1 Kapitel: Grundlagen

1.1 Mengen und Abbildungen

1.1.1 Die folgende Definition stammt von Georg Cantor, dem Begründer der Mengenlehre:

Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen.

1.1.2 Beispiele: $\mathbb{N} = \{1, 2, 3, 4, \ldots\} :=$ die Menge der natürlichen Zahlen.
$\mathbb{N}_0 := \{0, 1, 2, \ldots\}$.
$\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, \ldots\} =$ die Menge der ganzen Zahlen.
$\mathbb{Q} := \{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\}$ die Menge der rationale Zahlen.
$\mathbb{R} :=$ die Menge der reellen Zahlen.
$\emptyset =$ leere Menge.

Hier bedeutet $:=$ "Gleichheit per Definition".

Sei M eine Menge. Gehört ein Objekt x zu der Menge M (d.h. ist x ein Element von M) so schreiben wir

$$x \in M$$

($x \in M$ ist also eine abkürzende Schreibweise für die Aussage "x ist ein Element von M").

$x \notin M$ bedeutet: x ist kein Element von M.

Zum Beispiel gilt $-2 \in \mathbb{Z}$ aber $-2 \notin \mathbb{N}$.

Man kann Mengen dadurch beschreiben, dass man ihre Elemente zwischen zwei geschweifte Klammern schreibt. Dieses Hinschreiben kann auf verschiedene Weise geschehen. Beispielsweise bezeichnet $\{1, 2, 3\}$ die Menge die aus den Zahlen $1, 2$ und 3 besteht. Auf die Reihenfolge kommt es dabei nicht an

$$\{1,2,3\} = \{3,1,2\} = \{3,3,1,2,1\}.$$

Eine andere Möglichkeit besteht darin, Elemente die man nicht nennt durch Punkte anzudeuten (sofern klar ist was damit gemeint ist). Z.B. bezeichnet $\{1, 2, \ldots, 10\}$ bezeichnet die Menge $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

Schliesslich kann man Mengen beschreiben in dem man nach der Klammer $\{$ zun"achst einen Buchstaben, der die Elemente der Menge bezeichnen soll, hinschreibt, dann einen senkrechten Strich macht und dahinter genau hinschreibt welches die Elemente sind. Z.B. kann man statt $\{1, 2, 3\}$ auch $\{x \mid x \text{ ist ganze Zahl und } 1 \le x \le 3\}$ oder $\{x \in \mathbb{N} \mid x \le 3\}$ schreiben.

Seien A und B Mengen. $A \subseteq B$ bedeutet "A ist Teilmenge von B", d.h. jedes Element von A ist auch ein Element von B (d.h. für alle x gilt $x \in A \Rightarrow x \in B$).

1.1.3 Beispiele: $\emptyset \subseteq M$, $M \subseteq M$.
$\emptyset \subseteq \{1,2,3\} \subseteq \{1, 2, \ldots, 10\} \subseteq \mathbb{N} \subseteq \mathbb{Z} \subset \mathbb{Q} \subseteq \mathbb{R}$.

Es gilt

$$A = B \iff (A \subseteq B) \wedge (B \subseteq A)$$

wobei $\wedge$ das mathematische Symbol für "und" ist. Wir schreiben $A \not\subseteq B$ wenn A ist keine Teilmenge von B und $A \subsetneqq B$, wenn A eine echte Teilmenge von B ist, d.h. $A \subseteq B$ und $A \neq B$.

1.1.4 Definition: Es seien A und B Mengen. Als den Durchschnitt $A \cap B$ bezeichnet man die Menge der Elemente, die sowohl in A als auch in B enthalten sind.

$$A \cap B = \{x \mid x \in A \text{ und } x \in B\}$$

Die Vereinigung $A \cup B$ ("A vereinigt B") ist die Menge der Elemente, die in A oder in B (oder in beiden) enthalten sind.

$$A \cup B = \{x \mid x \in A \text{ oder } x \in B\}$$

$A \setminus B = \{x \in A \mid x \notin B\}$ heißt Differenzmenge oder auch das Komplement von B in A.

Als nächstes wird erläutert was mit dem Produkt von Mengen gemeint ist.

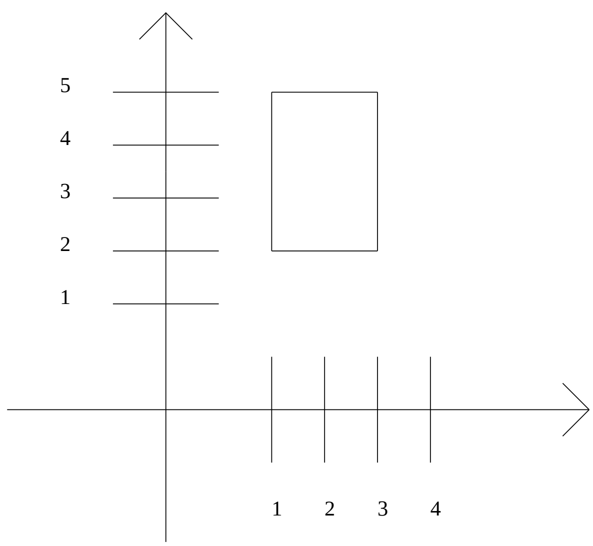
Ein *Paar* besteht in der Angabe eines ersten und eines zweiten Elements. Bezeichnet a das erste und b das zweite Element, so wird das Paar mit (a, b) bezeichnet. $(a, b) = (a', b')$ bedeutet also $a = a'$, $b = b'$.

Der Unterschied zwischen (a, b) und $\{a, b\}$ ist folgender: Beim Paar kommt es auf die Reihenfolge an, also $(a, b) \neq (b, a)$ (falls $a \neq b$) während $\{a, b\} = \{b, a\}$.

1.1.5 Definition: Sind A und B Mengen so heißt die Menge $A \times B = \{(a, b) \mid a \in A \text{ und } b \in B\}$ das (kartesische) Produkt der Mengen A und B.

Zur Veranschaulichung betrachten wir den Fall
$A = [1, 3] = \{x \in \mathbb{R} \mid 1 \leq x \leq 3\}$
$B = [2, 5] = \{y \in \mathbb{R} \mid 2 \leq y \leq 5\}$

$A \times B$ ist die Punkte des Rechtecks in dem folgenden Bild:



Analog zur Definition der Paare kann man auch *Triple* (a, b, c) oder allgemeiner *n-Tupel* $(a_1, \ldots, a_n)$ erklären (wobei n eine beliebige natürliche Zahl ist).

Sind $A_1, \ldots, A_n$ Mengen, so heißt die Menge

$$A_1 \times \ldots \times A_n = \{(a_1, \ldots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\}$$

das (kartesische) Produkt der Mengen $A_1, \ldots, A_n$. Anstatt

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

schreiben wir auch $\mathbb{R}^n$. $\mathbb{R}^2$ ist also die Menge der Punkte der (Koordinate-) Ebene.

Als nächstes möchte ich erklären was Abbildungen zwischen zwei Mengen sind. Dies ist eine Verallgemeinerung des Funktionsbegriffs.

1.1.6 Definition: Seien X und Y Mengen.
Eine Abbildung (oder auch Funktion genannt) f von X nach Y ist eine Vorschrift, die jedem $x \in X$ genau ein Element $f(x) \in Y$ zuordnet. Dabei heißt X der Definitionsbereich von f.
Statt "f ist eine Abbildung von X nach Y" schreibt man kurz

$$f : X \to Y, x \mapsto f(x).$$

1.1.7 Beispiele: (a) Sei $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$. Die Wurzelfunktion $w : \mathbb{R}_+ \to \mathbb{R}$ ist gegeben durch $x \mapsto \sqrt{x}$.

(b) $f : \mathbb{Z} \to \mathbb{N}_0, x \mapsto x^2$.

Man kann eine Abbildung $f : X \to Y$ auch mengentheoretisch verstehen als die Teilmenge

$$\Gamma_f = \{(x, f(x)) \mid x \in X\}$$

von $X \times Y$. Γ_f heißt Graph von f. Zwei Abbildungen $f_1, f_2 : X \to Y$ sind genau dann gleich, wenn die Mengen Γ_{f_1} und Γ_{f_2} übereinstimmen.

1.1.8 Definition: Sei $f : X \to Y$ eine Abbildung und $A \subseteq X$ und $B \subseteq Y$ Teilmengen. Dann heißt die Teilmenge

$$f(A) = \{f(x) \mid x \in A\}$$

die Bildmenge von A, oder das "Bild von A" und die Menge

$$f^{-1}(B) = \{x \mid f(x) \in B\}$$

heißt die Urbildmenge von B, oder einfach das "Urbild von B".

1.1.9 Definition: (a) Eine Abbildung $f : X \to Y$ heißt injektiv, wenn keine zwei Elemente von X auf dasselbe Element von Y abgebildet werden.

b) Eine Abbildung $f : X \to Y$ heißt surjektiv, wenn es für jedes $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$. Man sagt dann auch, daß f eine Abbildung von X *auf* Y ist.

c) Eine Abbildung $f : X \to Y$ heißt bijektiv, wenn sie injektiv und surjektiv ist.

1.1.10 Beispiel: Betrachte $q : \mathbb{R} \to \mathbb{R}$ mit $q(x) = x^2$. Wegen $q(\mathbb{R}) = \mathbb{R}_+ = \{y \in \mathbb{R} \mid y \geq 0\}$ ist q nicht surjektiv und wegen $q(-1) = q(1) = 1$ ist q auch nicht injektiv. Betrachten wir aber q als Funktion von $\mathbb{R}_+$ nach $\mathbb{R}_+$ (also $q : \mathbb{R}_+ \to \mathbb{R}_+, x \mapsto x^2$) so ist q sogar bijektiv.

Ist $y \in Y$ so definieren wir das Urbild von y unter f durch $\{x \in X \mid f(x) = y\}$ $(= f^{-1}(\{y\}))$. Dann gilt:

> f ist injektiv $\iff$ Das Urbild jedes $y \in Y$ hat höchstens ein Element,
>
> f ist surjektiv $\iff$ Das Urbild jedes $y \in Y$ hat mindestens ein Element,
>
> f ist bijektiv $\iff$ Das Urbild jedes $y \in Y$ hat genau ein Element.

1.1.11 Definition: Ist $f : X \to Y$ bijektiv, so heißt

$$\begin{aligned} f^{-1} : Y &\to X \\ y = f(x) &\mapsto x \end{aligned}$$

die Umkehrabbildung von f. Man bezeichnet f^{-1} auch als "f hoch minus 1" oder als "f invers".

Ist $f : X \to Y$ bijektiv und $f^{-1} : Y \to X$ die Umkehrabbildung, so gibt es einen Zusammenhang zwischen der Abbildung f^{-1} und dem Urbild $f^{-1}(B)$ einer Teilmenge $B \subseteq Y$. Es ist nämlich $f^{-1}(B)$ gleich dem Bild von B unter der Abbildung $f^{-1} : Y \to X$. Im allgemeinen (d.h. wenn f nicht bijektiv ist) hat $f^{-1}(B)$ aber nichts mit der Umkehrabbildung zu tun. Diese existiert nämlich nur für bijektives f während man das Urbild $f^{-1}(B)$ von B immer bilden kann.

Als nächstes erläutere ich was mit der Komposition von Abbildungen gemeint ist.

1.1.12 Definition: Sind $f : X \to Y$ und $g : Y \to Z$ Abbildungen so sei die zusammengesetzte Abbildung $g \circ f$ durch

$$X \longrightarrow Z, x \mapsto g(f(x))$$

definiert. Die Abbildung $g \circ f$ heißt die Komposition von f mit g oder einfach "f komponiert g" oder "g nach f".

Optisch einprägsamer kann man sich dies als Diagramm merken:

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & Y \arrow[r, "g"] & Z \end{tikzcd}$$

1.1.13 Beispiele: (a) Sei $f(x) = |x|$, $f : \mathbb{R} \to \mathbb{R}_+$ und $g(x) = \sqrt{x}$, $g : \mathbb{R}_+ \to \mathbb{R}$. Dann gilt:

$$(g \circ f)(x) = \sqrt{|x|} : \mathbb{R} \to \mathbb{R}$$
$$(f \circ g)(x) = f(\sqrt{x}) = \sqrt{x} : \mathbb{R}_+ \to \mathbb{R}_+$$

(b) Sei $f : X \to Y$ bijektiv mit Umkehrabbildung $f^{-1} : Y \to X$. Dann gilt:

$$f^{-1} \circ f = Id_X \text{ und } f \circ f^{-1} = Id_Y$$

Für eine Menge M bezeichnet dabei Id_M die Abbildung

$$\begin{aligned} Id_M : M &\to M \\ x &\mapsto x. \end{aligned}$$

Sie heißt die *Identität* auf M.

Zum Abschluss diese Paragraphen führe ich noch eine weitere Bezeichnung ein:

1.1.14 Definition: Ist $f : X \to Y$ bijektiv und $A \subseteq X$ eine Teilmenge von X so heißt

$$f\ |_A : A \to Y, a \mapsto f(a)$$

die Einschränkung von f auf A.

1.2 Reelle Zahlen

1.2.1 Set G eine Menge und $\circ : G \times G \to G$ eine Abbildung (eine solche Abbildung wird auch *Verknüpfung* genannt). Für $x, y \in G$ schreiben wir $x \circ y$ anstatt $\circ((x, y))$. Die Verknüpfung heißt *assoziativ*, falls für alle $x, y, z \in G$ gilt:

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

Sie heißt kommutativ, falls

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

für alle $x, y, z \in G$.

1.2.2 Beispiel: $(\mathbb{R}, +), (\mathbb{R}, \cdot)$ sind assoziativ und kommutativ. $(\mathbb{R}, -), (\mathbb{R} \setminus \{0\}, \div)$ sind weder assoziativ noch kommutativ.

Ein Element $e \in G$ heißt neutrales Element (bzgl. $\circ$) wenn $x \circ e = e \circ x = x$ für alle $x \in G$.

1.2.3 Lemma: Sind $e, e' \in G$ neutrale Elemente (bzgl. $\circ$), so gilt $e = e'$, d.h. das neutrale Element – sofern es existiert – ist eindeutig bestimmt.

Beweis: $e = e \circ e' = e'$ $\square$

Sei $(G, \circ, e,)$ ein Tripel bestehend aus einer Menge G, einer Verknüpfung $\circ : G \times G \to G$ und neutralem Element $e \in G$. Ist $x \in G$ so heißt $y \in G$ linksinverses (bzw. rechtsinverses) Element zu x falls $y \circ x = e$ (bzw. $x \circ y = e$).

1.2.4 Lemma: Sei $\circ : G \times G \to G$ eine assoziative Verknüpfung mit neutralem Element $e \in G$. Ist $y \in G$ linksinverses und $z \in G$ rechtsinverses Element zu $x \in G$, so gilt $y = z$. Insbesondere ist $x^{-1} := y = z$ eindeutig bestimmt. Es heißt das zu x inverse Element.

Beweis: $y \circ e = y \circ (x \circ z) = (y \circ x) \circ z = e \circ z = z$. $\square$

Wegen $e \circ e = e$ gilt: $e^{-1} = e$.

1.2.5 Definition: Ein Paar $(G, \circ)$ bestehend aus einer Menge G und einer Verknüpfung $\circ : G \times G \to G$ heißt eine Gruppe, wenn gilt:

(G1) $\circ$ ist assoziativ,

(G2) es gibt ein neutrales Element $e \in G$ mit $g \circ e = e \circ g = g$ für alle $g \in G$,

(G3) zu jedem $g \in G$ existiert ein inverses $g^{-1} \in G$ mit $g \circ g^{-1} = g^{-1} \circ g = e$.

Ist die Verknüpfung $\circ$ zusätzlich kommutativ, so heißt G abelsche Gruppe.

1.2.6 Beispiele: (a) $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +)$ und $(\mathbb{Q} \setminus \{0\}, 1)$ sind abelsche Gruppen.

(b) Sei M eine Menge $\neq \emptyset$ und $S(M)$ die Menge aller bijektiven Abbildungen $f : M \to M$ von M in sich. Dann ist $S(M)$ versehen mit der Komposition $\circ$ eine Gruppe. $(S(M), \circ)$ ist nicht abelsch falls M mindestens drei verschiedene Elemente besitzt.

(c) Sei $GL_2(\mathbb{R})$ die Menge der invertierbaren reellen 2×2-Matrizen. $GL_2(\mathbb{R})$ zusammen mit der Matrizenmultiplikation ist eine Gruppe. Sie ist nicht abelsch.

Erinnerung: Eine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ heißt invertierbar, falls es eine Matrix $B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ gibt mit $A \cdot B = B \cdot A = E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

1.2.7 Lemma: Sei $(G, \circ)$ eine Gruppe.

(i) Man kann in G "kürzen", d.h. sind $a, b, c \in G$ so gilt:

$$\begin{aligned} c \circ a = c \circ b &\Rightarrow a = b, \\ a \circ c = b \circ c &\Rightarrow a = b. \end{aligned}$$

(ii) Seien $a, b \in G$, so sind die Gleichungen $a \circ x = b$ und $y \circ a = b$ eindeutig durch ein $x \in G$ bzw. $y \in G$ lösbar.

Beweis: Zu (i):

$$\begin{aligned} c \circ a = c \circ b \Rightarrow & \\ &\Rightarrow c^{-1} \circ (c \circ a) = c^{-1} \circ (c \circ b) \\ &\Rightarrow (c^{-1} \circ c) \circ a = (c^{-1} \circ c) \circ b \\ &\Rightarrow e \circ a = e \circ b \\ &\Rightarrow a = b. \end{aligned}$$

Analog beweist man die zweite Aussage.

Zu (ii):

$$\begin{aligned} a \circ x = b &\iff a^{-1} \circ (a \circ x) = a^{-1} \circ b \\ &\iff (a^{-1} \circ a) \circ x = a^{-1} \circ b \\ &\iff e \circ x = a^{-1} \circ b \\ &\iff x = a^{-1} \circ b \end{aligned}$$

$\square$

1.2.8 Auf der Menge der reellen Zahlen $\mathbb{R}$ gibt es die beiden Verknüpfungen $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ und $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Für sie gelten:

(K1) $(\mathbb{R}, +)$ ist eine abelsche Gruppe mit neutralem Element $0 \in \mathbb{R}$ und zu $x \in \mathbb{R}$ inversem Element $-x$.

(K2) $(\mathbb{R} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1 und zu $x \neq 0$ inversem Element x^{-1} oder $\frac{1}{x}$.

(K3) $x(y + z) = xy + xz \quad \forall x, y, z \in \mathbb{R}$
Dies ist das *Distributivgesetz.*

Dabei ist "$\forall$" eine abkürzende Schreibweise für den Ausdruck "für alle" ($\forall$ wird als Allquantor bezeichnet).

Der Ausdruck $\exists$ bedeutet "es gibt" (Existenzquantor). Der Ausdruck "$\exists x \in \mathbb{R} : x^2 = 2$" ist also eine abkürzende Schreibweise für die Aussage:

"Es gibt eine reelle Zahl x mit $x^2 = 2$".

Die reellen Zahlen $\mathbb{R}$ bilden einen Körper.

1.2.9 Definition: Ein Tripel $(\mathbb{K}, +, \cdot)$ bestehend aus einer Menge $\mathbb{K}$ und Verknüpfungen $+$ und $\cdot$ heißt Körper, wenn die Bedingungen (K1) – (K3) für $(\mathbb{K}, +, \cdot)$ an Stelle von $\mathbb{R}$ gelten.

1.2.10 Beispiel: (a) $\mathbb{F}_2 = \{0, 1\}$ mit den Verknüpfungen

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

bildet einen Körper.

Sei $n \in \mathbb{N}$. Allgemein kann man zeigen, dass es genau dann einen Körper mit n Elementen gibt, wenn n eine Primzahlpotenz ist, d.h. $n = p^m$ für eine Primzahl p und $m \in \mathbb{N}$. Diese Körper sind eindeutig bestimmt (bis auf *Isomorphie*; dieser Begriff wird weiter unten erläutert).

(b) $(\mathbb{Q}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sind Körper.

1.2.11 Anordnungsaxiome In $\mathbb{R}$ sind gewisse Elemente als positiv ausgezeichnet (Schreibweise $x > 0$) so daß folgende Anordnungsaxiome erfüllt sind:

(A1) Für jedes $x \in \mathbb{R}$ gilt genau eine der drei Beziehungen: $x > 0$, $x = 0$, $-x > 0$

(A2) Sind $x > 0$ und $y > 0$, so folgt $x + y > 0$.

(A3) Sind $x > 0$ und $y > 0$, so folgt $xy > 0$.

1.2.12 Definition: Man schreibt $x > y$ falls $x - y > 0$. $x \geq y$ bedeutet $x > y$ oder $x = y$. Statt $x > y$ (bzw. $x \geq y$) schreibt man auch $y < x$ (bzw. $y \leq x$).

Man kann aus den Axiomen (A1) – (A3) die folgenden Aussagen ableiten:

1.2.13 Folgerungen: (1) $x < 0 \iff -x > 0$;
(2) $x < y$ und $y < z \Rightarrow x < z$;
(3) $x < y \Rightarrow x + z < y + z \quad \forall\, z \in \mathbb{R}$;
(4) $x_1 < x_2$, $y_1 < y_2 \Rightarrow x_1 + y_1 < x_2 + y_2$;
(5) $x < y$, $a > 0 \Rightarrow ax < ay$;
(6) $0 \leq x < y$, $0 \leq a < b \Rightarrow 0 \leq ax < by$;
(7) $x < y$, $a < 0 \Rightarrow ax > ay$;
(8) $\forall\, x \in \mathbb{R}$, $x \neq 0 : x^2 > 0$;
(9) $x > 0$ (bzw. $x < 0$) $\Rightarrow x^{-1} > 0$; (bzw. $x^{-1} < 0$)
(10) $1 > 0$.

Wir geben nur für einige dieser Aussagen einen Beweis. Die restlichen Beweise seien als Übungsaufgaben empfohlen:

zu (1): Nach Definition ist $x < 0$ äquivalent zu $0 > x$, was wiederum äquivalent zur Aussage $-x = 0 - x > 0$ ist.

zu (2): $x < y \wedge y < z \Longleftrightarrow y - x > 0 \wedge z - y > 0$ (nach Definition)
$\Rightarrow z - x = (y - x) + (z - y) > 0$ (nach (A2))
$\Longleftrightarrow x < z$ (nach Definition).

zu (3): $x < y \overset{def}{\Leftrightarrow} y - x > 0 \Rightarrow (y + z) - (x + z) = y - x > 0 \overset{def}{\Leftrightarrow} x + z < y + z$.

zu (4): $x_1 < x_2 \wedge y_1 < y_2 \Rightarrow x_1 + y_1 < x_2 + y_1 \wedge y_1 + x_2 < y_2 + x_2$ (nach (3))
$\Rightarrow x_1 + y_1 < y_2 + x_2$ (nach (2)).

zu (6): Wir betrachten zwei Fälle.
1. Fall: $x = 0$. Dann ist $ax = 0$ während $by > 0$ (nach (A3)), also insgesamt $ax = 0 < by$.
2. Fall: $x \neq 0$. Dann ist $0 \leq a < b$ und $0 < x < y$. Wegen $b - a > 0$ folgt $bx - ax = (b - a)x > 0$ (nach (A3)) und damit $ax < bx$. Ferner folgt aus $b > 0$ und $x < y$ ($\Leftrightarrow 0 < y - x$) auch $0 < b(y - x) = by - bx$ (wieder nach (A3)), also $bx < by$. Die beiden Aussagen $ax < bx$ und $bx < by$ implizieren nach (2) die Behauptung: $ax < by$.

zu (8): Für $x \in \mathbb{R}$, $x \neq 0$ gilt nach (A1): $x > 0$ oder $-x > 0$. Im ersten Fall folgt die Behauptung aus (A3): $x^2 = xx > 0$.
In zweiten Fall gilt: $x^2 = (-x)^2 > 0$ (die zweite Ungleichung folgt wieder aus (A3).

zu (10): Dies ist ein Spezialfall von (8) (für $x = 1$).

Um die reellen Zahlen vollständig axiomatisch zu charakterisieren benötigen wir noch zwei weitere Axiome.

(A4) (Archimedisches Axiom) Für je zwei positive reelle Zahlen x, y gibt es eine natürliche Zahl n mit $nx > y$.

1.2.14 Definition: Ein Körper $(\mathbb{K}, +, \cdot)$ in dem gewisse Elemente als positiv ausgezeinet sind, so daß die Axiome (A1) – (A3) gelten, heißt angeordneter Körper. Gilt überdies auch noch (A4) so heißt $\mathbb{K}$ archimedisch angeordnet.

1.2.15 Beispiele: (a) $\mathbb{R}$ und $\mathbb{Q}$ sind archimedisch angeordnete Körper.

(b) $\mathbb{F}_2$ und $\mathbb{C}$ sind nicht angeordnet da $-1 = 1 > 0$.

Beweis für $\mathbb{F}_2$: Wir führen einen indirekten Beweis. Angenommen es gibt in $\mathbb{F}_2$ gewisse als positiv ausgezeichnet Elemente (die wieder durch $x > 0$

gekennzeichnet werden), so dass die Axiome (A1) – (A3) gelten. Da in $\mathbb{F}_2$, $1+1=0$ gilt, ist $-1=1$ (d.h. in $\mathbb{F}_2$ ist das bzgl. der Addition inverse Element von 1 gleich 1). Das widerspricht aber dem Axiom (A1) wonach genau eines der Elemente 1 und -1 positiv ist.

Beweis für $\mathbb{C}$: Wieder durch indirekten Beweis. Angenommen die komplexen Zahlen lassen sich anordnen. Dann gelten die Regeln (1) – (10) auch für $\mathbb{C}$. Aus (8) und (10) folgt dann $-1 = i^2 > 0$ und $1 > 0$ im Widerspruch zu (A1). $\square$

1.2.16 Um das letzte Axiom formulieren zu können benötigen wir den Begriff der oberen Schranke einer Teilmenge von $\mathbb{R}$. Sei $M \subseteq \mathbb{R}$, $M \neq \emptyset$. Eine Zahl $s \in \mathbb{R}$ heißt *obere Schranke* von M falls

$$m \leq s \qquad \text{für alle } m \in M.$$

Besitzt M eine obere Schranke, so heißt M nach oben beschränkt.

Ein $s \in \mathbb{R}$ heißt *kleinste obere Schranke* von M, wenn gilt:

(1) s ist obere Schranke von M;

(2) Ist t eine weitere obere Schranke von M so ist $s \leq t$.

1.2.17 Lemma: Besitzt M eine kleinste obere Schranke, so ist diese eindeutig.

Beweis: Seien s_1, s_2 kleinste obere Schranken von M, so folgt $s_1 \leq s_2$ und $s_2 \leq s_1$, also $s_1 = s_2$. $\square$

1.2.18 Definition: Sei $M \subseteq \mathbb{R}$, $M \neq \emptyset$. Die kleinste obere Schranke von M bezeichnen wir mit $\sup M$, dem Supremum von M. Hat M keine obere Schranke, so setzen wir $\sup M := +\infty$.

(A5) (Vollständigkeitsaxiom): Jede nichtleere Teilmenge $M \subseteq \mathbb{R}$, die nach oben beschränkt ist, besitzt ein Supremum.

1.2.19 Definition: Ein archimedisch angeordneter Körper, der das Vollständigkeitsaxiom erfüllt, heißt vollständiger, archimedisch angeordneter Körper.

$\mathbb{R}$ ist also vollständig, während $\mathbb{Q}$ es nicht ist. Wir werden nämlich gleich sehen, dass das Axiom (A5) impliziert, dass $\sqrt{2} \in \mathbb{R}$, d.h. die Gleichung $x^2 = 2$ besitzt eine Lösung x in $\mathbb{R}$. Wäre $\mathbb{Q}$ vollständig, so würde folglich auch $\sqrt{2} \in \mathbb{Q}$ gelten. Bekanntlich ist $\sqrt{2}$ aber irrational.

