2a_3a_4) \oplus \mathfrak{c}(b_1b_2b_3b_4)$.
2. Einen Übertragungsfehler in einer Komponente eines Wortes kann man nicht nur feststellen, sondern sogar korrigieren.
Lösungshinweis: Schreiben Sie $a_1a_2a_3a_4$ in ein $2 \times 2-$ Quadrat, und addieren Sie die Einträge jeder Zeile und jeder Spalte!

[6] Starke Vereinfachung eines bei Magnetbändern verwendeten Codes !

2 Zur normierten Sprache der Mathematiker

Hinweis: Inhalte dieses Paragraphen werden zum Teil in der parallelen Vorlesung 'Elementare Algera/ Zahlentheorie I' behandelt und daher hier übersprungen.

Literatur:

Deiser, Oliver: Einführung in die Mengenlehre. 19.9.2018
http://www.aleph1.info/?call=Puc& permalink=mengenlehre1

Deiser, Oliver: Einführung in die Mengenlehre. Die Mengenlehre Georg Cantors und ihre Axiomatisierung durch Ernst Zermelo. Springer. Berlin 2002.

Loomis-Sternberg: Advanced Calculus (Einleitung). https://archive.org/details/LoomisL.H.SternbergS.AdvancedCalculusRevisedEditionJonesAndBartlett

Gebraucht erhaltbar:
Halmos, P.R.: Naive Mengenlehre, Göttingen 1968.
Felscher, W.: Naive Mengen und abstrakte Zahlen I, Zürich 1978.
Warlich,L.: Grundlagen der Mathematik für Studium und Lehramt I. 1996/ 2006.

2.1 Naive Logik

(a) Aussagen und Aussageformen

Mathematische **Aussagen** sind sprachliche Gebilde, denen entweder der Wert „wahr“ (w) oder der Wert „falsch“ (f) zu, geordnet werden kann. Eine andere Möglichkeit soll es für eine Aussage nicht geben („tertium non datur“, Zweiwertigkeitsprinzip).

Beispiel.

4 ist eine gerade Zahl	(w)
$4+3=6$	(f)
$4+x=6$	keine Aussage

Sprachliche Gebilde mit „Variablen“ („freien Variablen “, s.u.), die nach Ersetzen dieser Variablen in Aussagen (egal, ob wahr oder falsch) übergehen, heißen **Aussageformen** (Prädikate oder Formeln der Aussagenlogik).

Beispiel.

$4+x=6$ und x ist natürliche Zahl (Gleichung !)
$x \neq x$
$x < y$
sind Aussageformen.

Meist wird auch die Menge der Elemente, die man für die Variablen einsetzen will und kann, der Grundbereich M der Aussageform, angegeben. Ist $A(x)$ eine Aussageform in der nur die Variable x vorkommt (z.B. $A(x): 4+x=6$ mit Grundbereich $\mathbb{N}$ der Menge der natürlichen Zahlen), so erhalten wir also durch Ersetzen von x durch ein Element des Grundbereichs (z.B. durch 5) eine Aussage (z.B. $A(5): 4+5=6$), die dann wieder einen Wahrheitswert besitzt (hier: falsch). (Ist $A(x)$ Gleichung, so heißt x_1 aus M **Lösung**, falls $A(x_1)$ wahr ist).

Da Aussageformen durch jede Ersetzung in Aussagen übergehen, kann man mit ihnen meist hantieren wie mit Aussagen (s.u. Teil (c)).

(b) Quantoren

Die oben beschriebene Einsetzung von Elementen in eine Aussageform ist nur eine Möglichkeit, von einer Aussageform zu einer Aussage zu gelangen. Eine weitere ist es, die „Erfüllbarkeit“ einer Aussageform zu betrachten:
„Es existiert (mindestens) ein x aus M mit der Eigenschaft, dass $A(x)$ wahr ist.“
Dies ist dann wieder eine Aussage mit Wahrheitswert „wahr“ oder „falsch“, eine sogenannte
Existenzaussage.

Sprech- und Schreibweise: Es existiert ein x aus M mit $A(x)$;
ex. $x \in M : A(x)$;
$\boxed{\exists x \in M : A(x)}$,
$\bigvee_{x \in M} A(x)$.

Beispiele:

(i) Es existiert eine natürliche Zahl x mit $4 + x = 6$, kurz

$\exists x \in \mathbb{N} : 4 + x = 6$ (Wahrheitswert w)

(ii) $\exists x \in \mathbb{N} : x \neq x$ (Wahrheitswert f)

Anmerkungen.

1.) In „$\exists x : A(x)$“ kommt zwar noch eine Variable x formal vor, jedoch ist es nicht mehr erlaubt, ein Element des Grundbereichs einzusetzen. Wir sprechen von einer **gebundenen Variablen**

2.) Die Sprechweise „es existiert ein“ bedeutet stets „es existiert **mindestens** ein“; existiert **ein** $x \in M$ mit $A(x)$, aber **keine zwei** verschiedene in M, so sagen wir: „es existiert **genau ein** $x \in M$ mit $A(x)$“, kurz $\exists_1 x \in M : A(x)$ bzw. $\exists!\ x \in M$ oder $\overset{\bullet}{\bigvee}_{x \in M} : A(x)$.

Eine weitere Aussage mit gebundener Variablen ist die Allgemeingültigkeitsaussage für einen Bereich M (All-Aussage): „Für alle x aus M gilt: $A(x)$ ist wahr“.

Sprech- und Schreibweisen: Für alle x aus M gilt $A(x)$;
f.a. $x \in M : A(x)$;
$A(x)$ f.a. $x \in M$;
$\boxed{\forall x \in M : A(x)}$
$\bigwedge_{x \in M} A(x)$.

Beispiele.

$\forall x \in \mathbb{N} : 4 + x = 6$ (Wahrheitswert f)
$\forall x \in \mathbb{N} : x \geq 0$ (Wahrheitswert w)
$\forall x \in \mathbb{R} : x \geq 0$ (Wahrheitswert f)

Hierbei bezeichnet $\mathbb{R}$ die Menge aller reellen Zahlen.

An dem Beispiel zeigt sich die Wichtigkeit des Grundbereichs M für den Wahrheitswert einer All-Aussage.
Enthält eine Aussageform $A(x, y, ...)$ mehrere freie Variable, so können diese durch entsprechend viele Quantoren gebunden werden; dabei ist jeder Quantor auf die hinter ihm stehende Aussage bezogen; gegebenenfalls verdeutlicht man den Wirkungsbereich durch Klammern.

Beispiel. $\forall x \in \mathbb{N} : \forall y \in \mathbb{N} : x + y = y + x$ (w) .

Außer bei gleichen Quantoren ist unbedingt **auf die Reihenfolge zu achten**:

$$\begin{array}{lllll} & \forall x \in \mathbb{N} : \exists y \in \mathbb{N} : x < y & \text{(w)} & \text{(z.B. } y = x + 1) \\ \text{aber} & \exists y \in \mathbb{N} : \forall x \in \mathbb{N} : x < y & \text{(f)} & \text{(z.B. } x = y + 1 > y \text{ f.a. } y \in \mathbb{N}). \end{array}$$

Jedoch darf man zwei hintereinander stehende Allquantoren vertauschen, also von

$$\forall x \in M_1 : \forall y \in M_2 : A(x, y) \qquad \text{zu} \qquad \forall y \in M_2 : \forall x \in M_2 : A(x, y)$$

übergehen. Entsprechendes gilt für zwei hintereinander stehende Existenzquantoren.

(c) Junktoren

1. Negation einer Aussage $\boxed{\neg A}$, non A.

Beispiele. (i) A: $4 + 3 = 7$ (w)
$\neg A$: $4 + 3 \neq 7$ (f)[7]
(ii) A: $\exists x \in \mathbb{R} : x^2 = -1$ (f)
$\neg A$: $\forall x \in \mathbb{R} : x^2 \neq -1$ (w).

Der Zusammenhang zwischen Wahrheitswert einer Aussage und dem ihrer Negation lässt sich in einer Tabelle (**Wahrheitstafel**) darstellen:

A	$\neg A$
w	f
f	w

.

Umgekehrt präzisiert diese Wahrheitstafel den Begriff „Negation“, zumindest was sein Verhalten bzgl. des Wahrheitswertes angeht.

2. Das logische „und“ $\boxed{A \wedge B}$, A und B

Auch diese logische Verknüpfung beschreiben wir ganz formal durch eine Wahrheitstafel:

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

$A \wedge B$ ist also nur wahr, wenn **sowohl** A **als auch** B wahr sind.

[7] Hierbei ist $a \neq b$ eine Abkürzung für $\neg(a = b)$.

Beispiele. $(4+3=7) \wedge (\forall x \in \mathbb{R} : x^2 \geq 0)$ (w) ,
$(4+3=6) \wedge (\forall x \in \mathbb{R} : x^2 \geq 0)$ (f) .
Auch die Bildung der Aussageform $A(x, y, \dots) \wedge B(x, y, \dots)$ ist sinnvoll (Beispiel: Gleichungssysteme).

3. Das logische „oder“ $\boxed{A \vee B}$, A oder B

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

$A \vee B$ ist also nur wahr, wenn mindestens eine der beiden Aussagen A, B wahr ist. (Dieses „oder “ entspricht dem lateinischen vel, dessen Anfangsbuchstabe zum Zeichen $\vee$ führte.)

Beispiele. $(4+3=7) \vee (\forall x \in \mathbb{R} : x^2 \geq 0)$ (w) ,
$(4+3=6) \vee (\forall x \in \mathbb{R} : x^2 \geq 0)$ (w) .

4. Wenn – dann $\boxed{A \Rightarrow B}$ (Subjunktion)

Sprechweisen: „Wenn A, dann B“.
„A ist hinreichend für B“.
„B ist notwendig für A“.
„Aus A folgt B“; (wobei hier „folgt“ nicht inhaltlich, sondern **rein formal** gemeint ist).

Der Wahrheitswert der Aussage $A \Rightarrow B$ für Aussagen A und B wird definiert durch

A	B	$A \Rightarrow B$	
w	w	w	} aus Wahrem folgt nur Wahres.
w	f	f	
f	w	w	} „Ex falso quodlibet“.
f	f	w	

Wir verwenden das Zeichen „$\Rightarrow$“ auch zwischen Aussageformen

$$A(x, y, \dots) \Rightarrow B(x, y, \dots) \quad (\textbf{Implikation}).$$

Wenn für jede Einsetzung aus dem Grundbereich $A(x, y, \dots) \Rightarrow B(x, y, \dots)$ wahr ist, spricht man von einer **Folgerung**. Aber auch für umgangs- oder metasprachliche Herleitungen benutzt man das Zeichen „$\Rightarrow$“ und spricht ebenfalls von einer Folgerung. (In der mathematischen Praxis unterscheidet man meist nicht zwischen „Subjunktion“, „Implikation“ und „Folgerung“.)

M ↓

Bei der oben-stehenden Wahrheitstafel für $A \Rightarrow B$ sind die beiden letzten Zeilen beachtenswert. Besagen sie doch, dass ein Schluss formal richtig ist, wenn er von einer falschen Aussage ausgeht (Bsp. $9 = 5 \Rightarrow 14 = 14$ (w)). Jedoch ist diese Vereinbarung nicht abwegig. So wird z.B. der umgangssprachliche Satz (Aussageform mit Zeitvariable)

„Wenn es regnet, werden die Straßen nass. “

als gültig anerkannt, unabhängig davon, ob es gerade regnet oder ob die Straßen gerade nass sind. Auch ist es sinnvoll, die Aussage

$$\forall x \in \mathbb{R} : (3x + 1 = 4 \Rightarrow x^2 = 1)$$

als wahr anzuerkennen, obwohl z.B. für $x = -1$ aus einer falschen Aussage eine wahre impliziert wird und für $x = 3$ aus einer falschen Aussage eine falsche impliziert wird.

Damit bleiben die bekannten Umformungsregeln für Gleichungen richtig, auch wenn man von einer falschen Aussage ausgeht:

$$
\begin{aligned}
9 &= 5 && && \text{(f)}\\
9-7 &= 5-7 && &&\\
(9-7)^2 &= (5-7)^2 && &&\\
9^2-2\cdot 9\cdot 7+7^2 &= 5^2-2\cdot 5\cdot 7+7^2 && &&\\
9^2-5^2 &= 2\cdot 7(9-5) && \mid :(9-5) &&\\
9+5 &= 2\cdot 7 && &&\\
14 &= 14 && && \text{(w)}
\end{aligned}
$$

Aus der Tatsache, dass man aus einer Aussage A durch logische Schlüsse (formaler oder inhaltlicher Art) eine wahre Aussage B hergeleitet hat, folgt **nicht** die Richtigkeit der Aussage A. (Man kann heuristisch so vorgehen, muss dann aber die Schlüsse umzukehren versuchen, also zeigen, dass aus B die Aussage A folgt. Auf die Richtung kommt es an!) ↑ M

5. Genau dann – wenn $\boxed{A \Leftrightarrow B}$ (Abkürzung für $(A \Rightarrow B) \wedge (B \Rightarrow A)$)

Sprechweisen: A gilt genau dann, wenn B;
A ist notwendig und hinreichend für B;
A gilt dann und nur dann, wenn B gilt.

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

(Entsprechend definieren wir $A(x, y, \dots) \Leftrightarrow B(x, y, \dots)$, wenn für jede Einsetzung aus dem Grundbereich $A(x_1, y_1, \dots) \Leftrightarrow B(x_1, y_1, \dots)$ wahr ist).

Anmerkung. Um Klammern zu sparen, vereinbart man
„$\Leftrightarrow$“ trennt stärker als „$\Rightarrow, \wedge, \vee, \neg$“,
„$\Rightarrow$“ trennt stärker als „$\wedge, \vee, \neg$“ und
„$\wedge, \vee$“ trennt stärker als „$\neg$“.

(d) Einige allgemeingültige Gesetze für Aussagen (Tautologien)

Folgende Aussagen haben stets den Wahrheitswert (w) :

2.1.1 $\boxed{A \vee \neg A}$ **stets (w)** **Möglichkeit zur Fallunterscheidung**

Beweis. (Formal-logisch mit Hilfe einer Wahrheitstafel.)

A	$\neg A$	$A \vee \neg A$
w	f	w
f	w	w

„tertium non datur“ □

2.1.2 $\boxed{A \wedge \neg A \text{ ist immer falsch.}}$ **Satz vom Widerspruch**

Beweis.

A	$\neg A$	$A \wedge \neg A$
w	f	f
f	w	f

Ähnlich beweist man: □

2.1.3 $\boxed{\neg(\neg A) \Leftrightarrow A}$ **Satz von der doppelten Verneinung**

2.1.4 $\boxed{(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)}$ **Kontraposition**

2.1.5 $\boxed{[(A \Rightarrow B) \wedge (B \Rightarrow C)] \Rightarrow (A \Rightarrow C)}$ **Transitivität**

2.1.6 $\boxed{\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B)}$

Schlussbemerkungen: Die obigen Gesetze lassen sich schon von der Form her als wahr erkennen. Solche formal-logischen Schlüsse sind von den „inhaltlich begründeten“ logischen Schlüssen, zu unterscheiden, für die wir allerdings, wie schon angedeutet, ebenfalls das Zeichen „⇒“ benutzen werden. Tiefer-greifende Untersuchungen bleiben Vorlesungen über Grundlagen der Mathematik vorbehalten.

2.2 Mengen und Elemente

M ↓

Der in der Mathematik verwandte Begriff „Menge“ unterscheidet sich von dem umgangssprachlichen (mehr quantitativen).

Beispiel (Mengen im mathematischen Sinne).
Die Menge der am 1. Oktober 2006 an der FU immatrikulierten Studenten.
Die Menge $\mathbb{N}$ der natürlichen Zahlen.
Die Mengen $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ der rationalen, reellen bzw. komplexen Zahlen.
Die Menge aller zu einem gegebenen Pfeil vektorgleichen Pfeile der Ebene (Vektor).

Wie lässt sich der Begriff „Menge“ präzisieren? Nach Cantor (1845-1918) ist eine Menge eine „Zusammenfassung von bestimmten wohl-unterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.“

Diese Beschreibung ist jedoch zu unpräzise. Uneingeschränkte Erzeugung von Mengen führt zu Widersprüchen (s.u.)[8]. Eine erfolgreiche Präzisierung ist erst im Rahmen der axiomatischen Logik möglich. Wir werden im folgenden naiv vorgehen und nicht definieren, was eine Menge ist, sondern lediglich einige Eigenschaften anführen.

↑ M

Die Objekte, mit denen wir uns beschäftigen, bezeichnen wir durch Symbole, z.B. durch Buchstaben oder Ziffern. Dabei dürfen auch verschiedene Symbole dasselbe Objekt darstellen.

Beispiel. $\frac{2}{3}$ und $\frac{4}{6}$ bezeichnen dieselbe rationale Zahl: $\frac{2}{3} = \frac{4}{6}$. Dabei bedeutet $a = b$, dass die Symbole a und b dasselbe Objekt repräsentieren (Identität, Gleichheit der Objekte).

[8] Eine Analogie zu Problemen der Mengenbildung weist folgendes Beispiel auf: Der Katalog aller Bücher ist selbst ein Buch und müsste in sich erfasst sein; dann gäbe es ein weiteres Buch.

Anmerkungen.

1. Es sind auch die Schreibfiguren $\frac{2}{3}$ und $\frac{4}{6}$ als Objekte denkbar; und diese sind dann verschieden. Die Angabe von Symbolen allein reicht also nicht aus, vielmehr muss jeweils feststehen (ausgesprochen oder unausgesprochen), welche Symbole das selbe Objekt darstellen, für welche Symbole a und b von Objekten also $a = b$ und für welche $a \neq b$ gilt.

2. Von der Gleichheitsbeziehung erwarten wir, dass sie folgende Eigenschaften besitzt:

$$\left.\begin{array}{c} x = x \\ x = y \Leftrightarrow y = x \\ x = y \wedge y = z \Rightarrow x = z \end{array}\right\} \text{für alle Symbole } x,\ y,\ z.$$

3. Es ist zweckmäßig, bestimmte Gesamtheiten von (Einzel-) Objekten als neue Objekte anzusehen; diese neuen Objekte heißen Mengen, die Einzelobjekte, aus denen eine Menge besteht, heißen Elemente der Menge. Wie oben gesagt, ist die Beschreibung jedoch zu unpräzise für eine Definition. Sie soll uns nur als anschaulicher Hintergrund dienen. Wir fahren daher folgendermaßen fort:

   Gewisse Objekte sollen **Mengen** heißen. Zwischen einem Objekt x und einer Menge A kann die Beziehung

   $$x \text{ ist Element von } A,$$

   in Zeichen $\boxed{x \in A}$, bestehen. (Negation: $x \notin A$).

2.2.1 Grundeigenschaft der Element-Beziehung (Extensionalitätsaxiom)

Zwei Mengen A und B sind genau dann gleich, wenn sie dieselben Elemente haben. Für Mengen A und B gilt also:

$$\boxed{A = B \Leftrightarrow \forall x : (x \in A \Leftrightarrow x \in B)}$$

Anmerkung. Diese Eigenschaft präzisiert die Elemente-Beziehung mit Hilfe der Gleichheit, die ja zwischen Objekten schon gegeben ist. Umgekehrt lässt sich mit ihrer Hilfe die Gleichheit von Mengen bei Kenntnis ihrer Elemente nachprüfen.

Folgerung aus (2.2.1): **Eine Menge ist durch Angabe ihrer Elemente vollständig bestimmt** (z.B. durch Aufzählen ihrer Elemente, wobei es auf die Reihenfolge nicht ankommt.)
Schreibweise: $A = \{a, b, c, \dots\}$. *Beispiele.*

- Die Menge aller Primzahlen zwischen 1 und 12: $\{2, 3, 5, 7, 11\}$.
- $\{2, 4, 6, 8\} = \{6, 4, 8, 2\}$.
- Die Menge aller durch 2 teilbaren natürlichen Zahlen ist gleich der Menge aller natürlichen Vielfachen von 2.
- Die Menge aller (natürlichen) Teiler von 8 ist $T_8 = \{1, 2, 4, 8\}$.

Anmerkung (1). Dass umgekehrt durch konkretes Aufzählen von Elementen eine Menge beschrieben wird, ist nicht selbstverständlich, es folgt jedoch aus der folgenden Grundeigenschaft **„Paarmengenaxiom“** und der später geforderten Möglichkeit der Bildung von „Vereinigungsmengen“.

2.2.2 Paarmengenaxiom

Sind a und b Objekte, so gibt es eine Menge C mit

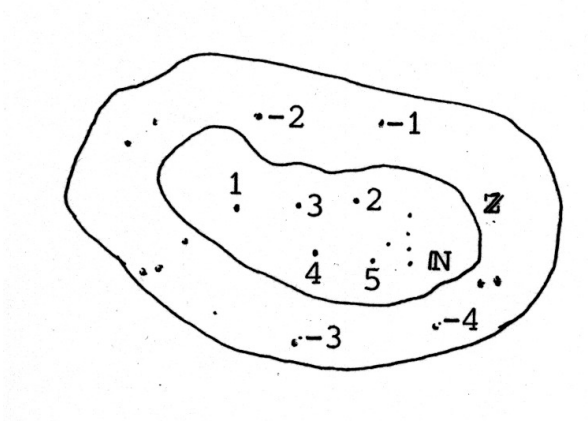
$$\forall d : (d \in C \Leftrightarrow d = a \vee d = b)$$

(Schreibweise: $C = \{a, b\}$)

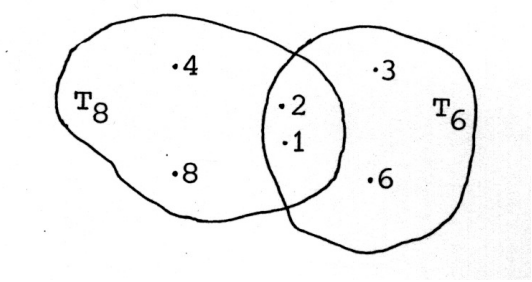
Anmerkung (2). In Übereinstimmung mit der von der Gleichheit von Objekten erwarteten Eigenschaften ergibt sich (aus den Eigenschaften der doppelten Implikation (2.1 c5)) für alle Mengen A, B, C

$A = A$	Reflexivität	(3)
$A = B \Rightarrow B = A$	Symmetrie	(4)
$(A = B \wedge B = C) \Rightarrow A = C$	Transitivität	(5)

Graphische Veranschaulichung von Mengen im sogenannten **Venn-Diagramm**: *Beispiele.*



Figur 2.1 a: Venn-Diagramm zu $\mathbb{N} \subseteq \mathbb{Z}$



Figur 2.1 b: Venn-Diagramm von T_6 und T_8 (mit T_n gleich der Menge der natürlichen Teiler von n).

2.2.3 Definition: Inklusion

Sind A und B Mengen, und ist jedes Element von A Element von B, so heißt

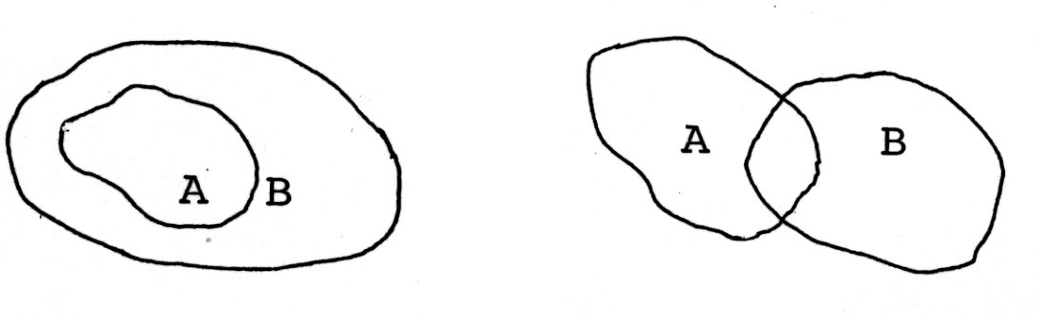
$$A \textbf{ Teilmenge von } B$$

bzw. B **Obermenge** von A; in Zeichen $A \subseteq B$ bzw. $B \supseteq A$. Also:

$$\boxed{A \subseteq B \Leftrightarrow \forall x : (x \in A \Rightarrow x \in B)} \; .$$

Beispiel. $\mathbb{N} \subseteq \mathbb{Z}$

Venn-Diagramme:



$A \subseteq B$ $\qquad\qquad$ $A \nsubseteq B$

Figur 2.2: Beispiele von Venn-Diagrammen zur Teilmengenbeziehung

Ist $A \subseteq B$ und $A \neq B$, so heißt A **echte Teilmenge** von B. Schreibweise: $A \subsetneq B$.

2.2.4 Eigenschaften der Inklusion

Für alle Mengen A, B, C gilt:

$A \subseteq A$	Reflexivität	(6)
$(A \subseteq B \wedge B \subseteq A) \Rightarrow A = B$	Antisymmetrie	(7)
$(A \subseteq B \wedge B \subseteq C) \Rightarrow A \subseteq C$	Transitivität	(8)

Beweis der Antisymmetrie.
Aus $A \subseteq B \Leftrightarrow \forall x : (x \in A \Rightarrow x \in B)$ und $B \subseteq A \Leftrightarrow \forall x : (x \in B \Rightarrow x \in A)$ ergibt sich

$(A \subseteq B \wedge B \subseteq A)$	
$\Leftrightarrow \forall x : (x \in A \Rightarrow x \in B) \wedge \forall x : (x \in B \Rightarrow x \in A)$	Einsetzen
$\Leftrightarrow \forall x : [(x \in A \Rightarrow x \in B) \wedge (x \in B \Rightarrow x \in A)]$	Eigenschaften des Allquantors (hier ohne Beweis)
$\Leftrightarrow \forall x : (x \in A \Leftrightarrow x \in B)$	
$\Leftrightarrow A = B$	(2.2.1) .

Beweisen Sie (6) und (8)! □

Anmerkung. Eine wichtige Methode, die Gleichheit zweier Mengen zu zeigen, besteht nach (2) in dem **Nachweis**, dass jede in der anderen enthalten ist. (Nachweis der **doppelten Inklusion**).

Die Angabe einer Teilmenge C von A erfolgt meist nicht durch Aufzählen der Elemente von C, sondern durch Angabe von Eigenschaften, die allen Elementen von C, aber keinen weiteren Elementen von A zukommen.

Beispiel. Die Menge $\mathbb{P}$ aller natürlichen Primzahlen: Ist $T(x)$ die Aussageform. „x ist Primzahl“, so ist $\mathbb{P}$ die Menge für die gilt:

$$\forall x : [x \in \mathbb{P} \Leftrightarrow (x \in \mathbb{N} \wedge T(x))].$$

Man könnte nun meinen, dass sich mit Hilfe von Aussageformen beliebige Mengen bilden lassen. Für **Teilmengen** wollen wir dies als weitere Grundeigenschaft des Mengen-Begriffes (also einer Eigenschaft, die wir mit dem Begriff „Menge“ verbinden, aber nicht beweisen können) fordern:

2.2.5 Erzeugungsprinzip für Teilmengen (Aussonderungsaxiom)

Sei A eine Menge und $T(x)$ eine Aussageform, zu deren Grundbereich die Elemente von A gehören. **Dann gibt es eine Teilmenge C** derart, dass für alle x gilt:

$$x \in C \Leftrightarrow [x \in A \wedge T(x)].$$

Schreibweisen: $C = \{x : x \in A \wedge T(x)\}$ oder

$$\boxed{C = \{x \in A : T(x)\}}$$

Beispiel. $A = \mathbb{N}$, $T(x) \Leftrightarrow x^2 - x - 2 = 0$
$L_1 = \{x \in \mathbb{N} : x^2 - x - 2 = 0)\}$ (Lösungsmenge der Gleichung $x^2 - x - 2 = 0$ über $\mathbb{N}$)
(*Anmerkung:* Aus der Theorie der quadratischen Gleichungen über $\mathbb{R}$ wissen wir, dass $x^2 - x - 2 = 0$ äquivalent ist mit $x = 2 \vee x = -1$. Wegen $-1 \notin \mathbb{N}$ und $\mathbb{N} \subseteq \mathbb{R}$ gilt dann $L_1 = \{2\}$.)

Denken wir an „unlösbare“ Gleichungen! Nach den Grundeigenschaften der Gleichheitsbeziehung geht z.B. $x \neq x$ (als $T(x)$) bei keiner Einsetzung in eine wahre Aussage über.

2.2.6 Hilfssatz:

(a) *Die Teilmenge* $\boxed{\emptyset_A := \{x \in A : x \neq x\}}$ *von A enthält kein Element. Es gilt also:* $\forall x : x \notin \emptyset_A$.

Ist $F(x)$ irgendeine Aussageform, so gilt für jedes x des Grundbereichs:
$x \in \emptyset_A \Rightarrow F(x)$; denn $x \in \emptyset_A$ ist stets falsch (vgl. (2.1)c4).
Ist B eine zweite Menge, so folgt daher $\forall x : (x \in \emptyset_A \Leftrightarrow x \in \emptyset_B)$. Nach (2.2.1) heißt das:

(b) *Für alle Mengen A, B gilt:* $\boxed{\emptyset_A = \emptyset_B}$.

$\emptyset_A$ ist also unabhängig von A. Es gibt nur eine Menge „ohne Elemente“. Wir nennen sie die **leere Menge**.

Schreibweise: $\emptyset$ $(= \emptyset_A = \emptyset_B)$. Für sie gilt:

(c) $\boxed{\forall x : x \notin \emptyset}$ und

(d) $\boxed{\emptyset \subseteq A}$ für jede Menge A.

Anmerkung: URL zur „leeren Menge“ : www.mathematik.de/fuenfminuten
„13. Die Existenz der leeren Menge macht Sinn.“

Wir haben ein Verfahren kennen gelernt, mit Hilfe von Aussageformen aus einer Menge neue (Teil-) Mengen zu bilden (auszusondern). Die Beschränkung auf Teilmengen ist dabei von Bedeutung; entgegen ersten Vermutungen kann eine Verallgemeinerung zu Widersprüchen führen. Als Beispiel dafür sei die **Antinomie von B. Russel** angeführt: Die Bildung $A = \{x : x \notin x)$ ist nicht erlaubt: Denn die beiden (einzigen) Fälle $A \in A (\Rightarrow A \notin A)$ und $A \notin A (\Rightarrow A \in A)$ führen zu Widersprüchen.

M ↓

Schon im alten Griechenland war bekannt: Wenn sich Aussagen auf sich selbst beziehen können, kommt es eventuell zu Problemen Ein Beispiel ist die Aussage: „ Ich lüge. “ (Warum ?)
Bekannt ist auch das „**Paradoxon des Dorfbarbiers**“. Dieser Barbier hat sich darauf spezialisiert, alle Männer des Dorfes zu rasieren, die sich nicht selbst rasieren, und auch keine anderen. Was ist aber dann mit ihm selbst ?
Siehe auch folgende URL: www.mathematik.de/fuenfminuten

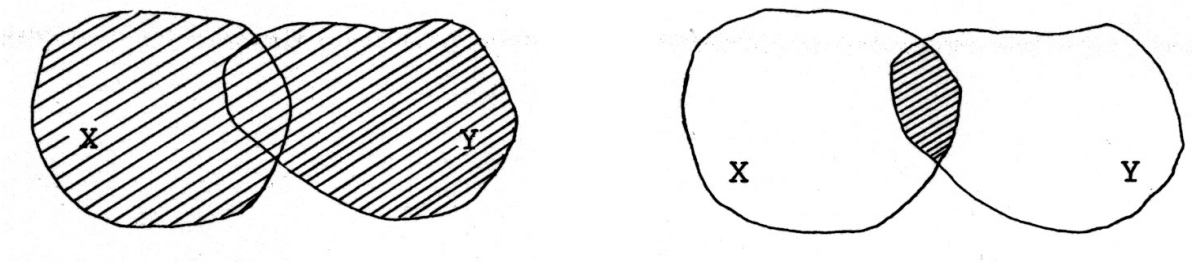
↑ M

2.2.7 Definition der Mengenoperationen $\cup$, $\cap$

Seien X, Y Teilmengen einer Menge M.

(a) $X \cup Y := \{z : z \in X \vee z \in Y\}$ heißt **Vereinigungsmenge** von X und Y.

(b) $X \cap Y = \{z : z \in X \wedge z \in Y\}$ heißt **Schnittmenge** von X und Y (**Durchschnitt**).



$X \cup Y$ $X \cap Y$

Figur 2.3: Venn-Diagramme zu Vereinigungs- und Durchschnitsmenge.

Beispiele.

1) $\left.\begin{array}{l} X = T_8 = \{1,2,4,8\} \\ Y = T_6 = \{1,2,3,6\} \end{array}\right\} \Rightarrow \left\{\begin{array}{lcl} X \cup Y & = & \{1,2,3,4,6,8\} \\ X \cap Y & = & \{1,2\} \end{array}\right.$.

2) Seien $L_1 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x, y \in \mathbb{R} \wedge 2x - 3y = 0 \right\}$ und $L_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x, y \in \mathbb{R} \wedge x + y = 1 \right\}$.

Dann ist $L = L_1 \cap L_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x, y \in \mathbb{R} \wedge 2x - 3y = 0 \wedge x + y = 1 \right\}$ Lösungsmenge des Gleichungssystems

$$\left\{\begin{array}{rcrcl} 2x & - & 3y & = & 0 \\ x & + & y & = & 1. \end{array}\right.$$

Anmerkung: Zur Existenz der Menge $\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x, y \in \mathbb{R} \right\}$ s. (2.2.11) !

3) $\left.\begin{array}{l} X = \mathbb{G} \quad \text{(Menge der geraden Zahlen)} \\ Y = \{1, 3, 9\} \end{array}\right\} \Rightarrow X \cap Y = \emptyset$.

Anmerkung. a) Vereinigung und Durchschnitt von Teilmengen von M existieren aufgrund des Aussonderungsaxioms. Für beliebige Mengen werden wir die Existenz der Vereinigungsmenge später fordern.

b) Falls $X \cap Y = \emptyset$ gilt, nennen wir X und Y **disjunkt**.

2.2.8 Eigenschaften von Vereinigung und Durchschnitt von Teilmengen

Seien X, Y Teilmengen von M. Dann gilt:

(a) Kommutativität:

$\boxed{X \cup Y = Y \cup X}$

Beweis-Andeutung:
Es genügt, $X \cup Y \subseteq Y \cup X$ zu zeigen; (warum?).
$\forall z \in M : (z \in X \cup Y \Rightarrow z \in X \vee z \in Y \Rightarrow$
$z \in Y \vee z \in X \Rightarrow z \in Y \cup X)$ [9]

$\boxed{X \cap Y = Y \cap X}$

Beweis ...

Beweis ...

(b) Assoziativität:

$\boxed{X \cup (Y \cup Z) = (X \cup Y) \cup Z}$

Beweis ...

$\boxed{X \cap (Y \cap Z) = (X \cap Y) \cap Z}$

Beweis ...

(c) Distributivität:

$\boxed{X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)}$

Beweis ...

$\boxed{X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)}$

Beweis ...

2.2.9 Definition und Eigenschaften von Differenz und Komplement

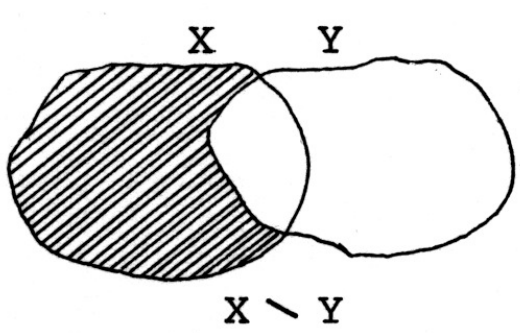
(a) **Definition Differenz**: Für Mengen X, Y definieren wir (siehe Figur 2.4 !):

$X \setminus Y := \{x \in X : x \notin Y\}$.

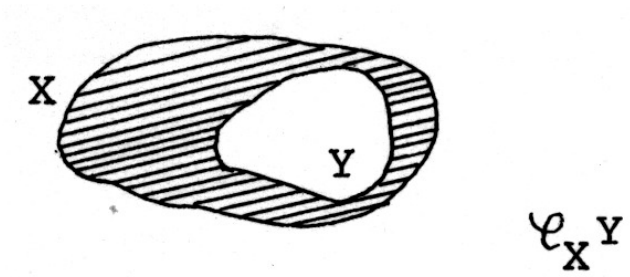
(b) **Definition Komplement** (Spezialfall von (a)):

Im Falle $Y \subseteq X$ heißt $X \setminus Y$ Komplement von Y in X, in Zeichen $\complement_X Y$.

[9] Den Beweis von $A \vee B = B \vee A$ für Aussagen A, B führt man mit Hilfe einer Wahrheitstafel!



Figur 2.4: Venn-Diagramm zur Mengendifferenz



Figur 2.5: Venn-Diagramm zum Mengen-Komplement $\complement_X Y$

(c) Zurückführung der Differenz-Bildung auf die Komplementbildung:
$X \setminus Y = \complement_X(X \cap Y)$.

(d) **Definition Symmetrische Differenz**
Für $A, B \subseteq X$ definiert man:
$A \bigtriangleup B := (A \cup B) \setminus (A \cap B)$.

(e) **Regeln von de Morgan**: Für $Y, Z \subseteq X$ und $\complement = \complement_X$ gilt:

$\complement(\complement Y) = Y$	$Z \subseteq Y \Leftrightarrow \complement Y \subseteq \complement Z$
$\complement(Y \cup Z) = \complement Y \cap \complement Z$	$\complement(Y \cap Z) = \complement Y \cup \complement Z$

Beweis ...

Ohne auf Widersprüche zu stoßen, fordern wir als weitere Grundeigenschaft, dass folgende Bildungen zu Mengen führen:

2.2.10 Bildung der Potenzmenge (Potenzmengenaxiom)

Zu jeder Menge A existiert die Menge[10]

$$\boxed{\wp(A) := \{X : X \text{ Teilmenge von A}\}}$$

[10] Anmerkung: Wir verwenden diese Schreibweise, auch wenn hier nicht aus einer umfassenden Menge ausgesondert wird.

$\wp(A)$ heißt **Potenzmenge** von A.

Beispiele:

(i) $A = \{a\}$ (einelementige Menge)$\Rightarrow \wp(A) = \{\emptyset, A\}$.

(ii) $A = \{a, b\} \Rightarrow \wp(A) = \{\emptyset, \{a\}, \{b\}, A\}$.

Anmerkung. 1.) Stets gilt $\emptyset \in \wp(A)$, $A \in \wp(A)$ sowie $\{x_1\} \in \wp(A)$, falls $x_1 \in A$.
2.) Beachten Sie auch Aufgabe 5.2 !

2.2.11 Bildung eines kartesischen Produktes zweier Mengen

(a) Seien X und Y Mengen. Zu Elementen x und y mit $x \in X$ und $y \in Y$ können wir ein neues Objekt bilden, das

geordnete Paar (x, y).

Näher beschrieben werden die geordneten Paare durch eine Bedingung für die Gleichheit:

$$\boxed{(x, y) = (\bar{x}, \bar{y}) \Leftrightarrow (x = \bar{x} \wedge y = \bar{y})}.$$

Anmerkung: Es ist möglich (x, y) als Menge $\{\{x\}, \{x, y\}\}$ zu definieren.

(b) (Grundeigenschaft: vgl. dazu aber die Anmerkung nach (2.6.1): Es existiert die Menge aller geordneten Paare (x, y) mit $x \in X$ und $y \in Y$; sie heißt „das **kartesische Produkt**“ von X und Y; in Zeichen:

$$\boxed{X \times Y := \{(x, y) : x \in X \wedge y \in Y\}}.$$

Beispiele:

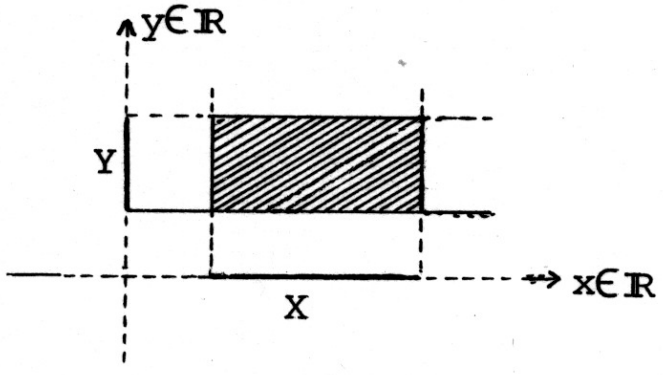
1.) Für $X = \{1, 2\}$ und $Y = \{1, 2, 3\}$ ist

$X \times Y = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\}$.

Tabellarische Darstellung:

$X \setminus Y$	1	2	3
1	$(1,1)$	$(1,2)$	$(1,3)$
2	$(2,1)$	$(2,2)$	$(2,3)$.

2.) Für $X, Y \subseteq \mathbb{R}$ ist $X \times Y \subseteq \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. (S. Figur 2.6 !)



Figur 2.6: Veranschauung eines Beispiels zu $X \times Y$.

Anmerkungen.

a) Bei einer zweielementigen **Menge** kommt es auf die Reihenfolge der Angabe der Elemente nicht an:

$$\{x, y\} = \{y, x\};$$

anders jedoch bei einem **geordneten Paar**:

$$x \neq y \;\Rightarrow\; (x, y) \neq (y, x).$$

b) $(x, y) \in X \times Y \Leftrightarrow (x \in X \wedge y \in Y)$.

c) $X^2 := X \times X$.

d) $X \times Y \times Z := (X \times Y) \times Z$; also $(x, y, z) := ((x, y), z)$.

e) $X^3 := X \times X \times X$.

Anmerkung: Die behandelten Prinzipien orientieren sich an dem Axiomensystem von Zermelo und Fraenkel (ZF).

2.3 Relationen

M ↓

Mit Hilfe von Paaren lassen sich Beziehungen, die zwischen den Elementen zweier Mengen bestehen, formal beschreiben:
Seien z.B. $X = \{a, b, c, d\}$ eine Menge von Personen und
$Y = \{s, t, u\}$ eine Menge von Vereinen.
Die Mitgliedschaftsbeziehung lässt sich dann u.a. in Form einer Tabelle angeben, z.B.:

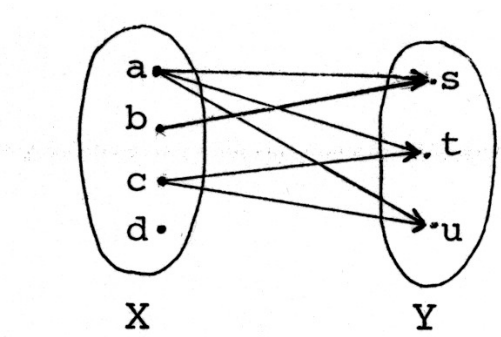
	s	t	u
a	×	×	×
b	×		
c		×	×
d			

(Hier ist also
a Mitglied der Vereine s, t, u
b Mitglied des Vereins s
c Mitglied der Vereine t, u
d Mitglied in keinem dieser Vereine.)

Ein Vergleich mit der tabellarischen Darstellung bei der Bildung des kartesischen Produktes legt auch folgende Beschreibungsform nahe: Angabe der „auftretenden“ Paare, d.h. derjenigen Paare (x, y) in $X \times Y$, für die gilt: x ist Mitglied von y.
Hier erhalten wir die Paare:

$$(a, s), (a, t), (a, u), (b, s), (c, t), (c, u).$$

Der beschriebene Sachverhalt lässt sich auch als „Zuordnung“ interpretieren, was zu Darstellung der Figur 2.7, einem „Pfeil-Diagramm“, führt.



Figur 2.7: Beispiel eines Pfeil-Diagramms

↑ M

2.3.1 Definition: Relation

Seien X, Y Mengen, sei $R \subseteq X \times Y$. Dann heißt (R, X, Y) zweistellige **Relation** zwischen X und Y. Die Menge R heißt **Graph** der Relation.

Anmerkung. Falls keine Verwechslungen zu befürchten sind, sprechen wir auch von R allein als Relation zwischen X und Y (bzw. im Falle $X = Y$ als **Relation auf** X).
Ist $\rho = (R, X, Y)$ Relation, so benutzen wir folgende Schreibweise:

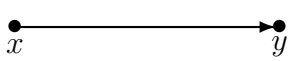
$$\boxed{x \,\rho\, y :\Leftrightarrow (x, y) \in R} \quad \text{für } x \in X, y \in Y.$$

(x steht in der Relation ρ zu y).
Wenn keine Verwechslungen zu befürchten sind, benutzen wir auch die Schreibeise xRy.

Möglichkeiten der Veranschaulichung einer Relation $\rho = (R, X, Y)$ u.a. durch

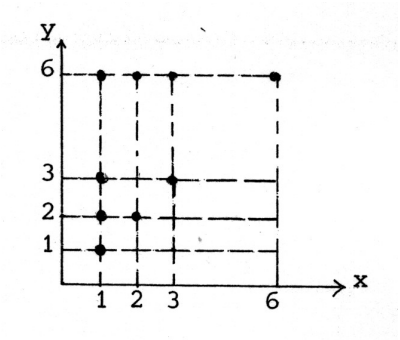
(a) Hervorhebung der Elemente von R im Diagramm von $X \times Y$ (**Darstellung des Graphen** von ρ);

(b) Verdeutlichung der Zuordnung mittels Pfeilen zwischen den Elementen von X und denen von Y in den entsprechenden Venn-Diagrammen (**Pfeil-Diagramm** zu ρ):

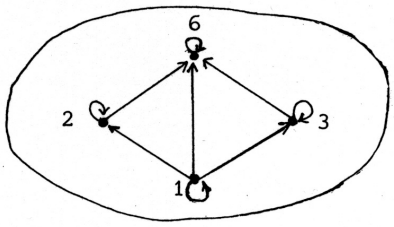


bedeute dabei $x \,\rho\, y$!

Beispiel (1). Seien $X = Y = T_6 := \{x \in \mathbb{N} : x \text{ ist Teiler von } 6\} = \{1, 2, 3, 6\}$ und
$R := \{(x, y) \in T_6 \times T_6 : x \text{ ist Teiler von } y\}$
$= \{(1,1), (1,2), (1,3), (1,6), (2,2), (2,6), (3,3), (3,6), (6,6)\}$;
dann heißt (R, T_6, T_6) Teilbarkeitsrelation auf T_6.



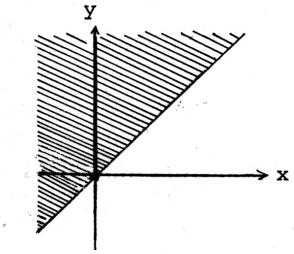
Figur 2.8a: Zu Beispiel (1):Darstellung des Graphen



Figur 2.8 b Pfeil-Diagramm zu Bsp. (1).

(Ähnlich lässt sich eine Teilbarkeitsrelation auf anderen Teilmengen von $\mathbb{N}$ definieren.)

Beispiel (2). Sei $X = Y = \mathbb{R}$ und $R := \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \leq y\}$. Die (hier nicht genau definierte) Relation $\leq = (R, \mathbb{R}, \mathbb{R})$ heißt natürliche Ordnungsrelation auf $\mathbb{R}$.



Figur 2.9: Teildarstellung des Graphen der Relation $\leq$

Das Pfeil-Diagramm ist, auch ausschnittsweise, nicht darstellbar.

Beispiel (3). Seien A eine beliebige Menge und $X = Y = A$; dann heißt

$$\triangle_A := \{(x, y) \in A \times A : x = y\}$$

Diagonale von $A \times A$. (Beachten Sie z.B. die Darstellung in einer Tabelle wie auf Seite 40 ! Welche graphische Darstellung hat $\triangle_{\mathbb{R}}$?)
Die Relation $\mathrm{id}_A := (\triangle_A, A, A)$ heißt Gleichheitsrelation oder auch **Identität** auf A; denn

$$x \ \mathrm{id}_A \ y \Leftrightarrow x = y.$$

Wie sieht das Pfeil-Diagramm von id_A aus?

2.3.2 Definition: Umkehrrelation

Sei $\rho = (R, X, Y)$ Relation. Dann heißt die Relation $\rho^{-1} := (R^{-1}, Y, X)$ mit $R^{-1} := \{(y, x) \in Y \times X : (x, y) \in R)\}$ inverse Relation oder **Umkehrrelation** von ρ. (Andere Schreibweise: ρ^{-}.)

Also:

$$(x,y) \in R \Leftrightarrow (y,x) \in R^{-1} \text{bzw.} \boxed{x\rho y \Leftrightarrow y\rho^{-1}x}.$$

Beispiele

- Umkehr-Relation von „$\leq$“ (siehe Beispiel 2 oben) ist „$\geq$“.
- Umkehr-Relation der Relation „ist Teiler von“ auf $\mathbb{N}$ ist die Relation „ist Vielfaches von“ auf $\mathbb{N}$.
- Umkehr-Relation der Relation „$\subseteq$“ auf $\wp(X)$, ist „$\supseteq$“.
- $(\mathrm{id}_A)^{-1} = id_A$

Bestimmen Sie $(\rho^{-1})^{-1}$!
Im Folgenden sei M eine Menge und $X = Y = M$; unter den Relationen auf M spielen zwei Typen eine fundamentale Rolle: die Äquivalenzrelationen und die Ordnungsrelationen; sie sind gekennzeichnet durch, Gesetzmäßigkeiten, die uns bei Gleichheitsrelationen bzw. bei den Relationen „$\leq$“ auf $\mathbb{R}$ bzw. „$\subseteq$“ auf $\wp(A)$ schon begegnet sind.

2.3.3 Definition: Äquivalenzrelation

Sei M Menge, ρ Relation auf M. Dann heißt ρ Äquivalenzrelation auf M, wenn gilt:

$$\left.\begin{array}{lll} (1) & \text{Reflexivität} & x\rho x \\ (2) & \text{Symmetrie} & x\rho y \Leftrightarrow y\rho x \\ (3) & \text{Transitivität} & (x\rho y \wedge y\rho z) \Rightarrow x\rho z \end{array}\right\} \text{ f.a. } x,y,z \in M.$$

Schreibweisen: Für eine Äquivalenzrelation verwendet man oft die Symbole $\sim$
(also $x \sim y$), oder $\approx$ oder $\equiv$, um damit die Ähnlichkeit zur Gleichheitsrelation zu betonen.

M ↓

Beispiele. id_M ist Äquivalenzrelation auf M.
Die Vektorgleichheit ist Äquivalenzrelation auf der Menge der Pfeile der Ebene.
Die Kongruenz von Figuren der Ebene „ $\overset{\wedge}{=}$ “ ist eine Äquivalenzrelation.

↑ M

Beispiel. Auf $M = \mathbb{Z}$ (Menge der ganzen Zahlen) definieren wir nach Auswahl eines festen $m \in \mathbb{N}$ eine Relation „ $\equiv$ “ durch

$$x \equiv y :\Leftrightarrow \exists t \in \mathbb{Z} : x - y = t \cdot m \quad (\Leftrightarrow m \text{ teilt } (x-y)).$$

$\equiv$ ist dann Äquivalenzrelation auf $\mathbb{Z}$ (nachprüfen!).
Bezeichnung: **Kongruenz modulo m**
Präzisere Schreibweise: $x \equiv y \pmod m$ oder auch nur $x \equiv y \ (m)$.
Anmerkung:
Haben die ganzen Zahlen x und y den gleichen Rest r bei Division durch m, existieren also t_1 und t_2 mit $x = t_1 m + r$ und $y = t_2 m + r$, so folgt ebenfalls $x \equiv y \pmod m$. Umgekehrt haben modulo m kongruente Zahlen den gleichen (nicht-negativen) Rest (kleiner m).

2.3.4 Definition: Äquivalenzklasse

Sei $\sim$ eine Äquivalenzrelation auf M. Dann heißt für jedes $a \in M$ die Menge

$$[a] := \{b \in M : b \sim a\}$$

die Äquivalenzklasse von a, und a ist ein **Repräsentant** von $[a]$.

Beispiele.

(a) Für die Äquivalenzrelation „Kongruenz mod 2 “ gilt mit $\bar{a} := [a]$ z.B.:

$\bar{1} = [1] = \{x \in \mathbb{Z} : x \text{ ungerade } \} = \{\ldots, -3, -1, 1, 3, 5, \ldots\} = -\bar{3} = \bar{3} = -\bar{5} = \bar{5} \ldots$

$\bar{0} = [0] = \{x \in \mathbb{Z} : x \text{ gerade } \} = -\bar{2} = \bar{2} = -\bar{4} = \bar{4} = \ldots$

(b) Äquivalenzrelation Vektorgleichheit: $[\overrightarrow{AB}] =$ Vektor mit Repräsentant $\overrightarrow{AB}$

2.3.5 Satz

Ist $\sim$ eine Äquivalenzrelation auf M, dann gilt für die Menge $M/\sim := \{[a] : a \in M\}$ der Äquivalenzklassen:

(i) $a \in [a]$ für alle $a \in M$;

(ii) $[a] \neq [b] \Rightarrow [a] \cap [b] = \emptyset$ für alle $[a], [b] \in M/\sim$.

Die Menge der Äquivalenzklassen $M/\sim$ bildet eine Partition von M.

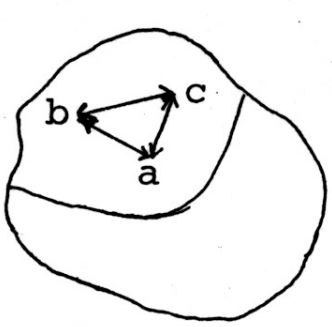
Dabei definieren wir den Begriff Partition folgendermaßen:

2.3.6 Definition

$\mathcal{T} \subseteq \wp(M)$ heißt **Partition** (**Faserung, disjunkte Zerlegung**), wenn gilt:
(1) $\forall x \in M\ \exists T \in \mathcal{T} : x \in T$ (Überdeckungseigenschaft)
(2) $\forall T_1, T_2 \in \mathcal{T} : (T_1 \neq T_2 \Rightarrow T_1 \cap T_2 = \emptyset)$ (paarweise disjunkte Mengen).
Oft fordert man noch:
(3) $\emptyset \notin \mathcal{T}$
Die Elemente von $\mathcal{T}$ heißen **Fasern** oder **Komponenten**.
Beweis von (2.3.5) :

(i) $\forall a \in M : a \sim a \Rightarrow \forall a \in M : a \in [a]$.

(ii) Seien $[a], [b]$ Äquivalenzklassen.
1. Fall: $a \sim b$
Dann $c \in [a] \Rightarrow c \sim a$
$\underset{a \sim b}{\Rightarrow} c \sim b$
$\underset{\text{Transitivität von } \sim}{\Rightarrow} c \in [b]$;
also $[a] \subseteq [b]$.



Figur 2.10: Zum Beweis von (2.3.5)

Aus Symmetriegründen, also wegen $a \sim b \Leftrightarrow b \sim a$, gilt auch $[b] \subseteq [a]$ damit $[a] = [b]$. Die Voraussetzung (ii) aus Satz 2.3.5 ist damit nicht erfüllt.
2. Fall: $a \not\sim b$
Sei $c \in [a] \cap [b]$; dann ist $c \sim a$ und $c \sim b$, wegen der Symmetrie und Transitivität von $\sim$ also $a \sim b$. Widerspruch. Also existiert kein $c \in [a] \cap [b]$, d.h. $[a] \cap [b] = \emptyset$.

□

Anmerkung. Umgekehrt gehört zu jeder Partition $\mathcal{T}$ eine Äquivalenzrelation, deren Äquivalenzklassen gerade die Fasern von $\mathcal{T}$ sind.

Beispiel. Für $M := \mathbb{Z}, m \in \mathbb{N}$ und die Äquivalenzrelation „**Kongruenz modulo m**“ erhält man

$$\mathbb{Z}_m := \mathbb{Z} \,/ \equiv = \{\bar{0}, \bar{1}, \ldots \overline{m-1}\},$$

wobei $\bar{r} = m \cdot \mathbb{Z} + r$ ist. Es gilt auch

$$\mathbb{Z} = \bigcup_{r=0}^{m-1} \bar{r}.$$

Anmerkung: Auf $\mathbb{Z}_m$ kann man eine Addition und eine Multiplikation durch folgende Festsetzung definieren:

$$\overline{r_1} \oplus \overline{r_2} = (r_1 + m\mathbb{Z}) \oplus (r_2 + m\mathbb{Z}) := r_1 + r_2 + m\mathbb{Z} = \overline{r_1 + r_2}$$

bzw.

$$\overline{r_1} \odot \overline{r_2} = (r_1 + m\mathbb{Z}) \odot (r_2 + m\mathbb{Z}) := r_1 \cdot r_2 + m\mathbb{Z} = \overline{r_1 \cdot r_2}.$$

An der gewählten Art der Definition sieht man sofort, dass Summe und Produkt wohldefiniert sind, d.h. dass sie eindeutig bestimmt sind, insbesondere unabhängig von den speziellen Repräsentanten der Äquivalenzklassen.

2.3.7 Definition: Ordnungsrelation

Sei A Menge, ρ Relation auf A. Dann heißt ρ Ordnungsrelation und (A, ρ) **geordnete Menge** (A bzgl. ρ geordnete Menge), wenn für alle $x, y, z \in A$ gilt:

(1)	$x \rho x$	Reflexivität
(2)	$(x \rho y \wedge y \rho x) \Rightarrow x = y$	**Antisymmetrie**
(3)	$(x \rho y \wedge y \rho z) \Rightarrow x \rho z$	Transitivität

(Konsequenzen für das Pfeil-Diagramm einer Ordnungsrelation ?)
Schreibweise: Falls über die Bezeichnung der Relation frei verfügt werden kann, wählt man als Symbol oft „ $\preceq$ “, um so den Charakter der Verallgemeinerung der natürlichen Ordnungsrelation „$\leq$“ auf $\mathbb{R}$ hervorzuheben.

Anmerkung. Manche Autoren sprechen hier auch von einer „teilweisen“ oder „Halb-Ordnung“ und behalten sich „Ordnung“ für eine speziellere Begriffsbildung vor („totale Ordnung“, s.u.).

Beispiele geordneter Mengen. $(\mathbb{R}, \leq)$, $(\wp(X), \subseteq)$, (M, id_M) .
(Beweis der beiden letzten Behauptungen durch Nachprüfen der Gesetze (1) bis (3); die erste Behauptung lässt sich noch nicht nachprüfen, da wir $(\mathbb{R}, \leq)$ nur als intuitiv bekannt voraussetzen.)

Beispiele.

(1) Auf $\mathbb{Z}$ definieren wir eine Relation $\leq_{\mathbb{Z}}$ durch

$$r \leq_{\mathbb{Z}} s :\Leftrightarrow (r = s \vee (\exists n \in \mathbb{N} : r + n = s) \ .$$

$\leq_{\mathbb{Z}}$ ist dann die Relation „kleiner gleich“ auf $\mathbb{Z}$. Mit obiger Definition werden dann, (falls die Eigenschaften der Addition auf $\mathbb{N}$ bekannt sind) die Gesetze (1) bis (3) nachprüfbar.

(2) Auf $\mathbb{N}$ definieren wir eine Relation „ $|$ “ durch

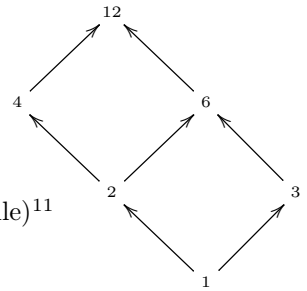
$$(n|m :\Leftrightarrow \exists b \in \mathbb{N} : n \cdot b = m) \text{ f. a. } n, m \in \mathbb{N}.$$

Sie ist die (Ordnungs-) Relation „**ist Teiler von**“ auf $\mathbb{N}$ (Teilbarkeitsrelation).

Anmerkung. Die Teilbarkeitsrelation auf T_k lässt sich dann durch Einschränkung der betrachteten Elemente m, n auf Elemente von T_k aus der Teilbarkeitsrelation auf $\mathbb{N}$ gewinnen.

M
↓

Ist $(M, \leq)$ eine geordnete Menge, so heißt das nicht, dass zwei Elemente $x, y \in M$ „vergleichbar“ sind, dass also $x \leq y$ oder $y \leq x$ gilt. In $(T_{12}, |)$ sind z.B. 4 und 6 nicht vergleichbar.



Figur 2.11: Pfeildiagramm von $(T_{12}, |)$
(ohne Schleifen und ohne Überbrückungspfeile)[11]

[11] Ordnet man, wie in den Figuren 2.11 und 2.12, die Elemente so an, dass die Pfeile von unten nach oben zeigen, so kann man auch die Pfeilrichtungen (also die Spitzen) weglassen. Ein solches Diagramm ohne Schleifen und Brücken heißt **Hasse-Diagramm**.

↑ M

Im Gegensatz dazu lassen sich in $(T_{12}, \leq)$ je zwei Elemente vergleichen. (Hierbei sei $\leq$ die natürlich gegebene $\leq$ Relation, die von $(\mathbb{N}, \leq)$ in T_{12} „induziert“ wird).

Figur 2.12: Pfeil-Diagramm von $(T_{12}, \leq)$ (ohne Schleifen und Brücken)

2.3.8 Definition

Eine Menge M heißt **total geordnet** (**linear geordnet**) und $\leq$ totale Ordnungsrelation, wenn gilt:

(1)	$(M, \leq)$ ist geordnete Menge .
(2)	$\forall x, y \in M : (x \leq y \vee y \leq x)$.

Beispiele. $(\mathbb{Z}, \leq)$ ist total geordnet.
$(\wp(X), \subseteq)$ ist nicht total geordnet, wenn X mindestens 2 verschiedene Elemente enthält.
$(T_n, |)$ ist nur dann total geordnet, wenn $n = p^s$ und p Primzahl ist.

2.4 Abbildungen

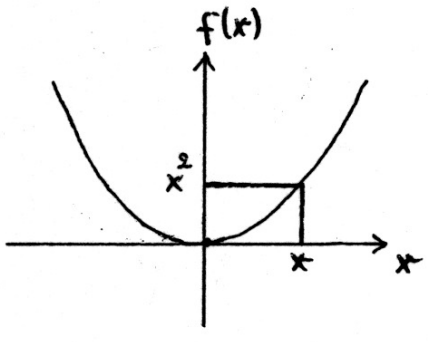
M ↓

Aus der Schulgeometrie sind uns schon viele Abbildungen bekannt, z.B. Kongruenz-Abbildungen (besonders Spiegelungen, Drehungen, Parallelverschiebungen), andere Ähnlichkeitsabbildungen (z.B. Streckungen) und andere affine Abbildungen. Auch funktionale Zusammenhänge lassen sich als Abbildungen beschreiben; so lässt sich $y = f(x)$ (für $x \in D$, dem Definitionsbereich, und $f(x) \in W$, dem Wertebereich oder Zielbereich) als Zuordnungs-Vorschrift interpretieren: „Jedem $x \in D$ wird durch f der Wert $y = f(x)$ zugeordnet.“

Beispiel. $f : y = x^2 \quad (D = W = \mathbb{R})$

Werte-Tabelle von f (Ausschnitt):

x	$f(x) = x^2$
-1	1
$\frac{1}{2}$	$\frac{1}{4}$
0	0
1	1
$\sqrt{2}$	2



Figur 2.13: Darstellung des Graphen von f („Graph“ von f)

Die Punkte des „Graphen“ haben die Koordinaten (x, x^2), allgemein $(x, f(x))$.
Obwohl bei den erwähnten Abbildungen und Funktionen der Zuordnungscharakter stärker im Vordergrund steht, ist es zweckmäßig, sie als spezielle Relationen einzuführen. Dabei sollte der Graph einer solchen Relation

$$G_f = \{(x, y) \in D \times W : y = f(x)\}$$

sein.
Umgekehrt werden wir eine Relation f nur dann Funktion oder Abbildung nennen, wenn es zu jedem $x \in D$ genau ein $y \in W$ gibt mit $(x, y) \in G_f$. (Interpretation im Pfeil-Diagramm?) ↑ M

2.4.1 Definition: linkstotal, rechtseindeutig

Seien X, Y Mengen und $\rho = (R, X, Y)$ eine Relation zwischen X und Y. Dann heißt

(i) ρ **linkstotal**, wenn $\forall x \in X \exists y \in Y : (x, y) \in R$ gilt.

(ii) ρ **rechtseindeutig**, wenn $(\forall x \in X)(\forall y_1, y_2 \in Y) : ((x \rho y_1 \wedge x \rho y_2) \Rightarrow y_1 = y_2)$.

2.4.2 Definition: Abbildung

Eine Relation f zwischen X und Y heißt **Abbildung von X in Y** (synonym: **Funktion** auf X mit Werten in Y), wenn f **linkstotal** und **rechtseindeutig** ist.

Anmerkung. Wir unterscheiden also nicht zwischen „Abbildung“ und „Funktion“.

Wir vereinbaren folgende Terminologie:

(a) X heißt **Definitionsbereich** und Y **Wertebereich** oder **Zielbereich** von f:
$$X = D(f),\ Y = W(f).$$

(b) Für eine Abbildung f gilt: Zu jedem $x \in X$ existiert genau ein $y \in Y$ mit $x\ f\ y$. Wir sagen: Jedem $x \in X$ ist durch f genau ein $y \in Y$ zugeordnet; dieses y bezeichnen wir auch mit $f(x)$.

Die **Zuordnungsvorschrift** geben wir meist in der Form

$$x \mapsto f(x)$$

oder $x \mapsto y$ mit $y = f(x)$ an, manchmal auch allein durch die Funktionsgleichung $y = f(x)$.

(c) $G_f = \{(x, f(x)) : x \in X\}$ heißt **Graph** von f; bei Zahlenmengen lässt, sich der Graph in einem Koordinatensystem darstellen. Auch bei dieser Darstellung spricht man vom Graphen der Abbildung (der Funktion).

(d) Statt $f = (G_f, X, Y)$ schreiben wir bei Abbildung meist $f : X \to Y$

Also: Eine Abbildung (Funktion) mit vorgegebenem Werte-Bereich wird durch drei Angaben festgelegt: Definitions-Bereich, Werte-Bereich, Zuordnungs-Vorschrift:

Schreibweise: $f : \begin{cases} X \to Y \\ x \mapsto f(x) \end{cases}$ oder $f : X \to Y$ mit $x \mapsto f(x)$.

Zwei Abbildungen (mit vorgegebenem Werte-Bereich)

$$f : \begin{cases} X \to Y \\ x \mapsto f(x) \end{cases} \qquad \text{und} \qquad g : \begin{cases} X' \to Y' \\ v \mapsto g(v) \end{cases}$$

sind also genau dann gleich, wenn gilt:

$$X = X' \wedge Y = Y' \wedge \forall x \in X (= X') : f(x) = g(x).$$

Beispiel. Bezeichne $\mathbb{R}_+$ die Menge der positiven reellen Zahlen! Die Abbildungen (mit vorgegebenem Werte-Bereich)

$f_1 : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases}$ und $g_1 : \begin{cases} \mathbb{R} \to \mathbb{R}_+ \\ x \mapsto x^2 \end{cases}$ sind verschieden, da $W(f_1) \neq W(g_1)$ ist.

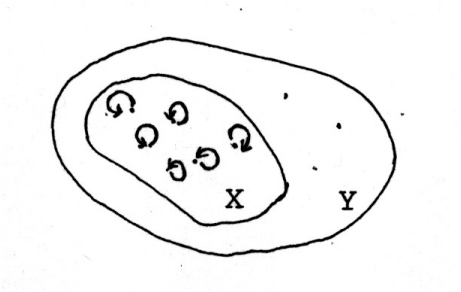
Anmerkung. Oft fasst man den Begriff „Abbildung“ bzw. „Funktion“ weiter, so dass im letzten Beispiel auch f_1 und g_1 als gleich angesehen werden können: **Eine alternative Möglichkeit ist es, eine Funktion f als ihren Graphen $\{(x, f(x)) : x \in X\}$ zu definieren**. (Deshalb sprachen wir von Abbildungen mit vorgegebenem Werte-Bereich.)

Spezialfälle:

2.4.3 Definition:

(a) Seien X, Y Mengen, $y_0 \in Y$. $f_{y_0} : \begin{cases} X \to Y \\ x \mapsto y_0 \end{cases}$ heißt **konstante Abbildung** mit Wert y_0. (Wie sieht das Pfeildiagramm von f_{y_0} aus?)

(b) Seien X, Y Mengen mit $X \subseteq Y$. Dann heißt $j_{X \to Y} : \begin{cases} X \to Y \\ x \mapsto x \end{cases}$ kanonische (natürliche) **Injektion** oder **Einbettung** von X in Y.



Figur 2.14:
Pfeil-Diagramm der kanonischen Injektion $j_{X \to Y}$

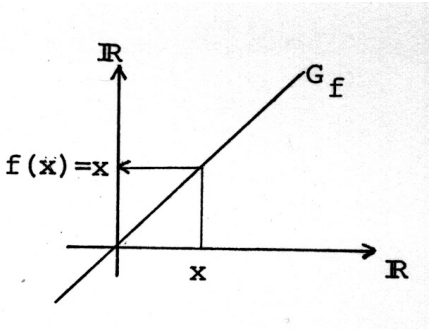
Die Injektion unterscheidet sich nur durch den Wertebereich von folgender Abbildung:

(c) $\mathrm{id}_X = j_{X \to X}$ heißt **Identität**, **identische Abbildung auf** X

Beispiele:

(i) Für $X = \mathbb{R}^2$ (als Punkte-Menge der euklidischen Ebene) ist id_X diejenige Abbildung, die jeden Punkt auf sich abbildet.

(ii) Für $X = \mathbb{R}$ ist $\mathrm{id}_{\mathbb{R}}$ ist die reelle Funktion mit der Funktionsgleichung $y = x$.



Figur 2.15:
Darstellung des Graphen von $\mathrm{id}_{\mathbb{R}}$ (Ausschnitt).

(d) Sei $f : X \to Y$ Abbildung und $A \subseteq X$. Dann heißt $g = f|_A : \begin{cases} A \to Y \\ x \mapsto f(x) \end{cases}$ die **Einschränkung** (oder **Restriktion**) von f auf A.

Umgekehrt heißt f eine **Fortsetzung** von g auf X.

Ist $f(A) := \{f(x) : x \in A\} \subseteq Y' \subseteq Y$, so ist auch $f_{A \to Y'} : \begin{cases} A \to Y' \\ x \mapsto f(x) \end{cases}$ definiert.