1.2.20 Bemerkung: Man kann zeigen, daß $\mathbb{R}$ der einzige vollständig, archimedisch angeordnete Körper ist (bis auf Isomorphie), d.h. die Axiome (K1) – (K3), (A1) – (A5) charakterisieren $\mathbb{R}$ eindeutig. Genauer gilt: Sei $(\mathbb{K}, +_{\mathbb{K}}, \cdot_{\mathbb{K}})$ ein vollständiger, archimedisch angeordneter Körper, dessen positive Elemente durch $x >_{\mathbb{K}} 0$ bezeichnet werden. (Um auch die Addition und Multiplikation in $\mathbb{K}$ von derjenigen in $\mathbb{R}$ zu unterscheiden habe ich diese mit dem Index $\mathbb{K}$ versehen). Dann gibt es eine bijektive Abbildung $\phi : \mathbb{R} \to \mathbb{K}$, so dass für alle $x, y \in \mathbb{R}$ gilt

$$(1) \qquad \phi(x+y) = \phi(x) +_{\mathbb{K}} \phi(y), \quad \phi(x \cdot y) = \phi(x) \cdot_{\mathbb{K}} \phi(y),$$

und ferner für alle $x \in \mathbb{R}$

$$x > 0 \Longleftrightarrow \phi(x) >_{\mathbb{K}} 0.$$

Eine bijektive Abbildung zwischen Körpern, die (1) erfüllt, heißt *Isomorphismus*.

1.2.21 Satz: $\sqrt{2} \in \mathbb{R}$, d.h. es gibt ein $x \in \mathbb{R}$, $x > 0$ mit $x^2 = 2$.

Beweis: Wir betrachten dazu die Menge

$$M = \{x \in \mathbb{R} \mid x^2 < 2\}.$$

M ist durch $s = 2$ nach oben beschränkt. Nach dem Vollständigkeitsaxiom besitzt M eine kleinste obere Schranke $x_0 \in \mathbb{R}$. Wegen $1 \in M$ gilt $1 \leq x_0 \leq 2$. Wir wollen zeigen, daß $x_0^2 = 2$.

Indirekter Beweis: Angenommen $x_0^2 \neq 2$, also entweder $x_0^2 < 2$ oder $x_0^2 > 2$.
1. Fall: $x_0^2 < 2$. Sei $\varepsilon := \frac{2-x_0^2}{5} < 1 \Rightarrow \varepsilon > 0$
$\Rightarrow (x_0+\varepsilon)^2 = x_0^2+\varepsilon(2x_0+\varepsilon) < x_0^2+\varepsilon(2x_0+1) \leq x_0^2+\varepsilon\cdot 5 = x_0^2+(2-x_0^2) = 2$
$\Rightarrow x_0 + \varepsilon \in M$, also $x_0 + \varepsilon \leq x_0$ da x_0 obere Schranke von M ist $\Rightarrow \varepsilon \leq 0$
Widerspruch!
2. Fall: $x_0^2 > 2$. Sei $\varepsilon = \frac{x_0^2-2}{4} < 1 \Rightarrow \varepsilon > 0$.

$\Rightarrow (x_0 - \varepsilon)^2 = x_0^2 - \varepsilon(2x_0 - \varepsilon) > x_0^2 - 4 = 2 \Rightarrow x_0 - \varepsilon \geq x$ für alle $x \in M$
$\Rightarrow x_0 - \varepsilon$ ist obere Schranke von $M \Rightarrow x_0 - \varepsilon \geq x_0$ (da x_0 die kleinste obere Schranke von M ist) $\Rightarrow \varepsilon \leq 0$, ein Widerspruch!

Also ist $x_0^2 = 2$. $\square$

1.2.22 Bemerkung: Analog kann man zeigen, daß es zu jeder positiven reellen Zahl x und jeder natürlichen Zahl n eine positive reelle Zahl y gibt mit $y^n = x$.

Analog zu oberen Schranken kann man auch untere Schranken behandeln. Für eine Teilmenge $M \subseteq \mathbb{R}$, $M \neq \emptyset$ heißt $s \in \mathbb{R}$ eine *untere Schranke* von M falls gilt:

$$m \geq s \qquad \forall m \in M.$$

Besitzt M eine untere Schranke, so heißt M nach unten beschränkt. $s \in \mathbb{R}$ heißt größte untere Schranke, wenn gilt (i) s ist untere Schranke und (ii) ist $t \in \mathbb{R}$ eine untere Schranke so ist $t \leq s$.

Es gilt wieder, dass die größte untere Schranke einer nichtleeren Menge M eindeutig bestimmt ist, falls sie existiert.

1.2.23 Definition: Sei $M \subseteq \mathbb{R}$, $M \neq \emptyset$. Die größte untere Schranke s von M wird als das Infimum von M bezeichnet (in Zeichen: $s =: \inf M$). Hat M keine untere Schranke so setzen wir $\inf M :\ = -\infty$.

Man könnte jetzt erwarten, daß man zusätzlich zum Axiom (A5) noch durch ein weiteres Axiom die Existenz von größten unteren Schranken fordern muss. Das ist aber nicht nötig, da (A5) dies bereits impliziert. Für eine Teilmenge $M \subseteq \mathbb{R}$, $M \neq \emptyset$ und $s \in \mathbb{R}$ gilt nämlich:

$$s \text{ ist untere Schranke von } M \iff$$

$$-s \text{ ist obere Schranke von } -M := \{-x \mid x \in M\}$$

Damit folgt sofort:

M ist genau dann nach unten beschränkt, wenn $-M$ nach oben beschränkt ist. In diesem Fall besitzt es eine größte untere Schranke nämlich $-\sup(-M)$. Es gilt also $-\sup(-M) = \inf(M)$. Diese Gleichheit gilt auch noch, wenn M nicht nach unten beschränkt ist, denn dann haben wir $-\sup(-M) = -(+\infty) = -\infty = \inf(M)$.

1.2.24 Beispiel:

M	$\inf M$	$\sup M$
(0,1)	0	1
(0,1]	0	1
$[0,1]$	0	1

Hierbei bezeichnet $(0,1)$ (bzw. $(0,1]$ bzw. $[0,1]$) das offene (bzw. halboffene, bzw. abgeschlossene) Intervall $\{x \in \mathbb{R} | \, 0 < x < 1\}$ (bzw. $\{x \in \mathbb{R} | \, 0 < x \leq 1\}$ bzw. $\{x \in \mathbb{R} | \, 0 \leq x \leq 1\}$).

1.2.25 Definition: Sei $M \subseteq \mathbb{R}$, $M \neq \emptyset$. Ein Element $m \in M$ heißt Maximum (bzw. Minimum) von M (in Zeichen: $m = \max M$ bzw. $m = \min M$), falls $m \geq x$ (bzw. $m \leq x$) für alle $x \in M$.

Hat M ein Maximum (bzw. Minimum) so gilt offensichtlich $\max M = \sup M$ (bzw. $\min M = \inf M$). Im Allgemeinen muss aber für beschränkte Mengen weder Maximum noch Minimum existieren, wie das Beispiel des offenen Intervalls $M = (0,1)$ zeigt.

Offenbar hat jede endliche Menge ein Maximum und ein Minimum. Für ein n-Tupel reeller Zahlen $(x_1, x_2, \ldots, x_n)$ setzen wir

$$\max(x_1, x_2, \ldots, x_n) : = \max\{x_1, x_2, \ldots, x_n\},$$

$$\min(x_1, x_2, \ldots, x_n) : = \min\{x_1, x_2, \ldots, x_n\}.$$

Zum Abschluss dieses Paragraphen geben wir noch die formale Definition der Betragsfunktion an.

1.2.26 Definition: Für eine reelle Zahl x ist der Betrag von x definiert als

$$|x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases}$$

Nach dem Axiom (A1) ist das wohldefiniert.

1.3 Natürliche Zahlen und vollständige Induktion

1.3.1 Eine grundlegende Eigenschaft der Menge $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ der natürlichen Zahlen ist das eine Teilmenge $M \subseteq \mathbb{N}$ schon dann gleich $\mathbb{N}$ ist, wenn gilt:

(1) $1 \in M$,

(2) $\forall\, n \in \mathbb{N} : n \in M \Rightarrow n + 1 \in M$

Dies ist eine Umformulierung des Beweisprinzip der *vollständigen Induktion*:

Für jede natürliche Zahl n sei eine Aussage $A(n)$ gegeben. Falls gilt:

(I) $A(1)$ richtig ist (Induktionsanfang).

(II) Für jedes n, für welches $A(n)$ richtig ist (Induktionsvoraussetzung), ist auch $A(n + 1)$ richtig (Induktionsschluß).

Dann gilt: $A(n)$ richtig für alle $n \in \mathbb{N}$.

Beweis: Sei $M = \{n \in \mathbb{N} \mid A(n) \text{ ist richtig}\}$. Nach (I) ist $1 \in M$ und nach (II) gilt $n \in M \Rightarrow n+1 \in M$. Folglich ist $M = \mathbb{N}$, d.h. $A(n)$ ist wahr für alle $n \in \mathbb{N}$. $\square$

1.3.2 Beispiel: Für jede natürliche Zahl n gilt:
$A(n): \quad 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$

Induktionsanfang: $A(1)$ ist richtig, da $1 = \frac{1(1+1)}{2}$.

Induktionsschluß: Wir nehmen an, dass $A(n)$ richtig ist (Induktionsvoraussetzung) und wollen daraus folgern, dass auch $A(n + 1)$ gilt.

$$1 + 2 + \ldots + n + (n + 1) = (1 + 2 + \ldots + n) + (n + 1) \overset{*}{=} \frac{1}{2} n(n + 1) + (n + 1)$$

Die Gleichheit $*$ folgt aus der Induktionsvoraussetzung. Da

$$\frac{1}{2} n(n + 1) + (n + 1) = (n + 1) \left(\frac{n}{2} + 1\right) = \frac{1}{2}(n + 1)(n + 2)$$

folgt $1 + 2 + \ldots + n + (n + 1) = \frac{1}{2}(n + 1)(n + 2)$, d.h. $A(n + 1)$ ist richtig. $\square$

1.3.3 Beispiel: Jede natürliche Zahl $n > 1$ ist Produkt von Primzahlen.

Erinnerung: Eine natürliche Zahl $p \geq 2$ heißt Primzahl, wenn sie nur durch 1 und sich selbst teilbar ist.

Beweis: Sei $A(n)$ die Aussage:

Jede natürliche Zahl m, $2 \leq m \leq n$ ist Produkt von Primzahlen.

$n = 1$: $A(1)$ ist richtig (die Aussage ist leer).

Induktionsschluss $n \mapsto n+1$: Sei also $A(n)$ richtig. Wir müssen zeigen, daß eine natürliche Zahl m, die $\leq n+1$ ist, Produkt von Primzahlen ist. Falls m sogar $\leq n$ ist, so folgt das aus der Induktionsvoraussetzung. Es bleibt also noch der Fall $m = n+1$ zu betrachten.

1. Fall: $n+1 = p$ ist Primzahl. Dann ist nichts zu zeigen.

2. Fall: $n+1$ hat einen Teiler d mit $d \neq 1$, $d \neq n+1$. Dann gilt:
$n+1 = dk, \quad$ mit $2 \leq d, k \leq n$.
$\Rightarrow m, k$ sind jeweils Produkte von Primzahlen.
$\Rightarrow n$ ist Produkt von Primzahlen. $\square$

1.3.4 Summen und Produkte: Seien $m \leq n$ ganze Zahlen. Für jede ganze Zahl k, $m \leq k \leq n$ sei eine reelle Zahl a_k gegeben. Dann setzt man

$$\sum_{k=m}^{n} a_k := a_m + a_{m+1} + \ldots + a_n$$
$$\prod_{k=m}^{n} a_k := a_m a_{m+1} \cdot \ldots \cdot a_n$$

Für $n = m-1$ führt man folgende Konventionen ein:

$$\sum_{k=m}^{m-1} a_k := 0$$
$$\prod_{k=m}^{m-1} a_k := 1$$

1.3.5 Fakultäten und Binomialkoeffizienten: Für $n \in \mathbb{N}_0$ ist $n!$ (n Fakultät) wie folgt definiert:

$$0! := 1, \qquad n! := 1 \cdot 2 \cdot \ldots \cdot n = \prod_{k=1}^{n} k$$

Es gilt also $(n+1)! = n!(n+1)$ für jedes $n \in \mathbb{N}_0$.

Wir wollen jetzt zeigen, dass $n!$ gerade die Anzahl der verschiedenen bijektiven Abbildung einer n-elementigen Menge in sich ist. Wir betrachten speziell die Menge $\{1, 2, \ldots, n\}$. Eine *Permutation* der Menge $\{, 1, 2, \ldots, n\}$ ist eine bijektive Abbildung

$$\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}.$$

Es gilt also $\pi(i) \neq \pi(j)$ für $i \neq j$. Sei $\mathfrak{S}_n$ die Menge aller Permutationen von $\{1, 2, \ldots, n\}$.

1.3.6 Bemerkung: $\mathfrak{S}_n$ bildet zusammen mit der Komposition $\circ$ eine Gruppe. Sie heisst *symmetrische Gruppe.*

Für eine Menge M mit endlich vielen Elementen bezeichnet $|M|$ die Anzahl der Elemente von M.

1.3.7 Satz: $|\mathfrak{S}_n| = n!$.

Man kann diesen Satz auch wie folgt interpretieren:

Es gibt genau $n!$ Möglichkeiten n Gegenstände $G_1, \ldots, G_n$ nacheinander anzuordnen.

Ist nämlich $G_{i_1}, \ldots, G_{i_n}$ eine solche Anordnung, so kommt jede der Zahlen $1, 2, \ldots, n$ unter den Zahlen $i_1, i_2, \ldots, i_n$ genau einmal vor, d.h. $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}, k \mapsto i_k$ ist eine Permutation. Umgekehrt liefert jede Permutation $\pi \in \mathfrak{S}_n$ eine Anordnung von $G_1, \ldots, G_n$, nämlich $G_{\pi(1)}, \ldots, G_{\pi(n)}$. Also entsprechen die möglichen Anordnungen von $G_1, \ldots, G_n$ eineindeutig den Permutationen von $\{1, 2, \ldots, n\}$.

Beweis des Satzes durch Induktion: Der Fall $n = 1$ ist klar.

$n \to n+1$: Für $i \in \{1, \ldots, n+1\}$ sei $\mathfrak{S}_{n+1}^{(i)}$, die Menge der Permutationen $\pi : \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}$ mit $\pi(n+1) = i$.

Ein Element $\pi \in \mathfrak{S}_{n+1}^{(i)}$ entspricht also einer Anordnung von $n+1$ Gegenständen $G_1, \ldots, G_{n+1}$, für die der letzte Gegenstand G_{n+1} an der i-ten Stelle steht.

Da wir die übrig gebliebenen n Gegenstände auf den restlichen n Plätzen immer noch frei anordnen können hat $\mathfrak{S}_{n+1}^{(i)}$ genauso viele Elemente wie $\mathfrak{S}_n$, also nach Induktionsvoraussetzung genau $n!$.

Da jedes $\pi \in \sigma_{n+1}$ in genau einer der Menge $\mathfrak{S}_{n+1}^{(1)}, \ldots, \mathfrak{S}_{n+1}^{(n+1)}$ liegt folgt:

$$|\mathfrak{S}_{n+1}| = |\mathfrak{S}_{n+1}^{(1)}| + \ldots + |\mathfrak{S}_{n+1}^{(n+1)}| = n! + \ldots + n! \;((n+1)\text{-mal}) = (n+1)!$$

. $\square$

1.3.8 Definition: Seien $n, k \in \mathbb{N}_0$, $0 \leq k \leq n$. Wir definieren den Binomialkoeffizient $\binom{n}{k}$ durch

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots 1}.$$

Es gilt die folgende Rekursionsformel:

1.3.9 Lemma: $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$ für $0 \leq k \leq n$.

Beweis:

$$\begin{aligned}\binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-(k+1))!} \\ &= \frac{(n+1)!}{(k+1)!(n-k)!}\left(\frac{k+1}{n+1} + \frac{n-k}{n+1}\right) = \binom{n+1}{k+1}\end{aligned}$$

$\square$

1.3.10 Satz: Seien $n \in \mathbb{N}, k \in \mathbb{N}_0$ mit $0 \leq k \leq n$. Die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge M ist $\binom{n}{k}$.

Beweis: Es bezeichne C_k^n die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge $M = \{a_1, \ldots, a_n\}$.

Vorbemerkung: Es gilt $C_0^n = 1$, da die leere Menge die einzige Teilmenge mit 0 Elementen ist und $C_n^n = 1$, da M die einzige n-elementige Teilmenge von M ist.

Die Gleichheit $C_k^n = \binom{n}{k}$ wird durch Induktion über n gezeigt.

$n = 1: \quad C_0^1 = 1 = \binom{1}{0}, \quad C_1^1 = 1 = \binom{1}{1}.$

$n \to n+1$: Ist $k = 0$ oder $k = n+1$ so gilt nach der Vorbemerkung $C_k^{n+1} = 1 = \binom{n+1}{k}$.

Sei jetzt $1 \leq k \leq n$. Die k-elementigen Teilmengen von $M = \{a_1, \ldots, a_n, a_{n+1}\}$ zerfallen in zwei Klassen $\mathcal{K}_0$ und $\mathcal{K}_1$, wobei $\mathcal{K}_0$ alle Teilmengen umfasst, die a_{n+1} nicht enthalten und $\mathcal{K}_1$ alle Teilmengen die a_{n+1} enthalten. Die Anzahl der Teilmengen in $\mathcal{K}_0$ ist gleich der Anzahl der k-elementigen Teilmengen in $\{a_1, \ldots, a_n\}$, also nach Induktionsvoraussetzung $= C_k^n = \binom{n}{k}$. Die Mengen der Klasse $\mathcal{K}_1$ bestehen aus $k-1$ Elementen aus $\{a_1, \ldots, a_n\}$ und aus dem Element a_{n+1}. Nach Induktionsvoraussetzung besteht $\mathcal{K}_1$ aus $C_{k-1}^n = \binom{n}{k-1}$ Mengen. Damit ergibt sich mit Hilfe des obigen Lemmas:

$$C_k^{n+1} = \binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

$\square$

1.3.11 Satz: Seien $n \in \mathbb{N}$, $x, y \in \mathbb{R}$. Dann gilt die verallgemeinerte Binomische Formel

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Beweis durch Induktion: Der Fall $n = 1$ ist klar.

$n \to n+1$:

$$(x+y)^{n+1} = (x+y)^n (x+y) \stackrel{*}{=} (\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k})(x+y) =$$

$$= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} =$$

$$= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n-(k-1)} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} =$$

$$= x^{n+1} y^0 + \sum_{k=1}^{n} (\binom{n}{k-1} + \binom{n}{k}) x^k y^{n+1-k} + x^0 y^{n+1} =$$

$$\stackrel{**}{=} \sum_{k=1}^{n} \binom{n+1}{k} x^k y^{n+1-k}$$

wobei die Gleichheit $*$ aus der Induktionsvoraussetzung und $**$ aus obigem Lemma folgt. $\square$

1.4 Komplexe Zahlen

Die komplexen Zahlen sind aus dem Wunsch heraus entstanden den Zahlbereich der reellen Zahlen so zu vergrössern, dass die Gleichung $x^2+1=0$ ein Lösung besitzt. Eine Lösung dieser Gleichung wird mit i bezeichnet. Komplexen Zahlen sind Ausdrücke der Form $x+iy$ mit denen man "genauso rechnet" wie mit reellen Zahlen wobei man berücksichtigt, dass $i^2=-1$. Die formale Definition der komplexen Zahlen ist wie folgt:

1.4.1 Definition: Eine komplexe Zahl ist ein Element $z=(x,y)$ der Menge $\mathbb{R}^2=\mathbb{R}\times\mathbb{R}$.

Komplexe Zahlen werden wie folgt addiert und multipliziert:

(A) $(x,y)+(u,v):=(x+u,y+v)$

(M) $(x,y)\cdot(u,v):=(xu-yv,xv+yu)$

1.4.2 Satz: Die Menge der komplexen Zahlen zusammen mit der Addition (A) und der Multiplikation (M) bildet einen Körper der mit $\mathbb{C}$ bezeichnet wird.

Beweis: Man rechnet leicht nach, dass die Addition (A) (bzw. die Multiplikation (M)) assoziativ und kommutativ sind. Das Distributivgesetz gilt wegen:

$$\begin{aligned}
&(x,y)\cdot((u_1,v_1)+(u_2,v_2))= \\
=&(x,y)\cdot(u_1+u_2,v_1+v_2)= \\
=&(x(u_1+u_2)-y(v_1+v_2),x(v_1+v_2)+y(u_1+u_2))= \\
=&(xu_1-yv_1+xu_2-yv_2,xv_1+yu_1+xv_2+yu_2)= \\
=&(xu_1-yv_1,xv_1+yu_1)+(xu_2-yv_2,xv_2+yu_2)= \\
=&(x,y)\cdot(u_1,v_1)+(x,y)\cdot(u_2,v_2).
\end{aligned}$$

Das neutrale Element bzgl. der Addition ist $(0,0)$ und das Inverse von (x,y) ist $(-x,-y)$. Das neutrale Element bzgl. der Multiplikation ist $(1,0)$, denn

$$(x,y)\cdot(1,0)=(x\,1-y\,0,x\,0+y\,1)=(x,y)=(1,0)\cdot(x,y)$$

für alle komplexe Zahlen (x,y). Schliesslich bleibt noch zu zeigen, dass eine komplexe Zahl $(x,y)\neq(0,0)$ ein Inverses bzgl. der Multiplikation besitzt. Wegen $x\neq0$ oder $y\neq0$ ist $x^2+y^2>0$. Da

$$(x,y)\cdot\left(\frac{x}{x^2+y^2},\frac{-y}{x^2+y^2}\right)=\left(\frac{x^2}{x^2+y^2}+\frac{y^2}{x^2+y^2},\frac{-xy}{x^2+y^2}+\frac{yx}{x^2+y^2}\right)=(1,0)$$

ist $(x,y)^{-1} :\ = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right)$ das Inverse von (x,y) bzgl. der Multiplikation.

Wir haben also gesehen, dass $\mathbb{C}$ mit der Addition (A) und der Multiplikation (M) den Körperaxiomen (K1) – (K3) genügt. $\square$

1.4.3 $\mathbb{R}$ als Unterkörper von $\mathbb{C}$: Für komplexe Zahlen der Gestalt $(x,0)$ gilt:

$$\begin{aligned} (x,0)+(y,0) &= (x+y,0) \\ (x,0)\cdot(y,0) &= (xy,0) \end{aligned}$$

Die komplexen Zahlen der Gestalt $(x,0)$ werden also wie die entsprechenden reellen Zahlen x addiert und multipliziert. Wir wollen im Folgenden nicht mehr zwischen $x \in \mathbb{R}$ und der komplexen Zahl $(x,0)$ unterscheiden. Wir schreiben also einfach x für $(x,0)$ und fassen auf diese Weise $\mathbb{R}$ als Teilmenge von $\mathbb{C}$. Insbesondere schreiben wir einfach 0 für $(0,0)$ und 1 statt $(1,0)$. Für jede komplexe Zahl gilt dann $z+0=z$, $z\cdot 1 = z$.

Da für reelle Zahlen die Addition (A) und Multiplikation (M) mit der gewöhnlichen Addition und Multiplikation übereinstimmt spricht man davon, dass $\mathbb{R}$ ein *Unterkörper* von $\mathbb{C}$ ist.

1.4.4 Bemerkung: Wenn man sehr genau sein möchte, kann man den Zusammenhang zwischen $\mathbb{R}$ und $\mathbb{C}$ wie folgt beschreiben: Die Abbildung

$$\begin{aligned} \iota : \mathbb{R} &\to \widetilde{\mathbb{R}} := \{(x,0) \mid x \in \mathbb{R}\} \\ x &\mapsto (x,0) \end{aligned}$$

ist ein Isomorphismus, d.h. ι ist bijektiv und es gilt: $\iota(x+y) = \iota(x)+\iota(y)$, $\iota(xy) = \iota(x)\iota(y)\ \forall\ x,y \in \mathbb{R}$. Also ist $\mathbb{R}$ zum Unterkörper $\widetilde{\mathbb{R}}$ von $\mathbb{C}$ *isomorph* (d.h. $\mathbb{R}$ und $\widetilde{\mathbb{R}}$ sind gleichstrukturiert).

1.4.5 Die imaginäre Einheit: Darunter versteht man die komplexe Zahl $i := (0,1)$.

Es gilt: $i^2 = i \cdot i = (0,1)(0,1) = (-1,0) = -1$, d.h. i und $-i$ sind die Lösungen der Gleichung $z^2+1=0$.

Für eine komplexe Zahl $z=(x,y)$ gilt:

$$z = (x,0)+(0,y) = (x,0)+(0,1)(y,0) = x+iy$$

Dies ist die *kartesische Darstellung* von z. Die reellen Zahlen x und y heißen der Real- bzw. der Imaginärteil von z (in Zeichen: $x := \mathrm{Re}(z)$, $y := \mathrm{Im}(z)$). Man kann eine komplexe Zahl z also geometrisch als einen Punkt der Ebene $\mathbb{R}^2$ mit x-Koordinate $\mathrm{Re}(z)$ und y-Koordinate $\mathrm{Im}(z)$ deuten. Man spricht deshalb auch von der *komplexen Zahlenebene*.

1.4.6 Die komplexe Konjugation: Für eine komplexe Zahl $z = (x, y) = x + iy$ setzt man

$$\overline{z} = (x, -y) = x - iy$$

Es gilt: $z\overline{z} = (x + iy)(x - iy) = x^2 - i^2y^2 = x^2 + y^2$ d.h. $z\overline{z}$ ist reell und > 0 falls $z \neq 0$.

1.4.7 Definition: Für $z \in \mathbb{C}$ heißt $|z| := \sqrt{z\overline{z}}$ der Betrag von z.

Geometrisch $|z|$ ist der Abstand von $z = x + iy$ vom Ursprung (oder anders ausgedrückt: $|z|$ ist die Länge des Vektors (x, y)). Das ergibt sofort aus dem Satz des Pythagoras.

Es gelten die folgenden Rechenregeln:

(1) $\mathrm{Re}\,(z) = \frac{1}{2}(z + \overline{z})$, $\mathrm{Im}\,(z) = \frac{1}{2i}(z - \overline{z})$;

(2) $z = \overline{z} \iff z \in \mathbb{R}$;

(3) $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\overline{w}$;

(4) $\overline{\overline{z}} = z$;

(5) $|\overline{z}| = |z|$;

(6) $|z \cdot w| = |z|\,|w|$;

(7) $|(x, 0)| = |x| \quad \forall\, x \in \mathbb{R}$;

(8) $|\mathrm{Re}(z)| \leq |z|$, $|\mathrm{Im}(z)| \leq |z|$;

(9) $|z + w| \leq |z| + |w|$ (Dreiecksungleichung).

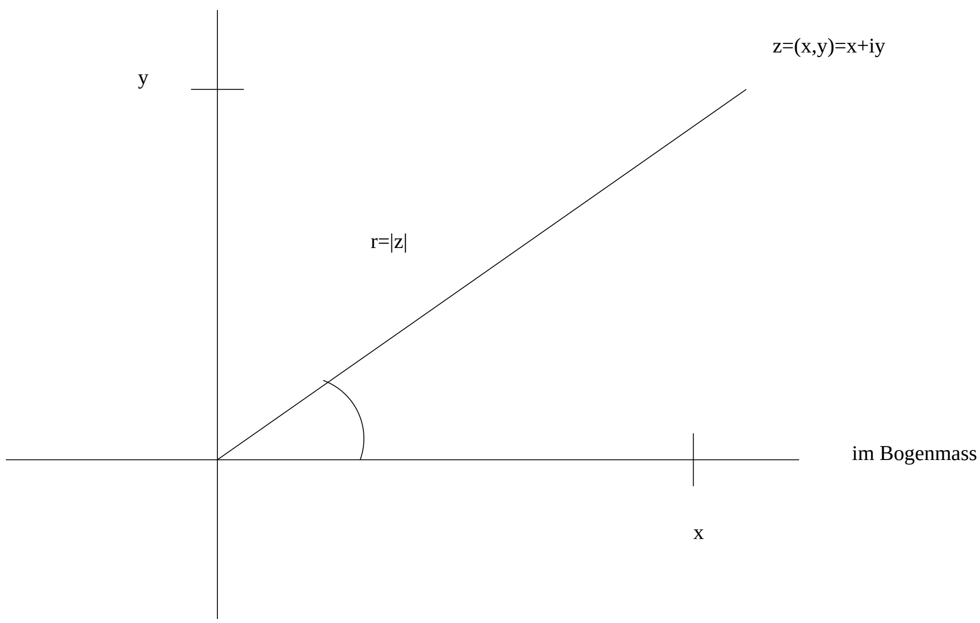
Wir beweisen nur (9) (der Nachweis von (1)–(8) sei als Übungsaufgabe empfohlen). Geometrisch ergibt sich (9) sofort aus der Tatsache, dass in einem

Dreick die Summe zweier Seiten immer grösser als die Dritte ist (das wendet man auf das Dreieck in der komplexen Zahlenebene mit den Eckpunkten $0, z, z+w$ an). Ein algebraischer Beweis verläuft wie folgt:

$$|z+w|^2 = (z+w)(\overline{z+w}) \overset{(3)}{=} (z+w)(\bar{z}+\bar{w}) =$$
$$= z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} \overset{(3)}{=} |z|^2 + z\bar{w} + \overline{z\bar{w}} + |w|^2 =$$
$$\overset{(1)}{=} |z|^2 + 2\mathrm{Re}(z\bar{w}) + |w|^2 \overset{(8)}{\leq} |z|^2 + 2|z\bar{w}| + |w|^2 =$$
$$\overset{(5),(6)}{=} |z|^2 + 2|z|\,|w| + |w|^2 = (|z|+|w|)^2$$

Es folgt $|z+w| \leq |z| + |w|$. □

1.4.8 Darstellung der komplexen Zahlen in Polarkoordinaten: Jede komplexe Zahl $z = x+iy \neq 0$ lässt sich in der Form $z = r(\cos\varphi + i\sin\varphi)$ mit eindeutig bestimmten $r \in \mathbb{R}, r > 0$, $\varphi \in [0, 2\pi)$. Dies ist die Darstellung von z in *Polarkoordinaten*. Dabei ist $r = |z|$ der Betrag von z. φ ist der Winkel zwischen der positiven Richtung der x-Achse und dem Vektor (x, y). Der Winkel φ wird als das Argument von z bezeichnet (in Zeichen $\varphi = \mathrm{Arg}\,(z)$).



Die folgenden Formeln geben die Umrechnung zwischen den kartesischen und Polarkoordinaten von z an:

$$\begin{array}{ll} x = r\cos\varphi & r = \sqrt{x^2+y^2} \\ y = r\sin\varphi & \tan\varphi = \frac{y}{x} \quad \text{(Quadranten beachten).} \end{array}$$

1.4.9 Beispiele: (a) Die Darstellung von $z = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ in Polarkoordinaten ist $\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)$, d.h. $r = 1$ und $\operatorname{Arg}(z) = \frac{\pi}{3}$.

(b) Sind umgekehrt die Polarkoordinaten von z, $r = 2$ und $\varphi = \frac{2\pi}{3}$, so ergibt sich für die kartesische Darstellung von z:

$$x = 2\cos\left(\frac{2\pi}{3}\right) = -1, \quad y = 2\sin\left(\frac{2\pi}{3}\right) = \sqrt{3}.$$

Der Vorteil der Darstellung in Poloarkoordinaten gegenüber der kartesischen Darstellung ist, dass sich in der ersten das Produkt zweier komplexer Zahlen einfacher beschreiben lässt. Sind $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1), z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$ zwei komplexe Zahlen $\neq 0$ in Polarkoordinatendarstellung so gilt:

$$z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)$$

Das folgt sofort aus den, aus der Schulmathematik bekannten, *Additionstheoremen*:

$$\begin{aligned}\cos(\varphi_1 + \varphi_2) &= \cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2,\\ \sin(\varphi_1 + \varphi_2) &= \sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2.\end{aligned}$$

Also gilt: Das Produkt zweier komplexer Zahlen $z_1 z_2$ erhält man, indem man ihre Beträge multipliziert und ihre Argumente addiert (bis auf ein Vielfaches von 2π, d.h. $\operatorname{Arg}(z_1 z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2) + 2\pi k$ für eine ganze Zahl k).

1.4.10 Beispiele: (a) Sei $z = \frac{1}{2} + \frac{\sqrt{3}}{2}i = \cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)$. Dann ist

$$z^6 = \cos\left(6\frac{\pi}{3}\right) + i\sin\left(6\frac{\pi}{3}\right) = 1.$$

Entsprechend gilt $(\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i)^8 = 1$, da $|\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i| = 1$ und $\operatorname{Arg}((\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i) = \frac{\pi}{4}$.

(b) Die Gleichung $z^n = 1$ für $n \in \mathbb{N}$ besitzt die n verschiedenen Lösungen $z = \cos\left(\frac{2\pi k}{n}\right) + i\sin\left(\frac{2\pi k}{n}\right)$ für $k = 0, 1, \ldots, n-1$.

Das letzte Beispiel ist ein Spezialfall des folgenden Satzes.

1.4.11 Satz: Sei $z \in \mathbb{C}$, $z \neq 0$ eine komplexe Zahl. Dann hat die Gleichung

$$w^n = z \tag{2}$$

genau n verschiedene Lösungen. Ist $z = r(\cos\varphi + i\sin\varphi)$ die Darstellung von z in Polarkoordinaten, so sind

$$w_k = \sqrt[n]{r}(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}), \quad k = 0, 1, \ldots, n-1$$

die verschiedenen Lösungen.

Dabei ist $\sqrt[n]{r}$ die eindeutig bestimmte n-te positive reelle Wurzel von reellen Zahl $r > 0$. Beweis: Ist $w = t(\cos\psi + i\sin\psi)$ eine Lösung der Gleichung (2), so gilt

$$t^n = r, \qquad n\psi = \varphi + 2\pi k$$

für ein gewisses $k \in \mathbb{Z}$. Wegen $\psi, \varphi \in [0, 2\pi)$ ist

$$2\pi k = n\psi - \varphi \leq n\psi < 2\pi n,$$

und

$$2\pi k = n\psi - \varphi \geq -\varphi > -2\pi$$

also $k \in \{0, 1, \ldots, n-1\}$.

Umgekehrt sind die $w_k := \sqrt[n]{r}\ (\cos\frac{\varphi+2\pi k}{n} + i\ \sin\frac{\varphi+2\pi k}{n})$ für $k = 0, 1, \ldots, n-1$ Lösungen von (2), da $|w_k^n| = \sqrt[n]{r}^n = r$ und

$$\mathrm{Arg}\,(w_k^n) \doteq n\mathrm{Arg}\,(w_k) \doteq \varphi + 2k\pi \doteq \varphi$$

wobei $\doteq$ bedeutet, dass eine Gleichheit bis auf ein ganzzahliges Vielfaches von 2π vorliegt. Damit folgt $\mathrm{Arg}\,(w_k^n) = \mathrm{Arg}\,(z)$, also $w_k^n = z$. □

1.4.12 Bemerkung: Üblicherweise versteht man unter der Polarkoordinatendarstellung einer komplexen Zahl $z \in \mathbb{C}$ $z \neq 0$, eine Darstellung der Form

$$z = r\,e^{i\varphi}, \qquad \text{mit } r \in \mathbb{R}, r > 0, \varphi \in [0, 2\pi).$$

Dabei ist $e = 2,718\ldots$ die *Eulersche Zahl.*

Diese ist äquivalent zu der Darstellung $z = r(\cos\varphi + i\sin\varphi)$, denn es gilt die *Eulersche Formel*:

$$e^{i\varphi} = \cos\varphi + i\ \sin\varphi.$$

Dazu ist aber zunächst einmal zu erklären, was der Ausdruck $e^{i\varphi}$ bedeutet. Aus der Schulmathematik sollte die folgende Formel für die Exponentialfunktion $\exp(x) := e^x$ bekannt sein:

$$e^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Für eine komplexe Zahl z definiert man analog

$$e^z = \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Dass man die Reihe $\sum\limits_{n=0}^{\infty} \frac{z^n}{n!}$ überhaupt bilden (d.h. dass sie *konvergiert*) werden wir in einem späteren Abschnitt über Folgen und Reihen zeigen.

1.5 Polynome

Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.

1.5.1 Definition: Ein Polynom (über $\mathbb{K}$) ist eine Funktion der Form

$$f : \mathbb{K} \to \mathbb{K}, z \mapsto f(z) = a_n z^n + \ldots + a_1 z + a_0 = \sum_{k=0}^{n} a_k z^k$$

mit $a_0, a_1, \ldots, a_n \in \mathbb{K}$. Ist $a_n \neq 0$, so heißt n der Grad des Polynoms (in Zeichen: $n = \deg f$)

Sind alle $a_k = 0$, so heißt f das Nullpolynom, dem wir den Grad $-\infty$ zuordnen.

Summen und Produkte von Polynomen sind wieder Polynome. Das Produkt der Polynome

$$\begin{aligned} f(z) &= a_n z^n + \ldots + a_1 z + a_0 \\ g(z) &= b_m z^m + \ldots + b_1 z + b_0 \end{aligned}$$

ist das Polynom

$$f(z)g(z) = c_{m+n} z^{m+n} + \ldots + c_1 z + c_0$$

mit

$$c_k = \sum_{r+s=k} a_r b_s \qquad k = 0, 1, \ldots, m+n.$$

1.5.2 Satz (Division mit Rest): Sei $g(z)$ ein Polynom $\neq 0$. Dann gibt es zu jedem Polynom $f(z)$ eindeutig bestimmte Polynome $q(z)$ und $r(z)$ mit

$$f(z) = g(z) \cdot q(z) + r(z)$$

und $\deg r(z) < \deg g(z)$.

Beweis der Existenz: Ist $\deg f < \deg g$ so ist $f = g \cdot 0 + f$ eine Zerlegung der gesuchten Art.

Im anderen Fall sei $n := \deg f \geq m := \deg g$, $f(z) = a_n z^n + \ldots + a_1 z + a_0$, $g(z) = b_m z^m + \ldots + b_1 z + b_0$.

$$\Rightarrow f(z) - a_n b_m^{-1} z^{n-m} g(z) =: f_1(z)$$

ist ein Polynom vom Grad $n_1 < n$.

Ist $n_1 \geq m$ so subtrahieren wir von f_1 wieder ein Vielfaches von g, so daß die Differenz ein Polynom vom Grad $n_2 < n_1$ ist.

So fortfahrend erhält man schließlich ein Polynom r dessen Grad $< m$ ist.

Mit geeignetem q gilt also $f - gq = r$.

Eindeutigkeit: Ist $f = gq' + r'$ eine weitere derartige Zerlegung mit $q \neq q'$, so folgt:

$$g(q - q') = r - r' \Rightarrow \deg\, g(q - q') = \deg\,(r - r') < m,$$

ein Widerspruch. $\square$

Die Menge der Polynome über $\mathbb{K}$ wird mit $\mathbb{K}[z]$ bezeichnet.

1.5.3 Definition: Ein Element $\alpha \in \mathbb{K}$ heißt Nullstelle von $f(z) \in \mathbb{K}[z]$, wenn $f(\alpha) = 0$.

1.5.4 Lemma: Sei $f(z)$ ein Polynom über $\mathbb{K}$ und $\alpha \in \mathbb{K}$. α ist genau dann Nullstelle von $f(z)$, wenn $f(z)$ durch $z - \alpha$ teilbar ist, d.h. es gibt ein Polynom $q(z)$ vom Grad $\deg\,(g) = \deg\,(f) - 1$ mit

$$f(z) = (z - \alpha)q(z)$$

Beweis: Nach dem Satz über die Division mit Rest gibt es $q(z), r(z) \in \mathbb{K}[z]$ mit

$$f(z) = (z - \alpha)q(z) + r(z) \text{ und } \deg\, r(z) < \deg\,(z - \alpha) = 1$$

also $r(z) = c$, $c \in \mathbb{K}$. Folglich gilt:

$$f(\alpha) = 0 \iff r(\alpha) = q(\alpha)(\alpha - \alpha) + r(\alpha) = 0 \iff c = 0$$

$\square$

1.5.5 Folgerung: Ein Polynom $\neq 0$ vom Grad n hat höchstens n verschiedene Nullstellen.

Beweis durch Induktion über n: Der Fall $n = 1$ ist klar.

$n \to n+1$: Sei f ein Polynom vom Grad $n+1$. Falls f keine Nullstelle hat ist nichts zu zeigen. Andernfalls sei α eine Nullstellen von f. Nach dem obigen Lemma gibt es ein Polynom g vom Grad n mit

$$f(z) = (z - \alpha_1)g(z).$$

Nach der Induktionsvoraussetzung hat g höchstens n Nullstellen. Damit hat f höchstens $n+1$ Nullstellen, denn jede solche ist entweder $= \alpha$ oder Nullstelle von g. $\square$

1.5.6 Folgerung: Stimmen die Werte der Polynome

$$\begin{aligned} f(z) &= a_n z^n + \ldots + a_1 z + a_0 \\ g(z) &= b_n z^n + \ldots + b_1 z + b_0 \end{aligned}$$

an $n+1$-Stellen überein so gilt $a_k = b_k$ für $k = 0, \ldots, n$ und damit $f(z) = g(z) \quad \forall\ z \in \mathbb{K}$.

In diesem Fall hat das Polynom $h\colon = f - g$ vom Grad $\leq n$ nämlich $n+1$ Nullstellen und muss daher das Nullpolynom sein.

1.5.7 Nullstellen von quadratischen und kubischen Gleichungen:

1.5.8 Satz: Ein quadratisches Polynom $f(z) = az^2 + bz + c$ hat in $\mathbb{C}$ eine Nullstelle.

Beweis: Die Nullstellen sind $\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ (das folgt aus der *p-q-Formel*).

Wir wenden uns jetzt Polynomen vom Grad 3 zu:

$$f(z) = az^3 + bz^2 + cz + d \in \mathbb{C}[z]$$

Wir betrachten statt f das einfachere Polynom

$$p(z) = \frac{1}{a} f\left(z - \frac{b}{3a}\right) = z^3 + pz + q$$

mit geeigneten $p, q \in \mathbb{C}$. Es gilt:

α ist Nullstelle von $f \iff \alpha + \frac{b}{3a}$ ist Nullstelle von $p(z)$.

Sei $D := \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2$ und

$$u_\pm := \sqrt[3]{-\frac{q}{2} \pm \sqrt{D}}.$$

Dann ist $\alpha = u_+ + u_-$ Nullstelle von $p(z)$ (ohne Beweis; dies ist die *Formel von Cardano*).

Beispiel: Sei $p(z) :\ = z^3 - 15z - 4$.
$\Rightarrow D = -125 + 4 = -121$
$\Rightarrow u_\pm = \sqrt[3]{+2 \pm 11i} = 2 \pm i \Rightarrow \alpha = 4$.
(denn $(2+i)^3 = (3+4i)(2+i) = 2 + 11i$).
$\Rightarrow \alpha = 4$ ist Nullstelle von $p(z)$.
$\Rightarrow p(z) = (z-4)q(z)$
mit $q(z) = az^2 + bz + c$ für geeignete $a, b, c \in \mathbb{C}$. Aus einer leichten Rechnung ergibt sich $q(z) = z^2 + 4z + 1$. Damit sind $-2 \pm \sqrt{3}$ die weiteren Nullstellen von $p(z)$ und es gilt:

$$p(z) = (z-4)(z+2+\sqrt{3}))(z+2-\sqrt{3})).$$

Man kann auch noch für Polynome vom Grad 4 die Nullstellen explizit durch seine Koeffizienten berechnen. Für Polynome vom Grad ≥ 5 ist das i.a. nicht mehr möglich wie der norwegische Mathematiker N.H. Abel 1826 gezeigt hat.

Der folgende Satz garantiert die Existenz von Nullstellen von Polynomen über $\mathbb{C}$

1.5.9 Satz (Fundamentalsatz der Algebra): Jedes Polynom $\neq 0$ über $\mathbb{C}$ besitzt eine Nullstelle.

Der Beweis wird zu einem späteren Zeitpunkt nachgeholt.

1.5.10 Folgerung (Satz von der Linearfaktorzerlegung): Jedes nicht konstante Polynom $f \in \mathbb{C}[z]$ (d.h. deg $f > 0$) besitzt eine Darstellung

$$f(z) = a(z-\alpha_1) \cdot \ldots \cdot (z-\alpha_n)$$

Dabei sind die Faktoren $(z-\alpha_1), \ldots, (z-\alpha_n)$ bis auf die Reihenfolge eindeutig bestimmt.

Beweis: Ist α_1 eine Nullstelle von f so gilt: $f(z) = (z-\alpha_1)f_1$ wobei f_1 eine Polynom vom Grad deg $f - 1$ ist. Ist $deg\, f_1 > 0$ so schreiben wir wieder $f_1 = (z-\alpha_2)f_2$ wobei α_2 Nullstelle von f_1 ist, also $f = (z-\alpha_1)(z-\alpha_2)f_2$. So fortfahrend erhält man schliesslich die gesuchte Darstellung. $\square$

1.5.11 Folgerung: Jedes reelle, nicht konstante Polynom $f(z) \in \mathbb{R}[z]$, besitzt eine Darstellung

$$f(z) = a(z-\alpha_1) \cdot \ldots \cdot (z-\alpha_s)P_1(z) \cdot \ldots \cdot P_r(z),$$

wobei die $P_1(z), \ldots, P_r(z)$ Polynome der Form $x^2 + p_i x + q_i$ sind, für deren Distkriminante $p_i^2 - 4q_i < 0$ gilt.

2 Kapitel: Lineare Algebra

2.1 1. Lineare Gleichungssysteme

Eine lineare Gleichung ist eine Gleichung der Form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b.$$

Dabei sind $a_1, a_2, \ldots, a_n, b$ gegebene Zahlen (in einem festgewählten Zahlbereich; also etwa in $\mathbb{R}$ oder $\mathbb{C}$) und $x_1, x_2, \ldots, x_n$ gesuchte n Zahlen. Das Adjektiv linear bezieht sich darauf, dass nicht etwa Quadrate oder Produkte oder noch kompliziertere Funktionen der Unbekannten in der Gleichung vorkommen. Hat man nur wenige Unbekannte, so bezeichnet man sie auch mit $x, y, z, \ldots$ statt mit $x_1, x_2, x_3, \ldots$. Wir betrachten nun lineare Gleichungssysteme, also mehrere solcher Gleichungen, wie etwa

$$\begin{array}{rcrcrcr} 3x & - & 2y & + & 5z & = & 2 \\ & & y & + & 9z & = & 0. \end{array}$$

Für lineare Gleichungssysteme gibt es ein allgemeines Verfahren, einen Algorithmus, der

- zur Lösung führt, wenn es eine gibt,
- alle Lösungen liefert, wenn es mehrere gibt,
- gegebenenfalls meldet, dass es keine Lösung gibt.

Dieses Verfahren heisst Gaußscher Algorithmus. Wir führen ihn zunächst an zwei Beispielen vor.

2.1.1 Beispiel:

$$\begin{array}{rcrcr} 5x & + & 7y & = & 3 \\ 2x & + & 3y & = & -1 \end{array}$$

Wir multiplizieren die erste Zeile mit $-\frac{2}{5}$ und addieren sie zur zweiten und erhalten das lineare Gleichungssystem

$$\begin{array}{rcrcr} 5x & + & 7y & = & 3 \\ & & \frac{1}{5}y & = & \frac{-11}{5} \end{array}$$

Die Lösung der zweiten Gleichung ist $y = -11$. Setzen wir diesen Wert in die erste Gleichung ein und lösen diese nach x auf, so erhalten wir $x = 16$. Das Gleichungssystem hat also genau eine Lösung, nämlich $(x, y) = (16, -11)$.

2.1.2 Beispiel:

$$(3)\qquad \begin{array}{rcrcrcrl} 2x_1 & + & x_2 & + & 3x_3 & = & 0 & \\ 4x_1 & - & x_2 & + & x_3 & = & 10 & \qquad -2 \times \text{ erste Zeile} \\ 3x_1 & + & 2x_2 & + & x_3 & = & 7 & \qquad -\tfrac{3}{2} \times \text{ erste Zeile} \end{array}$$

$$(4)\qquad \begin{array}{rcrcrcrl} 2x_1 & + & x_2 & + & 3x_3 & = & 0 & \\ & - & 3x_2 & - & 5x_3 & = & 10 & \\ & & \frac{1}{2}x_2 & - & \frac{7}{2}x_3 & = & 7 & \qquad +\tfrac{1}{6} \times \text{ zweite Zeile} \end{array}$$

$$(5)\qquad \begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 3x_3 & = & 0 \\ & - & 3x_2 & - & 5x_3 & = & 10 \\ & & & & \frac{13}{3}x_3 & = & \frac{26}{3} \end{array}$$

Das dritte Gleichungssystem ist sehr leicht zu lösen. Aus der letzten Gleichung folgt $x_3 = -2$. Einsetzen in die vorletzte Gleichung liefert $x_2 = 0$, und Einsetzen der gefundenen Werte in die erste Gleichung des dritten Systems liefert $x_1 = 3$. Also hat das System die einzige Lösung $(x_1, x_2, x_3) = (3, 0, -2)$.

Beachten Sie, dass man bei den Umformungen keine Information verliert: Addiert man z.B. im Gleichungssystem (4) zur zweiten Zeile das Doppelte der erste Zeile und zur dritten Zeile das $\frac{3}{2}$-fache der ersten Zeile so erhält man das ursprüngliche Gleichungssystem (3) zurück. Alle Umformungen sind umkehrbar.

Wir beschreiben nun den Gaußschen Algorithmus im allgemeinen. Gegeben sei ein lineares Gleichungssystem

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

Wir nehmen an, dass x_1 wirklich vorkommt, d.h. eines der a_{i1} ist $\neq 0$. Sonst nummerieren wir die Unbekannten um. Nach Vertauschung von zwei Zeilen können wir annehmen, dass sogar $a_{11} \neq 0$. Dann addieren wir von den folgenden Zeilen ein Vielfaches der ersten, so dass jeweils x_1 verschwindet. Genauer: wir addieren zur i-ten Zeile das $-\frac{a_{i1}}{a_{11}}$-fache der ersten. Wir erhalten ein neues, zum ersten System äquivalentes, in dem x_1 nur noch in der ersten Zeile vorkommt:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & a'_{22}x_2 & + & \dots & + & a'_{2n}x_n & = & b'_2 \\ & & \vdots & & & & \vdots & & \vdots \\ & & a'_{m2}x_2 & + & \dots & + & a'_{mn}x_n & = & b'_m \end{array}$$

Noch einmal die vorgenommen Umformungen:

- Vertauschung von zwei Zeilen,
- Addition eines Vielfachen einer Zeile zu einer anderen.

Bei den weiteren Umformungen des Systems betrachten wir nur noch die Zeile ab der zweiten, die erste Zeile führen wir unverändert mit. Wir betrachten in den folgenden Zeilen die Unbekannte mit dem kleinsten noch vorkommenden Index. In der Regel wird das x_2 sein, es könnte aber auch sein, dass alle Terme in denen x_2 vorkam, bei den Umformungen verschwunden sind. Nehmen wir an, x_j wäre diese Unbekannte. Nach Zeilenvertauschung können wir wieder annehmen, dass x_j in der zweiten Zeile wirklich vorkommt. Dann addieren wir wie eben zu den Zeilen $3, \ldots, m$ jeweils ein geeignetes Vielfaches der zweiten Zeile, so dass x_j in keiner der Zeilen $3, \ldots, m$ mehr vorkommt.

Dieses Verfahren setzen wir fort, bis in den nachfolgenden Zeilen gar keine Unbekannten mehr stehen. Das ist insbesondere dann der Fall, wenn es keine nachfolgenden Zeilen mehr gibt. Es kann aber auch sein, dass noch Gleichungen da sind, die aber auf der linken Seite keine Unbekannten mehr haben also von der Form

$$0 \quad = \quad b$$

für irgendeine rechte Seite b sind. Das System ist dann in *Zeilenstufenform.*

2.1.3 Definition: Ein lineares Gleichungssystem

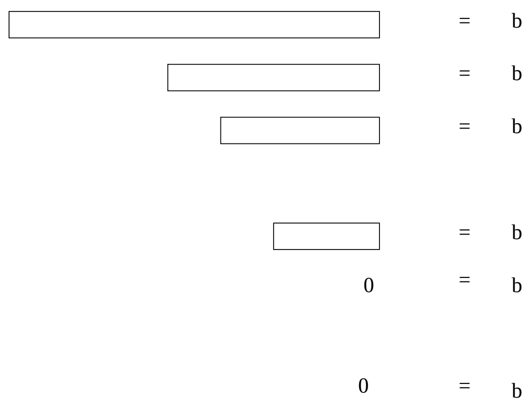
$$(*) \qquad \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

ist in Zeilenstufenform, falls für jeden Zeilenindex $i = 1, 2, \ldots, m$ ein Index $j \in \{1, \ldots, n\}$ existiert, so daß

(1) $a_{ij} \neq 0$ und $a_{ik} = 0 \quad \forall\ k = 1, \ldots, j-1$

(2) $a_{kl} = 0 \quad \forall\ k > i, l \leq j$.

Ein System in Zeilenstufenform sieht also so aus:



Da der Index j in (1) von i abhängt schreiben wir auch j_i dafür. Wegen (2) gilt:

$$j_1 < j_2 < \ldots < j_r \leq n.$$

Wir wollen jetzt untersuchen wann das System $(*)$ in Zeilenstufenform lösbar ist und beschreiben wie die Lösungsmenge gegebenenfalls aussieht.

1. Fall: $r < m$ und mindestens eine der Zahlen $b_{r+1}, b_{r+2}, \ldots, b_m$ ist $\neq 0$. Dann besitzt das System $(*)$ keine Lösung.

2. Fall: $b_{r+1} = \ldots = b_m = 0$ oder $r = m$.

Dann können wir die letzten $m - r$ Zeilen ignorieren, sie enthalten keine Information.

Das Gleichungssystem ist in diesem Fall lösbar.

Fall 2, (a): $r = n$. Dann sind alle Stufen von der Breite 1, also $j_i = i$ für $i = 1, 2, \ldots, n$. Die letzte Gleichung ist $a_{nn}x_n = b_n$ und wir erhalten daraus einen eindeutig bestimmten Wert für x_n. Diesen setzen wir in die vorletzte Gleichung

$$a_{n-1\,n-1}x_{n-1} + a_{n-1\,n}x_n = b_{n-1}$$

ein und erhalten daraus den Wert für x_{n-1}. So fortfahrend sehen wir, daß das Gleichungssystem eine eindeutig bestimmte Lösung $(x_1, x_2, \ldots, x_n)$ besitzt.

Fall 2, (b): $r < n$. Dann gibt es wenigstens eine breitere Stufe d.h. es gibt mindestens eine Variable x_j die nicht am Anfang einer Zeile steht. Eine solche Variable nennnen wir “freie Variable”. Davon gibt es offenbar genau $n - r$ Stück. Wir können sie beliebig wählen und für jede solche Wahl erhält man eine eindeutig bestimmte Lösung.