Anmerkung. (i) Für $a \in A$ gilt insbesondere $f|_A(a) = f(a)$.

(ii) Die Einschränkung von f auf A ist eindeutig bestimmt, jedoch existieren i.a. mehrere Fortsetzungen von $f|_A$ auf X.

(iii) Die Abbildung $f : \{0,1\} \times \{0,1\} \to \{0,1\}$ mit den Zuordnungen

$$\begin{aligned} (0,0) &\mapsto 0 \\ (0,1) &\mapsto 1 \\ (1,0) &\mapsto 1 \\ (1,1) &\mapsto 0 \end{aligned}$$

steht in engstem Zusammenhang mit der in ((1.5) definierten Addition auf $\{0,1\}$, nämlich $f : (x,y) \mapsto x \oplus y$.

Ähnlich lässt sich auch die Addition auf $\mathbb{R}$ als Abbildung interpretieren (wobei hier allerdings die Zuordnungsvorschrift ohne Kenntnis der Addition nur schwer formulierbar ist):

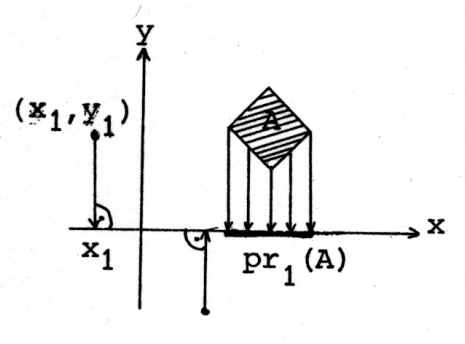
$$\text{„}+\text{“}: \begin{cases} \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ (x,y) \quad \mapsto x + y \end{cases}.$$

Auch die Multiplikation auf $\mathbb{R}$, die Addition von Vektoren,die Addition von Tripeln bzw. die diversen S-Multiplikationen „liefern“ Abbildungen.

(iv) Weitere wichtige Abbildungen von $\mathbb{R} \times \mathbb{R}$ in $\mathbb{R}$ sind die Projektionen

$$\mathrm{pr}_1 : \begin{cases} \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ (x,y) \quad \mapsto x \end{cases} \quad \text{bzw.} \quad \mathrm{pr}_2 : \begin{cases} \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ (x,y) \quad \mapsto y \end{cases}.$$

Veranschaulichung in der Ebene:



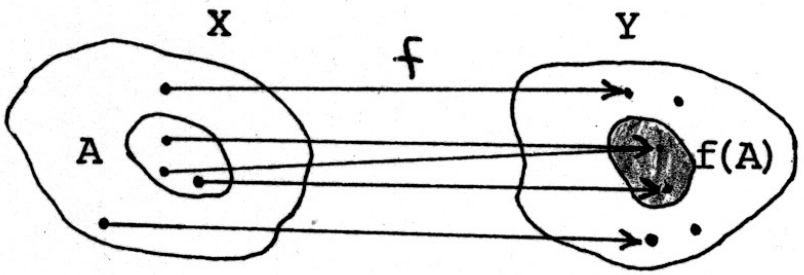
Figur 2.16: Zur Abbildung pr_1.

2.4.4 Definition: Bild, Urbild

Sei $f : X \to Y$ Abbildung !

(a) Für $A \subseteq X$ heißt die Menge $\boxed{f(A) := \{f(x) : x \in A\}}$ **Bild** von A unter f.

Diagramm:



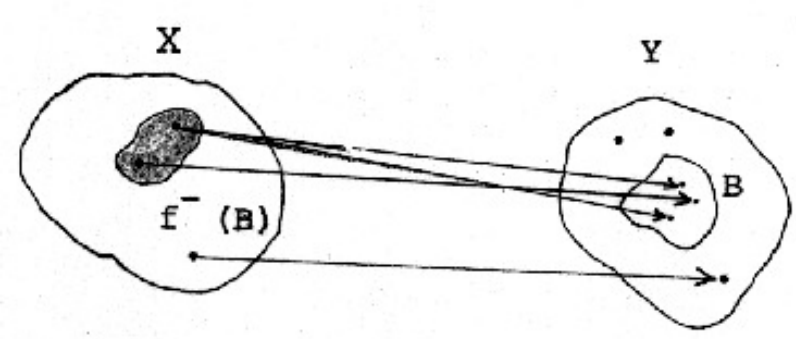
Figur 2.17: Bild $f(A)$ von A unter f.

Anmerkung. $A \subseteq X \Rightarrow f(A) \subseteq f(X) \subseteq Y$.

Beispiel. pr_1 (s.o.!)

(b) Für $B \subseteq Y$ heißt, die Menge $\boxed{f^-(B) := \{x \in X : f(x) \in B\}}$ (volles) **Urbild** von B unter f. (Oft schreibt man auch $f^{-1}(B)$ statt $f^-(B)$.)

Diagramm:



Figur 2.18: (Volles) Urbild $f^-(B)$ von B unter f.

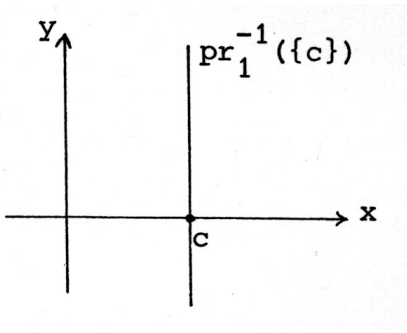
Anmerkung. (i) Urbilder einelementiger Teilmengen brauchen nicht einelementig zu sein:

Beispiel (1).

$$\text{Für } f: \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases} \qquad \text{ist} \quad f^-(\{b\}) = \begin{cases} \{\sqrt{b}, -\sqrt{b}\} & \text{für } b > 0 \\ \{0\} & \text{für } b = 0 \\ \emptyset & \text{für } b < 0\ . \end{cases}$$

Beispiel (2). $\mathrm{pr}_1^-(\{c\}) = \{(x,y) \in \mathbb{R}^2 : x = c\}$

Darstellung in der Ebene: Figur 2.19 : $\mathrm{pr}^-(\{c\})$



M ↓

(ii) Beim Aufsuchen des Urbildes von B im Diagramm verfolgt man die in B endenden Pfeile „rückwärts". Demgemäß erinnert die Schreibweise $f^-(B)$ an die Umkehrrelation f^{-1} zu f. Bei den Bildungen von f^- ordnen wir jedoch Teilmengen von Y solche von X zu.

Wann ist mit f auch f^- Abbildung?

↑ M

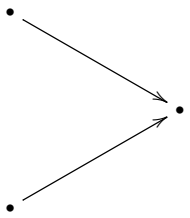
2.4.5 Definition:

Eine Abbildung $f : X \to Y$ (mit vorgegebenen Werte-Bereich Y) heißt

(i) **surjektiv** (rechtstotal), falls $f(X) = Y$ gilt;

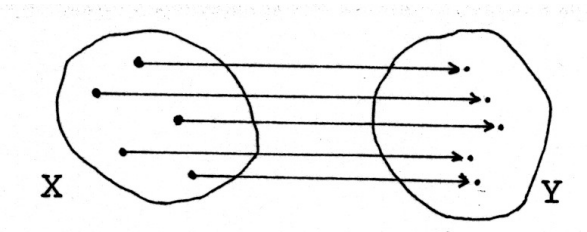
(ii) **injektiv** (linkseindeutig), falls gilt:

$$\forall x_1, x_2 \in X : [f(x_1) = f(x_2) \Rightarrow x_1 = x_2].$$



Figur 2.20: bei injektiven Abbildungen ausgeschlossen

(iii) **bijektiv**, falls sie surjektiv und injektiv ist.



Figur 2.21: Pfeildiagramm einer bijektiven Abbildung

Also: $\boxed{f : X \to Y \text{ bijektiv} \Leftrightarrow \forall y \in Y \exists 1\, x \in X : f(x) = y}$

Als Abbildung erfüllt f natürlich die Bedingung:

$$\forall x \in X\ \exists 1\ y \in Y : f(x) = y.$$

2.4.6 Folgerung für eine Abbildung f :

$$\boxed{f : X \to Y \text{ ist bijektiv.} \Leftrightarrow f^{-1} \text{ ist Abbildung von } Y \text{ nach } X.}$$

(Dabei identifizieren wir die „Sigleton-Menge“ $\{x\}$ mit dem Element x und schreiben f^{-1} statt $f^{-}|_{Y \to X}$, sofern die Einschränkung möglich ist.
Also: Die Umkehrrelation einer bijektiven Abbildung (bij. Abb.) $f : X \to Y$, ist eine Abbildung $f^{-1} : Y \to X$.
Bezeichnung: f^{-1} heißt „die zu f **inverse Abbildung**“.
Eigenschaft: $[f(x) = y \Leftrightarrow x = f^{-1}(y)]$ f.a. $x \in X, y \in Y$.

M
↓

Beispiele.
$f_1 : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x^2 \end{cases}$ ist weder surjektiv noch injektiv. $f_2 : \begin{cases} \mathbb{R} \to \mathbb{R}_+ \\ x \mapsto x^2 \end{cases}$ ist surjektiv, aber nicht injektiv.

$f_3 = f_1|_{\mathbb{R}_+} : \begin{cases} \mathbb{R}_+ \to \mathbb{R} \\ x \mapsto x^2 \end{cases}$ ist nicht surjektiv, aber injektiv.

$f_4 = f_2|_{\mathbb{R}_+} : \begin{cases} \mathbb{R}_+ \to \mathbb{R}_+ \\ x \mapsto x^2 \end{cases}$ ist bijektiv; $f_4^{-1} : \begin{cases} \mathbb{R}_+ \to \mathbb{R}_+ \\ y \quad \mapsto \sqrt{y} \end{cases}$.

↑
M

Frage: Wie sehen die Graphen von f_1, f_2, f_3 und f_4 aus?

M
↓

Häufig ist der Werte-Bereich einer Abbildung wichtiger als die Abbildung selbst. So lässt sich die Indizierung (bzw. Nummerierung) der Elemente einer Menge mit Hilfe des Abbildungs-Begriffes formalisieren. In solchen Fällen benutzt man eine veränderte Notation:

Beispiel. Durch die Abbildung

$$f : \begin{cases} \{1,2,3\} \to \{a,b\} \\ 1 \mapsto a \\ 2 \mapsto b \\ 3 \mapsto a \end{cases}$$

ist eine „Reihenfolge“ der (evtl. zu wiederholenden Elemente) von $Y = \{a, b\}$ festgelegt: Der Abbildung f entspricht das Tripel (a, b, a), ein Element aus Y^3. Umgekehrt lässt sich jedes Element von Y^3 als Abbildung $\{1, 2, 3\} \to Y$ auffassen. Verallgemeinerung:

↑
M

2.4.7 Definition: Familie, Folge

(a) Seien I und Y Mengen und $f : I \to Y$ Abbildung. Statt $f(i)$ für $i \in I$ schreiben wir auch y_i und nennen die Abbildung f, also

$$f : \begin{cases} I \to Y \\ i \mapsto y_i \end{cases} \quad \textbf{Familie} \text{ von Elementen aus } Y \text{ mit } \textbf{Index-Menge } I.$$

Schreibweise: $f = (y_i)_{i \in I}$.

Anmerkung. Dabei muss man die Familie f unterscheiden von der Teilmenge von Y, deren Elemente als Werte von f auftreten, d.h. von der indizierten Menge $\{y_i : i \in I\}$. f braucht ja nicht injektiv zu sein, sodass Elemente von Y mehrfach als Bilder vorkommen können.

(b) Spezialfall: $I = \mathbb{N}$

$(y_n)_{n \in \mathbb{N}}$ heißt **Folge** in Y.

(Auch für $I = \{n \in \mathbb{Z} : n \geq m\}$ mit $m \in \mathbb{Z}$ ist der Begriff der Folge verwendbar.)

Beispiel. $\left(\frac{1}{n}\right)_{n \in \mathbb{N}} : \begin{cases} \mathbb{N} \to \mathbb{R} \\ n \mapsto y_n = \frac{1}{n} \end{cases}$.

(c) **Spezialfall:** $I = \{1, \ldots, n\}$.

$(y_i)_{i \in \{1,\ldots,n\}}$ (auch $(y_i)_{i=1,\ldots,n}$ geschrieben) heißt **endliche Folge** in Y.

Anmerkung. Eine solche Folge kann mit dem entsprechenden n-**Tupel** $(y_1, \ldots, y_n)$ oder dem entsprechenden **Wort** $y_1 \ldots y_n$ identifiziert werden.

Beispiel.

Das Tripel $(x, y, z) \in \mathbb{R}^3$ lässt sich als Folge $\{1, 2, 3\} \to \mathbb{R}$ mit $\begin{cases} 1 \mapsto x \\ 2 \mapsto y \\ 3 \mapsto z \end{cases}$ auffassen.

2.5 Verknüpfung von Abbildungen

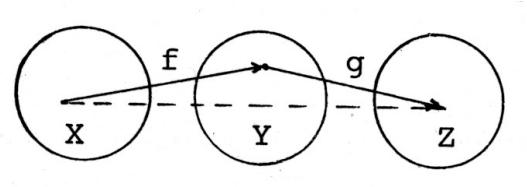
M ↓

Im Folgenden betrachten wir die Verkettung von Abbildungen. Zur Motivation ziehen wir die Hintereinanderausführung geometrischer Bewegungen heran: zwei Drehungen um dasselbe Zentrum in der Ebene z.B. lassen sich, sieht man von den Zwischenzuständen ab, durch eine einzige Drehung ersetzen; die Hintereinanderausführung zweier Spiegelungen lässt sich durch eine Drehung bzw. Parallelverschiebung beschreiben.

↑ M

2.5.1 Produkt (Komposition) von Abbildungen

Voraussetzung: Seien X, Y, Z Mengen und $f : X \to Y$, $g : Y \to Z$ Abbildungen.
Behauptung:
$x \mapsto g(f(x))$ ist Zuordnungsvorschrift einer Abbildung von X in Z. (Beweis: ?)



Figur 2.22: Zur Verknüpfung von Abbildungen

Definition: Diese Abbildung heißt das **Produkt** (Hintereinanderausführung, Verkettung, Verknüpfung, Komposition) der Abbildungen f und g, in Zeichen $g \circ f$.
Sprechweise: „ g nach f “

Also $$\boxed{g \circ f : \begin{cases} X \to Z \\ x \mapsto g(f(x)) \end{cases}}$$

Anmerkung. 1) Für $f, g \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ z.B. ist $f \circ g$ zu unterscheiden von $f \cdot g$ (mit $(f \cdot g)(x) = f(x) \cdot g(x)$).

2a) Man beachte das „Passen“ des Definitionsbereichs von g und des Werte-Bereichs von f.

2b) Eine analoge Definition ist auch für $g : Y' \to Z$ mit $Y \subseteq Y'$ möglich.

3) Oft schreibt man insbesondere in Geometrie und Algebra statt $f(x)$ auch x^f. Bei Benutzung dieser Schreibweise ist es sinnvoll die Faktoren eines Produktes in umgekehrter Reihenfolge zu schreiben also $f \cdot g$ statt $g \circ f$.

4) Die Produkt-Bildung ist bei Abbildungen i.a. nicht **kommutativ**, d.h. falls $g \circ f$ und $f \circ g$ überhaupt erklärt sind, ist $f \circ g = g \circ f$ nicht allgemeingültig.

Beispiel (1). Seien g die Drehung in der Ebene um den Ursprung um 90° gegen den Uhrzeigersinn und f die Spiegelung an der x-Achse! Dann ist
$g \circ f$ die Spiegelung an der Winkelhalbierenden des 1. und 3. Quadranten;
$f \circ g$ die Spiegelung an der Winkelhalbierenden des 2. und 4. Quadranten.

Beispiel (2). $X = Y = Z := \{1, 2, 3\}$; für $f : \begin{cases} X \to X \\ i \mapsto 1 \end{cases}$ und $g : \begin{cases} X \to X \\ j \mapsto 2 \end{cases}$ gilt:
$g \circ f = g \neq f = f \circ g$

Beispiel (3).
$X = Y = Z = \mathbb{R}$; für $f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto x + 1 \end{cases}$ und $g : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto 3x^3 \end{cases}$ gilt: $(g \circ f)(x) = 3(x+1)^3$
und $f(\circ g)(x) = 3x^3 + 1$ (für $x \in \mathbb{R}$).

5) Für jede Abbildung $f : X \to Y$ gilt
$$\boxed{f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f}.$$

6) Sind $f : X \to Y$ Abbildung und $A \subseteq X$, dann ist
$$\boxed{f|_A = f \circ j_{A \to X}}.$$

2.5.2 Assoziativgesetz für Abbildungen

Seien X, Y, Z, W Mengen, $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ Abbildungen[12].
Dann gilt:

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

Beweis. (i) Es gilt $(h \circ g) \circ f : X \to W$ und $h \circ (g \circ f) : X \to W$, also stimmen Definitions- und Werte-Bereich überein.

(ii) $[(h \circ g) \circ f](x) = (h \circ g)(f(x)) = h(g(f(x))) = h((g \circ f)(x)) = [h \circ (g \circ f)](x)$.

□

Weitere Eigenschaften:

2.5.3 Hilfssatz.

(a) (f injektiv $\wedge$ g injektiv) $\Rightarrow$ $g \circ f$ injektiv.

(b) (f surjektiv $\wedge$ g surjektiv) $\Rightarrow$ $g \circ f$ surjektiv.

Umgekehrt gilt (c) $g \circ f$ injektiv $\Rightarrow$ f injektiv; und (d) $g \circ f$ surjektiv $\Rightarrow$ g surjektiv.
Beweis. ...

2.5.4 Folgerung.

Voraussetzung: $f : X \to Y$, $g : Y \to X$ Abbildungen

(a)

$$\boxed{(g \circ f = \mathrm{id}_X \wedge f \circ g = \mathrm{id}_Y) \Leftrightarrow (f \text{ Bijektion } \wedge g = f^{-1})}.$$

(b) Sind f und g Bijektionen, so ist $g \circ f$ Bijektion und es gilt

$$\boxed{(g \circ f)^{-1} = f^{-1} \circ g^{-1}}$$

Beweis....

2.6 Weiteres zu Durchschnitt und Vereinigung

Wir wollen die Durchschnitt- und Vereinigungsmengen-Bildung auf Familien von Mengen verallgemeinern.

[12], also f, g, h Abbildungen mit „ passenden “ Definitions- und Wertebereichen.

2.6.1 Existenz der Vereinigungsmenge (Grundeigenschaft)

Sei $(A_i)_{i\in I}$ eine Familie von Mengen. Dann heißt die Menge

$$\boxed{\bigcup_{i\in I} A_i := \{x : (\exists i \in I : x \in A_i)\}}$$

die **Vereinigung** der A_i. Ihre Existenz fordern wir als weitere Grundeigenschaft.

Speziell: $A_1 \cup A_2 := \bigcup_{i\in\{1,2\}} A_i = \{x : x \in A_1 \vee x \in A_2\}$
(in Übereinstimmung mit der bereits definierten Vereinigung zweier Teilmengen einer Menge.)

Anmerkung. Fordert man zunächst nur die Existenz von $X \cup Y$ (für **beliebige** Mengen X und Y, so lässt sich die Existenz des kartesischen Produktes daraus und aus den anderen Grundeigenschaften herleiten, sodass dann auf die Grundeigenschaft (2.2.11) verzichtet werden kann.

2.6.2 Definition: Durchschnitt einer Familie von Mengen

Sei $(A_i)_{i\in I}$ eine Familie von Mengen mit $I \neq \emptyset$. Dann heißt die Menge

$$\boxed{\bigcap_{i\in I} A_i := \{x : (\forall i \in I : x \in A_i)\}}$$

Durchschnitt der A_i. (Ihre Existenz folgt aus den Grundeigenschaften.)

Speziell: $A_1 \cap A_2 := \bigcap_{i\in\{1,2\}} A_i = \{x : x \in A_1 \wedge x \in A_2\}$ (in Übereinstimmung mit dem bereits definierten Durchschnitt zweier Teilmengen einer Menge).

2.6.3 Verhalten bzgl. Abbildungen

Voraussetzung: Seien X, Y Mengen, $(A_i)_{i\in I}$ eine Familie von Teilmengen von X mit $I \neq \emptyset$, bzw. $A, B \subseteq X$, sowie $f : X \to Y$ Abbildung !

Behauptung:	(a)	$\boxed{A \subseteq B \Rightarrow f(A) \subseteq f(B)}$
	(b)	$f(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} f(A_i)$
speziell:	(b')	$\boxed{f(A \cup B) = f(A) \cup f(B)}$
	(c)	$f(\bigcap_{i\in I} A_i) \subseteq \bigcap_{i\in I} f(A_i)$
speziell:	(c')	$\boxed{f(A \cap B) \subseteq f(A) \cap f(B)}$

Beweis-Andeutung. Exemplarisch zeigen wir (c) (und damit auch (c')):

$$\begin{aligned} y \in f(\bigcap_{i\in I} A_i) &\Rightarrow \exists x : (x \in \bigcap_{i\in I} A_i \wedge y = f(x)) \\ &\Rightarrow \exists x : (\forall i \in I : x \in A_i \wedge y = f(x)) \\ &\Rightarrow \forall i \in I : \exists x \in A_i : y = f(x) \\ &\Rightarrow y \in \bigcap_{i\in I} f(A_i). \end{aligned}$$

□

Anmerkung. In (c) bzw. (c') gilt im Allgemeinen nicht das Gleichheitszeichen:
Beispiel. Für $A = \{(0,0)\} \subseteq \mathbb{R}^2$, $B = \{(0,1)\} \subseteq \mathbb{R}^2$, $f = \mathrm{pr}_1$ gilt:

$$\mathrm{pr}_1(A \cap B) = \mathrm{pr}_1(\emptyset) = \emptyset \quad \text{und} \quad \mathrm{pr}_1(A) \cap \mathrm{pr}_1(B) = \{0\} \neq \emptyset \ .$$

Ist f Bijektion, so gilt das Gleichheitszeichen.

2.7 Zwei Beweis-Prinzipien

M ↓

Im Folgenden behandeln wir zwei Beweis-Prinzipien, die von großer Wichtigkeit sind. Sie erlauben es, Beweise über unendlich viele Aussagen „endlich zu machen".
Wir beginnen zunächst mit dem Prinzip der vollständigen Induktion; es nutzt die Eigenschaften, die wir mit den natürlichen Zahlen verbinden (‚und wird bei deren Definition meist verlangt):

↑ M

2.7.1 Das Prinzip der vollständigen Induktion

Sei $A(n)$ eine Aussageform, deren Grundbereich $\mathbb{N}$ enthält. Gilt dann

(i) $A(1)$ ist wahr (Induktions-Verankerung, Induktions-Beginn)

(ii) $\forall m \in \mathbb{N} : (A(m) \Rightarrow A(m+1))$ (Induktions-Schritt, –Schluss)[13],

so ist $A(n)$ wahr für alle $n \in \mathbb{N}$.

Veranschaulichendes endliches Analogon: Domino-Kette
Anmerkung: Induktionsbeginn ist auch mit $A(0)$ möglich, falls dann $A(m) \Rightarrow A(m+1)$ für alle $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ gezeigt wird.
Zeigen Sie, dass auch $A(-k_0)$ mit $k_0 \in \mathbb{N}$ als (dann negativer) Induktionsbeginn dienen kann!
Eine weitere Variante ist der Beweis von $A(n)$ mit dem Induktionsbeginn $A(1)$ und dem Nachweis von $[\forall x \in \mathbb{N} \text{ mit } x \leq m : A(x)] \Longrightarrow A(m+1)$ für alle $m \in \mathbb{N}$.

Beispiel (1). $A(n)$: $\quad \sum_{i=1}^{n} i^3 = \left(\frac{n(n+1)}{2}\right)^2$

[13] Hierbei heißt $A(m)$ die **Induktions-Voraussetzung** (Induktions-Annahme) und $A(m+1)$ die **Induktions-Behauptung**.

(i) Induktions-Verankerung $n = 1$: $\sum_{i=1}^{1} i^3 = 1 = \left(\frac{n(n+1)}{2}\right)^2$.

(ii) Induktions-Voraussetzung: $A(m)$

Induktions-Schritt:

$$\sum_{i=1}^{m+1} i^3 = \sum_{i=1}^{m} i^3 + (m+1)^3$$
$$= \left(\frac{m(m+1)}{2}\right)^2 + (m+1)^3 \qquad \text{laut Induktions-Voraussetzung}$$
$$= \frac{(m+1)^2(m^2+4m+4)}{4} \qquad = \left(\frac{(m+1)(m+2)}{2}\right)^2, \text{ d.h. } A(m+1).$$

Beispiel (2). Seien $(M, \leq)$ total geordnete Menge und $X = \{x_1, \ldots, x_n\} \subseteq M$ (nichtleere) endliche Teilmenge. Dann existiert genau ein **kleinstes Element** von X , d.h. ein $x \in X$ mit

$$x \leq x_i \text{ für } i = 1, \ldots, n \ .$$

Beweis. Für $n = 1$ ist die Behauptung richtig; sie sei richtig für $n = m$.
Sei $\{x_1, \ldots, x_{m+1}\} \subseteq M$; dann existiert ein $j \in \{1, \ldots, m\} : x_j \leq x_i$ für $i = 1, \ldots, m$ (nach Induktions-Voraussetzung).
Da M total geordnet ist, gilt $x_{m+1} \leq x_j$ oder $x_{m+1} \geq x_j$, woraus die Existenz eines kleinsten Elementes für $\{x_1, \ldots, x_{m+1}\}$ folgt.
Beweis der Eindeutigkeit .... (ohne Induktion möglich). □

Das **zweite Beweis-Prinzip** ermöglicht Existenz-Beweise, wenn die zugrundeliegende Menge geordnet ist und das untersuchte Objekt durch „Maximalität“ charakterisiert ist. Den Wortlaut des Prinzips, das als **Lemma von Zorn** bekannt ist, geben wir weiter unten an. Zuvor vermerken wir, dass es nicht aus den bisher angeführten Grundeigenschaften hergeleitet werden kann, sodass wir seine Gültigkeit zusätzlich fordern, müssen. Oft wird auch statt des Zornschen Lemmas eine zu ihm **äquivalente** Aussage axiomatisch eingeführt, nämlich der
Wohlordnungssatz:
„Jede nicht-leere Menge lässt sich derart ordnen, dass jede nicht-leere Teilmenge T von M ein kleinstes Element besitzt, d.h. ein $x_0 \in T$ mit $x_0 \leq x$ für alle $x \in T$. “
oder das
Auswahlaxiom. „Ist $(A_\alpha)_{\alpha \in I}$ eine nicht-leere Familie nicht-leerer paarweise disjunkter Mengen, dann existiert eine Abbildung $f : I \to \bigcup_{\alpha \in I} A_\alpha$ mit $f(\alpha) \in A_\alpha$ für jedes $\alpha \in I$.“ (Die Abbildung f wählt aus jedem A_α ein Element aus).

Da keine Aussage über die Konstruktion von f gemacht werden kann, lehnen einige Mathematiker das Axiom und Folgerungen daraus ab.

2.7.2 Definition.

Sei $(M, \leq)$ eine geordnete Menge. Dann heißt

(i) $b \in M$ **obere Schranke**[14] von $N \subseteq M$, g.d.w.[15] $\forall a \in N : a \leq b$;

[14] Die obere Schranke von N muss also nicht zur Menge N selbst gehören
[15] genau dann, wenn

(ii) $b \in M$ **maximal** in M, g.d.w. $\forall a \in M : (a \geq b \Rightarrow a = b)$.

(Nicht unbedingt ist jedes Element $x \in M$ kleiner als ein maximales Element von M; denn x kann unvergleichbar mit b sein. Beispiel ? Dementsprechend kann „maximales[16] Element " und „größtes[17] Element " Verschiedenes bedeuten.

(Entsprechend wird „**untere Schranke**" und „**minimal**" definiert. Man beachte den (feinen, aber wesentlichen) Unterschied zwischen den Begriffen „ minimales Element" und „ kleinstes Element".)

Beispiele.

(i) Sei $(M, \leq) = (\mathbb{R}, \leq)$.

Jedes $x \in \mathbb{R}$ mit $x \geq \sqrt{2}$ ist obere Schranke von $\{x \in \mathbb{R} : x^2 < 2\}$.

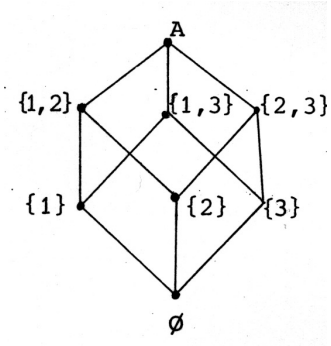
$\mathbb{R}$ besitzt keine maximalen Elemente.

(ii) Seien $(M, \leq) = (\wp(A), \subseteq)$, A Menge, $A \neq \emptyset$.

A ist maximales Element in $\wp(A)$ (bzgl. Inklusion),

$\emptyset$ ist minimales Element in $\wp(A)$ (bzgl. Inklusion).

Speziell: Für $A = \{1, 2, 3\}$ zeigt Figur 2.23 das Hasse-Diagramm (Pfeil-Diagramm ohne Schlingen, ohne Brücken, ohne Pfeilspitzen, Pfeile in aufsteigender Richtung) von $\wp(A)$.
Ist $\mathcal{X} = \{\{1\}, \{1, 2\}\}$, dann sind $\{1, 2\}$ und A obere Schranken von $\mathcal{X}$, und $\{1, 2\}$ ist maximales Element in $\mathcal{X}$.



Figur 2.23: Hasse-Diagramm zu $(\wp(A), \subseteq)$ für $A = \{1, 2, 3\}$.

2.7.3 Definition

Eine geordnete Menge $(M, \leq)$ heißt **induktiv geordnet**, wenn zu jeder nicht-leeren total-geordneten Teilmenge (**Kette**) von M eine obere Schranke in M existiert.

Beispiele.

(i) $(\wp(A), \subseteq)$ ist induktiv geordnet: Für ein nicht-leere Kette $\mathcal{X}$ von $\wp(A)$ ist $\bigcup_{X \in \mathcal{X}} X$ obere Schranke von $\mathcal{X}$.

(ii) $(\mathbb{N}, \leq)$ ist nicht induktiv geordnet: $\mathbb{N}$ ist nicht-leere Kette ohne obere Schranke in $\mathbb{N}$.

(iii) $([0, 1,], \leq)$ ist induktiv geordnet, $([0, 1), \leq)$ nicht [18]

[16] es gibt kein größeres

[17] vergleichbar mit und größer als alle anderen Elemente

[18] Die Menge der Elemente einer monoton gegen 1 konvergierenden Folge von $[0, 1)$ z.B. besitzt keine obere Schranke in $[0, 1)$.

2.7.4 Zornsches Lemma

In jeder nicht-leeren induktiv geordneten Menge existiert (mindestens) ein maximales Element.

Wir geben eine Anwendung des Zornschen Lemmas[19]:

2.7.5 Satz[20]

Seien X, Y nicht-leere Mengen. Dann gilt:

(Es existiert eine injektive Abbildung g von X in Y.)$\vee$ (Es existiert eine injektive Abbildung h von Y in X.)

Beweis. (i) Wir definieren $M := \{(A, f) : A \subseteq X \wedge f : A \to Y \text{ injektiv }\}$. M existiert (warum?). $M \neq \emptyset$, da für $x \in X, y \in Y$ und $f(x) := y$ das Paar $(\{x\}, f)$ aus M ist.

Weiter definieren wir auf M eine Relation „$\leq$“ durch

$$(A, f) \leq (B, g) :\Leftrightarrow A \subseteq B \wedge g|_A = f.$$

(ii) Wir zeigen, dass $(M, \leq)$ induktiv geordnet ist: Es ist „$\leq$“ Ordnungsrelation. Sei U eine nicht-leere total geordnete Teilmenge von M, indiziert durch Elemente von $I : U = \{(A_i, f_i) : i \in I\}$. Wir setzen $A = \bigcup_{i\in I} A_i$ und definieren $f : A \to Y$ durch $f(x) = f_i(x)$ für $x \in A_i$. Wir müssen zeigen, dass f als Abbildung „wohldefiniert“ ist, d.h. dass die durch $G_f = \{(x, y) | \exists i \in I : x \in A_i \wedge y = f_i(x)\}$ definierte Relation (G_f, a, Y) linkstotal und rechtseindeutig ist; die Linkstotalität folgt sofort, die Rechtseindeutigkeit folgendermaßen: seien $x \in A_i$ und $x \in A_k$; zu zeigen ist $f_i(x) = f_k(x)$; da U total geordnet ist, sind (A_i, f_i) und (A_k, f_k) vergleichbar; ohne Beschränkung der Allgemeinheit (o.B.d.A) gelte $(A_i, f_i) \leq (A_k, f_k)$, also $A_i \subseteq A_k$ und $f_k|_{A_i} = f_i$. Damit ist $f : A \to Y$ Abbildung, und es gilt $f|_{A_i} = f_i$ für alle $i \in I$. Wenn wir noch die Injektivität von f zeigen können, haben wir nachgewiesen, dass (A, f) obere Schranke von U ist. Seien also $x_1, x_2 \in A$ und $f(x_1) = f(x_2)$; dann existieren $i, j \in I$ mit $x_1 \in A_i, x_2 \in A_j$ und $f(x_1) = f(x_2)$; o.B.d.A. ist $A_i \subseteq A_j$ (da U total geordnet), also $x_1, x_2 \in A_j$; dann folgt $f_j(x_1) = f_j(x_2)$ und, weil f_j injektiv ist, $x_1 = x_2$.

(iii) $(M, \leq)$ ist also induktiv geordnet. Nach dem Lemma von Zorn existiert dann also ein maximales Element (B, g) in $(M, \leq)$. Wir wollen zeigen, dass $B = X$ oder $g(B) = Y$ gilt. *Annahme*: $B \neq X \wedge g(B) \neq Y$. Dann gibt es $x_0 \in X \setminus B$ und $y_0 \in Y \setminus g(B)$; die Abbildung $h : B \cup \{x_0\} \to Y$ mit $h(x) = \begin{cases} g(x) & \text{für } x \in B \\ y_0 & \text{für } x = x_0 \end{cases}$ ist injektiv. Damit ist $(B \cup \{x_0\}, h)$ eine Element von M, das echt größer ist als (B, g). Widerspruch zur Maximalität von (B, g).

Es gilt also $B = X$ oder $g(B) = Y$. Im ersten Fall ist $g : X \to Y$ eine injektive Abbildung (und die Behauptung ist gezeigt), im zweiten Fall ist $g : B \to Y$ surjektiv und nach

[19] Diese zeigt uns in (2.8), dass von je zwei Mengen eine der beiden die gleiche oder eine kleinere „Mächtigkeit“ als die andere hat.

[20] Falls Sie Schwierigkeiten beim Beweis dieses Satzes haben, die sich nicht sofort beseitigen lassen, so ist das nicht gravierend: Der Beweis ist recht formal und dient hier hauptsächlich als Beispiel zu (2.7.4).

Konstruktion injektiv, also bijektiv; damit existiert $g^{-1} : Y \to B$; mit g^{-1} ist auch $j_{B\to X} \circ g^{-1} : Y \to X$ injektiv, woraus die Behauptung folgt.

□

2.8 Mächtigkeit einer Menge (Kardinalzahlen)

M ↓

Für eine endliche Menge M können wir durch Abzählen der Elemente deren Anzahl n (zumindest theoretisch) feststellen. Das Abzählen entspricht dem „Herstellen" einer Bijektion von M auf $\{1, \ldots, n\}$. Jede andere Menge M' mit n Elementen lässt sich bijektiv auf $\{1, \ldots, n\}$ und damit auf M abbilden. Diese Tatsache nutzt man zur **Verallgemeinerung des Anzahl-Begriffs** auf unendliche Mengen. Man geht folgendermaßen vor:

↑ M

2.8.1 Definition.

Seien A, B Mengen. Existiert dann eine Bijektion von A auf B, so heißen A und B **gleichmächtig**, in Zeichen $A \sim B$.

Beispiele

(i) $\{0, 1, 2, 3\} \sim \{0, 5, 10, 15\}$; denn
$0 \mapsto 0, \quad 1 \mapsto 5, \quad 2 \mapsto 10, \quad 3 \mapsto 15$ liefert eine Bijektion zwischen diesen Mengen.

(ii) Bezeichne $\mathbb{G}$ die Menge der geraden ganzen Zahlen und $\mathbb{U}$ die Menge der ungeraden ganzen Zahlen! Es gilt:
$\mathbb{G} \sim \mathbb{U}$ (Eine mögliche Abbildungs-Vorschrift ist z.B. $x \mapsto x + 1$).

Satz (2.7.5) besagt dann, dass von zwei Mengen X und Y mindestens eine gleichmächtig zu einer Teilmenge der anderen ist.
Die Eigenschaft von Mengen, gleichmächtig zu sein, ist reflexiv ($A \sim A$), symmetrisch ($A \sim B \Leftrightarrow B \sim A$) und transitiv ($(A \sim B \wedge B \sim C) \Rightarrow A \sim C$).

M ↓

Auf jeder gegebenen Menge $\vec{M}$, deren Elemente selbst Mengen sind, (Mengen-System), ist die Gleichmächtigkeit daher eine Äquivalenzrelation und die Quotienten-Menge $\vec{M}/\sim$ besteht aus Klassen gleichmächtiger Mengen. Jedoch trifft man auf die Schwierigkeit, dass $\vec{M}$ evtl. nicht umfassend genug ist und die Menge aller Mengen nicht gebildet werden darf.
Eine Lösung dieses Problems bleibt der Vorlesung über Mengenlehre vorbehalten. Wir geben nur das Resultat wieder:

↑ M

2.8.2

Jeder Menge A lässt sich eine **Mächtigkeit** (**Kardinalzahl**), die wir mit $\mathrm{card}(A)$, $\aleph(A)$ oder auch $|A|$ (Kardinalzahl von A) bezeichnen, derart zuordnen, dass gilt:

(i) $|A| = |B| \Leftrightarrow A \sim B$. (Also: Zwei Mengen haben genau dann die gleiche Mächtigkeit, wenn es eine Bijektion zwischen ihnen gibt.)

(ii) Eine Vergleichsmölichkeit für Kardinalzahlen lässt sich auf folgende Weise (wohl-) definieren:
$|A| \leq |B| :\Rightarrow (\exists C : C \subseteq B \wedge A \sim C)$. Man kann zeigen, dass gilt:

a) $|A| \leq |A|$,

b) $(|A| \leq |B| \wedge |B| \leq |A|) \Rightarrow |A| = |B|$ (der Beweis ist nicht einfach) und

c) $|A| \leq |B| \wedge |B| \leq |C| \Rightarrow |A| \leq |C|$.

Aus Satz (2.7.5) folgt außerdem: $|A| \leq |B| \vee |B| \leq |A|$.

Jede Menge von Kardinalzahlen ist damit total geordnet.

2.8.3 Definition

Sei M eine Menge, $M \neq \emptyset$.

(a) M heißt **endlich**, falls gilt: $\exists n : M \sim \{x \in \mathbb{N} : 1 \leq x \leq n\} =: \mathbb{N}_n$.

(b) M heißt **abzählbar unendlich**, falls $M \sim \mathbb{N}$.

(c) M heißt **überabzählbar**, falls M weder endlich noch abzählbar unendlich ist.

Schreibweisen: (a) $|M| = n$.
(b) $|M| = |\mathbb{N}| = \aleph_0$ (Aleph Null).
(c) $|M| > \aleph_0$.
Ferner $|\emptyset| := 0$.

M
↓

Anmerkung. Die Definition (2.8.3) benutzt die Menge $\mathbb{N}$ der natürlichen Zahlen und Teilmengen von ihr; will man zunächst ohne $\mathbb{N}$ auskommen (und $\mathbb{N}$ mengentheoretisch begründen), so benutzt man die Definition:
Eine Menge M heißt (Dedekind-) **unendlich**, wenn es eine echte Teilmenge U von M gibt mit $M \sim U$, anderenfalls endlich.
Diese und andere Eigenschaften unendlicher Mengen sind ungewohnt.

Beispiel. **Hilberts Hotel**, das Hotel mit unendlich vielen Zimmern.[21]
Ausgangssituation: Gast G_i ist in Zimmer i (für $i \in \mathbb{N}$); alle Zimmer sind belegt.

1. Fall: Herr Z kommt.

Dann zieht Gast G_i in Zimmer $i+1$ und Gast Z in Zimmer 1.

2. Fall: Unendlich viele Gäste $Z_1, Z_2, \ldots$ kommen.

Neue Belegung: Gast G_m zieht in Zimmer $2m$ und

Gast Z_{m+1} in Zimmer $2m+1$ – und alle Gäste sind untergebracht.

Frage: Erkennen Sie die Anwendung auf $\mathbb{N}$ und $\mathbb{Z}$?

↑
M

[21] s. u.a. Funkkolleg Mathematik, Bd. 1, Frankfurt/M., 1971, oder O.Deiser: Einführung in die Mengenlehre, Berlin etc. 2002, oder
URL: http://www.mathematik.de/fuenfminuten/
Ehrhard Behrends: 15. In Hilberts Hotel ist für einzelne Gäste immer ein Zimmer frei (Volle Hotels gibt es im Unendlichen nicht).

2.8.4 Satz.

(a) $\boxed{\mathbb{Z} \text{ ist abzählbar .}}$

Beweis-Andeutung. $f : \mathbb{N} \to \mathbb{Z}$ mit $n \mapsto \begin{cases} \frac{n}{2} & \text{falls } n \text{ gerade} \\ -\frac{n-1}{2} & \text{falls } n \text{ ungerade .} \end{cases}$ □

(b) $\boxed{\mathbb{Q} \text{ ist abzählbar .}}$

Beweis-Andeutung.

(i) $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$ (Abbildungs-Vorschrift $(m, n) \mapsto 2^{m-1}(2n-1)$)

(ii) Jede Teilmenge einer abzählbar unendlichen Menge ist abzählbar unendlich oder endlich.

(iii) $\mathbb{Q}^+ \sim \varphi(\mathbb{Q}^+) \subseteq \mathbb{N} \times \mathbb{N} (\sim \mathbb{N})$, wobei

$$\varphi(x) : \begin{cases} \mathbb{Q}^+ \to \mathbb{N} \times \mathbb{N} \\ \frac{p}{q} \mapsto (p, q) \quad , p, q \text{ teilerfremd.} \end{cases}$$

Zusammenfassung: $\boxed{|\mathbb{Q}| = |\mathbb{Z}| = |\mathbb{N}| =: \aleph_0}$. □

(c) $\boxed{\mathbb{R} \text{ ist überabzählbar.}}$

Beweis-Andeutung. (2. Cantorsches Diagonal-Verfahren)

(i) $\mathbb{R} \sim J = \{x \in \mathbb{R} : 0 < x < 1\}$ (Intervall $(0, 1)$)

Beweis-Andeutung. $f : \begin{cases} J \to \mathbb{R} \\ x \mapsto \frac{2x-1}{1-\lvert 2x-1 \rvert} \end{cases}$ ist Bijektion.

(ebenso $g : \begin{cases} J \to \mathbb{R} \\ x \mapsto \frac{x-\frac{1}{2}}{x(x-1)} \end{cases}$ mit Umkehrfunktion $y \mapsto \frac{1}{2}\left(\frac{y}{1+\lvert y \rvert} + 1\right)$). □

(ii) Annahme: J ist abzählbar; dann existiert eine Bijektion $\mathbb{N} \to J$; sei diese gegeben durch [22]

$$\begin{array}{lll} 1 & \mapsto & 0,\boxed{z_{1\,1}}\ z_{1\,2}\ z_{1\,3}\ \dots \\ 2 & \mapsto & 0,z_{2\,1}\ \boxed{z_{2\,2}}\ z_{2\,3}\ \dots \\ 3 & \mapsto & 0,z_{3\,1}\ z_{3\,2}\ \boxed{z_{3\,3}}\ \dots \\ \vdots & & \end{array}$$

[22] Hierbei seien die Elemente von J als Dezimalbrüche dargestellt und (zum Zwecke der Eindeutigkeit dieser Darstellung) $\dots a\bar{9}$ als $\dots (a+1)\bar{0}$ geschrieben.

Die Zahl $r = 0, z'_{1\,1} z'_{2\,2} \ldots$ mit $z'_{i\,i} := \begin{cases} z_{i\,i} + 1 & \text{für } z_{i\,i} < 8 \\ z_{i\,i} - 1 & \text{für } z_{i\,i} = 8 \text{ oder } z_{i\,i} = 9 \end{cases}$

ist nicht aufgeführt, aber in J, ein Widerspruch!

□

Zusammenfassung:

$$\boxed{\mathbb{R} =: \vec{c} \neq \aleph_0.}$$

2.8.5 Anmerkungen

(i) Man kann sich nun fragen, ob es eine Kardinalzahl gibt, die echt zwischen $\aleph_0$ und $\vec{c}$ liegt. Die sogenannte **Kontinuumshypothese** beinhaltet die Annahme, dass keine solche Kardinalzahl existiert; die verallgemeinerte Kontinuumshypothese geht davon aus, dass es keine weitere Kardinalzahl zwischen $|M|$ und $|\wp(M)| = |2^M|$ gibt.Nach Gödel und Cohen sind sowohl die (verallgemeinerte) Kontinuumshypothese als auch deren Negation zu den übrigen Axiomen der Mengenlehre (Zermelo-Fraenkel + Auswahlaxiom) widerspruchsfrei, also auch nicht aus diesen beweisbar. (Siehe auch: http://de.wikipedia.org/wiki/Kontinuumshypothese)

(ii) Definiert man Summe und Produkt zweier Kardinalzahlen geeignet, (die Summe als die Mächtigkeit der Vereinigung von disjunkten Repräsentanten und das Produkt als Mächtigkeit des kartesischen Produkts zweier Repräsentanten), so lässt sich eine **Kardinalzahl-Arithmetik** entwickeln; z.B. gilt $\aleph_0 + 1 = \aleph_0 = \aleph_0 + \aleph_0$ (vgl. Hilberts Hotel !) und Folgendes:
Ist $(A_\mu)_{\mu \in I}$ eine Familie von Mengen mit $|A_\mu| \leq \aleph_0$ für alle $\mu \in I$; dann folgt

$$|\bigcup_{\mu \in I} A_\mu| \leq \sum_{\mu \in I} |A_\mu| \leq |I| \cdot \aleph_0 = \max(|I|, \aleph_0).$$

(Diese Tatsache werden wir in Satz 8.4 anwenden.)

Literatur-Hinweise zu (2.8).

Deiser, Oliver: Einführung in die Mengenlehre. 19.9.2018
http://www.aleph1.info/?call=Puc& permalink=mengenlehre1

Deiser, Oliver: Einführung in die Mengenlehre. Die Mengenlehre Georg Cantors und ihre Axiomatisierung durch Ernst Zermelo. Springer. Berlin 2002.

Kaplansky, Irwing: Set theory and metric spaces.Nachdruck AMS 2008:
Ausführliche Leseprobe im Netz.

Gebraucht erwerbbar:
dtv-Atlas zur Math., Bd. 1, München 1974, p.34/35
Dugundji, J.: Topology, Boston, Kap. I&II
Fischer-Lexikon Math. 1, Frankfurt/M. 1964 p. 166 ff

oder(andere) Bücher über Mengenlehre.

M
↓

Bei den Beispielen des §1 hatten wir Mengen betrachtet, auf denen eine Addition „+“ erklärt war sowie eine S-Multiplikation mit Elementen aus $\mathbb{R}$ bzw. $\{0,1\}$. Wir hatten gesehen, dass die Rechengesetze bei den aufgeführten Fällen weitgehend übereinstimmten. Im Folgenden wollen wir uns von den „Modellen“ lösen und diese gemeinsamen Eigenschaften (auf höherer Abstraktionsstufe) untersuchen. Damit behandeln wir indirekt auch die Beispiele aus §1.

↑
M

§ 3 Gruppen und Halbgruppen

In allen angeführten Beispielen genügte die „Verknüpfung“ + dem assoziativen Gesetz $(a+b)+c = a+(b+c)$. Wir wollen zunächst Mengen betrachten, auf denen es eine Verknüpfung dieser Eigenschaft gibt – und beginnen so mit einer relativ einfachen algebraischen Struktur.

3.1 Definition Halbgruppe

Sei H eine Menge, $H \neq \emptyset$. Weiter seien folgende „Axiome“ erfüllt.

(G1) Auf H ist eine **innere (binäre) Verknüpfung** definiert, d.h. eine Abbildung (Zuordnung) $* : H \times H \to H$.

Schreibweise: $a * b := *(a,b)$, also $(a,b) \mapsto a * b$.

Anmerkung. Die in §betrachteten Additionen sind Beispiele solcher inneren Verknüpfungen, weiter die Multiplikation auf $\mathbb{R}$ sowie die Multiplikation $\odot$ auf $\{0,1\}$.

(G2) $\forall a,b,c \in H : (a*b)*c = a*(b*c)$ (d.h. $*$ **ist assoziativ**).

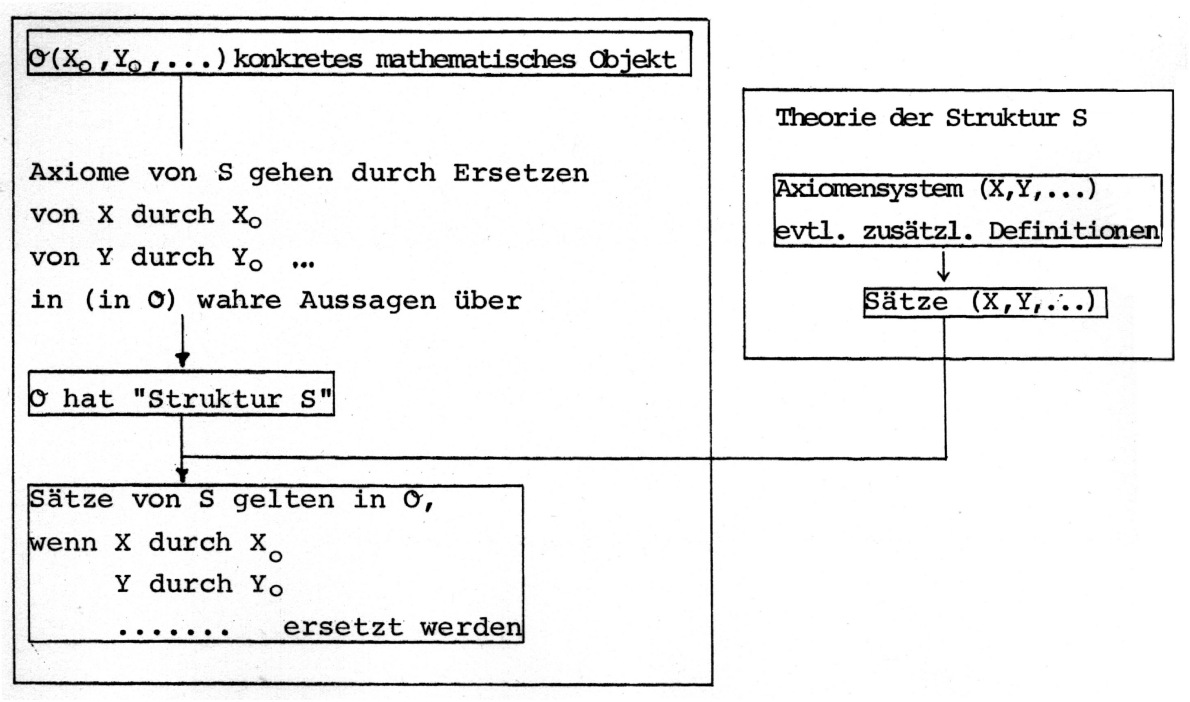
Dann heißt $(H,*)$ eine **Halbgruppe**.
Bemerkung zur Rolle der Axiome: In den „Axiomen“ kommen H und $*$ als „Variable“ vor. Wir dürfen sie in den Anwendungen ersetzen durch jede konkrete Menge H_0 und jede konkrete Verknüpfung $*_0$, für die (G1) und (G2) zu wahren Aussagen werden, also z.B.

- H durch die Menge $\vec{V}$ der Vektoren der Ebene und $*$ durch die auf dieser definierte Addition oder
- H durch $\{0,1\}$ und $*$ durch die auf $\{0,1\}$ erklärte Addition $\oplus$ (s. 1.5 !) oder
- H durch $\{g,u\}$ und $*$ durch die auf $\{g,u\}$ erklärte Addition $\oplus$ (s. 1.5 !) oder
- H durch $\{0,1\}$ und $*$ durch $\odot$ usw..

Jede der aus (G1) und (G2) allgemein hergeleiteten Aussagen gilt unter diesen Voraussetzungen auch für H_0 und $*_0$.

M
↓

Die Anwendungsmöglichkeit einer „axiomatischen Theorie“ auf ein konkretes mathematisches Objekt zeigt das folgende **vereinfachte Schema** (Figur 3.1):



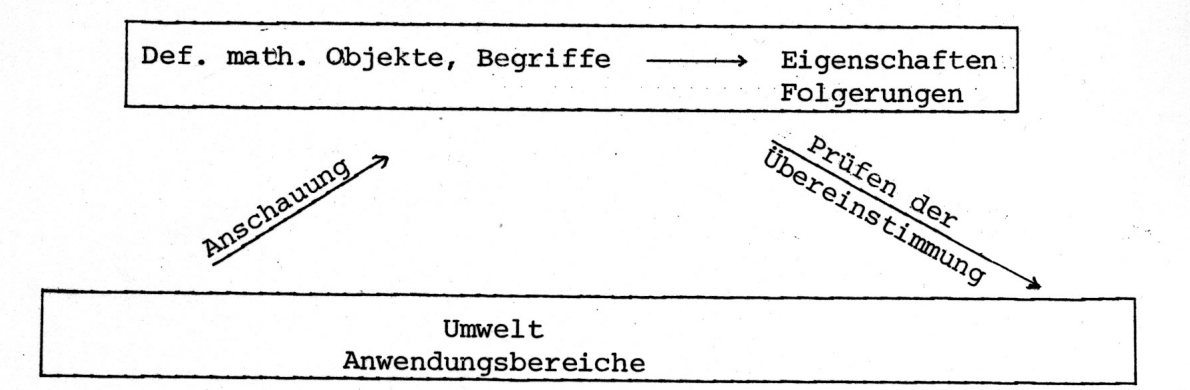
Figur 3.1: Schema für die Anwendung einer axiomatisch definierten Struktur

Beispiel. Hier gehen wir zuerst auf die Halbgruppen-Struktur S ein.

3.2 Exkurs: Mathematik und und die reale Welt

Wir vermerken, dass diese Möglichkeit der „gleichzeitigen“ Untersuchung der Eigenschaften einer Vielzahl mathematischer Objekte nur **eine** Seite des axiomatischen Arbeitens ist.

Eine andere, noch tiefer liegende, steht im engen Zusammenhang mit der Beschreibung unserer Umwelt in mathematischen Modellen. Hierbei werden (undefinierte) Grundbegriffe und Zusammenhänge zwischen diesen aus der Anschauung entnommen und in (unbeweisbaren) Axiomen an den Anfang einer Theorie gestellt. Ist das Axiomensystem geeignet gewählt, so ergeben sich auch weitere Übereinstimmungen mit dem untersuchten Sachverhalt der Umwelt, der damit inner-mathematisch „nachmodelliert“ ist. Ziel ist dann oft der Nachweis, dass bis auf Schreib- und Bezeichnungsweise **nur ein Modell** für dieses Axiomensystem existiert (Bsp.: Hilbert'sches Axiomensystem des Raumes; → Vorlesung Elementargeometrie).



Figur 3.2: Mathematik und die reale Welt

Beispiele.

(i) Elemente von $\mathbb{R}^2 \longleftrightarrow$ Punkte der Anschauungsebene.
(Die Pythagoräer versuchten mit $\mathbb{Q}^2$ „auszukommen“.)

(ii) Elemente von $\vec{V} \longleftrightarrow$ konkrete Parallelverschiebungen der Anschauungsebene.

(Diese Entsprechungen sind nicht beweisbar, da die Objekte der Anschauung entnommen sind).
Ende des Exkurses. ↑ M

Wir kommen zurück zur Halbgruppen-Struktur. Hierfür gibt es viele Beispiele.

M ↓

Anmerkung. Im folgenden benutzen wir bei Beispielen auch einige Eigenschaften der Verknüpfungen $+$ und $\cdot$ von $\mathbb{N}$ (der Menge der natürlichen Zahlen), von $\mathbb{Z}$ (der Menge der ganzen Zahlen), von $\mathbb{Q}$ (der Menge der rationalen Zahlen) und $\mathbb{R}$ (der Menge der reellen Zahlen). Diese **setzen wir als** aus der Schule oder der Vorlesung Analysis **bekannt voraus**. Da sie hier **lediglich bei Beispielen, Motivation, etc.** vorkommen, wird dadurch der exakte Aufbau nicht gestört.

↑ M

Beispiele.

(1) $(\mathbb{N}, +)$, $(\mathbb{N}, \cdot)$, $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R}, \cdot)$ sind Halbgruppen.

(2) $(\mathbb{N}, *_1)$ mit $*_1 : \begin{cases} \mathbb{N} \times \mathbb{N} \to \mathbb{N} \\ (a, b) \mapsto a^b \end{cases}$ ist keine Halbgruppe, da das Assoziativgesetz verletzt ist:[23] $2 *_1 (3 *_1 2) = 2^9 \neq 8^2 = (2 *_1 3) *_1 2$.

(3) $(\mathbb{N}, -)$ mit „ $-$ “ : $\begin{cases} \mathbb{N} \times \mathbb{N} \to \mathbb{N} \\ (a, b) \mapsto a - b \end{cases}$ ist keine Halbgruppe, da „$-$“ keine innere Verknüpfung auf $\mathbb{N}$ ist.

(4) Sei M Menge, $M \neq \emptyset$. Dann sind $(\wp(M), \cap)$, $(\wp(M), \cup)$, $(\wp(M), \bigtriangleup)$ Halbgruppen.Potenzmenge!Operationen auf der P.

(5) $(\vec{V}, +)$ ist Halbgruppe ($\vec{V}$ Menge der Vektoren der Ebene, soweit in (1.1) behandelt.)

(6) Sei M Menge und $\mathrm{Abb}(M, \mathbb{R}) := \{f | f : M \to \mathbb{R} \text{ Abbildung}\}$; (andere Schreibweise: $\mathbb{R}^M$). Für $f, g \in \mathrm{Abb}(M, \mathbb{R})$ definieren wir $\boldsymbol{f + g}$ durch $(f + g)(x) := f(x) + g(x)$ für alle $x \in M$. Dann gilt:

$(\mathrm{Abb}(M, \mathbb{R}), +)$ ist Halbgruppe.

Beweis. ... □

[23] wie auch bei „ Mädchen(handels schule)“ und „(Mädchen handels)schule“ Hinweis von Prof. Lenz, „(Fachwerk)stadt“ und „Fach(werkstadt)“, „(Vogelflug)linie“ und „Vogel(fluglinie), “ „(Dschungelfeuer)löscher“ und „Dschungel(feuerlöscher)“.

Analog kann man $(\mathrm{Abb}(M,\mathbb{R}),\cdot)$ definieren.

(7) Sei M Menge; dann ist $(\mathrm{Abb}(M,M),\circ)$ Halbgruppe.

Beweis. „$\circ$“ ist innere Verknüpfung nach (2.5.1), und das Assoziativgesetz gilt gemäß (2.5.2). □

3.3 Definition neutrales Element

Sei $(H,*)$ Halbgruppe, $e \in H$. Dann heißt e **neutrales Element** von $(H,*)$ g.d.w. gilt:

$$\forall a \in H : e * a = a = a * e \quad .$$

Beispiele.

(a) $(\mathbb{N},+)$ besitzt kein neutrales Element. (Denn: $e + a = a \Rightarrow e = 0; \quad 0 \notin \mathbb{N}$).[24]

0 ist neutrales Element von $(\mathbb{N}_0,+)$, wobei $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

1 ist neutrales Element von $(\mathbb{N},\cdot)$.

(b) M ist neutrales Element von $(\wp(M),\cap)$.

$\emptyset$ ist neutrales Element von $(\wp(M),\cup)$.

$\emptyset$ ist neutrales Element von $(\wp(M),\triangle)$.

(c) $\vec{o}$ ist neutrales Element von $(\vec{V},+)$

(d) Existiert ein neutrales Element von $(\mathrm{Abb}(M,\mathbb{R}),+)$, von $(\mathrm{Abb}(M,\mathbb{R}),\cdot)$?

Anmerkung. Bezeichnet man die Verknüpfung „$*$“ einer Halbgruppe als Addition „$+$“, so nennt man ein evtl. existierendes neutrales Element[25] auch **Nullelement** „0“, ähnlich im Falle der Multiplikation „$\cdot$“ auch **Einselement** „1“.

3.4 Satz (Eindeutigkeit des neutralen Elements)

In einer Halbgruppe gibt es höchstens ein neutrales Element.

Beweis. Seien e_1, e_2 neutrale Elemente der Halbgruppe $(H,*)$. Dann gilt:

$$\begin{aligned} e_1 &= e_1 * e_2 && \text{(da } e_2 \text{ neutrales Element ist)} \\ &= e_2 && \text{(da } e_1 \text{ neutrales Element ist).} \end{aligned}$$

□

[24] Ein Teil der Mathematiker bezeichnen mit $\mathbb{N}$ abweichend von uns die Menge/Halbgruppe der natürlichen Zahlen *einschließlich* der Null; es gibt sogar eine diesbezügliche DIN-Norm.

[25] falls keine Verwechslungen zu befürchten sind

3.5 Definition Inverse

Sei $(H, *)$ Halbgruppe mit neutralem Element e ; sei $a \in H$! Dann heißt a **invertierbar**, wenn gilt:

$$\exists \bar{a} \in H : a * \bar{a} = e = \bar{a} * a \quad .$$

$\bar{a}$ heißt dann **Inverse** zu a.

Beispiel. (i) In $(\mathbb{R}, +)$ ist $-r$ Inverse von r (für beliebiges $r \in \mathbb{R}$); in $(\mathbb{R}, \cdot)$ ist nur 0 nicht invertierbar; in $(\mathbb{N}_0, +)$ besitzt allein 0 eine Inverse (nämlich 0 selbst).

(ii) In $(\wp(M), \cap)$ hat kein Element $T \neq M$ eine Inverse: Sei $T \subseteq M$; dann existiert $\bar{T} \subseteq M$ mit $T \cap \bar{T} = M$ nur, wenn $T = M = \bar{T}$ gilt.

(iii) In $(\vec{V}, +)$ ist $-\vec{a}$ Inverse von $\vec{a}$.

Schreibweise der Inversen[26]: a^{-1} , falls $*$ als Multiplikation,
$-a$, falls $*$ als Addition geschrieben wird.

3.6 Satz (Zur Eindeutigkeit der Inversen)

Sei $(H, *)$ Halbgruppe mit neutralem Element e. Besitzt dann $a \in H$ ein inverses Element, so ist dieses eindeutig bestimmt.

Beweis. Seien $\bar{a}_1, \bar{a}_2$ Inverse von a, gelte also insbesondere $\bar{a}_1, \bar{a}_2 \in H$, $a * \bar{a}_2 = e$ und $\bar{a}_1 * a = e$. Dann folgt

$$\begin{aligned}
\bar{a}_1 &= \bar{a}_1 * e && (e \text{ ist neutrales Element der Halbgruppe}) \\
&= \bar{a}_1 * (a * \bar{a}_2) && (\text{nach Voraussetzung}) \\
&= (\bar{a}_1 * a) * \bar{a}_2 && (\text{nach dem Assoziativgesetz}) \\
&= e * \bar{a}_2 && (\text{laut Voraussetzung}) \\
&= \bar{a}_2 && (e \text{ ist neutrales Element der Halbgruppe}).
\end{aligned}$$

□

(iv) Wir behandeln nun den Spezialfall von Halbgruppen, in denen jedes Element invertierbar ist:

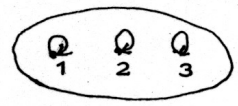
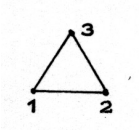
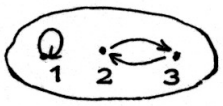
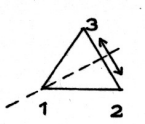
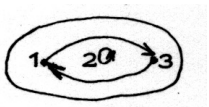
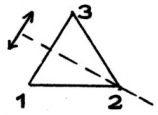
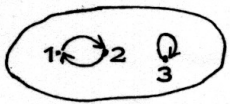
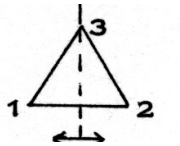
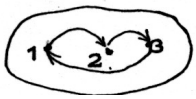
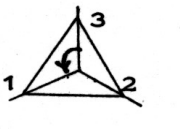
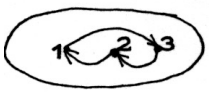
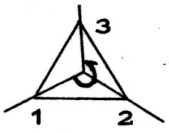
3.7 Definition Permutation

Sei M n.l. Menge. Eine bijektive Abbildung von M auf sich heißt auch **Permutation**; wir definieren $\vec{S}_M := \mathrm{Bij}(M, M) := \{f : f \text{ Permutation von } M\}$ und $\vec{S}_n := \vec{S}_{\{1,2,\ldots,n\}}$.

Schreibweise: Ist $f \in \vec{S}_n$, so schreiben wir auch $f = \begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}$.

Beispiel. Sei $M = \{1, 2, 3\}$; in Figur 3.3 betrachten wir $\vec{S}_3$. Man kann zeigen, dass $\vec{S}_3 = \{\mathrm{id}_M, \sigma_1, \sigma_2, \sigma_3, \delta_1, \delta_2\}$ ist.

[26] falls keine Verwechslungen zu befürchten sind

Elemente von $\vec{S}_3$:	Pfeildiagramm	Anwendung: Symmetrien eines gleichseitigen Dreiecks
$\delta_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \mathrm{id}_M$		
$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$		
$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$		
$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$		
$\delta_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$		
$\delta_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$		

Figur 3.3: Elemente von $\vec{S}_3$ und die entsprechenden Symmetrien eines Dreiecks.

3.8 Hilfssatz (S_M)

Sei M Menge, $M \neq \emptyset$. Dann gilt:

(a) $(\vec{S}_M, \circ)$ ist Halbgruppe.

(b) id_M ist neutrales Element von $(\vec{S}_M, \circ)$.

(c) Zu $f \in \vec{S}_M$ ist f^{-1} Inverse in $(\vec{S}_M, \circ)$.

(d) Für $|M| \geq 3$ ist $(\vec{S}_M, \circ)$ nicht kommutativ[27].

[27] Vgl. Definition (3.9)!

3.9 Definition kommutative/abelsche Halbgruppe

Eine Halbgruppe $(H, *)$ heißt **kommutativ** (**abelsch**), wenn für alle $a, b \in H$ gilt:

$$a * b = b * a.$$

Beweis von Hilfssatz (3.8). (a) „$\circ$“ ist innere Verknüpfung nach (2.5.4)b); das Assoziativgesetz gilt nach (2.5.2).

(b) Beispiel 5 nach (2.5.1)

(c) f^{-1} ist Abbildung nach (2.4.6), Inverse von f nach (2.5.4)a), Bijektion ebenfalls nach (2.5.4)a).

(d) Seien o.B.d.A. $\{1, 2, 3\} \subseteq M$; die Fortsetzungen von σ_1 und δ_1 (s. Bsp. oben) auf M kommutieren nicht.

Z.B. ist $\sigma_1 \circ \delta_1(1) = \sigma_1(2) = 3 \neq 2 = \delta_1(1) = \delta_1 \circ \sigma_1(1)$.

□

3.10 Definition Gruppe

Sei G eine Menge und $*$ eine Verknüpfung auf G; sei ferner $e \in G$ ein fest gewähltes Element. Weiter seien folgende Axiome erfüllt;

(G1) $*$ ist innere Verknüpfung, also $* : \begin{cases} G \times G \to G \\ (a, b) \mapsto a * b \end{cases}$;

(G2) $\forall a, b, c \in G : \ a * (b * c) = (a * b) * c$ (Assoziativgesetz) ;

(G3) $\forall a \in G : e * a = a = a * e$ (Existenz eines neutralen Elements);

(G4) $\forall a \in G \ \exists \bar{a} : a * \bar{a} = e = \bar{a} * a$ (Existenz aller Inversen).

Dann heißt $(G, *)$ **Gruppe mit neutralem Element** $\boldsymbol{e}$.
Anmerkung:
$(G, *)$ Gruppe $\Rightarrow (G, *)$ Halbgruppe mit neutralem Element e (nach (1), (2)).

Anmerkung (zur Schreibweise). Die (abstrakt definierte) Verknüpfung $*$ wird (wie oben schon angedeutet) oft als Multiplikation, im Fall einer kommutativen Gruppe auch als Addition geschrieben. Falls keine Verwechslungen (zwischen Menge und Gruppe) zu befürchten sind und klar ist, welche Verknüpfung gemeint ist, schreibt man oft nur G statt $(G, *)$.

3.11 Elementare Eigenschaften von Gruppen

Sei $(G, *)$ Gruppe. Dann gilt:

(a) In $(G, *)$ existiert genau ein neutrales Element.

(b) In $(G, *)$ hat jedes Element $a \in G$ genau eine Inverse
(Bezeichnung a^{-1} bzw. $-a$ im Falle additiver Schreibweise).

Beweis. Nach (G3) existiert mindestens ein neutrales Element, nach (3.3) ist dieses eindeutig bestimmt. Nach (G4) existiert zu jedem Element ein inverses Element; nach (3.6) ist die Inverse jeweils eindeutig bestimmt. □

Beispiele:

3.12 Korollar zu (3.8): S_M als Gruppe

Ist M n.l. Menge, so gilt:

$(\vec{S}_M, \circ)$ ist eine Gruppe.

Sie heißt symmetrische Gruppe auf M; ist speziell $M = \{1, \ldots, n\}$, so heißt sie **symmetrische Gruppe vom Grad n**.