Wir führen den Algorithmus noch einmal an einem Beispiel vor. Gegeben sei das Gleichungssystem

$$
(*) \qquad \begin{array}{rcrcrcl}
 & & x_2 & + & 2x_3 & = & 9 \\
3x_1 & + & 4x_2 & + & 5x_3 & = & 9 \\
6x_1 & + & 7x_2 & + & 8x_3 & = & 9 \\
9x_1 & + & 9x_2 & + & 9x_3 & = & 9.
\end{array}
$$

Wir benutzten dafür die abkürzende Schreibweise

$$
\begin{array}{ccc|c}
0 & 1 & 2 & 9 \\
3 & 4 & 5 & 9 \\
6 & 7 & 8 & 9 \\
9 & 9 & 9 & 9
\end{array}
$$

Als ersten Schritt vertauschen wir die 1. und 2. Zeile:

$$
\begin{array}{ccc|cl}
3 & 4 & 5 & 9 & \\
0 & 1 & 2 & 9 & \\
6 & 7 & 8 & 9 & \qquad \text{-2}\times \text{ 1. Zeile} \\
9 & 9 & 9 & 9 & \qquad \text{-3}\times \text{ 1. Zeile}
\end{array}
$$

$$\Updownarrow$$

$$
\begin{array}{cccc|l}
3 & 4 & 5 & 9 & \\
0 & 1 & 2 & 9 & \\
0 & -1 & -2 & -9 & \qquad 1\times \text{ 2. Zeile} \\
0 & -3 & -6 & -18 & \qquad 3\times \text{ 2. Zeile}
\end{array}
$$

$$\Updownarrow$$

$$
\begin{array}{ccc|c}
3 & 4 & 5 & 9 \\
0 & 1 & 2 & 9 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 9
\end{array}
$$

Also ist das System nicht lösbar.

Wenn wir stattdessen die letzte Zeile von $(*)$ durch die Gleichung

$$(*) \qquad 9x_1 + 9x_2 + 9x_3 = 0$$

ersetzen so erhalten wir am Ende auf das Diagramm

$$\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

gestoßen, d.h. in diesem Fall ist das System äquivalent zu dem Gleichungssystem

$$(**) \qquad \begin{array}{rcrcrcr} 3x_1 & + & 4x_2 & + & 5x_3 & = & 9 \\ & & x_2 & + & 2x_3 & = & 9 \end{array}$$

Wir erhalten unendlich viele Lösungen, da wir x_3 frei wählen können, d.h. zu jeder Wahl von x_3 gibt es genau eine Lösung. Als Lösungsmenge von $(**)$ erhalten wir

$$\{(t-9, 9-2t, t) \mid t \in \mathbb{R}\}.$$

2.1.4 Definition: 1. Ein lineares Gleichungssystem der Form

$$\begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & 0 \\ \vdots & & & & & \vdots & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array}$$

heißt homogen.

2. Setzt man in einem beliebigen linearen Gleichungssystem

$$\begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & & \vdots & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & b_n \end{array}$$

die rechte Seite $= 0$, so erhält man ein neues lineares Gleichungssystem, das das zugehörige homogene System genannt wird.

2.1.5 Bemerkung: Hat das homogene lineare Gleichungssystem

$$\begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & 0 \\ \vdots & & & & & \vdots & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array}$$

mehr Unbekannte als Gleichungen, d.h. ist $m < n$, so besitzt das System eine nicht-triviale Lösung

$$(x_1, \ldots, x_n) \neq (0, \ldots, 0).$$

Beweis: Wir bringen das System mit dem Gaußalgorithmus auf Zeilenstufenform mit r nicht-trivialen Zeilen. Da $r \leq m < n$ gibt es mindestens eine "freie Variablen". Sei x_j eine solche. Nach den obigen Überlegungen gibt es eine Lösung $(x_1, \ldots, x_n)$ des Systems bei der wir x_j beliebig vorgeben können (also etwa $x_j = 1$). $\square$

2.2 Vektorräume

Sei K ein Körper (etwa $K = \mathbb{R}$ oder $= \mathbb{C}$).

2.2.1 Definition: Ein K-Vektorraum ist ein Tripel $(V, +, \cdot)$ bestehend aus einer nichtleeren Menge V einer Verknüpfung (genannt Addition)

$$+ : V \times V \to V,$$

und einer Abbildung (genannt skalare Multiplikation)

$$\cdot : K \times V \to V$$

so daß gilt:

(V1) $x + y = y + x \qquad \forall\ x, y \in V$;

(V2) $x + (y + z) = (x + y) + z \qquad \forall\ x, y, z \in V$;

(V3) Es gibt ein Element $0 = 0_V \in V$ (genannt Nullvektor)

$$0 + x = x + 0 = x \qquad \forall\ x \in V;$$

(V4) Für alle $x \in V$ gibt es ein $-x \in V$, so dass

$$x + (-x) = 0;$$

(V5) $1 \cdot x = x \qquad \forall\ x \in V$;

(V6) $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x \qquad \forall\ \lambda, \mu \in K, x \in V$;

(V7) $\lambda(x + y) = \lambda x + \lambda y \qquad \forall\ \lambda \in K, x, y \in V$;

(V8) $(\lambda + \mu)x = \lambda x + \mu x \qquad \forall\ \lambda, \mu \in K, x \in V$.

Oft schreibt man auch einfach V anstatt $(V, +, \cdot)$. Die Axiome (V1) – (V4) bedeuten gerade, dass $(V, +)$ eine abelsche Gruppe ist. Die Elemente in V werden als Vektoren bezeichnet. Ist $K = \mathbb{R}$ (bzw. $= \mathbb{C}$), so spricht man auch von einem reellen (bzw. komplexen) Vektorraum.

Abkürzend schreibt man: $x - y :\ = x + (-y)$

2.2.2 Bemerkung: Sei V ein K-Vektorraum. Dann gilt:

(a) $0 \cdot x = 0_V \qquad \forall\, x \in V$

(b) $\lambda \cdot 0_V = 0_V \qquad \forall\, \lambda \in K$

(c) $(-1) \cdot x = -x \qquad \forall\, x \in V.$

Beweis: Es gilt

$$\begin{aligned} 0 \cdot x &\overset{(V3)}{=} 0 \cdot x + 0_V \overset{(V4)}{=} 0 \cdot x + (0 \cdot x - (0 \cdot x)) = \\ &\overset{(V2)}{=} (0 \cdot x + 0 \cdot x) - (0 \cdot x) \overset{(V8)}{=} (0+0) \cdot x - (0 \cdot x) = 0_V \end{aligned}$$

und damit (a). Der Beweis von (b) verläuft ganz analog. (c) ergibt sich wie folgt:

$$\begin{aligned} -x &\overset{(a)}{=} -x + (1 + (-1))x = -x + 1x + (-1)x = \\ &= -x + x + (-1)x = 0_V + (-1)x = (-1)x. \end{aligned}$$

□

2.2.3 Beispiele: (a) Das Standardbeispiel für einen Vektorraum ist die Menge $K^n = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in K\}$ mit der Addition und skalaren Multiplikation

$$\begin{aligned} (x_1, \ldots, x_n) + (y_1, \ldots, y_n) :&= (x_1 + y_1, \ldots, x_n + y_n) \\ \lambda \cdot (x_1, \ldots, x_n) :&= (\lambda x_1, \ldots, \lambda x_n). \end{aligned}$$

(b) Ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$, so ist die Menge der Polynome $\mathbb{K}[z]$ über $\mathbb{K}$ ein Vektorraum. Dabei ist die Addition und skalare Multiplikation definiert durch

$$f + g : \mathbb{K} \to \mathbb{K}, z \mapsto (f+g)(z) : = f(z) + g(z) \tag{6}$$

$$\lambda f : \mathbb{K} \to \mathbb{K}, z \mapsto (\lambda \cdot f)(z) : = \lambda f(z) \tag{7}$$

Ebenso ist $\{f \in \mathbb{K}[z] \mid \deg f \leq n\}$ ein $\mathbb{K}$-Vektorraum

(c) Sei M eine Menge und K ein Körper. Die Menge $\mathrm{Abb}\,(M, K)$ der Abbildungen $f : M \to K$ ist ein K-Vektorraum. Dabei ist die Addition und Skalramultiplikation wieder durch (7) definiert, d.h. es gilt $(f+g)(m) : = f(m) + g(m)$ und $(\lambda f)(m) : = \lambda f(m)$ für alle $f, g \in \mathrm{Abb}\,(M, K)$ und $\lambda \in K$.

2.2.4 Definition: Sei V ein K-Vektorraum. Eine Teilmenge U von V heisst Untervektorraum, wenn $U \neq \emptyset$ und für alle $x, y \in U$, und alle $\lambda \in K$ gilt:

$$x + y, \quad \lambda x \in U.$$

2.2.5 Beispiele: (a) Die Untervektorräume von $\mathbb{R}^2$ sind $\{0\}$, $\mathbb{R}^2$ und die Geraden durch den Nullpunkt $(0,0)$, d.h. Mengen der Form $\{(x,y) \in \mathbb{R}^2 \mid ax + by = 0\}$ wobei a und b reelle Zahlen sind die nicht beide $= 0$ sind.

(b) Sei K ein Körper und

$$(*) \qquad \begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & 0 \\ \vdots & & & & & \vdots & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array}$$

ein lineares Gleichungssystem (über K). Dann ist die Lösungsmenge U von $(*)$, d.h. die Menge der n-Tupel reeller Zahlen $(x_1, \ldots, x_n)$ die $(*)$ erfüllen ein Untervektorraum von K^n.

(c) Sei V der $\mathbb{R}$-Vektorraum der Funktionen $f : \mathbb{R} \to \mathbb{R}$. Sei

$$U := \{f \in V \mid f \text{ ist } 2\times \text{ stetig differenzierbar und } f'' + f = 0\}.$$

Dann ist U ein Untervektorraum von V. Zum Beispiel sind $\sin(x)$ und $\cos(x)$ Elemente von U, da $\sin' = \cos$ und $\cos' = -\sin$.

Allgemeiner ist der Lösungsraum einer homogenen linearen Differentialgleichung auf $\mathbb{R}$ ein Untervektorraum von V.

2.2.6 Bemerkung: Sei $(V, +, \cdot)$ ein K-Vektorraum und $U \subseteq V$ ein Untervektorraum. Dann ist U zusammen mit der induzierten Addition und Skalarmultiplikation wieder ein Vektorraum, (d.h. $(U, +\,|_{U\times U}, \cdot\,|_{K\times U})$ ist ein K-Vektorraum).

Beweis: Alle Axiome bis auf (V3) und (V4) folgen sofort aus der Tatsache, daß sie für V gelten.

Für $u \in U$ gilt

$$-u = (-1) \cdot u \in U,$$

und folglich

$$0_V = u + (-u) \in U.$$

Da $U \neq \emptyset$ folgen (V3) und (V4). $\square$

2.2.7 Bemerkung: Sind U_1, U_2 Untervektorräume von V, so ist auch $U_1 \cap U_2$ ein Untervektorraum.

Als nächstes diskutieren wir die Begriffe *lineare Unabhängigkeit, Basis* und *Dimension*.

Sei V ein K-Vektorraum. Seien $v_1, \ldots, v_r \in V$ Vektoren. Ein Vektor der Form

$$\lambda_1 v_1 + \ldots + \lambda_r v_r$$

mit $\lambda_1, \ldots, \lambda_r \in K$, heißt *Linearkombination* der Vektoren $v_1, \ldots, v_r$.

2.2.8 Definition: Die Menge aller Linearkombinationen von $v_1, \ldots, v_r$

$$L(v_1, \ldots, v_r) = \{\lambda_1 v_1 + \ldots + \lambda_r v_r \mid \lambda_1, \ldots, \lambda_r \in K\}$$

heißt lineare Hülle von $(v_1, \ldots, v_r)$ (oder auch der von den Vektoren $v_1, \ldots, v_r$ aufgespannten oder erzeugten Raum).

2.2.9 Bemerkung: $L(v_1, \ldots, v_r)$ ist ein Untervektorraum von V.

Beweis: Die Summe zweier Linearkombinationen und skalare Vielfache einer Linearkombination ist wieder eine Linearkombination denn

$$(\lambda_1 v_1 + \ldots + \lambda_r v_r) + (\mu_1 v_1 + \ldots + \mu_r v_r) = (\lambda_1 + \mu_1) v_1 + \ldots + (\lambda_r + \mu_r) v_r$$
$$\lambda(\lambda_1 v_1 + \ldots + \lambda_r v_r) = (\lambda\lambda_1) v_1 + \ldots + (\lambda\lambda_r) v_r.$$

$\square$

2.2.10 Beispiele: (a) Sei $v = (x, y) \in \mathbb{R}^2$ ein Vektor $\neq 0$. Dann ist $L(v)$ die Gerade in $\mathbb{R}^2$ durch die Punkte $(0, 0)$ und (x, y).

(b) In K^n betrachten wir die Vektoren

$$e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)$$

Der Vektor e_i hat also an der i-ten Stelle eine 1 und sonst lauter Nullen. Dann gilt für einen beliebigen Vektor $x = (x_1, \ldots, x_n) \in K^n$:

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Jeder Vektor in K^n ist also Linearkombination von $e_1, \ldots, e_n$, d.h. es gilt $L(e_1, \ldots, e_n) = K^n$.

(c) Das Polynom $x^3 \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ ist Linearkombination der Funktionen $(x-1)^k$, $k = 0, 1, 2, 3$. Es gilt nämlich

$$x^3 = ((x-1)+1)^3 = 1 \cdot (x-1)^3 + 3(x-1)^2 + 3(x-1) + 1 \cdot (x-1)^0$$

Umgekehrt ist die Funktion $(1+x)^n$ eine Linearkombination der *Monome* $1, x, x^2, \ldots, x^n$, denn

$$(1+x)^n = \binom{n}{0} 1 + \binom{n}{1} x^1 + \ldots + \binom{n}{n-1} x^{n-1} + \binom{n}{n} x^n.$$

Die Aufgabe einen (Unter-)Vektorraum durch möglichst wenige Vektoren zu erzeugen führt zum Begriff der *linearen Unabhängigkeit*.

2.2.11 Definition: Sei V ein K-Vektorraum. Ein k-Tupel von Vektoren $(v_1, \ldots, v_k)$ heißt linear unabhängig, falls für $\lambda_1, \ldots, \lambda_k \in K$ gilt:

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0 \Rightarrow \lambda_1 = \ldots = \lambda_k = 0.$$

Das Tupel $(v_1, \ldots, v_k)$ heisst linear abhängig, falls es nicht linear unabhängig ist.

Anders ausgedrückt: $(v_1, \ldots, v_k)$ ist linear abhängig, falls es eine Darstellung des Nullvektors als Linearkombination $0 = \lambda_1 \cdot v_1 + \ldots + \lambda_k \cdot v_k$ gibt, wobei die $\lambda_1, \ldots, \lambda_k \in K$ nicht alle gleich Nulll sind.

2.2.12 Beispiele: (a) Die Vektoren $x = (1, 0, 1)$, $y = (-1, 2, 1)$, $z = (0, 2, 2)$ sind linear abhängig, denn

$$3x + 3y - 3z = 0$$

(b) Die Vektoren $e_1, \ldots, e_n$ von K^n sind linear unabhängig. Aus

$$\lambda_1 \cdot e_1 + \ldots + \lambda_n \cdot e_n = 0$$

folgt nämlich $(\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0)$, also $\lambda_1 = \ldots = \lambda_n = 0$.

(c) Ein einzelner Vektor v ist genau dann linear unabhängig, wenn $v \neq 0$.

(d) Zwei Vektoren v_1, v_2 sind linear unabhängig, wenn keiner Vielfaches des anderen ist.

(e) Die Monome $1 = z^0, z, z^2, \ldots, z^n$ in $\mathbb{K}[z]$ sind linear unabhängig. Dies ergibt sich aus der Folgerung auf Seite 31.

(f) Drei Vektoren in $\mathbb{R}^3$ sind linear abhängig, wenn sie in einer Ebene liegen.

2.2.13 Lemma (Kriterien für lineare Unabhängigkeit): Sei V ein K-Vektorraum und $(v_1, \ldots, v_k)$ ein k-Tupel von Vektoren. Die folgenden Bedingungen sind äquivalent:

(i) $(v_1, \ldots, v_k)$ sind linear unabhängig.

(ii) Jeder Vektor in $L(v_1, \ldots, v_k)$ läßt sich in eindeutiger Weise als Linearkombination der $v_1, \ldots, v_k$ schreiben.

(iii) Keiner der Vektoren $v_1, \ldots, v_k$ ist Linearkombination der anderen.

Beweis: (i) $\Rightarrow$ (ii): Sei $v \in L(v_1, \ldots, v_k)$ und seinen

$$v = \lambda_1 v_1 + \ldots + \lambda_k v_k = \mu_1 v_1 + \ldots + \mu_k v_k$$

zwei Darstellungen von v als Linearkombination von $(v_1, \ldots, v_k)$.

$$\begin{array}{ll} \Rightarrow & 0 = x - x = (\lambda_1 - \mu_1)v_1 + \ldots + (\lambda_k - \mu_k)v_k \\ \overset{(i)}{\Rightarrow} & \lambda_1 - \mu_1 = \ldots = \lambda_k - \mu_k = 0 \\ \Rightarrow & \lambda_1 = \mu_1, \ldots, \lambda_k = \mu_k. \end{array}$$

(i) $\Rightarrow$ (iii): Angenommen es existiert ein Index $i \in \{1, 2, \ldots, k\}$, so dass v_i Linearkombination von $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ ist. Sei $v_i = \lambda_1 v_1 + \ldots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \ldots + \lambda_k v_k$ eine solche. Dann erhält man eine nicht-triviale Darstellung des Nullvektors als Linearkombination von $(v_1, \ldots, v_k)$ nämlich

$$0 = \lambda_1 v_1 + \ldots + \lambda_{i-1} v_{i-1} + (-1)v_i + \lambda_{i+1} v_{i+1} + \ldots + \lambda_k v_k$$

im Widerspruch zur linearen Unabhängigkeit von $(v_1, \ldots, v_k)$.

(iii) $\Rightarrow$ (i): Seien $\lambda_1, \ldots, \lambda_k \in K$ mit

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0.$$

Angenommen $\lambda_i \neq 0$ für einen Index $i \in \{1, \ldots, k\}$. Dann folgt

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \ldots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \ldots - \frac{\lambda_k}{\lambda_i} v_k$$

im Widerspruch zur Annahme (iii). $\square$

2.2.14 Beispiel (Unabhängigkeitstest in K^m): Seien $x_1 = (x_{11}, \ldots, x_{n1}), \ldots, x_n = (x_{n1}, \ldots, x_{nm})$ n Vektoren in K^m. Um diese auf lineare Unabhängigkeit zu testen betrachten wir

$$\lambda_1 x_1 + \ldots + \lambda_n x_n = 0.$$

Das ist äquivalent zum linearen Gleichungssystem

$$(*) \qquad \begin{array}{ccccccc} x_{11}\lambda_1 & + & \ldots & + & x_{n1}\lambda_n & = & 0 \\ \vdots & & & & & & \vdots \\ x_{m1}\lambda_1 & + & \ldots & + & x_{mn}\lambda_n & = & 0 \end{array}$$

Die Vektoren sind also genau dann linear unabhängig genau, wenn $(*)$ nur die Lösung $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$ besitzt. Zum Beispiel sind mehr als m Vektoren im K^m stets linear abhängig.

2.2.15 Definition: Sei V ein K-Vektorraum und $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren von V.

1. $(v_1, \ldots, v_n)$ heißt Erzeugendensystem, wenn jeder Vektor von V Linearkombination von $v_1, \ldots, v_n$ ist, d.h. wenn $V = L(v_1, \ldots, v_n)$

2. $(v_1, \ldots, v_n)$ heißt Basis, wenn $(v_1, \ldots, v_n)$ ein Erzeugendensystem und linear unabhängig ist.

3. V heißt endlich erzeugt, wenn es endlich viele Vektoren $v_1, \ldots, v_n$ gibt die ein Erzeugendensystem bilden.

2.2.16 Bemerkung: Ist $(b_1, \ldots, b_n)$ ein Erzeugendensystem von V, so lässt sich jeder Vektor $v \in V$ schreiben als

$$v = \lambda_1 v_1 + \ldots + \lambda_n b_n$$

für gewisse $\lambda_1, \ldots, \lambda_n \in K$. Ist $(b_1, \ldots, b_n)$ eine Basis, so sind $\lambda_1, \ldots, \lambda_n$ nach dem Kriterium (ii) des obigen Lemmas eindeutig bestimmt.

2.2.17 Beispiel: (a) $(e_1, \ldots, e_n)$ ist eine Basis von K^n. Sie heisst Standardbasis des K^n.

(b) Sei $\mathbb{K} = \mathbb{R}$ oder $= \mathbb{C}$. Der $\mathbb{K}$-Vektorraum $\mathbb{K}[z]$ der Polynome über $\mathbb{K}$ ist nicht endlich erzeugt.

Betrachten wir stattdessen den Untervektorraum der Polynome vom Grad $\leq n$:

$$\mathbb{K}[z]_{\deg \leq n} = \{f(z) \in \mathbb{K}[z] \mid \deg f(z) \leq n\}.$$

Die Monome $1 = z^0, z, z^2, \ldots, z^n$ bilden eine Basis von $\mathbb{K}[z]_{\deg \leq n}$.

(c) Sei V ein reeller Vektorraum mit Basis (v_1, v_2). Dann ist auch (v_1+v_2, v_1-v_2) eine Basis.

2.2.18 Lemma: Es seien $k+1$ Vektoren $v_1, \ldots, v_k, v_{k+1}$ in V gegeben. Sind $(v_1, \ldots, v_k)$ linear unabhänig und ist $v_{k+1} \in V \setminus L(v_1, \ldots, v_k)$ so ist $(v_1, \ldots, v_k, v_{k+1})$ linear unabhängig.

Beweis: Seien $\lambda_1, \ldots, \lambda_{k+1} \in K$ mit

$$\lambda_1 v_1 + \ldots + \lambda_{k+1} v_{k+1} = 0. \tag{8}$$

Dann ist zunächst $\lambda_{k+1} = 0$, denn wäre $\lambda_{k+1} \neq 0$, so könnte man die Gleichung nach v_{k+1} auflösen

$$v_{k+1} = -\frac{\lambda_1}{\lambda_{k+1}} v_1 - \ldots - \frac{\lambda_k}{\lambda_{k+1}} v_k$$

und v_{k+1} wäre Linearkombination von $(v_1, \ldots, v_k)$ im Widerspruch zur Voraussetzung.

Also können wir den Summanden $\lambda_{k+1} v_{k+1}$ in (8) weglassen und erhalten

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0.$$

Da $(v_1, \ldots, v_k)$ nach Voraussetzung linear unabhänig ist, folgt $\lambda_1 = \ldots = \lambda_k = 0$. $\square$

2.2.19 Lemma (Schrankenlemma): Sei V ein K-Vektorraum und $n \in \mathbb{N}$. Dann sind die folgenden beiden Aussagen äquivalent:

(i) V wird von n Vektoren erzeugt.

(ii) Je $n+1$ Vektoren von V sind linear abhängig.

Beweis: Ist $V = \{0\}$ so ist nichts zu zeigen (beide Aussagen sind erfüllt). Wir nehmen also an, dass $V \neq \{0\}$.

(i) $\Rightarrow$ (ii): Seien $v_1, \ldots, v_{n+1} \in V$. Wir wollen zeigen, dass es $\lambda_1, \ldots, \lambda_{n+1} \in K$ gibt, die nicht alle $= 0$ sind, so dass

$$\lambda_1 v_1 + \ldots + \lambda_{n+1} v_{n+1} = 0.$$

Nach Voraussetzung gibt es ein Erzeugendensystem $(b_1, \ldots, b_n)$ von V und damit $a_{ij} \in K, i = 1, \ldots, n, j = 1, \ldots, n+1$, so dass

$$v_j = a_{1j} b_1 + \ldots + a_{nj} b_n = \sum_{i=1}^{n} a_{ij} b_i \qquad \text{für } j = 1, \ldots, n+1.$$

Damit ist

$$\sum_{j=1}^{n+1} \lambda_j v_j = \sum_{j=1}^{n+1} \lambda_j \left(\sum_{i=1}^{n} a_{ij} b_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n+1} a_{ij} \lambda_j \right) b_i.$$

Es genügt also zu zeigen, dass es Werte λ_j gibt, die nicht alle $= 0$ sind, so dass

$$\sum_{j=1}^{n+1} a_{ij} \lambda_j \qquad \text{für alle } i = 1, \ldots, n.$$

Das ist ein homogenes lineares Gleichungssystem mit n Gleichungen für $n+1$ Unbekannte. Es hat also immer eine nicht-triviale Lösung.

(ii) $\Rightarrow$ (i): Wir wählen uns ein linear unabhängiges Tupel $(b_1, \ldots b_r)$ von Vektoren von grösstmöglicher Länge r. Das existiert nach Voraussetzung und es ist $1 \leq r \leq n$. Sei $U := L(b_1, \ldots, b_r)$. Ist $U = V$, so sind wir fertig. Andernfalls gibt es ein $b_{r+1} \in V \setminus U$. Nach dem obigen Lemma ist dann $(b_1, \ldots b_r, b_{r+1})$ linear unabhängig. Das ist ein Widerspruch zur Wahl von r. $\square$

2.2.20 Satz (Basissatz): Sei V ein endlich erzeugter K-Vektorraum. Dann gilt:

(i) V besitzt eine Basis. Genauer: Ist $(v_1, \ldots, v_m)$ ein Erzeugendensystem, so gibt es eine Basis $(b_1, \ldots, b_n)$ mit $b_i \in \{v_1, \ldots, v_m\}$ für alle $i \in \{1, \ldots, n\}$, d.h. man erhält eine Basis aus $(v_1, \ldots, v_m)$ indem man gegebenenfalls einige der Vektoren weglässt.

(ii) Je zwei Basen haben die gleiche Anzahl von Elementen.

(iii) (Ergänzungssatz) Sind $b_1, \ldots, b_r \in V$ linear unabhängig und keine Basis von V, so gibt es Vektoren $b_{r+1}, \ldots, b_n$, so dass $(b_1, \ldots, b_n)$ eine Basis von V ist.

Beweis: Zu (i): Sei $(b_1, \ldots, b_n)$ ein Erzeugendensystem von V mit minimalem n, gebildet aus Vektoren in $\{v_1, \ldots, v_m\}$. Wir wollen zeigen, dass $(b_1, \ldots, b_n)$ linear unabhängig ist und damit eine Basis. Es genügt zu zeigen dass keiner der Vektoren $b_1, \ldots, b_n$ Linearkombination der anderen ist (siehe dazu die "Kriterien für lineare Unabhängigkeit").

Wäre z.B. b_n Linearkombination von $(b_1, \ldots, b_{n-1})$, also $b_n \in L(b_1, \ldots, b_{n-1})$, so würde für eine beliebige Linearkombination $\sum_{i=1}^{n} \lambda_i b_i$ von $b_1, \ldots, b_n$ gelten

$$\sum_{i=1}^{n} \lambda_i b_i = (\lambda_1 b_1 + \ldots + \lambda_{n-1} b_{n-1}) + \lambda_n b_n \in L(b_1, \ldots, b_{n-1})$$

da $L(b_1, \ldots, b_{n-1})$ ein Untervektorraum ist. Also würde folgen

$$V = L(b_1, \ldots, b_n) \subseteq L(b_1, \ldots, b_{n-1}),$$

d.h. $(b_1, \ldots, b_{n-1})$ wäre ein Erzeugendensystem. Das steht im Widerspruch zur Minimalität von n.

Zu (ii): Seien $(b_1, \ldots, b_n)$ und $(b'_1, \ldots, b'_{n'})$ zwei Basen von V. Wir wollen zeigen, dass $n = n'$. Da $(b_1, \ldots, b_n)$ ein Erzeugendensystem ist, ist $n' \leq n$, den das n'-Tupel $(b'_1, \ldots, b'_{n'})$ wäre sonst nach dem Schrankenlemma linear abhängig. Analog, indem man die Rollen von $(b_1, \ldots, b_n)$ und $(b'_1, \ldots, b'_{n'})$ vertauscht, ergibt sich $n \leq n'$, also $n = n'$.