M
↓

Beispiele weiterer Gruppen. $(\mathbb{R}, +)$, $(\mathbb{R}^*, \cdot)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}^*, \cdot)$, $(\mathbb{Z}, +)$, $(\wp(M), \triangle)$, $(\vec{V}, +)$, $(\{0, 1\}, \oplus)$, $(\{1\}, \odot)$, $(\mathrm{Abb}(M, \mathbb{R}), +)$ und
'die' Menge der Drehungen eines regelmäßigen n-Ecks bzgl. Hintereinanderausführung als Verknüpfung.[28]

Bei den eben aufgeführten Beispielen stehen einige in engem Zusammenhang, z.B. $(\mathbb{R}, +)$ und $(\mathbb{Q}, +)$; es ist ja $\mathbb{Q} \subseteq \mathbb{R}$ und „die Addition auf $\mathbb{Q}$" (, die wir kurzfristig mit „$+_{\mathbb{Q}}$" bezeichnen,) stimmt mit der auf $\mathbb{R}$ („$+_{\mathbb{R}}$") überein". Es gilt also für $q_1, q_2 \in \mathbb{Q}$

$$q_1 +_{\mathbb{Q}} q_2 = q_1 +_{\mathbb{R}} q_2 \ .$$

↑
M

3.13 Definition induzierte Verknüpfung, Untergruppe

Seien $(G, *)$ Gruppe (insbesondere $* : G \times G \to G$) und $U \subseteq G$.

(a) Ist dann $*(U \times U) \subseteq U$, so heißt U **abgeschlossen** bzgl. der Verknüpfung $*$; und

$$*_U : \begin{cases} U \times U \to U \\ (u, v) \mapsto u * v \end{cases}$$

heißt die von G auf U **induzierte (innere) Verknüpfung**.

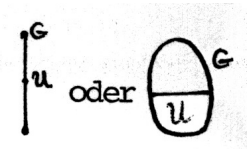
[28] Zu jedem regelmäßigen n-Eck gibt es eine solche Gruppe; nur ist es so, dass sich (für fesates n) diese Gruppen „in ihrer Struktur nicht wesentlich unterscheiden".

Anmerkung. $*_U$ entsteht dabei aus $*$ durch Einschränkung des Definitions-Bereichs und des Werte-Bereichs auf $U \times U$ bzw. U.

Beispiele: $\mathbb{N}$ und $\mathbb{Z}$ sind abgeschlossen in $\mathbb{Q}$ bzgl. der Addition; $\mathbb{N}$ und $\mathbb{Z}^*$ sind abgeschlossen in $(\mathbb{Q}^*, \cdot)$.

Ist zusätzlich $(U, *_U)$ Gruppe, so heißt $(U, *_U)$ **Untergruppe** von $(G, *)$ (abgekürzt „ U UG von G “), in Zeichen $U \leq G$.

(b) *Anmerkung.* Ist G **endlich**, so folgt schon aus der Abgeschlossenheit von U bzgl. $*$, dass $(U, *_U)$ Untergruppe ist. (Beweis?)
Ein weiteres Untergruppen-Kriterium ist in (3.16) zitiert (s.u.). Sind keine Verwechslungen zubefürchten, so schreiben, wir statt $*_U$ auch nur $*$.



Figur 3.5: Symbolische Darstellung der Untergruppen-Beziehung

Beispiele.

$(\mathbb{Z}, +)$ ist Untergruppe (UG) von $(\mathbb{Q}, +)$;

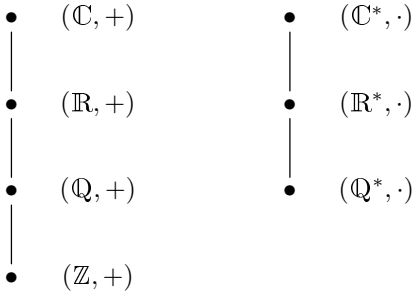
$(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ sind UG'n von $(\mathbb{R}, +)$;

$(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ sind UG'n von $(\mathbb{C}, +)$;

$(\mathbb{Q}^*, \cdot)$ ist UG von $(\mathbb{R}^*, \cdot)$;

$(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$ sind UG'n von $(\mathbb{C}^*, \cdot)$.

M ↓ ↑ M



Figur 3.6: Diagramm für die additiven bzw. multiplikativen Gruppen einiger Zahlbereiche

3.14 Hilfssatz: neutrales Element, Inversen einer Untergruppe

[29]

Seien $(G, \cdot)$ Gruppe, U Untergruppe von G und $a \in U$. Dann gilt:

(i) e ist neutrales Element von G. $\iff$ e ist neutrales Element von U.

(ii) a^{-1} ist Inverse von a in G. $\iff$ a^{-1} ist Inverse von a in U.

Beweis. (i) Seien e_U bzw. e_G die neutralen Elemente von U bzw. G, dann gilt:

$$\begin{aligned} e_U \cdot e_G &= e_U && \text{(da } e_G \text{ neutrales Element von } G \text{ ist)} \\ &= e_U \cdot e_U && \text{(da } e_U \text{ neutrales Element von } U \text{ ist)} \end{aligned}$$

[29] Bei abstrakt gegebenen Gruppen schreiben wir im Folgenden fast immer die Verknüpfung als Multiplikation, also mit oder ohne Malpunkt.

und damit $e_U \cdot e_U = e_U \cdot e_G$, also $e_U^{-1} \cdot (e_U \cdot e_U) = e_U^{-1} \cdot (e_U \cdot e_G)$, mit der Inversen e_U^{-1} von e_U in G, woraus $e_G \cdot e_U = e_G \cdot e_G$ und somit $e_U = e_G$ folgt.

(ii) Sei a_G^{-1} Inverse von a in G, a_U^{-1} von a in U. Dann ergibt sich

$$a_U^{-1} \cdot a = a \cdot a_U^{-1} \underset{\text{Def.}}{=} e_U \underset{(i)}{=} e_G \underset{(3.11b)}{\Longrightarrow} a_G^{-1} = a_U^{-1}.$$

□

3.15 Definition: Komplexschreibweise

Seien $(G, \cdot)$ Gruppe, $A, B \subseteq G$ und $a_0, b_0 \in G$. Dann vereinbaren wir folgende abkürzende Schreibweise („Komplexschreibweise“):

$$\begin{aligned} A \cdot B &:= \{a \cdot b : a \in A \wedge b \in B\} \\ a_0 \cdot B &:= \{a_0 \cdot b : b \in B) = \{a_0\} \cdot B \\ A \cdot b_0 &:= A \cdot \{b_0\} \\ A^{-1} &:= \{a^{-1} : a \in A\} \end{aligned}$$

Damit können wir leicht hinreichende (und notwendige) Bedingungen dafür formulieren, dass U eine UG von G bildet:

3.16 Satz (Untergruppenkriterium)

Sei $(G, \cdot)$ eine Gruppe und $U \subseteq G$. Dann gilt

$$U \text{ ist Untergruppe von } G. \iff U \neq \emptyset \wedge U \cdot U^{-1} \subseteq U.$$

(Machen Sie sich klar, wo in der Formulierung von (3.16) G und U die Gruppe $(G, \cdot)$ bzw. Untergruppe $(U, \cdot_U)$ bezeichnen und wo die diesen zugrunde liegenden Mengen!)

Beweis. „$\Rightarrow$“ U UG $\Longrightarrow \exists e \in U \Longrightarrow U \neq \emptyset$

U UG $\underset{(3.13),\ (3.10),\ (3.14)(ii)}{\Longrightarrow} U^{-1} \subseteq U \underset{(3.13),\ (G1)}{\Longrightarrow} UU^{-1} \subseteq U$.

„$\Leftarrow$“ $(G3)$ $U \neq \emptyset \Rightarrow \exists a \in U \Rightarrow \exists a^{-1} \in G \underset{(a \in U)}{\Longrightarrow} \exists a^{-1} \in U^{-1}$,

also $e = a \cdot a^{-1} \in U \cdot U^{-1} \underset{\text{Vor.}}{\subseteq} U \Rightarrow e \in U$

und e neutrales Element in U.

$(G4)$ Sei $a \in U; a^{-1}$ existiert in G und damit $a^{-1} \in U^{-1}$;

ferner $e \in U$, also $a^{-1} = e \cdot a^{-1} \in U \cdot U^{-1} \underset{\text{Vor.}}{\subseteq} U$.

$(G1)$ $[a, b \in U \underset{\text{s.o.}}{\Longrightarrow} a, b^{-1} \in U \Rightarrow a(b^{-1})^{-1} \in U \cdot U^{-1} \subseteq U$

$\underset{(b^{-1})^{-1}=b}{\Longrightarrow} a \cdot b \in U] \Rightarrow U$ abgeschlossen (bzgl. $\cdot$).

$(G2)$ ist erfüllt, da schon die Obermenge $(G, \cdot)$ assoziativ ist. □

Beispiel. Seien $n \in \mathbb{N}$ und $n\mathbb{Z} := \{n \cdot z : z \in \mathbb{Z}\}$. Dann gilt:

$$(n\mathbb{Z}, +) \text{ ist Untergruppe von } (\mathbb{Z}, +) \ .$$

Beweis. $(\mathbb{Z}, +)$ ist Gruppe, $n\mathbb{Z} \subseteq \mathbb{Z}$; $\quad n\mathbb{Z} \neq \emptyset$; ferner
$nz_1, nz_2 \in n\mathbb{Z} \Rightarrow nz_1 + (-nz_2) = n(z_1 - z_2) \in n\mathbb{Z} \underset{(3.16)}{\Longrightarrow}$ Behauptung. □

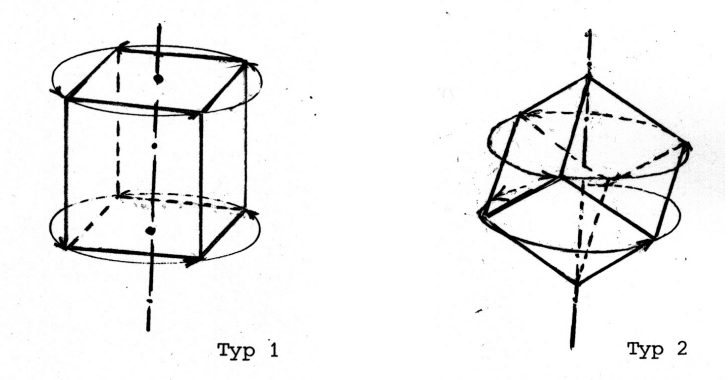
Anmerkung. Für $m, n \in \mathbb{N}$ gilt $\boxed{m\mathbb{Z} \subseteq n\mathbb{Z} \iff n|m}$. Beweis?

M
↓

3.17 Anhang: Gruppe der Deckbewegungen eines Würfels

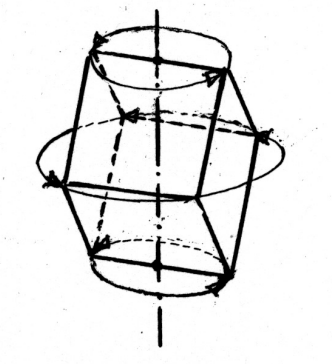
Literatur: Grundkurs Mathematik II 2, DIFF-Studienbrief, Tübingen 1972.
Eigentliche Bewegungen des Raumes sind solche „Kongruenzabbildungen", die sich aus Translationen (Parallelverschiebungen) und Drehungen zusammensetzen lassen. Unter einer (eigentlichen) **Deckbewegung** eines Körpers im Raum versteht man eine eigentliche Bewegung, die den Körper in sich überführt. Die (eigentlichen) Deckbewegungen eines Körpers beschreiben alle seine Symmetrie-Eigenschaften, wenn man von der Spiegelsymmetrie absieht.
Man kann zeigen, dass die Menge der (eigentlichen) Deckbewegungen eines Körpers bzgl. Hintereinanderausführung eine Gruppe bildet. Als Beispiel betrachten wir die **Würfelgruppe** $\vec{W}$, d.h. die Gruppe der eigentlichen Bewegungen des Würfels auf sich. Der Würfel hat die folgenden drei Typen von Dreh-Symmetrieachsen:

1. Achsen durch die Schwerpunkte gegenüberliegender Flächen; Drehung um 90°, 180°, 270°. Es gibt $3 \cdot 3 = 9$ entsprechende Drehungen.



Figur 3.7 a/b: Drehachsen von Deckabbildungen eines Würfels

2. Die räumlichen Diagonalen; Drehung um 120° und 240°. Zu den 4 Raumdiagonalen gibt es $4 \cdot 2 = 8$ entsprechende Drehungen.
3. Achsen durch die Mitte gegenüberliegender Kanten; Drehung um 180°. Von diesem Typ gibt es 6 Drehungen.



Typ 3

Figur 3.7 c: Drehachse einer weiteren Deckabbildung eines Würfels

Zusammen mit der Identität haben wir 24 Deckbewegungen gefunden. Da bei einer Deckbewegung des Würfels jede räumliche Diagonale wieder in eine solche abgebildet wird und umgekehrt jede bijektive Abbildung der Raumdiagonalen zu einer eigentlichen Deckabbildung gehört, ist 24 ($= |\vec{S}_4|$) auch die genaue Anzahl der eigentlichen Deckabbildungen. Damit haben wir alle Elemente der Würfelgruppe $\vec{W}$ angegeben.

Literaturhinweis. Zu Symmetrien (u.a. in Natur und Kunst) allgemein sind die folgenden Abhandlungen besonders lesenswert.

- H.Weyl: Symmetrie, Birkhäuser Verlag, 1981[2].
- E.Quaisser: Die Symmetriestruktur von Figuren. In A.Beutelspacher et al.(Hrsg.): Jahrbuch Überblicke Mathematik 1995, Vieweg Verlag, p.147-161, 1995.
- R.Wille (Hrsg.): Symmetrie in Geistes- und Naturwissenschaft. Springer Verlag, Berlin etc. 1988.

Weitere Literatur:

P.B:Yale: Geometry and Symmetry. Dover Publ., New York 1968.

H.F.Verheyen: Symmetry Orbits. Birkhäuser V., Boston etc. 1996.

- Beachten sie auch Bücher über Kristallographie!

↑ M

Übungsaufgaben:

Aufgabe 3.1: Bezeichne $\vec{W}$ die Würfelgruppe!

a) Bestimmen Sie zu jedem Element $x \in \vec{W}$ die kleinste Zahl $n \in \mathbb{N}$ mit

$$x^n = \mathrm{id},$$

(die sogenannte Ordnung von x)!

b) Nummerieren Sie die Raumdiagonalen und geben zu jedem Element von $\vec{W}$ die entsprechenden Permutation der Raumdiagonalen an!

c) Bestimmen Sie Elemente $x_1, x_2, x_3 \in \vec{W}$ derart, dass jedes Element von $\vec{W}$ sich als Produkt mit Faktoren aus $\{x_1, x_2, x_3\}$ schreiben lässt! (Man sagt dann, dass x_1, x_2, x_3 die Gruppe $\vec{W}$ erzeugen).

Aufgabe 3.2: Bezeichne $\vec{D}_5$ die Gruppe der Symmetrieabbildungen (Dreh- und Spiegelungssymmetrien) „des“ regelmäßigen 5−Ecks! Zeigen Sie, dass gilt:

a) $\vec{D}_5$ enthält 5 Spiegelungen und 5 Drehungen (einschließlich der Drehung id um 0°).

b) $|\vec{D}_5| = 10$. (es gibt also keine weiteren Deckabbildungen).

c) $\vec{D}_5$ enthält genau eine Untergruppe der Elemente-Anzahl (Ordnung) 5 und genau 5 weitere Untergruppen (außer $\{\mathrm{id}\}$ und $\vec{D}_5$ selbst).

Anmerkung: 1.) Diese Gruppe spielte eine Rolle bei der Bestimmung von Prüfziffern für die Nummern der letzten DM-Banknoten.
2.) Diese Aufgabe wird mit Aufgabe 4.2 fortgesetzt.

Aufgabe 3.3 Das Schema $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$ rationaler Zahlen nennen wir ein **magisches**

Quadrat (über $\mathbb{Q}$ mit 3×3 Feldern), wenn alle Zeilen- und Spaltensummen sowie die Summen beider Diagonalen den gleichen Wert s haben, wenn also gilt:

$$s = a+b+c = d+e+f = g+h+j = a+d+g = b+e+h = c+f+j = a+e+j = c+e+g.$$

Zu zwei magischen Quadraten $M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$ und $N = \begin{pmatrix} a' & b' & c' \\ d' & e' & f' \\ g' & h' & j' \end{pmatrix}$ definieren wir eine Summe durch $M + N := \begin{pmatrix} a+a' & b+b' & c+c' \\ d+d' & e+e' & f+f' \\ g+g' & h+h' & j+j' \end{pmatrix}$.

a) Zeigen Sie: M ist durch (a, b, c) bestimmt, und zu jedem Tripel $(a, b, c) \in \mathbb{Q}^3$ gibt es ein magisches Quadrat mit erster Zeile (a, b, c) !
b) Beweisen Sie: Die Menge G aller magischen $3 \times 3-$Quadrate über $\mathbb{Q}$ bildet bzgl. der oben definierten Addition eine kommutative Gruppe !

§ 4 Homomorphismen von Halbgruppen und Gruppen

M
↓

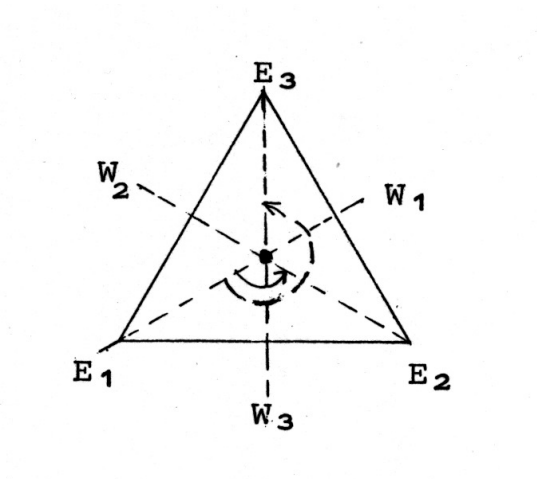
Bei der Betrachtung der Gruppe $\vec{S}_3$ hatten wir auf die Ähnlichkeit im Verhalten der Permutationen von $\{1,2,3\}$ mit dem der Symmetrien (Deckbewegungen) eines gleichseitigen Dreiecks hingewiesen. Diesen Zusammenhang zwischen der Gruppe

$$\vec{S}_3 = (\{\delta_0, \delta_1, \delta_2, \sigma_1, \sigma_2, \sigma_3\}, \circ) \qquad \text{(s. Figur 3.3 !)}$$

und der Gruppe

$$\vec{D}_3 = (\{D_0, D_{120}, D_{240}, S_{W_1}, S_{W_2}, S_{W_3}\}, \circ)$$

(mit D_i als der Drehung um den Schwerpunkt des Dreiecks um $i°$ und S_{W_j} als der Spiegelung an der Halbierenden des Winkels im Punkt E_j, vgl. Figur 4.1)



Figur 4.1: Zur Symmetriegruppe eines gleichseitigen Dreiecks

wollen wir nun präzisieren:
Es existiert eine Bijektion $f : \{D_0, \ldots, S_{w_3}\} \to \{\delta_0, \ldots, \sigma_3\}$, definiert durch $D_{i\cdot 120} \mapsto \delta_i$ $(i = 0, 1, 2)$ und $S_{W_j} \mapsto \sigma_j$. Diese Zuordnung allein reicht nicht zur Beschreibung aus; es wären nur Eigenschaften der zugrunde liegenden Mengen in Beziehung gebracht („gleiche Kardinalzahl"), nicht die der Verknüpfungen, wie z.B.

$$\begin{array}{ccccc} D_{120} & \circ & S_{W_1} & = & S_{W_3} \\ \updownarrow & & \updownarrow & & \updownarrow \\ \delta_1 & \circ & \sigma_1 & = & \sigma_3 \end{array}$$

oder allgemeiner $f(B_1 \circ B_2) = f(B_1) \circ f(B_2)$ für $B_1, B_2 \in \{D_0, \ldots, S_{W_3}\}$.
Mit Abbildungen, die in diesem Sinne „mit der Gruppenstruktur verträglich" sind, wollen wir uns beschäftigen. Derartige Bijektionen bestehen zwischen solchen Gruppen, die mit den Mitteln der Gruppentheorie nicht unterschieden werden können. Aber auch nicht-bijektive strukturbeachtende Abbildungen von (Halb-)Gruppen sind von Bedeutung, übertragen sich doch auch so viele Eigenschaften auf die Bildstruktur; z.B. werden Untergruppen auf Untergruppen abgebildet.
Als **Beispiel** für eine **strukturverträgliche Abbildung**, die keine Bijektion ist, betrachten wir die „kanonische" Abbildung von $\mathbb{Z}$ auf die Menge der Äquivalenzklassen $\mod m$ (für $m \in \mathbb{N}$ fest gewählt): Bezeichnen wir mit $\overline{x}$ die „Restklasse[32] " $\{y \in \mathbb{Z} : y \equiv x \pmod m\}$ von $x \in \mathbb{Z}$, mit $\mathbb{Z}_m$ die Menge all

[32]Zur Erklärung des Namens: $y \equiv x \pmod m$ bedeutet, dass m Teiler von $y - x$ ist und damit x und y den gleichen Rest bei der Division durch m haben.
Beispiel: $5 \equiv 8 \pmod 3$ heißt u.a., dass 5 und 8 den gleichen Rest, nämlich 2, bei der Division durch 3 besitzen.

dieser Restklassen (also $\mathbb{Z}_m := \mathbb{Z}/_{\equiv \pmod m}$), und definieren wir durch

$$\overline{x} + \overline{y} := \overline{x+y} \quad \text{für alle } x, y \in \mathbb{Z}$$

eine Addition auf $\mathbb{Z}_m$, so gilt (Übungsaufgabe): Diese Addition ist wohldefiniert, d.h. $\overline{x}+\overline{y}$ ist – wegen $\overline{x} = \overline{x}_1 \wedge \overline{y} = \overline{y}_1 \Rightarrow \overline{x}+\overline{y} = \overline{x}_1 + \overline{y}_1$ – von den Repräsentanten x und y unabhängig, und

$$(\mathbb{Z}_m, +) \text{ ist eine kommutative Gruppe.}$$

Für die kanonische Abbildung

$$\kappa : \begin{cases} \mathbb{Z} \to \mathbb{Z}_m \\ x \ \mapsto \overline{x} \end{cases}$$

ist nach Definition der Addition die Gleichung $\kappa(x+y) = \kappa(x) + \kappa(y)$ für alle $x, y \in \mathbb{Z}$ erfüllt.

Beispiel. $\mathbb{Z} \to \mathbb{Z}_2$ mit $z \mapsto \overline{z} = \begin{cases} \overline{0} & \text{für } z \text{ gerade} \\ \overline{1} & \text{für } z \text{ ungerade} \end{cases}$ ist eine strukturverträgliche Abbildung (Homomorphismus) von $(\mathbb{Z}, +)$ auf $(\mathbb{Z}_2, +)$.

Ähnlich gilt für die durch $\overline{x} \cdot \overline{y} := \overline{x \cdot y}$ (wohl-)definierte Multiplikation:

$$(\mathbb{Z}_m, \cdot) \text{ ist eine kommutative Halbgruppe,}$$

und κ ist eine auch mit dieser Struktur verträgliche Abbildung von $\mathbb{Z}$ auf $\mathbb{Z}_m$.

Beispiel (Anwendung). „**Neunerprobe**“ Gegeben sei $n \in \mathbb{N}$! Wir fragen: Ist 9 ein Teiler von n ?

Seien $a_0, \ldots, a_k$ die Ziffern der Dezimaldarstellung von n, also $n = \sum_{i=0}^{k} a_i 10^i$ (nach Identifikation der Ziffern $0, \ldots, 9$ mit den entsprechenden Zahlen). Gehen wir zur Kongruenz mod 9 über, so lautet die Frage: $\overline{n} = \overline{0}$? Nun ist $\overline{10^i} = \overline{10}^i = (\overline{1})^i = \overline{1^i} = \overline{1}$ und

$$\overline{n} = \overline{\sum_{i=0}^{k} a_i \cdot 10^i} = \sum_{i=0}^{k} \overline{a_i \cdot 10^i} = \sum_{i=0}^{k} \overline{a_i} \cdot \overline{10^i} = \sum_{i=0}^{k} \overline{a_i} \cdot \overline{1} = \overline{\sum_{i=0}^{k} a_i}.$$

Daher gilt: ***n* ist genau dann durch 9 teilbar, wenn die Quersumme von *n* (in der Dezimaldarstellung) durch 9 teilbar ist.**

Zahlenbeispiel. 9 teilt 49068, da $9 | (4+9+0+6+8)$.

Untersuchen Sie analog die „ **Dreierprobe** “ und die „ **Elferprobe** “ !

↑ M

4.1 Definition. (Halb-) Gruppen-Homomorphismen, -Isomorphismen

Seien $(H_1, *_1)$ und $(H_2, *_2)$ Halbgruppen!

(a) Eine Abbildung $f : H_1 \to H_2$ heißt **Homomorphismus** (Hom) von H_1 in H_2, falls gilt:
$$\forall a, b \in H_1 : f(a *_1 b) = f(a) *_2 f(b).$$

(b) Ist f zusätzlich bijektiv, so heißt f **Isomorphismus** (Iso). Ein Isomorphismus einer Halbgruppe auf sich selbst (also für $H_1 = H_2$) heißt **Automorphismus** (Aut).

(c) $(H_1, *_1)$ und $(H_2, *_2)$ heißen **isomorph**, falls es einen Isomorphismus von $(H_1, *_1)$ auf $(H_2, *_2)$ gibt.
In Zeichen : $H_1 \cong H_2$.

M
↓

Beispiele.

(1) Die am Anfang dieses Paragraphen angegebene Abbildung zwischen der Gruppe $\vec{D}_3$ der Deckabbildungen eines gleichseitigen Dreiecks und der symmetrischen Gruppe $\vec{S}_3$ ist ein (Gruppen-) Isomorphismus. Also gilt $\vec{D}_3 \cong \vec{S}_3$.

(2) Für die Würfelgruppe $\vec{W}$ (vgl. Anhang zu §3) gilt $\vec{W} \cong \vec{S}_4$.

(3) Die Abbildung „$^-$“ : $\begin{cases} \mathbb{Z} \to \mathbb{Z}_m \\ x \mapsto \overline{x} \end{cases}$ liefert einen (surjektiven) Homomorphismus der Gruppe $(\mathbb{Z}, +)$ auf die Gruppe $(\mathbb{Z}_m, +)$, sowie einen (surjektiven) Homomorphismus der Halbgruppe $(\mathbb{Z}, \cdot)$ auf die Halbgruppe $(\mathbb{Z}_m, \cdot)$.

(4) Die Abbildung $\begin{cases} \overline{0} \mapsto g \\ \overline{1} \mapsto u \end{cases}$ liefert einen Isomorphismus von $(\mathbb{Z}_2, +)$ auf $(\{g, u\}, \oplus)$ (s. §1 !).

(5) $g : \begin{cases} \mathbb{Z} \to \mathbb{R}^* \\ z \mapsto r^z \end{cases}$ (für $r \in \mathbb{R}^*$ fest) ist ein Homomorphismus von $(\mathbb{Z}, +)$ in $(\mathbb{R}^*, \cdot)$.

Beweis. g ist wohldefiniert, $g(z_1 + z_2) = r^{z_1+z_2} = r^{z_1} \cdot r^{z_2} = g(z_1) \cdot g(z_2)$. □

(6) $h : \begin{cases} \mathbb{C} \to \mathbb{C} \\ z = a + bi \mapsto \overline{z} = a - bi \end{cases}$ ist ein Isomorphismus von $(\mathbb{C}, +)$ auf sich (Automorphismus).
$h|_{\mathbb{C}^* \to \mathbb{C}^*}$ ist ein Isomorphismus von $(\mathbb{C}^*, \cdot)$ auf sich.

(7) Seien $(\mathbb{R}^2, +) := (\{(x, y) : x, y \in \mathbb{R}\},$ komponentenweise Addition) und
$(\mathcal{V}, +) := \left(\left\{ \begin{pmatrix} x \\ y \end{pmatrix}_{\vec{e}_1, \vec{e}_2} : x, y \in \mathbb{R} \right\},$ komponentenweise Addition $\right)$; dann ist
$f : \begin{cases} \mathbb{R}^2 \to \mathcal{V} \\ (x, y) \mapsto \begin{pmatrix} x \\ y \end{pmatrix}_{\vec{e}_1, \vec{e}_2} \end{cases}$ ein Isomorphismus von $(\mathbb{R}^2, +)$ auf $(\mathcal{V}, +)$.

(8) $\mathrm{pr}_1 : \begin{cases} \mathcal{V} \to \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix}_{\vec{e}_1,\vec{e}_2} \mapsto x \end{cases}$ ist Homomorphismus von $(\mathcal{V},+)$ auf $(\mathbb{R},+)$.

Beweis.
$\mathrm{pr}_1 \left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} \right] = \mathrm{pr}_1 \begin{pmatrix} x_1+x_2 \\ y_1+y_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} = x_1 + x_2 = \mathrm{pr}_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}_{\vec{e}_1,\vec{e}_2} + \mathrm{pr}_1 \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}_{\vec{e}_1,\vec{e}_2}$. □

↑ M

4.2 Hilfssatz. Bilder von neutralem Element und von Inversen

Seien $(G,*)$ Gruppe mit neutralem Element e, $(G',*')$ Gruppe mit neutralem Element e' und $f : G \to G'$ ein Homomorphismus. Dann gilt:

(a) $f(e) = e'$;

(b) $\forall a \in G : f(a^{-1}) = [f(a)]^{-1}$.

D.h.: Bei einem Gruppenhomomorphismus wird das neutrale Element auf das neutrale Element und das Inverse eines Elements auf das Inverse des Bildelements abgebildet.

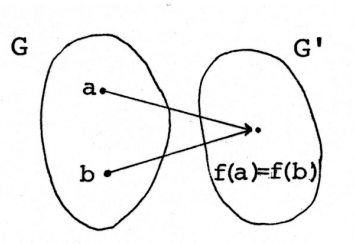
Beweis. (a) $f(e) *' e' = f(e) = f(e*e) = f(e) *' f(e) \underset{\text{Mult. mit } [f(e)]^{-1}}{\Rightarrow} e' = f(e)$;

(b) $f(a^{-1}) *' f(a) \underset{(4.1(a))}{=} f(a^{-1} * a) = f(e) \underset{(a)}{=} e' \Rightarrow f(a^{-1}) = [f(a)]^{-1}$.

□

M ↓

Bei einem Homomorphismus werden u.U. mehrere Elemente der Urbildgruppe auf dasselbe Element der Bildgruppe abgebildet. (Beispiele?) Wir untersuchen „Urbildmengen“ der Elemente von G'.



Figur 4.2: Bildelement mit zwei Urbildern

Sei also $f(a) = f(b)$ (vgl. Figur 4.2 !); dann gilt
$e' = f(a)^{-1} *' f(b) \underset{4.2(b)}{=} f(a^{-1}) *' f(b) = f(a^{-1} * b)$, d.h.
$f(a) = f(b) \Rightarrow a^{-1} * b \in f^-(\{e'\})$ (und umgekehrt).

Es ist also sinnvoll, sich näher mit $f^-(\{e'\})$ zu beschäftigen; Ziel dieses Paragraphen ist der Nachweis, dass (erstaunlicher Weise) durch diese Menge f schon weitgehend festgelegt ist.

↑ M

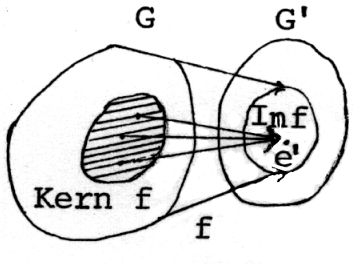
4.3 Definition. Kern, Bild eines Homomorphismus

Seien $(G, *)$, $(G', *')$ Gruppen mit neutralem Element e bzw. e' und $f : G \to G'$ Homomorphismus. Dann heißt

$\operatorname{Kern} f := \{x \in G : f(x) = e'\} = f^{-}(\{e'\})$ der **Kern** von f (vgl. Figur 4.3 !) und

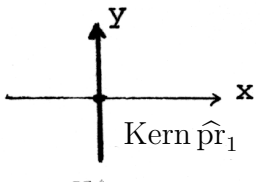
$\operatorname{Bild} f := \operatorname{Im} f := \{f(x) : x \in G\} = f(G)$ das **Bild** von f.

Beispiele.

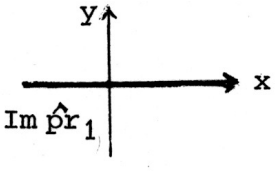


Figur 4.3: Kern f und Bild f

1. $\widehat{\mathrm{pr}}_1 : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2 \\ (x, y) \mapsto (x, 0) \end{cases}$ ist Homomorphismus von $(\mathbb{R}^2, +)$ in $(\mathbb{R}^2, +)$.



Figur 4.4: $\operatorname{Kern} \widehat{\mathrm{pr}}_1 = \{(0, y) : y \in \mathbb{R}\}$



Figur 4.5: $\operatorname{Bild} \widehat{\mathrm{pr}}_1 = \{(x, 0) : x \in \mathbb{R}\}$

2. Für $m \in \mathbb{N}$ ist $\kappa : \begin{cases} \mathbb{Z} \to \mathbb{Z}_m \\ x \mapsto \overline{x} \end{cases}$ Homomorphismus von $(\mathbb{Z}, +)$ auf $(\mathbb{Z}_m, +)$ (s.o.!).

   Es gilt $\operatorname{Bild} \kappa = \mathbb{Z}_m$ (da κ surjektiv ist) und

   $\operatorname{Kern} \kappa = \kappa^{-}(\overline{0}) = \{x \in \mathbb{Z} : x \equiv 0 \bmod m\} = \{x \in \mathbb{Z} : m | x\} = \{my : y \in \mathbb{Z}\} =: m\mathbb{Z}$.

4.4 Hilfssatz: Kern und Bild als Untergruppen

Seien $(G, *)$, $(G', *')$ Gruppen und $f : G \to G'$ Homomorphismus. Dann gilt:

(1) $\operatorname{Kern} f$ ist UG von $(G, *)$.

(2) $\operatorname{Bild} f$ ist UG von $(G', *')$.

Beweis. (1) $f(e) = e' \Rightarrow e \in \operatorname{Kern} f \Rightarrow \operatorname{Kern} f \neq \emptyset$.

Seien $u, v \in \operatorname{Kern} f$, also $f(u) = e' = f(v)$. Dann gilt

$$f(u * v^{-1}) \underset{(4.1)\mathrm{a}}{=} f(u) *' f(v^{-1}) \underset{(4.2)\mathrm{b}}{=} f(u) *' f(v)^{-1} = e' *' e'^{-1} = e'$$

Also $u * v^{-1} \in \operatorname{Kern} f$. Aus dem Untergruppen-Kriterium (3.16) folgt die Behauptung.

(2) $e' = f(e) \in f(G) = \text{Bild}\, f \Rightarrow \text{Bild}\, f \neq \emptyset$.

Seien $u', v' \in \text{Bild}\, f$; dann $\exists u, v \in G : f(u) = u'$ und $f(v) = v'$; damit gilt
$$u' *' (v')^{-1} = f(u) *' (f(v))^{-1} = f(u) *' f(v^{-1}) = f(u * v^{-1}) \in \text{Bild}\, f.$$
Wieder nach (3.16) folgt die Behauptung.

□

4.5 Hilfssatz: Kern bei injektiven Homomorphismen

Ist f Gruppenhomomorphismus, dann gilt

$$f \text{ injektiv } \iff \text{Kern}\, f = \{e\} \quad .$$

Beweis. „⇒“ Ist f injektiv, so hat e' höchstens ein Urbild (und wegen $f(e) = e'$ genau eins).

„⇐“ Seien $a, b \in G$ mit $f(a) = f(b)$, dann folgt
$f(a^{-1} * b) = f(a)^{-1} *' f(b) = f(b)^{-1} *' f(b) = e'$,
also $a^{-1} * b \in \text{Kern}\, f \underset{\text{(nach Voraussetzung)}}{=} \{e\}$ und damit $a = b$.

□

Weiter oben (nach (4.2)) und eben hatten wir gesehen, dass für einen Homomorphismus f gilt:
$$f(a) = f(b) \Rightarrow a^{-1} * b \in f^{-}(\{e'\}) = \text{Kern}\, f \quad (\text{d.h. } b \in a * \text{Kern}\, f).$$
Ist umgekehrt $b \in a * \text{Kern}\, f$, so folgt

$$f(b) \in f(a * \text{Kern}\, f) = f(a) *' f(\text{Kern}\, f) = f(a) *' \{e'\} = \{f(a)\}, \text{ also } f(b) = f(a).$$

Somit gilt (vgl. Figur 4.6 !) in multiplikativer Schreibweise der erste Teil von:

4.6 Hilfssatz: Volle Urbilder

Seien $(G, \cdot)$ und $(G', \cdot')$ Gruppen und $f : G \to G'$ Homomorphismus. Dann gilt für jedes $a \in G$:

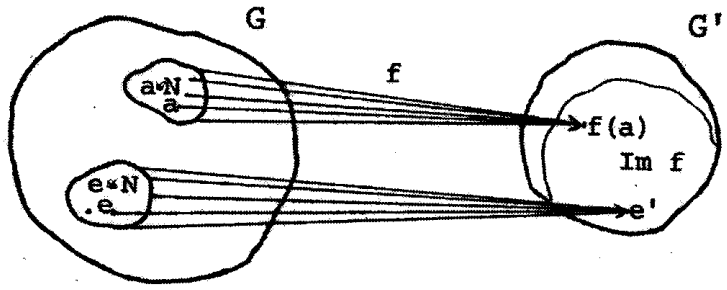
Das volle Urbild von $f(a)$ ist gleich $a \cdot \text{Kern}\, f$ und (s.u.) gleich $\text{Kern}\, f \cdot a$.

4.7 Definition: Nebenklasse nach einer Untergruppe

Sei $(G, \cdot)$ Gruppe und U Untergruppe von G. Für jedes $a \in G$ heißt

$a \cdot U$ **linke Nebenklassee** von U in G
$U \cdot a$ **rechte Nebenklassee** von U in G

Achtung: Manche Autoren benutzen umgekehrte Bezeichnungen: $a \cdot U$ rechte Nebenklassen etc.



Figur 4.6: Volle Urbilder (mit $N := \operatorname{Kern} f$)

Anmerkung. Was für die Nebenklassen von Kern f als volle Urbildmengen offensichtlich ist, gilt auch für die Nebenklassen einer beliebigen Untergruppe U:
Die rechten (bzw. linken) Nebenklassen von U in G bilden eine Partition der Menge G. (Beweis ...)

Jedes Element einer Nebenklasse heißt **Repräsentant** dieser Nebenklasse.

Beispiel. Für $m \in \mathbb{Z}$ ist $(m\mathbb{Z}, +)$ UG von $(\mathbb{Z}, +)$; die linken und rechten Nebenklassen von $m\mathbb{Z}$ in $(\mathbb{Z}, +)$ haben die Form $r + m\mathbb{Z}$ $(= \overline{r})$. (Ist $r \in \{0, \ldots, m-1\}$, so hat, wie schon erwähnt, jedes Element aus $r+m\mathbb{Z}$ bei der Division durch m den Rest r, weswegen man auch von **Restklassen** spricht). Ferner gilt:

$$\mathbb{Z} = \dot{\bigcup_{r=0}^{m-1}} (r + m\mathbb{Z}) = \dot{\bigcup_{r=0}^{m-1}} \overline{r} \quad \text{(Überdeckung von } \mathbb{Z} \text{ durch Restklassen mod } m\text{)}.$$

M ↓

Unser nächstes Ziel ist die Konstruktion von neuen Gruppen (aus einer gegebenen Gruppe G), die in engem Zusammenhang mit Homomorphismen von G stehen.
Auf den Restklassen mod m lässt sich, wie wir sahen, eine Addition definieren, die mit der Addition auf $\mathbb{Z}$ zusammenhängt. Für Nebenklassen einer nicht-kommutativen Gruppe nach einer beliebigen UG ist es im allgemeinen nicht möglich, auf ähnliche Weise wie im oben erwähnten Fall eine Verknüpfung zu definieren. Dafür benötigt man eine zusätzliche Eigenschaft der UG, die aber von Kern f erfüllt wird.

↑ M

Aus $f(a) = f(b)$ erhält man nämlich auch $f(b \cdot a^{-1}) = f(b) \cdot' f(a^{-1}) = f(b) \cdot' f(a)^{-1} = e'$, also $b \in \operatorname{Kern} f \cdot a$; da $f(\operatorname{Kern} f \cdot a) = f(a)$ ist, folgt daraus: Auch $(\operatorname{Kern} f) \cdot a$ ist das volle Urbild von $f(a)$. Es gilt also:

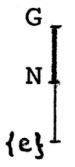
$$(\operatorname{Kern} f) \cdot a = a \cdot (\operatorname{Kern} f).$$

4.8 Definition: Normalteiler

Eine Untergruppe N der Gruppe $(G, \cdot)$ heißt **Normalteiler** von G, wenn gilt:

$$\forall a \in G: \; a \cdot N = N \cdot a.$$

Schreibweise: $N \trianglelefteq G$. Symbolische Darstellung s. Figur 4.7 !



Figur 4.7: Symbolische Darstellung von $N \trianglelefteq G$

Beispiel. Bei einer kommutativen Gruppe ist jede Untergruppe Normalteiler.

4.9 Satz: Kern als Normalteiler

Sind G, G' Gruppen und ist $f : G \to G'$ Homomorphismus, dann gilt:

$$\text{Kern}\, f \quad \text{ist Normalteiler von } G.$$

Bei Nebenklassen nach einem Normalteiler N ist es möglich, mit Hilfe der Verknüpfung von G eine Verknüpfung auf G/N sinnvoll zu definieren:

4.10 Definition: G/N

Sei $(G, *)$ Gruppe und N Normalteiler von $(G, *)$. Dann definiert man:

1. $G/N := \{a * N : a \in G\}$.
2. Für $a, b \in G$ sei
   $(a * N)\tilde{*}(b * N) := (a * N) * (b * N)$ (in Komplexschreibweise, vgl. (3.15) !)

Man beachte, dass die Bildung von $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ und die Definition der Addition auf $\mathbb{Z}_m$ einen Spezialfall der eben erfolgten darstellt.

4.11 Satz: Faktorgruppe

Seien $(G, *)$ Gruppe, N Normalteiler von $(G, *)$ und G/N und $\tilde{*}$ wie eben definiert. Dann gilt:

$$(G/N, \tilde{*}) \text{ ist eine Gruppe,}$$

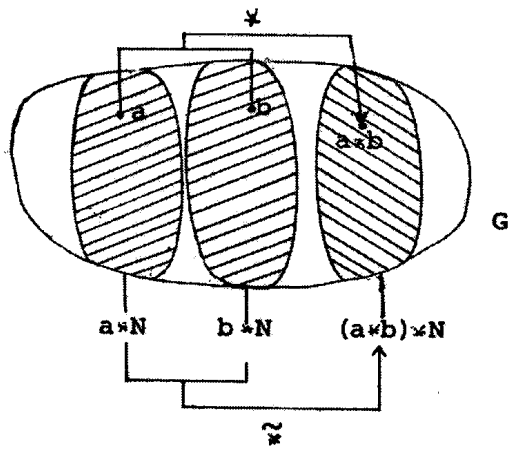
die so genannte **Faktorgruppe** von G nach N.

Anmerkung. Es gilt $\boxed{(a * N)\tilde{*}(b * N) = (a * b) * N}$ (s.u.). Auch diese Beziehung wird oft zur Definition von $\tilde{*}$ herangezogen; dann ist jedoch nachzuweisen, dass $\tilde{*}$ wohldefiniert ist.

Beweis. (G1) Seien $a * N, b * N \in G/N$ gegeben. Dann gilt (gleichzeitig ein Beweis der Anmerkung):

$$\begin{aligned}(a * N)\tilde{*}(b * N) &\underset{\text{Def.}}{=} (a * N) * (b * N) \underset{\text{Assoz.}}{=} a * (N * b) * N \\ &\underset{\text{N Nt.}}{=} a * b * N * N \underset{(N,*)\text{ abgeschl.}}{=} (a * b) * N\ .\end{aligned}$$

Insbesondere ist
$(a * N)\tilde{*}(b * N) \in G/N$.
Also:
$\tilde{*}$ ist innere Verknüpfung auf G/N
und
$\tilde{*}$ entspricht der Verknüpfung $*$ der Repräsentanten.



Figur 4.8: Zur Verknüpfung der Faktorgruppe

(G2) Unter Verwendung von (G1) unmittelbare Folgerung aus der Assoziativität von $(G, *)$.

(G3) $N = e * N$ ist neutrales Element: $N \in G/N$

$$N\tilde{*}(a * N) = (e * N)\tilde{*}(a * N) \underset{\text{s.o.}}{=} (e * a) * N = a * N$$

$$(a * N)\tilde{*}N \underset{\text{analog}}{=} a * N.$$

(G4) Sei $M \in G/N$; dann existiert $a \in G : M = a * N$, und es existiert $a^{-1} \in G$.

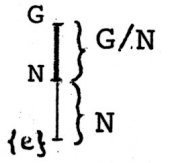
Behauptung: $a^{-1} * N$ ist Inverse von M.

Beweis. $a^{-1} * N \in G/N$;

$$M\tilde{*}(a^{-1} * N) = (a * N)\tilde{*}(a^{-1} * N) = (a * a^{-1}) * N = N$$

$$(a^{-1} * N)\tilde{*}M = (a^{-1} * N)\tilde{*}(a * N) = (a^{-1} * a) * N = N.$$

Damit ist die Behauptung gezeigt.
Also erfüllt $(G/N, \tilde{*})$ die Gruppenaxiome. □



Figur 4.9: Symbolische Darstellung der Faktorgruppe (vgl. auch Figur 4.7 !)

Wir betrachten nun die Abbildung, die jedem Element x von G diejenige Nebenklasse nach N zuordnet, in der x liegt. (Vgl. $r \mapsto \overline{r}$ im Falle $\mathbb{Z} \longrightarrow \mathbb{Z}_m$).

4.12 Hilfssatz: kanonischer Homomorphismus

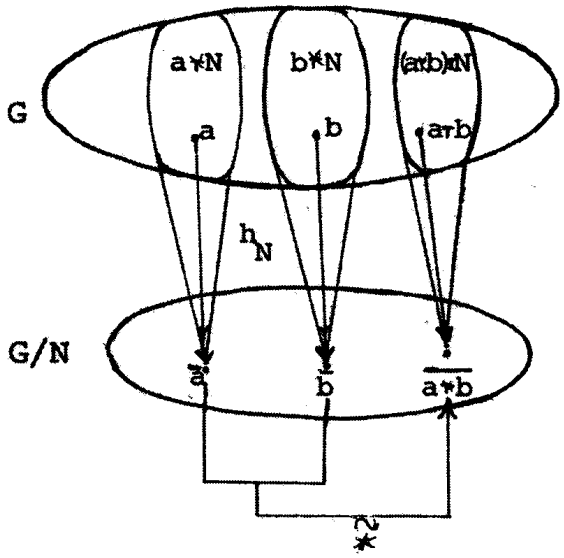
Sei N Normalteiler der Gruppe $(G, *)$. Dann ist

$$h_N : \begin{cases} G \to G/N \\ x \ \mapsto x * N \end{cases}$$

ein Homomorphismus von $(G, *)$ nach $(G/N, \tilde{*})$, der so genannte **natürliche (kanonische) Homomorphismus** von G auf G/N, und es gilt:

$$\operatorname{Kern} h_N = N.$$

Anmerkung (1). Mit $\overline{a} := a * N = h_N(a)$ gilt $\overline{a} \,\tilde{*}\, \overline{b} = \overline{a * b}$; (vgl. Figur 4.10 !).



Figur 4.10: Zur Wirkung des natürlichen Homomorphismus

Anmerkung (2). Aus (4.9) und (4.12) folgt: Jeder Kern eines Gruppenhomomorphismus ist Normalteiler, und umgekehrt ist jeder Normalteiler einer Gruppe Kern eines geeigneten Homomorphismus.

Beweis von 4.12. (a) $h_N : G \to G/N$ mit $a \mapsto \overline{a} = a * N$ ist wohldefiniert, und es gilt:

$h_N(a * b) = (a * b) * N = (a * N) \tilde{*} (b * N) = h_N(a) \tilde{*} h_N(b)$.

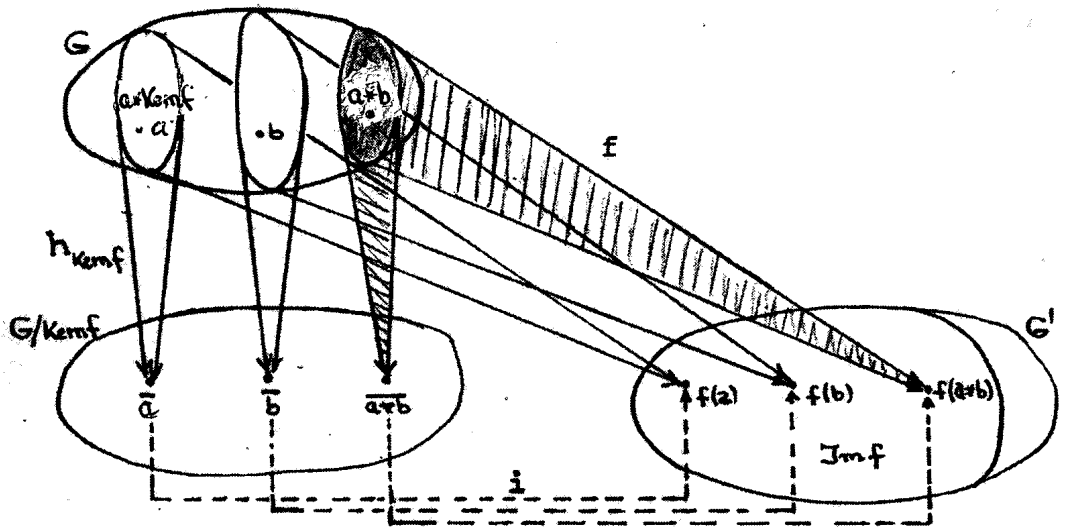
Also ist h_N Homomorphismus von G auf G/N.

(b) $x \in \operatorname{Kern} h_N \Leftrightarrow x * N = N \Leftrightarrow x \in N$.

□

Anmerkung (3). Damit ist ein Zwischenziel erreicht: Zu jeder Gruppe G können wir homomorphe Bilder konstruieren. Nun ist die Frage: Gibt es weitere homomorphe Bilder?

Sei uns daher ein beliebiger Homomorphismus $f : G \to G'$ gegeben. Wählen wir nun speziell N als $\operatorname{Kern} f$, so ergibt sich das Diagramm der Figur 4.11.



Figur 4.11 Zum Homomorphiesatz

4.13 Homomorphiesatz für Gruppen

Seien G, G' Gruppen und $f : G \to G'$ ein Homomorphismus. Dann gilt:

$$\boxed{G/\operatorname{Kern} f \cong \operatorname{Bild} f}$$

Insbesondere ist $\operatorname{Bild} f$ bis auf Isomorphie durch G und $\operatorname{Kern} f$ bestimmt.

Beweis. (a) Die Zuordnung $i : G/\operatorname{Kern} f \to \operatorname{Bild} f$ mit $a * \operatorname{Kern} f \mapsto f(a)$ ist wohldefiniert (d.h. insbesondere rechtseindeutig); denn aus zwei möglichen Darstellungen einer Nebenklasse $a * \operatorname{Kern} f = b * \operatorname{Kern} f$ folgt

$$\{f(a)\} = f(a * \operatorname{Kern} f) = f(b * \operatorname{Kern} f) = \{f(b)\}.$$

(b) i ist Homomorphismus von $(G/\operatorname{Kern} f, *)$ in $(\operatorname{Bild} f, *'|_{\operatorname{Bild} f})$; denn es gilt:

$$\begin{aligned} i[(a * \operatorname{Kern} f)\tilde{*}(b * \operatorname{Kern} f)] &\underset{\text{Eigensch. von } \tilde{*}}{=} i[(a * b) * \operatorname{Kern} f] \underset{\text{Def. v. } i}{=} f(a * b) \\ &\underset{f \text{ Hom.}}{=} f(a) *' f(b) = i[a * \operatorname{Kern} f] *' i[b * \operatorname{Kern} f]. \end{aligned}$$

(c) i ist surjektiv:

Sei $a' \in \operatorname{Bild} f$. Dann ex. $a \in G : f(a) = a'$; daraus folgt $a' = f(a) = i(a * \operatorname{Kern} f)$.

(d) i ist injektiv:

Nach Hilfssatz (4.5) reicht es zu zeigen, dass $\operatorname{Kern} i$ nur das neutrale Element, hier von $G/\operatorname{Kern} f$, enthält.

$$\begin{aligned}\operatorname{Kern} i &= \{a * \operatorname{Kern} f : a \in G \wedge i(a * \operatorname{Kern} f) = e'\} \\ &= \{a * \operatorname{Kern} f : a \in G \wedge f(a) = e'\} \\ &= \{a * \operatorname{Kern} f : a \in \operatorname{Kern} f\} = \{\operatorname{Kern} f\}.\end{aligned}$$

$\operatorname{Kern} f$ ist neutrales Element von $(G/\operatorname{Kern} f, \tilde{*})$ (s.o.).

Insgesamt ist also $(G/\operatorname{Kern} f, *) \cong (\operatorname{Bild} f, *'|_{\operatorname{Bild} f})$

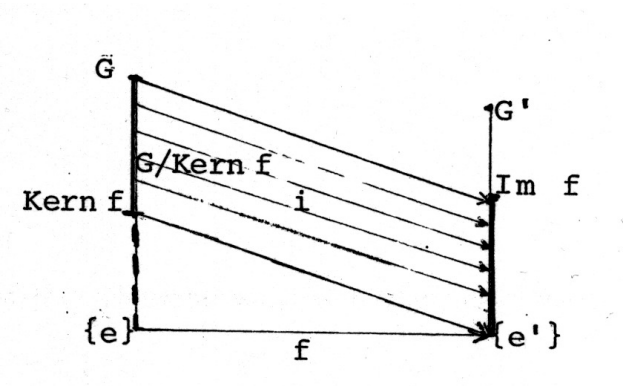
□

M ↓

Damit ist geklärt, dass alle homomorphen Bilder einer Gruppe G isomorph zu Faktorgruppen von G sind. Kennt man also alle Normalteiler von G, so auch (zumindest theoretisch) die Struktur aller homomorphen Bilder von G. Homomorphismen sind insbesondere schon weitgehend durch ihren Kern festgelegt.

↑ M

Eine symbolische Darstellung des Sachverhalts beim Homomorphiesatz zeigt Figur 4.12.



Figur 4.12: Symbolische Darstellung zum Homomorphiesatz

Übungsaufgaben

Aufgabe 4.1

(a) Zeigen Sie, dass eine ganze Zahl genau dann durch 3 teilbar ist, wenn ihre Quersumme durch 3 teilbar ist. („**Dreierprobe**“).

(b) Wie kann (analog) eine „**Elferprobe**“ aussehen?

Aufgabe 4.2

Sei G die Menge aller Elemente aus $\mathbb{Z}_9$, die eine multiplikative Inverse besitzen.
a) Zeigen Sie, dass $(G, \cdot)$ eine Gruppe ist, die von einem Element erzeugt wird.
b) Geben Sie zu jedem Element g aus G seine Ordnung[33] an!

[33] Dabei ist die Ordnung eines Gruppenelementes g definiert als die kleinste natürliche Zahl k mit $g^k = 1$.

Aufgabe 4.3 (Fortsetzung von Aufg. 3.2)
Definieren Sie eine Verknüpfung $\star$ auf $\mathbb{Z}_{10}$ derart, dass folgende Abbildung von der Gruppe der Symmetrien eines regelmäßigen 5−Ecks [34] in $(\mathbb{Z}_{10}, \star)$ ein Gruppenisomorphismus ist: $\vec{D}_5 \longrightarrow \mathbb{Z}_{10}$ mit $D_{j\cdot 72^\circ} \mapsto \overline{j}$ und $D_{j\cdot 72^\circ} \circ S \mapsto \overline{j+5}$ (für $j = 0, \ldots 4$).
Lösungshinweis: Unterscheiden Sie 4 Fälle!

Aufgabe 4.4 Sei C_n eine zyklische Gruppe (d.h. eine von einem Element erzeugte Gruppe) der endlichen Ordnung (Elementeanzahl) n und a ein erzeugendes Element von C_n ! Für jedes $k \in \{1, \ldots, n\}$ ist dann folgende Abbildung ein Homomorphismus:

$$\varphi_k : C_n \to C_n \text{ mit } \varphi_k(x) = x^k \text{ für alle } x \in C_n.$$

Zeigen Sie, dass jeder Homomorphismus $\varphi : C_n \to C_n$ von dieser Form ist!
Lösungshinweis: Beachten Sie, dass φ schon durch k mit $\varphi(a) = a^k$ festgelegt ist.

[34] mit den Drehungen $D_{j\cdot 72^\circ}$ um $j \cdot 72^\circ$ um den Mittelpunkt und einer „Klappung“ S des 5−Ecks auf sich.

§ 5 Ringe und Körper

Auf den Zahlbereichen $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, aber auch auf $\{0,1\}$ bzw. $\{g,u\}$, sind jeweils zwei innere Verknüpfungen definiert. Deren Gesetzmäßigkeit wollen wir nun erfassen.
Statt $*_1$, $*_2$ schreiben wir für die beiden Verknüpfungen auch im allgemeinen Fall „$+$“ und „$\cdot$“ (in Anlehnung an die angegebenen Beispiele).

5.1 Definition: Ring

Sei R eine Menge. Ferner seien auf R zwei innere Verknüpfungen definiert:

$$„+“ : \begin{cases} R \times R \to R \\ (a,b) \quad \mapsto a+b \end{cases} \quad \text{und} \quad „\cdot“ : \begin{cases} R \times R \to R \\ (a,b) \quad \mapsto a \cdot b =: ab \end{cases}$$

Gelten dann folgende Aussagen:

(R1) $(R,+)$ ist kommutative Gruppe (mit neutralem Element $0 \in R$);

(R2) $(R,\cdot)$ ist Halbgruppe;

(R3) die Distributivgesetze:

$\forall a,b,c \in R : (a+b)\cdot c = a\cdot c + b\cdot c$ und

$\forall a,b,c \in R : a\cdot(b+c) = a\cdot b + a\cdot c$,

dann heißt das Tripel $(R,+,\cdot)$ **Ring** (engl.: ring).

Anmerkung. 1.) Sind keine Verwechslungen (zwischen Ring und Menge) zu befürchten und ist klar, um welche Verknüpfungen es sich handelt, so schreibt man nur R statt $(R,+,\cdot)$.
2.) Oft wird (wie oben schon angegeben) der Mal-Punkt beim Produkt weggelassen.
3.) Um Klammern zu sparen, vereinbaren wir, dass die Multiplikation stärker bindet als die Addition; also z.B. $ab + c := (a\cdot b) + c$.

Beispiele.

(a) $(\mathbb{Z},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, $(\mathbb{C},+,\cdot)$ sind Ringe.

(b) $(\mathbb{Z}/m\mathbb{Z},+,\cdot)$ ist für $m \in \mathbb{N}$ ein Ring, der sogenannte **Restklassenring mod m**.

Anmerkung. Ist $m \in \mathbb{N}\setminus\{1\}$ keine Primzahl, so hat $(\mathbb{Z}/m\mathbb{Z}),+,\cdot)$ sogenannte „**Nullteiler**“ (engl.: zero divisor), also Elemente $s,t \neq 0$ mit $s\cdot t = 0$. Ist z.B. $m = r\cdot s$ mit $r,s \in \mathbb{N}\setminus\{m\}$, dann gilt $\overline{r} \neq \overline{0}$, $\overline{s} \neq \overline{0}$, aber $\overline{r}\cdot\overline{s} = \overline{m} = \overline{0}$.
Insbesondere kann daher $(\mathbb{Z}_m{}^*,\cdot)$ keine Gruppe sein. *Beispiel:* $\overline{2}\cdot\overline{3} = \overline{0}$ in $(\mathbb{Z}_6,\cdot)$.

5.2 Anmerkung und Definition: Integritätsbereich

1. $(R,+,\cdot)$ heißt **nullteilerfrei**, wenn gilt

   (R4) $\forall a,b \in R : [a\cdot b = 0 \Rightarrow a = 0 \vee b = 0]$.

   Anm.: Gilt in einem Ring die Nullteilerfreiheit, so folgt die Kürzungsregel

   $$\forall a,bc,d \in R : \quad ac = dc \Longrightarrow (a=d) \vee (c=0).$$

Beweis ?

2. Ein Ring $(R, +, \cdot)$ heißt **kommutativ**, wenn $(R, \cdot)$ kommutativ ist.

3. Ein Ring $(R, +, \cdot)$ heißt **Integritätsbereich**(Integritätsring; engl.: integer domain), wenn er nullteilerfrei und kommutativ ist und mindestens zwei Elemente enthält.

Die oben unter (a) aufgeführten Beispiele von Ringen sind auch Beispiele von Integritätsbereichen; ist $m \neq 1, m \in \mathbb{N}$ keine Primzahl, so ist $\mathbb{Z}/m\mathbb{Z}$ wegen seiner Nullteiler (s. obige Anmerkung zu (b)) kein Integritätsbereich.
Beispiele von Ringen (Fortsetzung):

(c) Sei X Menge, R kommutativer Ring mit 1 (neutralem Element der Multiplikation), z.B. $R = \mathbb{R}$. Zu $f, g \in \mathrm{Abb}(X, R)$ definiert man (in Verallgemeinerung zu den Beispielen (6) zu Halbgruppen, s. §3 !):

$$f + g : \begin{cases} X \to R \\ x \mapsto f(x) + g(x) \end{cases} \quad \text{und} \quad f \cdot g : \begin{cases} X \to R \\ x \mapsto f(x) \cdot g(x) \end{cases}.$$

Dann gilt:

$(\mathrm{Abb}(X, R), +, \cdot)$ **ist ein kommutativer Ring**; dieser ist i.a. nicht nullteilerfrei. (Beweis. ...)

(d) Sei R kommutativer Ring mit 1 (z.B. $R = \mathbb{R}$). Wir definieren (mit $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$):

$\vec{P}(R) := \{f \in \mathrm{Abb}(R, R) | \exists n \in \mathbb{N}_0 \exists a_0, \ldots, a_n \in R : f(x) = \sum_{i=0}^{n} a_i x^i \text{ für alle } x \in R\}$.

Jedes Element $f \in \vec{P}(R)$ heißt *Polynomabbildung*. $(\vec{P}(R), +|_{\vec{P}}, \cdot|_{\vec{P}})$ ist ein Unterring von $(\mathrm{Abb}(R, R)+, \cdot)$. Er heißt **Ring der Polynomabbildungen von R**.

Anmerkung. Definiert man $f_a : R \to R$ durch $f_a(x) = a$ für alle $x \in R$ und $(\mathrm{id})^i : R \to R$ durch $(\mathrm{id})^i(x) = x^i$, und ist $f \in \vec{P}(R)$ mit $f(x) = \sum_{i=0}^{n} a_i x^i$, so gilt für alle $x \in R$:

$(\sum_{i=0}^{n} f_{a_i} \cdot (\mathrm{id}_R)^i)(x) \underset{\substack{\text{Def. + in} \\ \mathrm{Abb}(R,R)}}{=} \sum_{i=0}^{n} [f_{a_i} \cdot (\mathrm{id}_R)^i)](x) \underset{\text{Def.}}{=} \sum_{i=0}^{n} [f_{a_i}(x) \cdot (\mathrm{id}_R)^i)(x)] = \sum_{i=0}^{n} a_i x^i$.
Daher ist $f = \sum_{i=0}^{n} f_{a_i} \cdot (\mathrm{id}_R)^i$.

(d') Man kann zeigen: Die Menge $\mathbb{R}[x] := \{\sum_{i=0}^{n} a_i x^i | n \in \mathbb{N}_0, a_i \in \mathbb{R}\}$ mit den Verknüpfungen

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{m} b_i x^i := \sum_{i=0}^{max(m,n)} (a_i + b_i) x^i$$

und

$$\sum_{i=0}^{n} a_i x^i \cdot \sum_{j=0}^{m} b_j x^j := \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j x^{i+j} = \sum_{k=0}^{n+m} (\sum_{i=0}^{k} a_i b_{k-i}) x^k$$

(wobei $a_i := 0 =: b_j$ für $i > n$ und $j > m$) ist ein Ring, der reelle **Polynomring**.

(e) Sei $(G,+)$ eine abelsche Gruppe und $\mathrm{End}(G) := \{f | f : G \to G \wedge f \text{ Homomorphismus}\}$.

Definiert man für $f, g \in \mathrm{End}\, G$: $\quad f+g : x \mapsto f(x)+g(x)$
$f \circ g : x \mapsto f\big(g(x)\big)$,

dann ist $(\mathrm{End}(G), +, \circ)$ ein Ring, der sogenannte **Endomorphismenring der kommutativen Gruppe G**.

5.3 Definition: Schiefkörper, Körper

Sei K Menge und seien auf K zwei innere Verknüpfungen

$$„+“ : \begin{cases} K \times K \to K \\ (a,b) \quad \mapsto a+b \end{cases} \quad \text{und} \quad „\cdot“ : \begin{cases} K \times K \to K \\ (a,b) \quad \mapsto a \cdot b (=: ab) \end{cases}$$

definiert. Für diese gelte[35]:

(K1) $(K,+)$ ist kommutative Gruppe (mit neutralem Element $0 \in K$).

(K2) $(K^*, \cdot)$ ist Gruppe[36](mit neutralem Element $1 \in K$).

(K3) die Distributivgesetze:

$\forall a,b,c \in K : (a+b)\cdot c = a\cdot c + b \cdot c$ und

$\forall a,b,c \in K : a \cdot (b+c) = a \cdot b + a \cdot c$!

Dann heißt $(K,+,\cdot)$ ein **Schiefkörper** (engl.: skew field). Gilt zusätzlich, dass $(K\setminus\{0\},\cdot)$ kommutativ ist, so heißt $(K,+,\cdot)$ **Körper**(engl.: field).

Anmerkung. 1) Viele Autoren nennen auch Schiefkörper schon Körper und sprechen bei Körpern in unserem Sinne von kommutativen Körpern. Besondere Vorsicht ist also geboten !

2) Die besondere Bedeutung von Körpern liegt für unsere Vorlesung darin, dass die meisten Betrachtungen, die wir an (1.1) bis (1.4) anschließen wollen, nicht von spezielleren Eigenschaften der reellen Zahlen, sondern nur von ihren Körpereigenschaften abhängen. Indem wir als Skalarbereich einen beliebigen Körper zulassen, werden unsere Betrachtungen an zusätzlicher Allgemeinheit (und vielleicht auch Klarheit) gewinnen, ohne wesentlich erschwert zu werden. Unsere Betrachtungen lassen sich dann auch auf „Vektorräume“ über $\mathbb{Q}$, $\mathbb{C}$ oder $\{0,1\}$ (vgl. 1.5) anwenden.

3) Es ist auch möglich, „Vektorräume“ über Schiefkörpern (allgemein) zu definieren. Da hierbei jedoch zusätzliche Schwierigkeiten auftreten (z.B. zu unterscheiden ist, ob man vereinbart mit den Skalaren „von links“ oder „von rechts“ zu multiplizieren), werden wir in dieser Vorlesung davon absehen.

[35] mit $K^* := K \setminus \{0\}$

[36] genauer für die Verknüpfung $\cdot|_{K^*}$

5.4 Beispiele von Körpern

(a) $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$, sind Körper.

(b) $(\{0,1\}), \oplus, \odot)$ mit

$\oplus$	0	1
0	0	1
1	1	0

und

$\odot$	0	1
0	0	0
1	0	1

(vgl. 1.5) ist ein Körper.

(c) Für jede Primzahl p ist $\mathrm{GF}(p) := (\mathbb{Z}_p, +, \cdot) := (\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ ein Körper, das sogenannte **Galoisfeld p** (engl.: Galois field p, $\mathbb{F}_p$) (vgl. auch Bsp. (b) nach (5.1)!).

Beweis: Übungsaufgabe 5.1, s.u. !

Spezialfall: $p = 2$

Die GF(2) zugrunde liegende Menge ist:
$\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ mit $\overline{0} = \{\ldots, -2, 0, 2, 4, \ldots\}$ und $\overline{1} = \{\ldots, -3, -1, 1, 3, \ldots\}$.

Zu den Verknüpfungen! Die Definitionen $\overline{x} + \overline{y} := \overline{x + y}$ und $\overline{x} \cdot \overline{y} := \overline{x \cdot y}$ führen zu folgenden Verknüpfungstafeln:

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\cdot$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

.

Ein Vergleich mit den Verknüpfungstafeln des Beispiels 5.4(b) zeigt, dass die Bijektion

$$i : \begin{cases} \mathbb{Z}_2 \longrightarrow \{0,1\} \\ \overline{0} \mapsto 0 \\ \overline{1} \mapsto 1 \end{cases}$$

sowohl mit den Additionen auf $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$ und $\{0,1\}$ als auch mit den Multiplikationen auf diesen Mengen „verträglich“ ist.

5.5 Definition: Ring-Homomorphismus,Ring-Isomorphie

Seien $(R, +, \cdot)$ und $(S, \hat{+}, \hat{\cdot})$ Ringe. Eine Abbildung $f : R \to S$ heißt **(Ring-) Homomorphismus**, falls für alle $x, y \in R$

$$f(x + y) = f(x) \hat{+} f(y) \text{ und } f(x \cdot y) = f(x) \hat{\cdot} f(y)$$

gilt.[37] Ein bijektiver (Ring-) Homomorphismus heißt **(Ring-) Isomorphismus**.
Im Falle der Existenz eines Isomorphismus des Rings $(R, +, \cdot)$ auf den Ring $(S, \hat{+}, \hat{\cdot})$ heißen die beiden Ringe **isomorph**, in Zeichen $(R, +, \cdot) \cong (S, \hat{+}, \hat{\cdot})$, kurz $R \cong S$.

M
↓

Wie bei Gruppen, so ist auch bei Ringen die Isomorphie als eine „Strukturgleichheit“ aufzufassen. Isomorphe Ringe kann man also als „im Wesentlichen gleich“ ansehen, jedenfalls insofern man sich auf ihre Ringstruktur beschränkt.