Zu (iii): V sei erzeugt von n Elementen. Sind $b_1, \ldots, b_r$ linear unabhängig aber keine Basis, so sind sie kein Erzeugendensystem, d.h. $L(b_1, \ldots, b_r) \neq V$. Dann gibt es ein $b_{r+1} \in V \setminus L(b_1, \ldots, b_r)$. Nach einem der vorangegangenen Lemmata ist $(b_1, \ldots, b_r, b_{r+1})$ linear unabhängig. Diesen Prozess können wir fortsetzen, solange wir nicht bei einem Erzeugendensystem angekommen sind. Aber nach dem Schrankenlemma geht das allenfalls, solange $r < n$ ist. Also bricht der Prozess ab, und wir erhalten eine Basis der gesuchten Form. □

2.2.21 Definition: Sei V ein K-Vektorraum. Wir definieren die Dimension von V

$$\dim(V) \in \{0, 1, 2, \ldots\} \cup \{\infty\}$$

wie folgt:

1. Ist $V = \{0\}$, so sei $\dim V = 0$.
2. Ist $V \neq \{0\}$ endlich erzeugt und $(b_1, \ldots, b_n)$ eine Basis von V, so sei

$$\dim V = n.$$

3. Ist V nicht endlich erzeugt, so sei

$$\dim V = \infty.$$

Beispielsweise ist $\dim(K^n) = n$. Für den reellen Vektorraum $\mathbb{R}[x]$ aller Polynome über $\mathbb{R}$ gilt $\dim(\mathbb{R}[x]) = \infty$ und $\dim\{f(x) \in \mathbb{R}[x] \mid \deg f(x) \leq n\}) = n + 1$.

Ein Vektorraum V heisst endlich-dimensional falls $\dim(V) < \infty$, also wenn V endlich erzeugt ist. Im Folgenden betrachten wir haupsächlich endlich-dimensionale Vektorräume. Als Folgerung aus dem Basissatz erhalten wir:

2.2.22 Folgerung: Sei V ein endlich-dimensionaler Vektorraum der Dimension n und $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren V. Dann sind die folgenden Bedingungen äquivalent:

(i) $(v_1, \ldots, v_n)$ ist eine Basis.

(ii) $(v_1, \ldots, v_n)$ ist linear unabhängig.

(iii) $(v_1, \ldots, v_n)$ ist ein Erzeugendensystem.

Beweis: (ii) $\Rightarrow$ (i): Wäre $(v_1, \ldots, v_n)$ keine Basis, so könnte man es nach dem Ergänzungssatz zu einer Basis auffüllen, die dann aber mehrals n Elemente enthielte im Widerspruch zu $\dim(V) = n$.

(iii) $\Rightarrow$ (i): Wäre $(v_1, \ldots, v_n)$ keine Basis, so würde aus $(v_1, \ldots, v_n)$ durch Weglassen einiger Vektoren eine Basis mit weniger als n Elementen entstehen. Das ist wieder ein Widerspruch zu $\dim(V) = n$. □

2.2.23 Folgerung: (Dimension von Unterräumen) Sei $U \subseteq V$ ein Untervektorraum des endlich-dimensionalen Vektorraums V. Dann ist auch U endlich-dimensional und

$$\dim U \leq \dim V.$$

Gleichheit gilt genau dann, wenn $U = V$ ist.

Beweis: Sei $n := \dim(V)$. Nach dem Schrankenlemma sind dann je $n + 1$ Vektoren in V linear abhängig. Also sind auch je $n+1$ Vektoren aus U linear abhängig. Nach dem Schrankenlemma (angewendet auf U) hat U daher ein Erzeugendensystem aus n Elementen und folglich, nach Teil (i) des Basissatz, eine Basis aus höchstens n Vektoren. Das zeigt, dass $\dim(U) \leq \dim(V)$.

Ist $\dim(U) = \dim(V) = n$ und $(v_1, \ldots, v_n)$ eine Basis von U, so ist $(v_1, \ldots, v_n)$, aufgrund der Implikation (ii) $\Rightarrow$ (i) in der obigen Folgerung 1, auch eine Basis von V und damit $U = L(v_1, \ldots, v_n) = V$. $\square$

Als Beispiel wollen wir eine Basis der Lösungsmenge $U \subseteq \mathbb{R}^4$ des Gleichungssystems

$$\begin{array}{rcrcrcrcl} x_1 & - & 4x_2 & + & 2x_3 & & & = & 0 \\ 2x_1 & - & 3x_2 & - & x_3 & - & 5x_4 & = & 0 \\ 3x_1 & - & 7x_2 & + & x_3 & - & 5x_4 & = & 0 \\ & & x_2 & - & x_3 & - & x_4 & = & 0 \end{array}$$

bestimmen. Mit dem Gaußschen Algorithmus bringen wir es in Zeilenstufenform:

$$\begin{array}{rcrcrcrcl} x_1 & - & 4x_2 & + & 2x_3 & & & = & 0 \\ & & x_2 & - & x_3 & - & x_4 & = & 0 \end{array}$$

Also sind x_3, x_4 freie Variable. Für $x_3 = 1, x_4 = 0$ (bzw. $x_3 = 0, x_4 = 1$) erhalten wir die speziellen Lösung $u\colon = (2, 1, 1, 0)$ (bzw. $v\colon = (4, 1, 0, 1)$).

2.2.24 Behauptung: (u, v) ist eine Basis von U, d.h. für jede Lösung (x_1, x_2, x_3, x_4) des Gleichungssystems gibt es eindeutig bestimmte reelle Zahlen s, t mit

$$(x_1, x_2, x_3, x_4) = su + tv = (2s + 4t, s + t, s, t).$$

Beweis: Ist $su + tv = 0$ für $s, t \in \mathbb{R}$ so folgt $(2s + 4t, s + t, s, t) = (0, 0, 0, 0)$ also $s = 0 = t$. Das zeigt, dass (u, v) linear unabhängig ist.

Ist $x = (x_1, x_2, x_3, x_4) \in U$, dann ist $x - x_3 u - x_4 v$ ein Vektor in U dessen dritte und vierte Komponente beide $= 0$ sind. Da die einzige Lösung des Gleichungssystems mit dieser Eigenschaft die triviale Lösung $(0, 0, 0, 0)$ ist folgt $x - x_3 u - x_4 v = (0, 0, 0, 0)$, d.h. $x = x_3 u + x_4 v$. Das zeigt, dass (u, v) auch ein Erzeugendensystem von U ist und damit eine Basis. $\square$

Ganz allgemein kann man mit dem obigen Verfahren eine Basis der Lösungsmenge eines homogenen linearen Gleichungssystems bestimmen. Sei

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

ein solches (über einem Körper K) und $U \subseteq K^n$ die Lösungsmenge. Zunächst einmal können wir das Gleichungssystems mit Hilfe des Gaußschen Algorithmus vereinfachen. Danach bleiben r von Null verschiedene Zeilen in Zeilenstufenform übrig. Ist $r = n$, so ist $U = \{(0, \ldots, 0)\}$.

Andernfalls zerfallen die Komponenten x_i der Lösungen $(x_1, \ldots, x_n)$ in zwei Klassen: Die r Variablen $x_{j_1}, \ldots, x_{j_r}$, die am Anfang einer Zeile stehen und die $n - r$ freien Variablen. Wir können jeweils eine "freie" Komponente $= 1$ setzen und alle anderen $= 0$ wählen und dann daraus $x_{j_1}, \ldots, x_{j_r}$ bestimmen. Das gibt $n - r$ verschiedene spezielle Lösungen des Gleichungssystem, die gesuchte Basis von U.

2.3 Matrizenrechnung

2.3.1 Sei K ein Körper. Eine $m \times n$-Matrix (oder Matrix vom Typ $m \times n$) über K ist ein rechteckiges Schema der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Die Elemente $a_{ij} \in K$ $1 \leq i \leq m, 1 \leq j \leq n$ heißen Komponenten oder Einträge der Matrix A. Wir schreiben abkürzend

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

oder auch nur $A = (a_{ij})$, wenn der Typ feststeht. Matrizen vom Typ $n \times n$ heißen quadratische Matrizen.

Die $1 \times n$–Matrizen (bzw. $m \times 1$–Matrizen) heißen Zeilen- (bzw. Spalten-) vektoren. Sie haben die Form

$$z = (a_1\ a_2\ \dots\ a_n) \quad (\text{bzw. } s = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix})$$

Wir identifizieren K^n mit $M(1 \times n, K)$ (Bei nur einer Zeile oder Spalte benötigt man i.a. keinen zusätzlichen Zeilenindex oder Spaltenindex). Häufig identifizieren wir Elemente $(a_1, \dots, a_n) \in K^n$ mit Zeilenvektoren $(a_1 \dots a_n)$.

Mit $M(m \times n, K)$ bezeichnen wir die Menge aller $m \times n$-Matrizen mit Komponenten aus K. Wir definieren die Addition "+" von zwei Matrizen vom gleichen Typ und skalare Multiplikation "$\cdot$" einer Matrix mit einer Zahl komponentenweise: Für $A = (a_{ij}), B = (b_{ij}) \in M(m \times n, K)$ und $\lambda \in K$ setzten wir

$$(*) \qquad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} :=$$

$$:= \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

$$(**) \qquad \lambda \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix}$$

2.3.2 Bemerkung: $M(m \times n, K)$ zusammen mit der Addition $(*)$ und skalaren Multiplikation $(**)$ ist ein K-Vektorraum.

Das neutrale Element bzgl. der Addition in $M(m \times n, K)$ ist die Nullmatrix

$$0 := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}.$$

2.3.3 Bemerkung: $\dim M(m \times n, K) = mn$.

Beweis: Die Matrizen

$$E_{ij} = \begin{array}{cc} & \begin{array}{ccccccc} & & & j & & & \end{array} \\ i & \begin{pmatrix} & & & 0 & & & \\ & 0 & & \vdots & & 0 & \\ & & & 0 & & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & & 0 & & & \\ & 0 & & \vdots & & 0 & \\ & & & 0 & & & \end{pmatrix} \end{array} \qquad i = 1, \dots, m, \quad j = 1, \dots, n,$$

mit einer Eins an j-ten Stelle der i-ten Zeile und sonst Nullen bilden eine Basis von $M(m \times n, K)$. $\square$

Man kann zwei Matrizen vom Typ $m \times n$ und. $n \times r$ miteinander multiplizieren und erhält dabei eine $m \times r$ Matrix als Produkt: Ist $A = (a_{ij}) \in M(m \times n, K)$ und $B = (b_{jk}) \in M(n \times r, K)$, so definieren wir das Produkt $AB \in M(m \times r, K)$ durch

$$A \cdot B = \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right)_{\substack{i=1,\dots,m \\ k=1,\dots,r}}$$

2.3.4 Beispiele: (a) $(2\ 3\ 0\ 5\ 1) \begin{pmatrix} 0 \\ 1 \\ 4 \\ 1 \\ 0 \end{pmatrix} = 8$. Allgemeiner ist das Produkt eines Zeilenvektors $x = (x_1, \dots, x_n) \in \mathbb{R}^n = M(1 \times n, \mathbb{R})$ mit einem Spaltenvektor $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in M(n \times 1, \mathbb{R})$ eine 1×1 Matrix, d.h. eine Zahl, nämlich $xy = \sum_{i=1}^{n} x_i y_i$.

(b) $\begin{pmatrix} 4 & 3 & 0 & 1 & 2 \\ 2 & 1 & 4 & 0 & 1 \\ 0 & 0 & 4 & 1 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 21 & 17 \\ 14 & 8 & 10 \\ 13 & 1 & 7 \\ 5 & 4 & 17 \end{pmatrix}$.

Sind $a_1, \dots, a_m \in M(1 \times n, K)$ die m Zeilen von $A = (a_{ij}) \in M(m \times n, K)$ aufgefasst als Zeilenvektoren und $b_1, \dots, b_r \in M(n \times 1, K)$ die r Spaltenvektoren von $B = (b_{jk}) \in M(n \times r, K)$ so gilt: $A \cdot B = (a_i \cdot b_j)_{\substack{i=1,\dots,m \\ j=1,\dots,r}}$.

Die quadratische Matrix

$$E_n = \begin{pmatrix} 1 & 0 & \dots & & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & \vdots \\ & & & & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix} \in M(n \times n, K)$$

heisst $n \times n$-Einheitsmatrix. Besteht keine Unklarheit über n, so schreiben wir kurz E statt E_n.

2.3.5 Satz: (Rechenregeln für die Matrizenmultiplikation).

(i) $(A_1+A_2)B = A_1B+A_2B, \quad A(B_1+B_2) = AB_1+AB_2 \qquad \forall\, A, A_1, A_2, \in M(m\times n, K), B, B_1, B_2 \in M(n \times r, K)$;

(ii) $\alpha(AB) = (\alpha A)B = A(\alpha B) \qquad \forall\, \alpha \in K,\ A \in M(m\times n, K),\ B \in M(n\times r, K)$;

(iii) $A(BC) = (AB)C \qquad \forall\, A \in M(m\times n, K), B \in M(n\times r, K), C \in M(r \times s, K)$;

(iv) $E_m A = AE_n = A \qquad \forall\, A \in M(m\times n, K)$.

2.3.6 Bemerkung: Die Matrizenmultiplikation ist i.a. nicht kommutativ. Für $A \in M(m\times n, K)$, $B \in M(n \times r, K)$ kann man zwar AB bilden nicht aber BA falls $r \neq m$. Wenn $r = n$ ist und $m \neq n$, so ist AB eine $m \times m$ Matrix während BA eine $n \times n$ Matrix ist. Selbst wenn A und B beide $n \times n$ Matrizen sind gilt i.a. $AB \neq BA$. Beispielsweise ist

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

2.3.7 Bemerkung: Eine nützliche Beobachtung, die im Folgenden noch häufiger benutzt wird, ist, dass die j-te Spalte der Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n)$$

gleich dem Produkt

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow j\text{-te Stelle} \\ \\ \\ \end{matrix}$$

ist.

Entsprechend sind $e_1A, \dots, e_mA$ die Zeilen von A, wobei $(e_1, \dots, e_m)$ die Standardbasis von K^m ist.

2.3.8 Die Transponierte einer Matrix. Jeder $m \times n$ Matrix $A = (a_{ij})$ ist die an der Diagonalen $i = j$ gespiegelte Matrix A^t (die *Transponierte von* A) zugeordnet, deren i-te Zeile aus den Koeffizienten der i-ten Spalte von A besteht $(1 \leq i \leq n)$. Explizit:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \Rightarrow \quad A^t = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} \in M(n \times m, K)$$

(d.h. $A^t = (b_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ mit $b_{ij} = a_{ji}$).

2.3.9 Beispiel: $A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 5 & -1 \end{pmatrix} \Rightarrow A^t = \begin{pmatrix} 1 & 2 \\ 0 & 5 \\ 3 & -1 \end{pmatrix}$.

Es gelten die folgenden Rechenregeln

(1) $(A+B)^t = A^t + B^t \qquad \forall\ A, B \in M(m \times n, K)$;

(2) $(\alpha A)^t = \alpha A^t \qquad \forall\, \alpha \in K,\ A, \in M(m \times n, K)$;

(3) $(A^t)^t = A \qquad \forall\, A, \in M(m \times n, K)$;

(4) $(AB)^t = B^t A^t \qquad \forall\, A \in M(m \times n, K),\ B \in M(n \times r, K)$.

2.3.10 Definition: Eine $n \times n$ Matrix $A = (a_{ij})$ heißt symmetrisch falls $A^t = A$, d.h. falls $a_{ij} = a_{ji}\ \forall\ i, j \in \{1, \dots, n\}$.

2.3.11 Invertierbare Matrizen. Im Folgenden werden quadratische $n \times n$-Matrizen über K betrachtet. Mit $E = E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ sei die Einheitsmatrix bezeichnet.

2.3.12 Definition: Eine $n \times n$ Matrix A heißt invertierbar, wenn es eine $n \times n$ Matrix B gibt, so daß

$$AB = BA = E.$$

In diesem Fall ist B eindeutig bestimmt und wird mit A^{-1} bezeichnet (d.h. $A^{-1} := B$). Sie heißt inverse Matrix oder die Inverse von A.

2.3.13 Lemma: Wenn es zu $A \in M(n \times n, K)$ zwei Matrizen $B, C \in M(n \times n, K)$ gibt mit $BA = AC = E$, dann ist A invertierbar und es ist $B = C = A^{-1}$.

Das ist ein Spezialfall von 1.2.4.

2.3.14 Beispiel: (a) E ist invertierbar.

(b) Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, K)$ mit $ad - bc \neq 0$. Dann ist A invertierbar und

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2.3.15 Satz: (i) Das Inverse einer invertierbaren Matrix A ist invertierbar und $(A^{-1})^{-1} = A$.

(ii) Das Produkt zweier invertierbarer Matrizen AB ist invertierbar und $(AB)^{-1} = B^{-1}A^{-1}$.

(ii) Die Transponierte einer invertierbaren Matrix A ist invertierbar und es gilt $(A^t)^{-1} = (A^{-1})^t$.

2.3.16 Folgerung: Sei $GL_n(K) = \{A \in M(n \times n, K) \mid A \text{ ist invertierbar}\}$. Dann ist $GL_n(K)$ zusammen mit der Matrizenmultiplikation eine Gruppe. Sie ist nicht abelsch falls $n > 1$ ist.

2.3.17 Satz: Für eine $n \times n$ Matrix $A = (a_{ij})$ sind die folgenden Aussagen äquivalent:

(i) A ist invertierbar;

(ii) Es gibt eine Matrix $B \in M(n \times n, K)$ mit $AB = E$;

(iii) Es gibt eine Matrix $C \in M(n \times n, K)$ mit $CA = E$.

Beweis: (i) $\Rightarrow$ (ii) und (i) $\Rightarrow$ (iii) sind klar.

Vorbemerkung: Es seien

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, a_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix}, \ldots, a_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} \in M(n \times 1, K)$$

die Spalten von A. Vorübergehend identifizieren wir K^n mit $M(n \times 1, K)$, d.h. wir fassen die Elemente aus K^n als Spaltenvektoren $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ auf.

Sei $A = (a_{ij}) \in M(m \times n, K)$. Multiplikation mit A definiert eine Abbildung $K^n \to K^m$ die wir mit l_A bezeichnen:

$$l_A : K^n \to K^m, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \Sigma a_{1j} x_j \\ \vdots \\ \Sigma a_{mj} x_j \end{pmatrix}.$$

l_A ist ein Beispiel für einer sogenannte *lineare Abbildung*, die in einem späteren Abschnitt noch genauer untersucht werden.

Ist $B \in M(n \times r, K)$ so gilt:

$$l_A \circ l_B(x) = A(Bx) = (AB)x = l_{AB}(x) \quad \forall\ x \in K^r$$

d.h. $l_A \cdot l_B = l_{AB}$ Es gilt also

$$l_A(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = x_1 a_1 + \ldots + x_n a_n.$$

Folglich gilt:

l_A ist injektiv $\iff$ $(a_1, \ldots, a_n)$ ist linear unabhängig.

l_A ist surjektiv $\iff$ $(a_1, \ldots, a_n)$ ist ein Erzeugendensystem.

l_A ist bijektiv $\iff$ $(a_1, \ldots, a_n)$ ist eine Basis.

Mit 2.2.22 ergibt sich:

$$l_A \text{ ist injektiv} \quad \iff \quad l_A \text{ ist surjektiv} \quad \iff \quad l_A \text{ ist bijektiv.}$$

(ii) $\Rightarrow$ (i):

$$AB = E \quad \Rightarrow \quad (l_A \circ l_B)(x) = l_A(Bx) = A(Bx) = (AB)x = Ex = x \quad \forall\, x \in K^n.$$

$$\Rightarrow \quad l_A \text{ ist surjektiv, also bijektiv.}$$

$$\Rightarrow \quad l_B = ((l_A)^{-1} \circ l_A) \circ l_B = l_A^{-1} \circ (l_A \circ l_B) = l_A^{-1}$$

$$\Rightarrow \quad x = l_B \circ l_A(x) = B(Ax) \quad \forall\, x \in K^n$$

$$\Rightarrow e_i = (BA)e_i \quad \text{für } i = 1, \ldots, n$$

Also stimmen die Spalten von BA und E überein und es folgt $E = BA$. Das zeigt, dass A invertierbar ist.

(c) $\Rightarrow$ (a) Analog impliziert $CA = E$, dass $l_C \circ l_A = Id_{K^n}$. Damit ist l_A injektiv, also bijektiv und $l_C = (l_A)^{-1}$. Wie vorher folgt $l_A \circ l_C = Id_{K^n}$ und damit $AC = E$. □

2.3.18 Bemerkung: Da wir strenggenommen Zeilenvektoren und Spaltenvektoren der Länge n unterscheiden sollten und K^n aus Zeilenvektoren besteht, sollte man besser l_A als die Abbildung

$$(x_1, \ldots, x_n) \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \Sigma a_{1j} x_j \\ \vdots \\ \Sigma a_{mj} x_j \end{pmatrix} \mapsto (\Sigma a_{1j} x_j, \ldots, \Sigma a_{mj} x_j)$$

definieren, d.h. $l_A : K^n \to K^m, x \mapsto (Ax^t)^t = xA^t$.

2.3.19 Wie invertiert man eine Matrix? Es gibt einen Algorithmus, der es einem erlaubt zu entscheiden, ob eine gegebene Matrix $A \in M(n \times n, K)$ invertierbar ist und der gegebenenfalls ihr Inverses berechnet.

Wir wollen ein $B = (b_{ij}) \in M(n \times n, K)$ finden mit $AB = E$. Betrachten wir nur die j-te Spalte so müssen wir also das lineare Gleichungssystem

$$(*) \qquad A \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-te Stelle}$$

in den Unbekannten $(b_{1j}, \ldots, b_{nj})$ lösen. Dazu wenden wir Gaußschen Algorithmus simultan auf alle Gleichungssysteme $(*)$ für $j = 1, \ldots, n$ an.

Wir führen den Algorithmus zuerst an zwei Beispielen vor.

2.3.20 Beispiel: (a) Sei $A = \begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix}$. Wir versuchen A mit Hilfe von Zeilenumformungen in die Einheitsmatrix umzuwandeln und führen alle Umformungen parallel auch an E_3 durch.

$$\begin{array}{rrr|rrr} 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & -1 & 1 \end{array}$$

$$\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}$$

$$\begin{array}{rrr|rrr} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$$\begin{array}{rrr|rrr} 1 & 0 & 7 & -2 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$$\begin{array}{rrr|rrr} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}$$

$\Rightarrow A$ ist invertierbar mit Inversem $A^{-1} = \begin{pmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{pmatrix}$.

(b)

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}$$

$\Rightarrow B\colon = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ ist nicht invertierbar.

Wir haben also das folgende Verfahren zur Bestimmung des Inversen von $A \in M(n \times n, K)$: Man schreibe die Matrizen A und E_n nebeneinander. Alle Zeilenumformungen, die an A vorgenommen werden, führt man parallel auch an E_n durch. Zunächst bringt man A durch Zeilenumformungen auf Zeilenstufenform. Dabei stellt sich heraus ob $r = n$ ist. Ist $r < n$, so ist A nicht invertierbar. Ist $r = n$, so führt man weitere Zeilenumformungen durch bis aus A links die Matrix E_n geworden ist. Rechts steht dann die inverse Matrix A^{-1}. Das möchte ich jetzt begründen.

Die Zeilenumformungen, die wir in jedem Schritt vorgenommen haben waren jeweils von einem der folgenden Typen:

Typ 1: Vertauschung zweier Zeilen;

Typ 2: Multiplikation einer Zeile mit einem Zahlenfaktor $\neq 0$;

Typ 3: Addition des Vielfachen einer Zeile zu einer anderen.

Diese heißen elementare Zeilenumformungen. Analog definiert man elementare Spaltenumformungen des Typs 1, 2 und 3.

Eine $m \times m$ Matrix $\tilde{E}$ heißt Elementarmatrix vom Typ i wenn sie aus der Einheitsmatrix E_m durch eine elementare Zeilenumformung vom Typ i hervorgeht. Die Elementarmatrizen haben die folgende Gestalt:

$$\begin{pmatrix}
1 & & & & & & & & & & 0 \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
0 & & & & & & & & & & 1
\end{pmatrix}
\begin{matrix}
\\ \\ \\ i\text{-te Zeile} \\ \\ \\ \\ j\text{-te Zeile} \\ \\ \\ \\
\end{matrix}
\qquad \text{Typ 1}$$

$$\begin{pmatrix}
1 & & & & & & 0 \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & \lambda & & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
0 & & & & & & 1
\end{pmatrix}
\qquad \text{Typ 2}$$

$$\begin{pmatrix}
1 & & & & & & 0 \\
 & \ddots & & & & & \\
 & & 1 & \dots & 0 & & \\
 & & \vdots & \ddots & \vdots & & \\
 & & \lambda & \dots & 1 & & \\
 & & & & & \ddots & \\
0 & & & & & & 1
\end{pmatrix}
\begin{matrix}
\\ \\ i\text{-te Zeile} \\ \\ j\text{-te Zeile} \\ \\ \\
\end{matrix}
\qquad \text{Typ 3}$$

2.3.21 Satz: (i) Entsteht $\tilde{A} \in M(m \times n, K)$ aus $A \in M(m \times n, K)$ durch eine elementare Zeilenumformung (bzw. Spaltenumformung) dann gilt:

$$\tilde{A} = \tilde{E}A \quad (\text{bzw. } \tilde{A} = A\tilde{E}^t)$$

mit der zugehörigen Elementarmatrix $\tilde{E}$.

(ii) Die Elementarmatrizen sind invertierbar, die Inversen sind ebenfalls Elementarmatrizen.

Beweis: (i) Wir betrachten den Fall einer elementare Zeilenumformung vom Typ 3. Seien $z_1, \ldots, z_m \in K^n$ die Zeilen von A. $\tilde{A}$ entstehe aus A durch Addition des λ-fachen der i-ten Zeile zur j-ten:

$$\Rightarrow \qquad \tilde{A} = \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j + \lambda z_i \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} e_1 A \\ \vdots \\ e_i A \\ \vdots \\ (e_j + \lambda e_i)A \\ \vdots \\ e_m A \end{pmatrix} = \begin{pmatrix} e_1 \\ \vdots \\ e_j + \lambda e_i \\ \vdots \\ e_m \end{pmatrix} A = \tilde{E}A.$$

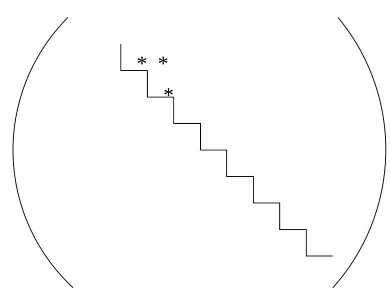
$\square$

2.3.22 Folgerung: (i) Jede invertierbare Matrix ist ein Produkt von Elementarmatrizen.

(ii) Erhält man E_n durch elementare Zeilenumformungen aus A, so verwandeln dieselben Zeilenumformungen die Matrix E_n in A^{-1}.

Beweis: Durch Zeilenumformungen, d.h. durch Multiplikation von links mit Elementarmatrizen läßt sich A in eine Matrix M in Zeilenstufenform

M=



r

überführen, die ebenfalls invertierbar ist. Daher ist $r = n$. Weitere Zeilenumformungen überführen M in die Einheitsmatrix. Also gibt es Elementarmatrizen $\tilde{E}_1, \ldots, \tilde{E}_s$ mit

$$\tilde{E}_s \cdot \ldots \cdot \tilde{E}_1 \cdot A = E_n.$$

$$\Rightarrow A^{-1} = \tilde{E}_s \cdot \ldots \cdot \tilde{E}_1 = \tilde{E}_s \cdot \ldots \cdot \tilde{E}_1 E_n$$

Das beweist (b). Ferner gilt:

$$A = (A^{-1})^{-1} = \tilde{E}_1^{-1} \cdot \ldots \cdot \tilde{E}_s^{-1}.$$

Da das Inverse einer Elementarmatrix wieder eine Elementarmatrix ist, ist A Produkt von Elementarmatrizen. $\square$

2.3.23 Der Rang einer Matrix. Sei $A = (a_{ij})$ eine $m \times n$ Matrix über einem Körper K mit den Zeilen $z_1, \ldots, z_m \in K^n$ und Spalten $a_1, \ldots, a_n \in M(m \times 1, K)$. Die lineare Hülle von $z_1, \ldots, z_m$ heißt *Zeilenraum* von A und $L(a_1, \ldots, a_n)$ heißt *Spaltenraum* von A. Bezeichnet

$$l_A : K^n \to K^m$$

wieder die Abbildung $x \mapsto (Ax^t)^t = xA^t$ so gilt:

$$\begin{aligned} \text{Spaltenraum } A &= \{Ax^t \mid x \in K^n\} = \text{ Bild } (l_A)^t \\ \text{Zeilenraum } A &= \{yA \mid y \in K^m\} = \text{ Bild } (l_{A^t}). \end{aligned}$$

Der Zeilen- (bzw. Spalten-)Rang von A ist die Dimension des Zeilen- (bzw. Spalten-) raums von A. Wir werden gleich sehen, daß der Zeilenrang immer gleich dem Spaltenrang ist.

2.3.24 Definition: Rang $A :=$ Zeilenrang von $A = \dim\ L(z_1, \ldots, z_m)$.

Die folgende Bemerkung ist nützlich für die Bestimmung des Rangs einer Matrix.

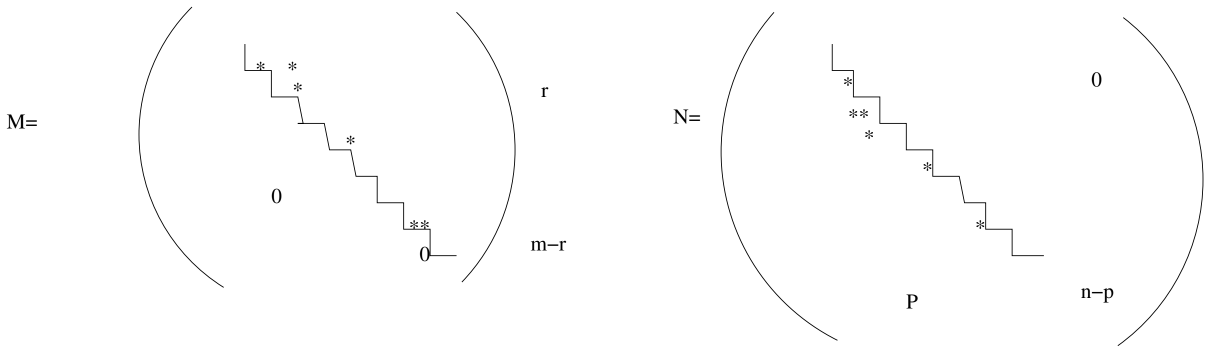
2.3.25 Bemerkung: Elementare Zeilen- (bzw. Spalten-)umformungen ändern den Zeilen- (bzw. Spalten-)rang nicht.

Beweis für eine Zeilenumformung vom Typ 3: Sind $z_1, \ldots, z_m \in K^n$ die Zeilen von $A = (a_{ij}) \in M(m \times n, K)$ und entsteht $\tilde{A}$ aus A durch Addition des λ-fachen der i-ten Zeile zur j-ten Zeile so gilt:

$$L(z_1, \ldots, z_i, \ldots, z_j + \lambda z_i, \ldots, z_m) = L(z_1, \ldots, z_i, \ldots, z_j, \ldots, z_m),$$

d.h. der Zeilenraum – und folglich auch der Zeilenrang – ändert sich nicht. □

Wird nun die $m \times n$ Matrix A durch elementare Zeilenumformungen in eine Matrix M in Zeilenstufenform umgewandelt (und in analoger Weise mit Spaltenumformungen in eine Matrix N in *Spaltenstufenform*),



dann zeigt die folgende Bemerkung, daß der Zeilenrang von $A = r$ ist (und entsprechend der Spaltenrang von $A = p$ ist).

2.3.26 Bemerkung: Sei $A = (a_{ij}) \in M(m \times n, K)$ eine Matrix in Zeilenstufenform, so sind die von Null verschiedenen Zeilen von A linear unabhängig.

Eine entsprechende Aussage gilt für Matrizen in Spaltenstufenform.

Beweis: Sind $z_1 = (0, \ldots, 0, a_{1j_1}, \ldots, a_{1n}), z_2 = (0, \ldots, 0, a_{2j_2}, \ldots, a_{2n}), \ldots, z_r = (0, \ldots, 0, a_{rj_r}, \ldots, a_{rn})$ die von Null verschiedenen Zeilen mit $a_{1j_1}, \ldots, a_{rj_r} \neq 0$ und $j_1 < \ldots < j_r$ dann folgt aus

$$\lambda_1 z_1 + \ldots + \lambda_r z_r = 0$$

durch Betrachtung der j_1-ten Komponente zunächst $\lambda_1 a_{1j_1} = 0$, also $\lambda_1 = 0$; dann $\lambda_2 a_{2j_2} = 0$, also $\lambda_2 = 0$ etc. □

Für eine Matrix $A \in M(m \times n, K)$ bezeichne $\operatorname{Kern} A \subseteq K^n$ den Untervektorraum der Lösungen des homogenen linearen Gleichungssystems

$$Ax^t = 0$$

Es ist also $\operatorname{Kern} A = \{x \in K^n \mid l_A(x) = 0\} = l_A^{-1}(\{0\})$.

2.3.27 Satz (Dimensionsformel): $\operatorname{Rang} A + \dim(\operatorname{Kern} A) = n$.

Beweis: Da sich die linke Seite der Gleichung unter elementaren Zeilenumformungen nicht ändert, können wir o.B.d.A. annehmen, daß A in Zeilenstufenform ist.

Ist r die Anzahl der von 0 verschiedenen Zeilen von A, so gilt: $\operatorname{Rang} A = r$ und $\dim(\operatorname{Kern} A) = n-r$ (letzteres, da das lineare Gleichungssystem $Ax^t = 0$ $n-r$ freie Variable besitzt). □

2.3.28 Folgerung: "Zeilenrang = Spaltenrang".

Beweis: Sei $r = \operatorname{Rang} A$ (= Zeilenrang A) und sei $(b_1, \ldots, b_n)$ eine Basis von K^n, so daß die letzten $n-r$ Vektoren $(b_{r+1}, \ldots, b_n)$ eine Basis von $\operatorname{Kern} A$ sind (eine solche Basis existiert nach dem Ergänzungssatz; wähle zuerst eine Basis von $\operatorname{Kern} A$ und ergänze diese zu einer Basis von K^n).

Stellen wir einen beliebigen Vektor $x \in K^n$ als Linearkombination von $(b_1, \ldots, b_n)$ dar

$$x = \lambda_1 b_1 + \ldots + \lambda_n b_n,$$

so gilt:

$$\begin{aligned} Ax^t &= A(\lambda_1 b_1^t + \ldots + \lambda_n b_n^t) \\ &= \lambda_1 A b_1^t + \ldots + \lambda_n A b_n^t \\ &= \lambda A b_1^t + \ldots + \lambda_r A b_r^t \end{aligned}$$

denn $Ab_{r+1}^t = \ldots = Ab_n^t = 0$.

$$\Rightarrow \qquad \{Ax^t \mid x \in K^n\} = L(Ab_1^t, \ldots, Ab_r^t)$$

Folglich ist $(Ab_1^t, \ldots, Ab_r^t)$ ein Erzeugendensystem des Spaltenraums. Da ein Vektorraum der von r Vektoren erzeugt wird höchstens die Dimension r hat ist der Spaltenrag von A höchstens $= r$, also

$$\text{Spaltenrang } A \quad \leq \quad \text{Zeilenrang } A.$$

Diese Ungleichung gilt fuer eine beliebige Matrix, also auch für A^t. Es folgt

$$\text{Zeilenrang } A = \text{ Spaltenrang } A^t \leq \text{ Zeilenrang } A^t = \text{ Spaltenrang } A.$$

und damit Zeilenrang $A =$ Spaltenrang A. □

2.3.29 Satz (P–Q–Normalform): Sei A eine $m \times n$ Matrix mit Rang $A = r$.

(i) Es gibt invertierbare Matrizen $P \in M(m \times m, K)$ und $Q \in M(n \times n, K)$, derart dass
$$PAQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

(ii) Sind $P \in M(m \times m, K)$, $Q \in M(n \times n, K)$ invertierbare Matrizen mit
$$PAQ = \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix}$$
dann gilt $r = s$ und die letzten $n - r$ Spalten von Q bilden eine Basis von Kern A.

Beweis: (i) Da elementare Zeilenumformungen der Multiplikation von links mit Elementarmatrizen entspricht gibt es eine invertierbare Matrix $P \in M(m \times m, K)$, so dass PA in Zeilenstufenform ist. Offenbar kann man eine Matrix in Zeilenstufenform durch elementare Spaltenumformungen, d.h. durch Multiplikation von rechts mit Elementarmatrizen, umwandeln in $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Folglich gibt eine invertierbare $n \times n$ Matrix Q mit $PAQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$.

(ii) Nach 2.3.25 und 2.3.28 ändert sich der Rang einer Matrix nicht, wenn man sie von links oder rechts mit einer Elementarmatrix multipliziert. Da jede invertierbare Matrix Produkt von Elementarmatrizen ist und ändert sich der Rang folglich nicht bei Multiplikation von links oder rechts mit einer invertierbaren Matrix.
$$\Rightarrow \qquad r = \text{ Rang } A = \text{ Rang } (P \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} Q) = \text{ Rang } \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} = s.$$
$\square$

2.3.30 Folgerung: Für eine $n \times n$ Matrix A sind äquivalent:

(i) A ist invertierbar,

(ii) Rang $A = n$,

(iii) Kern $A = \{0\}$.

Beweis: Die Äquivalenz von (ii) und (iii) folgt aus der Dimensionsformel.

Sind $P, Q \in M(n \times n, K)$ invertierbar mit,

$$PAQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

so gilt:

$$A \text{ ist invertierbar} \quad \Leftrightarrow \quad PAQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \text{ ist invertierbar} \quad \Leftrightarrow \quad r = n.$$

□

Beispiel: $P - Q$–Normalform für $A = \begin{pmatrix} 1 & -4 & 2 & 0 \\ 2 & -3 & -1 & -5 \\ 3 & -7 & 1 & -5 \end{pmatrix}$. Wir bringen A zunächst durch Zeilenumformungen auf Zeilenstufenform:

$$\begin{array}{cccc|ccc} 1 & -4 & 2 & 0 & 1 & 0 & 0 \\ 2 & -3 & -1 & -5 & 0 & 1 & 0 \\ 3 & -7 & 1 & -5 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|ccc} 1 & -4 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{5} & -\frac{1}{5} & \frac{1}{5} \end{array}$$

Jetzt verwandeln wir die Matrix rechts durch Spaltenumformungen:

$$\begin{array}{cccc|cccc} 1 & -4 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 4 & -2 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 4 & 2 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}$$

Mit

$$P := \begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{5} & \frac{1}{5} & 0 \\ -\frac{1}{5} & -\frac{1}{5} & \frac{1}{5} \end{pmatrix}, \quad Q := \begin{pmatrix} 1 & 4 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

gilt dann

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

2.3.31 Wir können ein lineares Gleichunssystem

$$(*) \qquad \begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

auch als Matrizengleichung $Ax = b$ mit $x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ schreiben. Die Matrix

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M(m \times n, K)$$

heisst Koeffizientenmatrix. Ist $m = n$ und A invertierbar, so hat $(*)$ die eindeutig bestimmte Lösung $x = A^{-1}b$.

Im allgemeinen gilt das folgende Lösbarkeitskriterium:

2.3.32 Bemerkung: $(*)$ ist genau dann lösbar, wenn gilt:

$$\operatorname{Rang}(A, b) = \operatorname{Rang} A$$

Dabei entsteht (A, b) aus der Matrix A, indem man noch die Spalte b anhängt:

$$(A, b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

Die Matrix (A, b) heisst erweiterte Koeffizientenmatrix.

2.4 Determinanten

Einer quadratischen Matrix kann man eine Zahl zuordnen – ihre Determinante – die genau dann nicht null ist wenn die Matrix invertierbar ist. Man kann mit Hilfe von Determinanten die Lösung eines linearen Gleichungssystems mit invertierbarer Koeffizientenmatrix direkt als Funktion der Koeffizienten darstellen. Im Falle einer Matrix mit reellen Komponenten misst die Determinante das Volumen des von den Spalten (oder von Zeilen) erzeugten *Parallepiped* (ein 2-dimensionales Parallepiped ist ein Parallelogramm, ein 3-dimensionales ein Spat).

2.4.1 2-reihige Determinanten Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine 2×2-Matrix. Die Zahl

$$\det A := ad - bc$$

heißt Determinante von A.

Wir wollen die Nützlichkeit dieses Begriffs an den oben erwähnten Beispielen – lineare Gleichungssysteme und Flächeninhalt eines Parallelogramms – dokumentieren.

2.4.2 Sei

$$(*) \qquad \begin{array}{rcrcl} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \end{array}$$

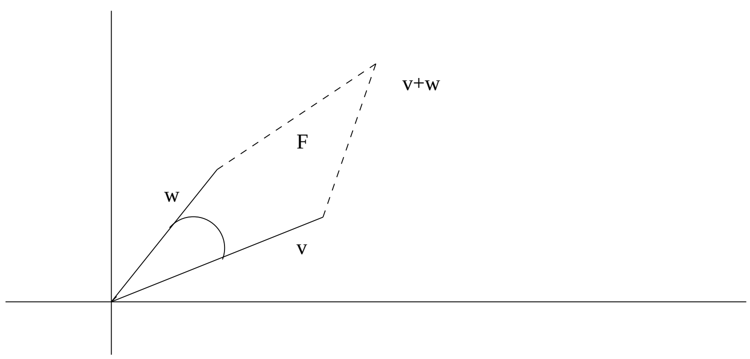
ein lineares Gleichungssystemvon dem wir annehmen, dass die Koeffizientenmatrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ invertierbar ist. Also ist $\det A = a_{11}a_{22} - a_{12}a_{21} \neq 0$ und das Inverse ist $A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

Wir suchen eine allgemeine Formel für die Lösung. Multiplizieren wir die Matrizengleichung $A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ mit A^{-1}, so erhalten wir für die Lösung von $(*)$:

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{\det(A)} = \frac{\det\begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

$$x_2 = \frac{-a_{21}b_1 + a_{11}b_2}{\det(A)} = \frac{\det\begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

2.4.3 Seien $v = (a, b), w = (c, d) \in \mathbb{R}^2$. Wir wollen den Flächeninhalt F des von v und w aufgespannten Parallelogramms $\{sv + tw \mid 0 \leq s, t \leq 1\}$ bestimmen. Im obigen Bild bezeichnet φ den Winkel um den man v drehen



muss (entgegen dem Uhrzeigersinn), um einen zu w parallelen Vektor zu erhalten. Der Flächeninhalt ist Grundfläche mal Höhe. Schreiben wir v und w in Polarkoordinaten

$$\begin{aligned} v &= r_1\ (\cos\alpha, \sin\alpha) \\ w &= r_2\ (\cos\beta, \sin\beta) \end{aligned}$$

so gilt also

$$F = r_1 h = r_1 r_2 \sin\varphi$$

Dabei ist h der Abstand von w von der Geraden durch v und den Ursprung, also $h = r_2 \sin\varphi = r_2 \sin(\beta - \alpha)$.

$$\Rightarrow \qquad F = hr = r_1 r_2 (\sin\beta\cos\alpha - \sin\alpha\cos\beta) = \det\begin{pmatrix} r_1\cos\alpha & r_1\sin\alpha \\ r_2\cos\beta & r_2\sin\beta \end{pmatrix} = \det\begin{pmatrix} v \\ w \end{pmatrix}.$$

2.4.4 Sei K ein beliebiger Körper. Wir listen jetzt einige Eigenschaften der Abbildung

$$\det : M(2 \times 2, K) \longrightarrow K, A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc = \det A$$

auf. Dabei benutzen wir die Notation $A = \begin{pmatrix} v \\ w \end{pmatrix}$ wobei $v = (a, b), w = (c, d)$ die Zeilenvektoren von V sind.

(a) Für $\lambda, \mu \in K$ ist

$$\det \begin{pmatrix} \lambda v \\ w \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} v \\ w \end{pmatrix}, \quad \det \begin{pmatrix} v \\ \mu w \end{pmatrix} = \mu \cdot \det \begin{pmatrix} v \\ w \end{pmatrix}.$$

Im Falle $K = \mathbb{R}$ folgt dass aus der Tatsache, dass der Fläche eines Parallelogramms so gestreckt wird wie die einzelnen Seiten.

(b) Für $w_1, w_2, v \in K^2$ ist

$$\det \begin{pmatrix} v \\ w_1 + w_2 \end{pmatrix} = \det \begin{pmatrix} v \\ w_1 \end{pmatrix} + \det \begin{pmatrix} v \\ w_2 \end{pmatrix}$$

Ebenso gilt für $v_1, v_2, w \in K^2$

$$\det \begin{pmatrix} v_1 + v_2 \\ w \end{pmatrix} = \det \begin{pmatrix} v_1 \\ w \end{pmatrix} + \det \begin{pmatrix} v_2 \\ w \end{pmatrix}.$$

(c)

$$\det \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = 1.$$

Das bedeutet (für $K = \mathbb{R}$) dass das Einheitsquadrat den Flächeninhalt 1 hat.

(d)

$$\det \begin{pmatrix} v \\ v \end{pmatrix} = 0.$$

(e)

$$\det \begin{pmatrix} v \\ w \end{pmatrix} = -\det \begin{pmatrix} w \\ v \end{pmatrix}.$$

(f) $\det \begin{pmatrix} v \\ w \end{pmatrix} \neq 0 \iff \operatorname{Rang} \begin{pmatrix} v \\ w \end{pmatrix} = 2 \iff \begin{pmatrix} v \\ w \end{pmatrix}$ ist invertierbar.

2.4.5 Bevor wir Determinanten von 3×3-Matrizen einführen soll noch einmal an die aus der Schulmathematik bekannten Skalar-, Vektor- und Spatprodukte im $\mathbb{R}^3$ erinnert werden. Wir fassen dabei Vektoren im $\mathbb{R}^3$ vorübergehend als Spaltenvektoren auf.

Das Skalarprodukt $a \bullet b$ zweier Vektoren $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ ist definiert durch

$$a \bullet b = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Sind weder a noch b gleich dem Nullvektor und ist $\varphi \in [0, \pi]$ der Winkel zwischen a und b, so gilt

$$a \bullet b = \|a\| \cdot \|b\| \cdot \cos(\varphi).$$

Dabei bezeichnet $\|a\|$ (bzw. $\|b\|$) die Länge des Vektors a (bzw. b). Es gilt also $\|a\| = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{a \bullet a}$.

Das Vektorprodukt $a \times b$ ist definiert durch

$$a \times b = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Der Vektor $a \times b \in \mathbb{R}^3$ ist der Nullvektor, falls a und b linear abhängig sind. Andernfalls ist er durch die folgenden Eigenschaften eindeutig charakterisiert:

- $a \times b$ steht senkrecht auf a und b;
- $a, b, a \times b$ bildet ein Rechtssystem;
- Die Länge von $a \times b$ ist gleich dem Flächinhalt des von a und b aufgespannten Parallelogramms.

Das Spatprodukt von drei Vektoren $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ und $c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ ist gegeben durch

$$[a, b, c] := a \bullet (b \times c).$$

Bis auf das Vorzeichen, ist es gleich dem Volumen des von a, b und c aufgespannten Spats $\{t_1 a + t_2 a + t_3 a \mid 0 \leq t_1, t_2, t_3 \leq 1\}$. Das sieht man wie folgt. Der Flächeninhalt der Grundfläche des Spats mit den Kanten b und c ist

$F = \|b \times c\|$. Der Vektor $v\colon = \frac{1}{F} \cdot b \times c$ hat die Länge 1 und steht senkrecht auf dieser Grundfläche. Die Höhe h des Spats bzgl. dieser Grundfläche ist $h = |a \bullet v|$. Damit ergibt sich

$$V = hF = |Fa \bullet v| = |a \bullet (b \times c)|.$$

2.4.6 3-reihige Determinanten Sei

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

eine relle 3×3 Matrix. Wir bezeichnen die 3 Spaltenvektoren von A mit

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad a_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

Die Determinante von A ist das Spatprodukt $[a_1, a_2, a_3]$ also

$$\begin{aligned} \det A :=& a_1 \bullet (a_2 \times a_3) = a_1 \bullet \begin{pmatrix} a_{22}a_{33} & -a_{23}a_{32} \\ -a_{12}a_{33} & +a_{13}a_{32} \\ a_{12}a_{23} & -a_{13}a_{22} \end{pmatrix} \\ =& a_{11}(a_{21}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) \\ =& a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \end{aligned}$$

Bis auf das Vorzeichen ist $\det A$ gleich dem Volumen des von des Spalten aufgespannten Spats.

Es gilt $\det A = 0$, genau dann wenn der Spat entartet ist, d.h. wenn (a_1, a_2, a_3) linear abhängig sind. Also:

$$\begin{aligned} \det A = 0 &\iff \operatorname{Rang} A < 3 \\ \det A \neq 0 &\iff \operatorname{Rang} A = 3 \iff A \text{ ist invertierbar.} \end{aligned}$$

Berechnung der Determinante einer 3×3 Matrix mit Hilfe der *Formel von Sarrus*:

$$\begin{array}{ccccc} + & + & + & & \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & & \end{array}$$

Man schreibt die ersten beiden Spalten noch einmal rechts neben A, addiert längs der $\searrow$-Diagonalen zu bildenden Produkte und subtrahiert die längs der $\nearrow$ berechneten Produkte. Also:

$$\begin{aligned}\det A =& a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}\\ & - a_{31}a_{22}a_{13} - a_{32}a_{22}a_{11} - a_{33}a_{21}a_{12}\end{aligned}$$

2.4.7 Beispiel: $A = \begin{pmatrix} 0 & -2 & 3 \\ -2 & 1 & -2 \\ 3 & 6 & 5 \end{pmatrix}$

$$\begin{matrix} 0 & -2 & 3 & 0 & -2 \\ -2 & 1 & -2 & -2 & 1 \\ 3 & 6 & 5 & 3 & 6 \end{matrix}$$

$$\det A = 12 - 36 - 9 - 20 = -53.$$

Wir betrachten jetzt $n \times n$ Matrizen über einem beliebigen Körper K. Ist $A \in M(n \times n, K$ mit den Zeilenvektoren $a_1, \ldots, a_n$ so schreiben wir

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

2.4.8 Satz und Definition: Es gibt genau eine Abbildung

$$\det : M(n \times n, K) \to K$$

mit den folgenden Eigenschaften:

(D1) Für jeden Zeilenindex $i \in \{1, \ldots, n\}$ gilt:

(a) Ist $a_i = a'_i + a''_i$ so ist

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a'_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a''_i \\ \vdots \\ a_n \end{pmatrix}.$$

(b) Ist $a_i = \lambda a'_i, \lambda \in K$ so ist

$$\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a'_i \\ \vdots \\ a_n \end{pmatrix}.$$

(D2) Hat A zwei gleiche Zeilen, so ist

$$\det A = 0.$$

(D3) $\det E_n = 1$.

det heißt *Determinante*, $\det A$ die Determinante von A.

Die Eigenschaft (D1) bedeutet daß det *linear* in den Zeilen ist. Dieser Begriff wird im nächsten Abschnitt eingeführt und untersucht.

Bevor wir zum Beweis des Satzes kommen sollen einige Folgerungen aus den Axiomen (D1) – (D3) abgeleitet werden.

2.4.9 Lemma (Änderung von det bei elementaren Zeilenumformungen): Sei $\det : M(n \times n, K) \to K$ eine Abbildung mit den Eigenschaften (D1) und (D2). Seien $A, \widetilde{A} \in M(n \times n, K)$.

(i) Entsteht $\widetilde{A}$ aus A durch Vertauschung zweier Zeilen so ist $\det \widetilde{A} = -\det A$.

(ii) Entsteht $\widetilde{A}$ aus A durch Multiplikation einer Zeile mit $\lambda \in K$, so gilt $\det \widetilde{A} = \lambda \det A$.

(iii) Entsteht $\widetilde{A}$ aus A durch Addition des Vielfachen einer Zeile zu einer anderen so gilt: $\det \widetilde{A} = \det A$.

Beweis: (ii) folgt direkt aus (D1).

zu (iii): Entsteht $\widetilde{A}$ aus $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ durch Addition des λ-fachen der i-ten

Zeile zur j-ten, so folgt aus der Eigenschaft (D1) für die j-te Zeile:

$$\det \widetilde{A} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ \lambda a_i + a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ \lambda a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}$$

$$= \lambda \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det A \overset{(D2)}{=} \det A.$$

zu (i): Seien die i-te und die j-te die beiden zu vertauschenden Zeilen. Nach (ii) und (iii) ergibt sich durch mehrmaliges Anwenden von Zeilenumformungen:

$$\det A = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j + a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i - (a_i + a_j) \\ \vdots \\ a_j + a_i \\ \vdots \end{pmatrix}$$

$$= - \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j + a_i \\ \vdots \end{pmatrix} = - \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i + a_j + (-1)a_j \\ \vdots \end{pmatrix} = - \det \widetilde{A}$$

□

2.4.10 Folgerung: Sei $\det : M(n \times n, K) \to K$ eine Abbildung mit den Eigenschaften (D1) – (D3).

(i) Entsteht $\widetilde{A}$ aus $A \in M(n \times n, K)$ durch eine elementare Zeilenumformung und ist $\widetilde{E} \in M(n \times n, K)$ die zugehörige Elementarmatrix (also $\widetilde{A} = \widetilde{E}A$) so gilt:

$$\det \widetilde{A} = \det \widetilde{E} \det A$$

(ii) Ist $A \in M(n \times n, K)$ ein Produkt von Elementarmatrizen $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + - \ldots = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}$ konvergiert (man kann zeigen dass ihre Summe $\frac{\pi}{4}$ ist). $A = \widetilde{E}_1 \cdot \ldots \cdot \widetilde{E}_s$ so gilt:

$$\det A = \det \widetilde{E}_1 \cdot \ldots \cdot \det \widetilde{E}_s.$$

Beweis: (i) Nach dem Lemma unterscheiden sich $\det \widetilde{A}$ und $\det A$ durch einen Zahlenfaktor α (d.h. $\det \widetilde{A} = \alpha \det A$) der nur vom Typ der Zeilenumformung und nicht von A selber abhängt. Also gilt auch $\det \widetilde{E} = \alpha \det E_n$ und daher wegen (D3) $\det \widetilde{E} = \alpha$.

$$\Rightarrow \qquad \det \widetilde{A} = \det \widetilde{E} \det A.$$

(ii) folgt sofort aus (i).