Wie wir gleich zeigen werden, ist ein Körper ein Ring mit zusätzlichen Eigenschaften, die unter Ring-Isomorphismen erhalten bleiben. In Fortführung unseres speziellen Beispiels können wir also festhalten: i ist ein (Körper-) Isomorphismus; insbesondere gilt also:

$$(\mathbb{Z}_2, +, \cdot) \cong (\{0,1\}, \oplus, \odot).$$

Auch im Falle von Beispiel (b) sprechen wir daher vom Galoisfeld 2. Ähnlich folgt

$$(\mathbb{Z}_2, +, \cdot) \cong (\{g,u\}, \oplus, \odot).$$

5.6 Zusammenhang Ring – Körper

$$(K,+,\cdot) \text{ ist Körper } \iff \begin{cases} (K,+,\cdot) \text{ ist kommutativer Ring mit}^{38} \quad 1 \text{ und } 1 \neq 0, \\ \text{und zu jedem Element } x \in K \setminus \{0\} \text{ existiert eine} \\ \text{multiplikative Inverse.} \end{cases}$$

Beweisskizze.

„⇒“ Sei $(K,+,\cdot)$ Körper. Nach Definition sind $+$, $\cdot$ innere Verknüpfungen, ist $(K,+)$ kommutative Gruppe und gelten die Distributivgesetze. Da $(K\setminus\{0\},\cdot)$ kommutative Gruppe ist, interessiert nur noch das Verhalten von 0 bei der Multiplikation.

Sei $a \in K$. Dann folgt

$$0\cdot a + 0\cdot a \underset{\substack{\text{Distribut.}\\ \text{Gesetz}}}{=} (0+0)\cdot a \underset{\substack{0 \text{ neutr.Elt}\\ \text{d.Add.}}}{=} 0\cdot a \underset{\substack{\text{Add.v. } -(0\,\cdot\, a)\\ \text{(ex.)}\\ \text{Assoz.-Ges. der}\\ \text{Addition}}}{\Rightarrow} 0\cdot a = 0.$$

Analog lässt sich $a\cdot 0 = 0$ für alle $a \in K$ zeigen. Das 0-Element verletzt daher weder die Assoziativität noch die Kommutativität der Multiplikation (und es gilt $1\cdot 0 = 0$).

Bevor wir die Rückrichtung zeigen, bemerken wir, dass bei der Herleitung von $0\cdot a = 0 = a\cdot 0$ im oberen Beweisteil nur solche Eigenschaften benutzt wurden, die auch Ringe allgemein besitzen, und halten fest:

5.7 Hilfssatz: Produkt mit Null

In jedem Ring $(R,+,\cdot)$ gilt $0\cdot a = 0 = a\cdot 0$ für alle $a \in R$.

Fortsetzung des Beweises von 5.6:

[37] Vgl. diese Definition mit der des Gruppen-Homomorphismus/Isomorphismus in 4.1 !

[38] wobei 1 das neutrale Element der Multiplikation bezeichnet und 0 das der Addition.

„⇐“ Auch hier interessiert nur noch die multiplikative Struktur.

Die Nullteilerfreiheit von K und damit die Abgeschlossenheit von $(K \setminus \{0\}, \cdot)$ erhält man dann folgendermaßen:

$$x \cdot y = 0 \wedge x \neq 0 \Rightarrow y = x^{-1}(xy) = x^{-1} \cdot 0 = 0.$$

Die zusätzlichen Eigenschaften von K bewirken, dass $(K \setminus \{0\})$ kommutative Gruppe sein muss. □

5.8 Definition: Char K

Sei $(K, +, \cdot)$ Körper.

(a) Für $a \in K$ und $n \in \mathbb{N}$ definieren wir $n \cdot a := \underbrace{a + a + \cdots + a}_{n \text{ Summanden}} = (\underbrace{1 + 1 + \cdots + 1}_{n \text{ Summanden}})a =$ $(n \cdot 1) \cdot a$ (mit $1 \in K$).

(b) Ist $m \cdot 1 \neq 0$ für alle $m \in \mathbb{N}$, so sagen wir: K hat **Charakteristik** 0; in Zeichen: $\operatorname{Char} K = 0$.

Andernfalls nennen wir die kleinste natürliche Zahl n mit $n \cdot 1 = 0$ die Charakteristik von K, in Zeichen: $\operatorname{Char} K = n$.

Beispiel. $\mathrm{GF}(p)$ hat Charakteristik p; $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ haben Charakteristik 0.

5.9 Satz: Werte von Char K

K Körper $\Rightarrow \operatorname{Char} K = 0 \vee \operatorname{Char} K = p$ (mit geeigneter Primzahl p).

Beweis: Übungsaufgabe 5.3.

Übungsaufgaben:

Aufgabe 5.1 Sei $(R, +, \cdot)$ ein kommutativer Ring mit

$$(*) \quad x^2 = x \quad \text{für alle} \quad x \in R.$$

(Diese Eigenschaft hat z.B. jeder Potenzmengenring[39].) Zeigen Sie:
a) Für alle $x \in R$ gilt $x + x = 0$.
b) Ist $|R| > 2$, so ist $(R, +, \cdot)$ kein Integritätsbereich.
Lösungshilfe: Betrachten Sie die Terme $(a + b)^2$ und $a(a - 1)$!

Aufgabe 5.2
Zeigen Sie, dass $(\mathbb{Z}_p, +, \cdot)$ für jede Primzahl p ein Körper ist !
Lösungshinweis: Für den Nachweis der Existenz der Inversen zu $\overline{a}$ betrachtet man die injektive und damit bijektive Abbildung $\overline{x} \mapsto \overline{a} \cdot \overline{x}$.

Aufgabe 5.3 Beweisen Sie Satz 5.9 !

[39] Zur Definition des Potenzmengenringes siehe 2.2.10 und Aufgabe 5.4 !

Lösungshinweis: Zu $\operatorname{Char} K = p \cdot q$ berechnet man $(p \cdot 1)(q \cdot 1)$!

Aufgabe 5.4 Untersuchen Sie, unter welcher (notwendigen und hinreichenden) Bedingung an die nicht-leere Menge T der Potenzmengenring $(\wp(T), \bigtriangleup, \cap)$ ein Körper ist ! (Vgl. 2.2.9 d und 2.2.10 !)

Kapitel III: Vektorräume

§ 6 Vektorräume und Unterräume – Definitionen und Beispiele

In diesem Paragraphen kommen wir nun endlich zurück auf die in §1 erörterten Beispiele und behandeln diejenige Struktur, die durch die dort herausgestellten Gesetze definiert ist, für die also jene Beispiele Modelle liefern.

6.1 Definition: Vektorraum über einem Körper

Sei $(K, +, \cdot)$ ein Körper und V eine Menge. Ferner seien Verknüpfungen $\oplus : V \times V \to V$ (Addition) und $\underset{K}{\odot} : V \times K \to V$ (S-Multiplikation) definiert. Für diese seien folgende Axiome erfüllt:

(V1) $(V, \oplus)$ ist kommutative Gruppe.

(V2) Für alle $v, w \in V$, $\lambda, \mu \in K$ gelten die 'gemischten Distributivgesetze':
$(v \oplus w) \odot \lambda = (v \odot \lambda) \oplus (w \odot \lambda)$ und $v \odot (\lambda + \mu) = (v \odot \lambda) \oplus (v \odot \mu)$.

(V3) Für alle $v \in V$, für alle $\lambda, \mu \in K$ gilt das 'gemischte Assoziativgesetz':
$v \odot (\lambda \cdot \mu) = (v \odot \lambda) \odot \mu$.

(V4) Für alle $\forall v \in V$ ist $v \odot 1 = v$.

Dann heißt $(V, \oplus, \underset{K}{\odot})$ **Vektorraum** über K, K-Vektorraum (Abkürzung: K-VR) oder linearer Raum.
Da wir uns auf Vektorräume über Körpern beschränkt haben (eine analoge Definition von 'rechten' Vektorräumen über Schiefkörpern ist möglich), brauchen wir nicht streng zu unterscheiden, ob wir von links oder von rechts mit den Skalaren multiplizieren. Wir definieren

$$\lambda \odot v := v \odot \lambda.$$

Zur Schreibweise: Durch die verwendeten Zeichen $+, \oplus, \cdot, \odot$ wollen wir (aus didaktischen Gründen) vorläufig zwischen den beiden Additionen ($+$ in K, $\oplus$ in V) und den beiden Multiplikationen ($\cdot$ in K und $\odot$, der so genannten **S-Multiplikation**) unterscheiden.. Falls keine Missverständisse zu befürchten sind, schreiben wir anstelle von $\oplus$ ebenfalls $+$, anstelle von $v \odot \lambda$ nur $v \cdot \lambda$ oder $v\lambda$, verwenden also dieselben Symbole für verschiedene Verknüpfungen. Später schreiben wir statt $(V, +, \cdot)$ auch nur V und weiterhin statt $(K, +, \cdot)$ nur K. Zur Ersparnis von Klammern definieren wir, dass die S-Multiplikation mehr als die Addition binde, also dass z.B. $\lambda v + w := (\lambda v) + w$ gilt.

Beispiele:

M
↓

- Die Menge der Vektoren der Ebene mit den in (1.1) definierten Verknüpfungen (Addition, S-Multiplikation) bildet einen $\mathbb{R}$-Vektorraum $\mathcal{V}$.

Ähnlich bilden

- die Vektoren des Raumes (vgl. 1.2)
- die Lösungstripel der linearen homogen Gleichung $(*)$ von (1.3)
- die Lösungsfunktionen der linearen homogenen Differenzialgleichung von (1.4) mit den entsprechenden. Additionen bzw. S-Multiplikationen

jeweils einen $\mathbb{R}$-Vektorraum.

- Die Menge der Wörter der Länge 5 über dem Alphabet $\{0,1\}$ bildet bzgl. der in (1.5) definierten (komponentenweisen) Addition und S-Multiplikation einen Vektorraum über dem Körper GF(2).

↑ M

Fortsetzung der Beispiele von Vektorräumen:

(a) Sei $(K,+,\cdot)$ ein Körper und $n \in \mathbb{N}$. Wir definieren:

$V := K^n = \{(\xi_1, \xi_2, \ldots, \xi_n) : \xi_1, \xi_2, \ldots, \xi_n \in K\}$,

$\oplus : (\alpha_1, \alpha_2, \ldots, \alpha_n) \oplus (\beta_1, \beta_2, \ldots, \beta_n) := (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$,

$\underset{K}{\odot} : (\alpha_1, \alpha_2, \ldots, \alpha_n) \odot \lambda := (\alpha_1 \cdot \lambda, \alpha_2 \cdot \lambda, \ldots, \alpha_n \cdot \lambda)$

(also Addition und S-Multiplikation sind komponentenweise erklärt).

Dann gilt: $\boxed{(K^n, \oplus, \underset{K}{\odot}) \text{ ist ein } K\text{-Vektorraum.}}$ (1. Standardbeispiele)

Spezialfälle:

(1) $n = 1$: Jeder Körper K ist auch K-Vektorraum.

(2) $n = 2$, $K = \mathrm{GF}(2)$, $V = \{(0,0),(0,1),(1,0),(1,1)\}$

$\oplus$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,1)$	$(0,1)$	$(0,0)$	$(1,1)$	$(1,0)$
$(1,0)$	$(1,0)$	$(1,1)$	$(0,0)$	$(0,1)$
$(1,1)$	$(1,1)$	$(1,0)$	$(0,1)$	$(0,0)$

$\underset{K}{\odot}$: $(\alpha,\beta) \odot 0 := (0,0)$

$(\alpha,\beta) \odot 1 := (\alpha,\beta)$

für $\alpha, \beta \in \{0,1\}$.

(3) $n = 5$, $K = \mathrm{GF}(2)$, $V =$ Menge der Wörter der Länge 5 über $\{0,1\}$; $\oplus$, $\underset{K}{\odot}$ komponentenweise.

(4) $\mathbb{R}^2 = \{(\xi_1,\xi_2) : \xi_1, \xi_2 \in \mathbb{R}\}$; $(\xi_1,\xi_2) \oplus (\eta_1,\eta_2) := (\xi_1 + \eta_1, \xi_2 + \eta_2)$ und

$(\xi_1,\xi_2) \odot \lambda = (\xi_1 \cdot \lambda, \xi_2 \cdot \lambda)$ für alle $\xi_1, \xi_2, \eta_1, \eta_1, \lambda \in \mathbb{R}$

.

(5) Analog zu (4): $(\mathbb{R}^3, \oplus, \underset{\mathbb{R}}{\odot})$.

Anmerkung. Unter einer Familie $(\lambda_i)_{i \in I}$ über K mit Index- Menge I verstanden wir in 2.4.7 die Abbildung $I \to K$ mit $i \mapsto \lambda_i$.

Damit definieren wir nun Folgendes:

(c) Seien $(K,+,\cdot)$ ein Körper, I nicht-leere Menge sowie

$K^I := \{(\lambda_i)_{i \in I} : \lambda_i \in K \text{ für alle } i \in I\}$;

$\oplus : (\lambda_i)_{i\in I} + (\mu_i)_{i\in I} := (\lambda_i + \mu_i)_{i\in I}$;

$\underset{K}{\odot} : (\lambda_i)_{i\in I} \odot \mu := (\lambda_i \cdot \mu)_{i\in I}$.

Dann gilt $\boxed{(K^I, \oplus, \underset{K}{\odot}) \text{ ist ein } K\text{-Vektorraum.}}$

(2.Standardbeispiele; umfassen die 1.Strandardbeispiele)

Beweis. ...

Anmerkung. Jede Familie $f = (\lambda_i)_{i\in I}$ ist ja definitionsgemäß eine Abbildung von I in K (s.o.!); Umgekehrt lässt sich jede solche Abbildung als Familie schreiben. Also gilt

$$K^I = \mathrm{Abb}(I, K).$$

Für $f = (\lambda_i)_{i\in I}$ und $g = (\mu_i)_{i\in I}$ gilt dann $f \oplus g = (\lambda_i + \mu_i)_{i\in I}$ und $f \odot \mu = (\lambda_i \cdot \mu)_{i\in I}$, also (in Übereinstimmung mit früheren Definitionen):

$f \oplus g : (f \oplus g)(i) = \lambda_i + \mu_i = f(i) + g(i)$,

$f \odot \mu : (f \odot \mu)(i) = \lambda_i \cdot \mu = f(i) \cdot \mu$.

Damit lassen sich die angeführten Beispiele auch in der Form

$$\big(\mathrm{Abb}(I, K), \oplus, \underset{K}{\odot}\big)$$

schreiben.

Speziell:

- Für $I = \{1, 2, \ldots, n\}$ gilt $K^I = K^n$; da auch Addition und S-Multiplikation übereinstimmen, sind die Beispiele aus (a) auch unter den eben in (b) definierten.
- Für $I = \mathbb{N}$ und $K = \mathbb{R}$ erhält man den **Vektorraum der reellen Zahlfolgen**.

(d) Sei $(K, +, \cdot)$ Körper und I nicht-leere Menge.

$$\boxed{K^{(I)}} := \{(\lambda_i)_{i\in I} | (\lambda_i)_{i\in I} \in K^I \text{ und } \lambda_i = 0 \text{ für fast alle } i \in I\} \ ;$$

(„$\lambda_i = 0$ **für fast alle** $i \in I$“ bzw. „$\lambda_i \neq 0$ für höchstens endlich viele $i \in I$“ sagen wir, wenn $\{i \in I | \lambda_i \neq 0\}$ endlich ist; wir sprechen auch von einer **Abbildung von I in Kmit endlichem Träger**).

Addition $\oplus$ und S-Multiplikation $\underset{K}{\odot}$ definieren wir wie bei K^I (eingeschränkt auf Urbild $K^{(I)} \times K^{(I)}$ bzw. $K^{(I)} \times K$ und Bild $K^{(I)}$). Dann gilt

$\boxed{(K^{(I)}, \oplus, \underset{K}{\odot}) \text{ ist ein } K\text{-Vektorraum .}}$ Weitere Standardbeispiele

Ist speziell $|I| = n \in \mathbb{N}$, so gilt $K^{(I)} = K^I$, im Falle $I = \{1, \ldots, n\}$ also insbesondere $K^{(I)} = K^n$.

6.2 Hilfssatz: Eigenschaften der neutralen und inversen Elemente

Sei $(V, \oplus, \underset{K}{\odot})$ Vektorraum über dem Körper $(K, +, \cdot)$, bezeichne ferner
0 das neutrale Element von $(K, +)$ (Nullelement des Körpers),
1 das neutrale Element von $(K, \cdot)$ (Einselement des Körpers),
$-\lambda$ die additive Inverse von λ, $\lambda \in K$,
$\mathfrak{o}$ das neutrale Element von $(V, \oplus)$ (Nullvektor),
$-v$ das additive Inverse von v, $v \in V$ (auch Gegenvektor von v genannt.)
Dann gilt für alle $v \in V$ und alle $\lambda \in K$:

(1) $v \odot 0 = \mathfrak{o}$,
$\mathfrak{o} \odot \lambda = \mathfrak{o}$;

(2) $v \odot \lambda = \mathfrak{o} \Rightarrow v = \mathfrak{o} \vee \lambda = 0$;

(3) $(-v) \odot \lambda = -(v \odot \lambda) = v \odot (-\lambda)$;
speziell: $-v = v \odot (-1)$.

Beweis. (1) $\mathfrak{o} \oplus (v \odot 0) \underset{\text{(V1)}}{=} v \odot 0 \underset{\text{(K2)}}{=} v \odot (0+0) \underset{\text{(V2)}}{=} (v \odot 0) \oplus (v \odot 0) \underset{\text{Kürzen (V1)}}{\Rightarrow} \mathfrak{o} = v \odot 0$,

$\mathfrak{o} \oplus (\mathfrak{o} \odot \lambda) = \mathfrak{o} \odot \lambda = (\mathfrak{o} \oplus \mathfrak{o}) \odot \lambda = (\mathfrak{o} \odot \lambda) \oplus (\mathfrak{o} \odot \lambda) \Rightarrow \mathfrak{o} = \mathfrak{o} \odot \lambda$.

(2) Sei $v \odot \lambda = \mathfrak{o}$ und $\lambda \neq 0$; dann existiert λ^{-1} mit $\lambda \cdot \lambda^{-1} = 1$.

$v \underset{\text{(V4)}}{=} v \odot 1 = v \odot (\lambda \cdot \lambda^{-1}) \underset{\text{(V3)}}{=} (v \odot \lambda) \odot \lambda^{-1} \underset{\text{Vor.}}{=} \mathfrak{o} \odot \lambda^{-1} \underset{\text{(1)}}{=} \mathfrak{o}$.

(3) $[v \odot (-\lambda)] \oplus [v \odot \lambda] \underset{\text{(V2)}}{=} v \odot [(-\lambda) + \lambda] \underset{\text{(K2)}}{=} v \odot 0 \underset{\text{(1)}}{=} \mathfrak{o} \Rightarrow v \odot (-\lambda) = -(v \odot \lambda)$.

Für $\lambda = 1$ ergibt sich $-v = -(v \odot 1) = v \odot (-1)$ und damit

$v \odot (-\lambda) \underset{\substack{\text{Eigenschaft} \\ \text{von } -1 \in K}}{=} v \odot [(-1) \cdot \lambda] \underset{\text{(V3)}}{=} [v \odot (-1)] \odot \lambda = (-v) \odot \lambda$.

□

6.3 Definition: Unterraum

Sei $(V, +, \underset{K}{\cdot})$ ein K-Vektorraum und $U \subseteq V$. Dann heißt U (linearer) **Unterraum** (**UR**, **Teilraum**) von $(V, +, \underset{K}{\cdot})$, wenn $+$ bzw. $\underset{K}{\cdot}$ auf U Verknüpfungen $+_U := +|_{U \times U \to U}$ bzw. $\underset{K_U}{\cdot} = \underset{K}{\cdot}\Big|_{U \times K \to U}$ induzieren (– insbesondere also U abgeschlossen bzgl. $+$ und $\underset{K}{\cdot}$ ist –) und

$$(U, +_U, \underset{K_U}{\cdot}) \text{ ein } K\text{-Vektorraum ist.}$$

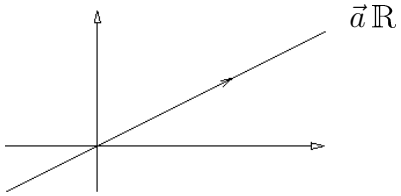
Beispiel. (a) Sei $\mathcal{V}$ der Vektorraum der Vektoren der Ebene (s. 1.1).

Für jedes $\vec{a} \in \mathcal{V} \setminus \{\vec{o}\}$ ist

$$\vec{a}\mathbb{R} := \{\vec{a}k | k \in \mathbb{R}\}$$

(s. Figur 6.1 !) ein Unterraum von $\mathcal{V}$.

Es lässt sich zeigen, dass dies die einzigen von $\mathcal{V}$ und $\{\vec{o}\}$ verschiedenen Unterräume von $\mathcal{V}$ sind.



Figur 6.1: $\vec{a}\,\mathbb{R}$ als Gerade (für $\vec{a} \neq \vec{o}$).

(b) In einem Vektorraum V sind stets V und $\{\mathfrak{o}\}$ Unterräume.

$\{\mathfrak{o}\}$ heißt Nullraum oder trivialer Unterraum. Ein von V verschiedener Unterraum von V heißt echter Unterraum von V.

(c) $K^{(I)}$ ist Unterraum des K-Vektorraumes $(K^I, +, \underset{K}{\cdot})$ (vgl. Bsp.(c) und (d) nach (6.1)), allerdings für $|I| < \infty$ kein echter Unterraum, da dann $K^{(I)} = K^I$ gilt.

Um nicht jedesmal sämtliche VR-Eigenschaften eines UR-„Kandidaten“ nachprüfen zu müssen, zeigen wir

6.4 Satz (Unterraum-Kriterium)

Sei V ein K-Vektorraum und $U \subseteq V$. Dann gilt:

$$U \text{ ist Unterraum von } V \Leftrightarrow \begin{cases} \text{(i)} & U \neq \emptyset. \quad !! \\ \text{(ii)} & U \text{ ist abgeschlossen bzgl. Addition und} \\ & \text{S-Multiplikation.} \end{cases}$$

Beweis.

„$\Rightarrow$“ U UR $\Rightarrow (U, +_U)$ Gruppe $\Rightarrow \mathfrak{o} \in U$ (also $U \neq \emptyset$) und U abgeschlossen bzgl. $+$.

Ferner ist $\underset{K_U}{\cdot} = \underset{K}{\cdot}\Big|_{U \times K \to U}$ (per definitionem), also U abgeschlossen bzgl. S-Multiplikation.

„$\Leftarrow$“ Die Einschränkung von $+$ auf $U \times U \to U$ und $\underset{K}{\cdot}$ auf $U \times K \to U$ ist wegen (ii) möglich.

(V2) bis (V4) und die Kommutativität der Addition gelten für V und damit auch für Teilmengen von V. Zu zeigen bleibt also aus (V1), dass $(U, +_U)$ Gruppe, also Untergruppe von $(V, +)$, ist:

Nach (i) ist $U \neq \emptyset$; nach (6.2)(3) gilt $-u = u \cdot (-1)$ für jedes $u \in U$, wegen (ii) also $-u \in U$ für jedes $u \in U$ und ebenfalls nach (ii) damit $U + (-U) \subseteq U$. Nach dem UG-Kriterium (3.16) folgt die Behauptung.

□

Fortsetzung der Beispiele von Unterräumen:

(d) $V = \mathbb{R}^3$; $U = \{(x, y, z) | x, y, z \in \mathbb{R} \wedge -2x + 5y + z = 0\}$ ist Unterraum von $\mathbb{R}^3$. (Vgl. dazu (1.3) !)

Beweis. $(0, 0, 0) \in U \Rightarrow U \neq \emptyset$;
$(x_1, y_1, z_1), (x_2, y_2, z_2) \in U \Rightarrow -2x_1 + 5y_1 + z_1 = 0$ und $-2x_2 + 5y_2 + z_2 = 0 \Rightarrow$
$-2(x_1+x_2)+5(y_1+y_2)+(z_1+z_2) = 0$ und $\forall \lambda \in \mathbb{R} : -2(x_1 \cdot \lambda)+5(y_1 \cdot \lambda)+(z_1 \cdot \lambda) = 0 \Rightarrow$
$(x_1, y_1, z_1)+(x_2, y_2, z_2) \in U$ und $(x_1, y_1, z_1) \cdot \lambda \in U$ für alle $\lambda \in \mathbb{R} \underset{(6.4)}{\Rightarrow}$ Behauptung. □

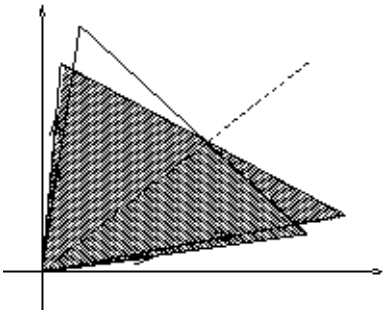
(e) $\mathcal{C}([0, 1], \mathbb{R})$ (die Menge der stetigen reellen Funktionen auf $[0, 1]$) bildet einen Unterraum von $(\mathbb{R}^{[0,1]}, +, \cdot_{\mathbb{R}})$, des Vektorraumes aller reellen Abbildungen mit Definitionsbereich $[0, 1]$.

(f) Im Vektorraum der Vektoren des Raumes (vgl. 1.2) erzeugen linear unabhängige Vektoren $\vec{a}$ und $\vec{b}$ die Menge

$$\vec{a}\,\mathbb{R} + \vec{b}\,\mathbb{R} = \{\vec{a}\,k + \vec{b}\,l \mid k, l \in \mathbb{R}\};$$

diese Mengen bilden Unterräume, ebenso $\vec{c}\,\mathbb{R} = \{\vec{c}\,k | k \in \mathbb{R}\}$ für jeden Vektor $\vec{c} \neq \vec{o}$.
Die Unterräume der ersten Art bestehen jeweils aus den Ortsvektoren einer „Ebene durch den Nullpunkt“ (s. Figur 6.2 !), die der zweiten Art aus denen einer „Nullpunktsgeraden“. Je zwei verschiedene Ebenen durch den Nullpunkt schneiden sich in einer Geraden. Allgemeiner vermuten wir, dass der Durchschnitt zweier Unterräume wieder ein Unterraum ist.
Anmerkung: $\vec{p} + \vec{m}\mathbb{R}$ ist für $\vec{p} \notin \vec{m}\mathbb{R}$ kein Unterraum.

M ↓ ↑ M



Figur 6.2: Schnitt zweier Ebenen durch den Nullpunkt

6.5 Hilfssatz: Durchschnitt von Unterräumen

Seien V ein K-Vektorraum und $(U_i)_{i \in I}$ eine nicht-leere Familie von Unterräumen von V.
Dann ist auch $\bigcap_{i \in I} U_i$ ein Unterraum von V.

Beweis. Für alle $i \in I$ gilt: $\mathfrak{o} \in U_i$, daher $\mathfrak{o} \in \bigcap_{i\in I} U_i$ und somit $\bigcap_{i\in I} U_i \neq \emptyset$.
Seien $v, w \in \bigcap_{i\in I} U_i$ und $\lambda \in K$. Dann gilt für alle $i \in I$ zunächst $v, w \in U_i$ und, da U_i Unterraum ist, auch $\left\{\begin{matrix} v+w \in U_i \\ v\lambda \in U_i \end{matrix}\right\}$, somit $v + w, v\lambda \in \bigcap_{i\in I} U_i$; nach (6.4) folgt die Behauptung. □

Eine wichtige Folgerung ist:

6.6 Korollar

Seien $(V, +, \cdot_K)$ ein K-Vektorraum und $T \subseteq V$.
Dann existiert genau ein Unterraum U_0 von V mit

(a) $T \subseteq U_0$.

(b) Für jeden Unterraum W von V mit $T \subseteq W$ gilt $U_0 \subseteq W$.

d.h. U_0 ist (eindeutig bestimmter) kleinster T enthaltender Unterraum.

6.7 Bezeichnung: Erzeugnis, Erzeugendensystem

Der Unterraum U_0 der in (6.6) aufgeführten Eigenschaften heißt der von T **erzeugte Unterraum**, von T **aufgespannte Unterraum** oder die **lineare Hülle** von T, in Zeichen $U_0 = \langle T\rangle$. T heißt **Erzeugendensystem** von $\langle T\rangle$.

Beweis von (6.6). (Allgemeineres Beweisprinzip!) Betrachte
$\mathcal{U} := \{X | X \text{ UR von } V \wedge T \subseteq X\}$! Es gilt: $V \in \mathcal{U}$, also $\mathcal{U} \neq \emptyset$. Wir definieren:

$$U_0 := \bigcap_{X\in\mathcal{U}} X \subseteq V.$$

Behauptung: U_0 hat die geforderten Eigenschaften.
Nach (6.5) ist U_0 Unterraum von V, nach Konstruktion ist $T \subseteq U_0$.
Sei W Unterraum von V mit $T \subseteq W$. Dann ist $W \in \mathcal{U}$ und damit

$$U_0 = \bigcap_{X\in\mathcal{U}} X \subseteq W.$$

Eindeutigkeit:
Seien U_1, U_2 Unterräume der Eigenschaften (a) und (b). Dann gilt $U_1 \subseteq U_2$ (mit $U_0 = U_1$ und $W = U_2$) und $U_2 \subseteq U_1$ (mit $U_0 = U_2$ und $W = U_1$), insgesamt also $U_1 = U_2$. □

M
↓

Beispiel. Sei $V = \mathcal{V}$ (vgl. (1.1)), der Vektorraum der Vektoren der Ebene.

(i) Setzen wir $T = \{\vec{a}\}$ mit $\vec{a} \in \mathcal{V} \setminus \{\vec{o}\}$, so erhalten wir
$\langle\vec{a}\rangle := \langle\{\vec{a}\}\rangle = \vec{a}\,\mathbb{R}$ (geometrische Deutung: Nullpunktsgerade in Richtung von $\vec{a}$, s.o.!).

Beweis. $\vec{a}\,\mathbb{R}$ ist Unterraum (Bsp. (a), s.o.); ferner gilt
$\forall r \in \mathbb{R} : \vec{a}r \in \langle \vec{a} \rangle$ (da mit $\vec{a}$ auch alle skalaren Vielfache von $\vec{a}$ in einem Unterraum liegen). $\square$

(ii) Sind $\vec{a}_1$ und $\vec{a}_2$ linear unabhängige Vektoren aus $\mathcal{V}$, so gilt: $\langle \vec{a}_1, \vec{a}_2 \rangle := \langle \{\vec{a}_1, \vec{a}_2\} \rangle = \vec{a}_1\mathbb{R} + \vec{a}_2\mathbb{R} = \mathcal{V}$.

(iii) Es gilt allgemein: $\langle \emptyset \rangle = \{\mathfrak{o}\}$.

↑ M

Im Folgenden werden wir versuchen, auch allgemein einen Zusammenhang zwischen dem Erzeugnis von T und den Linearkombinationen von Elementen von T herzustellen und so $\langle T \rangle$ direkter zu beschreiben.

6.8 Definition: Linearkombination

Sei V ein K-Vektorraum!

(a) Seien $v_1, \ldots, v_n \in V$ und $n \in \mathbb{N}$. Dann heißt jeder Vektor

$$v = \sum_{i=1}^{n} v_i \lambda_i \quad (\text{ mit } \lambda_i \in K \quad \text{für } i \in \{1, \ldots, n\})$$

eine **Linearkombination** (LK) der Vektoren $v_1, \ldots, v_n$.

(b) Sei $\emptyset \neq T \subseteq V$. Unter einer Linearkombination von (Elementen von) T versteht man eine Linearkombination von *endlich vielen* Vektoren $v_1, \ldots, v_n \in T$.

(c) Für $T \subseteq V$ definieren wir:
$\mathrm{LK}(T) := \{v \in V | v \text{ Linearkombination von } T\}$, falls $T \neq \emptyset$, und
$\mathrm{LK}(\emptyset) := \{\mathfrak{o}\}$

Beispiele

(a) Jeder Vektor von $\mathcal{V}$ aus (1.1) ist Linearkombination von $\vec{e}_1$ und $\vec{e}_2$.

(b) Sei K Körper und $V = K^n$; mit

$$\begin{aligned} e_1 &:= (1, 0, 0, \ldots, 0) \\ e_2 &:= (0, 1, 0, \ldots, 0) \\ &\vdots \\ e_n &:= (0, 0, \ldots, 0, 1) \end{aligned}$$

gilt $V = LK(\{e_1, \ldots, e_n\})$ wegen $(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} e_i \xi_i$.

(c) Sei K Körper und $\mathcal{P}(K)$ die Menge der Polynomabbildungen $f = \sum_{i=0}^{n} \alpha^i (\mathrm{id}_K)^i$. Bei der Einführung von $\mathcal{P}(K)$ (§5 Bsp. (d)) haben wir, u.a. eine Addition $+$ auf $\mathcal{P}(K)$ definiert: $f+g: \ (f+g)(x) := f(x)+g(x)$. Eine skalare Multiplikation $\underset{K}{\cdot}$ ist gegeben durch die Festsetzung

$$\alpha f = \alpha \underset{K}{\cdot} f : \begin{cases} K \to K \\ x \mapsto \alpha \cdot f(x). \end{cases}$$

Es gilt also $\alpha \underset{K}{\cdot} f = f_\alpha \cdot f$ sowie $(\alpha \underset{K}{\cdot} f)(x) = \alpha \cdot \sum_{i=0}^{n} \alpha_i (\mathrm{id}_K)^i(x) = \alpha \cdot \sum_{i=0}^{n} \alpha_i x^i = \sum_{i=0}^{n} \alpha\alpha_i x^i$ und $f = \sum_{i=0}^{n} f_{\alpha_i} (\mathrm{id}_K)^i = \sum_{i=0}^{n} \alpha_i \underset{K}{\cdot} (\mathrm{id}_K)^i$.

Man kann zeigen, dass $(\mathcal{P}(K), +, \underset{K}{\cdot})$ ein K-Vektorraum ist.

Es gilt: $\mathcal{P}(K) = \mathrm{LK}\{(\mathrm{id}_K)^i \mid i \in \mathbb{N}_0\}$ (wobei das Produkt von Funktionen wieder argumentweise zu bilden ist).

(c') Für den Vektorraum $K[X] := \{\sum_{i=0}^{n} a_i X^i \mid n \in \mathbb{N}_0, \quad a_i \in K\}$ der Polynome (mit geeigneter Addition und S-Multiplikation, s.LINA II) gilt $K[X] = LK\{X^i \mid i \in \mathbb{N}_0\}$.

6.9 Satz (Darstellung des Erzeugnisses)

Sei V ein K-Vektorraum und $T \subseteq V$. Dann gilt:

$$\langle T \rangle = \mathrm{LK}(T).$$

Beweis. 1. Fall $T = \emptyset$

$\langle \emptyset \rangle = \bigcap_{\substack{U \text{ UR} \\ \text{von } V}} U = \{o\} = \mathrm{LK}(\emptyset)$, da $\{o\}$ Unterraum ist und $\emptyset \subseteq U$ für jeden Unterraum U von V gilt.

2. Fall $T \neq \emptyset$

(a) $\underline{T \subseteq \mathrm{LK}(T).}$

$t \in T \Rightarrow t = t \cdot 1 \in \mathrm{LK}(T)$.

(b) $\underline{\mathrm{LK}(T) \text{ ist Unterraum von } V.}$

Nach (a) ist $\mathrm{LK}(T) \supseteq T \neq \emptyset$.

Wir zeigen die Abgeschlossenheit von $\mathrm{LK}(T)$ bzgl. Addition und S-Multiplikation:

Seien $v, w \in \mathrm{LK}(T)$ und $\lambda \in K$. Nach Definition (6.8) (b) haben v und w dann folgende Darstellung:

$$v = \sum_{i=1}^{r} v_i\lambda_i \text{ und } w = \sum_{j=1}^{s} w_j\mu_j \text{ mit geeigneten } r, s \in \mathbb{N},$$
$$v_1, \ldots, v_r, w_1, \ldots, w_s \in T,\ \lambda_1, \ldots, \lambda_r, \mu_1, \ldots, \mu_s \in K.$$

Damit gilt

$$v + w = (\sum_{i=1}^{r} v_i\lambda_i + \sum_{j=1}^{s} w_j\mu_j) \in \mathrm{LK}(T)$$

($v_1, \ldots, v_r, w_1, \ldots, w_s$ sind ebenfalls nur endlich viele Vektoren aus T) sowie

$$v \cdot \lambda = (\sum_{i=1}^{r} v_i\lambda_i) \cdot \lambda = \sum_{i=1}^{r} v_i(\lambda_i\lambda) \in \mathrm{LK}(T).$$

Nach dem UR-Kriterium (6.4) folgt die Zwischenbehauptung.

(c) $\underline{\langle T\rangle \subseteq \mathrm{LK}(T)\ .}$

Diese Tatsache folgt aus (a) und (b), da $\mathrm{LK}(T)$ unter den Unterräumen vorkommt, über die der Durchschnitt gebildet wird.

(d) $\underline{\mathrm{LK}(T) \subseteq \langle T\rangle\ .}$

Sei $v \in \mathrm{LK}(T)$; dann existieren $n \in \mathbb{N}$, $v_1, \ldots, v_n \in T$, $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \sum_{i=1}^{n} v_i\lambda_i$. Da $v_i \in T$ für $i = 1, \ldots, n$ und $T \subseteq \langle T\rangle$ (vgl. 6.6 !) ist, gilt nach (6.4): $v_i\lambda_i \in \langle T\rangle$ für $i = 1, \ldots, n$ und damit $v = \sum_{i=1}^{n} v_i\lambda_i \in \langle T\rangle$.

Damit ist die „doppelte Inklusion“ gezeigt.

□

§ 7 Lineare Unabhängigkeit, Basis – Existenzsatz

M
↓

Am Ende des vorigen Paragraphen betrachteten wir bei vorgegebener Teilmenge T eines K-Vektorraumes V das Erzeugnis U von T in V. Die Bildung des Erzeugnisses ist dabei ein wichtiges Prinzip, das auch in anderen Gebieten der Algebra in ähnlicher Weise herangezogen wird. Die Information über das Erzeugnis ist dabei jeweils schon ganz in dem Erzeugendensystem enthalten. In unserem Falle lässt sich $U = \langle T\rangle$ mit Hilfe von Linearkombinationen von T beschreiben.
Wir verlagern entsprechend unseren Standpunkt: Wir geben den Unterraum U vor und fragen nach erzeugenden Mengen von U. Zunächst stellen wir fest:

$$U = \langle U\rangle,$$

so dass also jeder Unterraum als Erzeugnis (seiner selbst) geschrieben werden kann. Es ist aber, zweckmäßig, das **Erzeugendensystem** T von U **möglichst klein auszuwählen**, also ohne solche Vektoren, die sich aus den übrigen durch Linearkombination ergeben.
Diese treten auf, wenn es $v_1, \ldots, v_n \in T$ gibt mit

$$\sum_{i=1}^{n} v_i\lambda_i = \vec{o} \text{ und (z.B.) } \lambda_1 \neq 0.$$

Denn dann ist v_1 selbst Linearkombination der übrigen v_i:

$$v_1 = -(\sum_{i=2}^{n} v_i\lambda_i)\lambda_1^{-1} = \sum_{i=2}^{n} v_i(-\lambda_i\lambda_1^{-1}).$$

↑
M

7.1 Definition: Lineare Unabhängigkeit, lineare Abhängigkeit

(a) (i) Eine **endliche Teilmenge** T von V heißt **linear abhängig** (lin.abh., l.a.) g.d.w. gilt: Es existieren Elemente $\lambda_1, \ldots, \lambda_n \in K$, nicht alle $\lambda_i = 0$, und $v_1, \ldots, v_n \in T$ mit $v_i \neq v_j$ für $i \neq j$ derart, dass $\sum_{i=1}^{n} v_i\lambda_i = \vec{o}$. (Es existiert dann also eine „nicht-triviale Darstellung des Nullvektors“.)

(ii) Das n-Tupel von Vektoren $(v_1, \ldots, v_n) \in V^n$ heißt **linear abhängig** g.d.w. gilt:

$$\{v_1, \ldots, v_n\} \text{ lin.abh. } \vee \ v_1, \ldots, v_n \text{ nicht paarweise verschieden.}$$

(In diesem Falle sagen wir auch: die Vektoren $v_1, \ldots, v_n$ sind linear abhängig.)

(iii) v_1 heißt linear abhängig von $v_2, \ldots, v_n$, wenn $v_1 \in \mathrm{LK}(\{v_2, \ldots, v_n\})$ gilt; dann sind insbesonder $v_1, v_2, \ldots v_n$ linear abhängig.

(iv) Eine **Teilmenge** M von V (nicht notwendig endlich) heißt **linear abhängig**, wenn gilt: Es existiert eine endliche Teilmenge von M, die linear abhängig ist.

(b) **linear unabhängig** (lin.unabh., l.u.) $\equiv$ nicht linear abhängig, also für $M \subseteq V$:

$$M \text{ lin.unabh. Menge } \Leftrightarrow \text{Jede } \textbf{endliche} \text{ Teilmenge von } M \text{ ist lin.unabh. .}$$

und:

$$v_1, \ldots, v_n \text{ lin.unabh.} :\Leftrightarrow v_1, \ldots, v_n \text{ nicht lin.abh.}$$
$$\Leftrightarrow \{v_1, \ldots, v_n\} \text{ lin.unabh. und } v_1, \ldots, v_n \text{ paarweise versch.}$$
$$\Leftrightarrow \left[\sum_{i=1}^{n} v_i\lambda_i = \vec{o} \Rightarrow \lambda_i = 0 \text{ für alle } i \in \{1, \ldots, n\}\right]$$

7.2 Beispiele

(a) $V = \mathbb{R}^2$

(i) $(1,0)$ und $(2,2)$ sind linear unabhängig.

Beweis. Zu zeigen ist $(1,0) \cdot \lambda + (2,2) \cdot \mu = 0 \Rightarrow \lambda = \mu = 0$ für alle $\lambda, \mu \in \mathbb{R}$.

$(1,0)\lambda + (2,2)\mu = (0,0) \Rightarrow (\lambda, 0) + (2\mu, 2\mu) = (0,0) \Rightarrow (\lambda + 2\mu, 2\mu) = (0,0) \Rightarrow$
$\lambda + 2\mu = 0 \wedge \mu = 0 \Rightarrow \lambda = \mu = 0.$ □

(ii) $(1,0), (2,2), (-2,1), (1,\sqrt{2})$ sind linear abhängig.

Beweis. $(1,0) \cdot \underline{3} + (2,2) \cdot \underline{(-\frac{1}{2})} + (-2,1) \cdot \underline{1} + (1,\sqrt{2}) \cdot 0 = \vec{o}$ □

(Aus dem Beweis sehen wir, dass sogar $(1,0)$, $(2,2)$, $(-2,1)$ linear abhängig sind).

(b) $\boxed{V = K^n}$, K Körper; $e_i := (0, 0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \ldots, 0)$ (vgl. Beispiel (b) nach (6.9)

(i) $\{e_1, \ldots, e_n\}$ ist linear unabhängig.

Beweis. $\sum_{i=1}^{n} e_i\lambda_i = \vec{o} \Rightarrow (\lambda_1, \ldots, \lambda_n) = \vec{o} \Rightarrow \forall i = 1, \ldots, n : \lambda_i = 0.$ □

(ii) Die Menge $\{(\lambda_1, \ldots, \lambda_n), e_1, \ldots, e_n\}$ ist linear abhängig für jedes n-Tupel $(\lambda_1, \ldots, \lambda_n) \in K^n \setminus \{e_1, \ldots, e_n\}$.

Beweis. $(\lambda_1, \ldots, \lambda_n) \cdot (-1) + \sum_{i=1}^{n} e_i\lambda_i = \vec{o}$ und $-1 \neq 0.$ □

Anmerkung. $(\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} e_i\lambda_i \in \mathrm{LK}(\{e_1, \ldots, e_n\})$

(c) Verallgemeinerung von (b):

In $\boxed{K^I}$ definieren wir e_i für $i \in I$ durch

$$e_i := (\delta_{ij})_{j\in I} : \begin{cases} I \to K, \\ j \mapsto \delta_{ij}. \end{cases} \qquad \text{wobei} \quad \delta_{ij} = \begin{cases} 0 \in K & \text{für } i \neq j \\ 1 \in K & \text{für } i = j \end{cases}.$$

Dann ist $\{e_i | i \in I\}$ linear unabhängig.

Beweis. Für jede endliche Teilmenge $\{e_{i_1}, \ldots, e_{i_n}\}$ folgt aus $\sum\limits_{k=1}^{n} e_{i_k}\lambda_k = \vec{o}$ zunächst $\vec{o} = \sum\limits_{k=1}^{n} (\delta_{i_k j})_{j\in I} \cdot \lambda_k = \left(\sum\limits_{k=1}^{n} \delta_{i_k j}\lambda_k\right)_{j\in I}$, also $\sum\limits_{k=1}^{n} \delta_{i_k j}\lambda_k = 0$ für jedes $j \in I$. Wählen wir für $l \in \{1, \ldots, n\}$ nun $j = i_l$, so erhalten wir daraus und aus $\delta_{i_k j} = 0$ für $i_k \neq i_l = j$ sofort $\delta_{i_l i_l}\lambda_l = 0$, also $\lambda_l = 0$. □

(d) In $\boxed{\mathcal{P}(\mathbb{R})}$, dem Vektorraum der Polynomabbildungen über $\mathbb{R}$, ist $\{(\mathrm{id}_{\mathbb{R}})^i | i \in \mathbb{N}_0\}$ linear unabhängig.

Beweis. Wäre die Menge linear abhängig, so gäbe es eine linear abhängige endliche Teilmenge, insbesondere ein $m \in \mathbb{N}_0$ und Elemente $\lambda_0, \ldots, \lambda_m \in K$ mit

$$\vec{o} = \sum_{i=0}^{m} (\mathrm{id})^i \lambda_i \quad \text{und} \quad \lambda_m \neq 0.$$

Für alle $x \in \mathbb{R}$ gilt somit $\sum\limits_{i=0}^{m} x^i \lambda_i = 0$. Durch m-maliges Differenzieren erhält man

$$\lambda_m \cdot m! = 0. \quad \lightning \quad \text{(Widerspruch)}$$

□

(d') $\{X^i \mid i \in \mathbb{N}_0\}$ ist linear unabhängige Teilmenge von $V = K[X]$.

Anmerkung. In $\mathcal{P}\big(\mathrm{GF}(2)\big)$ sind schon $(\mathrm{id}_{\mathrm{GF}(2)}), (\mathrm{id}_{\mathrm{GF}(2)})^2$ linear abhängig:
$\big[\mathrm{id}_{\mathrm{GF}(2)} + (\mathrm{id}_{\mathrm{GF}(2)})^2\big](x) = x + x^2 = 0$ für alle $x \in \mathrm{GF}(2)$.
Es gilt sogar: $\mathrm{id}_{\mathrm{GF}(2)} = (\mathrm{id}_{\mathrm{GF}(2)})^2$.

7.3 Hilfssatz: Teilemengen linear unabhängiger Mengen

Sei V ein K-Vektorraum.

(i) Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

(ii) Jede Obermenge einer linear abhängigen Menge ist linear abhängig.

Beweis. Die Behauptung folgt unmittelbar aus Definition (7.1). □

Anmerkung. Insbesondere gilt: $\boxed{\emptyset \text{ ist linear unabhängig}}$!
Nun können wir auf Erzeugendensysteme zurückkommen.

7.4 Satz: Entbehrliche Elemente in linear abhängigen Mengen

Voraussetzung: V K-Vektorraum, $\emptyset \neq M \subseteq V$.
Behauptung: $\boxed{M \text{ lin.abh. } \Leftrightarrow \exists a \in M : \langle M \setminus \{a\}\rangle = \langle M\rangle}$.

Insbesondere sind in einem linear abhängigen Erzeugendensystem Elemente entbehrlich.

Beweis. „$\Rightarrow$“ M lin.abh. $\Rightarrow \exists v_1, \ldots, v_n \in M,\ \lambda_1, \ldots, \lambda_n \in K : \sum_{i=1}^{n} v_i\lambda_i = \vec{o}$ und $\lambda_k \neq 0$ für mindestens ein $k \in \{1, \ldots, n\}$; für diese Elemente folgt $v_k\lambda_k = -\sum_{\substack{i=1\\ i\neq k}}^{n} v_i\lambda_i$, also

$v_k = -\big(\sum_{\substack{i=1\\ i\neq k}}^{n} v_i\lambda_i\big)\lambda_k^{-1} = \sum_{\substack{i=1\\ i\neq k}}^{n} v_i(-\lambda_i\lambda_k^{-1})$, d.h. $v_k \in \langle M \setminus \{v_k\}\rangle$ und $\langle M\rangle = \langle M \setminus \{v_k\}\rangle$.

„$\Leftarrow$“ $a \in \langle M\rangle = \langle M \setminus \{a\}\rangle \Rightarrow a \in \mathrm{LK}(M \setminus \{a\}) \Rightarrow$
$\exists\, b_1, \ldots, b_n \in M \setminus \{a\},\ \lambda_1, \ldots, \lambda_n \in K : a = \sum_{i=1}^{n} b_i\lambda_i \Rightarrow a \cdot 1 + \sum_{i=1}^{n} b_i(-\lambda_i) = \vec{o} \Rightarrow$
$\{a, b_1, \ldots, b_n\}$ l.a. $\Rightarrow M$ lin.abh.

□

Im Folgenden interessieren wir uns hauptsächlich für linear unabhängige Erzeugendensysteme.

7.5 Definition: Basis

Sei V ein K-Vektorraum und $B \subseteq V$. Dann heißt B **Basis** von V g.d.w. B linear unabhängiges Erzeugendensystem ist, also gilt:

(1) B ist linear unabhängig,

(2) $V = \langle B\rangle$.

7.6 Beispiele

(a) In K^n ist $B = \{e_1, \ldots, e_n\}$ mit $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)$ eine Basis, die sogenannte **kanonische Basis** von K^n. Denn B ist linear unabhängig (Bsp. (b) s.o.) und jedes $x = (\xi_1, \ldots, \xi_n)$ aus K^n ist Linearkombination von B: $x = \sum_{i=1}^{n} e_i\xi_i \in \langle B\rangle$.

(b) Verallgemeinerung:

In $K^{(I)}$ ist $B = \{e_i | i \in I\}$ mit $e_i = (\delta_{ij})_{j\in I}$ Basis:

B ist linear unabhängige Teilmenge von K^I (s. Beispiel (c), s.o.) und damit von $K^{(I)}$. Ist $x = (\lambda_i)_{i\in I} \in K^{(I)}$, so gibt es eine endliche Teilmenge J von I mit $\lambda_i = 0$ für

$i \in I \setminus J$; damit gilt $x = \sum_{i \in J} e_i \lambda_i$ (wobei nur endlich viele Summanden auftreten).[40]
Also $K^{(I)} = \langle B \rangle$.

Anmerkung. Ist I unendlich, so ist B keine Basis von K^I. Zum Beispiel lässt sich die konstante Folge $(a_i)_{i \in \mathbb{N}}$ in $K^{\mathbb{N}}$ mit $a_i = a \neq 0$ nicht als Summe von e_i mit endlich vielen Summanden darstellen.

(c) $\{\vec{n}_1, \vec{n}_2\}$ ist (für 'orthogonale Einheitsvektoren' $\vec{n_1}, \vec{n_2}$) eine Basis von $\mathcal{V}$ (dem $\mathbb{R}$-Vektorraum der Ebene), ebenso $\{\vec{a}, \vec{b}\}$, wenn $\vec{a}, \vec{b}$ linear unabhängige Vektoren sind; (vgl. (1.1)!)

(d) $\emptyset$ ist Basis des Nullraumes eines Vektorraums; denn:

$\emptyset$ ist lin.unabh. (es existiert keine lin.abh. Teilmenge von $\emptyset$) und $\langle \emptyset \rangle = \{\vec{o}\}$, s.o.!

(e) $B = \{\mathrm{id}_{\mathbb{R}}^i \,|\, j \in \mathbb{N}_0\}$ ist Basis von $\mathcal{P}(\mathbb{R})$. (B ist lin.unabh. gemäß Beispiel (d) nach (7.1); $\mathcal{P}(\mathbb{R}) = \langle B \rangle$ gemäß Bsp. (c) nach (6.9)

(e') $B = \{X^i | i \in \mathbb{N}_0\}$ ist Basis von $K[X]$, dem Vektorraum der Polynome über K.

7.7 Satz (Charakterisierung von Basen)

Sei V ein K-Vektorraum. Dann sind folgende Aussagen äquivalent:

(a) B ist **Basis** von V; (d.h. B ist lin.unabh. Erzeugendensystem).

(b) B ist eine **maximale linear unabhängige Teilmenge** von V
(d.h. B lin.unabh. und für alle $x \in V \setminus B : B \cup \{x\}$ lin.abh.)

(c) B ist ein **minimales Erzeugendensystem** von V
(d.h $V = \langle B \rangle \wedge \forall x \in B : \langle B \setminus \{x\} \rangle \subsetneqq V$.)

(d) **Jeder Vektor** $v \in V$ lässt sich (abgesehen von Summanden der Form $b_j \cdot 0$, Reihenfolge und Aufspalten von Summanden) **auf genau eine Weise als Linearkombination von B** darstellen.

Beweis.

(a)⇒(b) Ist B Basis, so definitionsgemäß B lin.unabh. $\wedge V = \langle B \rangle \underset{(6.10)}{=} \mathrm{LK}(B)$. Sei $x \in V \setminus B$, dann $x \in \mathrm{LK}(B)$, also $x \cdot 1 = \sum b_i \lambda_i$ und $x \notin B$, und daher $B \cup \{x\}$ lin.abh..

(b)⇒(a) B maximal linear unabhängige Teilmenge ⇒
(B linear unabh. $\wedge \forall x \in V \setminus B : B \cup \{x\}$ lin.abh.) ⇒ (B linear unabhängig und $x \in \mathrm{LK}(B)$ für alle $x \in V$) ⇒ (B lin.unabh. $\wedge V = \mathrm{LK}(B) \underset{(6.10)}{=} \langle B \rangle$).

[40] Man schreibt dann auch $x = \sum_{i \in I} e_i \lambda_i$, wobei zu beachten ist, dass nur endlich viele der formal auftretenden Summanden von $\vec{o}$ verschieden sind.

(a)⇔(c) Nach Kontraposition von Satz 7.4 ist ein lin.unabh. Erzeugendensystem minimal und umgekehrt.

(a)⇒(d) Sei B Basis und $v \in V$. Dann gilt $v \in \langle B\rangle = \mathrm{LK}(B)$. Wir zeigen die Eindeutigkeit der Darstellung von v als Linearkombination von B.

Seien $v = \sum_{i=1}^{n} b_i\lambda_i$ und $v = \sum_{j=1}^{m} c_j\mu_j$ zwei solche Darstellungen (mit $n, m \in \mathbb{N}$, $b_i, c_j \in B$, $\lambda_i, \mu_j \in K$ und paarweise verschiedenen b_i bzw. c_j).

Durch Addieren von Summanden $b_i\mu_i$ mit $\mu_i = 0$ zur zweiten Darstellung im Falle, dass b_i nicht unter den c_j vorkommt, und von Summanden $c_j\lambda_j$ mit $\lambda_j = 0$ zur ersten Darstellung, wenn c_j nicht unter den b_i vorkommt, sowie durch Änderung der Nummerierung (falls $b_i = c_j$) und der Bezeichnungen (z.B. $b_{n+1} := c_1$, falls $c_1 \notin \{b_1, \ldots, b_n\}$ usw.), erhält man

$$v = \sum_{k=1}^{l} b_k\lambda_k \quad \text{und} \quad v = \sum_{k=1}^{l} b_k\mu_k.$$

Zu zeigen ist nun: $\lambda_k = \mu_k$ f.a. $k \in \{1, \ldots, l\}$. Zunächst erhält man

$$\vec{o} = v - v = \sum_{k=1}^{l} b_k\lambda_k - \sum_{k=1}^{l} b_k\mu_k = \sum_{k=1}^{l} b_k(\lambda_k - \mu_k).$$

Als endliche Teilmenge von B ist $\{b_1, \ldots, b_l\}$ linear unabhängig. Damit gilt $\lambda_k - \mu_k = 0$ f.a. $k \in \{1, \ldots, l\}$, was zu zeigen war.

(d)⇒(a) $V = \mathrm{LK}(B) \underset{(6.10)}{\Rightarrow} V = \langle B\rangle$.

Wäre B linear abhängig, so ließe sich $\vec{o}$ auf verschiedene Weisen als Linearkombination von B darstellen.(Wieso ?)

□

7.8 Definition von Koordinaten

Sei V ein K-Vektorraum, der eine endliche Basis $\overline{B} = \{b_1, \ldots, b_n\}$ besitzt. Durch die Nummerierung der Elemente von $\overline{B}$ legen wir eine Reihenfolge fest. In diesem Fall sprechen wir von der **geordneten Basis** $B = (b_1, \ldots, b_n)$.
Nach (7.7) lässt sich jedes $x \in V$ auf genau eine Weise in der Form

$$x = \sum_{i=1}^{n} b_i\xi_i$$

darstellen. ξ_i heißt dann die i-te Koordinate von x bzgl. $B = (b_1, \ldots, b_n)$.

Schreibweise: $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B$.

Es gilt

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B + \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}_B = \sum b_i \xi_i + \sum b_i \eta_i = \sum b_i (\xi_i + \eta_i) = \begin{pmatrix} \xi_1 + \eta_1 \\ \vdots \\ \xi_n + \eta_n \end{pmatrix}_B$$

Entsprechend wird die S-Multiplikation komponentenweise ausgeführt.

Anmerkung. Die Zuordnung

$$M_B : \begin{cases} V \to K^n \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B \mapsto \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \end{cases}$$

ist ein Isomorphismus von $(V, +)$ auf $(K^n, +)$, der mit der S-Multiplikation verträglich ist, also eine Bijektion mit

$$M_B(v + w) = M_B(v) + M_B(w) \text{ für alle } v, w \in V$$

und

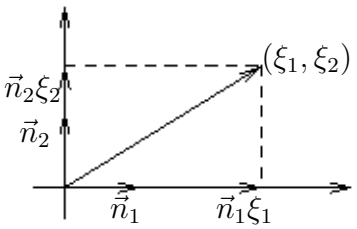
$$M_B(v \cdot \lambda) = M_B(v) \cdot \lambda \text{ für alle } v \in V \text{ und } \lambda \in K$$

(VR-Isomorphismus s.u.).

$(\xi_1, \ldots, \xi_n)$ oder auch $M_B(x) := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ heißt **Koordinatenvektor** von $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B$

Beispiele

(1) $\mathcal{V}$ hat die geordnete Basis $(\vec{n}_1, \vec{n}_2)$. Jedem $\vec{x} \in \mathcal{V}$ mit $\vec{x} = \vec{n}_1\xi_1 + \vec{n}_2\xi_2 = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}_{\vec{n}_1, \vec{n}_2}$ ist durch $M_{(\vec{n}_1, \vec{n}_2)}$ das Paar $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ bzw. $(\xi_1, \xi_2) \in \mathbb{R}^2$ zugeordnet; dies ist auch das Koordinatenpaar des Punktes mit Ortsvektor $\vec{x}$. (Siehe Figur 7.1 !)



Figur 7.1: Koordinaten eines Vektors der Ebene.

(2) Sei $V = K^n$, $B = (e_1, \ldots, e_n)$. Dann gilt $(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_B$.

Eine Verallgemeinerung auf Vektorräume, die eine unendliche Basis besitzen, ist mit Koordinaten-Familien möglich.
Es erhebt sich die Frage nach der Existenz einer Basis in einem beliebigen Vektorraum. Zunächst betrachten wir Vektorräume, die ein endliches Erzeugendensystem besitzen:

7.9 Satz (Basis-Ergänzungs-Satz bei endlichem Erzeugendensystem)

Ist V endlich erzeugter Vektorraum, A linear unabhängige Teilmenge und S ein endliches Erzeugendensystem von V mit $A \subseteq S$, dann existiert eine Basis B von V mit $A \subseteq B \subseteq S$.

7.10 Korollar (Basis-Existenz-Satz für endlich erzeugte Vektorräume)

Jeder endlich erzeugte Vektorraum V besitzt eine Basis.

Beweis von (7.10): Man setze $A = \emptyset$, wähle S als endliches Erzeugendensystem von V und wende (7.9) an! □

Beweis von (7.9): Sei S endliches Erzeugendensystem von V, A linear unabhängige Teilmenge von S. Betrachte

$$\mathfrak{X} := \{C \mid A \subseteq C \subseteq S \wedge C \text{ linear unabhängig}\} \text{ !}$$

Da S endlich ist, enthält S nur endlich viele Teilmengen; somit ist $\mathfrak{X}$ endlich. Daher existiert bzgl. der Ordnung "$\subseteq$" mindestens ein maximales Element B in $\mathfrak{X}$. Laut Definition ist $A \subseteq B$. Ist $x \in S$ und $x \notin B$, dann gilt $B \cup \{x\} \subseteq S$ und wegen der Maximalität von B dann $B \cup \{x\} \notin \mathfrak{X}$; daraus folgt die Lineare Abhängigkeit von $B \cup \{x\}$. Da B linear unabhängig ist, erhält man $x \in < B >$. Da x beliebig in $S \setminus B$ gewählt war, ergibt sich $S \subseteq < B >$ und daraus $V = < S > = < B >$. Also ist B Basis. □

Anmerkung: Man kann zeigen, dass die Existenz einer Basis für jeden $K-$Vektorrraum äquivalent zur Gültigkeit des *Auswahlaxioms* bzw. zum *Zornschen Lemma* ist. Wir zeigen hier unter Voraussetzung des Zornschen Lemmas:

7.11 Satz (Basis-Ergänzung und Basis-Existenz allgemein)

Voraussetzung: "Zermelo-Fraenkelsches Axiomensystem + Auswahlaxiom" (ZF+AC bzw. ZFC) bzw. Zornsches Lemma"

(a) Ist V Vektorraum, A linear unabhängige Teilmenge und S ein Erzeugendensystem von V mit $A \subseteq S$, dann existiert eine Basis B von V mit $A \subseteq B \subseteq S$.
(b) Jeder Vektorraum V besitzt eine Basis.

Beweis von (7.11 (b)): (analog zu 7.10): Man setze $A = \emptyset$ und $S = V$ und wende (7.11 (a)) an! □

Beweis von (7.11 (a)): (a) Definition: Sei

$$\mathfrak{X} := \{C | A \subseteq C \subseteq S \wedge C \text{ lin.unabh.}\} \subseteq \wp(V).$$

Wegen $A \in \mathfrak{X}$ ist $\mathfrak{X} \neq \emptyset$ und daher $(\mathfrak{X}, \subseteq)$ nicht-leere geordnete Menge.

(b) Zwischenbehauptung: $(\mathfrak{X}, \subseteq)$ ist induktiv geordnet (vgl. Def. 2.7.3):

Sei $\mathfrak{T} \subseteq \mathfrak{X}$ nicht leer und total geordnet, eine sogenannte **Kette**! Wir wollen zeigen, dass $C_{\mathfrak{T}} := \bigcup_{C \in \mathfrak{T}} C$ obere Schranke von $\mathfrak{T}$ in $(\mathfrak{X}, \subseteq)$ ist, also $(\mathfrak{X}, \subseteq)$ induktiv geordnet.

(i) Da konstruktionsgemäß $\forall C \in \mathfrak{T} : C \subseteq C_{\mathfrak{T}}$ gilt, ist $C_{\mathfrak{T}}$ obere Schranke; es reicht daher der Nachweis von $C_{\mathfrak{T}} \in \mathfrak{X}$.

(ii) Laut Definition gilt $\forall C \in \mathfrak{T} : A \subseteq C \subseteq S$ und daher $A \subseteq \bigcup_{C \in \mathfrak{T}} C = C_{\mathfrak{T}} \subseteq S$.

(iii) Behauptung: $C_{\mathfrak{T}}$ ist lin.unabh..
Sei $U = \{c_1, \ldots, c_n\}$ endliche Teilmenge von $C_{\mathfrak{T}} = \bigcup_{C \in \mathfrak{T}} C$.
Dann gilt: $\forall i \in \{1, \ldots, n\} \exists C_i \in \mathfrak{T} : c_i \in C_i$. Da $\mathfrak{T}$ total geordnet ist, existiert ein größtes Element C_k in $\{C_1, \ldots, C_n\}$; damit gilt $C_i \subseteq C_k$ für $i \in \{1, \ldots, n\}$ und $U \subseteq \bigcup_{i=1}^{n} C_i = C_k$. C_k ist linear unabhängig, da $C_k \in \mathfrak{T}$. Daher ist U als endliche Teilmenge linear unabhängig. Also: Jede endliche Teilmenge von $C_{\mathfrak{T}}$ ist linear unabhängig, d.h. $C_{\mathfrak{T}}$ ist linear unabhängig.

Aus (i), (ii) und (iii) folgt: $C_{\mathfrak{T}}$ ist obere Schranke von $\mathfrak{T}$ in $\mathfrak{X}$. $(\mathfrak{X}, \subseteq)$ erfüllt daher die Voraussetzung des **Zornschen Lemmas** (2.7.4); dessen Anwendung liefert:

(c) Es existiert ein maximales Element B in $\mathfrak{X}$.

(d) Behauptung: B ist Basis der geforderten Eigenschaft.

(i) Wegen $B \in \mathfrak{X}$ ist B linear unabhängig und $A \subseteq B \subseteq S$.

(ii) Ist $B = S$, so sind wir fertig. Sei andernfalls $s \in S \setminus B$! Dann gilt $B \cup \{s\} \supsetneq B$. Da B maximales Element von $\mathfrak{X}$ ist, gilt $B \cup \{s\} \notin \mathfrak{X}$. Aus $A \subseteq B \cup \{s\} \subseteq S$ ergibt sich daher $B \cup \{s\}$ als linear abhängig, also $s \in \langle B \rangle$. Mit $S \subseteq \langle B \rangle$ erhalten wir $V = \langle S \rangle \subseteq \langle B \rangle$ und daraus $V = \langle B \rangle$. □

§ 8 Dimension und Isomorphie von Vektorräumen

M ↓

Bei den Vektorräumen der Beispiele (1.1)und (1.2) ist es anschaulich klar, dass jede Basis die gleiche Elementeanzahl enthält: Je zwei linear unabhängige Vektoren spannen die Ebene, je drei den Raum auf. Wir fragen nach einer Verallgemeinerung: Sind bei einem gegebenen Vektorraum alle Basen von der gleichen Mächtigkeit?

↑ M

Bei dieser Untersuchung beschränken wir uns zunächst auf **endlich erzeugbare** Vektorräume, d.h. solche, die ein endliches Erzeugendensystem und nach (7.7) damit eine endliche Basis besitzen.

8.1 Hilfssatz.

Seien $B = \{b_1, \ldots, b_n\}$ eine Basis des K-Vektorraumes V und $a = \sum_{i=1}^{n} b_i\lambda_i$ mit $\lambda_i \in K$ und $\lambda_j \neq 0$ für ein $j \in \{1, \ldots, n\}$. Dann ist auch

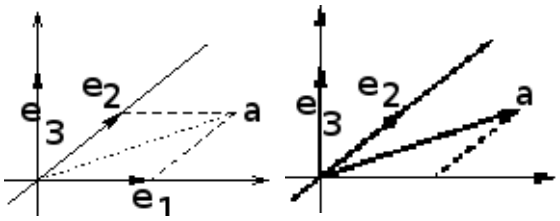
$$B' = \{b_1, \ldots, b_{j-1}, a, b_{j+1}, \ldots, b_n\}$$

eine Basis von V; außerdem gilt $a \notin \{b_1, \ldots, b_{j-1}, b_{j+1}, \ldots, b_n\}$.

Beispiel.

Seien $V = \mathbb{R}^3$,
$e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$,
$B = \{e_1, e_2, e_3\}$, $a = (1,1,0)$.

Wegen
$a = (1,0,0) \cdot \underline{1} + (0,1,0) \cdot \underline{1} + (0,0,1) \cdot 0$
sind auch $B' = \{a, e_2, e_3\}$ und
$B'' = \{e_1, a, e_3\}$ Basen von $\mathbb{R}^3$; (s. auch Figur 8.1 !)



Figur 8.1: Austausch eines Basisvektors (Ersetzen von $\vec{e_1}$ durch $\vec{a}$).

Beweis von 8.1. Zunächst gilt:

(i) $a = \sum_{\substack{i=1 \\ i\neq j}}^{n} b_i\lambda_i + b_j\lambda_j \underset{\lambda_j\neq 0}{\Rightarrow} b_j = \big(a - \sum_{\substack{i=1 \\ i\neq j}}^{n} b_i\lambda_i\big)\lambda_j^{-1}$.

Sei $v \in V$; wegen $V = \langle B\rangle$ existieren $\mu_1, \ldots, \mu_n \in K : v = \sum_{i=1}^{n} b_i\mu_i$; damit folgt

$$v = \sum_{\substack{i=1 \\ i\neq j}}^{n} b_i\mu_i + b_j\mu_j = \sum_{\substack{i=1 \\ i\neq j}}^{n} b_i\mu_i + \big(a - \sum_{\substack{i=1 \\ i\neq j}}^{n} b_i\lambda_i\big)\lambda_j^{-1}\mu_j = \sum_{\substack{i=1 \\ i\neq j}}^{n} b_i(\mu_i - \lambda_i\lambda_j^{-1}\mu_j) + a(\lambda_j^{-1}\mu_j).$$

Also ist $V = \langle B'\rangle$.

(ii) Wir zeigen: $a, b_1, \dots b_{j-1}, b_{j+1}, \dots b_n$ und damit B' sind linear unabhängig.

$$a\mu + \sum_{\substack{i=1\\ i\neq j}}^{n} b_i\mu_i = \mathfrak{o} \Rightarrow \big(\sum_{\substack{i=1\\ i\neq j}}^{n} b_i\lambda_i + b_j\lambda_j\big)\mu + \sum_{\substack{i=1\\ i\neq j}}^{n} b_i\mu_i = \mathfrak{o}$$

$$\Rightarrow \sum_{\substack{i=1\\ i\neq j}}^{n} b_i(\lambda_i\mu + \mu_i) + b_j(\lambda_j\mu) = \mathfrak{o}$$

$$\underset{B \text{ l.u.}}{\Rightarrow} (\forall i \in \{1,\dots,n\} \setminus \{j\} : \lambda_i\mu + \mu_i = 0) \wedge \lambda_j\mu = 0$$

$$\underset{\lambda_j\neq 0}{\Rightarrow} \mu = 0 \wedge \forall i \in \{1,\dots,n\} \setminus \{j\} : \mu_i = 0.$$

(iii) In (ii) wurde gezeigt, dass $a, b_1, \dots b_{j-1}, b_{j+1}, \dots b_n$ linear unabhängig sind; daher gilt $a \neq b_k$ für alle $k \neq j$. □

Verallgemeinerung:

8.2 Satz. (Austausch-Satz/Lema von Steinitz)

Seien $B = \{b_1, \dots, b_n\}$ Basis des (endlich erzeugbaren) K-Vektorraumes V und $A = \{a_1, \dots, a_m\}$ eine linear unabhängige Teilmenge von V (mit $|B| = n$ und $|A| = m$).
Dann gilt: (1) $m = |A| \leq |B| = n$.

(2) $\exists C \subseteq B :$ $B' = C \cup A$ ist Basis von V und $|B'| = n$.

Beweis. Durch vollständige Induktion nach m:
Induktions–Verankerung: $m = 1$. Gegeben sei: $\{a_1\}$.
B Basis $\Rightarrow a_1 = \sum_{i=1}^{n} b_i\lambda_i$ für geeignete $\lambda_i \in K \underset{\{a_1\} \text{ l.u.}}{\Rightarrow} \exists j \in \{1,\dots,n\} : \lambda_j \neq 0$;
mit (8.1) folgt $B' = \{b_1, \dots, b_{j-1}, a_1, b_{j+1}, \dots, b_n\}$ ist Basis der Mächtigkeit n und damit $m = 1 \leq n$.
Induktions–Voraussetzung: Die Behauptung sei richtig für $m = k-1$, d.h.:
$\{a_1, \dots, a_{k-1}\}$ linear unabhängig $\Rightarrow k-1 \leq n$ und $\{a_1, \dots, a_{k-1}\} \cup \{b_{i_k}, \dots, b_{i_n}\}$ ist Basis der Mächtigkeit n für eine geeignete Teilmenge $\{b_{i_k}, \dots, b_{i_n}\} \subseteq B$.
Induktions–Schluss:
$\{a_1, \dots, a_k\}$ l.u. $\Rightarrow \{a_1, \dots, a_{k-1}\}$ l.u. $\underset{\text{Vor.}}{\overset{\text{Ind.}}{\Rightarrow}} k-1 \leq n$ und $B'' = \{a_1, \dots, a_{k-1}, b_{i_k}, \dots, b_{i_n}\}$
Basis für eine geeignete Teilmenge $\{b_{i_k}, \dots, b_{i_n}\}$ von B.
Wäre $k-1 = n$, so $B'' = \{a_1, \dots, a_{k-1}\}$, damit $a_k \in \langle a_1, \dots, a_{k-1}\rangle$ und A linear abhängig, ein Widerspruch. Also ist $k-1 < n$, d.h. $k \leq n$.
Da $a_k \in V = \langle B''\rangle$ ist, existieren $\lambda_k, \dots, \lambda_n \in K$ mit $a_k = \sum_{i=1}^{k-1} a_i\lambda_i + \sum_{j=k}^{n} b_{i_j}\lambda_j$; wären $\lambda_k, \dots, \lambda_n$ sämtlich 0, so A linear abhängig; somit existiert ein $\lambda_0 \in \{\lambda_k, \dots, \lambda_n\}$ mit $\lambda_0 \neq 0$. Nach (8.1) gilt: $(B'' \setminus \{b_{i_0}\}) \cup \{a_k\}$ ist Basis von V der Mächtigkeit n. □

8.3 Korollar: Elementeanzahl in Basen

Je zwei Basen eines endlich erzeugbaren K-Vektorraumes V enthalten dieselbe (endliche) Anzahl von Elementen.

Beweis. Sei B Basis mit $|B| = n \in \mathbb{N}_0$ (eine solche existiert nach Voraussetzung, da V endlich erzeugbar ist); sei B' eine weitere Basis von V. (B' könnte unendlich sein.)
Ist M endliche Teilmenge von B', so ist M linear unabhängig; nach (8.2) gilt $|M| \leq n$ und damit $|B'| \leq n = |B|$. Analog folgt $|B| \leq |B'| < \infty$, also $|B| = |B'|$. □

Die Beschränkung auf endlich erzeugbare Vektorräume in (8.3) ist unnötig, falls man das Auswahlaxiom oder äquivalente Bedingungen heranzieht[41]; es gilt dann folgende Verallgemeinerung:

8.4 Satz von Löwig (Basis-Gleichmächtigkeits-Satz).

Je zwei Basen eines K-Vektorraumes V sind gleichmächtig.

Beweis-Andeutung.
Seien B_1 und B_2 Basen von V. Ist B_1 oder B_2 endlich, so folgt $|B_1| = |B_2|$ nach (8.3). Seien also B_1, B_2 unendlich.
Ist $x \in B_1$, so bezeichne $B_{2,x}$ die Menge der Basisvektoren aus B_2, die zur Darstellung von x „benötigt“ wird; diese Menge ist endlich. Außerdem gilt $B_2 = \bigcup\limits_{x \in B_1} B_{2,x}$, anderenfalls wären Elemente von B_2 überflüssig.
Daher gilt $|B_2| = |\bigcup\limits_{x \in B_1} B_{2,x}| \underset{\substack{\text{wegen} \\ |B_{2,x}| \leq \aleph_0}}{\leq} |B_1| \cdot \aleph_0 \underset{\substack{\text{wegen} \\ B_1 \text{unendlich}}}{=} |B_1|$.
(Zur Kardinalzahl-Arithmetik s. z.B. Dugundji, Topology, Chap.II, Boston 1966 oder Scheja/Storch - s. Literaturliste - oder Bücher über Mengenlehre.)
Analog folgt $|B_1| \leq |B_2|$. Insgesamt ergibt sich $|B_1| = |B_2|$. □

(Einen hiervon etwas abweichenden Beweis findet man z.B. in H. Lüneburg: Einführung in die Algebra etc. 1973 p. 169).

Wie wir später sehen werden, ist die (konstante) Mächtigkeit einer Basis von V eine der wesentlichen charakteristischen Eigenschaften von V.

8.5 Definition: Dimension eines Vektorraums

Sei V ein K-Vektorraum. Die Mächtigkeit m einer (und nach (8.4) damit jeder) Basis von V heißt die

Dimension des Vektorraumes V über dem Körper K,

oft auch **Rang** von V über K.
Schreibweise: $\dim_K V = m$
Falls keine Verwechslungen zu befürchten sind, schreiben wir auch $\dim V = m$.

[41] Zur Definition der Kardinalzahlen wird die Wohlordnung der sogenannten Ordinalzahlen verwandt.

Anmerkung 1. Ein endlich erzeugter Vektorraum besitzt definitionsgemäß eine endliche Basis. Ist deren Elementeanzahl n ($\in \mathbb{N}_0$), so gilt $\dim V = n$. Wir nennen V dann auch einen n-dimensionalen Vektorraum. Statt von einem endlich erzeugten Vektorraum werden wir im folgenden von einem **endlich-dimensionalen** Vektorraum sprechen.

Beispiel.

(i) $V = \mathcal{V}$, der Vektorraum der Vektoren der reellen Ebene: $\mathcal{V} = \langle \vec{n}_1, \vec{n}_2 \rangle$. Also:

$\dim_{\mathbb{R}} \mathcal{V} = 2$ (in Übereinstimmung mit der umgangssprachlichen Dimensionsvorstellung).

Dimension von Unterräumen:

$\dim_{\mathbb{R}}(\vec{a}\,\mathbb{R}) = 1$, für $\vec{a} \in \mathcal{V} \setminus \{\mathfrak{o}\}$,
$\dim_{\mathbb{R}}(\{\mathfrak{o}\}) = 0$, (da $\emptyset$ Basis von $\{\mathfrak{o}\}$ ist).

(ii) $V = \mathcal{V}_{\mathrm{r}}$, der Vektorraum der Vektoren des reellen Raumes: $\mathcal{V}_r = \langle \vec{n}_1, \vec{n}_2, \vec{n}_3 \rangle$. Es folgt: $\dim_{\mathbb{R}} \mathcal{V}_{\mathrm{r}} = 3$.

Dimension von Unterräumen:

$\dim_{\mathbb{R}}(\vec{a}\mathbb{R} + \vec{b}\mathbb{R}) = 2$, falls $\vec{a}, \vec{b}$ linear unabhängig sind,
$\dim_{\mathbb{R}} \vec{a}\mathbb{R} = 1$, falls $\vec{a} \neq \mathfrak{o}$,
$\dim_{\mathbb{R}}(\{\mathfrak{o}\}) = 0$.

(iii) Für $V = K^n$, $n \in \mathbb{N}$ gilt

$$\boxed{\dim_K K^n = n} \quad (\text{da } \{e_1, \ldots, e_n\} \text{ Basis ist}).$$

Anmerkung 2. Besitzt der K-Vektorraum V keine endliche Basis, so nennt man ihn auch oft unendlich-dimensional, in Zeichen $\dim_K V = \infty$.
Hierbei ist die (unendliche) Mächtigkeit nicht angegeben, es handelt sich also um eine globalere Angabe als in Definition (8.5).

Beispiel.
$\{{\mathrm{id}_{\mathbb{R}}}^i \mid i \in \mathbb{N}_0\}$ ist eine Basis von $\mathcal{P}(\mathbb{R})$, dem Vektorraum der reellen Polynomabbildungen. Daher gilt: $\dim_{\mathbb{R}} \mathcal{P}(\mathbb{R}) = \infty$, genauer: $\dim_{\mathbb{R}} \mathcal{P}(\mathbb{R}) = \aleph_0$.

Anmerkung 3. Obwohl $\emptyset$ kein Vektorraum (insbesondere auch kein Unterraum) ist, erweist es sich (für später behandelte Formeln) als zweckmäßig,

$$\dim_K \emptyset := -1$$

zu definieren.

Weitere Folgerungen:

8.6 Hilfssatz: Anzahl linear unabhängiger Vektoren

Seien V ein n-dimensionaler Vektorraum, $n \in \mathbb{N}$ und $A \subseteq V$. Dann gilt

(i) A linear unabhängig $\Rightarrow |A| \leq n$.

(ii) A Basis $\Leftrightarrow$ A linear unabhängig und $|A| = n$.

(iii) U Unterraum von $V \Rightarrow \dim_K U \leq n$.

Beweis.

(i) Nach (7.7) existiert eine Basis B mit $A \subseteq B$. Nach (8.3) ist $|B| = n$.

(ii) Ist A Basis, so nach (8.3) $|A| = \dim_K V = n$ und definitionsgemäß A linear unabhängig.

Ist A linear unabhängig und $|A| = n$, so folgt wie in (i) $A \subseteq B$ für eine Basis B und aus Anzahlgründen $A = B$.

(iii) Ist C Basis von U, so gilt: C linear unabhängig in U und damit C linear unabhängig in V. Aus (i) folgt

$$\dim_K U = |C| \leq n = \dim_K V.$$ □

Wir wollen im folgenden untersuchen, inwieweit die Struktur eines Vektorraumes schon durch seine Dimension festgelegt ist. Zunächst beachten wir folgende

Anmerkung: Ist $(V, +, \cdot)$ ein Vektorraum über dem Körper K und K' ein Unterkörper von K, so ist $\left(V, +, \underset{K}{\cdot}\Big|_{V \times K' \to V}\right)$ ein K'-Vektorraum.

Beispiele.

Jeder $\mathbb{R}$-Vektorraum wird durch Einschränkung des Skalarbereichs von $\mathbb{R}$ auf $\mathbb{Q}$ zum $\mathbb{Q}$-Vektorraum. Jeder $\mathbb{C}$-Vektorraum wird bei Beschränkung auf reelle Skalare zum $\mathbb{R}$-Vektorraum.

Da bei der Einschränkung des Skalarbereichs eine linear abhängige Menge linear unabhängig werden kann (in den Linearkombinationen dürfen einige Koeffizienten nicht mehr verwandt werden), gilt i.A.

$$\dim_K V \neq \dim_{K'} V.$$

Beispiele.

$\dim_{\mathbb{R}} \mathbb{R} = 1$,	$\dim_{\mathbb{Q}} \mathbb{R} = \infty$	genauer: $\dim_{\mathbb{Q}} \mathbb{R} = \mathfrak{c} = \lvert\mathbb{R}\rvert$
$\dim_{\mathbb{C}} \mathbb{C} = 1$,	$\dim_{\mathbb{R}} \mathbb{C} = 2$	$(1, i$ sind linear unabhängig über $\mathbb{R}$,
		aber linear abhängig über $\mathbb{C})$.

Da sich mit Wechsel des Körpers auch weitere Eigenschaften (z.B. Dimension von Unterräumen) der entsprechenden Vektorräume ändern, scheint es zweckmäßig, sich bei Betrachtungen über Struktur-Gleichheit auf Vektorräume über dem selben Körper zu beschränken, insbesondere also K und die Dimension festzuhalten.

8.7 Definition: Vektorraum-Isomorphismus

Seien V und V' Vektorräume über dem selben Körper K.
(a) (Vorgriff auf 10.1:) Eine Abbildung $f : V \to V'$ heißt **lineare Abbildung** (VR-Homomorphismus, Operator, lineare Transformation), falls gilt:

$$(1)\ \forall v, w \in V : f(v + w) = f(v) + f(w);$$

$$(2)\ \forall v \in V : \forall \lambda \in K : f(v\lambda) = f(v) \cdot \lambda.$$

(b) V und V' heißen **isomorph** (in Zeichen $V \cong V'$), wenn es eine bijektive lineare Abbildung $f : V \to V'$ gibt. Eine solche Bijektion dieser Eigenschaft heißt **Vektorraum-Isomorphismus** (VR-Iso)
Anmerkung. Ein Vektorraum-Homomorphismus $f : V \to V'$ ist wegen (1) insbesondere Gruppenhomomorphismus von $(V, +)$ in $(V', +)$. Zusätzlich beachtet er die beiden Multiplikationen mit Skalaren. Analog ist ein VR-Isomorphismus auch ein Gruppenisomorphismus.

M ↓ ↑ M

Isomorphe Vektorräume sind von der Vektorraum-Struktur her nicht unterscheidbar.

M ↓

Beispiele: (i) Sei $\mathcal{V}$ der in (1.1) behandelte Vektorraum der Klassen vektorgleicher Pfeile. Dann gilt

$$\mathcal{V} \cong \mathbb{R}^2.$$

Wählen wir die Basis $(\vec{e}_1, \vec{e}_2)$ von $\mathcal{V}$, so lässt sich diese Isomorphie als Übergang von den Vektoren zu den durch die entsprechenden Ortsvektoren bestimmten Punkten (in kartesischen Koordinaten) geometrisch deuten: $\vec{v} = \vec{e}_1\xi_1 + \vec{e}_2\xi_2 \mapsto (\xi_1, \xi_2) \in \mathbb{R}^2$
Aufgrund der Isomorphie können wir von dem zum Teil intuitiv eingeführten $\mathcal{V}$ zu dem Modell $\mathbb{R}^2$ übergehen (siehe auch § 9 !)

↑ M

(ii) Sei V ein K-Vektorraum der Dimension n. Dann existiert definitionsgemäß eine Basis B, die wir als geordnet annehmen können.
Die Abbildung

$$i_B : \begin{cases} V & \to K^n \\ \sum_{i=1}^{n} b_i\xi_i = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B & \mapsto (\xi_1, \ldots, \xi_n) \end{cases}$$

ist dann ein Vektorraum-Isomorphismus (vgl. (7.6)); es folgt nämlich

$$i_B(x + y) = i_B(x) + i_B(y) \quad \text{und} \quad i_B(x\lambda) = i_B(x)\lambda\ .$$

Es gilt also $V \cong K^n$.
Oft werden wir (nach Auswahl einer Basis B) die Elemente von V mit den Elementen von K^n identifizieren.
(iii) Sei $V = V' = \mathbb{R}^n$ und $\lambda \in \mathbb{R} \setminus \{0\}$. Dann ist $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mit $f(v) = v\lambda$ ein Vektorraum-Isomorphismus von V auf sich, ein Vektorraum-Automorphismus. Die Inverse ist ebenfalls ein Automorphismus: $f^{-1}(w) = w\lambda^{-1}$.
Beweis?
Wir betrachten Eigenschaften der Isomorphie. Ähnlich wie bei der Gruppenisomorphie gilt:

8.8 Hilfssatz: Eigenschaften von VR-Isomorphismen

(a) Auf jeder Menge von Vektorräumen ist die Vektorraum-Isomorphie eine Äquivalenzrelation.
(b) Ist $f : V \to V'$ ein Vektorraum-Isomorphismus, so auch $f^{-1} : V' \to V$.
(c) Die Hintereinanderausführung von zwei VR-Isomorphismen ist wieder ein VR-Isomorphismus.

Beweis.

(i) $V \cong V$: id_V ist ein Isomorphismus.

(ii) $V \cong V' \Leftrightarrow V' \cong V$.

Ist $f : V \to V'$ Isomorphismus, so existiert f^{-1}; wir zeigen, dass f^{-1} wieder Vektorraum-Isomorphismus ist: die Additivität von f^{-1} sieht man folgendermaßen: Für alle $v', w' \in V'$ existieren $v, w \in V$ mit $f(v) = v'$ und $f(w) = w'$, also gilt:
$f^{-1}(v'+w') = f^{-1}\big(f(v)+f(w)\big) \underset{f \text{ Isom.}}{=} f^{-1}\big(f(v+w)\big) = v+w = f^{-1}(v')+f^{-1}(w')$;
die Verträglichkeit mit der S-Multiplikation ergibt sich so:

$\forall v' \in V' \, \exists v \in V : f(v) = v'$, damit gilt für alle $\lambda \in K$:

$f^{-1}(v'\lambda) = f^{-1}\big(f(v)\lambda\big) \underset{f \text{ Isom.}}{=} f^{-1}\big(f(v\lambda)\big) = (f^{-1} \circ f)(v\lambda) = v\lambda = f^{-1}(v')\lambda.$

(iii) $V \cong V' \wedge V' \cong V'' \Rightarrow V \cong V''$.

Seien $f : V \to V'$ und $g : V' \to V''$ Isomorphismen; dann ist $g \circ f$ definiert, Gruppen-Isomorphismus und wegen

$$(g \circ f)(v\lambda) = g\big(f(v\lambda)\big) = g\big(f(v)\lambda\big) = g\big(f(v)\big)\lambda = (g \circ f)(v)\,\lambda$$

Vektorraum-Isomorphismus.

(b) und (c) folgen aus dem obigen Beweis. □

Aus obigem Beispiel (ii) ergibt sich die eine Implikation von

8.9 Satz: Vektorräume der dim n

Sei V ein K-Vektorraum und $n \in \mathbb{N}$. Dann gilt:

$$\dim_K V = n \Leftrightarrow V \cong K^n.$$

Bis auf Isomorphie gibt es nur einen $K-$Vektorraum der Dimension n.

Beweis. "⇒" s.o.(Beispiel)
"⇐" $B = (e_1, \ldots, e_n)$ ist (geordnete) Basis von K^n; damit $\boxed{\dim_K K^n = n}$, (s.o.). Laut Voraussetzung existiert ein Isomorphismus $f : K^n \to V$. Wir zeigen, dass $f(B) = (f(e_1), \ldots, f(e_n))$ Basis von V ist:

(i) $V = \langle f(B)\rangle$:

Sei $v \in V$; da f Bijektion ist, gilt: $\exists w = (\lambda_1, \ldots, \lambda_n) \in K^n : f(w) = v$.

$v = f(w) = f\big((\lambda_1, \ldots, \lambda_n)\big) = f\big(\sum_{i=1}^{n} e_i\lambda_i\big) \underset{f\text{Isom.}}{=} \sum_{i=1}^{n} f(e_i)\lambda_i \in \langle f(B)\rangle$.

(ii) $f(B)$ linear unabhängig:

$\sum_{i=1}^{n} f(e_i)\mu_i = \mathfrak{o} \underset{f\text{Isom.}}{\Rightarrow} f\big(\sum_{i=1}^{n} e_i\mu_i\big) = \mathfrak{o} \underset{\substack{\text{inj.Hom.}\\ \text{von}(K^n,+)\\ \text{auf } (V,+)}}{\Rightarrow} \sum_{i=1}^{n} e_i\mu_i = \mathfrak{o} \underset{B \text{ l.u.}}{\Rightarrow} \forall i = 1, \ldots, n : \mu_i = 0.$

□

M
↓

Damit ist es uns gelungen, alle endlich erzeugten Vektorräume zu beschreiben; es sind dies (bis auf Isomorphie) gerade die Vektorräume $(K^n, +, \cdot)$ mit $n \in \mathbb{N}$.

↑
M

Ohne Beweis erwähnen wir eine Verallgemeinerung.

8.10 Satz (Isomorphie und Gleichmächtigkeit der Basen)

Seien V und W zwei Vektorräume über demselben Körper K. Dann sind V und W genau dann isomorph, wenn V und W gleichmächtige Basen besitzen.

Zwei Vektorräume über demselben Körper K sind also genau, dann isomorph, wenn sie dieselbe Dimension besitzen. **Achtung**: Hierbei ist aber die präzise Definition (8.5) und nicht die aus Anmerkung 2 zu (8.5) zu verwenden:

$$\dim V = \infty \wedge \dim W = \infty \underset{\text{i.A.}}{\not\Rightarrow} V \cong W.$$

Insbesondere lässt sich zeigen:

8.11 Satz : Isomorphie zu $K^{(I)}$

Sei V ein K-Vektorraum und B Basis von V. Dann gilt $$V \cong K^{(I)} \Leftrightarrow |I| = |B| \ (= \dim_K V).$$

Anmerkung :
Die $K^{(I)}$ (für Indexmengen I) sind also bis auf Isomorphie sämtliche $K-$Vektorräume.

§ 9 Die affine Geometrie eines Vektorraums

M ↓

Nach Auszeichnung eines kartesischen Koordinatensystems lassen sich die Punkte der Zeichenebene (vgl. (1.1)) durch Koordinatenpaare, also Elemente von $\mathbb{R}^2$ repräsentieren.

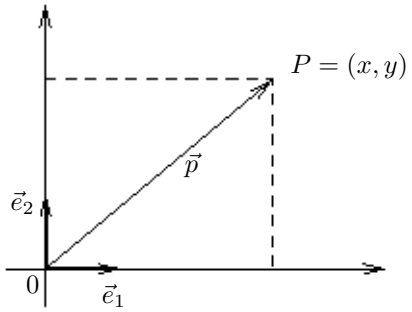
Jedem Punkt P lässt sich — wie in (1.1) ausgeführt — **ein Element von $\mathcal{V}$ als Ortsvektor zuordnen**.

Sind $\vec{e}_1$ bzw. $\vec{e}_2$ die Ortsvektoren von $(1,0)$ bzw. $(0,1)$, so ist $\vec{p} = \begin{pmatrix} x \\ y \end{pmatrix}_{(\vec{e}_1,\vec{e}_2)} = \vec{e}_1 x + \vec{e}_2 y$ Ortsvektor von $P = (x,y)$.

Die Ortsvektoren der Geraden $0P$ durch 0 und $P \neq 0$ sind die Elemente von

$$\vec{p}\,\mathbb{R} = \{\vec{p}\lambda | \lambda \in \mathbb{R}\}.$$

Umgekehrt bestimmt jeder 1-dimensionale Unterraum von $\mathcal{V}$ eine Gerade durch den Nullpunkt.



Figur 9.1 : Zuordnung Punkt-Ortsvektor

Nun ergibt sich eine Möglichkeit, die Unsicherheiten bei der Einführung von $\mathcal{V}$, (Begriffe wie Punkt, Gerade, parallel, Länge, Orientierung waren undefiniert verwandt worden), zu eliminieren:

Als **Modell** für die **Zeichenebene/Anschauungsebene (mit Auszeichnung eines Nullpunktes)** wählen wir $\mathbb{R}^2$, genauer: als Punkte die Elemente (Vektoren) von $\mathbb{R}^2$, als Geraden durch den Nullpunkt die 1-dimensionalen Unterräume des $\mathbb{R}$-Vektorraums $\mathbb{R}^2$. (Weitere Definitionen folgen im Laufe der Vorlesung). Wir sprechen dabei von der **reellen Ebene** später auch von der **euklidischen Ebene**.

Diese Präzisierung erkaufen wir durch „Aufgabe der ontologischen Bindung“: Wir sind uns bewusst, dass es sich lediglich nur um ein Modell für die Zeichenebene handelt. Letzteres hat sich bisher bewährt (– im Gegensatz zu dem Modell $\mathbb{Q}^2$ der Pythagoräer, das zu Widersprüchen mit der Anschauung führte). Ähnlich wählen wir $\mathbb{R}^3$ als Modell für den Anschauungsraum.

Anmerkung. Diese Vorgehensweise innerhalb der Geometrie nennt man „analytisch“ (daher „analytische Geometrie“) im Gegensatz zur „synthetischen“ Vorgehensweise, bei der Punkte, Geraden, Inzidenz, Parallelität etc. axiomatisch eingeführt werden (vgl. das Axiomensystem des 3-dimensionalen Raumes nach HILBERT !) und bei der die Koordinatisierbarkeit durch $\mathbb{R}^3$ erst nach längeren Betrachtungen erhalten wird.

In unserem Modell haben wir zunächst nur die Punkte und die Geraden durch den Nullpunkt beschrieben.

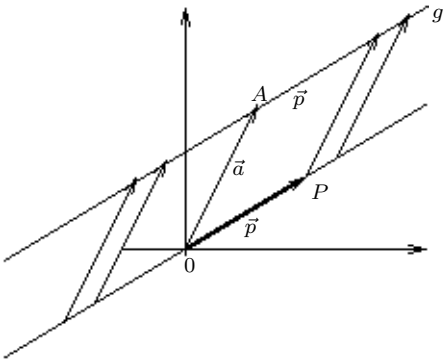
Welche Form haben die übrigen Geraden? Dazu gehen wir nochmals zur Zeichenebene zurück:

Eine zur Geraden $g' := 0P$ parallele Gerade g durch den Punkt A mit Ortsvektor $\vec{a}$ entsteht durch Translation (Parallelverschiebung) um den Vektor $\vec{a}$, also Addition von $\vec{a}$:

$$g = \vec{a} + \vec{p}\,\mathbb{R} = \{\vec{a} + \vec{p}\lambda | \lambda \in \mathbb{R}\}$$

(vgl. (1.1) d(ii) ist die Menge der Ortsvektoren der Punkte von g. Also ist g „Nebenklasse“ nach einem Unterraum der Dimension 1.

Umgekehrt stellt jede solche Nebenklasse $\vec{a}+U$ (mit $\dim_{\mathbb{R}} U = 1$) eine zu U parallele Gerade durch den Punkt mit Ortsvektor $\vec{a}$ dar.



Figur 9.2: Zu einer Nullpunkts-Geraden parallele Gerade als Nebenklasse

↑ M

9.1 Definition: Lineare Mannigfaltigkeit, affiner Unterraum

Sei V ein Vektorraum über dem Körper K. Eine Teilmenge M von V heißt **lineare Mannigfaltigkeit** (LM) oder **affiner Unterraum** von V, wenn es zu ihr einen Unterraum U_M und ein $a \in V$ gibt mit $M = a + U_M = \{a + u | u \in U_M\}$.
Anmerkungen.

(i) Unterräume sind spezielle lineare Mannigfaltigkeiten: $U = \mathfrak{o} + U$.

(ii) Ist $U_M = \{\mathfrak{o}\}$, so identifizieren wir $M = a + \{\mathfrak{o}\} = \{a\}$ mit $a \in V$.

(iii) Der Repräsentant a von M ist durch M i.A. nicht eindeutig festgelegt (s.u.):

$$\boxed{b \in a + U \Leftrightarrow a + U = b + U \Leftrightarrow a - b \in U} .$$

Beweisskizze: $b = a + u \Rightarrow b + U = a + (u + U) = a + U$. Die Umkehrung ist klar ?

Die linearen Mannigfaltigkeiten der Form $a + U$ sind ja auch Nebenklassen nach der Untergruppe U von $(V, +)$.

M
↓

(iv) Geometrisch erhält man den Unterraum U_M aus $M = a + U_M$ durch Parallelverschiebung (Translation) um den Vektor $-a$ in den Nullpunkt.

↑
M

Allgemein ist also nicht verwunderlich, dass U_M durch M eindeutig bestimmt ist.

9.2 Hilfssatz: Zur Eindeutigkeit der Darstellung eines affinen Unterraums

Sei V K-Vektorraum, seien U, U' Unterräume von V und $a, a' \in V$. Dann gilt:

$$a + U = a' + U' \Leftrightarrow a' - a \in U \wedge U = U' .$$

Beweis.

„⇒“ Sei $a + U = a' + U'$. Dann gilt: $\forall u \in U\ \exists u' \in U' : a + u = a' + u'$.
Setzt man $u = 0$, so ergibt sich $a - a' \in U'$.
Sei nun $u \in U$ mit $u \neq 0$; dann folgt $\exists u' \in U' : u = (a' - a) + u' \in U'$, also $U \subseteq U'$.
Analog folgt $U' \subseteq U$, insgesamt also $U = U'$ sowie $a' - a \in U$.

„⇐“ $U = U' \wedge a' \in a + U \Rightarrow a + U = a' + U = a' + U'$. □

Geometrisch motiviert definieren wir:

9.3 Definition: Dimension einer linearen Mannigfaltigkeit

Sei V K-Vektorraum und $M = a + U$ lineare Mannigfaltigkeit von V mit Unterraum U.

$$\boxed{\dim_K(a + U) := \dim_K U} .$$

Beispiele:

(a) $V = \mathcal{V}$ (bzw. $\mathbb{R}^2$).

LM'en der Dimension 0 sind die Punkte (Vektoren) von V (s. Anm. (ii)),

LM'en der Dimension 1 sind die Geraden von V (s. Motivation!),

LM der Dimension 2 ist V selbst.

(b) $V = \mathcal{V}_\mathrm{r}$ (bzw. $\mathbb{R}^3$).

LM'en der Dimension 0 sind die Punkte (Vektoren) von V,

LM'en der Dimension 1 sind die Geraden von V,

LM'en der Dimension 2 sind die Ebenen von V,

LM der Dimension 3 ist V selbst.

M ↓

Betrachten wir wieder die Situation in der durch $\mathbb{R}^2$ beschriebenen Zeichenebene bzw. dem durch $\mathbb{R}^3$ beschriebenen Anschauungsraum! Der Punkt P liegt auf der Geraden g bzw. die Gerade g geht durch den Punkt P, wenn $\{P\} \subseteq g$ gilt. Wir sagen auch: P inzidiert mit g. Analog inzidiert die Gerade g mit der Ebene E, wenn $g \subseteq E$ ist.

Zwei Geraden $g = a + \vec{p}\mathbb{R}$ und $h = b + \vec{q}\mathbb{R}$ sind **parallel**, wenn $\vec{q} \in \vec{p}\mathbb{R}$, also $U_g = \vec{p}\mathbb{R} = \vec{q}\mathbb{R} = U_h$, gilt. Zwei Ebenen E und F sind parallel, wenn sie durch Parallelverschiebungen in den Nullpunkt zur selben Ebene führen, also

$$U_E = U_F$$

gilt. Für eine zu E parallele Gerade g haben wir $U_g \subseteq U_E$.

Diese algebraische Beschreibung stellt einen Zusammenhang zwischen Sachverhalten im Vektorraum und geometrischen Gegebenheiten dar. Eine Übertragung auf einen beliebigen Vektorraum V führt zum Begriff der **affinen Geometrie** $\mathrm{AG}(V)$ von V.

Unter affiner Geometrie versteht man allgemein das Studium geometrischer Begriffe wie Punkte, Geraden, Ebenen, Inzidenz, Parallelismus **ohne** Verwendung der Begriffe Länge und Winkel.

↑ M

9.4 Definition: Affine Geometrie eines Vektorraums

Sei V K-Vektorraum mit $1 \leq \dim_K V = n < \infty$.

(a) $\mathrm{AG}(V) := \{M | M \text{ LM von } V\}$

Wir benutzen geometrische Sprechweise:

0–dim LM'en heißen	Punkte (wir identifizieren sie mit den Vektoren von V),
1–dim LM'en heißen	Geraden,
2–dim LM'en heißen	Ebenen
⋮	⋮
i–dim LM'en heißen	**affine Unterräume** der Dimension i,
⋮	⋮

$(n-1)$-dim LM'en heißen **Hyperebenen**.

(b) Für lineare Mannigfaltigkeiten $L, M \in \mathrm{AG}(V)$ definieren wir **Inzidenz** durch

$$L \,\mathrm{I}\, M \Leftrightarrow L \subseteq M \vee M \subseteq L.$$

(Insbesondere: $\{a\} \,\mathrm{I}\, L \Leftrightarrow a \in L$)

Auch hier verwendet man geometrische Sprechweise: Ein Punkt P liegt auf einer Geraden g (statt $P \,\mathrm{I}\, g$) bzw. g geht durch P. Punkte $P_1, P_2, \ldots, P_s$, die mit einer Geraden g inzidieren, heißen **kollinear**; Punkte einer Ebene heißen **komplanar**.

(c) Zwei lineare Mannigfaltigkeiten $L = a + U_L$ und $M = b + U_M$ (mit Unterräumen U_L und U_M) heißen **parallel**, in Zeichen

$L \parallel M$ g.d.w. gilt $U_L \,\mathrm{I}\, U_M$; also

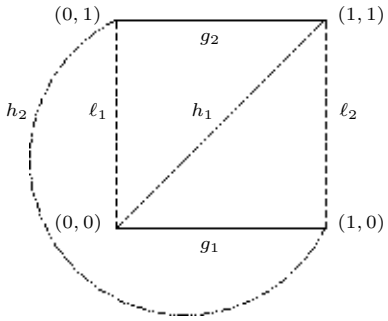
$L \parallel M \Leftrightarrow U_L \subseteq U_M \vee U_M \subseteq U_L$.

(d) $(\mathrm{AG}(V), \mathrm{I}, \parallel)$ heißt **affine Geometrie** von V.

Beispiel (Fortsetzung).
Sei $K = \mathrm{GF}(2)$ und $V = K^2$!

Punkte von $\mathrm{AG}(V)$ sind
$$(0,0),\ (1,0),\ (0,1),\ (1,1)$$

Geraden sind z.B.
$g_1 = \langle (1,0) \rangle = \{(0,0), (1,0)\}$
$g_1 \parallel g_2 = (0,1) + \langle (1,0) \rangle$
$\quad = \{(0,1), (1,1)\}$
$h_1 = \langle (1,1) \rangle \parallel h_2 = h_1 + (1,0)$
$\ell_1 = \langle (0,1) \rangle \parallel \ell_2 = \ell_1 + (1,0)$.



Figur 9.3: Veranschaulichung von $\mathrm{AG}(\mathrm{GF}(2)^2)$; (parallele Geraden sind mit gleicher Strichart dargestellt).

Wie in $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ erhält man allgemein den folgenden Hilfssatz:

9.5 Hilfssatz: Geraden- und Ebenen-Gleichungen

Sei V K-Vektorraum. Dann gilt

(a) Zu je zwei verschiedenen Punkten p und q von $\mathrm{AG}(V)$ existiert genau eine mit p und q inzidierende Gerade g, nämlich

$$g = p + (q - p)K.$$

(b) Ist $\dim_K V \geq 3$ und sind p, q, r drei nicht-kollineare Punkte von $\mathrm{AG}(V)$, so existiert genau eine mit p, q, r inzidierende Ebene E, nämlich

$$E = r + (p - r)K + (q - r)K.$$

9.6 Hilfssatz: Parallelität in der Ebene

Sei V K-Vektorraum der Dimension 2, seien ferner g und h Geraden in $\mathrm{AG}(V)$. Dann gilt

$$g \parallel h \Leftrightarrow (g = h \vee g \cap h = \emptyset) \Leftrightarrow U_g = U_h \ .$$

Anmerkung: Die Aussage dieses Satzes ist falsch für $\dim V \geq 3$; es existieren dann sogenannte windschiefe Geraden. Daher ist bei $\dim V \geq 3$ zusätzlich zu fordern, dass g und h komplanar sind.

Beweis von Satz 9.6: Seien $g = a + U_g$ und $h = b + U_h$.

(i) Behauptung: $g \parallel h \Leftrightarrow U_g = U_h$.

$$g \parallel h \Leftrightarrow U_g \subseteq U_h \vee U_h \subseteq U_g \underset{\substack{\dim U_g \\ = \dim U_h = 1}}{\overset{\Leftarrow}{\Rightarrow}} U_g = U_h.$$

(ii) $g \parallel h \underset{\text{(i)}}{\Rightarrow} g, h$ sind Nebenklassen nach dem gleichen Unterraum $\Rightarrow g = h \vee g \cap h = \emptyset$.

(iii) Beh.: $g \cap h = \emptyset \Rightarrow g \parallel h$. Beweis durch Kontraposition: $g \nparallel h \Rightarrow g \cap h \neq \emptyset$.

Sei $g \nparallel h$. Nach (i) ist $U_g = \langle p \rangle \neq \langle q \rangle = U_h$; somit gilt $V = \langle p, q \rangle$ (wegen $\dim_K = 2$) und daher $\exists \lambda, \mu \in K$ mit $a - b = p\lambda + q\mu$; dann ist $a + p(-\lambda) = b + q\mu \in g \cap h$.□

Aufgabe 9.1:
Zeigen Sie für einen beliebigen K-Vektorraum V: Zu jeder Geraden g und jedem Punkt P aus $\mathrm{AG}(V)$ existiert genau eine Gerade $h \in \mathrm{AG}(V)$ mit $P \,\mathrm{I}\, h$ und $g \parallel h$ (**Euklidisches Parallelenaxiom**).

Zusammenfassende Bemerkung

In der affinen Geometrie eines Vektorraums V werden außer den Unterräumen von V auch alle Bilder dieser Unterräume unter Translationen, also die zu ihnen parallelen affinen Unterräume betrachtet.
Dadurch gelangt man zu Modellen (z.B. der Zeichenebene bzw. des Anschauungsraumes), die sich nicht auf die Unterräume durch einen festen Punkt (Ursprung) beschränken; die Bindung an den Ursprung wird so abgemildert.

Ziel

Wir wollen den Zusammenhang von affinen Unterräumen und den Lösungsräumen von linearen Gleichungssystemen untersuchen. Dazu benötigen wir auch die Theorie der Vektorraum-Homomorphismen (linearen Abbildungen).

§ 10 Lineare Abbildungen und ihre Anwendung auf lineare Gleichungssysteme I

Generalvoraussetzung: Seien K Körper und V, V' K-Vektorräume.
Um Strukturähnlichkeiten auch bei nicht-isomorphen Vektorräumen beschreiben und untersuchen zu können, definieren wir Vektorraum-Homomorphismen. Diese lassen sich darüber hinaus auf die Theorie Linearer Gleichungssysteme, auf geometrische Abbildungen und auf Abbildungen aus der Analysis anwenden.

10.1 Definitionen: Lineare Abbildung, VR-Isomorphismus u.ä.

(Wiederholung von 8.7)

(i) Eine Abbildung $f : V \to V'$ heißt **lineare Abbildung** (lineare Transformation, ***K*-Homomorphismus**, **Vektorraum-Homomorphismus**, **linearer Operator**), wenn gilt:

$$(1)\ \forall v, w \in V : f(v + w) = f(v) + f(w) \qquad \text{(Additivität)}$$

$$(2)\ \forall v \in V\ \forall \lambda \in K : f(v \cdot \lambda) = f(v) \cdot \lambda \qquad \text{(Homogenität)}$$

Anmerkung.
Ein Vektorraum-Homomorphismus ist insbesondere Gruppen-Homomorphismus von $(V, +)$ in $(V', +')$.

(ii) Ein **Vektorraum-Isomorphismus** ist ein *bijektiver* Vektorraum-Homomorphismus.

(iii) Ein *surjektiver* Homomorphismus heißt **Epimorphismus**.

Ein *injektiver* Homomorphismus heißt **Monomorphismus**.

(iv) Ein Homomorphismus von V *in sich* heißt **Endomorphismus**.

Ein Isomorphismus von V *auf sich* heißt **Automorphismus**.
Automorphismen dienen insbesondere der Beschreibung von Symmetrien in Vektorräumen.

10.2 Beispiele

(a) Für $1 \leq k \leq n$ ist die k-te Projektion $\mathrm{pr}_k : \begin{cases} K^n \to K \\ (\xi_1, \ldots, \xi_k, \ldots, \xi_n) \mapsto \xi_k \end{cases}$ linear.

(b) Für $\alpha \in K$ ist $\sigma_\alpha : \begin{cases} K^n \to K^n \\ (\xi_1, \ldots, \xi_n) \mapsto (\xi_1\alpha, \ldots \xi_n\alpha) = (\xi_1, \ldots, \xi_n)\alpha \end{cases}$
eine lineare Abbildung (geometrische Interpretation: Streckung mit Zentrum $\mathfrak{o}$ und Streckfaktor α).

(c) Nach folgender Anmerkung definieren wir eine weitere lineare Abbildung.

10.3 Anmerkung: Matrixschreibweise

Seien $n, m \in \mathbb{N}$ und $\alpha_{ij} \in K$ für $i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$!
a_{ij} ist mit „doppeltem Index“ versehen; diese Schreibweise benutzen wir abkürzend für $\alpha_{(i,j)}$; die Familie $A = (\alpha_{ij})_{(i,j)\in\{1,\ldots,m\}\times\{1,\ldots,n\}}$ gibt man oft in „Matrizenform“ an.

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

und nennt sie **Matrix**. Die Menge aller dieser Matrizen über K mit festem n und m bezeichnen wir mit $K^{(m,n)}$.

Besonders eignet sich die Matrixschreibweise für die Angabe der Koeffizienten eines linearen Gleichungssystems:

$$(*) \quad \begin{cases} \alpha_{1\,1}\xi_1 + \alpha_{1\,2}\xi_2 + \cdots + \alpha_{1\,n}\xi_n = \beta_1 \\ \vdots \\ \alpha_{m\,1}\xi_1 + \alpha_{m\,2}\xi_2 + \cdots + \alpha_{m\,n}\xi_n = \beta_m \end{cases}$$

bzw. abgekürzt

$$\begin{cases} \sum_{j=1}^{n} \alpha_{1\,j}\xi_j = \beta_1 \\ \vdots \qquad\qquad \vdots \\ \sum_{j=1}^{n} \alpha_{m\,j}\xi_j = \beta_m \end{cases}$$

10.4 Beispiele (Fortsetzung)

(a) (i) (Fortsetzung von 10.2(c))
Durch Anmerkung 10.3 motiviert definieren wir eine Abbildung

$$f_A : \begin{cases} K^n \to K^m \\ (\xi_1, \ldots, \xi_n) \mapsto \left(\sum_{j=1}^{n} \alpha_{1\,j}\xi_j, \ldots, \sum_{j=1}^{n} \alpha_{m\,j}\xi_j\right) \end{cases}$$

und stellen fest, dass $(\xi_1, \ldots, \xi_n)$ genau dann eine Lösung von $(*)$ ist, wenn

$$f_A(\xi_1, \ldots, \xi_n) = (\beta_1, \ldots, \beta_m)$$

gilt.

Es ist also sinnvoll, sich mit Abbildungen der Form f_A zu beschäftigen.

(ii) Behauptung: Es gilt folgender Satz:

f_A ist ein Vektorraum-Homomorphismus, also linear.

Beweis.

$$\begin{aligned} f_A\big((\xi_1,\dots,\xi_n)+(\eta_1,\dots,\eta_n)\big) &= \big(\sum_{j=1}^{n}\alpha_{1\,j}(\xi_j+\eta_j),\dots,\sum_{j=1}^{n}\alpha_{m\,j}(\xi_j+\eta_j)\big)\\ &= \big(\sum_{j=1}^{n}\alpha_{1\,j}\xi_j+\sum_{j=1}^{n}\alpha_{1\,j}\eta_j,\dots,\sum_{j=1}^{n}\alpha_{m\,j}\xi_j+\sum_{j=1}^{n}\alpha_{m\,j}\eta_j\big)\\ &= f_A\big((\xi_1,\dots,\xi_n)\big)+f_A\big((\eta_1,\dots,\eta_n)\big);\\ f_A\big((\xi_1,\dots,\xi_n)\cdot\lambda\big) &= f_A\big((\xi_1\lambda,\dots,\xi_n\lambda)\big)\\ &= \big(\sum_{j=1}^{n}\alpha_{1\,j}\xi_j\lambda,\dots,\sum_{j=1}^{n}\alpha_{m\,j}\xi_j\lambda\big)\\ &= f_A\big((\xi_1,\dots,\xi_n)\big)\cdot\lambda \quad . \end{aligned}$$

Alternativer Beweis (Idee):
Das Skalarprodukt $z_i \cdot x$ für jede Zeile z_i von A ist linear.

Anmerkungen:

(1) In Beispiel (a) sind die Beispiele 10.2 (a) und (b) als Spezialfälle enthalten:

(i) Wählt man in (a) $m = 1$, $\alpha_{1\,k} = 1$ und $\alpha_{1\,j} = 0$ für $j \neq k$, dann wird die Zuordnungsvorschrift zu $(\xi,\dots,\xi_n) \mapsto (1 \cdot \xi_k)$. Und (ξ_k) identifizieren wir mit ξ_k. Die zugehörige Matrix ist $A = (0\dots0\;1\;0\dots0)$ (mit 1 an der k-ten Stelle). (Projektion, s.o.)

(ii) Setzt man in (a) $n = m$ und [42] $\alpha_{i\,j} = \alpha\delta_{i\,j}$, so ergibt sich $(\xi_1,\dots,\xi_n) \mapsto (\alpha\delta_{1\,1}\xi_1,\dots,\alpha\delta_{n\,n}\xi_n) = (\alpha\xi_1,\dots,\alpha\xi_n) = (\xi_1,\dots,\xi_n)\cdot\alpha$. Die zugehörige Matrix ist $\begin{pmatrix} \alpha & 0 & \dots & 0 & 0\\ 0 & \alpha & \dots & 0 & 0\\ \vdots & & \ddots & & \vdots\\ 0 & 0 & \dots & \alpha & 0\\ 0 & 0 & \dots & 0 & \alpha \end{pmatrix}$; (zentische Streckung, s.o.)

(2) Wählt man $\alpha_{i\,j} = 0$ für alle $i \in \{1,\dots,m\}, j \in \{1,\dots,n\}$, so erhält man die „Nullabbildung“ $K^n \to K^m$ mit $(\xi,\dots,\xi_n) \mapsto \mathfrak{o}$.

(3) (Beispiel aus der Codierungstheorie; vgl.(1.5) !)

(i) *Zufügen eines Kontrollbits*

Für $K = \mathrm{GF}(2)$ ist die Codierung $K^4 \to K^5$, die durch $a_1a_2a_3a_4 \mapsto a_1a_2a_3a_4a_5$ mit $a_5 = a_1 \oplus a_2 \oplus a_3 \oplus a_4$ (Paritätskontrollbit) definiert ist, eine lineare Abbildung; zu ihr gehört die Matrix

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 1 & 1 & 1 & 1 \end{pmatrix}$$

[42] Erinnerung: $\delta_{i\,j} := \begin{cases} 0 & \text{für} \quad i \neq j\\ 1 & \text{für} \quad i = j \end{cases}$.

(vgl. Anmerkung 10.3 !)

(ii) *Paritätskontrolle*

Auch die lineare Abbildung $f_{A_2} : K^5 \to K$ definiert durch $a_1a_2a_3a_4a_5 \mapsto a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5$ mit Matrix $A_2 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$ ist von Bedeutung: Wegen $a_1 \oplus a_2 \oplus \ldots \oplus a_4 \oplus a_5 = 2a_1 \oplus \ldots \oplus 2a_4 = 0$ ist die Menge der Codewörter gerade gleich Kern f_{A_2}. (Vgl. Bemerkung 10.6, s.u.!)

(4) (Ein Beispiel aus der Populationsdynamik):

Ein einfaches Modell für die Entwicklung eines Käfers sieht folgendermaßen aus: Aus den Eiern schlüpfen nach einem Monat Larven, nach einem weiteren Monat werden diese zu Käfern, die dann wieder Eier legen und anschließend sterben. Die Übergänge bei jedem „Generationswechsel“ seien gegeben durch folgende Tabelle

	Eier	Larven	Käfer	(vor dem Generationenwechsel)
Eier	0	0	s_3	
Larven	s_1	0	0	
Käfer	0	s_2	0	
(nach einem Monat)				

Dabei sei s_1 die Wahrscheinlichkeit, dass aus einem gelegten Ei eine Larve schlüpft, s_2 die Wahrscheinlichkeit, dass aus einer Larve ein Käfer entsteht und s_3 die Anzahl der Eier, die ein Käfer im Monat durchschnittlich legt.

Die Abbildung, $f : \mathbb{R}^3 \to \mathbb{R}^3$ definiert durch

$$(n_1, n_2, n_3) \mapsto (n_1', n_2', n_3') = (s_3n_3, s_1n_1, s_2n_2)$$

gibt die Veränderung der Anzahlen der „Generationen“ wieder.

Die Abbildung f ist linear; zu ihr gehört die Matrix

$$\begin{pmatrix} 0 & 0 & s_3 \\ s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{pmatrix}$$

(vgl. Bsp. (c) und obige Tabelle).

Weitere Literatur:
- J. Lighthill: Newer uses of mathematics, Penguin 1978
- E. Lehmann: Matrizenrechnung, BSV 1975
- E. Lehmann: Endliche homogene Markoffsche Ketten, BSV 1973.

(b) (Ein Beispiel aus der Analysis):

Sei V der $\mathbb{R}$-Vektorraum aller konvergenten reellen Zahlenfolgen, also $V = \left\{(\alpha_n)_{n\in\mathbb{N}} \middle| \lim_{n\to\infty} \alpha_n \text{ existiert}\right\}$ mit komponentenweiser Addition und S-Multiplikation:

$$(\alpha_n)_{n\in\mathbb{N}} + (\beta_n)_{n\in\mathbb{N}} := (\alpha_n + \beta_n)_{n\in\mathbb{N}} \text{ und } (\alpha_n)_{n\in\mathbb{N}} \cdot \lambda := (\alpha_n\lambda)_{n\in\mathbb{N}}.$$

V ist dann Unterraum von $\mathbb{R}^{\mathbb{N}}$. Wegen $\lim\limits_{n\to\infty}(\alpha_n + \beta_n) = \lim\limits_{n\to\infty}\alpha_n + \lim\limits_{n\to\infty}\beta_n$ und $\lim\limits_{n\to\infty}(\alpha_n\alpha) = \big(\lim\limits_{n\to\infty}\alpha_n\big)\alpha$ ist die folgende Abbildung ebenfalls eine lineare Abbildung:

$$\lim : \begin{cases} V \to \mathbb{R} \\ (\alpha_n)_{n\in\mathbb{N}} \mapsto \lim\limits_{n\to\infty}\alpha_n \end{cases}$$

(c) (Weiteres Beispiel aus der Analysis):

Sei $V = V' := \mathcal{P}(\mathbb{R}) = \{f \mid f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n\}$, der $\mathbb{R}$-Vektorraum der Polynomabbildungen, sowie

$$\mathrm{D} : \begin{cases} \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R}) \\ f \mapsto \mathrm{D}\,f \end{cases} \quad \text{mit } (\mathrm{D}\,f)(x) := \frac{d}{dx}f(x) = f'(x) \quad .$$

Nach den Sätzen der Differential-Rechnung gilt

$$\big[\mathrm{D}(f_1 + f_2)\big](x) = (f_1 + f_2)'(x) = f_1'(x) + f_2'(x) = \big[(\mathrm{D}\,f_1) + (\mathrm{D}\,f_2)\big](x)$$

und

$$\big[\mathrm{D}(\alpha f)\big](x) = (\alpha f)'(x) = \alpha f'(x) = \big[\alpha(\mathrm{D}\,f)\big](x) \text{ für alle } x \in \mathbb{R}.$$

Daher ist D linear, also Vektorraum-Homomorphismus.

(d) **Beispiel** (in Fortsetzung von (1.4)):

Sei V der $\mathbb{R}$-Vektorraum der zweimal differenzierbaren reellen Funktionen und $\mathrm{D} : V \to \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ definiert durch $\mathrm{D}\,f = f'' + 4f$. Dann ist D linear:

$$\mathrm{D}(\lambda f + g) = (\lambda f + g)'' + 4(\lambda f + g) = \lambda(f'' + 4f) + (g'' + 4g) = \lambda\,\mathrm{D}\,f + \mathrm{D}\,g\ .$$

Der Lösungsraum der homogenen Differentialgleichung $y'' + 4y = 0$ (vgl. (1.4)) ist gerade gleich dem Kern von D (s. auch 10.6 !).

(e) Für $V = \mathcal{C}[a, b]$ definiert $f \mapsto \int\limits_a^b f(t)\mathrm{d}t$ eine lineare Abbildung von V in $\mathbb{R}$.

10.5 Anmerkung: Multiplikation einer Matrix mit einem Vektor

Im Fall endlicher Dimensionen von V und V' ist es üblich, die Abbildung $x \mapsto f_A(x)$ als eine Multiplikation von x mit der Matrix A zu interpretieren. Dazu beachten wir, dass wir das lineare Gleichungssystem $(*)$ auch in folgender Form schreiben können:

$$\begin{pmatrix} \sum\limits_{j=1}^{n}\alpha_{1j}\xi_j \\ \vdots \\ \sum\limits_{j=1}^{n}\alpha_{mj}\xi_j \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \vdots \\ \beta_m \end{pmatrix}.$$

Bei dieser Schreibweise ist die linke Seite gleich dem als Spalte geschriebenen Vektor $f_A(x) = (\sum_{j=1}^{n} \alpha_{1j}\xi_j, \ldots, \sum_{j=1}^{n} \alpha_{mj}\xi_j)$. Konsequenter Weise schreiben wir dann auch die Vektoren b und x als Spalten:

$$b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \quad \text{und} \quad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

Eine **Multiplikation zwischen der Matrix A und dem Spaltenvektor x** wird dann definiert, indem man $A \cdot x$ als den Spaltenvektor $f_A(x)$ nimmt; den $i-$ten Eintrag von $A \cdot x$ erhält man also als kanonisches Skalarprodukt der $i-$ten Zeile von A mit dem Spaltenvektor x. (Dies setzt natürlich voraus, dass die Anzahl der Spalten von A gleich der der Zeilen von x ist.)
Schematisch:

$$\begin{pmatrix} \underline{\text{1.Zeile}} \\ \\ \end{pmatrix} \begin{pmatrix} S \\ p \\ a \\ l \\ t \\ e \end{pmatrix} = \begin{pmatrix} \bullet \\ \end{pmatrix}, \ldots \begin{pmatrix} \underline{\text{i.te Zeile}} \\ \\ \end{pmatrix} \left(\begin{array}{c|} \\ \\ \end{array} \right) = \begin{pmatrix} \bullet\, i. \\ \end{pmatrix}, \ldots, \begin{pmatrix} \\ \underline{\text{letzte Zeile}} \end{pmatrix} \left(\begin{array}{c|} \\ \\ \end{array} \right) = \begin{pmatrix} \\ \bullet \end{pmatrix}.$$

Der $i-$te Schritt $(i = 1, \ldots, m)$ bei dieser Multiplikation wird beschrieben durch das folgende Schema:

$$\begin{array}{l} \\ \\ i\text{-te} \\ \text{Zeile} \end{array} \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & & \vdots \\ \hline \alpha_{i1} & \ldots & \alpha_{in} \\ \hline \vdots & & \vdots \\ \alpha_{r1} & \ldots & \alpha_{mn} \end{pmatrix} \rightarrow \begin{array}{c} \text{mal Spalte} \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \\ \downarrow \\ \left(\begin{array}{c} \vdots \\ \hline \gamma_i \\ \hline \vdots \end{array} \right) \end{array} \text{ mit } \gamma_i = \sum_{j=1}^{n} \alpha_{ij}\xi_j$$

Dabei ist, wie oben schon erwähnt, γ_i das kanonische Skalarprodukt zwischen der $i-$ten Zeile $a_{i\bullet} := (\alpha_{i1} \ldots \alpha_{in})$ von A und dem Vektor $x = (\xi_1 \ldots \xi_n)$.
Das lineare Gleichungssystem $(*)$ können wir nun in der Form

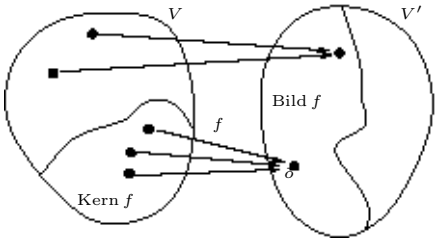
$$\boxed{A \cdot x = b}$$

schreiben.
Nach etwas Theorie werden wir diese auf lineare Gleichungssyteme anwenden:

10.6 Bemerkung: lineare Abbildung als Gruppen-Homomorphismus

Ein Vektorraum-Homomorphismus $f : V \to V'$ ist (wegen der Additivität) insbesondere ein Gruppen Homomorphismus $(V,+) \to (V',+)$.(Vgl. auch Anmerkung (i) nach 10.1 !) Damit sind Definitionen und Aussagen aus §4 auf f anwendbar.

U.a. sind definiert
Kern $f = \{v \in V | f(v) = \mathfrak{o}\}$ und Bild$f$ = Im$f = \{v' \in V' | \exists v \in V : f(v) = v'\} = f(V)$.
Nach (4.4) sind Kern f und Bild f Untergruppen von $(V,+)$ bzw. $(V',+)$. Es gilt darüber hinaus:



Figur 10.1: Kern und Bild einer linearen Abbildung.

10.7 Satz (Kern und Bild einer linearen Abbildung)

Ist $f : V \to V'$ Vektorraum-Homomorphismus, so gilt

(a) Kern f ist ein Unterraum von V.

(b) Bild f ist ein Unterraum von V'.

Beweis.

(a) Kern f ist Untergruppe von $(V,+)$ (s.o.), insbesondere additiv abgeschlossen und nicht leer. Außerdem gilt

$$v \in \text{Kern}\, f \Rightarrow f(v) = \mathfrak{o} \Rightarrow f(v\lambda) = f(v) \cdot \lambda = \mathfrak{o} \Rightarrow v \cdot \lambda \in \text{Kern}\, f.$$

Damit ist Kern f abgeschlossen bzgl. Addition und S-Multiplikation. Aus dem Unterraum-Kriterium (6.4) folgt die Behauptung.

(b) Für $\lambda \in K$ und $v' \in$ Bild f gilt $\exists v \in V : f(v) = v'$ und $f(v\lambda) = f(v) \cdot \lambda = v'\lambda \in$ Bild f. Im Übrigen verläuft der Beweis wie in (a). □

10.8 Anwendung

(i) Fortsetzung von Beispiel (10.4) (a), jetzt für Spalten- statt Zeilenvektoren:

$$\text{Seien } A = \begin{pmatrix} \alpha_{1\,1} & \dots & \alpha_{1\,n} \\ \vdots & & \vdots \\ \alpha_{m\,1} & \dots & \alpha_{m\,n} \end{pmatrix} \text{ mit } \alpha_{ij} \in K \text{ und } f_A : \begin{cases} K^{(n,1)} \to K^{(m,1)} \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \mapsto \begin{pmatrix} \sum_{j=1}^{n} \alpha_{1\,j}\xi_j, \\ \vdots \\ \sum_{j=1}^{n} \alpha_{m\,j}\xi_j \end{pmatrix} \end{cases},$$

also $f_A(x) = Ax$. Dann gilt:

$$\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} \in \operatorname{Kern} f_A \quad \Leftrightarrow (**) \quad \begin{cases} \alpha_{1\,1}\xi_1 + \cdots + \alpha_{1\,n}\xi_n = 0 \\ \vdots \qquad\qquad\qquad \vdots\,\vdots \\ \alpha_{m\,1}\xi_1 + \cdots + \alpha_{m\,n}\xi_n = 0 \end{cases}$$

Das System $(**)$ heißt **System von m homogenen linearen Gleichungen in n Unbekannten mit Koeffizientenmatrix A**. Die Menge aller Lösungen eines Gleichungssystems, d.h. die Menge aller n-Tupel $\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} \in K^{(n,1)}$, die das Gleichungssystem erfüllen, nennen wir **Lösungsraum**. Aus (10.7) folgt wegen $L_0 = \operatorname{Kern} f_A$ damit:

10.9 Satz (Lösungsraum eines homogenen linearen Gleichungssystems)

Der Lösungsraum L_0 eines homogenen linearen Gleichungssystems in n Unbekannten über dem Körper K ist ein Unterraum von K^n, also selbst K-Vektorraum.

Anmerkung. Vgl. dazu Beispiel (1.3).
Fortsetzung von Beispiel (10.4) (a)

$$\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix} \in \operatorname{Bild} f_A \Leftrightarrow \exists \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} \in K^{(n,1)} : \begin{pmatrix}\sum_{j=1}^{n} \alpha_{1\,j}\xi_j\\ \vdots\\ \sum_{j=1}^{n} \alpha_{m\,j}\xi_j\end{pmatrix} = \begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}$$

$$\Leftrightarrow \exists \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} \in K^{(n,1)} : \begin{cases} \alpha_{1\,1}\xi_1 + \cdots + \alpha_{1\,n}\xi_n = \beta_1 \\ \vdots \qquad\qquad \vdots \quad \vdots \\ \alpha_{m\,1}\xi_1 + \cdots + \alpha_{m\,n}\xi_n = \beta_m \end{cases} \qquad (*)$$

Für das lineare Gleichungssystem $(*)$ in n Unbekannten über dem Körper K mit Koeffizientenmatrix A halten wir fest:

10.10 Satz (Lösbarkeit eines LGS)

Genau dann ist das lineare Gleichungssystem

$$(*) \quad \begin{cases} \alpha_{1\,1}\xi_1 + \cdots + \alpha_{1\,n}\xi_n = \beta_1 \\ \vdots \qquad\qquad \vdots \quad \vdots \\ \alpha_{m\,1}\xi_1 + \cdots + \alpha_{m\,n}\xi_n = \beta_m \end{cases}$$

über K lösbar, wenn gilt:

$$\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix} \in \operatorname{Bild} f_A \quad \text{mit} \quad f_A : \begin{cases} K^{(n,1)} \to K^{(m,1)} \\ x = \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} \mapsto \begin{pmatrix}\sum_{j=1}^{n} \alpha_{1\,j}\,\xi_j\\ \vdots\\ \sum_{j=1}^{n} \alpha_{m\,j}\,\xi_j\end{pmatrix} = Ax \quad (= f_A(x)) \end{cases} .$$

Weitere Anwendungs-Beispiele

(ii) Vgl. mit Beispiel (10.4c) !

$$\text{Für} \qquad D : \begin{cases} \mathbb{R}[x] \to \mathbb{R}[x] \\ \sum_{i=0}^{n} \alpha_i x^i \mapsto \sum_{i=1}^{n} i\alpha_i X^{i-1} \end{cases} \qquad \text{gilt}$$

$$\operatorname{Kern} D = \{f(x) \in \mathbb{R}[x] \mid f(x) \text{ konstant}\}.$$

Beweisskizze: Aus $\sum i\alpha_i x^i = 0$ für alle $x \in \mathbb{R}$ folgt $\alpha_1 = \ldots \alpha_n = 0$, also $f(x) = \alpha_0$ für alle $x \in \mathbb{R}$.
Alternativ kann man dem Mittelwertsatz der Differentialrechnung argumentieren:

$$f' \equiv 0 \Rightarrow \exists\, \xi \in (x, y) : \quad 0 = f'(\xi) = \frac{f(x) - f(y)}{x - y} \Rightarrow f(x) = f(y).$$

(iii) Fortsetzung von Beispiel (10.4) (b):

Ist V der Vektorraum aller konvergenten reellen Zahlenfolgen und $f = \lim\limits_{n\to\infty}$, so ist $\operatorname{Kern} f$ der Unterraum aller reellen Nullfolgen.

(iv) Fortsetzung von Beispiel (10.4d) und (1.4):

In Beispiel (10.4d) hatten wir gesehen, dass der Lösungsraum L der homogenen Differentialgleichung $y'' + 4y = 0$ Kern einer linearen Abbildung vom Vektorraum V der zweimal differenzierbaren Funktionen in den Vektorraum $\operatorname{Abb}(\mathbb{R}, \mathbb{R})$ ist. Daher ist L ein Unterraum von V; (vgl. (1.4)!).

Als weitere Aussage erhalten wir aus §4 unmittelbar:

10.11 Hilfssatz (Kern und Injektivität)

Ist $f : V \to V'$ Vektorraum-Homomorphismus, so gilt:

(i) $v + \operatorname{Kern} f$ ist das volle Urbild von $f(v)$.

(ii) f injektiv $\Leftrightarrow \operatorname{Kern} f = \{\mathfrak{o}\} \Leftrightarrow |\operatorname{Kern} f| = 1$.

Beweis. s. (4.5) und (4.6).

Wegen der Kommutativität von $(V, +)$ ist jeder Unterraum von V, als Untergruppe von $(V, +)$ aufgefasst, Normalteiler. Satz (4.9) liefert also keine über (10.7) (a) hinausgehende Information.

Fortsetzung von **Beispiel** (10.4) (a): Aus (10.11) (i) erhalten wir u.a.:

10.12 Satz (Lösungsraum eines linearen Gleichungssystems)

Für den Lösungsraum L eines linearen Gleichungssystems

$$(*) \quad \begin{cases} \alpha_{1\,1}\xi_1 + \cdots + \alpha_{1\,n}\xi_n = \beta_1 \\ \vdots \qquad\qquad \vdots \quad \vdots \\ \alpha_{m\,1}\xi_1 + \cdots + \alpha_{m\,n}\xi_n = \beta_m \end{cases}$$

über K gilt entweder $L = \emptyset$ oder $L = p + L_0$;

hierbei bezeichnet $p \in K^{(n,1)}$ eine spezielle Lösung von $(*)$, (eine sogenannte **Partikulärlösung**), und L_0 den Lösungsraum des zu $(*)$ gehörenden homogenen Systems

$$(**) \quad \begin{cases} \alpha_{1\,1}\xi_1 + \cdots + \alpha_{1\,n}\xi_n = 0 \\ \vdots \qquad\qquad \vdots \quad \vdots \\ \alpha_{m\,1}\xi_1 + \cdots + \alpha_{m\,n}\xi_n = 0 \end{cases} .$$

Beweis.

Ist $A := \begin{pmatrix} \alpha_{1\,1} & \dots & \alpha_{1\,n} \\ \vdots & & \vdots \\ \alpha_{m\,1} & \dots & \alpha_{m\,n} \end{pmatrix}$, $b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$, so folgt

$$L = \{ \begin{pmatrix} \xi_1 \\ \vdots \\ xi_n \end{pmatrix} \in K^{(n,1)} \mid f_A\big(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \big) = b \} = f_A^-(b).$$

Falls $b \notin$ Bild f_A gilt, erhält man $L = \emptyset$. Ist $L \neq \emptyset$, so existiert ein $p \in K^n$ mit $f_A(p) = b$; es folgt $L = p + \text{Kern } f_A$ (s. (10.11) (i)). Dabei ist Kern $f_A = \{ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n | f_A\big(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \big) = \{\mathfrak{o}\}\} =: L_0$ der Lösungsraum des zugehörigen homogenen Gleichungssystems $(**)$.

Beispiel: Sei $K = \mathbb{R}$, $n = 3$, $m = 2$

$$(*) \; : \; \begin{cases} 2\xi_1 + 3\xi_2 - \xi_3 = 4 \\ \xi_1 + \xi_2 + \xi_3 = 3 \end{cases}$$

Wir suchen zunächst eine Partikulärlösung. Aus $(*)$ ergibt sich (2 mal 2. Zeile minus 1. Zeile)

$$\begin{cases} 2\xi_1 + 3\xi_2 - \xi_3 = 4 \\ -\xi_2 + 3\xi_3 = 2 \end{cases}$$

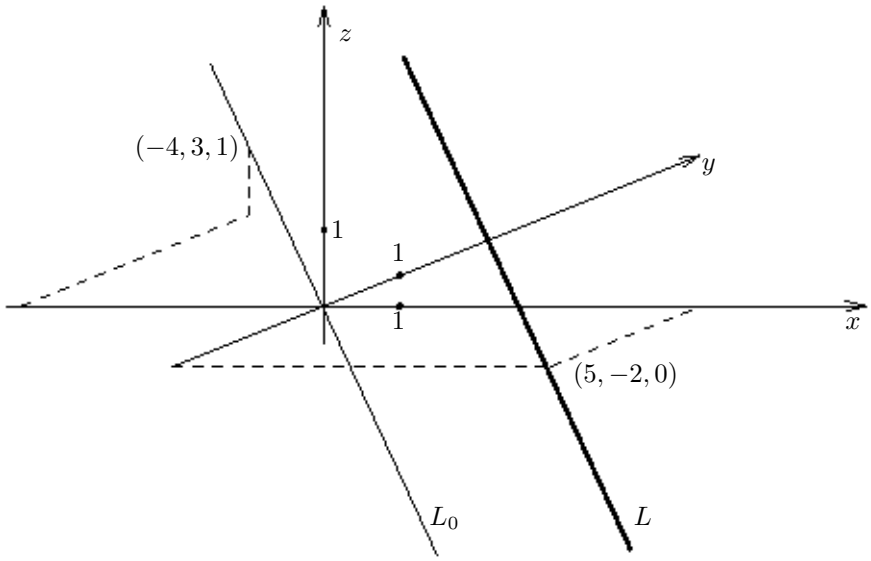
Setzt man nun z.B. $\xi_3 = 0$, so folgt $\xi_2 = -2$ und $\xi_1 = 5$. Tatsächlich ist $p = (5, -2, 0)$ (Partikulär-) Lösung. (Probe wegen Beweisrichtung**!**)
Das zu $(*)$ gehörige homogene System ist

$$(**) \; : \; \begin{cases} 2\xi_1 + 3\xi_2 - \xi_3 = 0 \\ \xi_1 + \xi_2 + \xi_3 = 0 \end{cases} \qquad \text{bzw.} \qquad (**') \; : \; \begin{cases} 2\xi_1 + 3\xi_2 - \xi_3 = 0 \\ -\xi_2 + 3\xi_3 = 0 \end{cases}$$

Setzt man $\xi_3 = 0$, so folgt $\xi_1 = \xi_2 = 0$. Daher können wir o.B.d.A. von $\xi_3 \neq 0$ ausgehen. Der Lösungsraum ist invariant bzgl. Multiplikation mit einem Skalar. Bis auf ein skalares Vielfaches des Lösungsvektors ist $\xi_3 = 1$; so ergibt sich $\xi_2 = 3$ und $\xi_1 = -4$. Tatsächlich ist $\begin{pmatrix} -4 \\ 3 \\ 1 \end{pmatrix}$ Lösung (Wegen der Beweisrichtung: Probe!). Daher ist $L_0 = \left\langle \begin{pmatrix} -4 \\ 3 \\ 1 \end{pmatrix} \right\rangle$ Lösungsraum von $(**)$. Der Lösungsraum von $(*)$ ist damit

$$L = \begin{pmatrix} 5 \\ -2 \\ 0 \end{pmatrix} + \begin{pmatrix} -4 \\ 3 \\ 1 \end{pmatrix} \mathbb{R}.$$

Im affinen Raum $\mathrm{AG}(\mathbb{R}^3)$ ist L eine Gerade durch den Punkt $(5, -2, 0)$. Diese Gerade ist also durch das Gleichungssystem $(*)$ beschreibbar.