2.4.11 Folgerung: Ist $\det = M(n \times n, K) \to K$ eine Abbildung mit den Eigenschaften (D1) und (D2) und ist $A \in M(n \times n, K)$ so gilt:

$$\operatorname{Rang} A < n \Rightarrow \det A = 0.$$

Beweis: Durch elementare Zeilenumformungen vom Typ 1 und 3 läßt sich A in Matrix M in Zeilenstufenform umwandeln. Nach Lemma 2.4.9 gilt $\det A = \pm \det M$.
Ist $r = \operatorname{Rang} A < n$ so sind die letzten $n - r$ Zeilen von M alle $= 0$. Wenden wir (D1), (b) auf die $r + 1$-te Zeile von M und $\lambda = 0$ an so ergibt sich:

$$\det M = 0 - \det M = 0.$$

$\square$

Beweis des Satzes:

1. Beweis der Eindeutigkeit: Seien det und det$'$ zwei Abbildungen mit den Eigenschaften (D1) – (D3) und sei $A \in M(n \times n, K)$. Wir müssen zeigen, daß $\det A = \det' A$. Ist zunächst *Rang* $A < n$, so gilt nach obiger Folgerung:

$$\det A = 0 = \det{}' A.$$

Ist andererseits $\operatorname{Rang} A = n$, so ist A invertierbar und läßt sich durch elementare Zeilenumformungen in die Einheitsmatrix überführen. Nach Lemma

2.4.9 ändern sich det und det′ bei jeder dieser Umformungen um den gleichen Zahlenfaktor. Daher folgt $\det A = \det' A$ aus

$$\det E_n = 1 = \det{}' E_n$$

2. Beweis der Existenz: Die Existenz einer Abbildung $\det : M(n \times n, K) \to K$ mit den Eigenschaften (D1) – (D3) wird durch Induktion bewiesen. Für $n = 1$ hat offenbar $\det(a) := a$ diese Eigenschaften. Für den Induktionsschluß $n-1 \to n$ nehmen wir an, daß die Determinante für $(n-1) \times (n-1)$ Matrizen bereits existiert.

Für $A = (a_{ij}) \in M(n \times n, K)$ bezeichne A_{ij} die aus A durch weglassen der i-ten Zeile und j-ten Spalte entstandene $(n-1) \times (n-1)$ Matrix.

Wir fixieren jetzt einen Spaltenindex $j \in \{1, \ldots, n\}$ und definieren

$$\det A := \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

Es ist leicht zu sehen, dass die so definierte Abbildung $\det : M(n \times n, K) \to K$ die Eigenschaften (D1) – (D3) hat.

2.4.12 Satz (Rechenregeln für Determinanten): Für alle $A, B \in M(n \times n, K)$ gilt:

(i) (Multiplikationssatz)

$$\det AB = \det A \det B$$

(ii) $\det A = \det A^t$
Insbesondere ist $\det A$ auch linear in den Spalten.

(iii) A ist invertierbar $\iff$ $\det A \neq 0$; äquivalent dazu: $\operatorname{Rang} A < n \iff \det A = 0$.

(iv) (Entwicklung von $\det A$ nach einer beliebigen Zeile oder Spalte): Für $A = (a_{ij}) \in M(n \times n, K)$ gilt:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \qquad \text{(Entwicklung von } \det A \text{ nach der } j\text{-ten Spalte)}$$

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \qquad \text{(Entwicklung nach der } i\text{-ten Zeile)}$$

Beweis: (i) Ist $\operatorname{Rang} A < n$ oder $\operatorname{Rang} B < n$, so ist auch $\operatorname{Rang} AB < n$ (Übungsaufgabe!) und folglich (nach Folgerung 2.4.11)

$$\det AB = 0 = \det A \det B.$$

Andernfalls sind A und B invertierbar und lassen sich daher als Produkt von Elementarmatrizen schreiben:

$$A = \widetilde{E}_1 \dots \widetilde{E}_s, B = \widetilde{E}_{s+1} \dots \widetilde{E}_{s+t}$$

Mit 2.4.10, (b) ergibt sich

$$\det AB = \det \widetilde{E}_1 \cdot \ldots \cdot \det \widetilde{E}_{s+t} = \det A \det B$$

(ii) Ist $\operatorname{Rang} A < n$ so gilt nach 2.4.11

$$\det A = 0 = \det A^t$$

(die zweite Gleichheit gilt wegen $\operatorname{Rang} A^t = \operatorname{Rang} A < n$). Ist $\operatorname{Rang} A = n$, so können wir A als Produkt von Elementarmatrizen schreiben

$$A = \widetilde{E}_1 \cdot \ldots \cdot \widetilde{E}_s, \text{ also } A^t = \widetilde{E}_s^t \dots \widetilde{E}_1^t$$

und nach (i) genügt es daher für eine Elementarmatrix $\widetilde{E}$ zu zeigen:

$$\det \widetilde{E} = \det \widetilde{E}^t$$

Das ist klar falls $\widetilde{E}$ vom Typ 1 oder 2 ist, denn dann ist $\widetilde{E}$ symmetrisch. Ist $\widetilde{E}$ vom Typ 3 also von der Form

$$\widetilde{E} = \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & \dots & 0 & & \\ & & \vdots & \ddots & \vdots & & \\ & & \lambda & \dots & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ i\text{-te Zeile} \\ \\ j\text{-te Zeile} \\ \\ \\ \end{array}$$

so ist

$$\widetilde{E}^t = \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & \dots & \lambda & & \\ & & \vdots & \ddots & \vdots & & \\ & & 0 & \dots & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ i\text{-te Zeile} \\ \\ j\text{-te Zeile} \\ \\ \\ \end{array}$$

wieder eine Elementarmatrix vom Typ 3. Nach Lemma ... folgt:

$$\det \widetilde{E} = \det E_n = 1 = \det \widetilde{E}^t$$

(iii) Aufgrund von 2.4.11 ist nur noch zu zeigen:

$$\operatorname{Rang} A = n \Rightarrow \det A \neq 0.$$

Ist $\operatorname{Rang} A = n$ dann ist A invertierbar. Nach (i) gilt:

$$\det A \det A^{-1} = \det AA^{-1} = \det E_n = 1.$$

$\Rightarrow \qquad \det A \neq 0.$

(iv) Die Formel für die Entwicklung von $\det A$ nach der j-ten Spalte haben wir bereits gezeigt. Die Formel für die Entwicklung nach der i-ten Zeile folgt aus $\det A = \det A^t$. □

2.4.13 Beispiel: $A = \begin{pmatrix} 5 & -4 & 2 \\ 0 & 1 & 0 \\ 7 & 8 & 3 \end{pmatrix}$. Wir entwickeln $\det A$ nach der zweiten Zeile:

$$\det A = \det \begin{pmatrix} 5 & 2 \\ 7 & 3 \end{pmatrix} = 1.$$

2.4.14 Bemerkungen: (a) Für eine invertierbare $n \times n$ Matrix A gilt:

$$\det(A^{-1}) = (\det A)^{-1}$$

(b) Für $A \in M(n \times n, K)$ und eine invertierbare Matrix $C \in M(n \times n, K)$ gilt:

$$\det(CAC^{-1}) = \det A$$

(c) Ist $A = (a_{ij}) \in M(n \times n, K)$ eine *obere* (bzw. *untere*) *Dreiecksmatrix*, d.h. ist $a_{ij} = 0 \quad \forall\ i > j$ (bzw. $a_{ij} = 0 \quad \forall\ j > i$) so gilt:

$$\det A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

2.4.15 Die Adjunkte von A: Sei $A = (a_{ij}) \in M(n \times n, K)$. Setze für $i, j \in \{1, \ldots, n\}$

$$a_{ij}^* := (-1)^{i+j} \det A_{ji}.$$

Die Matrix $A^* = (a_{ij}^*)$ heißt *Adjunkte* von A. Nach Satz 2.4.12, (iv) gilt für die (ii)-te Komponente des Produkts AA^*:

$$\sum_{j=1}^{n} a_{ij} a_{ji}^* = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} = \det A$$

Für $j \neq i$ ergibt sich für die (ii)-te Komponente von AA^*:

$$\sum_{k=1}^{n} a_{ik} a_{kj}^* = \sum_{k=1}^{n} a_{ik} (-1)^{k+j} \det A_{jk}$$

Das ist die Determinante der Matrix $\widetilde{A}$ die man aus A erhält, wenn man die j-te Zeile durch die i-te ersetzt. Also hat $\widetilde{A}$ zwei gleiche Zeilen und $\sum\limits_{k=1}^{n} a_{ik} a_{kj}^* = 0$.

Für das Produkt AA^* gilt also

$$AA^* = \begin{pmatrix} \det A & & 0 \\ & \ddots & \\ 0 & & \det A \end{pmatrix} = \det(A) \cdot E_n$$

Analog zeigt man: $A^*A = \det(A) \cdot E_n$. Damit ergibt sich:

2.4.16 Satz: Sei $A \in M(n \times n, K)$ invertierbar. Dann gilt für die inverse Matrix A^{-1}:

$$A^{-1} = \det(A)^{-1} A^* = \left(\frac{(-1)^{i+j} \det A_{ji}}{\det A} \right)_{i,j}.$$

Man kann mit Hilfe des Determinantenkalküls eine Formel für die Lösung eines linearen Gleichungssystems mit n Gleichungen in n Unbestimmten und invertierbarer Koeffizientenmatrix angeben.

2.4.17 Satz (Cramersche Regel): Sei

$$(*) \qquad \begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \ldots & + & a_{nn}x_n & = & b_n \end{array}$$

ein lineares Gleichungssystem mit n Gleichungen in n Unbestimmten mit Koeffizientenmatrix $A = (a_{ij}) \in M(n \times n, K)$. Ist $\det A \neq 0$ so besitzt $(*)$ eine eindeutig bestimmte Lösung $(x_1, \ldots, x_n)$. Es gilt

$$x_i = \frac{\det \begin{pmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}}$$

wobei die Matrix im Zähler aus A entsteht, indem man die i-te Spalte von A durch $b \colon = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ ersetzt.

Beweis: Dass $(*)$ eine eindeutig bestimmte Lösung $(x_1, \ldots, x_n)$ besitzt, folgt aus der Invertierbarkeit von A. Seien $a_1, \ldots, a_n$ die Spalten von A. Die Vektoren $a_1, \ldots, a_{i-1}, x_i \cdot a_i - b, a_{i+1}, \ldots, a_n$ sind linear abhängig, denn

$$x_1 a_1 + \ldots + x_{i-1} a_{i-1} + 1 \cdot (x_i a_i - b) + \ldots + x_n a_n = Ax - b = 0.$$

Es folgt

$$\begin{aligned} 0 &= \det(a_1, \ldots, a_{i-1}, x_i \cdot a_i - b, a_{i+1}, \ldots, a_n) \\ &= x_i \det A - \det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) \end{aligned}$$

wobei wir bei der zweiten Gleichheit benutzt haben, dass die Determinante auch in den Spalten linear ist. □

2.4.18 Beispiel: Wir betrachten das Gleichungssystems

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_2 + x_3 &= 1 \\ 3x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

Für die Determinante der Koeffizientenmatrix gilt

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = 1 \det \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} + 3 \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = 2$$

Also ergibt sich:

$$x_1 = \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} = \frac{-2}{2} = -1$$

$$x_2 = \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} = \frac{1+3}{2} = 2$$

$$x_3 = \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 3 & 2 & 0 \end{pmatrix} = \frac{-2+0}{2} = -1$$

Für große n ist die Cramersche Regel unpraktisch, da der Rechenaufwand zur Berechnung von $n+1$ Determinanten von $n \times n$-Matrizen von zu gross ist.

2.4.19 Leibnizsche Formel Abschliessend soll noch eine Verallgemeinerung der Formel von Sarrus für n-reihige Determinanten erwähnt werden. Sei $A = (a_{ij}) \in M(n \times n, K)$. Dann gilt:

$$\det A = \sum_{i=(i_1,\ldots,i_n)} (-1)^{\varepsilon(i)} a_{i_1 1} a_{i_2 2} \cdot \ldots \cdot a_{i_n n} \tag{9}$$

Dabei erstreckt sich die Summe über alle möglichen Anordnungen $i = (i_1, \ldots, i_n)$ der Zahlen $1, \ldots, n$ (d.h. über alle Permutationen von $\{1, \ldots, n\}$). $\varepsilon(i)$ ist die Anzahl der Vertauschungen, die erforderlich ist um $(i_1, \ldots, i_n)$ in die natürliche Reihenfolge zu bringen.

Im allgemeinen gibt es verschiedene Möglichkeiten eine Anordnung $i = (i_1, \ldots, i_n)$ durch Vertauschungen in die natürliche Reihenfolge zu bringen. Man kann aber zeigen, dass die Anzahl der Vertauschungen dabei immer gerade oder immer ungerade ist. Daher ist das Vorzeichen $(-1)^{\varepsilon(i)}$ wohldefiniert. Für $n = 3$ und $i = (2, 3, 1)$ gilt beispielsweise $\varepsilon(i) = 2$ und für $i = (3, 2, 1)$ ist $\varepsilon(i) = 1$ oder $= 3$.

Die linke Seite der Formel (9) besteht aus $n!$ Summanden. Da mit wachsendem n die Fakultät $n!$ rapide steigt, ist die Formel zur Berechnung der Determinante für grosse n ungeeignet. Wir verzichten auf den Beweis von (9).

2.5 Lineare Abbildungen

Lineare Abbildungen sind Abbildungen zwischen Vektorräumen, die die Struktur erhalten. Für den K^n sind dies Abbildungen die durch Mulplikation mit einer festen Matrix gegeben sind.

2.5.1 Definition: Sei K ein Körper und V und W K-Vektorräume. Eine Abbildung

$$f : V \longrightarrow W$$

heisst linear oder Homomorphismus, wenn für alle $v, v_1, v_2 \in V$ und $\lambda \in K$ gilt

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2) \\ f(\lambda v) &= \lambda f(v). \end{aligned}$$

Dazu äquivalent ist die Bedingung

$$f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)$$

für alle $\lambda_1, \ldots, \lambda_n \in K$ und $v_1, \ldots, v_n \in V$.

2.5.2 Beispiele: (a) Die Abbildung

$$f : \mathbb{R}^3 \to \mathbb{R}^2, (x, y, z) \mapsto (3x - y + z, x + 2y)$$

ist linear, denn es ist

$$\begin{aligned} f((x_1 + x_2, y_1 + y_2, z_1 + z_2)) &= (3(x_1 + x_2 x) - (y_1 + y_2) + (z_1 + z_2), (x_1 + x_2 x) + 2(y_1 + y_2)) \\ &= (3x_1 - y_1 + z_1, x_1 + 2y_1) + (3x_2 - y_2 + z_2, x_2 + 2y_2) = f((x_1, y_1, z_1)) + f((x_2, y_2, z_2)) \end{aligned}$$

und entsprechend

$$f(\lambda(x, y, z)) = \lambda f((x, y, z)).$$

(b) Allgemein ist die Multiplikation mit einer festen Matrix $A \in M(m \times n, K)$ ist linear:

$$l_A : K^n \to K^m, x = (x_1 \ldots x_n) \mapsto \left(A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)^t = xA^t$$

Wir werden weiter unten zeigen dass jede lineare Abbildung $K^n \to K^m$ von dieser Form ist.

(c) Sei $\mathcal{C}^1([a,b])$ der reelle Vektorraum, der stetig differenzierbaren Funktionen $f : [a,b] \to \mathbb{R}$ und $\mathcal{C}^0([a,b])$ der Vektorraum aller stetigen Funktionen auf $[a,b]$. Der Differentiationsoperator

$$\frac{d}{dx} : \mathcal{C}^1([a,b]) \to \mathcal{C}^0([a,b]), f(x) \mapsto \frac{df}{dx} = f'(x)$$

ist linear. (d) Das Integral $f \mapsto \int_a^b f(x)dx, \quad f \in \mathcal{C}^0([a,b])$ ist linear.

Keine linearen Abbildungen sind z.B.:

(e) Die Parallelverschiebung $t_a : V \to V, v \mapsto v + a$ (mit einem festem Vektor $a \in V$).

(f) $f : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \mapsto (x^2, x+y)$.

Seien V und W K-Vektorräume. Sind $f, g : V \to W$ lineare Abbildungen und $\lambda \in K$ so verifiziert man leicht dass die Abbildungen

$$f + g : V \to W, v \mapsto f(v) + g(v), \tag{10}$$

$$\lambda f : V \to W, v \mapsto \lambda f(v) \tag{11}$$

wieder linear sind. Die Menge aller linearen Abbildungen von V in W bezeichnen wir mit $\mathrm{Hom}_K(V,W)$. Sie wird durch die in 10 definierte Addition und skalare Multiplikation selbst wieder zu einem K-Vektorraum.

2.5.3 Lineare Abbildungen und Matrizen: Seien V und W endlichdimensionale K-Vektorräume der Dimensionen n und m. Seien $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_m)$ Basen von V bzw. W. Wir können die Bilder $f(v_j), j = 1, \ldots, n$ in eindeutiger Weise als Linearkombination von $(w_1, \ldots, w_m)$ schreiben:

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

Die $m \times n$ Matrix

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = (a_{ij})$$

heisst die *Darstellungsmatrix* von f bzgl. der Basen $\mathcal{A}$ und $\mathcal{B}$.

2.5.4 Beispiel: Sei $\mathbb{R}[x]_{\deg \leq n}$ der Vektorraum der Polynome vom Grad $\leq n$. Wir wollen die Darstellungsmatrix von

$$D : \mathbb{R}[x]_{\deg \leq 3} \to \mathbb{R}[x]_{\deg \leq 2}, f(x) \mapsto \frac{df}{dx} = f'(x)$$

bzgl. der Bases $\mathcal{A} = (1, x, x^2, x^3)$ und $\mathcal{B} = (1, x, x^2)$ bestimmen. Wegen $D(x^3) = 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2$ sind die Komponenten der letzte Spalte von $M_{\mathcal{B}}^{\mathcal{A}}(D)$ der letzten Spalte $0, 0$ und 3. Entsprechend ergibt sich wegen $D(x^i) = ix^{i-1}$:

$$M_{\mathcal{B}}^{\mathcal{A}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

2.5.5 Im Fall einer linearen Abbildung $f : K^n \to K^m$ und der Standardbasen $\mathcal{K} = (e_1, \ldots, e_n)$, $\mathcal{K}' = (e_1, \ldots, e_m)$, gilt

$$f(e_j) = \sum_{i=1}^{m} a_{ij} e_i = (a_{1j}, \ldots, a_{mj}).$$

d.h. $f(e_j)^t$ ist die j-te Spalte $A\colon = (a_{ij} = M_{\mathcal{B}}^{\mathcal{A}}(f)$. Also gilt

$$f(e_j) = e_j A^t = l_A(e_j) \qquad \forall\, j \in \{1, 2, \ldots, n\}.$$

Wegen der Linearität von f und l_A ergibt sich damit für einen beliebigen Vektor $x = (x_1, \ldots, x_n) \in K^n$

$$\begin{aligned} f(x) = f(x_1 e_1 + \ldots + x_n e_n) &= x_1 f(e_1) + \ldots + x_n f(e_n) \\ &= x_1 l_A(e_1) + \ldots + x_n l_A(e_n) = l_A(x), \end{aligned}$$

also $f = l_A$. Daher ist jede lineare Abbildung $f : K^n \to K^m$ von der Form $f = l_A$ ist, wobei $A \in M(m \times n, K)$ die Darstellungsmatrix von f bzgl. der Standardbasen ist. Das zeigt ausserdem, dass die Abbildung

$$\mathrm{Hom}_K(K^n, K^m) \longrightarrow M(m \times n, K),\, f \mapsto M_{\mathcal{K}'}^{\mathcal{K}}(f)$$

bijektiv ist und die Umkehrabbildung durch $A \mapsto l_A$ gegeben ist.

Indem man die obige Argumentation leicht modifiziert kann man allgemein zeigen:

2.5.6 Satz: Seien V und W K-Vektorräume mit fest gewählten Basen $\mathcal{A} = (v_1, \ldots, v_n)$, $\mathcal{B} = (w_1, \ldots, w_m)$. Dann ist die Abbildung

$$Hom_K(V, W) \longrightarrow M(m \times n, K),\, f \mapsto M_{\mathcal{B}}^{\mathcal{A}}(f) \tag{12}$$

bijektiv.

Als nächstes wollen wir untersuchen we sich die Darstellungsmatrix ändert, wenn man andere Basen wählt.

2.5.7 Lemma: (i) Seien V, W, U drei K-Vektorräume und $f : U \to V$, $g : V \to W$ lineare Abbildungen. Dann ist auch die Komposition

$$g \circ f : U \to W, u \mapsto g \circ f(u) = g(f(u))$$

linear.

(ii) Seien $\mathcal{A} = (u_1, \dots, u_n)$, $\mathcal{B} = (v_1, \dots, v_m)$ und $\mathcal{C} = (w_1, \dots, w_r)$ Basen von U, V und W. Dann gilt

$$M_{\mathcal{C}}^{\mathcal{A}}(g \circ f) = M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f).$$

Beweis: (i) ist klar.

(ii) Sei $M_{\mathcal{C}}^{\mathcal{B}}(g) = (a_{ij})$ und $M_{\mathcal{B}}^{\mathcal{A}}(f) = (b_{jk})$. Die Behauptung folgt aus

$$g \circ f(u_k) = g(\sum_{j=1}^{m} b_{jk} v_j) = \sum_{j=1}^{m} b_{jk} g(v_j) = \sum_{j=1}^{m} b_{jk} \sum_{i=1}^{r} a_{ij} w_i = \sum_{i=1}^{r} (\sum_{j=1}^{m} a_{ij} b_{jk}) w_i.$$

Sei V ein n-dimensionaler K-Vektorraum und $\mathcal{A} = (v_1, \dots, v_n), \mathcal{B} = (w_1, \dots, w_n)$ zwei Basen von V. Die Darstellungsmatrix der Identität $Id_V : V \to V, v \mapsto v$ bzgl. $\mathcal{A}$ und $\mathcal{B}$ bezeichnen wir mit

$$T_{\mathcal{B}}^{\mathcal{A}} := M_{\mathcal{B}}^{\mathcal{A}}(Id_V).$$

Sie heisst *Transformationsmatrix* des Basiswechsels. Sei v ein beliebiger Vektor in V. Wir schreiben v als Linearkombination von $(v_1, \dots, v_n)$ und $(w_1, \dots, w_n)$

$$v = x_1 v_1 + \dots x_n v_n = y_1 w_1 + \dots y_n w_n$$

Die Kenntnis von $T_{\mathcal{B}}^{\mathcal{A}}$ erlaubt es, die *Koordinaten* $(x_1, \dots, x_n)$ von v bzgl. $(v_1, \dots, v_n)$ in die Koordinaten bzgl. $(w_1, \dots, w_n)$ umzurechnen:

$$T_{\mathcal{B}}^{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Ist $T_{\mathcal{B}}^{\mathcal{A}} = (a_{ij})$ so gilt nämlich

$$v_j = Id_V(v_j) = \sum_{i=1}^{n} a_{ij} w_i$$

und daher

$$\sum_{i=1}^{n} y_i w_i = v = \sum_{j=1}^{n} x_j v_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ij} w_i = \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} x_j) w_i$$

also $y_i = \sum_{j=1}^{n} a_{ij} x_j$.

2.5.8 Lemma: Sei V ein endlich-dimensionaler K-Vektorraum und $\mathcal{A}$ und $\mathcal{B}$ zwei Basen von V. Die Transformationsmatrix $T_{\mathcal{B}}^{\mathcal{A}}$ ist invertierbar mit Inversem

$$(T_{\mathcal{B}}^{\mathcal{A}})^{-1} = T_{\mathcal{A}}^{\mathcal{B}}.$$

Beweis: Offenbar gilt $T_{\mathcal{A}}^{\mathcal{A}} = E_n = T_{\mathcal{B}}^{\mathcal{B}}$ mit $n = \dim V$. Mit Lemma 2.5.7 ergibt sich

$$T_{\mathcal{B}}^{\mathcal{A}} \cdot T_{\mathcal{A}}^{\mathcal{B}} = T_{\mathcal{A}}^{\mathcal{A}} = E_n = T_{\mathcal{B}}^{\mathcal{B}} = T_{\mathcal{A}}^{\mathcal{B}} \cdot T_{\mathcal{B}}^{\mathcal{A}}.$$

2.5.9 Beispiel: Im $\mathbb{R}^2$ seien die Basen $\mathcal{A} = ((2,1),(3,2))$ und $\mathcal{B} = ((1,0),(3,1))$ und der Vektor

$$v = 5 \cdot (2,1) - 3 \cdot (3,2)$$

gegeben. Sei $\mathcal{K} = ((1,0),(0,1))$. Für die Transformationsmatrix $T_{\mathcal{B}}^{\mathcal{A}}$ gilt nach 2.5.7 und 2.5.8:

$$T_{\mathcal{B}}^{\mathcal{A}} = T_{\mathcal{B}}^{\mathcal{K}} \cdot T_{\mathcal{K}}^{\mathcal{A}} = (T_{\mathcal{K}}^{\mathcal{B}})^{-1} \cdot T_{\mathcal{K}}^{\mathcal{A}}.$$

Ferner ist

$$T_{\mathcal{K}}^{\mathcal{A}} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix},$$

denn die erste Spalten von $T_{\mathcal{K}}^{\mathcal{A}}$ besteht aus den Koordinaten von $(2,1)$ und die zweite Spalte aus den Koordinaten von $(3,2)$ bzgl. der Basis $((1,0),(0,1))$. Entsprechend ist

$$T_{\mathcal{K}}^{\mathcal{B}} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

also

$$T_{\mathcal{B}}^{\mathcal{A}} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & -3 \\ 1 & 2 \end{pmatrix}$$

Wegen $\begin{pmatrix} -1 & -3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ -3 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$ gilt also

$$v = 4 \cdot (1,0) - 1 \cdot (3,1).$$

Wir wollen nun die Frage beantworten, wie sich die Darstellungsmatrix $M_{\mathcal{B}}^{\mathcal{A}}$ einer linearen Abbildung ändert, wenn man von den Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W zu zwei neuen Basen $\mathcal{A}'$ und $\mathcal{B}'$ übergeht.

2.5.10 Satz: Es gilt:

$$M_{\mathcal{B}'}^{\mathcal{A}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot T_{\mathcal{A}}^{\mathcal{A}'}$$

Die Darstellungsmatrizen von f bezüglich zweier Basenpaare unterscheidet sich also durch Links- und Rechtsmultiplikation mit gewissen Transformationsmatrizen.

Beweis: Das folgt wegen $f = Id_W \circ f \circ Id_V$ aus 2.5.7 (b). $\square$

2.5.11 Definition: Seien V und W K-Vektorräume.

1. Eine bijektive, lineare Abbildung $f : V \to W$ heisst Isomorphismus (von Vektorräumen).

2. V heisst isomorph zu W (in Zeichen: $V \cong W$), wenn es einen Isomorphsmus $f : V \to W$ gibt.

Seien V und W endlich-dimensionale K-Vektorräume. Man kann leicht sehen, dass die Umkehrabbildung f^{-1} eines Isomorphismus $f : V \to W$ wieder linear ist. Also gilt: $V \cong W \Leftrightarrow W \cong V$.

Man kann leicht nachrechnen, dass die Abbildung (12) auch linear ist, d.h. (12) ist ein Isomorphismus.

2.5.12 Lemma: Seien $\mathcal{A}$ und $\mathcal{B}$ Basen von V und W. Eine lineare Abbildung $f : V \to W$ ist genau dann ein Isomorphismus, wenn $\dim V = \dim W$ ist und die Darstellungsmatrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$ invertierbar ist.

Beweis: Sei $n = \dim V$ und $m = \dim W$. Ist $f : V \to W$ ein Isomorphismus, so gilt für die Darstellungsmatrizen $A\colon = M_{\mathcal{B}}^{\mathcal{A}}(f)$, $B =\colon = M_{\mathcal{A}}^{\mathcal{B}}(f^{-1})$:

$$A \cdot B = M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot M_{\mathcal{A}}^{\mathcal{B}}(f^{-1}) = M_{\mathcal{B}}^{\mathcal{A}}(f \circ f^{-1}) = M_{\mathcal{B}}^{\mathcal{B}}(Id_W) = E_m$$

und ebenso $B \cdot A = E_n$. Also ist $m = n$ und A ist invertierbar.