Figur 10.2: Geometrische Interpretation des Lösungsraums L eines Linearen Gleichungssystems und des Lösungsraums L_0 des zugehörigen homogenen Systems.

Aus (10.12)(ii) folgt auch:

10.13 Satz (Eindeutigkeit der Lösung)

Das lineare Gleichungssystem

$$(*) \quad \begin{cases} \alpha_{1\,1}\xi_1 + \cdots + \alpha_{1\,n}\xi_n = \beta_1 \\ \vdots \qquad\qquad\qquad \vdots \quad \vdots \\ \alpha_{m\,1}\xi_1 + \cdots + \alpha_{m\,n}\xi_n = \beta_m \end{cases}$$

über K besitzt höchstens dann eine eindeutige Lösung, wenn das zugehörige homogene System nur die triviale Lösung besitzt.

§ 11 Lineare Gleichungssysteme II: Rang von Matrizen, elementare Zeilenumformungen und Gaußsche Elimination

M ↓

Im Laufe der Entwicklung der Vektorraumtheorie haben wir die Lösungsstruktur linearer Gleichungssysteme untersucht. Die dabei erhaltenen Ergebnisse erleichtern zwar die Angabe der Lösung, liefern aber kein Verfahren zur praktischen Bestimmung. Ein solches behandeln wir im Folgenden. (Weitere Verfahren können Sie u.a. in der Vorlesung über Numerische Mathematik kennenlernen.)

↑ M

Teil I: Rang und Lösbarkeit

11.1 Voraussetzungen und Schreibweisen

Gegeben sei (über dem Körper K) das lineare Gleichungssystem (LGS)

$$(*)\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n = \beta_1 \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \ldots + \alpha_{2n}\xi_n = \beta_2 \\ \vdots \qquad\qquad\qquad\qquad \vdots \quad \vdots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n = \beta_m \end{cases}$$

mit Koeffizientenmatrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix} .$$

Letztere ist eine $(m \times n)-$Matrix mit Einträgen $\alpha_{i,j} \in K$, in Zeichen:
$A \in K^{(m,n)} := K^{\{1,\ldots,m\}\times\{1,\ldots,n\}}$. (Vgl. Â§10 !)
Die Elemente $\beta_1, \ldots, \beta_m \in K$ der rechten Seite schreiben wir jetzt als Spaltenvektor, also als $(m \times 1)-$Matrix:

$$b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in K^{(m,1)} .$$

Auch eine Lösung von $(*)$ geben wir in Zukunft meist als Spaltenvektor an: Definitionsgemäß ist dann

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{(n,1)}$$

Lösung von $(*)$, falls (mit $a_{i\bullet} := (\alpha_{i1}, \ldots, \alpha_{in})$) gilt:

$$A \cdot x := \begin{pmatrix} a_{1\bullet} \cdot x \\ \vdots \\ a_{m\bullet} \cdot x \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^{n} \alpha_{1j}\xi_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{mj}\xi_j \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \vdots \\ \beta_m \end{pmatrix} = b.$$

Es erweist sich als sinnvoll, A und b zusammenzufassen zur **erweiterten Koeffizientenmatrix** $(A|b)$, definiert durch

$$(A|b) := \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\ \vdots & & & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m \end{array}\right) .$$

Diese ist eine Matrix über K mit m Zeilen und $n+1$ Spalten, also ein Element von $K^{(m,n+1)}$. Der Strich dient dabei der besseren optischen Gliederung und erinnert an die Stellung der Spalten im LGS:

$$\begin{array}{|c|}\hline \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \\ \hline \end{array}\ \xi_1 + \begin{array}{|c|}\hline \alpha_{12} \\ \alpha_{22} \\ \\ \alpha_{m2} \\ \hline \end{array}\ \xi_2 + \ldots + \begin{array}{|c|}\hline \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \\ \hline \end{array}\ \xi_n = \begin{array}{|c|}\hline \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \\ \hline \end{array} \quad .$$

Beispiel: Das lineare Gleichungssystem

$$\left\{\begin{array}{l} 2\xi_1 + 3\xi_2 - \xi_3 = 4 \\ \ \xi_1 + \ \xi_2 + \xi_3 = 3 \end{array}\right.$$

lässt sich schreiben als

$$\begin{pmatrix} 2 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

Die erweiterte Koeffizientnematrix ist

$$\left(\begin{array}{ccc|c} 2 & 3 & -1 & 4 \\ 1 & 1 & 1 & 3 \end{array}\right) .$$

Wir behandeln nun ein **Kriterium für die Lösbarkeit** des LGS $(*)$: Es gilt:

$Ax = b$ besitzt eine Lösung $x \iff f_A((\xi_1, \dots, \xi_n)) = (\beta_1, \dots, \beta_m)$ ist lösbar.

$\iff (\beta_1, \dots, \beta_m) \in$ Bild f_A (Vgl. Satz 10.10 !)

$$\iff \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in \left\{ Ax \mid x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{(n,1)} \right\} = \left\{ \begin{pmatrix} \alpha_{11}\xi_1 + \ldots + \alpha_{1n}\xi_n \\ \vdots \quad \vdots \\ \alpha_{m1}\xi_1 + \ldots + \alpha_{mn}\xi_n \end{pmatrix} \mid \xi_j \in K \right\}$$

$$= \left\{ \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix} \xi_1 + \ldots + \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \xi_n \mid \xi_1, \dots, \xi_n \in K \right\} = \left\langle \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, \dots, \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \right\rangle .$$

Das LGS $(*)$ist also genau dann lösbar, wenn der Vektor b der rechten Seite in dem von den Spalten der Koeffizientenmatrix A aufgespannten Unterraum liegt, sich also durch Hinzunahme von b dieser Unterraum , d.h. auch dessen Dimension, nicht vergrößert. Wir vereinbaren folgende Bezeichnung:

11.2 Definition: Rang einer Matrix

Zu $A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}$ definiert man als (Spalten-) **Rang von A** die Dimension des von den Spaltenvektoren von A aufgespannten Unterraumes von $K^{(m,1)}$:

$$\operatorname{Rang} A := \operatorname{rg} A := \dim_K(\left\langle \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, \dots, \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \right\rangle) \ .$$

11.3 Definition (Rang einer linearen Abbildung) und Anmerkung

Für eine lineare Abbildung $f : V \to W$ definiert man:

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

Ist nun A eine Matrix über K und f_A wie oben definiert (d.h. durch Multiplikation mit der Matrix A gegeben), so gilt wegen Bild $f_A = \langle f_A(e_1), \dots, f_A(e_n)\rangle = \langle a_{\bullet 1}, \dots, a_{\bullet n}\rangle$:

$$\boxed{\operatorname{Rang} f_A = \operatorname{Rang} A.}$$

Beispiel: Sind $K = \mathbb{R}, m = 2, n = 3$ und $A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix}$, so erhält man

$$\operatorname{Rang} A = \dim_K < \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} >= 2,$$

da die angegebenen Spalten als Vektoren des 2-dim Vektorraum $\mathbb{R}^{(2,1)}$ einen Unterraum der Dimension höchstens und damit genau 2 erzeugen.

Mit den eben eingeführten Bezeichnungen können wir dann die vorigen Überlegungen zusammenfassen zu folgendem

11.4 Satz (Lösbarkeitskriterium)

(manchmal als Satz von Kronecker-Capelli bezeichnet, aber laut Wikipedia schon vor Leopold Kronecker und Alfredo Capelli von anderen Mathematikern verwendet)

Das lineare Gleichungssystem $(*)$ $Ax = b$ ist genau dann lösbar, wenn b von den Spaltenvektoren der Koeffizientenmatrix A linear abhängt, d.h. wenn der Rang der Koeffizientenmatrix A gleich dem der erweiterten Koeffizientenmatrix $A_{\text{erw}} := (A|b)$ ist, also:

$$\boxed{(*) \quad \text{ist lösbar} \iff \operatorname{Rang} A = \operatorname{Rang} A_{\text{erw}}}$$

11.5 Anmerkung (Dimension des Lösungsraums)

Wir werden später sehen, dass gilt:

$$\dim_K(\text{Kern} f_A) = n - \text{Rang}\, f_A \quad .$$

Im Fall der Lösbarkeit des LGS $(*)$ folgt daher (unter Verwendung von Satz 10.12) für den Lösungsraum L von $(*)$:
$\dim_K L = \dim_K(p + L_0) = \dim_K L_0 = \dim(\text{Kern} f_A) = n - \text{Rang}\, f_A = n - \text{Rang}\, A$.
Wir halten fest:

$$\boxed{\dim_K L = n - \text{Rang } A} \quad .$$

Teil II: Gaußsche Elimination

M
↓

Um zu einem Lösungsverfahren von Lineare Gleichungssystemen zu kommen, überlegen wir nun, welche Umformungen von (11.1) die Lösungsmenge von (11.1) „erhalten". Es sind dies u.a. :

(i) die Multiplikation der k–ten Gleichung mit einem Skalar $\lambda \in K \setminus \{0\}$ (für $k \in \{1, \ldots, m\}$)

(ii) die Addition der k–ten Gleichung zur i–ten Gleichung (für $i, k \in \{1, \ldots, m\},\ i \neq k$)

(iii) das Vertauschen zweier Gleichungen

(iv) die Addition des λ–fachen der k–ten Gleichung zur i–ten Gleichung (für $\lambda \in K$ und $i, k \in \{1, \ldots, m\},\ i \neq k$).

Wir bemerken, dass die Umformungen (iii) und (iv) durch mehrfache Anwendung von Umformungen der ersten beiden Typen erreicht werden kann, nämlich[43]

(iv) durch (i) und danach (ii) und

(iii) mittels $z_i \longrightarrow \overline{z_i} := z_i - z_k \longrightarrow \overline{z_k} := z_k + \overline{z_i} = z_i \longrightarrow \overline{\overline{z_i}} := -\overline{z_i} + \overline{z_k} = z_k$ (vgl.11.7!)

Untersucht man, wie sich die erwähnten Umformungen auf die erweiterte Koeffizientenmatrix von (11.1) auswirkt, so stößt man auf folgende sogenannte elementare Zeilenumformungen einer Matrix:

↑
M

11.6 Definition: Elementare Zeilenumformungen

Unter einer elementaren Zeilenumformung an einer $(m \times s)$–Matrix C über K (z.B. A oder A_{erw} für ein LGS) versteht man eine der folgenden Operationen:

(i) Die Multiplikation einer Zeile von C mit einem Skalar $\lambda \in K \setminus \{0\}$.

(ii) Die Addition der k–ten Zeile von C zur i–ten Zeile (für $i, k \in \{1, \ldots, m\},\ i \neq k$).

(iii) Das Vertauschen zweier Zeilen von C.

(iv) Die Addition des λ–fachen der k–ten Zeile von C zur i–ten Zeile (für $\lambda \in K$ und $i, k \in \{1, \ldots, m\}$, $i \neq k$).

[43] Dabei bezeichnen wir mit z_i die $i-$te Zeile (bzw. Gleichung), mit $\overline{z_i}$ die $i-$te Zeile nach der ersten Umformung dieser Zeile, mit $\overline{\overline{z_i}}$ die $i-$te Zeile nach der zweiten Umformung dieser Zeile usw..

11.7 Anmerkung

Die Umformungen (iii) und (iv) erhält man durch wiederholte Anwendungen der Umformungen der ersten beiden Typen. *Beweis* von (iii):

$$\begin{pmatrix} z_i \\ z_k \end{pmatrix} \longrightarrow \begin{pmatrix} z_i \\ z_k + z_i \end{pmatrix} \longrightarrow \begin{pmatrix} -z_i \\ z_k + z_i \end{pmatrix} \longrightarrow \begin{pmatrix} z_k + z_i - z_i \\ z_k + z_i \end{pmatrix} \longrightarrow \begin{pmatrix} z_k \\ z_i \end{pmatrix}.$$

Beweis von (iv): Übungsaufgabe.

Beispiel: Seien $K = \mathbb{R}$, $C = \begin{pmatrix} -1 & 2 & 3 & 1 \\ 1 & -1 & -3 & 0 \\ 0 & 2 & 0 & 4 \end{pmatrix}$ und $z_i, \bar{z}_i, \bar{\bar{z}}_i$ die Zeilenvektoren von C und den daraus erhaltenen Matrizen.

$$C \xrightarrow[z_2 \to \bar{z}_2 = z_2 + z_1]{} \begin{pmatrix} -1 & 2 & 3 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 4 \end{pmatrix} \xrightarrow[\bar{z}_3 \to \bar{\bar{z}}_3 = \bar{z}_3 - 2\bar{z}_2]{} \left(\begin{array}{cccc} -1 & 2 & 3 & 1 \\ \hline 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 2 \end{array}\right)$$

Wir vermerken, dass C die erweiterte Koeffizientenmatrix des LGS' s

$$\begin{cases} -\xi_1 + 2\xi_2 + 3\xi_3 = 1 \\ \xi_1 - \xi_2 - 3\xi_3 = 0 \\ 2\xi_2 = 4 \end{cases}$$

ist.
Die obigen elementaren Umformungen entsprechen dem Übergang zu folgenden LGS'en:

$$\begin{cases} -\xi_1 + 2\xi_2 + 3\xi_3 = 1 \\ \xi_2 = 1 \\ 2\xi_2 = 4 \end{cases} \quad \text{bzw.} \quad \begin{cases} -\xi_1 + 2\xi_2 + 3\xi_3 = 1 \\ \xi_2 = 1 \\ 0 = 2 \end{cases},$$

deren Unlösbarkeit offensichtlich ist.
Dass die Lösungsmenge eines LGS's durch elementare Umformungen erhalten bleibt, besagt der folgende Satz:

11.8 Satz (Lösungsräume bei elementaren Umfomungen)

Entsteht die Matrix $\hat{B} = \left(\begin{array}{c|c} B & \begin{matrix} \gamma_1 \\ \vdots \\ \gamma_m \end{matrix} \end{array}\right)$ aus $\hat{A} = \left(\begin{array}{c|c} A & \begin{matrix} \beta_1 \\ \vdots \\ \beta_m \end{matrix} \end{array}\right)$ durch elementare **Zeilen**umformungen, so stimmen die Lösungsmengen der Gleichungssysteme mit den erweiterten Koeffizientenmatrizen $\hat{A}$ bzw. $\hat{B}$, also von

$$f_A(x) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \quad \text{und} \quad f_B(x) = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix},$$

überein.

Beweis. Wie bemerkt entspricht einer elementaren Zeilenumformung der erweiterten Koeffizientenmatrix eine Multiplikation einer der Gleichungen mit einem Skalar bzw. die Addition einer der Gleichungen zu einer anderen oder eine daraus zusammengesetzte Operation. Damit sind die Lösungen des LGS mit erweiterter Koeffizientenmatrix $\hat{A}$, also von $f_A(x) = (\beta_1, \ldots, \beta_m)$, unter denen des LGS mit erweiterter Koeffizientenmatrix $\hat{B}$, also von $f_B(x) = (\gamma_1, \ldots, \gamma_m)$, enthalten.
Die umgekehrte Inklusion erhält man, wenn man beachtet, dass die **elementaren Zeilenumformungen durch ebensolche umgekehrt werden können,** nämlich z.B. die Multiplikation einer Zeile mit $\lambda \in K \setminus \{0\}$ durch die Multiplikation der entstandenen Zeile mit λ^{-1}, die Addition zweier Zeilen $z_i \longrightarrow \bar{z}_i = z_i + z_k$ durch $\bar{z}_i \longrightarrow z_i - \bar{z}_k = (z_i + z_k) - z_k = z_i$.

Ein weiteres Beipiel (mit $K = \mathbb{R}$) zeigt Figur 11.1.

$$\left.\begin{array}{rrrrr} \xi_1 + & \xi_2 - & \xi_3 & = & 0 \\ \xi_1 - 2\xi_2 + & & \xi_3 & = & 1 \\ \xi_1 - 2\xi_2 & & - \ \xi_4 & = & 2 \\ & & \xi_3 + 2\xi_4 & = & -1 \end{array}\right\} \longrightarrow A_{\text{erw.}} = \left(\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 \\ 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 2 & -1 \end{array}\right)$$

$$\downarrow \text{ elementare Zeilenumformung}$$

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 0 \\ 0 & -3 & 2 & 0 & 1 \\ 0 & -3 & 1 & -1 & 2 \\ 0 & 0 & 1 & 2 & -1 \end{array}\right)$$

$$\downarrow \text{ elementare Zeilenumformung}$$

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 0 \\ 0 & -3 & 2 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 2 & -1 \end{array}\right)$$

$$\downarrow \text{ elementare Zeilenumformung}$$

$$\left.\begin{array}{rrrrr} \xi_1 + \xi_2 - & \xi_3 & & = & 0 \\ - \ \xi_2 + \frac{2}{3}\xi_3 & & & = & \frac{1}{3} \\ - & \xi_3 - & \xi_4 & = & 1 \\ & & \xi_4 & = & 0 \end{array}\right\} \longleftarrow B_{\text{erw.}} = \left(\begin{array}{cccc|c} 1 & 1 & -1 & 0 & 0 \\ 0 & -1 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right)$$

Figur 11.1: Ein weiteres Beispiel zu elementaren Umformungen
mit $\overline{z_2} := z_2 - z_1,\ \overline{z_3} := z_3 - z_1, \overline{\overline{z_3}} := \overline{z_3} - \overline{z_2},\ \overline{\overline{z_4}} := \overline{\overline{z_4}} + \overline{\overline{z_3}},\ \overline{\overline{z_2}} := \frac{1}{3}\overline{\overline{z_2}}$.

Aus letzterem LGS in „Zeilen-Stufenform“ lassen sich nun leicht die Komponenten einer Lösung „von unten her“ berechnen:

$$\xi_4 = 0,\ \xi_3 = -1,\ \xi_2 = \frac{2}{3}\xi_3 - \frac{1}{3} = -1,\ \xi_1 = -\xi_2 + \xi_3 = 0;$$

die Lösungsmenge L, nach (11.8) auch des Ausgangsgleichungssystems, ist also

$$L = \{(0, -1, -1, 0)\}.$$

11.9 Bemerkung (Gaußsche Elimination)

Der im Beispiel geschilderte Prozess zum Auflösen eines LGS's (oder auch die entsprechenden Operationen mit den Gleichungen selbst) heißt **Gaußsches Eleminationsverfahren**. Bei ihm wird durch elementare Umformungen ein Gleichungssystem in Stufenform hergestellt, das sich dann leicht auflösen lässt.
Im vorhergehenden Beispiel (s. Figur 11.1) hat das betrachtete LGS eine eindeutig bestimmte Lösung. Wie geht man vor, wenn die Dimension der Lösungsmannigfaltigkeit größer als 0 ist?
Dazu erneut ein **Beispiel** (mit $K = \mathbb{R}$ und Umformungen $\overline{z_2} = z_2 - z_1, \overline{z_3} = z_3 + \overline{z_2}$), s. Figur 11.2 ! Es gilt hier: Rang $B = 2 =$Rang B_{erw}, woraus $\dim L = 4 - \text{Rang}B = 2$ folgt.

$$\left.\begin{array}{rrrrr} \xi_1 - 2\xi_2 + \xi_3 & & = & 1 \\ \xi_1 - 2\xi_2 & - \xi_4 & = & 2 \\ & \xi_3 + \xi_4 & = & -1 \end{array}\right\} \longrightarrow A_{\text{erw.}} = \left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right)$$

$$\downarrow \text{ elementare Zeilenumformung}$$

$$\left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right)$$

$$\downarrow \text{ elementare Zeilenumformung}$$

$$\left.\begin{array}{rrrrr} \xi_1 - 2\xi_2 + \xi_3 & & = & 1 \\ & \xi_3 + \xi_4 & = & -1 \end{array}\right\} \longleftarrow B_{\text{erw.}} = \left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = (B|c)$$

Figur 11.2: Zu einem Beispiel mit mehrdimensionalem Lösungsraum

Beim Beispiel der Figur 11.2 gilt Rang $B = 2 =$Rang B_{erw}, woraus $\dim L = 4 -$ Rang$B = 2$ folgt. Zur Bestimmung der Lösungen hat man nun u.a. zwei Möglichkeiten:

(a) Man löst das letztere LGS, in dem man „von unten her“ Variable „separiert“:

$$\begin{aligned} \xi_3 &= -1 - \xi_4 \\ \xi_1 &= 2\xi_2 - \xi_3 + 1 = 2\xi_2 + \xi_4 + 2. \end{aligned}$$

Die höhere Dimension von L macht sich in der „freien Verfügbarkeit über ξ_2 und ξ_4“ bemerkbar. Mit $\xi_2 = \lambda$ und $\xi_4 = \mu$ erhält man

$$\begin{aligned} L &= \{(2\lambda + \mu + 2, \lambda, -1 - \mu, \mu) \mid \lambda, \mu \in \mathbb{R}\} \\ &= (2, 0, -1, 0) + (2, 1, 0, 0)\mathbb{R} + (1, 0, -1, 1)\mathbb{R}\ . \end{aligned}$$

(b) Man wählt auf der Suche nach einer Partikulärlösung z.B.

$$(\xi_2, \xi_3) = (0, 0) \text{ und erhält } (1, 0, 0, -1) \in L \, ;$$

Zur Bestimmung von linear unabhängigen Lösungen des homogenen Systems setzt man z.B.
$(\xi_2, \xi_3) = (1, 0)$ [Damit erhält man $\rightarrow \ \xi_4 = 0, \xi_1 = 2$] bzw. $(\xi_2, \xi_3) = (0, 1)$ und gewinnt so $(2,\mathbf{1, 0}, 0)$ bzw. $(-1,\mathbf{0, 1}, -1)$ als linear unabhängige Vektoren von L_0, also

$$L = (1, 0, 0, -1) + (2, 1, 0, 0)\mathbb{R} + (1, 0, -1, 1)\mathbb{R} \, .$$

Abschließend präzisieren wir den Begriff der Zeilen-Stufenform einer Matrix und gehen auf die Möglichkeit ein, eine gegebene Matrix in eine Matrix von Stufenform überzuführen.

11.10 Satz (Zeilenstufenform durch elementare Umformungen)

Jede $(m \times n)$-Matrix A über K lässt sich durch endlich viele elementare Zeilenumformungen in eine Matrix der Form

$$B = \begin{pmatrix} \beta_{1j_1} & \dots & & & & \\ 0 & \beta_{2j_2} & \dots & & & \\ \vdots & & \ddots & & & \\ 0 & \dots & 0 & \beta_{kj_k} & \dots & \beta_{kn} \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

mit $\beta_{ij_i} \neq 0$ für alle $i \in \{1, \dots, k\}$ und $j_1 < j_2 < \dots < j_k$ überführen. Die Form von B nennen wir **Zeilenstufenform**. (Evtl. ist $\{1, \dots, k\} = \emptyset$, d.h. B Nullmatrix.)

Beweisskizze. Die Überführung der Matrix A in die Matrix B erfolgt in mehreren Schritten.
Nach q Schritten sei

$$B_q = \begin{pmatrix} \beta_{1j_1} & \dots & & & & \\ 0 & \beta_{2j_2} & \dots & & & \\ \vdots & & \ddots & & & \\ 0 & \dots & 0 & \beta_{qj_q} & \dots & \\ 0 & \dots & & 0 & * & \dots \\ \vdots & & & \vdots & \vdots & \\ 0 & \dots & & 0 & * & \dots \end{pmatrix}$$

(Im Falle $q = 0$ setzen wir $B_q := A$). Der „eingerahmte Teil“ habe o.B.d.A. in der ersten Spalte ein von 0 verschiedenes Element (andernfalls verkleinern wir ihn oder sind fertig). Eventuell durch Zeilenvertauschungen erreichen wir, dass in der entstandenen Matrix

$$\bar{B}_q = \begin{pmatrix} \boxed{\beta_{1j_1}} & \dots & & & \\ \boxed{0} & \ddots & & & \\ \vdots & & \boxed{\beta_{qj_q}} & \dots & \\ 0 & \dots & 0 & \boxed{\beta_{q+1j_{q+1}} \quad \dots} & \\ \vdots & & \vdots & \boxed{\vdots} & \\ 0 & \dots & 0 & \boxed{\beta_{mj_{q+1}} \quad \dots} & \end{pmatrix}$$

nun $\beta_{q+1\, j_{q+1}} \neq 0$ ist.
Durch Umformungen vom Typ (iv) mit $k = q+1$, $i \in \{q+2, \dots, m\}$ und $\lambda = -\big(\beta_{(q+1)j_{q+1}}\big)^{-1}\beta_{ij_{q+1}}$ gelangen wir zu

$$B_{q+1} = \begin{pmatrix} \boxed{\beta_{1j_1}} & \dots & & & & \\ \boxed{0} & \ddots & & & & \\ \vdots & & \boxed{\beta_{qj_q}} & \dots & & \\ 0 & \dots & \boxed{0} & \boxed{\beta_{q+1j_{q+1}} \quad \dots} & & \\ 0 & \dots & 0 & \boxed{*} & \dots & \\ \vdots & & \vdots & \boxed{\vdots} & & \\ 0 & \dots & 0 & \boxed{*} & \dots & \end{pmatrix}$$

mit $\beta_{q+1\, j_{q+1}} \neq 0$ und $\beta_{sj_{q+1}} = 0$ für jedes $s \geq q+2$, falls ein solches existiert.
Nach endlich vielen Schritten wird eine Matrix der Form B erreicht.

Anmerkung. Das Element $\beta_{q+1\, j_{q+1}}$ im Beweis zu 11.10 heißt **Pivot**-Element. Zwar ist theoretisch unerheblich, welches der dafür in Frage kommenden Elemente gewählt wird; für praktische Rechnungen spielt die Auswahl jedoch eine Rolle, da Rundungsfehler etc. je nach Wahl des Pivot-Elements verschieden ausfallen können. In folgendem Beispiel ist das Pivotelement jeweils eingekästelt:

Beispiel.

(i) $A_{\text{erw}} = \left(\begin{array}{cc|c} \boxed{10^{-4}} & 1 & 1 \\ 1 & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 0 & -9999 & -9998 \end{array}\right)$

(ii) $A_{\text{erw}} = \left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ \boxed{1} & 1 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 0,9999 & 0,9998 \end{array}\right)$
(Vertauschung von z_1 und z_2 und Addition von $-10^{-4}\bar{z}_1$ zu $\bar{z}_2$.)

Als „Lösung“ des LGS's mit erweiterter Koeffizientenmatrix A_{erw} erhält man bei Rechnung mit 3-stelliger Mantisse im Falle (i) aus $10^{-4}\xi_1 + \xi_2 = 1$ und $-10^4\xi_2 = -10^4$ das Ergebnis $\bar{\xi}_1 = 0$, $\xi_2 = 1$, im Falle (ii) aus $\xi_1 + \xi_2 = 2$ und $1 \cdot \xi_2 = 1$ die Werte $\xi_1 = \xi_2 = 1$.

12 Lineare Ungleichungen und Optimierung

M
↓

In diesem Paragraphen behandeln wir[44] die für Anwendungen u.a. in Wirtschaftswissenschaft und Soziologie sehr wichtigen Systeme linearer Ungleichungen und damit zusammenhängende Maximum- und Minimumprobleme.
Zur Einleitung betrachten wir folgendes **Beispiel** (s. Aigner, l.c. Seite 5): Ein Schnapsbrenner stellt zwei Schnäpse S_1, S_2 her. Neben Alkohol, Zucker und Wasser verwendet er zwei Zusätze Z_1, Z_2. Er will nun einen Produktionsplan aufstellen, der ihm einen maximalen Gewinn garantiert. Dabei hat er folgende Rahmenbedingungen zu beachten:

Anteil	S_1	S_2	Vorrat (in l)
Alkohol	0,4	0,25	30
Zucker	0,2	0,3	25
Z_1	0,15	0	10
Z_2	0	0,3	20
Wasser	0,25	0,15	100

Gewinn pro Liter:
S_1: 7
S_2: 3
(in Währungseinheiten)

Umformulierung: Sei ξ_i die zu produzierende Literzahl von S_i ($i \in \{1, 2\}$). Dann müssen folgende Ungleichungen erfüllt sein:

$$0,4\ \xi_1 + 0,25\xi_2 \leq 30 \tag{1}$$
$$0,2\ \xi_1 + 0,3\ \xi_2 \leq 25 \tag{2}$$
$$0,15\xi_1 \leq 10 \tag{3}$$
$$0,3\ \xi_2 \leq 20 \tag{4}$$
$$0,25\xi_1 + 0,15\xi_2 \leq 100 \tag{5}$$

sowie

$$\xi_1 \geq 0 \tag{6}$$

und

$$\xi_2 \geq 0. \tag{7}$$

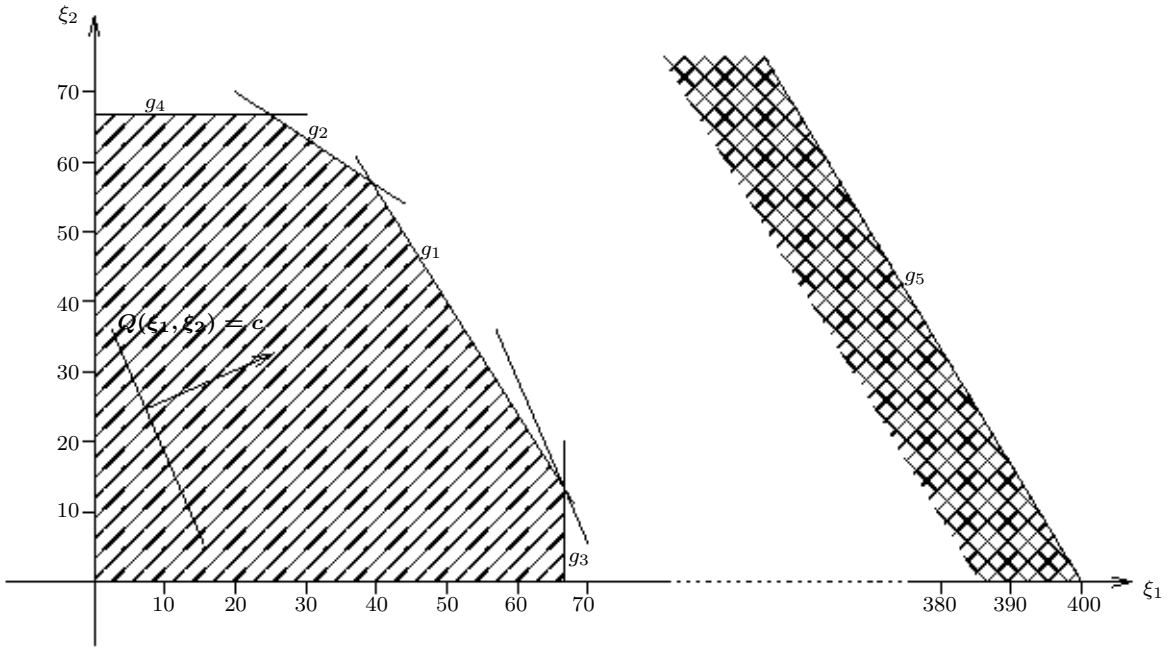
Zu maximieren ist

$$Q(\xi_1, \xi_2) = 7\,\xi_1 + 3\,\xi_2. \tag{8}$$

In diesem einfachen Fall (nur zwei Variablen) ist eine graphische Lösung möglich (s. Figur 12.1;); dabei beachtet man, dass $\alpha\xi_1 + \beta\xi_2 = \gamma$ (für $(\alpha, \beta) \neq (0, 0)$) eine Gerade in der reellen Ebene darstellt und $\alpha\xi_1 + \beta\xi_2 \leq \gamma$ eine „Halbebene".

Der schraffierte Teil in der Figur 12.1 ist der Bereich der „**zulässigen Lösungen**", also solcher Lösungen, für die die Ungleichungen (1) – (7) erfüllt sind. Dabei bezeichnet g_i die Randgerade des Gebiets, das durch Ungleichung (i) beschrieben wird; (das zu (5) gehörende Gebiet umfasst den schraffierten übrigen Teil). Jede Parallelverschiebung der Geraden mit der Gleichung $Q(\xi_1, \xi_2) = 7\xi_1 + 3\xi_2 = c_1$ in Pfeilrichtung führt zu einer Geraden der Gleichung $7\xi_1 + 3\xi_2 = c_2$ mit $c_2 \geq c_1$. Wie sich aus der Zeichnung ablesen lässt, hat unser Maximierungsproblem die eindeutige optimale Lösung $(\tilde{\xi}_1, \tilde{\xi}_2) = g_1 \cap g_3$,

[44] unter Beachtung der folgenden Quellen:
- Aigner, M.: Angewandte Mathematik, Vorlesungsskript FU 1979
- Schwarz, H. u.a.: Grundkurs Mathematik III 2, DIFF, Tübingen 1975
- Guber, S.: Lineare Algebra und Analytische Geometrie, I §15



Figur 12.1: Beispiel einer linearen Optimierung

die sich zu $\tilde{\xi}_1 = \frac{10}{0,15} = 66,\overline{6}$ und $\tilde{\xi}_2 = \frac{30-0,4\tilde{\xi}_1}{0,25} = 13,\overline{3}$ berechnet. Bei der Produktion von 66,6 l Schnaps S_1 und 13,3 l Schnaps S_2 wird der optimale Gewinn von $7 \cdot 66,6 + 3 \cdot 13,3$, d.h. rund 506 Währungseinheiten, erreicht.

Wir wollen im Folgenden zunächst lineare Ungleichungen betrachten; in diesem Zusammenhang definieren wir, was wir unter einem Halbraum verstehen wollen, und weisen eine seiner wichtigen Eigenschaften nach. ↑ M

12.1 Definition: Lineare Ungleichung; Lösungshalbraum

(a) Seien $n \in \mathbb{N}$ und $\lambda_1, \ldots, \lambda_n, \alpha \in \mathbb{R}$. Dann heißt

$$\sum_{i=1}^{n} \lambda_i \xi_i \leq \alpha \quad (\text{bzw. } \sum_{i=1}^{n} \mu_i \xi_i \geq \beta \text{ für } \mu_i = -\lambda_i,\ \beta = -\alpha) \qquad (*)$$

eine **lineare Ungleichung**; die Lösungsmenge

$$H = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n | \sum_{i=1}^{n} \lambda_i \xi_i \leq \alpha\} = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n | \sum_{i=1}^{n} (-\lambda_i) \xi_i \geq (-\alpha)\}$$

heißt **(Lösungs-) Halbraum** (von $(*)$).

(b) Verallgemeinerung:

Ist V $\mathbb{R}$-Vektorraum und $f : V \to \mathbb{R}$ linear, z.B. $f : x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \mapsto \begin{pmatrix} \lambda_1 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$

und $\alpha \in \mathbb{R}$, dann heißt

$$f(x) \leq \alpha \qquad (*')$$

eine **lineare Ungleichung** und $H = \{x \in V | f(x) \leq \alpha\}$ **Lösungshalbraum** zu $(*')$.

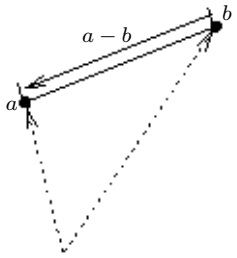
Anschaulich ist klar, wie ein Halbraum bei Dimension 2 (Halbebene) oder 3 aussieht. Wir zeigen nun allgemein, dass mit zwei Punkten eines Halbraums H auch die **Verbindungsstrecke** zu H gehört:

12.2 Definition: Verbindungsstrecke

Seien V ein $\mathbb{R}$-Vektorraum und $a, b \in V$. Dann heißt

$$\begin{aligned}[a,b] :=& \{a\lambda + b(1-\lambda) \mid \lambda \in [0,1] \subseteq \mathbb{R}\} \\ =& \{b + (a-b)\lambda \mid 0 \leq \lambda \leq 1\} \\ =& \{a\lambda + b\mu \mid \lambda, \mu \geq 0, \lambda + \mu = 1\}\end{aligned}$$

Verbindungsstrecke von a und b.

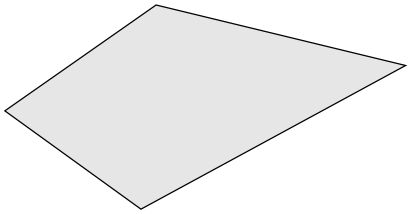


Figur 12.2:
Differenz der Ortsvektoren a und b und die Strecke $[a, b]$

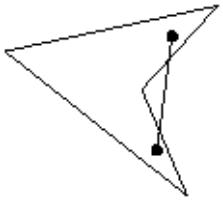
12.3 Definition: konvex

Sei V ein $\mathbb{R}$-Vektorraum und $A \subseteq V$. Dann heißt A **konvex**, falls gilt

$$\forall a, b \in A : [a, b] \subseteq A.$$



Figur 12.3: Konvexe Menge im $\mathbb{R}^2$
(z.B. Grundriss eines Museumsraums)



Figur 12.4: Nicht konvexe Menge in $\mathbb{R}^2$
(dto. mit schlechter Übersicht für den Museumswärter)

Beispiele.

(1) Jede lineare Mannigfaltigkeit in V, insbesondere jeder Lösungsraum eines linearen Gleichungssystems, ist konvex:
Seinen $L = v + U$ und $a, b \in L$, also $a = v + u_1$ und $b = v + u_2$ mit $u_1, u_2 \in U$; dann folgt $a \in b + U = L$ (bzw. $a - b \in U$) und für $x \in [a, b]$ damit $x = b + (a - b)\lambda \in b + U = L$ für ein geeignetes $\lambda \in \mathbb{R}$.

(2) Jedes Intervall von $\mathbb{R}$ ist eine konvexe Teilmenge von $\mathbb{R}$.

(3) Der Durchschnitt konvexer Mengen ist wieder konvex.

12.4 Satz (Konvexität des Lösungshalbraums)

Der Lösungshalbraum einer linearen Ungleichung ist konvex.

Beweis. Seien $f : V \to \mathbb{R}$ linear und $H = \{x \in V \mid f(x) \leq \alpha\}$ mit $\alpha \in \mathbb{R}$; seien ferner a und b Elemente von H. Dann gilt

$$f(a) \leq \alpha \quad \text{und} \quad f(b) \leq \alpha.$$

Nach (12.3) ist $[a, b] \subseteq H$ zu zeigen. Sei also $x \in [a, b]$; dann existiert ein $\lambda \in [0, 1]$ mit $x = a\lambda + b(1 - \lambda)$; es folgt

$$f(x) = f(a\lambda + b(1-\lambda)) \underset{f \text{ linear}}{=} f(a)\lambda + f(b)(1-\lambda) \underset{\substack{\lambda \geq 0 \\ 1-\lambda \geq 0}}{\leq} \alpha\lambda + \alpha(1-\lambda) = \alpha$$

und daraus $x \in H$.

Da der Durchschnitt konvexer Mengen wieder konvex ist, gilt

12.5 Korollar (Konvexität der Lösungsmenge)

Die Menge der Lösungen eines Systems linearer Ungleichungen

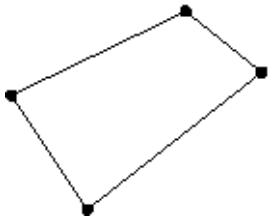
$$\begin{cases} f_1(x) \leq \alpha_1 \\ \vdots \\ f_m(x) \leq \alpha_m \end{cases}$$

(mit $f_i : V \to \mathbb{R}$ linear und $\alpha_i \in \mathbb{R}$ für $i \in \{1, \ldots, m\}$) ist konvex.

12.6 Definition: Extremalpunkt

Sei A eine konvexe Teilmenge des $\mathbb{R}$-Vektorraums V. Dann heißt $x \in A$ **Ecke** (oder **Extremalpunkt**) von A, falls x nicht im Inneren einer ganz in A enthaltenen Strecke liegt, d.h. falls gilt:

$$\forall a, b \in A : (x \in [a, b] \Rightarrow x = a \vee x = b).$$

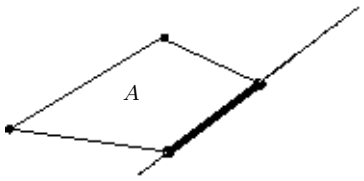


Figur 12.5: Extremalpunkte

Folgender Satz besagt: Nimmt eine lineare Abbildung $f : V \to \mathbb{R}$ auf einer konvexen Menge A ihr Maximum α an, und bezeichnet

$$A_{\text{opt}} := \{x \in A | f(x) = \alpha\}$$

die Menge der „**optimalen Punkte**“ von A, dann ist jeder Extremalpunkt von A_{opt} auch Extremalpunkt von A.



Figur 12.6: Die Menge A_{opt}

12.7 Satz (Eckpunkte der Menge optimaler Punkte)

Seien V ein $\mathbb{R}$-Vektorraum, A konvexe Teilmenge von V und $f : V \to \mathbb{R}$ linear; weiter existiere ein $x_0 \in A$ mit

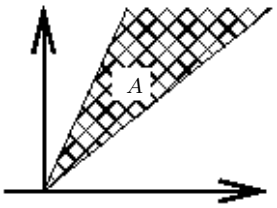
$$\forall x \in A : f(x) \leq f(x_0) =: \alpha.$$

Dann gilt:
Jeder Eckpunkt von $A_{\text{opt}} = \{x \in A | f(x) = \alpha\}$ ist ein Eckpunkt von A.

Beweis.

(i) $f^-(a) := \{x \in V | f(x) = \alpha\}$ ist für $f \neq 0$ (oder für $f = 0$ und $\alpha = f(x_0) = 0$) eine lineare Mannigfaltigkeit und als solche konvex (vgl. Bsp. 1 zu 12.3). Daher ist auch $A_{\text{opt}} = A \cap \{x \in V | f(x) = \alpha\}$ konvex (vgl. Bsp. 3 zu 12.3).

(ii) Sei y Eckpunkt von A_{opt} Wir nehmen an, y sei kein Eckpunkt von A. Es ist aber $y \in A$; verbindet man y mit einem weiteren Punkt von A, so sieht man: es existieren $a, b \in A$ und $\lambda \in (0,1)$ mit $y = a\lambda + b(1-\lambda)$). Da y Eckpunkt von $A_{\text{opt}} = \{x | f(x) = \alpha\}$ ist, gilt $f(a) \overset{=}{\neq} \alpha \overset{=}{\text{oder}} f(b) \neq \alpha$; andernfalls wäre $f(a) = f(b) = \alpha \wedge a, b \in A$, also $a, b \in A_{\text{opt}}$, damit wegen der Konvexität auch $[a, b] \subseteq A_{\text{opt}}$; es wäre dann y innerer Punkt von A_{opt} und nicht Eckpunkt. Also gilt: $f(a) < \alpha$ oder $f(b) < \alpha$. Damit ergibt sich
$f(y) = f(a\lambda + b(1-\lambda)) = f(a)\lambda + f(b)(1-\lambda) \underset{\substack{\lambda>0 \\ 1-\lambda>0}}{<} \alpha\lambda + \alpha(1-\lambda) = \alpha$, also $y \notin A_{\text{opt}}$, ein Widerspruch.

Anmerkung. Ein x_0 der im Satz 12.7 angegebenen Eigenschaften braucht nicht zu existieren.
Andeutung eines Beispiels: 

und f geeignet.

Es ist also richtig, **optimale Lösungen** (falls vorhanden) **einer linearen Optimierungsaufgabe**

$$\begin{cases} f_1(x) \leq \alpha_1 \\ \vdots \\ f_m(x) \leq \alpha_m \\ f_0(x) \text{ maximal} \end{cases} \quad (\text{mit } f_0, f_1, \ldots, f_m : V \to \mathbb{R} \text{ linear}, \alpha_1, \ldots, \alpha_m \in \mathbb{R})$$

unter den Ecken der (konvexen) Lösungsmenge des Systems linearer Ungleichungen

$$\begin{cases} f_1(x) \leq \alpha_1 \\ \vdots \\ f_m(x) \leq \alpha_m \end{cases}$$

zu suchen.
Bei zwei Variablen kann diese Suche graphisch geschehen (vgl. einleitendes Beispiel); für komplizierte Probleme gibt es Algorithmen zur rechnerischen Bestimmung, z.B. das **Simplexverfahren**, das – geometrisch interpretiert – im gezielten Durchlaufen von „Kanten“ des „Polyeders“ der Lösungen bis zum Erreichen einer „optimalen Ecke“ besteht (s. z.B. Aigner l.c. oder DIFF Studienbrief).

§ 13 Basisbezogene Darstellung von linearen Abbildungen, Matrizen

Wir setzen jetzt unsere Untersuchung von linearen Abbildungen fort.

13.1 $\mathrm{Hom}_K(V,W)$

(a) Sind V und W Vektorräume über dem Körper K, so bezeichnet man mit $\mathrm{Hom}_K(V,W)$ die Menge *aller linearen Abbildungen* von V in W.
Auf $\mathrm{Hom}_K(V,W) := \{f \mid f : V \to W \text{ linear }\}$ lassen sich durch
$f+g : V \to W$ definiert durch $(f+g)(v) = f(v)+g(v)$ und
$f\lambda : V \to W$ definiert durch $(f\lambda)(v) = f(v)\lambda$
eine Addition und eine S-Multiplikation derart definieren, dass $(\mathrm{Hom}_K(V,W), +, \cdot_K)$ ein K-Vektorraum ist. (Beweis ?)

M ↓

Um die Elemente von $\mathrm{Hom}_K(V,W)$, also lineare Abbildungen, „ökonomisch“ beschreiben zu können, untersuchen wir, wodurch eine lineare Abbildung (eindeutig) bestimmt ist. Dazu betrachten wir zunächst eine Abbildung

$f_A : K^n \to K^m$ mit $\begin{pmatrix}\xi_1\\ \dots\\ \xi_n\end{pmatrix} \mapsto \begin{pmatrix}\sum_{j=1}^n \alpha_{1j}\xi_j\\ \dots\\ \sum_{j=1}^n \alpha_{mj}\xi_j\end{pmatrix}$ und Matrix $A = \big(a_{ij}\big)_{(i,j)\in\{1,\dots,m\}\times\{1,\dots,n\}}$. Es ist f_A linear, also $f_A \in \mathrm{Hom}_K(K^n,K^m)$. Da jedes Element von K^n sich als Linearkombination von $e_1,\dots,e_n$ mit $e_i = (\delta_{ij})_{j\in\{1,\dots,n\}}$ schreiben lässt und f_A linear ist, reicht es, sich auf $f_A(e_i)$ zu beschränken; es zeigt sich

$$f_A(e_i) = \begin{pmatrix}\sum_{j=1}^n \alpha_{1j}\delta_{ij}\\ \dots\\ \sum_{j=1}^n \alpha_{mj}\delta_{ij}\end{pmatrix} = \begin{pmatrix}\alpha_{1i}\\ \dots\\ \alpha_{mi}\end{pmatrix} = a_{\bullet i} \quad .$$

Für $i \in \{1,\dots,n\}$ ist das Bild von e_i unter f_A (als Spalte geschrieben) gerade der i-te Spaltenvektor von A:

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1i-1} & \boxed{\alpha_{1i}} & \alpha_{1i+1} & \dots & \alpha_{1n}\\ \vdots & & & \vdots & & & \vdots\\ \alpha_{m1} & \dots & \alpha_{mi-1} & \boxed{\alpha_{mi}} & \alpha_{mi+1} & \dots & \alpha_{mn}\end{pmatrix}$$

↑ M

Dabei war $A = \big(\alpha_{ij}\big)_{(i,j)\in\{1,\dots,m\}\times\{1,\dots,n\}}$ beliebig in $K^{(m,n)}$ gewählt.

Wiederholung:
Mit $I := \{1,\dots,m\}$ und $J := \{1,\dots,n\}$ definieren wir (wie in §11) für $m,n \in \mathbb{N}$ die Menge der **$(m \times n)$-Matrizen** über dem Körper K:

$$K^{(m,n)} := K^{I\times J} := \big\{(\alpha_{ij})_{(i,j)\in I\times J} | \alpha_{ij} \in K \text{ für alle } i \in \{1,\dots,m\}, j \in \{1,\dots,n\}\big\} \quad .$$

M ↓

Wir haben gesehen, dass die Abbildungen f_A mit Matrizen A lineare Abbildungen sind. Gehören umgekehrt im Endlich-Dimensionalen zu linearen Abbildungen auch Matrizen?

↑ M

Vor Beantwortung dieser Frage halten wir fest: Durch die Auswahl einer geeigneten Matrix $A \in K^{(m,n)}$ kann man, wie aus obigen Überlegungen folgt, eine lineare Abbildung f angeben, die die Vektoren $e_1, \ldots, e_n$ der kanonischen Basis von K^n auf vorgeschriebene Vektoren $(\alpha_{11}, \ldots, \alpha_{m1}), \ldots, (\alpha_{1n}, \ldots, \alpha_{mn})$ von K^m abbildet. Dies gilt auch allgemein für Vektorräume. Darüberhinaus ist eine lineare Abbildung schon durch ihre Wirkung auf die Elemente einer Basis des Urbildraumes eindeutig bestimmt:

13.2 Satz von der Linearen Fortsetzung (Basissatz für lineare Abbildungen)

Seien V und W K-Vektorräume, $(b_i)_{i\in I}$ eine geordnete Basis von V und $(w_i)_{i\in I}$ eine Familie von Vektoren aus W (mit gleicher Indexmenge I). Dann gilt: Es existiert genau ein $f \in \mathrm{Hom}_K(V, W)$ mit $f(b_i) = w_i$ für alle $i \in I$.

Beweis. Übungsaufgabe A13.1.

M ↓

Für die angestrebte Beschreibung einer linearen Abbildung betrachten wir zunächst die Bilder einer fest gewählten Basis von V. Wir beschränken uns dabei auf endlich-dimensionale Vektorräume:

↑ M

13.3 Beschreibung von linearen Abbildungen bzgl. fester Basen

Voraussetzungen: Seien V, W K-Vektorräume mit $\dim V < \infty$ und $\dim W < \infty$. Seien $B = (b_1, \ldots, b_n)$ (geordnete) Basis von V und $C = (c_1, \ldots, c_m)$ (geordnete) Basis von W.

(a) **Definition**: Ist $f \in \mathrm{Hom}_K(V, W)$, so heißt die durch

$$f(b_j) = \sum_{i=1}^{m} c_i \alpha_{ij} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}_C \qquad (j = 1, \ldots, n)$$

definierte Matrix

$$M(f) := M_C^B(f) := (a_{ij})_{\substack{i\in\{1,\ldots,m\} \\ j\in\{1,\ldots,n\}}}$$

die Matrix von f bzgl. des Basispaares (B, C).

Dabei ist die **j−te Spalte** von $M_C^B(f)$, d.h. $\begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$, gleich $M_C(f(b_j))$, also **gleich dem Koordinatenvektor des Bildes des j−ten Basisvektors des Urbildraums bzgl. der Basis des Bildraums** (für $j = 1, \ldots, n$).

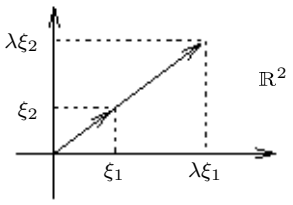
(b) **Beispiele**

(i) $K = \mathbb{R}$, $V = W = \mathbb{R}^2$

$s_\lambda : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2 \\ (\xi_1, \xi_2) \mapsto (\lambda\xi_1, \lambda\xi_2) \end{cases}$

(**Streckung** um den Faktor λ vom Nullpunkt aus).

Wir wählen $B = C = (e_1, e_2)$. Es folgt:



Figur 13.1:
Zur zentrischen Streckung

$$s_\lambda(e_1) = s_\lambda\big((1,0)\big) = (\lambda, 0) = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}_{(e_1,e_2)}$$

$$s_\lambda(e_2) = s_\lambda\big((0,1)\big) = (0, \lambda) = \begin{pmatrix} 0 \\ \lambda \end{pmatrix}_{(e_1,e_2)} \quad \text{und damit} \quad M_C^B(s_\lambda) = \begin{pmatrix} \boxed{\lambda} & \boxed{0} \\ \boxed{0} & \boxed{\lambda} \end{pmatrix}.$$

Speziell: $M_C^B(\mathrm{id}_{\mathbb{R}^2}) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ („Einheitsmatrix“).

(ii) Seien K, V, W, s_λ wie in Beispiel (b)(i) gewählt.

$B' = \big((1,1),(0,1)\big)$, $C' = (e_2, e_1)$

$$s_\lambda\big((1,1)\big) = (\lambda,\lambda) = \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}_{C'}, \qquad s_\lambda\big((0,1)\big) = (0,\lambda) = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}_{C'}$$

$$M_{C'}^{B'}(s_\lambda) = \begin{pmatrix} \lambda & \lambda \\ \lambda & 0 \end{pmatrix}.$$

(**Beachten Sie die Abhängigkeit der darstellenden Matrix von den gewählten Basen!**)

(iii) Sei f die „**Drehung** um 0 um den Winkel vom Maß α “ in der reellen euklidischen Ebene, also in $V = W = \mathbb{R}^2$ versehen mit dem kanonischen Skalarprodukt (vgl. §1 !).(Siehe Figur 13.2 !)

Sei $f(x) = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}$ Bildpunkt des Vektors $x = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$. (Wir indentifizieren wieder die Punkte mit ihren Ortsvektoren bzgl. des Ursprungs 0). Sei ferner y_1 der Fußpunkt des Lots von $f(x)$ auf $0\vec{x}$. Wegen $|x| = |f(x)|$ ergibt sich

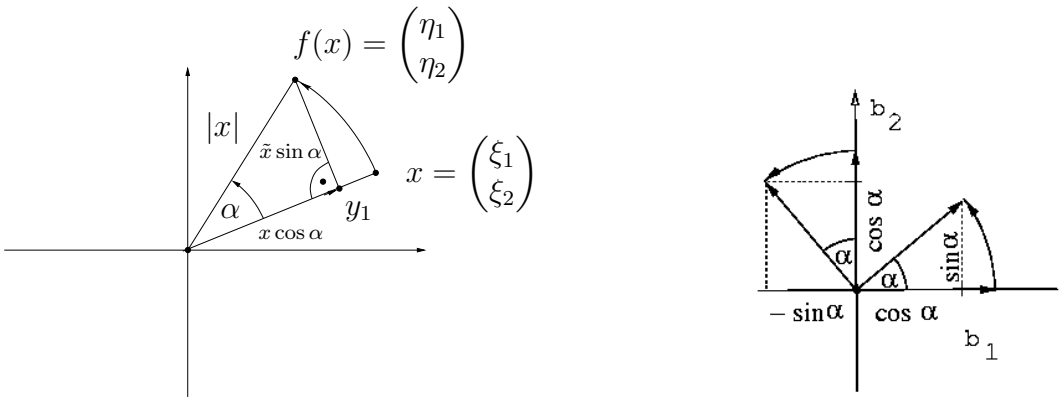
$y_1 = \big(|x|\cos\alpha\big)\frac{x}{|x|} = x\cos\alpha$ und $y_1\overrightarrow{f(x)} = \big(|x|\sin\alpha\big)\frac{\tilde{x}}{|x|} = \tilde{x}\sin\alpha$, (mit $\tilde{x} := \begin{pmatrix} -\xi_2 \\ \xi_1 \end{pmatrix} \perp x$, also $|\tilde{x}| = |x|$), folglich $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = f(x) = x\cos\alpha + \tilde{x}\sin\alpha = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\cos\alpha + \begin{pmatrix} -\xi_2 \\ \xi_1 \end{pmatrix}\sin\alpha = \begin{pmatrix} \xi_1\cos\alpha - \xi_2\sin\alpha \\ \xi_2\cos\alpha + \xi_1\sin\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$

insbesondere (mit f linear und $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ bzw. $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $B = (b_1, b_2) = (e_1, e_2)$):

$$f(e_1) = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}_B, \quad f(e_2) = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}_B.$$

Ergebnis: In der reellen euklidischen Ebene ist die Drehung um 0 mit Drehwinkelmaß α linear, und es gilt (bzgl. der kanonischen Basis B):

$$M_B^B(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad .$$



a) Bild $f(x)$ von x b) Koordinaten der Bilder der Basisvektoren

Figur 13.2: Zur Drehung um 0.

(iv) Ist $V = K^n$, $W = K^m$ und sind B und C die kanonischen Basen, so gilt

$$M_C^B(f_A) = A.$$

(v) Eine **Scherung** S von $AG(\mathbb{R}^2)$ ist eine Abbildung von $\mathbb{R}^2$ auf sich, bei der

(i) eine Gerade g punktweise fest bleibt (*Fixpunktgerade*) – sie heißt *Affinitätsachse* –,

(ii) der Bildpunkt $S(Q)$ jeden Punktes Q auf der Parallelen zur Affinitätsachse g durch den Urbildpunkt Q liegt und

(iii) Kollinearität erhalten bleibt.

Wir wählen nun den Ursprung 0 auf g; sei $B = (b_1, b_2)$ eine Basis derart, dass b_1 Einheitsvektor auf der Achse g und b_2 dazu orthogonaler Einheitsvektor ist (s. Figur 13.3)

Heuristik: Ist P ein fester Punkt mit zweiter Koordinate 1 und Bildpunkt $P' = S(P)$, so sei $\alpha \in \mathbb{R}$ bestimmt durch $\vec{PP'} = b_1\alpha$. Z.B. nach dem Strahlensatz folgt dann für einen Punkt $Q \in 0P$ mit 2. Koordinate η die Gleichung

$$S(Q) = S(\begin{pmatrix} \xi \\ \eta \end{pmatrix}_B) = \begin{pmatrix} \xi \\ \eta \end{pmatrix}_B + \begin{pmatrix} 1 \\ 0 \end{pmatrix}_B \cdot \eta\alpha = \begin{pmatrix} \xi + \eta\alpha \\ \eta \end{pmatrix}_B .$$

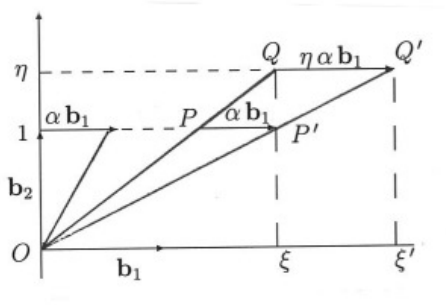
Die Abbildung S bildet daher Q ab auf den Punkt mit Koordinaten

$$\begin{pmatrix} \xi' \\ \eta' \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \xi \\ \eta \end{pmatrix} .$$

Die lineare Abbildung mit der Abbildungsvorschrift (bzgl. Basis B)

$$\begin{pmatrix}\xi\\ \eta\end{pmatrix} \mapsto \begin{pmatrix}1 & \alpha\\ 0 & 1\end{pmatrix} \cdot \begin{pmatrix}\xi\\ \eta\end{pmatrix}, \quad \text{also mit Matrix} \quad M_B^B(S) = \begin{pmatrix}1 & \alpha\\ 0 & 1\end{pmatrix}$$

und $\alpha \in \mathbb{R}$ erhält die Kollinearität und Parallelität, lässt die $\xi-$Achse punktweise und die jeweilige $\eta-$Koordinate fest. Damit erfüllt sie die Bedingungen (i) bis (iii) für die Scherung und ist somit die durch P und P' eindeutig bestimmte Scherung.



Figur 13.3 Zur Scherung

(c) **Existenz und Eindeutigkeit der darstellenden Matrix**

$M_C^B(f)$ ist bei festen endlichen geordneten Basen B und C für jedes $f \in \mathrm{Hom}_K(V,W)$ definiert und eindeutig bestimmt.

Folgerung aus Satz 13.2.

(d) **Umkehrung**

Ist umgekehrt $A = (\alpha_{ij})$ eine $(m \times n)$-Matrix über K, so existiert (für B und C fest) genau eine Abbildung $f \in \mathrm{Hom}_K(V,W)$ mit $M_C^B(f) = A$.

Beweis.

Ist $A = (\alpha_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ gegeben, so definieren wir $w_j := \sum_{i=1}^{m} c_i\alpha_{ij}$ $(j = 1,\dots,n)$. Genau dann ist $M_C^B(f) = A$, wenn $f(b_j) = w_j$ ist. Nach (13.2) existiert ein eindeutig bestimmtes $f \in \mathrm{Hom}_K(V,W)$ mit $f(b_j) = w_j$ für alle $j \in \{1,\dots,n\}$. Es folgt $M_C^B(f) = A$. Umgekehrt folgt aus $M_C^B(f) = A$ die Gleichung $f(b_j) = \sum_{i=1}^{m} c_i\alpha_{ij} = w_j$. Nach dem Fortsetzungssatz ist f eindeutig.

(e) **Zusammenfassung**

Die Abbildung $M_C^B : \begin{cases}\mathrm{Hom}_K(V,W) \to K^{(m,n)}\\ f \mapsto M_C^B(f)\end{cases}$ ist wohldefiniert und bijektiv.

13.4 Erinnerung und Anmerkung zu Koordinaten

In (7.8) hatten wir jedem Vektor x eines endlich-dimensionalen K-Vektorraums V bzgl. einer festen geordneten Basis $B = (b_1, \dots, b_n)$ einen „**Koordinatenvektor**“ zugeordnet; dieser lässt sich auch als $n \times 1$-Matrix über K auffassen:

Zu $x = \sum_{i=1}^{n} b_i \xi_i = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B \in V$ hatten wir definiert: $M_B(x) := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{(n,1)}$.

Die Zuordnung $M_B \colon \begin{cases} V \to K^{(n,1)} \\ x \mapsto M_B(x) \end{cases}$ ist ein Vektorraum-Isomorphismus (vgl. (7.8)).

Die Elemente von $K^n, K^{(1,n)}, K^{(n,1)}$ unterscheiden sich zwar formal:

$(\xi_1, \dots, \xi_n) \in K^n$ ist eine Abbildung von $(1, \dots, n)$ in K, $(\xi_1, \dots, \xi_n) \in K^{(1,n)}$ eine Abbildung von $\{(1,1), \dots, (1,n)\}$ in K, und $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{(n,1)}$ ist eine Abbildung von $\{(1,1), \dots, (n,1)\}$ in K. Jedoch gibt es eine natürliche Zuordnung zwischen ihnen, die mit der Addition und der S-Multiplikation verträglich ist:

$$K^{(1,n)} \cong K^n \cong K^{(n,1)}.$$

Die Unterscheidung ist also mehr formaler als inhaltlicher Art. Sie muss aber trotzdem beachtet werden.

Für den $K-$Vektorraum $V = K^{(n,1)}$ und die kanonische Basis

$$B = \Big(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \Big) \quad \text{gilt:} \quad M_B\Big(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \Big) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Im Fall eines beliebigen n-dimensionalen K-Vektorraums V müssen wir jedoch zwischen $x \in V$ und dem Spaltenvektor $M_B(x) \in K^{(n,1)}$ unterscheiden. Wir verwenden dabei wie bisher folgende Schreibweise: $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_B = \sum_{j_1}^{n} b_j \xi_j$.

Im Hinblick auf den Übergang von Zeilen- zu Spaltenvektor definieren wir noch

$$(\xi_1, \dots, \xi_n)^T := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

als Spezialfall der Definition $(\alpha_{ij})^T_{i\in I, j\in J} := (\alpha_{ji})_{i\in I, j\in J}$.
A^T heißt die zu A **transponierte Matrix**, a^T der zum Vektor a transponierte Vektor.
Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{pmatrix}.$$

Wie wirkt nun ein Homomorphismus, dessen Matrix bzgl. Basen B und C gegeben ist, auf die entsprechenden Koordinatenvektoren?

Seien $B = (b_1, \ldots, b_n)$ und $C = (c_1, \ldots, c_m)$ geordnete Basen der endlich-dimensionalen $K-$Vekrorräume V bzw. W; sei ferner $f \in \mathrm{Hom}_K(V, W)$ mit Matrix $M_C^B(f) = (\alpha_{ij})_{\substack{i\in\{1,\ldots,m\} \\ j\in\{1,\ldots,n\}}}$. Für $x \in V$ und $y = f(x)$ mit $M_B(x) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$, also $x = \sum_{j=1}^{n} b_j\xi_j$, sowie $M_C(y) = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}$, also $y = \sum_{i=1}^{m} c_i\eta_i$, folgt aus $y = f(x) = f\big(\sum_{j=1}^{n} b_j\xi_j\big) = \sum_{j=1}^{n} f(b_j)\xi_j = \sum_{j=1}^{n}\big(\sum_{i=1}^{m} c_i\alpha_{ij}\big)\xi_j = \sum_{i=1}^{m}\sum_{j=1}^{n} c_i\alpha_{ij}\xi_j = \sum_{i=1}^{m} c_i\big(\sum_{j=1}^{n} \alpha_{ij}\xi_j\big)$ die Gleichung

$$M_C(y) = \begin{pmatrix} \sum_{j=1}^{n} \alpha_{1j}\xi_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{mj}\xi_j \end{pmatrix}$$

und damit

$$\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} = M_C\big(f(x)\big) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = M_C^B(f) \cdot M_B(x)\ .$$

Wir halten fest:

13.5 Satz (Abbildungsgleichung in Matrixform)

Seien V und W endlich-dimensionale K-Vektorräume mit geordneter Basis B bzw. C und sei $f \in \mathrm{Hom}_K(V, W)$. Dann gilt für alle $x \in V$ die Gleichung

$$M_C(f(x)) = M_C^B(f) \cdot M_B(x).$$

Anmerkungen: 1.) Der Übergang zu Koordinaten sieht dabei schematisch wie folgt aus:

$$\begin{array}{ccccccccc} f: & V & \longrightarrow & W & \text{linear} & x & \mapsto & y & \\ \downarrow & \downarrow M_B & & \downarrow M_C & & \downarrow & & \downarrow & \\ M_C^B(f) & K^{(n,1)} & \longrightarrow & K^{(m,1)} & \text{mit} & M_B(x) & \mapsto & M_C(y) & = M_C^B(f) \cdot M_B(x) \end{array}$$

2.) *Alternative Schreibweise:* Bezeichnen wir mit
- $\vec{x}$ den Koordinatenvektor[45] von $x \in V$ zur Basis B,
- $\vec{y}$ den Koordinatenvektor [46] von $y \in W$ zur Basis C,
- M_f die Matrix von f bzg. der Basen B und C,

dann gilt analog zur "Abbildungsgleichung" $y = f(x)$ die Gleichung

$$\boxed{\vec{y} = M_f \cdot \vec{x}}$$

Beispiel: (Vgl. 13.3 Bsp. (b)(iii) :)

Seien $V = W = \mathbb{R}^2$, $K = \mathbb{R}$, $B = C = (e_1, e_2)$ sowie d_α die Drehung um 0 um den Winkel vom Maß α. Dann haben wir folgende Entsprechungen:

$$\begin{array}{ccccc} y & = & d_\alpha & & (x) \\ \updownarrow & & \updownarrow & & \updownarrow \\ \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} & = & \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} & \cdot & \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \end{array}$$

wegen

$$\begin{array}{rcll} \eta_1 & = & (\cos\alpha)\,\xi_1 & +(-\sin\alpha)\,\xi_2 \quad \text{und} \\ \eta_2 & = & (\sin\alpha)\,\xi_1 & +(\cos\alpha)\,\xi_2 \end{array} .$$

Anmerkung. Die Koordinatenvektoren haben wir hier in Spaltenform geschrieben. Bei einer modifizierten Definition der Matrix von d_α mit dem Ziel der Multiplikation mit der Matrix "von rechts" können auch Zeilenvektoren benötigt werden:

$$(\eta_1, \ldots, \eta_m) = (\xi_1, \ldots, \xi_n) \cdot M'(\text{ mit } M' = M^T).$$

[45] in Spaltenform
[46] ebenfalls in Spaltenform

13.6 Spezialfall: Linearform

Sei V ein K-Vektorraum der Dimension n mit geordneter Basis B, sei ferner $f \in V^* := \mathrm{Hom}_K(V,K)$ (f heißt dann **Linearform** und V^* **Dualraum**) und $M_B(x) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ Koordinatenvektor von $x \in V$ bzgl. B; dann existiert eine Matrix $\begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} \in K^{(1,n)}$ mit

$$f(x) = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \quad \text{für alle } x \in V \ .$$

Beweis. Die Behauptung folgt aus 13.5, wenn wir die Elemente von K mit denen von $K^{(1,1)}$ identifizieren; dann ist $M_{\{1\}}\big(f(x)\big) = \big(f(x)\big) = f(x)$.

Anmerkung:

1.) Wir werden später (als Spezialfall von (13.10)) sehen, dass die Abbildung

$$\begin{cases} V^* \to K^{(1,n)} \\ f \mapsto M^B_{\{1\}}(f) = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} \end{cases}$$

im Falle endlicher Dimension n von V ein Vektorraum-Isomorphismus ist. Insbesondere gilt also im Endlich-Dimensionalen : $V^* \cong V$.

2.) Die Aussage $V^* \cong V$ gilt nicht für unendlich-dimensionale Vektorräume.

3.) Ein Beispiel einer Linearform eines unendlich-dimensionalen Vektorraums[47] ist die Abbildung $f : \mathcal{C}[0,1] \to \mathbb{R}$ mit $g \mapsto \int_0^1 g(t)\,dt$.

13.7 Anwendung: Lineare Gleichung

Die lineare Gleichung (also das lineare Gleichungssystem mit einer Gleichung)

$$\alpha_1\xi_1 + \ldots + \alpha_n\xi_n = \beta$$

über K lässt sich mit Hilfe der einreihigen Koeffizientenmatrix $a = (\alpha_i)_{\{i\in\{1,\dots,n\}}$ und den Vektoren $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{(n,1)}$ bzw. $b = \big(\beta\big) \in K^{(1,1)}$ folgendermaßen schreiben

[47] des Vektorraums der reellen stetigen Funktionen auf dem Intervall $[0,1]$

(vgl. §11):

$$\begin{pmatrix}\alpha_1 & \dots & \alpha_n\end{pmatrix} \cdot \begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} = (\beta) \quad \text{bzw.} \quad \boxed{a \cdot x = b \quad \text{bzw.} \quad f_a(x) = b}$$

Dabei ist f_a Linearform, also ein Element von V^*. Ist $a \neq 0$, so gilt für den Lösungsraum L von $f_a(x) = b$, dass $\dim L = \dim \operatorname{Kern} f_a = n - \operatorname{Rang} a = n - 1$ ist. **Der Lösungsraum einer linearen Gleichung ist daher Urbild eines Skalars unter einer Linearform** und als solcher (im Falle $\boldsymbol{a \neq 0}$) **eine Hyperebene in $\mathbf{AG}(\boldsymbol{K^n})$.**

Wir hatten zu Anfang dieses Paragraphen erwähnt, dass $\operatorname{Hom}_K(V, W)$ ein Vektorraum (mit der üblichen Addition und S-Multiplikation von Abbildungen) ist.
Auch $K^{(m,n)}$ $(= K^{I'}$ mit $I' = \{1, \dots, m\} \times \{1, \dots, n\})$ trägt eine Vektorraum-Struktur (s. §6 Bsp. b). Wir wiederholen die Definition der Addition und S-Multiplikation, spezialisiert auf Matrizen, und wenden uns danach der Frage zu, ob M_C^B linear ist.

13.8 Definition (Addition und S-Multiplikation bei Matrizen)

Spezialisierung von §6 Bsp. b)
Seien $(\alpha_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}}$ und $(\beta_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}}$ $m \times n$-Matrizen über dem Körper K und $\lambda \in K$.

(a) Addition „+“ (komponentenweise)

$$(\alpha_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}} + (\beta_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}} := (\alpha_{ij} + \beta_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}},$$

also

$$\begin{pmatrix}\alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn}\end{pmatrix} + \begin{pmatrix}\beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn}\end{pmatrix} := \begin{pmatrix}\alpha_{11} + \beta_{11} & \dots & \alpha_{1n} + \beta_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} + \beta_{m1} & \dots & \alpha_{mn} + \beta_{mn}\end{pmatrix}$$

(b) S-Multiplikation „$\cdot_K$“ (komponentenweise)

$$\lambda \cdot (\alpha_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}} := (\lambda\alpha_{ij})_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}} =: (\alpha_{ij}) \cdot \lambda \ ,$$

also

$$\lambda \cdot \begin{pmatrix}\alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn}\end{pmatrix} = \begin{pmatrix}\lambda\alpha_{11} & \dots & \lambda\alpha_{1n} \\ \vdots & & \vdots \\ \lambda\alpha_{m1} & \dots & \lambda\alpha_{mn}\end{pmatrix} = \begin{pmatrix}\alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn}\end{pmatrix} \cdot \lambda \ .$$

13.9 Satz ($K^{(m,n)}$ als Vektorraum)

(i) $(K^{(m,n)}, +, \underset{K}{\cdot})$ ist ein K-Vektorraum.

(ii) Eine Basis von $K^{(m,n)}$ über K wird gebildet von den Matrizen

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 0 \end{pmatrix}, \quad \dots \quad , \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \end{pmatrix}, \quad \dots \quad , \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Beweis. Übung ! Hinweis: Es gilt $K^{(m,n)} \cong K^{m \cdot n}$.

Anmerkung.

(a) Die Summe zweier Matrizen über K ist nur definiert, wenn sie **vom gleichen Typ** sind, also gleiche Zeilenzahl und gleiche Spaltenzahl haben.

(b) Neutrales Element der Addition ist die „Nullmatrix“ $\begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$.

13.10 Satz (Matrizenzuordnung als Isomorphismus)

Seien V ein n-dimensionaler K-Vektorraum und W ein m-dimensionaler K-Vektorraum ($n, m \in \mathbb{N}$) mit (geordneter) Basis B bzw. C. Dann gilt

(a) $M_C^B(f+g) = M_C^B(f) + M_C^B(g)$ und $M_C^B(f \cdot \lambda) = M_C^B(f) \cdot \lambda$
für alle $f, g \in \mathrm{Hom}_K(V,W)$ und $\lambda \in K$, d.h. M_C^B ist **linear** und damit ein Vektorraumisomorphismus.

(b) $\boxed{\mathrm{Hom}_K(V,W) \cong K^{(m,n)}}$ (als Vektorräume).

Beweis. (a) Seien $f, g \in \mathrm{Hom}_K(V,W)$, $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$, ferner seien $f(b_j) = \sum_{i=1}^m c_i \alpha_{ij}$ und $g(b_j) = \sum_{i=1}^m c_i \beta_{ij}$ (für $j \in \{1, \dots, n\}$), sowie $I := \{1, \dots, m\}$ und $J := \{1, \dots, n\}$.

Bezüglich der Basen B und C gilt dann mit $M := M_C^B$ definitionsgemäß

$$M(f) = (\alpha_{ij})_{\substack{i\in I\\ j\in J}} \text{ und } M(g) = (\beta_{ij})_{\substack{i\in I\\ j\in J}}.$$

Aus

$$(f+g)(b_j) = f(b_j) + g(b_j) = \sum_{i=1}^{m} c_i\alpha_{ij} + \sum_{i=1}^{m} c_i\beta_{ij} = \sum_{i=1}^{m} c_i(\alpha_{ij}+\beta_{ij})$$

und

$$(f\lambda)(b_j) = f(b_j)\lambda = \big(\sum_{i=1}^{m} c_i\alpha_{ij}\big)\lambda = \sum_{i=1}^{m} c_i(\alpha_{ij}\lambda)$$

ergibt sich

$$M(f+g) = \big(\alpha_{ij}+\beta_{ij}\big)_{\substack{i\in I\\ j\in J}} = \big(\alpha_{ij}\big)_{\substack{i\in I\\ j\in J}} + \big(\beta_{ij}\big)_{\substack{i\in I\\ j\in J}} = M(f) + M(g)$$

und

$$M(f\lambda) = \big(\alpha_{ij}\lambda\big)_{\substack{i\in I\\ j\in J}} = \big(\alpha_{ij}\big)_{\substack{i\in I\\ j\in J}} \cdot \lambda = M(f)\cdot\lambda.$$

(b) Aus (a) und (13.3)(e) folgt, dass M_C^B ein Isomorphismus zwischen den Vektorräumen $(\mathrm{Hom}_K(V,W), +, \cdot_K)$ und $(K^{(m,n)}, +, \cdot_K)$ ist.

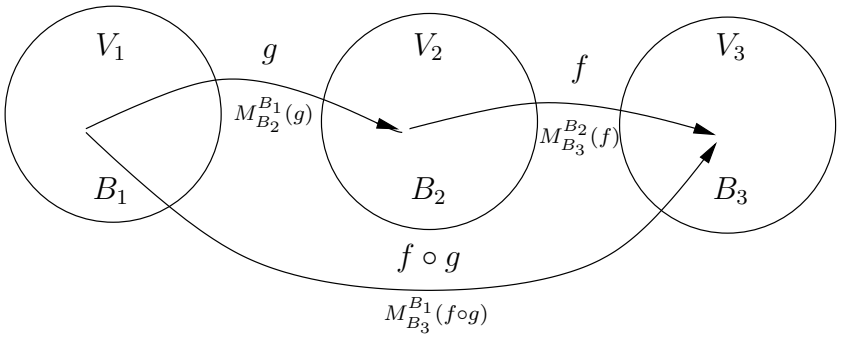
M ↓

Der Isomorphismus M_C^B hängt, wie schon mehrfach betont, von den Basen B und C ab. Nach deren Festlegung kann man dann von den linearen Abbildungen zu den entsprechenden Matrizen übergehen (oder umgekehrt, vgl. 13.3(d)).

Eine weitere Verknüpfung von Homomorphismen müssen wir allerdings noch berücksichtigen: die Hintereinanderausführung (Verkettung).

↑ M

Seien V_1, V_2 und V_3 K-Vektorräume mit (geordneter) Basis $B_1 = (b_1,\dots,b_n)$, $B_2 = (c_1,\dots,c_m)$ bzw. $B_3 = (d_1,\dots,d_r)$, ferner $g \in \mathrm{Hom}_K(V_1,V_2)$ und $f \in \mathrm{Hom}_K(V_2,V_3)$ sowie $M_{B_3}^{B_2}(f) = (\alpha_{ij})_{i=1,\dots,r,j=1,\dots,m}$ und $M_{B_2}^{B_1}(g) = (\beta_{jk})_{j=1,\dots,m,k=1,\dots,n}$. Es gilt also $f(c_j) = \sum_{}^{r} d_i\alpha_{ij}$ und $g(b_k) = \sum_{}^{m} c_j\beta_{jk}$. Wir wollen γ_{ik} mit $M_{B_3}^{B_1}(f\circ g) = (\gamma_{ik})_{i=1,\dots,r,k=1,\dots,n}$ bestimmen. (Vgl. Figur 13.4 !)



Figur 13.4: Venndiagramm zur Verkettung von linearen Abbildungen

$$\text{Aus}\quad (f\circ g)(b_k) \underset{\substack{\text{Def.}\\ f\circ g}}{=} f(g(b_k)) \underset{\substack{\text{Def.}\\ M(g)}}{=} f\Big(\sum_{j=1}^{m} c_j\beta_{jk}\Big)$$
$$\underset{\substack{f\\ \text{linear}}}{=} \sum_{j=1}^{m} f(c_j)\beta_{jk} \underset{\substack{\text{Def.}\\ M(f)}}{=} \sum_{j=1}^{m}\Big(\sum_{i=1}^{r} d_i\alpha_{ij}\Big)\beta_{jk} = \sum_{i=1}^{r}\sum_{j=1}^{m} d_i\alpha_{ij}\beta_{jk}$$
$$= \sum_{i=1}^{r}\Big(d_i \underbrace{\sum_{j=1}^{m}\alpha_{ij}\beta_{jk}}_{\gamma_{ik}}\Big) \qquad (k=1,\dots,n)$$

ergibt sich $\gamma_{ik} = \sum\limits_{j=1}^{m} \alpha_{ij}\beta_{jk}$ (für $i \in \{1,\dots,r\}$, $k \in \{1,\dots,n\}$).
Demgemäß definieren wir das Produkt zweier Matrizen:

13.11 Definition (Matrizenmultiplikation)

Sind $(\alpha_{ij})_{\substack{i\in\{1,\dots,r\}\\ j\in\{1,\dots,m\}}} \in K^{(r,\underline{m})}$ und $(\beta_{jk})_{\substack{j\in\{1,\dots,m\}\\ k\in\{1,\dots,n\}}} \in K^{(\underline{m},n)}$, so definieren wir:

$$(\alpha_{ij})_{\substack{i\in\{1,\dots,r\}\\ j\in\{1,\dots,m\}}} \cdot (\beta_{jk})_{\substack{j\in\{1,\dots,m\}\\ k\in\{1,\dots,n\}}} := \Big(\sum_{j=1}^{m}\alpha_{ij}\beta_{jk}\Big)_{\substack{i\in\{1,\dots,r\}\\ k\in\{1,\dots,n\}}} \in K^{(r,n)}$$

Schematisch:

$$\begin{array}{c} \\ \\ i\text{-te}\\ \text{Zeile} \end{array} \left(\begin{array}{ccc} \alpha_{11} & \dots & \alpha_{1m}\\ \vdots & & \vdots\\ \hline \alpha_{i1} & \dots & \alpha_{im}\\ \hline \vdots & & \vdots\\ \alpha_{r1} & \dots & \alpha_{rm} \end{array}\right) \left(\begin{array}{c|c|c} \beta_{11}\dots & \beta_{1k} & \dots\ \beta_{1n}\\ \vdots & & \vdots\\ \beta_{m1}\dots & \beta_{mk} & \dots\ \beta_{mn} \end{array}\right) = \left(\begin{array}{c|c|c} & \downarrow & \\ \vdots & & \vdots\\ \hline \rightarrow & \gamma_{ik} & \\ \hline \vdots & & \vdots \end{array}\right)$$
$$\text{mit } \gamma_{ik} = \sum_{j=1}^{m}\alpha_{ij}\beta_{jk}$$

$$k\text{-te Spalte}$$

Insbesondere: $\begin{pmatrix} \alpha_1 & \dots & \alpha_m \end{pmatrix} \begin{pmatrix} \beta_1\\ \vdots\\ \beta_m \end{pmatrix} = \sum\limits_{i=1}^{m}\alpha_i\beta_i$.

Anmerkung. 1.) Die Matrizenmultiplikation ist nach dieser Definition nur ausführbar, wenn die **Spaltenzahl des ersten Faktors gleich der Zeilenzahl des zweiten** ist. – Entsprechend müssen ja bei der Komposition linearer Abbildungen der Wertebereich der zuerst angewandten Abbildung mit dem Definitionsbereich der zweiten übereinstimmen.

2.) Besteht (β_{ik}) nur aus einem Spaltenvektor (d.h. $k = 1$), so geht die Definition 13.11 über in die Definition der Multiplikation einer Matrix mit einem Spaltenvektor aus §10.

Beispiel:

Für $K = \mathbb{Q}$, $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{Q}^{(2,3)}$ und $B = \begin{pmatrix} 2 & -3 & 3 & 0 \\ 3 & -1 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{(3,4)}$ gilt:

$$A \cdot B = \begin{pmatrix} 1\cdot 2+2\cdot 3+3\cdot 0 & 1\cdot(-3)+2(-1)+3\cdot 1 & 1\cdot 3+2\cdot 2+3\cdot 0 & 1\cdot 0+2\cdot 0+3\cdot 1 \\ 4\cdot 2+5\cdot 3+6\cdot 0 & 4(-3)+5(-1)+6\cdot 1 & 4\cdot 3+5\cdot 2+6\cdot 0 & 4\cdot 0+5\cdot 0+6\cdot 1 \end{pmatrix}$$
$$= \begin{pmatrix} 8 & -2 & 7 & 3 \\ 23 & -11 & 22 & 6 \end{pmatrix} \in \mathbb{Q}^{(2,4)}.$$

Gemäß der Überlegung, die zu Definition 13.11 führten, können wir festhalten:

13.12 Satz (Matrix eines Produkts von linearen Abbildungen)

Sind V_i $(i = 1, 2, 3)$ endlich-dimensionale K-Vektorräume mit (geordneten) Basen B_i, und sind $g \in \mathrm{Hom}_K(V_1, V_2)$, $f \in \mathrm{Hom}_K(V_2, V_3)$, dann gilt:

$$\boxed{M_{B_3}^{B_1}(f \circ g) = M_{B_3}^{B_2}(f) \cdot M_{B_2}^{B_1}(g)}.$$

Der Komposition der Abbildungen entspricht die Multiplikation der darstellenden Matrizen.

Beispiel. Seien $K = \mathbb{R}$, $V_i = \mathbb{R}^2$, $B_i = (e_1, e_2) =: B$ (für $i = 1, 2, 3$), ferner
$f = D_\alpha$ die "Drehung um $\mathfrak{o}$ um den Winkel vom Maß α" und
g die "Spiegelung an der $x-$Achse" (s. Figur 13.5 !)
Nach Beispiel (b)(iii) aus 13.3 ist f linear und es gilt:

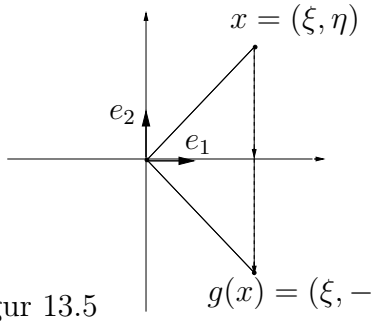
$$M_B^B(D_\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Für g erhalten wir folgende Beschreibung:

$$g((\xi_1, \xi_2)) = (\xi_1, -\xi_2) = \begin{pmatrix} \xi_1 \\ -\xi_2 \end{pmatrix}_B.$$

Also ist auch g linear und es gilt

$$M_B^B(g) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$



Figur 13.5
Zur Spiegelung an der $x-$Achse

Für die Verknüpfung $D_\alpha \circ g$ folgt:

$$M_B^B(D_\alpha \circ g) = M_B^B(D_\alpha) \cdot M_B^B(g) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix},$$

für $g \circ D_\alpha$ hingegen:

$$M_B^B(g \circ D_\alpha) = M_B^B(g) \cdot M_B^B(D_\alpha) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & -\cos\alpha \end{pmatrix}.$$

Dass D_α und g i.A. nicht vertauschbar sind, findet eine Entsprechung darin, dass $M(D_\alpha)$ und $M(g)$ i.A. nicht kommutieren, nämlich für $\alpha \notin \pi\mathbb{Z}$.

Die Matrizen–Multiplikation ist also i.A. nicht kommutativ.

Als Eigenschaft der Verknüpfung quadratischer Matrizen (also von $(m \times n)-$Matrizen mit $m = n$) erwähnen wir:

13.13 Satz (Ring der $n \times n$ Matrizen)

(a) Für jeden Körper K und jedes $n \in \mathbb{N}$ ist $(K^{(n,n)}, +, \cdot)$ ein Ring, der **Ring der quadratischen $n-$reihigen Matrizen über K.**.

(b) Ist V ein K-Vektorraum der Dimension n mit $n \in \mathbb{N}$, so gilt:

$$(\mathrm{End}_K(V), +, \circ) := (\mathrm{Hom}_K(V,V), +.\circ) \cong (K^{(n,n)}, +, \cdot)$$

Beweis-Andeutung. $\mathrm{End}_K V$ ist ein Ring, wie man durch Nachprüfen der Ring-Axiome feststellt. Für eine Basis B ist die Abbildung

$$M_B^B : \left\{ \begin{array}{ccc} \mathrm{Hom}_K(V,V) & \to & K^{(n,n)} \\ f & \mapsto & M_B^B(f) \end{array} \right.$$

ein Ring-Isomorphismus.

13.14 Anmerkung (K-Algebra)

$K^{(n,n)}$ und $\mathrm{End}_K V$ tragen sowohl Vektorraum- als auch Ring-Struktur. Allgemein heißt $(V, +, \underset{K}{\cdot}, \cdot)$ eine **K–Algebra**, wenn $(V, +, \underset{K}{\cdot})$ ein $K-$Vektorraum und $(V, +, \cdot)$ ein Ring ist und folgende Verträglichkeitsbedingung gilt:

$$\forall a, b \in V \quad \forall \lambda \in K : \quad (a \cdot b) \underset{K}{\cdot} \lambda = (a \underset{K}{\cdot} \lambda) \cdot b = a \cdot (b \underset{K}{\cdot} \lambda) \quad .$$

Nach den Sätzen 13.9, 13.10 und 13.13 sowie durch Nachprüfen der erwähnten Verträglichkeitsbedingung erhält man:

$(\ \mathrm{End}_K(V), +, \underset{K}{\cdot}, \circ)$ und $(K^{(n,n)}, +, \underset{K}{\cdot}, \cdot)$ sind $K-$Algebren und als solche isomorph.

Die Abbildung M_B^B ist dabei ein **bijektiver $K-$Algebren-Homomorphismus**, d.h. sowohl Ring- als auch Vektorraum-Homomorphismus, also vertäglich mit den Operationen "+","$\underset{K}{\cdot}$" und "$\circ$" sowie "$\cdot$".

Weitere Beispiele von Algebren:

$\mathbb{Q}, \mathbb{R}, \mathbb{C}$ mit den üblichen Verknüpfungen sind $\mathbb{Q}-$Algebren.

$\mathbb{R}, \mathbb{C}$ mit den üblichen Verknüpfungen sind $\mathbb{R}-$Algebren.

Ebenso bilden die Polynomabbildungen von K eine $K-$Algebra (bzgl. $+, \underset{K}{\cdot}, \cdot$).

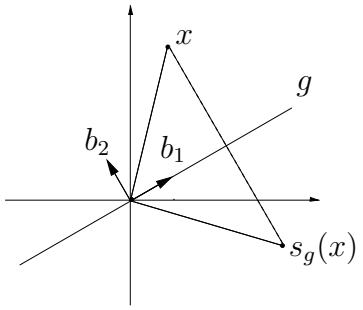
Auch die Menge $\mathcal{C}[0,1]$ der auf dem Intervall $[0,1]$ stetigen reellen Funktionen ist eine $\mathbb{R}-$ Algebra (bzgl. $+, \underset{\mathbb{R}}{\cdot}, \cdot$).

Übungsaufgaben

A 13.1 Führen Sie den Beweis von Satz 13.2 aus!

A 13.2 In der reellen euklidischen Ebene sei s_g die Geradenspiegelung an einer Nullpunktsgeraden g (s. Figur 13.6 !). Zeigen Sie, dass s_g linear ist.

Lösungshilfe: Wählen Sie eine Basis $B = (b_1, b_2)$ wie in Figur 13.6 angedeutet! x habe die Darstellung $x = b_1\xi + b_2\eta$; welche Linearkombination beschreibt dann $s_g(x)$?



Figur 13.6: Zur Spiegelung an einer Nullpunktsgeraden

§ 14 Affine Abbildungen

M
↓

Lineare Abbildungen lassen den Nullvektor fest. Wenn man daher z.B. bei geometrischen Abbildungen den Nullpunkt nicht frei wählen kann oder wenn, wie bei nicht-trivialen Translationen, gar kein Fixpunkt existiert, ist ein Verallgemeinerung der Begriffsbildung nötig.

↑
M

14.1 Definition: affin-lineare Abbildung, Affinität

a) Sei V ein K-Vektorraum. Dann heißt $F : V \to V$ *affin-lineare Abbildung* von V (bzw. von $AG(V)$), auch *affine Abbildung*, falls es eine *lineare* Abbildung $f : V \to V$ und ein $t \in V$ gibt mit

$$F(x) = f(x) + t.$$

b) Eine affin-lineare Abbildung F heißt *Affinität*, falls F zusätzlich bijektiv ist.

Anmerkungen :

(1) Eine affin-lineare Abbildung in $\mathrm{AG}(V)$ ist also eine Translation verknüpft mit einer linearen Abbildung. Oft ist es möglich, durch geeignete Wahl des Ursprungs (als einen Fixpunkt der Abbildung) $t = 0$ zu erreichen und damit zu einer linearen Abbildung zu gelangen.

(2) Ist $\dim_K V = n < \infty$, so hat F bzgl. einer Basis B von V die Darstellung $\vec{x} \mapsto A \cdot \vec{x} + \vec{t}$ mit Matrix $A := M_B^B(f)$ und Koordinatenvektoren $\vec{x} := M_B(x)$ und $\vec{t} := M_B(t)$ von x bzw. t. Bei einer Affinität ist A regulär, d.h. Rang $A = n$, und umgekehrt.

Im Folgenden behandeln wir neben Affinitäten von $\mathrm{AG}(\mathbb{R}^1)$ unf $\mathrm{AG}(\mathbb{R}^2)$ auch **spezielle Typen von Affinitäten** in $\mathrm{AG}(\mathbb{R}^n)$: Ähnlichkeitsabbildungen und Kongruenzabbildungen (Bewegungen).

14.2 Eigenschaften affin-linearer Abbildungen

Eine affin-lineare Abbildung bildet affine Unterräume auf affine Unterräume ab und erhält Inzidenz und Parallelität. Eine Affinität ist damit insbesondere eine **Kollineation** von $\mathrm{AG}(V)$, d.h. eine Bijektion der Punktmenge von $\mathrm{AG}(V)$ auf sich, die die Menge der Geraden von $AG(V)$ auf sich abbildet.

Beweis. Es gilt $F(a + U) = f(a + U) + t = f(a) + f(U) + t = (f(a) + t) + f(U)$. Da f linear ist, wird der Unterraum U auf einen Unterraum $f(U)$ abgebildet. Sind L und M parallele Unterräume, so folgt $U_L \; I \; U_M$ und damit $f(U_M) \; I \; f(U_L)$, woraus sich die Parallelität von $F(L)$ und $F(M)$ ergibt.

Anmerkungen :

1.) Umgekehrt werden die Kollineationen von $AG(K^n)$ für $K = \mathbb{R}$ ausschließlich von Affinitäten induziert; (dies ergibt sich aus dem sogenannten 2. Hauptsatz der Projektiven Geometrie, da $\mathbb{R}$ keine nicht-trivialen Automorphismen zulässt.)

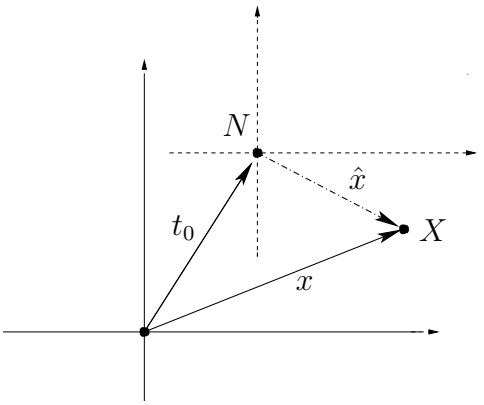
2.) Im Gegensatz zum reellen Fall liefert für $K = \mathbb{C}$ zum Beispiel die Abbildung $x = (\xi_1, \ldots, \xi_n) \mapsto \overline{x} = (\overline{\xi_1}, \ldots, \overline{\xi_n})$ (mit $\overline{\xi} = \alpha - \beta\, i$ für $\xi = \alpha + \beta\, i$) (Übergang zu konjugiert komplexen Koordinaten) eine Kollineation, die keine Affinität ist.

14.3 Beispiele

(i) Im Fall von $V = K$ (also $\dim_K(V) = 1$) ist jede Affinität von der Form $K \to K$ mit $x \mapsto ax + b$ für $a \neq 0$, $a, b \in K$; umgekehrt ist jede Abbildung mit dieser Zuordnung eine Affinität (Beweis ?). Die Menge aller dieser Abbildungen bildet bzgl. der Hintereinanderausführung eine Gruppe, die *affine Gruppe* $\mathrm{AGL}(1, K)$.

(ii) Wir betrachten nun **Drehungen**, **Geradenspiegelungen Scherungen** von $AG(\mathbb{R}^2)$, bei denen nicht der Nullpunkt fest bleibt. Diese sind dann keine linearen Abbildungen.

Ist t_0 der Ortsvektor eines Fixpunkts N einer sochen Abbildung F, so verschieben wir das Koordinatensystem um den Vektor t_0. Sei x der Ortsvektor eines Punktes X im urspünglichen Koordinatensystem und $\hat{x}$ derjenige im neuen Koordinatensystem. Es gilt dann (vgl. Figur 14.1 !): $\boxed{x = \hat{x} + t_0.}$



Figur 14.1 Zur Koordinatentransformation

Im neuen Koordinatensystem hat F die Darstellung wie in §13; insbesondere gilt $\hat{y} = F(\hat{x}) = f(\hat{x})$ (mit linearer Abbildung f). Damit folgt
$y - t_0 = \hat{y} = F(\hat{x}) = f(\hat{x}) = f(x - t_0) = f(x) - f(t_0)$, also

$$y = f(x) + t \quad \text{mit} \quad t = t_0 - f(t_0).$$

Jede Drehung, Spiegelung und Scherung in $\mathbf{AG(\mathbb{R}^2)}$ist also eine Affinität; und die zugehörende lineare Abbildung ist die einer Drehung, Spiegelung oder Scherung mit Fixpunkt 0.

14.4 Abstand, Orthogonalität (Erinnerung)

Im Folgenden sei $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt versehen (vgl. §1 für $n = 2$ bzw. $n = 3$), also

$$(\xi_1, \ldots, \xi_n) \cdot (\eta_1, \ldots, \eta_n) := \big(\xi_1, \ldots, \xi_n\big) \cdot \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} = \sum_{i=1}^{n} \xi_i \eta_i.$$

Durch dieses Skalarprodukt ist
1.) ein Abstand d zwischen je zwei Punkten definiert (**euklidischer Abstand**) mittels

$$d(x, y) = d((\xi_1, \ldots, \xi_n), (\eta_1, \ldots, \eta_n)) := \sqrt{\sum_{i=1}^{n} (\xi_i - \eta_i)^2} = \sqrt{(x-y)^2} =: ||x - y||$$

2.) eine Orthogonalitätsrelation $\perp$ erklärt durch: $x \perp y \Leftrightarrow x \cdot y = 0$.
$\mathrm{AG}(\mathbb{R}^n)$ mit diesem Skalarprodukt, dieser Abstandsfunktion und dieser Orthogonalitätsrelation heißt $n-$dim **reeller euklidischer Raum**, in Zeichen $\mathrm{EG}(\mathbb{R}^n)$.

14.5 Ähnlichkeitsabbildungen

a) *Definition*:
Sei F eine affin-lineare Abbildung von $\mathrm{EG}(\mathbb{R}^n)$. Dann heißt F *Ähnlichkeitsabbildung* oder *äquiforme Abbildung*, falls ein $\gamma \in \mathbb{R}$ mit $\gamma > 0$ existiert derart, dass gilt:

$$d\,(F(x), F(y)) = \gamma \cdot d\,(x, y) \qquad \text{für alle } x, y \in \mathbb{R}^n.$$

Diese Formel bedeutet, dass Streckenlängen-Verhältnisse konstant bleiben.

b) *Beispiel* (vgl. 13.3 b (i)):
Die **zentrische Streckung** $S_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ mit $x \mapsto \alpha \cdot x$ (für $\alpha \neq 0$) ist eine Ähnlichkeitsabbildung. [48]

Es gilt nämlich:

$$d(S_\alpha(x), S_\alpha(y)) = d(\alpha x, \alpha y) = \sqrt{(\alpha x - \alpha y)^2} = \sqrt{\alpha^2 (x-y)^2} = |\alpha| d(x, y).$$

Ferner gilt $(S_\alpha)^{-1} = S_{\alpha^{-1}}$ (für $\alpha \neq 0$).

c) *Weitere Eigenschaften*:
Ist F wie in 14.5 a) definiert, so ist $\tilde{F} := S_\gamma^{-1} \circ F$ wegen

$$d\,(\tilde{F}(x), \tilde{F}(y)) = |\gamma^{-1}| d\,(F(x), F(y)) = |\gamma^{-1}\gamma| d\,(x, y) = d(x, y)$$

[48] Die zugehörige Matrix bzgl. der kanonischen Basis ist $A = \alpha \cdot E_n$.

eine längenerhaltende Affinität und damit Kongruenzabbildung (Bewegung)(s.u. 14.6). Jede Ähnlichkeitsabbildung ist damit *Produkt einer Bewegung und einer zentrischen Streckung*: $F = S_\gamma \circ \tilde{F}$.

Nach geeigneter Definition von Winkelgrößen kann man zeigen, dass auch Winkelgrößen bei Ähnlichkeitsabbildungen invariant sind.

d) *Anmerkung:*
Ähnliche Figuren, d.h. solche, die durch Ähnlichkeitsabbildungen ineinander übergeführt werden können, stimmen daher in der Größe entsprechender Winkel und dem Verhältnis entsprechender Streckenlängen überein.

14.6 Bewegungen (Kongruenzabbildungen)

a) *Definition*:
Unter einer *Bewegung* (*Kongruenzabbildung*) des reellen euklidischen Raumes $\mathrm{EG}(\mathbb{R}^n)$ versteht man eine Affinität von $\mathrm{EG}(\mathbb{R}^n)$, die den Abstand je zweier Punkte invariant lässt, also **längentreu** ist.

b) *Beispiele*:
Translationen, Spiegelungen an einer Geraden (in $\mathrm{EG}(\mathbb{R}^2)$) bzw. an einer Ebene (in $\mathrm{EG}(\mathbb{R}^3)$), Drehungen.

c) *Eigenschaften:*
Man kann zeigen, dass im 2- (bzw. 3-) dimensionalen Fall eine Bewegung

(i) Produkt einer Translation mit einer Drehung ist (Bezeichnung: *eigentliche Bewegung*, *gleichsinnige Kongruenzabbildung*), und es ist $\det f = 1$; (im 3-Dimensionalen: Schraubung oder Translation)

oder

(ii) Produkt einer Translation mit einer Drehung verknüpft mit einer Achsenspiegelung (*uneigentliche Bewegung*), $\det f = -1$; (im 3-Dimensionalen: Drehspiegelung, Punktspiegelung, Gleitspiegelung).

d) *Anmerkungen*:

(i) **Kongruente Figuren** sind Teilmengen von $\mathbb{R}^n$, die durch eine Kongruenzabbildung aufeinander abgebildet werden können. (Vgl. die Vorlesung Elementargeometrie).

(ii) U.a. die folgenden Mengen von Abbildungen des reellen euklidischen Raumes $\mathcal{R} = \mathrm{EG}(\mathbb{R}^n)$ bilden bzgl. Verkettung eine **Gruppe**:

$\mathcal{A}$: die Menge der Affinitäten von $\mathcal{R}$
$\ddot{\mathcal{A}}$: die Menge der Ähnlichkeitsabbildungen von $\mathcal{R}$
$\mathcal{K}$: die Menge der Kongruenzabbildungen (Bewegungen) von $\mathcal{R}$
$\mathcal{K}^+$: die Menge der gleichsinnigen Kongruenzabbildungen von $\mathcal{R}$
$\mathcal{T}$: die Menge der Translationen von $\mathcal{R}$
$\mathcal{T}_g$: die Menge der Translationen von $\mathcal{R}$ längs einer (festen) Geraden g.

Dabei gilt: $$\mathcal{T}_g \leq \mathcal{T} \leq \mathcal{K}^+ \leq \mathcal{K} \leq \ddot{\mathcal{A}} \leq \mathcal{A}.$$

§ 15 Determinanten

M
↓

Motivation

Die Definition der „Determinante“ einer Matrix bzw. einer linearen Abbildung wird motiviert durch folgende Ziele:

1. **Bestimmung des Volumens** des durch die Vektoren $v_1, v_2, v_3 \in \mathbb{R}^3$ aufgespannten „Parallelflachs“ $\{\sum \alpha_i v_i \mid \alpha_i \in [0,1]\} \subseteq \mathbb{R}^3$ bzw. *Bestimmung der Fläche* des durch Vektoren $v_1, v_2 \in \mathbb{R}^2$ aufgespannten Parallelogramms und damit verbunden:
2. **Test der linearen Unabhängigkeit** von n Vektoren eines n-dimensionalen $K-$Vektorraums: bei linearer Abhängigkeit ist das Volumen bzw. das Flächenmaß gleich 0.
3. Test der **Regularität** einer quadratischen Matrix $A \in K^{(n,n)}$, d.h. eine Möglichkeit zum Feststellen, ob $\text{Rang} A = n$ gilt und damit A eine multiplikative Inverse besitzt.
4. Maß für die **Änderung des Volumens** von Körpern bei linearen Abbildungen.

↑
M

(A) Steilkurs

15.1 Definition: Volumen (Determinantenform)

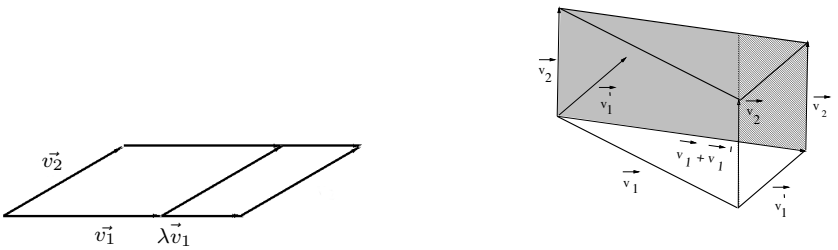
Sei V ein $n-$dimensionaler K-Vektorraum. Eine Abbildung $\Delta : V^n \to K$ heißt *Volumen (Determinantenform)*, falls gilt:

(i) Δ ist n-fache Linearform, d.h. $\Delta : V^n \to K$ ist linear in jeder Komponente.

(ii) $\Delta(v_1, \ldots, v_n) = 0$ für beliebige linear abhängige Vektoren $v_1, \ldots, v_n$ aus V.

(iii) $\Delta(b_1, \ldots, b_n) \neq 0$ für mindestens eine Basis $B_1 = (b_1, \ldots, b_n)$ von V.

Anmerkung:
Die Forderung (i) ergibt sich ebenso wie (ii) und (iii) u. a. aus dem Ziel der Bestimmung eines (gerichteten) Volumens, (s. Figur 15.1 für $n = 2$).



a) b)

Figur 15.1 Eigenschaften eines Volumens im Fall $n = 2$
a) $\Delta(\lambda v_1, v_2) = \lambda\Delta(v_1, v_2)$
b) Additivität (Flächenumwandlungen durch 2 Scherungen !)

15.2 Spezialfall (n=2):

Seien $n = 2$ und $\Delta : V \times V \to K$ Volumen sowie $C = (c_1, c_2)$ Basis von V ! Für $v_i = c_1\xi_{i1} + c_2\xi_{i2} \quad (i = 1, 2)$ folgt aus der Multilinearität von Δ wegen $\Delta(c_i, c_i) = 0$ und wegen $\Delta(c_2, c_1) = -\Delta(c_1, c_2)$ (s.u., 15.3) die Darstellung:
$\Delta(c_1\xi_{11} + c_2\xi_{12},\, c_1\xi_{21} + c_2\xi_{22})$

$$= \xi_{11}\xi_{21}\Delta(c_1, c_1) + \xi_{11}\xi_{22}\Delta(c_1, c_2) + \xi_{12}\xi_{21}\Delta(c_2, c_1) + \xi_{12}\xi_{22}\Delta(c_2, c_2)$$
$$= (\xi_{11}\xi_{22} - \xi_{12}\xi_{21}) \cdot \Delta(c_1, c_2) = \begin{vmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{vmatrix} \cdot \Delta(c_1, c_2) \quad \text{mit}$$

$$\begin{vmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{vmatrix} := \xi_{11}\xi_{22} - \xi_{12}\xi_{21} =: \det \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix}.$$

15.3 Hilfssatz und Definition (alternierende Multilinearform)

(a) Jedes Volumen Δ ist eine **alternierende Multilinearform** , d.h. dass die Multilinearform Δ die folgende Eigenschaft hat:

(ii') $\Delta(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\Delta(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$
(für alle $v_1, \ldots, v_n \in V$).

(b) Umgekehrt ist im Falle [49] char $K \neq 2$ eine Abbildung Δ mit (i), (ii') und (iii) ein Volumen.

Beweisskizze.

(a) $\Delta(\ldots, v_i + v_j, \ldots, v_i + v_j, \ldots) = 0 \Longrightarrow$
$$0 + 0 + \Delta(\ldots, v_i, \ldots, v_j, \ldots) + \Delta(\ldots, v_j, \ldots, v_i, \ldots) = 0.$$

(b) $\Delta(x_1, \ldots x_i, \ldots, x_i, \ldots, x_n) = -\Delta(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) \Longrightarrow$
$$2\Delta(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) = 0 \Longrightarrow \Delta(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) = 0.$$
Ist z.B. x_j linear abhängig von den übrigen x_i, so folgt:
$$\Delta(x_1, \ldots, \sum_{i \neq j}^{n} x_i\lambda_i, \ldots, \ldots, x_n) = \sum_{i \neq j}^{n} \Delta(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n)\lambda_i = 0. \qquad \square$$

Angeregt durch den Fall $n = 2$ definieren wir:

15.4 Definition: Determinante

(i) Als **Determinante (normiertes Volumen)** bezeichnet man ein Volumen Δ_0 von $K^{(n,1)}$, das für die Basis $B = (e_1, \ldots, e_n)$ (mit den kanonischen Spaltenvektoren e_i) den Wert 1 hat, also für das gilt:

$$\Delta_0(e_1, \ldots, e_n) = 1.$$

[49] d.h. im Falle eines Körpers K mit $1 + 1 \neq 0$

Anmerkung: (1) Weiter unten werden wir zeigen, dass ein solches normiertes Volumen existiert und eindeutig bestimmt ist.
(2) Durch die Determinante ist aus der Menge aller Determinantenformen gerade die des orientierten Inhalts ausgewählt: c_1 und c_2 (bzw. c_1, c_2 und c_3) spannen ja gerade das Einheitsquadrat (den Einheitswürfel) auf; und 15.3 zeigt die Abhängigkeit von der Reihenfolge der Vektoren c_i.

(ii) Ist nun $A = (\alpha_{ij})_{i,j=1,\dots,n} \in K^{(n,n)}$ eine quadratische Matrix mit Spalten $a_{\bullet 1} := \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{n1} \end{pmatrix}, \dots, a_{\bullet n} := \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{nn} \end{pmatrix}$, so definiert man $\det A = \Delta_0 (a_{\bullet 1}, \dots, a_{\bullet n})$.

A 15.1 *Übungsaufgabe:*
Zeigen Sie die Existenz und Eindeutigkeit der Determinante im Fall $n = 2$!

Unmittelbar aus der Definition des Volumens erhält man eine Richtung von (i), ferner (ii) und (iii) (für die Spalten) des folgenden Satzes; die Aussagen über die Zeilen folgt aus (iv); die restlichen Aussagen werden weiter unten, (v) erst in Teil II der Vorlesung gezeigt.

15.5 Eigenschaften von Determinanten

Eigenschaften von Determinanten

(i) Sei $A \in K^{(n,n)}$. Es gilt dann $\det A \neq 0$ genau dann, wenn Rang $A = n$ gilt, also A regulär ist.

(ii) $\det : A \mapsto \det A$ ist linear in jeder Spalte von A.
(Man kann zeigen, dass $\det A$ auch linear in jeder Zeile von A ist.)

(iii) Verhalten bei elementaren Umformungen:
- Die Determinante bleibt unverändert bei Addition einer Linearkombination von Spalten (Zeilen) zu einer anderen Spalte (bzw. Zeile).
- $\det B = \alpha \cdot \det A$, falls B aus A durch Multiplikation einer Spalte (Zeile) mit $\alpha \in K$ hervorgeht.
- $\det B = -\det A$, falls B aus A durch Vertauschen zweier Spalten (Zeilen) hervorgeht.

(iv) $\det A = \det A^T$ für $(\alpha_{ij})^T_{i,j=1,\dots,n} := (\alpha_{ji})_{i,j=1,\dots,n}$.
(Determinante der „**transponierten**“ Matrix).

(v) $\det(A \cdot B) = \det A \cdot \det B$ für $A, B \in K^{(n,n)}$ **(Multiplikationssatz)**.

15.6 Berechnung von $3 \times 3-$Determinanten

(i) Spezialfall: **Dreiecksmatrix**[50]

$$\begin{vmatrix} \alpha_{11} & & \star \\ & \alpha_{22} & \\ \mathrm{O} & & \alpha_{33} \end{vmatrix} := \det \begin{pmatrix} \alpha_{11} & & \star \\ & \alpha_{22} & \\ \mathrm{O} & & \alpha_{33} \end{pmatrix} = \prod_{i=1}^{3} \alpha_{ii}.$$

Beweis:
Wegen der definierenden Eigenschaften von Determinanten folgt:

$\det(e_1\alpha_{11}, e_1\alpha_{12} + e_2\alpha_{22}, e_1\alpha_{13} + e_2\alpha_{23} + e_3\alpha_{33})$
$= \alpha_{11}\Delta_0(e_1, e_1\alpha_{12} + e_2\alpha_{22}, e_1\alpha_{13} + e_2\alpha_{23} + e_3\alpha_{33}) =$
$\alpha_{11}\alpha_{12}\Delta_0(e_1, e_1, e_1\alpha_{13} + e_2\alpha_{23} + e_3\alpha_{33}) + \alpha_{11}\alpha_{22}\Delta_0(e_1, e_2, e_1\alpha_{13} + e_2\alpha_{23} + e_3\alpha_{33}) =$
$0 + \alpha_{11}\alpha_{22}[\alpha_{13}\Delta_0(e_1, e_2, e_1) + \alpha_{23}\Delta_0(e_1, e_2, e_2) + \alpha_{33}\Delta_0(e_1, e_2, e_3)] = \alpha_{11}\alpha_{22}\alpha_{33}.$ □

Anmerkungen: 1.) Ein alternativer Beweis benutzt elementare Umformungen (s.u.) zu einer Diagonalmatrix.
2.) Eine Verallgemeinerung auf $n \times n-$Dreiecksmatrizen ist möglich.

(ii) *Regel von Sarrus* (gesprochen Sarrü) (nur für 3 × 3-Matrizen !) :

$$\overbrace{\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}}^{+} \begin{matrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{matrix}$$
$$\underbrace{\qquad\qquad\qquad}_{-}$$

$= \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{31}\alpha_{22}\alpha_{13} - \alpha_{32}\alpha_{23}\alpha_{11} - \alpha_{33}\alpha_{21}\alpha_{12}.$

Beweisandeutung: Spezialisierung von 15.7, s. auch (iv) !
Beispiel:

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \end{vmatrix} \begin{matrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{matrix} = 1\cdot 2\cdot 3^2 + 1\cdot 2^2\cdot 1 + 1\cdot 1\cdot 3 - 1\cdot 2\cdot 1 - 3\cdot 2^2\cdot 1 - 3^2\cdot 1\cdot 1 = 2$$

(iii) Durch **elementare Umformungen** und Zurückführung auf (i) :

Beispiel:

$$\begin{vmatrix} \lambda & k & k \\ k & \lambda & k \\ k & k & \lambda \end{vmatrix} = \begin{vmatrix} \lambda - k & 0 & k \\ k - \lambda & \lambda - k & k \\ 0 & k - \lambda & \lambda \end{vmatrix} = \begin{vmatrix} \lambda - k & 0 & k \\ 0 & \lambda - k & 2k \\ 0 & 0 & \lambda + 2k \end{vmatrix} \underset{(i)}{=}$$

1. Spalte minus 2. Spalte	2. Zeile plus 1. Zeile
2. Spalte minus 3. Spalte	3. Zeile plus neue 2. Zeile

[50] Dabei sind die Einträge des oberen Dreiecks beliebige Körperelemente, durch $\star$ angedeutet, die Einträge des unteren Dreiecks sämtlich 0, durch O angedeutet.

$$= (\lambda - k)^2(\lambda + 2k).$$

(iv) **Laplace'sche Entwicklung (im Fall n=3)** nach der ersten Zeile (Zurückführung auf $2 \times 2-$Matrizen) :

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12} \begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13} \begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix}$$

Beispiel

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 2^2 \\ 1 & 3 & 3^2 \end{vmatrix} = 1 \cdot \begin{vmatrix} 2 & 2^2 \\ 3 & 3^2 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 2^2 \\ 1 & 3^2 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = (2 \cdot 3^2 - 2^2 \cdot 3) - (1 \cdot 3^2 - 1 \cdot 2^2) + (1 \cdot 3 - 1 \cdot 2) = 2$$

Anmerkung: 15.6 (iv) ist eine Spezialisierung von 15.7, s.u..
Wir erwähnen hier schon die Verallgemeinerung:

15.7 Laplace'sche Entwicklung (nach einer Zeile):

Allgemein gilt für $A = (\alpha_{ij})_{i,j=1,\ldots,n}$ mit $A_{kj} := (-1)^{k+j}$ $\begin{vmatrix} \alpha_{11} \cdots & \alpha_{1j} \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{k1} & \alpha_{kj} & \alpha_{kn} \\ \vdots & & \vdots \\ \alpha_{n1} \cdots & \alpha_{nj} \cdots & \alpha_{nn} \end{vmatrix}$

die Formel

$$\boxed{\det A = \sum_{j=1}^{n} \alpha_{kj} A_{kj}} \quad (\text{für } k \in \{1, \ldots, n\} \text{ fest}).$$

Beispiel: siehe 15.6 (iv) !
Diese Formel werden wir in Teil II der Vorlesung beweisen. Wegen 15.5 (iv) ist auch die Entwicklung nach einer Spalte möglich:

$$\det A = \sum_{i=1}^{n} \alpha_{ik} A_{ik} \quad (\text{für } k \in \{1, \ldots, n\} \text{ fest}).$$

Im Folgenden gehen wir fast zurück zum Anfang und diskutieren das Thema:

(B) Zur Existenz und Eindeutigkeit des Volumens

Analyse: Sei eine Determinantenform Δ gegeben. Aus den definierenden Eigenschaften folgt mit $x_j = \sum_{i_j=1}^{n} b_{i_j}\xi_{i_j j}$ für $j = 1, \ldots, n$ und Basis $B = (b_1, \ldots, b_n)$:

$$\begin{aligned}\Delta(x_1, x_2, \ldots, x_n) &= \Delta(\sum_{i_1=1}^{n} b_{i_1}\xi_{i_1 1}, \ldots, \sum_{i_n=1}^{n} b_{i_n}\xi_{i_n n}) \\ &= \sum_{i_1=1}^{n} \xi_{i_1 1}\, \Delta(b_{i_1}, \sum_{i_2=1}^{n} b_{i_2}\xi_{i_2 2}, \ldots, \sum_{i_n=1}^{n} b_{i_n}\xi_{i_n n}) \\ &= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \xi_{i_1 1}\xi_{i_2 2}\, \Delta(b_{i_1}, b_{i_2}, \sum_{i_3=1}^{n} b_{i_3}\xi_{i_3 3}, \ldots, \sum_{i_n=1}^{n} b_{i_n}\xi_{i_n n}) \\ &= \sum_{i_1}^{n}\sum_{i_2}^{n}\ldots\sum_{i_n}^{n} \xi_{i_1 1}\xi_{i_2 2}\cdots\xi_{i_n n}\, \Delta(b_{i_1}, b_{i_2}, \ldots, b_{i_n}).\end{aligned}$$

Zu summieren ist zunächst über alle Kombinationen $(i_1, i_2, \ldots, i_n) \in \{1, 2, \ldots, n\}^n$, also über die Bilder aller Abbildungen $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ mit $k \mapsto i_k$. Kommt jedoch ein b_j an mehreren Positionen vor, so ist der Summand definitionsgemäß gleich Null. Damit kann man sich auf diejenigen $n-$Tupel $(i_1, i_2, \ldots, i_n)$ beschränken, die eine Permutation $k \mapsto i_k$ von $\{1, 2, \ldots, n\}$ darstellen, also Elemente der symmetrischen Gruppe S_n sind. Es folgt daher

$$(*) \qquad \Delta(x_1, x_2, \ldots, x_n) = \sum_{\pi \in S_n} \xi_{\pi(1)1}\xi_{\pi(2)2}\cdots\xi_{\pi(n)n}\, \Delta(b_{\pi(1)}, b_{\pi(2)}, \ldots, b_{\pi(n)}).$$

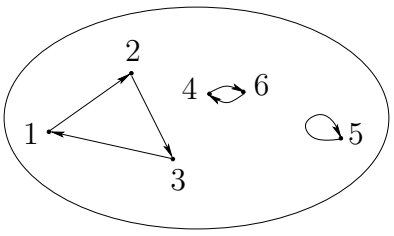
Wir wissen nach Satz 15.3 (a), dass die Vertauschung zweier Vektoren das Vorzeichen der Determinante ändert. Die Frage ist nun, als Produkt von wievielen solcher Vertauschungen eine Permutation $\pi \in S_n$ geschrieben werden kann.

Exkurs: Gerade und ungerade Permutationen

Sei $\pi \in S_n$, also π Permutation von $\{1, \ldots, n\}$. Wir wollen wiederholte Anwendungen von π betrachten; im Folgenden bezeichne wieder π^i die $i-$fache Hintereinanderausführung von π, also $\underbrace{\pi \circ \pi \circ \pi \ldots \circ \pi}_{i \text{ mal}}$.

Beispiel: $n = 6$; $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 6 & 5 & 4 \end{pmatrix}$ (Vgl. Figur 5.1 !)

$$\begin{aligned}\pi^0 &= \mathrm{id}_{\{1,\ldots,6\}} \\ \pi^1(1) &= \pi(1) = 2 \\ \pi^2(1) &= \pi(\pi((1)) = \pi(2) = 3 \\ \pi^3(1) &= \pi(\pi^2(1)) = \pi(3) = 1 \\ \pi^4(1) &= \pi(1) \ldots \\ \pi(4) &= 6 \\ \pi^2(4) &= \pi(6) = 4 \\ \pi(5) &= 5 \ \text{(Fixpunkt)}\end{aligned}$$



Figur 15.1 Bahnen einer Permutation

15.8 Bahnen einer Permutation

(a) *Definition:* Sei $\pi \in S_n$; für $a, b \in \{1, \ldots, n\}$ definieren wir eine Relation

$$a \underset{\pi}{\sim} b :\Leftrightarrow \exists\, k \in \mathbb{Z} : b = \pi^k(a).$$

(b) *Hifssatz:* Für $\pi \in S_n$ ist $\underset{\pi}{\sim}$ eine Äquivalenzrelation auf $\{1, \ldots, n\}$. (Beweis $\cdots$).

(c) *Definition:* Die Äquivalenzklasse von $a \in \{1, \ldots, n\}$ bzgl. $\underset{\pi}{\sim}$, also

$$\{a, \pi(a), \pi^2(a), \ldots\},$$

heißt **Bahn** oder auch **Transitivitätsgebiet** von a unter π.

Fortsetzung des letzten Beispiels (siehe wieder Figur 15.1):
Bahnen von π sind in diesem Fall $\{1, 2, 3\}, \{4, 6\}$ und $\{5\}$.

15.9 Definition: Zyklus

(a) Eine Permutation, deren Bahnen bis auf höchstens eine einelementig sind, (also außerhalb einer Bahn die Punkte fest läßt,) heißt *Zyklus*. Ein Zyklus ist also von der Form

$$\begin{pmatrix} a_1 & a_2 & a_3 & \ldots & a_m & a_{m+1} & \ldots & a_n \\ a_2 & a_3 & a_4 & \ldots & a_1 & a_{m+1} & \ldots & a_n \end{pmatrix}.$$

Wir schreiben den Zyklus auch in der Form $(a_1\ a_2\ a_3 \ldots a_m)(a_{m+1}) \ldots (a_n)$ oder noch kürzer

$$(a_1\ a_2\ a_3 \ldots a_m) \in S_n.$$

(b) **Beispiel:**

(i) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 5 & 6 \end{pmatrix} = (1\ 2\ 3)(4)(5)(6) = (1\ 2\ 3) \in S_6.$

(ii) $(2\ 5\ 1\ 3) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 4 & 1 & 6 \end{pmatrix}$, falls $(2\ 5\ 1\ 3) \in S_6$ gilt.

(c) *Definition:* Zwei Zyklen $(a_1 \ldots a_s)$ und $(b_1 \ldots b_t)$ aus S_n heißen **disjunkt**, falls

$$\{a_1, \ldots, a_s\} \cap \{b_1, \ldots, b_t\} = \emptyset.$$

15.10 Hilfssatz (Disjunkte Zyklen)

Disjunkte Zyklen aus S_n kommutieren.

Beweisskizze. $\begin{pmatrix} b_1 & \ldots & b_t \end{pmatrix} \circ \begin{pmatrix} a_1 & \ldots & a_s \end{pmatrix} = \begin{pmatrix} a_1 & \ldots & a_s & b_1 & \ldots & b_t & \ldots & \ldots \\ a_2 & \ldots & a_1 & b_2 & \ldots & b_1 & \ldots & \ldots \end{pmatrix}$

$$= \begin{pmatrix} a_1 & \ldots & a_s \end{pmatrix} \circ \begin{pmatrix} b_1 & \ldots & b_t \end{pmatrix}.$$

□

15.11 Hilfssatz (Zerlegung in Zyklen)

Seien $n \geq 2$ und $\pi \in S_n \setminus \{\mathrm{id}\}$; dann lässt sich π (bis auf die Reihenfolge der Faktoren) eindeutig als Produkt paarweise disjunkter Zyklen schreiben.

Beweisandeutung: Die Zyklenzerlegung entspricht der Partition in Bahnen. □

Beispiel (Fortsetzung): $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 6 & 5 & 4 \end{pmatrix} = (123) \circ (46) =: (123)(46).$

15.12 Definition: Transposition

Ein Zyklus der Länge 2, d.h. ein solcher von der Form $(a_1 a_2)$, heißt *Transposition*. (Eine Transposition vertauscht also genau 2 Elemente aus $\{1, \ldots, n\}$ und lässt die übrigen fest.)

15.13 Satz (Darstellung als Produkt von Transpositionen)

Jede Permutation aus S_n (mit $n \geq 2$) lässt sich als Produkt von (nicht notwendig disjunkten) Transpositionen schreiben.

Beweis-Andeutung. Nach 15.11 existiert eine Zerlegung in disjunkte Zyklen; jeder Zyklus $(a_1\ a_2\ \ldots\ a_t)$ lässt sich schreiben als[51] $(a_1\ a_t) \circ (a_1\ a_{t-1}) \circ \ldots \circ (a_1\ a_3) \circ (a_1\ a_2)$. □

Beispiel: $\pi_1 := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 1 & 6 & 4 & 2 \end{pmatrix} = (1\ 3) \circ (2\ 5\ 4\ 6) = (1\ 3) \circ (2\ 6) \circ (2\ 4) \circ (2\ 5).$
Andereseits gilt auch $\pi_1 = (1\ 2) \circ (1\ 6) \circ (1\ 4) \circ (1\ 5) \circ (2\ 3) \circ (1\ 2)$. Die Darstellung als Produkt von Transpositionen ist also nicht eindeutig.
Jedoch gilt: (hier ohne Beweis zitiert):

15.14 Satz (Anzahl der Faktoren)

Sei $\pi \in S_n$ mit $n \geq 2$. Dann ist die Anzahl der Faktoren bei Darstellungen von π als Produkt von Transpositionen entweder stets gerade oder stets ungerade.

Damit werden folgende Definitionen sinnvoll:

15.15 Definition: gerade Permutation; Signum

(a) Ist π Produkt einer geraden Anzahl von Transpositionen, so heißt π *gerade Permutation*, andernfalls ungerade Permutation.
(b) Wir definieren das *Signum* einer Permutation wie folgt:

$$\operatorname{sgn} \pi := \begin{cases} +1 & \text{falls } \pi \text{ gerade} \\ -1 & \text{falls } \pi \text{ ungerade.} \end{cases}$$

[51] hier von rechts aus zu lesen; d.h. der letzte Faktor wird zuerst angewandt

Beispiele: sgn id$= +1$, sgn $\tau = -1$ für jede Transposition τ;

$$\operatorname{sgn}\pi_1 = \operatorname{sgn}\begin{pmatrix}1&2&3&4&5&6\\3&5&1&6&4&2\end{pmatrix} = +1 \quad \text{(vgl. voriges Beispiel; 4 bzw. 6 Faktoren!).}$$

Anmerkungen: Die Komposition

- einer geraden und einer ungeraden Permutation ist eine ungerade Permutation
- zweier geraden Permutationen ist eine gerade Permutation
- zweier ungeraden Permutationen ist eine gerade Permutation.

Die Definition von 'sgn' ermöglicht daher folgende Aussage:

15.16 Hilfssatz (Eigenschaften des Signums)

Für $\pi, \rho \in S_n$ gilt $\boxed{\operatorname{sgn}(\pi \circ \rho) = \operatorname{sgn}\pi \cdot \operatorname{sgn}\rho.}$

Tabellarische Darstellung:

$\pi \circ \rho$	π gerade	π ungerade
ρ gerade	gerade	ungerade
ρ ungerade	ungerade	gerade

entsprechend
für das Signum

$\operatorname{sgn}(\pi \circ \rho)$	$\operatorname{sgn}(\pi) = +1$	$\operatorname{sgn}(\pi) = -1$
$\operatorname{sgn}(\rho) = +1$	$+1$	-1
$\operatorname{sgn}(\rho) = -1$	-1	$+1$

Für $n > 1$ ist 'sgn' daher ein Gruppenhomomorhismus von $(S_n, \circ)$ auf die Untergruppe $(\{+1, -1\}, \cdot)$ von $(\mathbb{R}^*, \cdot)$.
Der Kern dieses Homomorphismus, also die Menge aller geraden Permutationen von S_n, bildet einen Normalteiler von S_n. Da die Multiplikation mit einer Transposition τ die Menge der geraden Permutationen bijektiv auf die der ungeraden Permutationen abbildet, folgt insgesamt:

15.17 Satz und Definition: alternierende Gruppe

(i) Die Menge der geraden Permutationen von S_n bilden eine Untergruppe A_n von S_n.
A_n ist Normalteiler von S_n, und es gilt $S_n = A_n \cup A_n \circ \tau$ für jede Transposition τ sowie $|A_n| = \frac{1}{2}|S_n| = \frac{n!}{2}$.

(ii) A_n heißt **alternierende Gruppe** vom Grad n.

Beispiel:
A_3 besteht aus folgenden $\frac{3!}{2} = 3$ Elementen (– den geraden Permutationen von S_3):
id$= (1)$, $(1\ 2\ 3) = (1\ 3) \circ (1\ 2)$, $(1\ 3\ 2) = (1\ 2) \circ (1\ 3) = (1\ 2\ 3)^{-1}$.
Ungerade Permutationen von S_3 sind die Transpositionen: $(1\ 2), (1\ 3), (2\ 3)$.

Ende des Exkurses

Zurück zur Determinante: Kann man eine Permutation π in ein Produkt von t Transpositionen zerlegen, so gilt nach Definition $\operatorname{sgn}\pi = (-1)^t$. Daraus folgt mit 15.3 (ii'):

15.18 Hilfssatz

Seien Δ Determinantenform auf V, $\dim_K V = n$ und $\pi \in S_n$; dann gilt

$$\Delta(x_{\pi(1)}, \ldots, x_{\pi(n)}) = \operatorname{sgn}\pi \cdot \Delta(x_1, \ldots, x_n).$$

Wir können damit die oben erhaltene Beziehung $(*)$ weiter umformen und erhalten:

15.19 Eindeutigkeitssatz

Ist Δ eine Determinantenform auf dem $K-$ Vektorraum V mit Basis $C = (c_1, \ldots, c_n)$, und sind $x_j = \sum_{i=1}^{n} c_i \xi_{ij}$ für $(j = 1, \ldots n)$ n beliebige Vektoren aus V (mit $\xi_{ij} \in K$ für $i, j = 1, \ldots, n$), so gilt:

$(**)$ $$\Delta(x_1, \ldots, x_n) = \gamma \cdot \Big[\sum_{\pi \in S_n} (\operatorname{sgn}\pi) \xi_{\pi(1)1} \cdots \xi_{\pi(n)n} \Big] \quad \text{mit} \quad \gamma = \Delta(c_1, \ldots, c_n).$$

Beispiel: Seien $\dim_K V = 2$ und $C = (c_1, c_2)$ Basis von V. Wegen $S_2 = \{\mathrm{id}, (1\ 2)\}$ ist jede Determinantenform von V von folgender Art:

$$\Delta\Big(\begin{pmatrix} \xi_{11} \\ \xi_{21} \end{pmatrix}_C, \begin{pmatrix} \xi_{12} \\ \xi_{22} \end{pmatrix}_C \Big) = \Big[\underbrace{1 \cdot \xi_{11}\xi_{22}}_{\text{von } \mathrm{id}_{\{1,2\}}} + \underbrace{(-1)\xi_{21}\xi_{12}}_{\text{von } (1\ 2) \in S_2} \Big] \cdot \Delta(c_1, c_2)$$

$$= (\xi_{11}\xi_{22} - \xi_{21}, \xi_{12}) \cdot \gamma \quad \text{mit } \gamma := \Delta(c_1, c_2) \text{ konstant.}$$

15.20 Korollar (Determinantenform und lineare Unabhängigkeit)

Sind Δ Determinantenform auf dem $n-$dim Vektorraum V und $c_1, \ldots, c_n \in V$, so gilt:

$$\Delta(c_1, \ldots, c_n) \neq 0 \quad \Leftrightarrow \quad \{c_1, \ldots, c_n\} \text{ ist Basis von } V \ .$$

Anmerkung: Dies zeigt, dass 15.1 (ii) nicht nur für eine, sondern für jede Basis von V gilt. Außerdem beweist dieses Korollar die Eigenschaft aus 15.5 (i).

Beweis von 15.20.

„$\Rightarrow$ “ Wäre $\{c_1, \ldots, c_n\}$ linear abhängig, so definitionsgemäß $\Delta(c_1, \ldots, c_n) = 0$.
„$\Leftarrow$ “ Seien $\{c_1, \ldots, c_n\}$ Basis von V und $B_1 = \{b_1, \ldots, b_n\}$ die Basis von V, für die $\Delta(b_1, \ldots, b_n) \neq 0$ gemäß 15.1 (iii) gilt. Nach 15.19 existiert ein $\mu \in K$ mit $0 \neq \Delta(b_1, \ldots, b_n) = \mu\Delta(c_1, \ldots, c_n)$, woraus die Behauptung $\Delta(c_1, \ldots, c_n) \neq 0$ folgt. $\square$

15.21 Korollar (Zur Frage der Eindeutigkeit)

Sind Δ_1 und Δ_2 Determinantenformen auf V, so existiert ein $\lambda \in K \setminus \{0\}$ mit

$$\Delta_1 = \lambda \cdot \Delta_2.$$

Bei festem Vektorraum V sind also Determinantenformen bis auf skalare Vielfache bestimmt.

Anmerkung: Wir werden sehen, dass umgekehrt eine Determinantenform Δ existiert und mit Δ auch $\lambda \cdot \Delta$ Determinantenform ist. Dies zeigt:

Die Determinantenformen von V bilden einen $1-$dimensionalen Unterraum des Vektorraums aller $n-$fachen Linearformen von V.

Beweis von 15.21.
Ist $C = (c_1, \ldots, c_n)$ eine geordnete Basis V, so existieren nach 15.19 und 15.20 Elemente $\gamma_i \in K \setminus \{0\}$ mit $\Delta_i(x_1, \ldots, x_n) = \gamma_i \cdot \big[\sum\limits_{\pi \in S_n} \operatorname{sgn} \xi_{\pi(1)1} \cdots \xi_{\pi(n)n} \big]$ für $(i = 1, 2)$. Die Behauptung folgt mit $\lambda := \gamma_1 \cdot \gamma_2^{-1}$; denn damit ergibt sich

$$\Delta_1(x_1, \ldots, x_n) = \lambda\gamma_2 \cdot \big[\sum_{\pi \in S_n} \operatorname{sgn} \pi \ \xi_{\pi(1)1} \cdots \xi_{\pi(n)n} \big] = \lambda \cdot \Delta_2(x_1, \ldots, x_n). \qquad \square$$

Jetzt können wir auch 15.5 (iv))zeigen:

15.22 Korollar (Determinante der transponierten Matrix)

Für $A \in K^{(n,n)}$ gilt $\det A = \det A^T$.

Beweis. Mit $\pi \in S_n$ durchläuft auch ρ für $\rho := \pi^{-1}$ alle Elemente aus S_n. Zu jedem j gibt es außerdem genau ein i mit $j = \pi(i)$, also $i = \rho(j)$. Damit gilt $\prod\limits_{i=1}^{n} \alpha_{\pi(i)i} = \prod\limits_{j=1}^{n} \alpha_{j\rho(j)}$. Wegen $\operatorname{sgn} \rho = \operatorname{sgn} \pi^{-1} = \operatorname{sgn} \pi$ erhält man:

$$\det A = \Delta_0(a_{\bullet 1}, \ldots, a_{\bullet n}) = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^{n} \alpha_{\pi(i)i} = \sum_{\rho = \pi^{-1} \in S_n} \operatorname{sgn} \rho \prod_{j=1}^{n} \alpha_{j\rho(j)} = \det A^T.$$

Nachzuholen ist nun noch der Nachweis der Existenz von Determinantenformen für endlich-dimensionale Vektorräume.

15.23 Existenzsatz

Sei V ein $n-$dimensionaler $K-$ Vektorraum; seien ferner $C = (c_1, \ldots, c_n)$ eine fest gewählte Basis von V und $\gamma \in K \setminus \{0\}$ sowie $x_j = \sum_{i=1}^{n} c_i \xi_{ij} \in V$ für $j = 1, \ldots, n$. Dann wird durch

$$\Delta(x_1, \ldots, x_n) := \gamma \cdot \Big[\sum_{\pi \in S_n} (\operatorname{sgn} \pi) \xi_{\pi(1)1} \cdots \xi_{\pi(n)n} \Big]$$

(vgl. 15.19 $(**)$!) eine Determinantenform auf V mit $\Delta(c_1, \ldots, c_n) = \gamma$ definiert.

Beweisskizze. Sei $\Delta(x_1, \ldots, x_n) := \gamma \cdot \Big[\sum_{\pi \in S_n} (\operatorname{sgn} \pi) \xi_{\pi(1)1} \cdots \xi_{\pi(n)n} \Big]$ für $x_j = \begin{pmatrix} \xi_{1,j} \\ \vdots \\ \xi_{nj} \end{pmatrix}_C$.
Man prüft leicht nach, dass Δ in jedem seiner Argumente linear ist. Weiter gilt

$$\Delta(c_1, \ldots, c_n) := \gamma \cdot \Big[\sum_{\pi \in S_n} (\operatorname{sgn} \pi) \underbrace{\prod_{j=1}^{n} \delta_{\pi(j)j}}_{\neq 0 \text{ nur für } \pi = \mathrm{id}} \Big] = \gamma \cdot 1 = \gamma \neq 0 \ .$$

Zu zeigen bleibt noch, dass aus der linearen Abhängigkeit von $x_1, \ldots, x_n$ die Aussage $\Delta(x_1, \ldots, x_n) = 0$ folgt. Seien also $x_1, \ldots, x_n$ linear abhängig; dann existieren ein $i \in \{1, \ldots, n\}$ und Elemente $\lambda_1, \ldots, \lambda_n \in K$ mit $x_i = \sum_{j=1, j \neq i}^{n} x_j \lambda_j$, und es gilt (wegen der Linearität in der $i-$ten Komponente):

$$\Delta(x_1, \ldots, x_n) = \sum_{j=1, j \neq i}^{n} \lambda_j \cdot \Delta(x_1, \ldots, \underbrace{x_j}_{i-\text{te Stelle}}, \ldots, \underbrace{x_j}_{j-\text{te Stelle}}, \ldots, x_n).$$

Es reicht also, $\Delta(x_1, \ldots, x_n) = 0$ für $x_i = x_j$ für ein Indexpaar $i \neq j$ zu zeigen. Mit der Transposition $\tau := (i\ j) \in S_n$ folgt dann nach 15.17, dass $S_n = A_n \cup A_n \circ \tau$ und daher

$$\Delta(x_1, \ldots, x_n) = \gamma \sum_{\sigma \in A_n} \Big(\underbrace{1}_{\operatorname{sgn} \sigma} \cdot \prod_k \xi_{\sigma(k)k} + \underbrace{(-1)}_{\operatorname{sgn}(\sigma \circ \tau)} \cdot \prod_k \xi_{\sigma \circ \tau(k)k} \Big)$$

gilt. Für $k \notin \{i, j\}$ ist $(\sigma \circ \tau)(k) = \sigma(k)$ und somit $\xi_{\sigma \circ \tau(k)k} = \xi_{\sigma(k)k}$. Wegen $x_i = x_j$ gilt $\xi_{ei} = \xi_{ej}$ für alle $e = 1, \ldots, n$; somit ist

$$\xi_{(\sigma \circ \tau)(i)i} = \xi_{\sigma(j)i} = \xi_{\sigma(j)j} \quad \text{und} \quad \xi_{(\sigma \circ \tau)(j)j} = \xi_{\sigma(i)j} = \xi_{\sigma(i)i}.$$

Insgesamt folgt daher die Gleichheit der Faktoren von $\prod_k \xi_{\sigma(k)k}$ und $\prod_k \xi_{\sigma \circ \tau(k)k}$ für alle $\sigma \in A_n$, woraus sich die Behauptung ergibt.

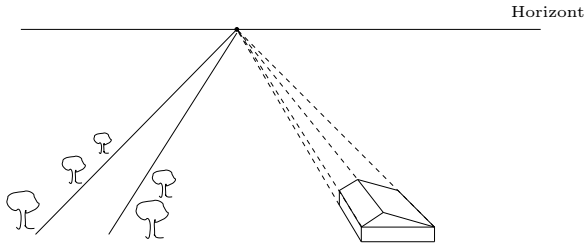
15.24 Anmerkung: Existenz und Eindeutigkeit der Determinante

Aufgrund des Existenz- und des Eindeutigkeits-Satzes gibt es zu gewählter Basis $C = (c_1, \ldots, c_n)$ genau eine Determinantenform Δ_0 mit $\Delta_0(c_1, \ldots, c_n) = 1$; insbesondere gilt dies für $V = K^n$ und Wahl der kanonischen Basis von K^n als C. Damit ist die **Determinante** einer Matrix $A = (a_{\bullet 1}, \ldots a_{\bullet n}) \in K^{(n,n)}$ durch $\det A := \Delta_0(a_{\bullet 1}, \ldots, a_{\bullet n})$ eindeutig definiert (vgl. 15.4).

§ 16 Etwas projektive Geometrie

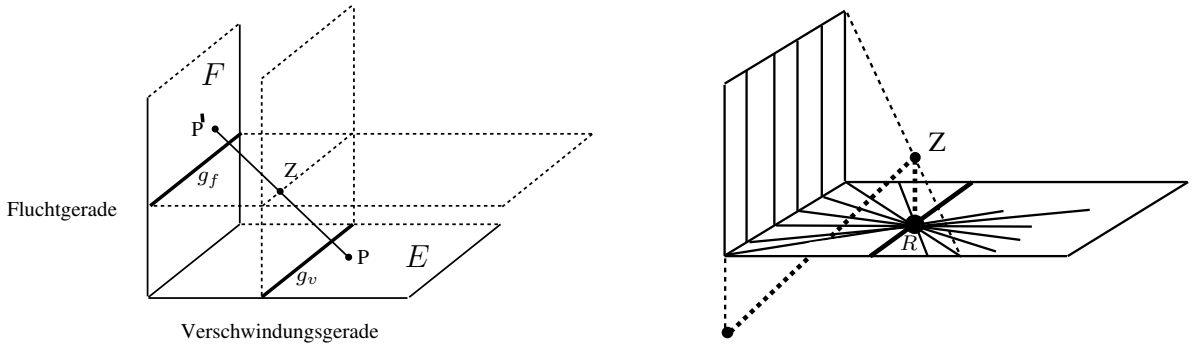
16.1 Motivation: Zentralprojektion

Auf Fotografien und Bildern treffen sich parallele Geraden oft in einem Punkt am Horizont. Bei perspektivischen Zeichnungen sind solche Punkte, sogenannte *Fluchtpunkte*, wesentliche Konstruktionshilfen.



Figur 16.1: Parallele Geraden treffen sich im „Unendlichen“.

Wir behandeln nun zuerst eine **Zentralprojektion** ζ in $\mathbb{R}^3$ mit Zentrum Z von einer Ebene E auf eine dazu nicht parallele Ebene F. Dabei wird einem Punkt P von E der Schnittpunkt P' der Geraden ZP mit F zugeordnet. (Siehe Figur 16.2 a) !)



a) b)

Figur 16.2: Zentralprojektion der Ebene E auf die Ebene F mit Zentrum Z

Nicht alle Punkte von E haben einen Bildpunkt und nicht alle Punkte von F ein Urbild: Die Schnittgerade g_v der durch Z gehenden zu F parallelen Ebene mit E besteht genau aus denjenigen Punkten, die kein Bild besitzen. Analog besitzen die Punkte der Schnittgeraden g_f der mit Z inzidierenden zu E parallelen Ebene mit F kein Urbild. (Siehe ebenfalls Figur 16.2 a) !). Die Gerade g_v heißt **Verschwindungsgerade**, die Gerade g_f **Fluchtgerade**,
Bei der Abbildung $\zeta : E \setminus g_v \to F \setminus g_f$ werden Geraden von $E \setminus g_v$ (die ja mit Z eine Ebene aufspannen) auf Geraden von $F \setminus g_f$ (die Schnitte dieser Ebenen mit F) abgebildet – und umgekehrt haben Geraden von $F \setminus g_f$ als Urbilder Geraden von $E \setminus g_v$.

Betrachtet man nun, was mit einem Geradenbüschel durch einen Punkt R von g_v passiert, so sieht man, dass die Bilder dieser Geraden in F keinen Punkt gemeinsam haben und damit parallel sind. (Siehe Figur 16.2 b)!) „Der Bildpunkt von R liegt also auf dieser Parallelenschar im Unendlichen“. Dadurch motiviert definiert man:

16.2 Projektive Erweiterung der reellen affinen Ebene

Als *projektive Erweiterung* $\mathrm{PG}(\mathcal{A})$ der reellen affinen Ebene $\mathcal{A} = \mathrm{AG}(\mathbb{R}^2) =: \mathrm{AG}(2, \mathbb{R})$ bezeichnen wir folgende Geometrie:

(i) *Punkte* von $\mathrm{PG}(\mathcal{A})$ sind die Punkte von $\mathcal{A}$ (die **eigentlichen Punkte**) und die Parallelenscharen der Geraden von $\mathcal{A}$. Also: Jede Menge aller zu einer Geraden parallelen Geraden bildet per definitionem einen neuen Punkt, einen sogenannten „**uneigentlichen oder idealen Punkt**“.

(ii) *Geraden* von $\mathrm{PG}(\mathcal{A})$ sind die Geraden von $\mathcal{A}$, jede erweitert um den eindeutig bestimmten uneigentlichen Punkt in ihrer Richtung, und eine weitere Gerade, bestehend aus allen uneigentlichen Punkten, die **uneigentliche Gerade** g_∞.

(iii) Ein eigentlicher Punkt *inzidiert* mit einer eigentlichen Geraden, falls das auch in $\mathcal{A}$ der Fall ist. Ein uneigentlicher Punkt inzidiert mit einer eigentlichen Geraden, falls diese zu seiner Parallenschar gehört. Und g_∞ inzidiert genau mit den uneigentlichen Punkten.

Übungsaufgabe 16.1 Zeigen sie, dass in $\mathrm{PG}(\mathcal{A})$ je zwei Punkte durch genau eine Gerade verbunden sind und je zwei Geraden sich in genau einem Punkt schneiden.

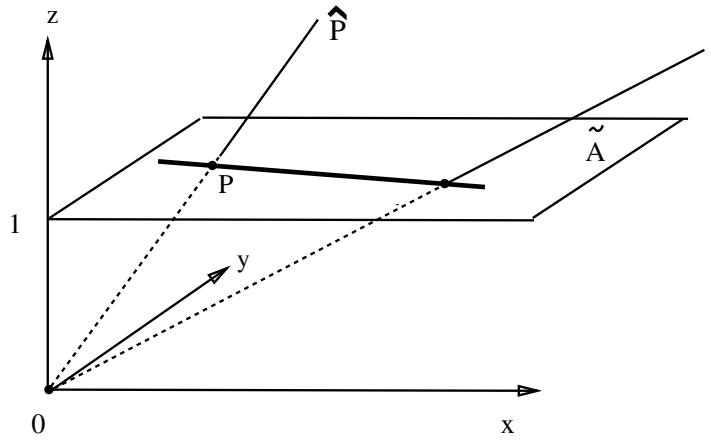
Um die neue Geometrie koordinatisieren zu können, wählen wir Z als den Nullpunkt O unseres Koordinatensystem und E als die (zu $\mathcal{A}$ isomorphe) Ebene $\tilde{A}$ mit der Gleichung $z = 1$; nun ordnen wir jedem Punkt P von $\tilde{A}$ die Gerade $\hat{P} := PO$ zu und jeder Geraden g von $\tilde{A}$ die Ebene, die von g und O aufgespannt wird. (Siehe Figur 16.3).

16.3 Definition: homogene Koordinaten

Ein eigentlicher Punkt der affinen Ebene $\mathcal{A}$ mit den Koordinaten $(x, y) \in \mathcal{A}$ (bzw. $(x, y, 1) \in \tilde{A}$) entspricht der Geraden $(x, y, 1)\mathbb{R}$; die uneigentlichen Punkte entsprechen den Geraden der Form $(x, y, 0)\mathbb{R}$. Umgekehrt definiert jede Gerade einen eigentlichen oder uneigentlichen Punkt.
Jedem Koordinatentripel (ξ_1, ξ_2, ξ_0) mit $\xi_0 \neq 0$ ordnen wir daher einen affinen Punkt zu durch:

$$\hat{P} := (\xi_1, \xi_2, \xi_0)\mathbb{R} \mapsto (\frac{\xi_1}{\xi_0}, \frac{\xi_2}{\xi_0}, 1) = P.$$

(ξ_1, ξ_2, ξ_0) heißt *homogenes Koordinatentripel* von P; es ist nur bis auf einen Faktor eindeutig bestimmt.
Analog lässt sich jedem $(\xi_1, \xi_2, 0)$ mit $\xi_1 \neq 0 \vee \xi_2 \neq 0$ ein uneigentlicher Punkt zuordnen. Man kommt damit zu der in folgender Definition beschriebenen Geometrie:



Figur 16.3: Zur Einführung homogener Koordinaten
$(x, y) \mapsto P = (x, y, 1) \mapsto \hat{P} = (x, y, 1)\mathbb{R})$.

16.4 Definition: Reelle projektive Ebene

Die *projektive Ebene* $\mathrm{PG}(2, \mathbb{R})$ =$\mathrm{PG}(\mathbb{R}^3)$ ist wie folgt definiert:

(i) Punkte sind die $1-$dim Unterräume von $\mathbb{R}^3$.

(ii) Geraden sind die $2-$dim Unterräume von $\mathbb{R}^3$.

(iii) Inzidenz ist Enthaltensein.

16.5 Anmerkungen zur projektiven Geometrie

a.) Eine Verallgemeinerung auf beliebige Schiefkörper und auf andere Dimensionen ist möglich.

b.) Der Vorteil der projektiven Geometrie ist u.a., dass

- die Punkte, Geraden, Ebenen etc. jetzt lineare Unterräume und keine anderen affinen Unterräume sind
- affin-lineare Abbildungen, wie man zeigen kann, nun durch lineare Abbildungen dargestellt werden
- Fallunterscheidungen bzgl. Existenz von Schnittpunkten von Geraden in einer Ebene nicht nötig sind
- Punkte und Hyperebenen „dual“ zueinander sind
- die reelle affine Ebene (nach Auszeichnung einer Geraden als uneigentliche Gerade) in $\mathrm{PG}(2, \mathbb{R})$ wiederzufinden ist (bzw. der $n-$dim affine Raum über K durch Auszeichnung einer Hyperebene in $\mathrm{PG}(K^{n+1})$ als uneigentliche Hyperebene erhalten wird.)
- Kurven, z.B. Kegelschnitte, sich “vereinheitlichen” lasssen: Hyperbeln und Parabeln als 'Ellipsen' mit uneigentlicher Sekante bzw. Tangente.

16.6 Anwendung bei Elliptischen Kurven

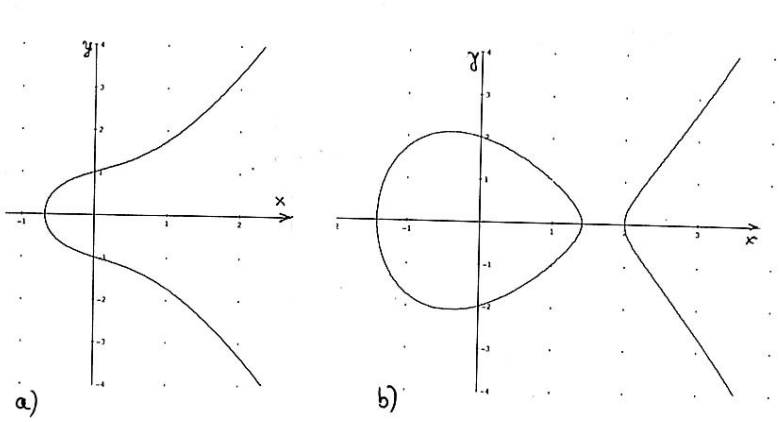
(a) *Definition*
Sei K ein Körper; oft wird hier $K \in \{\mathbb{R}, \mathrm{GF}(p)\}$ genommen. Seien ferner $b, c \in K$ fest gewählt. Wir betrachten zunächst in der affinen Eben $\mathrm{AG}(K^2)$ über K die kubische Kurve (vgl. Figur 16.4) der Punkte, die die Gleichung

$$(*) \quad y^2 = x^3 + bx + c$$

erfüllen, also

$$\mathcal{C} := \big\{(x, y) \in K^2 \mid y^2 = x^3 + bx + c\big\}.$$

$\mathcal{C}$ ist eine *elliptische Kurve* (vom sogenannten 'Weierstrass-Typ') (vgl. Figuren 16.4/16.5)



Figur 16.4: Beispiele kubischer Kurven im Reellen mit Gleichung $y^2 = f(x)$, wobei $f(x)$ ein Polynom vom Grad 3 ist mit a) einer reellen Nullstelle b) drei reellen Nullstellen ist

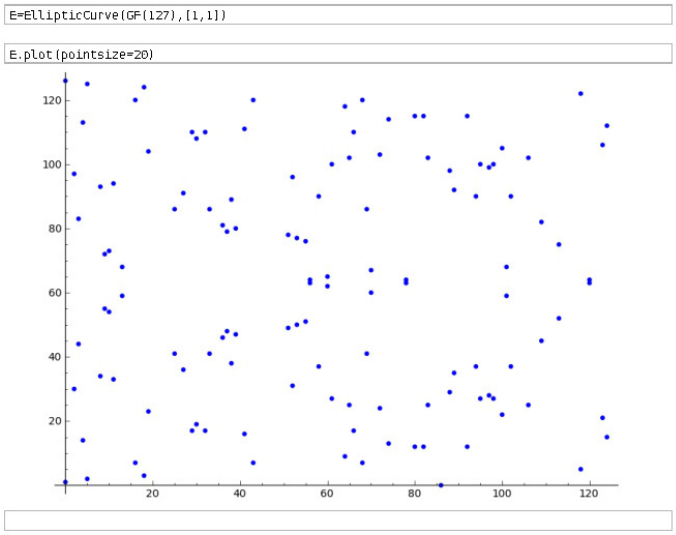
(b) *Projektive Koordinaten*
Ersetzt man x und y in Gleichung $(*)$ durch

$$x = \frac{\xi_1}{\xi_0} \quad \text{und} \quad y = \frac{\xi_2}{\xi_0} \quad \text{mit} \quad \xi_0 \neq 0,$$

so ergibt sich (nach Multiplikation mit ξ_0^3)

$$(*') \quad \xi_2^2 \xi_0 = \xi_1^3 + b\xi_1 \xi_0^2 + c\xi_0^3.$$

Für die affinen Punkte ist $(*)$ zu $(*')$ äquivalent. Für einen uneigentlichen Punkt (mit Koordinate $\xi_0 = 0$) ist $(**)$ genau für $\mathcal{O} := K(0, 1, 0)$ erfüllt. Dieser Punkt kommt also als einziger projektiver Punkt zu $\mathcal{C}$ hinzu.
$\mathcal{O}$ liegt auf den Geraden mit Gleichung $\xi_1 - c\xi_0 = 0$, also auf den Geraden mit affiner Gleichung $x = c$ (für $c \in K$). Daher ist $\mathcal{O}$ der Fernpunkt in $y-$Richtung.



Figur 16.5: Elliptische Kurve über $\mathbb{Z}_{127}$ (erstellt mithilfe von Sage; vgl. Schulz/Witten/Esslinger: Rechnen mit Punkten einer elliptischen Kurve. LOG IN Heft 181/182 (2015))

(c) *Anmerkung 1:*

Auf $\overline{\mathcal{C}} := \mathcal{C} \cup \{O\}$ lässt sich eine Addition definieren, im Wesentlichen wie folgt:
$P + Q$ ist definiert als der Punkt, den man durch Spiegelung des Schnittpunktes der Geraden PQ mit $\mathcal{C}$ an der $x-$Achse erhält (s.z.B. Figur 16.6a), und $R + R$ als der an der $x-$Achse gespiegelte Schnittpunkt der Tangente an $\mathcal{C}$ in R (s.Figur 16.6b). Wie man zeigen kann, erhält man so zusammen mit $\mathcal{O}$ als neutralem Element eine kommutative Gruppe.

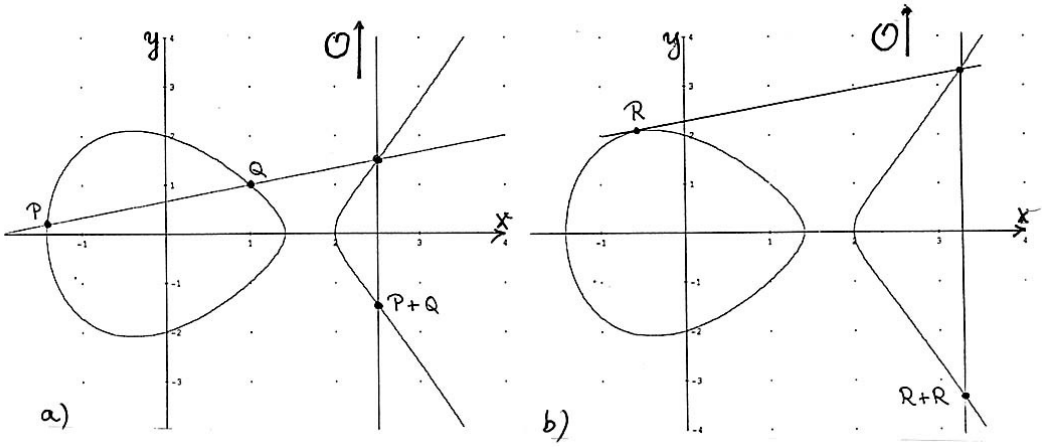
(d)*Anmerkung 2:*

Für einen geeigneten fest gewählten Punkt $P \in \mathcal{C}$ definiert man

$$Q = mP := P + P + \ldots + P \quad \text{(mit } m \text{ Summanden)}.$$

Dabei ist die Berechnung von Q aus m und P vergleichsweise leicht, umgekehrt die Berechnung von $m := \log_P Q$ viel schwerer. Dieser Sachverhalt ist zum Chiffrieren in der Kryptographie ausnutzbar (–> Problem des diskreten Logarithmus.) Bei der Elliptischen-Kurven-Kryptographie (ECC) ist die vom Bundesamt für Sicherheit in der Informationstechnik (BSI) empfohlene Schlüssellänge[52]. von mindestens 250 Bit für die Ordnung des Basispunktes P (bei Verwendung bis 2022) deutlich niedriger als die empfohlene Schlüssellänge von 2000 Bit beim RSA-Verfahren[53]

[52]https://www.bsi.bund.de>BSI>TechnischeRichtlinien>BSI-TR-02102
[53]s.z.B. R.-H.Schulz: Codierungstheorie. Eine Einführung, Verlag Fr.Vieweg&Sohn, 2.Auflage 2003.

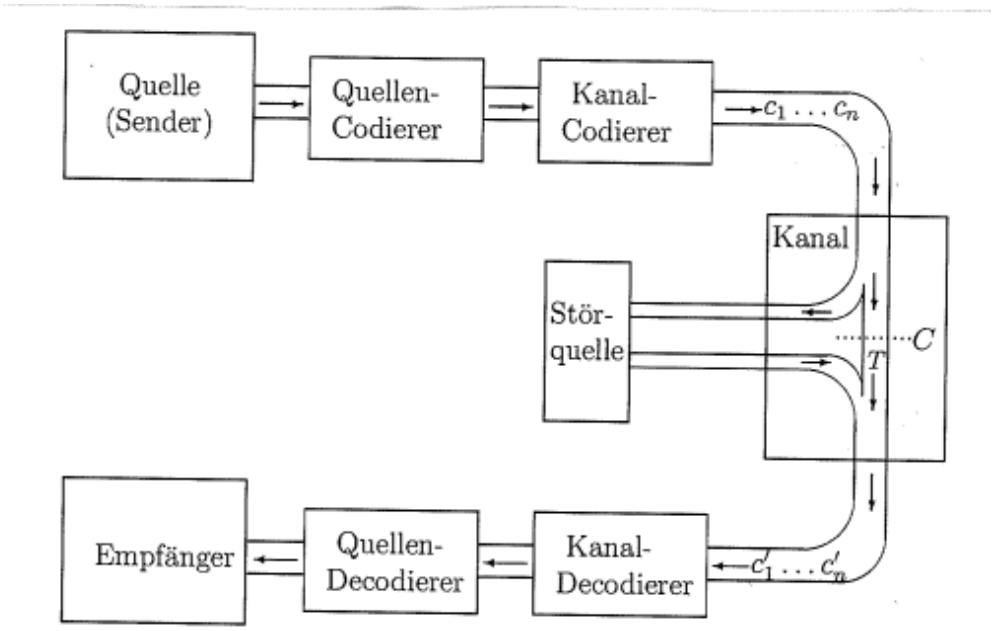


Figur 16.6: Zur Definition der Summe von Punkten
einer elliptischen Kurve ('Sehnen/Tangenten-Addition')
(vgl. Schulz/Witten/Esslinger, l.c.)

E: Anwendungsbeispiele in der Codierungstheorie

E 1.0 Schema eines Nachrichtenübertragungssystems

Das Schema eines Nachrichenübertragungssystems zeigt Figur E 1:



Figur E.1: Schema eines Nachrichtenübertragungssystems
mit gestörtem Kanal

Wir beschreiben zunächst ein Codier- und ein Decodierverfahren. Dass dieses eine Korrektur von Übertragungsfehlern gestattet, erläutern wir dann im Anschluss.

1.1 Definition Wort

In der Codierungstheorie schreit man ein $m-$Tupel $(\alpha_1, \ldots, \alpha_m) \in A^m$ (mit $m \in \mathbb{N}$) über der Menge A meist ohne Klammern und Kommata und nennt es **Wort** (Plural: Wörter, nicht Worte) der Läge m über dem **Alphabet** A. Nach der Quellencodierung seien die Nachrichten Wörter der Länge k über einem Körper K, also von der Form

$$\alpha_1 \ldots \alpha_k := (\alpha_1, \ldots, \alpha_k) \in K^k.$$

1.2 Lineare Codierung

Unter einer **linearen Codierung** dieser Nachrichten versteht man die durch Multiplikation mit einer Matrix $G \in K^{(k,n)}$ vom Rang k gegebene lineare Abbildung:

$$c_G : \begin{cases} K^k & \longrightarrow & K^n \\ (\alpha_1, \ldots, \alpha_k) & \mapsto & (\alpha_1, \ldots, \alpha_k) \cdot G = \sum_{i=1}^{k} \alpha_i g_{i\bullet} \end{cases}$$

(Hierbei bezeichnet $g_{i\bullet}$ die $i-$te Zeile von G.)

1.3 Beispiel

Seien K =GF(2) $= (\{0,1\}, \oplus, \odot) = \mathbb{Z}_2$, ferner $k = 4, n = 7$ und

$$G_1 := \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right) \in \mathrm{GF}(2)^{(4,7)}.$$

Die Nachricht 1100 wird dann codiert zu

$$1100 \cdot G_1 = g_{1\bullet} \oplus g_{2\bullet} = 1100010.$$

Da G_1 als linke $4 \times 4-$Untermatrix die $4 \times 4-$ Einheitsmatrix I_4 hat, ist beim Codewort

$$c_G(\alpha_1\alpha_2\alpha_3\alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_1 + \alpha_2 + \alpha_3, \alpha_2 + \alpha_3 + \alpha_4, \alpha_1 + \alpha_2 + \alpha_4)$$

das Anfangsstück (*Präfix*) der Länge 4 gleich der ursprünglichen Nachricht. (Zu diesen **Informationsbits** kommen noch drei **Kontrollbits** .) Eine solche Codierung heißt **systematische Codierung**.

1.4 Generatormatrix eines linearen Codes

Der **Code** $\mathcal{C}$, also die Menge der Codewörter

$$\text{Bild}\, c_G = \{(\alpha_1, \ldots, \alpha_k) \cdot G \mid \alpha_1, \ldots, \alpha_k \in K\} \subseteq K^n$$

ist ein Unterraum von K^n. Daher heißt C **linearer Code** oder **Linearcode**. Die Bilder der Einheitsvektoren von K^k unter c_G sind die Zeilen von G; diese erzeugen daher $\mathcal{C}$:

$$\mathcal{C} = < g_{1\bullet}, \cdots, g_{k\bullet} > .$$

G heißt dementsprechend eine **Generatormatrix**, oft auch **Basismatrix** von $\mathcal{C}$.

Beispiel (Fortsetzung):
Der Code $\mathcal{C}_1$ mit Generatormatrix G_1 hat Dimension 4 und Wörter der Länge 7. Er heißt **(7, 4)−Hammingcode** (nach R.W. Hamming, dem Erfinder einer ganzen Serie von Codes).

1.5 Kontrollmatrix eines linearen Codes

$\mathcal{C}$ kann (wie jeder Unterraum von K^n) als Lösunsraum eines homogenen linearen Gleichungssystems dargestellt werden:

$$c \in \mathcal{C} \quad \Leftrightarrow \quad H \cdot c^T = 0.$$

Dazu sucht man als Zeilen der Koeffizientenmatrix H, der sogenannten **Kontrollmatrix**, $n - k$ linear unabhängige Vektoren $h_1, \ldots, h_{n-k}$, die auf $\mathcal{C}$ senkrecht stehen, für die also gilt

$$h_i \cdot g_{j\bullet}^T = 0 \qquad (i = 1, \ldots, n-k, \quad j = 1, \ldots, k).$$

Wegen der Linearität der Abbildung

$$S_H : \begin{cases} K^n & \longrightarrow & K^{(n-k,1)} \\ v & \mapsto & H \cdot v^T \end{cases}$$

wird durch S_H dann jeder Codevektor als Linearkombination von $\{g_{1\bullet}, \ldots, g_{k\bullet}\}$ annulliert. Aus Dimensionsgründen gilt sogar

$$\mathcal{C} = \text{Kern}\, S_H.$$

Spezialfall: Ist $G = (I_k \mid B)$, so kann man $H = (-B^T \mid I_{n-k})$ wählen.
Beispiel (Fortsetzung)
Ist $G = G_1$, so wählen wir z.B. die $(4 \times 7)-$ Matrix

$$H_1 = \begin{pmatrix} 1\,1\,1\,0 & 1\,0\,0 \\ 0\,1\,1\,1 & 0\,1\,0 \\ 1\,1\,0\,1 & 0\,0\,1 \end{pmatrix}.$$

Für diese gilt insbesondere $H_1 \cdot {G_1}^T = 0$ und $H_1 \cdot c^T = 0$ für jeds $c \in \mathcal{C}_1$.
Z.B. folgt:

$$H_1 \cdot (1,1,0,0,0,1,0)^T = h_{\bullet 1} \oplus h_{\bullet 2} \oplus h_{\bullet 6} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

1.6 Fehlerkorrektur

Nach der Übertragung über einen Kanal seien die

empfangene Nachricht	y	$=$	η_1	$\dots$	η_n
gesendete Nachricht	c	$=$	γ_1	$\dots$	γ_n

und der sogenannte **Fehlervektor** $e := y - c = \eta_1 - \gamma_1 \quad \dots \quad \eta_n - \gamma_n$.

Wieder wegen der Linearität von S_H und wegen $S_H(c) = 0$ gilt:

$$S_H(y) = S_H(c+e) = S_H(c) = S_H(e).$$

Der Vektor $S_H(y)$ ist daher unabhängig von der gesendeten Nachricht c und nur vom Fehervektor abhängig; er heißt **Syndrom** von y. Die Abbildung S_H wird daher **Syndrom-Abbidung** genannt.
Sind nun die Syndrome $S_H(e_1), \dots, H(e_n)$, d.h. die Spalten von H paarweise verschieden, so kann man uner der Voraussetzung, dass ein Fehler in nur einer Komponente von c vorgekommen ist, diese Komponente bestimmen und im Falle $K = \mathrm{GF}(2)$ korrigieren. Ein solcher Code heißt 1-fehlererkennend bzw. **1-fehlerkorrigierend**. (Sind allerdings in einem Codewort mehrere Stellen verfälscht worden, so kann die Fehlerkorrektur sogar deren Anzahl vergrößern.)

Beispiel: $(7,4)-$Hammingcode (Fortsetzung):

In $H_1 = \begin{pmatrix} 1\,1\,1\,0 & 1\,0\,0 \\ 0\,1\,1\,1 & 0\,1\,0 \\ 1\,1\,0\,1 & 0\,0\,1 \end{pmatrix}$ sind die 7 Spalten paarweise verschieden. Daher ist der $(7,4)-$Hammingcode $\mathcal{C}_1 = \text{Kern } S_{H_1}$ 1-fehlerkorrigierend.
Ist etwa das übertragene Codewort $c_1 = 1100010$ zu $y_1 = 1101010$ verfälscht, so liefert y_1 das Syndrom

$$S_{H_1}(y_1) = H_1 \cdot y_1 = \begin{pmatrix} 1\,1\,1\,0 & 1\,0\,0 \\ 0\,1\,1\,1 & 0\,1\,0 \\ 1\,1\,0\,1 & 0\,0\,1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = h_{\bullet 1} \oplus h_{\bullet 2} \oplus h_{\bullet 4} \oplus h_{\bullet 6}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = h_{\bullet 4};$$

Der Sequenzraum: Codes und Kugelpackungen

Wir betrachten im Folgenden Codes fester Länge n, sogenannte **Blockcodes**, also Teilmengen von A^n.Wir führen einen *Abstandsbegriff* ein:

E1.7 Hamming-Abstand, Sequenzraum

(a) **Definition: Hamming-Abstand**

Seien $\mathbf{a} = a_1 \dots a_n$ und $\mathbf{b} = b_1 \dots b_n$ Wörter der Länge n über dem Alphabet A; dann bezeichnet $d_H(\mathbf{a}, \mathbf{b})$ die Anzahl der unterschiedlichen Komponenten von $\mathbf{a}$ und $\mathbf{b}$. Diese Zahl $d_H(\mathbf{a}, \mathbf{b}) = \mid \{i \mid a_i \neq b_i\} \mid$ heißt *Hamming-Abstand (Hamming-Distanz)* zwischen $\mathbf{a}$ und $\mathbf{b}$. (Wir schreiben oft nur d statt d_H).

Beispiel:

$$\begin{array}{rcl} d(0001\underline{0}, 0001\underline{1}) & = & 1 \\ d(000\underline{11}, 000\underline{00}) & = & 2 \end{array} \quad = \quad \begin{array}{l} d(000\underline{1}0, 000\underline{0}0). \\ d(\underline{1}000\underline{1}, \underline{0}000\underline{0}). \end{array}$$

(b) *Eigenschaft:* Der Hamming-Abstand $d_H : A^n \times A^n \to \mathbb{N}_0$ ist eine *Metrik* auf A^n,

d.h. es gilt (für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in A^n$):

(i) $d(\mathbf{x}, \mathbf{y}) \geq 0$ (Positivität) (ii) $d(\mathbf{x}, \mathbf{y}) = 0 \Longrightarrow \mathbf{x} = \mathbf{y}$

(iii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (Symmetrie)

(iv) $d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z})$ (Dreiecksungleichung) .

Beweisskizze
(i), (ii) und (iii) sind offensichtlich: Die Dreiecksungleichung (iv) sieht man folgendermaßen ein: Ist i die Nummer einer Komponente, in der sich $\mathbf{x}$ und $\mathbf{z}$ unterscheiden ($x_i \neq z_i$), so gilt mindestens eine der Beziehungen $x_i \neq y_i$ oder $y_i \neq z_i$. Eine Komponente, die einen Beitrag 1 zu $d(\mathbf{x}, \mathbf{z})$ liefert, gibt einen solchen auch für $d(\mathbf{x}, \mathbf{y})$ oder für $d(\mathbf{y}, \mathbf{z})$. □

(c) **Definition: Sequenzraum**

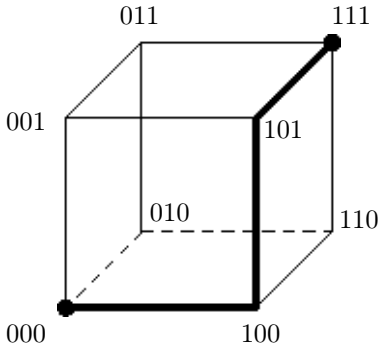
Der metrische Raum (A^n, d_H), bestehend aus der Menge aller Wörter der Länge n über A und dem Abstand d_H, heißt *Sequenzraum*. Ist speziell $A = \{0, 1\}$, so spricht man auch vom $n-$dimensionalen **Einheitswürfel**.

(Der Sequenzraum spielt u.a. auch in der Evolutionstheorie eine Rolle; die Hamming-Distanz zwischen zwei Genen ist dort die minimale Anzahl der Punktmutationen, die nötig ist, um das Muster eines der beiden Gene in das des anderen umzuwandeln.)

E 1.8 Beispiel:

$(\{0, 1\}^3, d_H)$ lässt sich mittels eines Würfels der Kantenlänge 1 im 3-dim reellen Raum darstellen; die Wörter von $\{0, 1\}^3$ sind den Ecken des Würfels zugeordnet; der Abstand

zweier Wörter ist die Länge des kürzesten Weges zwischen den zugeordneten Ecken längs einer Würfelkante (s. Figur E.2).



Figur E.2: Darstellung des Sequenzraums $(\{0,1\}^3, d_H)$.
(Markiert ist ein kürzester Weg von 000 zu 111;
seine Länge ist 3 – entsprechend dem Hamming-Abstand.)

Aufgabe E.1
Zeigen Sie: Zu gegebenem Wort $\mathbf{z} \in A^n$ gibt es $\binom{n}{m}(|A|-1)^m$ Wörter in A^n, die Abstand m von $\mathbf{z}$ haben (für $0 \leq m \leq n$).

Das Begehen von t Fehlern bei Übermittlung eines Codeworts $\mathbf{c}$ führt zu einem Wort $\mathbf{x}$ des Sequenzraums, das sich von $\mathbf{c}$ in t Komponenten unterscheidet. Dieser Sachverhalt legt es nahe, Sphären oder Kugeln um Codewörter zu betrachten; dabei sind Kugeln wie in metrischen Räumen üblich definiert:

E 1.9 Definition: Kugel

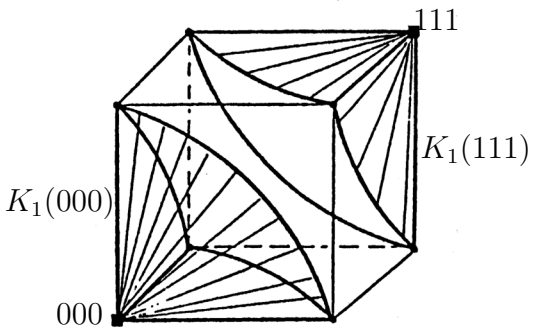
Ist $\mathbf{z} \in A^n$ und $t \in \mathbb{N} \cup \{0\}$, dann heißt

$$K_t(\mathbf{z}) := \{\mathbf{x} \in A^n \mid d_H(\mathbf{x}, \mathbf{z}) \leq t\}$$

Kugel vom Radius t um den Mittelpunkt $\mathbf{z}$.

E 1.10 Beispiel (Fortsetzung):

Figur E.3 skizziert die Kugeln $K_1(000) = \{000,\ 100,\ 010,\ 001\}$ und $K_1(111) = \{111,\ 011,\ 101,\ 110\}$ vom Radius 1 in $(\{0,1\}^3, d_H)$.



Figur E.3: Zwei Kugeln vom Radius 1 in $(\{0,1\}^3, d_H)$.

E 1.11 Bemerkungen: Kugeln und Fehlerkorrektur

(i) Die Elemente von A^n, die aus $\mathbf{z} \in A^n$ durch *Fehler in bis zu t Komponenten* entstehen, liegen in $K_t(\mathbf{z})$. Umgekehrt kann jedes Wort von $K_t(\mathbf{z})$ aus $\mathbf{z}$ durch t oder weniger Fehler entstanden sein.

(ii) Ist $\mathbf{b} \in K_t(\mathbf{z}_1) \cap K_t(\mathbf{z}_2)$, so kann $\mathbf{b}$ durch bis zu t Fehler aus $\mathbf{z}_1$ oder aus $\mathbf{z}_2$ entstanden sein.

(iii) Ist $\mathcal{C}$ ein Blockcode mit der Eigenschaft

$$(*) \qquad K_t(\mathbf{c}_i) \cap K_t(\mathbf{c}_j) = \emptyset \quad \text{für alle } \mathbf{c}_i, \mathbf{c}_j \in \mathcal{C} \text{ mit } \mathbf{c}_i \neq \mathbf{c}_j,$$

so lassen sich (durch Decodierung zum Mittelpunkt derjenigen Kugel, in der das fehlerhafte Wort liegt) Wörter mit bis zu t Fehlern korrigieren.

(iv) Ein einfaches *Decodierschema* besteht dabei aus der Auflistung der Kugeln um Codewörter und ihrer Elemente mit der Vorschrift, zum Kugelmittelpunkt hin zu decodieren. Alle nicht erfassten Wörter gehören zum *Decodierungsausfall*, sind nicht decodierbar und werden als "fehlerhaft" gemeldet.

Durch (10.5)(ii) und (iii) wird die folgende Definition motiviert:

E 1.12 Definition: Fehlerkorrigierender Code

Ein Code $\mathcal{C}$, der die Eigenschaft $(*)$ von E 1.11(iii) besitzt, heißt

t-fehlerkorrigierender Code.

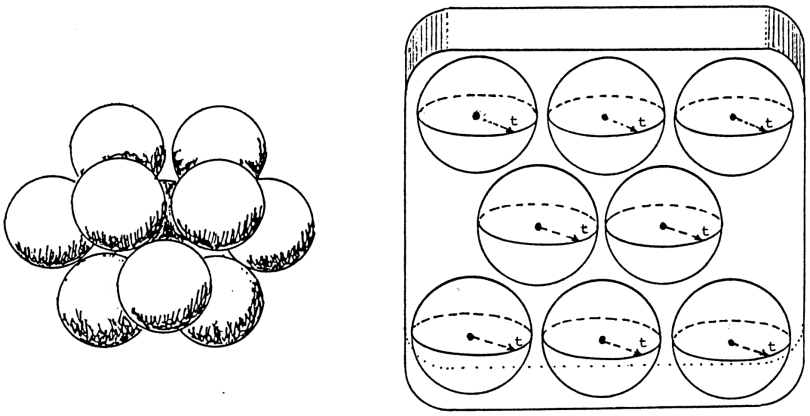
Er hat die Eigenschaft, dass bei ihm fehlerhafte Wörter mit Fehlern in m Komponenten für $m \leq t$ auf die in E 1.11 (iii) angegebene Weise richtig decodiert werden können.
Jeder $t-$ fehlerkorrigierende Code ist also auch $(t-1)-$fehlerkorrigierend, $(t-2)-$fehlerkorrigierend, usw..

Beispiel (Fortsetzung):

$\mathcal{C}_1 = \{000, 111\}$ ist nach E 1.10 ein 1-fehlerkorrigierender Code. Die fehlerhaften Wörter (in Figur E.3 durch • markiert) werden zum Kugelmittelpunkt (mit ■ markiert) korrigiert.

E 1.13 Kugelpackungen

Es ist also in unserem Zusammenhang sinnvoll, nach einer disjunkten Überdeckung von A^n (bzw. der Überdeckung einer Teilmenge von A^n) durch Kugeln vom Radius t zu fragen, nach sogenannten *Kugelpackungen.* Die Kugel-Mittelpunkte können dann als Elemente eines Codes $\mathcal{C}$ gewählt werden, der t-fehlerkorrigierend ist. (Vgl. Figur E.4 !)



a) b)

Bild E.4: Kugelpackung a) modellhaft b) schematisch

Wir behandeln *Schranken* für die Möglichkeit, solche Packungen zu finden. Dazu betrachten wir zunächst die Mächtigkeit einer Kugel und dann die Anzahl der durch die Kugeln abgedeckten Wörter.

E 1.14 Anzahlen

(a) **Kugel-Mächtigkeit**

Für eine Kugel vom Radius t in A^n gilt (mit $\mid A \mid = q$ und $\mathbf{c} \in \mathcal{C}$):

$$\mid K_t(\mathbf{c}) \mid = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$$

(b) **Kugelpackungs-Schranke** (Hamming-Volumenschranke)

Existiert ein t-fehlerkorrigierender Code $\mathcal{C}$ in A^n, so gilt folgende Ungleichung (mit $\mid A \mid = q$):

$$(**) \qquad \sum_{i=0}^{t} \binom{n}{i} (q-1)^i |\mathcal{C}| \quad \leq \quad q^n.$$

Beweis:

(a) Nach Aufgabe E.1 gibt es $\binom{n}{i}(q-1)^i$ Wörter vom Abstand i von $\mathbf{c}$; in $K_t(\mathbf{c})$ liegen alle diese Wörter für $i = 0$ bis t.

(b) Definitionsgemäß sind die $|\mathcal{C}|$ Kugeln vom Radius t um Codewörter disjunkt; es gilt also

$$| \dot{\bigcup_{\boldsymbol{c}\in\mathcal{C}}} K_t(\mathbf{c}) | = \sum_{\boldsymbol{c}\in\mathcal{C}} | K_t(\mathbf{c}) | = | \mathcal{C} | \cdot K_t(\mathbf{c}) =: s \ .$$

Alle betrachteten Wörter liegen in A^n. Es gilt also $s \leq | A^n | = q^n$. □

(c) **Kugelüberdeckung**

In der Ungleichung $(**)$ *gilt die Gleichheit genau dann*, wenn die zum Code gehörige Kugelpackung die Eigenschaft hat, dass jedes Wort von A^n in einer (eindeutig bestimmten) Kugel vom Radius t liegt. Ein Code dieser Eigenschaft heißt *perfekter Code.*

Genauer:

E 1.15 Definition: Perfekter Code

Ein Code $\mathcal{C} \subseteq A^n$ heißt **t–perfekt** (mit $t \in \mathbb{N} \cup \{0\}$), falls gilt:

(i) $K_t(\mathbf{c}_1) \cap K_t(\mathbf{c}_2) = \emptyset$ für alle $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ mit $\mathbf{c}_1 \neq \mathbf{c}_2$ (d.h. $\mathcal{C}$ ist t-fehlerkorrigierend) und

(ii) $A^n = \bigcup_{\boldsymbol{c}\in\mathcal{C}} K_t(\mathbf{c})$ (d.h. die Kugeln vom Radius t um Codewörter überdecken A^n).

Statt $t-$perfekt sagt man auch nur **perfekt.**

E 1.16 Anmerkung zu perfekten Codes

(a) **Beispiele von perfekten Codes:**
Triviale perfekte Codes über dem Alphabet A sind:
• A^n (mit $t = 0$),
• nur aus einem einzigen Codewort der Länge n bestehende Codes (mit $t = n$) und
• im Falle $A = \{0, 1\}$ die Wiederholungscodes $\{00\ldots0, 11\ldots1\}$ der (ungeraden) Länge $2m + 1$ (mit $t = m$).

Wir erwähnen noch folgende perfekten Codes:

- einen binären Code der Länge $n = 23$ mit 2^{12} Wörtern (für $t = 3$), den sogenannten *binären Golay-Code* $\mathcal{G}_{23}$,
- einen ternären Code der Länge $n = 11$ mit 3^6 Wörtern (für $t = 2$), den sogenannten *ternären Golay-Code*,

- eine Serie von Codes $\mathcal{H}_{q,r}$ mit $n = (q^r - 1)/(q - 1), |\mathcal{H}_{q,r}| = q^{n-r}$ (mit $t = 1$) für jede Primzahlpotenz q, die sogenannten *Hamming-Codes*

(b) Perfekte Codes sind deswegen interessant, weil sie den Raum A^n gut ausschöpfen und weil die Decodierung für jedes Wort erklärt ist. Auch in anderem Zusammenhang gibt es Anwendungen; (s. Teil d)!).

(c) Nach weitgehender Vorarbeit von J.H. van Lint (s. auch [1971]) ist es A. Tietäväinen (und unabhängig davon V.A. Zinovjev und K.V. Leontjev) gelungen, die $t-$perfekten Codes über Alphabeten von Primzahlpotenz-Ordnungen zu kennzeichnen, (s. van Lint [1982], Pless [1982]):

Satz über perfekte Codes:
Ist $\mathcal{C}$ ein nicht-trivialer perfekter Code über einem Alphabet A mit $|A| = p^m$, p prim, dann hat $\mathcal{C}$ die gleichen Parameter wie ein Hamming- oder ein Golay-Code.

Für $t > 2, t \neq 6$ ist nach M.R. Best die Voraussetzung " q Primzahlpotenz" überflüssig.

(d) Mit Hilfe eines perfekten Codes über dem Alphabet $\{0, 1, 2\}$ lässt sich das Problem von **Garantiesystemen bei Toto-Wetten** behandeln. Sei der Ausgang von 13 Fußballspielen vorauszusagen; unter Verwendung der Codierung:

$$0 \triangleq \text{Unentschieden}, \quad 1 \triangleq \text{Heimsieg}, \quad 2 \triangleq \text{Heimniederlage}$$

wird das Ergebnis einer Ausspielung zu einem Wort $\mathbf{v}$ aus $\{0, 1, 2\}^{13}$. Von einem Tipp $\mathbf{c}$ mit 12 Richtigen hat $\mathbf{v}$ den Abstand 1; die richtige Tippreihe $\mathbf{v}$ liegt daher in $K_1(\mathbf{c})$. Wie aus Teil (a) zu ersehen, gibt es einen perfekten Code in $\{0, 1, 2\}^{13}$, den *ternären Hamming-Code $\mathcal{H}_{3,3}$ der Länge 13*. Man kann dann (theoretisch) die Mittelpunkte der (den Raum überdeckenden) Kugeln vom Radius 1 um Codewörter als Tippreihen wählen und hat dann mit Sicherheit 12 Richtige zu erwarten. Nach 10.8(c) gibt es jedoch $3^{13}/(1 + 13 \cdot 2) = 3^{10}$ solcher Punkte, für praktische Zwecke eine zu große Anzahl.

Allgemein kann man bei einer Toto-Wette mit n Tippreihen fragen, ob es ein Wettsystem gibt, das $n-t$ Richtige garantiert. Eine Lösung würden t-perfekte ternäre Codes mit Wortlänge n liefern. Für $t = 1$ sind nicht wesentlich andere Zahlen zu erwarten als beim eben behandelten Beispiel. Für $t > 1$ gibt es (nach Tietäväinen und Pless), im Wesentlichen nur einen einzigen nicht-trivialen ternären $t-$perfekten Code, und zwar mit $n = 11$, $|\mathcal{C}| = 3^6 = 729$ und $t = 2$ (d.h. 9 Richtige garantiert), den bereits erwähnten *ternären Golay-Code*. Wegen der zu großen Anzahl von nötigen Tipps und lediglich garantiertem Nebengewinn ist das Verfahren nicht lukrativ. Somit ist das Toto-System noch nicht zusammengebrochen.

Zwar ist die Voraussetzung einer *disjunkten* Überdeckung mit Kugeln gleichen Radius nicht *notwendige* Voraussetzung für ein Garantiesystem. Jedoch ist nicht zu erwarten, dass andere geeignete Überdeckungen mit wesentlich weniger Mittelpunkten auskommen als die der erwähnten perfekten Codes.

(e) Wie bemerkt, ist bei obigen Beispielen die Disjunktheit der Kugeln nicht wesentlich (ausser für deren Anzahl). Wir definieren in diesem Zusammenhang:

Ist $\mathcal{C} \subseteq A$, dann heißt $\rho(\mathcal{C})$ **Überdeckungsradius** von $\mathcal{C}$, falls $\rho(\mathcal{C})$ der kleinste Radius ρ ist mit der Eigenschaft, dass die Kugeln vom Radius ρ um Codewörter die Menge A^n überdecken, also $A^n = \bigcup_{\boldsymbol{c}\in\mathcal{C}} K_\rho(\mathcal{C})$ gilt. Bei einem $t-$perfekten Code $\mathcal{C}$ ist $\rho(\mathcal{C}) = t$, und zusätzlich sind die erwähnten Kugeln disjunkt. Ja es gilt sogar:

$\mathcal{C}$ ist $t-$perfekt genau dann, wenn $\rho(\mathcal{C}) = t$ ist und je zwei Codewörter mindestens den Hamming-Abstand $2t+1$ haben.

Um bei gegebenen Code $\mathcal{C}$ dessen Fehlerkorrektur-Eigenschaften zu bestimmen, ist es nicht nötig, die Kugeln vom Radius t um Codewörter auf Disjunktheit zu untersuchen. Es reicht dazu, den Abstand je zweier Codewörter zu kennen, genauer den minimalen Wert dieses Abstands:

E 1.17 Minimalabstand

(a) *Definition:* Sei $\mathcal{C} \subseteq A^n$; dann ist der Minimalabstand von $\mathcal{C}$ definiert als

$$d_{\min}(\mathcal{C}) := \min_{\boldsymbol{c}_1,\boldsymbol{c}_2\in\mathcal{C},\boldsymbol{c}_1\neq\boldsymbol{c}_2} d_H(\mathbf{c}_1,\mathbf{c}_2)\ .$$

(b) *Eigenschaften:*

$\mathcal{C} \subseteq A^n$ ist t-fehlerkorrigierend genau dann, wenn gilt: $\quad d_{\min}(\mathcal{C}) \geq 2t+1$.

Beweis: "$\Rightarrow$" Sei $d_{\min}(\mathcal{C}) =: s \leq 2t$. Dann existieren $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ mit $d_H(\mathbf{c}_1,\mathbf{c}_2) = s$. Ist $s < t$, so gilt $\mathbf{c}_2 \in K_t(\mathbf{c}_1)$. Ist $s \geq t$, so kann man bei $\mathbf{c}_1$ unter den s Komponenten, in denen sich $\mathbf{c}_1$ und $\mathbf{c}_2$ unterscheiden, t Komponenten in solche von $\mathbf{c}_2$ umändern. Es entsteht ein Wort $\mathbf{z}$ mit $d_H(\mathbf{c}_1,\mathbf{z}) = t$ und $d_H(\mathbf{c}_2,\mathbf{z}) = s-t$. Wegen $s-t \leq t$ folgt $K_t(\mathbf{c}_1) \cap K_t(\mathbf{c}_2) \neq \emptyset$. Damit ist $\mathcal{C}$ nicht $t-$fehlerkorrigierend (s. 10.6).

"$\Leftarrow$" Ist $\mathcal{C}$ nicht t-fehlerkorrigierend, so existieren $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ und $\mathbf{v} \in A^n$ mit $\mathbf{v} \in K_t(\mathbf{c}_1) \cap K_t(\mathbf{c}_2)$. Aus der Dreiecksungleichung (10.1b) ergibt sich

$$d_{\min}(\mathcal{C}) \leq d_H(\mathbf{c}_1,\mathbf{c}_2) \leq d_H(\mathbf{c}_1,\mathbf{v}) + d_H(\mathbf{v},\mathbf{c}_2) \leq t+t. \qquad \square$$

(c) **Beispiel:** Beim Wiederholungscode $\mathcal{C}_1 = \{000, 111\}$ aus Beispiel E 1.10 ist trivialerweise $d_{\min}$ gleich 3, in Übereinstimmung damit, dass $\mathcal{C}_1$ 1-fehlerkorrigierend ist.

Einige Symbole und Bezeichnungen

$\forall x \in X \exists 1\ y \in Y : A(x,y)$	Für alle x aus X existiert genau ein y aus Y derart, dass A(x,y) gilt
$\Rightarrow$	daraus folgt
$\Leftrightarrow$	dies ist äquivalent zu
$\neg A$	Non A
$X \subseteq Y \subsetneqq Z$	X ist Teilmenge von Y und Y ist echte Teilmenge von Z
$\wp(M)$	Potenzmenge von M
$f : \begin{cases} X \to Y \\ x \mapsto f(x) \end{cases}$	Funktion f mit Definitionsbereich X, Wertebereich Y, Zuordnungsvorschrift $x \mapsto f(x)$
$f(X) = \{f(x) \mid x \in X\}$	Bild von X unter f
$G_f := \{(x, f(x)) \mid x \in X\}$	Graph von f
$f\|_Z := \mathrm{Rest}_Z f$; $f\|_{Z \to W}$	Einschränkung von f auf Definitionsbereich Z bzw. Wertebereich W
id_M	Identische Abbildung auf M
$\mathrm{Abb}(I,K)$ oder K^I	Menge aller Abbildungen von I in K (aller Familien aus K mit Indexmenge I)
$K^{(I)}$	Menge aller Familien aus K mit endlichem Träger
$\mathrm{Bij}(X,Y)$	Menge aller Bijektionen von X auf Y
$\mathcal{B}(E,\mathbb{R})$	Menge aller beschränkten reellen Funktionen auf E
$\mathcal{C}(E,\mathbb{R})$	Menge aller stetigen reellen Funktionen auf E
$\mathbb{N}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$	Zahlbereiche
$p^n \parallel m \Leftrightarrow (p^n \mid m \wedge p^{n+1} \not\mid m)$	
$\mathbb{H}$	reelle Quaternionenalgebra
$\mathbb{Z}_m := \mathbb{Z}/(m) := \mathbb{Z}/m\mathbb{Z}$	
Körper	kommutativer Schiefkörper
$\mathrm{GF}(q) = \mathbb{F}_q$	Galoisfeld mit q Elementen

$\mathrm{Aut}(G)$	volle Automorphismengruppe von G
$\mathrm{End}(G)$	Endomorphismenring der abelschen Gruppe
$\mathcal{S}_n$	symmetrische Gruppe auf n Elementen
$\mathcal{A}_n$	alternierende Gruppe auf n Elementen
$\mathcal{P}(K)$	Algebra der Polynomabbildungen über K
$K[x]$	Polynomalgebra über K
$V(d, K)$	Vektorraum der Dimension d über K
$\dim_K V$	Dimension des K-Vektorraums V
$V/\operatorname{Kern} f \cong \operatorname{Bild} f$	
$\mathrm{Hom}_K(V, W)$	Vektorraum der linearen Abbildungen von V in W (als K-Vektorräume)
$\mathrm{End}_K V = \mathrm{Hom}_K(V, V)$	
$K^{(n,m)}$	Vektorraum der $n \times m$-Matrizen über K
$\Gamma\mathrm{L}(V)$	Gruppe der n.s. semilinearen Abbildungen von V auf sich
$\mathrm{GL}(V) = \mathrm{Aut}_K V$	Gruppe der n.s. linearen Abbildungen von V auf sich
$\mathrm{SL}(V)$	Gruppe der nicht-singulären linearen Abbildungen von V auf sich mit det 1
$\mathrm{AG}(V)$ bzw. $\mathrm{AG}(d, K)$	affiner Raum über V bzw. über $V(d, K)$
V^* bzw. V^{d}	algebraischer Dualraum von V
δ_{ij}	$\delta_{ij} = 1$ für $i = j$ und $\delta_{ij} = 0$ für $i \neq j$ (Kronecker-Symbol)

Das Alphabet

A	$\mathfrak{a}$	$\mathfrak{A}$
B	$\mathfrak{b}$	$\mathfrak{B}$
C	$\mathfrak{c}$	$\mathfrak{C}$
D	$\mathfrak{d}$	$\mathfrak{D}$
E	$\mathfrak{e}$	$\mathfrak{E}$
F	$\mathfrak{f}$	$\mathfrak{F}$
G	$\mathfrak{g}$	$\mathfrak{G}$
H	$\mathfrak{h}$	$\mathfrak{H}$
I	$\mathfrak{i}$	$\mathfrak{I}$
J	$\mathfrak{j}$	$\mathfrak{J}$
K	$\mathfrak{k}$	$\mathfrak{K}$
L	$\mathfrak{l}$	$\mathfrak{L}$
M	$\mathfrak{m}$	$\mathfrak{M}$
N	$\mathfrak{n}$	$\mathfrak{N}$
O	$\mathfrak{o}$	$\mathfrak{O}$
P	$\mathfrak{p}$	$\mathfrak{P}$
Q	$\mathfrak{q}$	$\mathfrak{Q}$
R	$\mathfrak{r}$	$\mathfrak{R}$
S	$\mathfrak{s}$	$\mathfrak{S}$
T	$\mathfrak{t}$	$\mathfrak{T}$
U	$\mathfrak{u}$	$\mathfrak{U}$
V	$\mathfrak{v}$	$\mathfrak{V}$
W	$\mathfrak{w}$	$\mathfrak{W}$
X	$\mathfrak{x}$	$\mathfrak{X}$
Y	$\mathfrak{y}$	$\mathfrak{Y}$
Z	$\mathfrak{z}$	$\mathfrak{Z}$

alpha	α	A	xi	ξ	Ξ
beta	β	B	omikron	o	O
gamma	γ	Γ	pi	π	Π
delta	δ	Δ	rho	ρ	R
epsilon	ε	E	sigma	σ	Σ
zeta	ζ	Z	tau	τ	T
eta	η	H	ypsilon	υ	Υ
theta	ϑ	Θ	phi	φ	Φ
jota	ι	I	chi	χ	X
kappa	κ	K	psi	ψ	Ψ
lambda	λ	Λ	omega	ω	Ω
my	μ	M			
ny	ν	N	aleph	$\aleph$	(hebräisches a)

Literatur-Auswahl zu „Lineare Algebra“

[Fis17] Gerd Fischer: *Lernbuch Lineare Algebra und Analytische Geometrie.* Das Wichtigste ausführlich für das Lehramts- und Bachelorstudium. Springer Spektrum, 3.Auflage 2017.

[Fis05] Gerd Fischer. *Lineare Algebra:* Eine Einführung für Studienanfänger (Grundkurs Mathematik) Springer Spektrum 18.Auflage 2014.

[StoGRi17] Stoppel, Hannes, Griese, Birgit: *Übungsbuch zur Linearen Algebra.* Aufgaben und Lösungen. (Übungsbuch zu Fischer: Lineare Algebra). Springer Spektrum 9.Auflage 2017.

[AblHer13] Christoph Ableitinger, Angela Herrmann: *Lernen an Musterlösungen zur Analysis und Linearen Algebra.* Springer Spektrum 2011, 2013.

[ABJSSW03] Alpers, Bielig-Schulz, Th.Jahnke, Janßen, Siekmann, Simanovsky, Wuttke: *Mathematik. Analytische Geometrie, Lineare Algebra,* (Schulbuch) Cornelsen Verlag 2003

[Beut14] Albrecht Beutelspacher. *Lineare Algebra.* Springer Spektrum 8. Auflage 2014

[DeiLa15] O. Deiser, C.Lasser *Erste Hilfe in Linearer Algebra:Überblick und Grundwissen mit vielen Abbildungen und Beispielen.* Springer Spektrum, 2015.

[Gro69] Karl Peter Grotemeyer. *Analytische Geometrie.* Sammlung Göschen. Walter de Gruyter 1969

[HupWill10] Bertram Huppert,Wolfgang Willems :*Lineare Algebra.* Mit zahlreichen Anwendungen in Kryptographie, Codierungstheorie, Mathematischer Physik und Stochastischen Prozessen. 2.Auflage Springer Vieweg 2010

[ScheiSchw17] Harald Scheid, Wolfgang Schwarz *Elemente der Linearen Algebra und der Analysis.* Spektrum Akademischer Verlag Heidelberg 2009

Weitere (relativ neue) deutsche Bücher zur Linearen Algebra/Analytischen Geometrie bzw. Übungsbücher zur Linearen Algebra
(ohne Anspruch auf Vollständigkeit):
Christian Bär 2018, Siegfried Bosch 2014, Thomas Epp/Rolf Busam: Prüfungstrainer LA, 2009/2011, Carsten Gellrich/Regina Gellrich 2016, Laurenz Göllann 2017, Günter M.Gramlich 2014, Ernst Georg Haffner 2018, Hans Havlicek 2006, Klaus Jänich 2013, Daniel Jung, E.-G.Haffner 2012, Theo de Jong 2020, Christian Karpfinger/Hellmuth Stachel 2020, Peter Knabner/Wolf Barth 2013/2018, Gregor Kemper/Fabian Reimers 2022, Hans-Joachim Kowalsky/Gerhard O.Michler 2003, Burkhard Lenze 2.Aufl.2020, Jörg Liesen/Volker Mehrmann 2015, Thomas C.T. Michaels/Marcel Liechti: Prüfungstraining LA, 2022, Florian Modler/Martin Kreh 4.Aufl.2018, Thoralf Räsch 2015; Michael Ruhrländer 2017, Schichl-Steinbauer 2018 (Kap.7), Gilbert Strang (übersetzt aus den Englischen) 2003, Stefan Waldmann 2017, Guido Walz 2019, Dirk Werner 2022

Ältere Literaturliste:

Literatur-Auswahl zu „Lineare Algebra“

[Aig74a] Martin Aigner. *Lineare Algebra I.* Skript, FU Berlin FB Mathematik, 1974.

[Aig74b] Martin Aigner. *Lineare Algebra II.* Skript, FU Berlin FB Mathematik, 1974.

[Art60a] E Artin. *Analytische Geometrie und Algebra I.* Vorlesungsausarbeitung, Universität Hamburg, 1960.

[Art60b] E. Artin. *Analytische Geometrie und Algebra II.* Vorlesungsausarbeitung, Universität Hamburg, 1960.

[Art60c] E Artmann. *Lineare Algebra.* Birkäuser Verlag, Basel, 1991^3.

[Beut01] Albrecht Beutelspacher und Marc-Alexander Zschiegner. *Lineare Algebra interaktiv.* Vieweg Verlag, 2001. CD-ROM.

[Bos06] Siegfried Bosch. *Lineare Algebra.* Springer Verlag, 2006^3.

[Bou62] N. Bourbarki. *Eléments de mathématique.* Paris, 1962.

[Bri83] E Brieskorn. *Lineare Algebra und analytische Geometrie.* Vieweg, Braunschweig etc., 1983.

[Bro04] Theodor Bröcker. *Lineare Algebra und Analytische Geometrie.* Birkhäuser Verlag, 2004^2.

[DHK72] K. Doerk, B. Huppert, and Kroll. *Lineare Algebra I.* FB Mathematik, Universiät Mainz, 1972.

[Edw95] H.M. Edwards. *Linear Algebra.* Birkhäuser Verlag, 1995 (engl.).

[End84] K. Endl. *Analytische Geometrie und Algebra.* Würfel Verlag, Gießen, 1984.

[Fle72] T. J. Fletcher. *Linear Algebra through its applications.* London, 1972.

[Gra99] Hans Grauert und Hans-Christoph Grunau. *Lineare Algebra und Analytische Geometrie.* Oldenbourg Verlag, München, 1999.

[Gre67] W. H. Greub. *Linear algebra.* Berlin, 1967.

[Gro70] Karl Peter Grotemeyer. *Lineare Algebra.* Number 732. BI, Mannheim, 1970.

[Gru72a] S. Gruber. *Lineare Algebra und Analytische Geometrie I.* Erlangen, 1972.

[Gru72b] S. Gruber. *Lineare Algebra und Analytische Geometrie II.* Erlangen, 1972.

[Hal74] P.R. Halmos. *Finite-Dimensional Vector Spaces.* Springer Verlag, New York , 1974 (engl.).

[Hav06] Hans Havlicek. *Lineare Algebra für Technische Mathematiker.* Heldermann Verlag, 2006.

[Hof61] R. Hoffman, K.and Kunze. *Linear Algebra*. Englewood Cliffs, 1961.

[HR73] J Heinhold and B Riedmüller. *Lineare Algebra und Analytische Geometrie*. Müchen, 1973.

[Jac53] N. Jacobson. *Lecture in Abstract Algebra II*. Toronto etc., 1953.

[Jän04] K. Jänisch. *Lineare Algebra*. Springer Verlag, Berlin etc., 2004^{10}.

[Ker05] Ina Kersten. *Analytische Geometrie und Lineare Algebra I*. Univ. Verlag, Göttingen. 2005.

[KK71] W. Klingenberg and P. Klein. *Lineare Algebra und Analytische Geometrie*. Number 748. BI, Mannheim, 1971.

[Koe83] M. Koecher. *Lineare Algebra und Analytische Geometrie*. Springer Verlag, Berlin etc., 1983.

[Kow03] Hans-Joachim Kowalsky und Gerhard O. Michler. *Lineare Algebra*. De Gruyter Verlag, 2003^{12}.

[Lan70] Serge Lang. *Linear Algebra*. Addison-Wesley, Reading, 1970 (engl.).

[Lan72] Serge Lang. *Introduction to Linear Algebra*. Springer, 1986 (engl.)

[Leh83] E. Lehmann. *Lineare Algebra mit dem Computer*. Teubner Verlag, Stuttgart, 1983.

[LM15] J. Liesen, V. Mehrmann *Lineare Algebra:*. Springer Spektrum, 2015.

[Lin69] Lingenberg. *Lineare Algebra*. Number 828/828a. BI, Mannheim, 1969.

[Lip77] S. Lipschutz. *Theory and problems of linear algebra*. Schaum's Überblick/Aufgaben. McGraw Hill, 1977.

[Lor82a] F. Lorenz. *Lineare Algebra I*. Number 605. BI, Mannheim, 1982.

[Lor82b] F. Lorenz. *Lineare Algebra II*. Number 605. BI, Mannheim, 1982.

[lR80] H. le Roy. *Prinzipien der Linearen Algebra*. Number 984 in UTB. Bern, 1980.

[Lün73] H. Lüneburg. *Einfürung in die Algebra*. Springer, Berlin, 1973.

[Mir65] L. Mirsky. An introduction to linear algebra. London, 1965.

[Sch76] H. Schaal. *Lineare Algebra und Analytische Geometrie I*. Vieweg, Braunschweig, 1976.

[Sch80] H. Schaal. *Lineare Algebra und Analytische Geometrie II*. Vieweg, Braunschweig, 1980^{2}.

[Sch81] H. Schaal. *Lineare Algebra und Analytische Geometrie III*. Vieweg, Braunschweig, 1981.

[SSch09] H. Scheid, W. Schwarz *Elemente der Linearen Algebra und der Analysis*. Spektrum Akademischer Verlag. Heidelberg 2009.

[Str08] K. Strambach, H. Vöelklein, R. Straszewski. *Lineare Algebra.* Vieweg, Braunschweig, 1981.

[Str76] G. Strong. *Linear Algebra and its applications.* Academic Press, 1976.

[Tie73] H. Tietze. *Lineare Geometrie.* Number 248 in UTB. Göttingen, 1973.

[Tis] G. Tischel *Introduction to Lineare Algebra. Aufgaben, Lösungswege und Lösungen.* Diesterweg-Salle u. Verlag Sauerländer, 1975.

[Wag81] R. Wagner. *Grundzüge der linearen Algebra.* Stuttgart, 1981.

[Wal17] Stefan Waldmann *Lineare Algebra I.* Springer Spektrum 2017.

Weiterführende oder ergänzende Literatur

[Bae52] R. Baer. *Linear Algebra and projective geometry.* Academic Press, New York, 1952.

[Bri76] F. Brickell. *Matrizen und Vektorräume.* Verlag Chemie, Weinheim, 1976.

[Dre76] K. D. Drews. *Lineare Gleichungssysteme und lineare Optimierung.* Number 520. UTB, 1976.

[Fis78] Gerd Fischer. *Analytische Geometrie.* Vieweg, Braunschweig, 1978.

[Her64] J. Herstein. Topics in algebra. Waltham, 1964.

[Jac74] N. Jacobson. *Basic Algebra I.* San Francisco, 1974.

[Lan67] Serge Lang. *Algebra Structures.* Addison-Wesley, Reading, 1967.

[Len75] H. Lenz. Grundlagen der Elementarmathematik. Berlin, 1975.

[LP82] R Lidl and G Pilz. *Angewandte abstrakte Algebra.* BI, Zürich, 1982.

[Pic67] G. Pickert. Analytische Geometrie. Leipzig, 1967.

[Spe61] E. Sperner. *Einführung in die Analytische Geometrie und Algebra.* Vandenhoeck & Ruprecht, Göttingen, 1951,1961.

[SS81] G Scheja and U Storch. *Lehrbuch der Algebra*, volume 3. Teubner Verlag, Stuttgart, 1981.