Ist umgekehrt $m = n$ und $A = M_{\mathcal{B}}^{\mathcal{A}}(f)$ invertierbar mit Inversem B, so gibt es aufgrund der Surjektivität von (12) eine lineare Abbildung $g : W \to V$ mit Darstellungsmatrix $M_{\mathcal{A}}^{\mathcal{B}}(g) = B$. Wegen

$$M_{\mathcal{A}}^{\mathcal{A}}(g \circ f) = M_{\mathcal{A}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) = B \cdot A = E_n = M_{\mathcal{A}}^{\mathcal{A}}(Id_V)$$

folgt $g \circ f = Id_V$ (da die Abbildung (12) injektiv ist). Analog zeigt man $f \circ g = Id_W$. Also ist f ein Isomorphismus mit Umkehrabbildung g. $\square$

2.5.13 Satz: Seien V und W zwei endlich-dimensionale K-Vektorräume. Dann gilt:

$$\dim V = \dim W \quad \Longleftrightarrow V \cong W$$

Beweis: "$\Rightarrow$" folgt aus Lemma 2.5.12.

"$\Leftarrow$": Sei $n = \dim V = \dim W$. Wir wählen Basen $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_n)$ von V und W. Nach Satz 2.5.6 gibt es eine lineare Abbildung $f : V \to W$ mit $M^{\mathcal{A}}_{\mathcal{B}}(f) = E_n$. Nach 2.5.12 ist f ein Isomorphismus. Explizit ist f auf Linearkombinationen von $(v_1, \ldots, v_n)$ gegeben durch

$$f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n$$

$\square$

2.5.14 Bemerkung: Explizit ist die Abbildung $f : V \to W$ im zweiten Teil des obigen Beweises auf Linearkombinationen $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ durch

$$f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n$$

gegeben. Man kann leicht direkt sehen, dass die so definierte Abbildung $V \to W$ ein Isomorphismus ist.

2.5.15 Definition: Sei $f : V \to W$ eine lineare Abbildung.

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

heisst der Kern von f.

$$\operatorname{Bild} f = \{w \in W \mid \exists\ v \in V : f(v) = w\} = f(V)$$

heisst das Bild von f. Ist $\operatorname{Bild} f$ endlich-dimensional, so heisst $\operatorname{Rang} f = \dim \operatorname{Bild} f$ der Rang der linearen Abbildung f.

2.5.16 Bemerkung: Das Bild einer linearen Abbildung der Form $f = l_A : K^n \to K^m$ ist offenbar gleich dem Spaltenraum von A. Also ist $\operatorname{Rang} l_A = \operatorname{Rang} A$.

2.5.17 Lemma: Sei $f : V \to W$ eine lineare Abbildung. Dann sind $\mathrm{Kern}\, f$ und $\mathrm{Bild}\, f$ Untervektorräume von V bzw. W.

Beweis: Für $v, w \in \mathrm{Kern}\, f$ und $\lambda \in K$ gilt

$$f(v + w) = f(v) + f(w) = 0 + 0 = 0, \quad f(\lambda v) = \lambda f(v) = 0$$

also $v + w, \lambda v \in \mathrm{Kern}\, f$. Die Unterraum-Eigenschaft für $\mathrm{Bild}\, f$ ist genauso leicht zu zeigen. □

2.5.18 Satz (Dimensionsformel für lineare Abbildungen): Sei $f : V \to W$ eine lineare Abbildung von endlich-dimensionalen K-Vektorräumen. Dann gilt:

$$\dim V = \dim \mathrm{Kern}\, f + \mathrm{Rang}\, f.$$

Beweis: Sei $r := n - \dim \mathrm{Kern}\, f$, und sei $(v_{r+1}, \ldots, v_n)$ eine Basis von $\mathrm{Kern}\, f$, die wir zu einer Basis $(v_1, \ldots, v_n)$ von V ergänzen. Sei $v = \lambda_1 v_1 + \ldots + \lambda_n v_n \in V$. Wegen

$$f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = \lambda_1 f(v_1) + \ldots + \lambda_r f(v_r)$$

ist $(f(v_1), \ldots, f(v_r))$ ein Erzeugendensystem von $\mathrm{Bild}\, f$. Die Vektoren $f(v_1), \ldots, f(v_r)$ sind überdies linear unabhängig, denn aus

$$\lambda_1 f(v_1) + \ldots + \lambda_r f(v_r) = 0$$

folgt $f(\lambda_1 v_1 + \ldots + \lambda_r v_r) = 0$, also $\lambda_1 v_1 + \ldots + \lambda_r v_r \in \mathrm{Kern}\, f$ und damit

$$\lambda_1 v_1 + \ldots + \lambda_r v_r = \lambda_{r+1} v_{r+1} + \ldots + \lambda_n v_n$$

für gewisse $\lambda_{r+1}, \ldots, \lambda_n \in K$. Da $(v_1, \ldots, v_n)$ eine Basis ist, folgt

$$\lambda_1 = \ldots = \lambda_r = \lambda_{r+1} = \ldots = \lambda_n = 0.$$

Also ist $(f(v_1), \ldots, f(v_r))$ eine Basis von $\mathrm{Bild}\, f$ und es folgt

$$n - \dim \mathrm{Kern}\, f = r = \mathrm{Rang}\, f.$$

□

2.5.19 Bemerkung: Sei $f : V \to W$ eine lineare Abbildung von endlich-dimensionalen K-Vektorräumen. Wir wählen eine Basis $\mathcal{A} = (v_1, \ldots, v_n)$ von V wie oben und ergänze $w_1 = f(v_1), \ldots, w_r = f(v_r)$ ($r :\ = \operatorname{Rang} f$) zu einer Basis $(w_1, \ldots, w_m)$ von W. Da $f(v_j) = w_j$ falls $j \leq r$ und $f(v_j) = 0$ falls $j > r$ hat die Darstellungsmatrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$ die Gestalt

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Wir betrachten speziell den Fall einer lineare Abbildung der Form $f = l_A : K^n \to K^m$ für eine Matrix $A \in M(m \times n, K)$. Sei $P = T_{\mathcal{B}}^{\mathcal{K}'}$ und $Q = T_{\mathcal{K}}^{\mathcal{A}}$ wobei $\mathcal{K}$ und $\mathcal{K}'$ wieder die Standardbasen von K^n bzw. K^m sind. Mit Satz 2.5.10 ergibt sich:

$$PAQ = T_{\mathcal{B}}^{\mathcal{K}'} \cdot M_{\mathcal{K}'}^{\mathcal{K}}(l_A) \cdot T_{\mathcal{K}}^{\mathcal{A}} = M_{\mathcal{B}}^{\mathcal{A}}(l_A) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

d.h. wir erhalten einen neuen Beweis für die Existenz der P–Q–Normalform.

3 Kapitel: Analysis

3.1 Konvergente Folgen

3.1.1 Definition: Sei M eine Menge. Eine Folge in M ist eine Abbildung $\mathbb{N}_0 \to M, n \mapsto a_n$. Hierfür schreibt man meistens einfach $\{a_n\}_{n\in\mathbb{N}_0}$ oder $\{a_n\}_{n\geq 0}$ oder auch nur $\{a_n\}$. Oft schreibt man auch die ersten Werte der Folge:

$$a_0, a_1, a_2, \ldots.$$

Die Elemente a_n heißen Glieder der Folge.

Ist $M = \mathbb{R}$ (oder $M = \mathbb{C}$) so spricht man von einer reellen (bzw. komplexen) (Zahlen-)Folge.

3.1.2 Bemerkung: Das erste Glied einer Folge braucht nicht immer a_0 zu sein. Durch Umbenennung, etwa $b_0 = a_5, b_1 = a_6, \ldots, b_n = a_{n+5}$ erreicht man, daß auch $a_5, a_6, \ldots$ eine Folge darstellt. Wir bezeichnen sie mit $\{a_n\}_{n\geq 5}$

3.1.3 Beispiele: (a) Sei $a_n := c$ für alle $n \in \mathbb{N}_0$ für ein fest gewähltes $c \in \mathbb{R}$ (konstante Folge);

(b) Ein Folge der Form $a_n = a_0 + nd, n \geq 0$ für ein fest gewähltes $d \in \mathbb{R}$ heisst arithmetische Folge. Sie ist dadurch charakterisiert, dass die Differenz zweier aufeinander folgender Glieder konstant ist:

$$a_{n+1} - a_n = d \qquad \forall n \geq 0.$$

(c) Ein Folge der Form $a_n := a_0 q^n, n \geq 0$ ($q \in \mathbb{R}, q \neq 0$ fest gewählt) heisst geometrische Folge. In diesem Fall ist der Quotient zweier aufeinander folgender Glieder konstant (falls $a_0 \neq 0$):

$$\frac{a_{n+1}}{a_n} = q \qquad \forall n \geq 0.$$

Häufig kommen auch sogenannte rekursiv definierte Folgen vor. Man gibt einen oder mehrere Anfangswerte vor und eine Vorschrift, wie sich die Folgenglieder aus den vorangehenden Gliedern berechnen lassen.

(d) Sei $\{a_n\}_{n\geq 0}$ gegeben durch

$$a_0 = 1, \quad a_{n+1} = (n+1)a_n \qquad \text{für alle } n \geq 0$$

(Das ist also die Folge $a_n = n!$).

(e) Die Folge der Fibonacci-Zahlen ist gegeben durch

$$a_0 = a_1 = 1, \quad a_n = a_{n-1} + a_{n-2}, \qquad \text{für } n \geq 2$$

Die ersten Glieder sind $1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$.

Eine Folge reeller Zahlen $\{a_n\}_{n \in \mathbb{N}_0}$ heisst *beschränkt*, wenn die Menge der Folgenglieder beschränkt ist, wenn also sämtliche Folgenglieder in einem endlichen Intervall liegen:

$$K_1 \leq a_n \leq K_2$$

(oder dazu äquivalent $|a_n| \leq K$) gilt für alle $n \geq 0$ mit gewissen Konstanten K_1, K_2 (bzw. K).

Die Folgen im Beispiel (a) sind beschränkt. Im Beispiel (c) ist die Folge genau dann beschränkt, wenn $|q| \leq 1$ ist (oder $a_0 = 0$). Die Folgen im Beispiel (b) (für $d \neq 0$), (d) und (e) sind nicht beschränkt.

3.1.4 Definition: Eine reelle Zahlenfolge $\{a_n\}_{n \geq 0}$ heisst konvergent, wenn es eine Zahl $a \in \mathbb{R}$ mit der folgenden Eigenschaft gibt: Zu jedem (beliebig kleinen) $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}_0$ so daß

$$|a_n - a| < \varepsilon$$

für alle $n \geq N$.

Die Zahl a heißt Grenzwert oder Limes der Folge und man schreibt

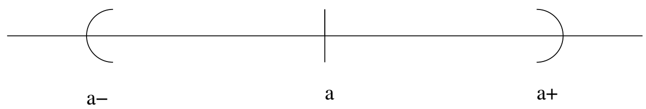
$$\lim_{n \to \infty} a_n = a$$

(oder $a_n \to a$ für $n \to \infty$ oder auch nur $a_n \to a$)

Man sagt auch, daß $\{a_n\}_n$ gegen a konvergiert. Eine Folge die gegen 0 konvergiert heißt Nullfolge.

Eine Folge $\{a_n\}_n$ die nicht konvergent ist, heißt divergent.

Man kann sich die Konvergenz Folge $\{a_n\}$ mit Limes a wie folgt veranschaulichen: jedes (noch so kleine) Intervall $I = (a - \varepsilon, a + \varepsilon)$ mit Mittelpunkt a enthält fast alle Folgenglieder (d.h. alle bis auf endliche viele).



3.1.5 Satz (i) (Eindeutigkeit des Grenzwerts): Ist $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} a_n = b$, so gilt $a = b$.

(ii) Jede konvergente Zahlenfolge ist beschränkt.

Beweis: (i) Angenommen es ist $a \neq b$. Dann ist $\varepsilon :\ = \frac{|a-b|}{2} > 0$. Dann gibt es $N_1, N_2 \in \mathbb{N}_0$, so dass:

$$|a_n - a| < \varepsilon \quad \forall n \geq N_1, \qquad |a_n - b| < \varepsilon \quad \forall n \geq N_2.$$

Für ein fest gewähltes $n \geq \max(N_1, N_2)$ gilt deshalb

$$|a_n - a| < \varepsilon \quad \text{und} |a_n - b| < \varepsilon$$

also nach der Dreiecksungleichung

$$|a - b| = |(a - a_n) + (a_n - b)| \leq |a - a_n| + |a_n - b| < \varepsilon + \varepsilon = |a - b|,$$

ein Widerspruch.

(ii) Sei $\lim_{n\to\infty} a_n = a$ und $\varepsilon > 0$ beliebig (wir können z.B. $\varepsilon = 1$ wählen). Dann gibt es ein $N \in \mathbb{N}_0$, so dass $|a_n - a| < \varepsilon$ für alle $n \geq N$ d.h.

$$a - \varepsilon < a_n < a + \varepsilon \qquad \forall n \geq N,$$

d.h. die Menge $\{a_N, a_{N+1}, \ldots\}$ ist beschränkt ($a - \varepsilon$ ist eine untere und $a + \varepsilon$ eine obere Schranke). Fügen wir zu dieser Menge noch die endlich vielen Elemente $a_0, a_1, \ldots, a_N$ hinzu, so erhalten wir offenbar wieder eine beschränkte Menge (eine obere Schranke ist z.B. $\max(a_0, a_1, \ldots, a_N, a + \varepsilon)$). $\square$

3.1.6 Beispiele: (a) Die Folge $\{\frac{1}{n}\}_{n\geq 1}$ konvergiert gegen 0.

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

Sei $\varepsilon > 0$ eine beliebig kleine aber fest vorgegebene reelle Zahl. Nach dem Archimedischen Axiom (A4) gibt es ein $N \in \mathbb{N}_0$ mit $N = N > \frac{1}{\varepsilon}$ (man wende das Archimedische Axiom für $x = 1$ und $y = \frac{1}{\varepsilon}$ an). Für $n \geq N$ gilt dann

$$0 < \frac{1}{n} \leq \frac{1}{N} < \varepsilon$$

also $|0 - \frac{1}{n}| < \varepsilon$ für alle $n \geq N$.

(b) Für eine reelle Zahl x mit $|x| < 1$ ist die geometrische Folge $\{x^n\}_{n\geq 0}$ eine Nullfolge (das kann man leicht mit Hilfe der Aufgabe 2 (b) des zweiten Übungsblatts zeigen).

Für $|x| \geq 1$ ist $\{x^n\}_{n\geq 0}$ divergent. Ist $x > 1$ so sind die Folgenglieder x^n für hinreichend grosse n, grösser als jede vorgegebene Schranke $K > 0$. Man sagt, dass $\{x^n\}_{n\geq 0}$ *uneigentlich gegen* $+\infty$ *konvergiert.*

3.1.7 Definition: Eine Folge reeller Zahlen $\{a_n\}_{n\in\mathbb{N}}$ konvergiert uneigentlich gegen ∞ (bzw. $-\infty$), wenn es zu jedem $K > 0$ ein $N \in \mathbb{N}_0$ gibt mit

$$a_n \geq K$$

(bzw. $a_n \leq -K$) für alle $n \geq N$. Dann schreibt man auch $\lim\limits_{n\to\infty} a_n = \infty$ (bzw. $= -\infty$) oder auch $a_n \to \infty$ (bzw. $a_n \to -\infty$) für $n \to \infty$.

3.1.8 Satz (Rechenregeln für Grenzwerte): Seien $\{a_n\}_{n\in\mathbb{N}_0}, \{b_n\}_{n\in\mathbb{N}_0}$ konvergente reelle Zahlenfolgen mit $\lim\limits_{n\to\infty} a_n = a$, $\lim\limits_{n\to\infty} b_n = b \in \mathbb{R}$. Dann gilt:

(i) $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n$.

(ii) $\lim\limits_{n\to\infty} (a_n b_n) = \lim\limits_{n\to\infty} a_n \lim\limits_{n\to\infty} b_n$.

(iii) Ist $a \neq 0$, dann gibt es ein $n_1 \in \mathbb{N}$ mit $a_n \neq 0$ für alle $n \geq n_1$ und für die Folgen $\{a_n\}_{n\geq n_1}, \{b_n\}_{n\geq n_1}$ gilt:

$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{a}, \quad \lim_{n\to\infty} \frac{b_n}{a_n} = \frac{b}{a}.$$

(iv) $\lim\limits_{n\to\infty} |a_n| = |a|$.

(v) Ist $a_n \geq 0$ für alle n so gilt:

$$\lim_{n\to\infty} \sqrt{a_n} = \sqrt{a}.$$

Beweis: Wir geben einen ausführlichen Beweis von (ii) und skizzieren nur die Beweise der anderen Aussagen.

(i) folgt aus

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| \to 0$$

(ii) Es gilt

$$|a_n b_n - ab| = |a_n(b_n - b) + b(a_n - a)| \leq |a_n||b_n - b| + |b||a_n - a|$$

Nach Satz 3.1.5 ist die Folge $\{a_n\}$ beschränkt. Sei etwa

$$|a_n| \leq K \qquad \text{für alle } n.$$

Sei $\varepsilon > 0$ vorgegeben. Dann gibt es $N_1, N_2 \in \mathbb{N}_0$, so dass

$$|a_n - a| < \frac{\varepsilon}{2|b|} \qquad \text{für alle } n \geq N_1,$$
$$|b_n - b| < \frac{\varepsilon}{2K} \qquad \text{für alle } n \geq N_2.$$

Für $n \geq N := \max(N_1, N_2)$ folgt:

$$|a_n b_n - ab| < K\frac{\varepsilon}{2K} + |b|\frac{\varepsilon}{2|b|} = \varepsilon$$

(iii) Ist $a \neq 0$, so enthält das Intervall $(a - \frac{|a|}{2}, a + \frac{|a|}{2})$ nicht den Nullpunkt, aber sämtliche Folgenglieder ab einem Index n_1. Für $n \geq n_1$ gilt

$$|\frac{1}{a_n} - \frac{1}{a}| = \frac{|a_n - a|}{|a_n a|} \leq \frac{2}{|a|^2}|a_n - a| \to 0$$

Also gilt $\frac{1}{a_n} \to \frac{1}{a}$. Die zweite Behauptung folgt aus (ii).

(iv) folgt aus

$$||a_n| - |a|| \leq |a_n - a| \to 0$$

(v) 1. Fall: $a \neq 0$. Dann gibt es ein $n_1 \in \mathbb{N}_0$, so dass $a_n \geq \frac{a}{2}$ für alle $n \geq n_1$ gilt, also

$$|\sqrt{a_n} - \sqrt{a}| = |\frac{(\sqrt{a_n} - \sqrt{a})(\sqrt{a_n} + \sqrt{a})}{\sqrt{a_n} + \sqrt{a}}| \leq \frac{1}{\sqrt{\frac{a}{2}} + \sqrt{a}}|a_n - a| \to 0.$$

2. Fall: $a = 0$. Für ein vorgegebenes (beliebig kleines) $\varepsilon > 0$ wähle $N \in \mathbb{N}_0$, so dass $a_n \leq \varepsilon^2$ für alle $n \geq N$ gilt. Für diese n ist dann $\sqrt{a_n} < \varepsilon$, also $\sqrt{a_n} \to 0$. $\square$

3.1.9 Beispiele: (a) $\lim\limits_{n\to\infty} \frac{n+1}{n} = \lim\limits_{n\to\infty} (1+\frac{1}{n}) = 1 + \lim\limits_{n\to\infty} \frac{1}{n} = 1.$

(b) Aus $\lim\limits_{n\to\infty} \frac{3}{n^2} = 3\left(\lim\limits_{n\to\infty} \frac{1}{n}\right)^2 = 0$, $\lim\limits_{n\to\infty} \frac{2}{n^4} = 0$, $\lim\limits_{n\to\infty} \frac{12}{n} = 0$, $\lim\limits_{n\to\infty} \frac{6}{n^4}$ folgt:

$$\lim_{n\to\infty} \sqrt{\frac{6n^4+3n^2+2}{7n^4+12n^3+6}} = \lim_{n\to\infty} \sqrt{\frac{6+\frac{3}{n^2}+\frac{2}{n^4}}{7+\frac{12}{n}+\frac{6}{n^4}}} = \sqrt{\frac{6}{7}}$$

Eine wichtige Methode, das Konvergenzverhalten und gegebenenfalls den Grenzwert einer Folge $\{a_n\}_n$ zu bestimmen, besteht darin, die Folgenglieder a_n geeignet in der Form $b_n \le a_n \le c_n$ abzuschätzen mit leichter zugängliche Folgen b_n und c_n.

3.1.10 Lemma: Sei $\{a_n\}_n$ eine reelle Zahlenfolge. Lassen sich für $n \ge n_1$ die Glieder der Folge nach oben und unten abschätzen durch $b_n \le a_n \le c_n$ mit $\lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} c_n = c$, so ist die Folge $\{a_n\}_n$ konvergent und es gilt $\lim\limits_{n\to\infty} a_n = c$.

Beweis: Für jedes $\varepsilon > 0$ gilt

$$c-\varepsilon < b_n \le a_n \le c_n < c+\varepsilon$$

für alle hinreichend grossen n, also $a_n \to c$. $\square$

3.1.11 Beispiel: (a) $\lim\limits_{n\to\infty} \frac{n}{2^n} = 0$, denn für $n \ge 4$ ist $n^2 \le 2^n$ (das lässt sich leicht durch vollständige Induktion beweisen) und daher

$$0 \le \frac{n}{2^n} \le \frac{1}{n}.$$

(b) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1.$

Beweis: Sei $a_n = \sqrt[n]{n} - 1$. Dann gilt für $n \ge 2$

$$n = (a_n+1)^n = \sum_{k=0}^{n} \binom{n}{k} a_n^k \ge 1 + \binom{n}{2} a_n^2 = 1 + \frac{n(n-1)}{2} a_n^2$$

und damit $0 \le a_n^2 \le \frac{2}{n}$. Es folgt $a_n^2 \to 0$ für $n \to \infty$ und daher $a_n = \sqrt{a_n^2} \to 0$ (nach Satz 3.1.8, (v)).

(c) Für $a > 0$ ist $\lim\limits_{n\to\infty} \sqrt[n]{a} = 1$. Nach Satz 3.1.8, (iii) genügt es dass im Fall $a \ge 1$ zu betrachten. Dann gilt aber $1 \le \sqrt[n]{a} \le \sqrt[n]{n}$ für $n \ge a$ und die Behauptung folgt ergibt sich mit dem vorangegangenen Beispiel.

3.1.12 Satz: Für zwei konvergente Folgen $\{a_n\}_n, \{b_n\}_n$ mit $a_n \leq b_n$ für alle $n \geq n_1$ gilt $\lim\limits_{n\to\infty} a_n \leq \lim\limits_{n\to\infty} b_n$.

Beweis: Für $c_n = b_n - a_n$ ist $\lim\limits_{n\to\infty} c_n \geq 0$ zu zeigen. Wäre $\lim\limits_{n\to\infty} c_n < 0$ dann müssten die c_n ab einem Index N negativ sein, ein Widerspruch. $\square$

3.1.13 Folgerung: Sei $\{a_n\}_n$ eine konvergente Folge reeller Zahlen mit $A \leq a_n \leq B$ für alle $n \geq n_1$. Dann gilt auch $A \leq \lim\limits_{n\to\infty} a_n \leq B$.

3.1.14 Definition: Sei $\{a_n\}_n$ Folge reelle Zahlen.

(i) $\{a_n\}_n$ heisst monoton wachsend, wenn für alle n gilt: $a_n \leq a_{n+1}$.

(ii) $\{a_n\}_n$ heisst monoton fallend, wenn für alle n gilt: $a_n \geq a_{n+1}$.

(iii) $\{a_n\}_n$ heisst monoton, wenn sie monoton wachsend oder fallend ist.

Gilt in (i) (bzw. (ii)) die strikte Ungleichung $a_n < a_{n+1}$ (bzw. $a_n > a_{n+1}$) für alle n so heisst $\{a_n\}_n$ streng monoton wachsend (bzw. fallend).

3.1.15 Satz: Jede monotone und beschränkte Folge von reellen Zahlen ist konvergent.

Beweis: Sei $\{a_n\}_{n\geq 0}$ monoton wachsend und $a = \sup\{a_n; n \geq 0\}$ die kleinste obere Schranke der a_n (sie existiert nach dem Vollständigkeitsaxiom (A5)). Sei $\varepsilon > 0$ vorgegeben. Da $a-\varepsilon$ keine obere Schranke ist, gibt es ein Folgenglied a_N mit $a_N > a - \varepsilon$. Aufgrund der Monotonie gilt dann $a - \varepsilon < a_N \leq a_n \leq a$ für alle $n \geq N$, also $a_n \to a$.

Der Fall einer monoton fallenden Folge $\{a_n\}$ wird analog behandelt, jedoch mit $a = \inf\{a_n | n \geq 0\}$, oder man geht zur monoton wachsenden Folge $\{-a_n\}_n$ über und benutzt, dass $-\lim\limits_{n\to\infty}(-a_n) = \lim\limits_{n\to\infty} a_n$. $\square$

3.1.16 Beispiel: Wir betrachten die rekursiv definierte Folge $a_0 = 1, a_{n+1} = \frac{6(1+a_n)}{7+a_n}$ für $n \geq 0$.

Wir zeigen durch vollständige Induktion, dass $\{a_n\}$ monoton wachsend ist und das 2 und 0 eine obere bzw. untere Schranke von $\{a_n | n \geq 0\}$ sind. Zum Beweis benötigen wir das folgende einfache Tatsache, die sich durch einfache Umformungen zeigen lässt:

Sind $x, y \in \mathbb{R}$ reelle Zahlen mit $0 \leq x \leq y \leq 2$ so gilt $0 \leq \frac{6(1+x)}{7+x} \leq \frac{6(1+y)}{7+y} \leq 2$.

Induktionsanfang: Es ist $0 < a_0 = 1 < \frac{12}{7} = \frac{6(1+1)}{7+1} = a_1$.

Induktionsschritt $n-1 \to n$: Ist $0 \leq a_{n-1} < a_n \leq 2$ für ein $n > 0$, so gilt auch

$$0 \leq \frac{6(1+a_{n-1})}{7+a_{n-1}} = a_n \leq \frac{6(1+a_n)}{7+a_n} = a_{n+1} \leq 2.$$

Nach Satz 3.1.15 ist $\{a_n\}$ konvergent. Sei a der Grenzwert. Der Übergang zum Limes in der Rekursionsformel liefert dann die Gleichung

$$a = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{6(1+a_n)}{7+a_n} = \frac{6(1+a)}{7+a},$$

wobei die letzte Gleichheit aus den Rechenregeln für den Grenzwert folgt. Also gilt $a^2 + a - 6 = 0$, d.h. $a = 2$ oder $a = -3$. Nach 3.1.13 ist $0 \leq a \leq 2$ und damit $a = 2$.

3.1.17 Beispiel: Als weitere Anwendung von Satz 3.1.15 betrachten wir die Folge $c_n = \left(1 + \frac{1}{n}\right)^n, n \geq 1$.

Wegen

$$\begin{aligned}
\frac{c_n}{c_{n-1}} &= \frac{(n+1)^n}{n^n} \frac{(n-1)}{n^{n-1}} = \frac{n}{n-1} \frac{(n^2-1)^n}{n^{2n}} \\
&= \frac{n}{n-1} \left(1 - \frac{1}{n^2}\right)^n \overset{(*)}{\geq} \frac{n}{n-1} \left(1 - n\frac{1}{n^2}\right) \\
&= \frac{n}{n-1} \left(\frac{n-1}{n}\right) = 1.
\end{aligned}$$

ist $\{c_n\}$ monoton wachsend. Dabei folgt $(*)$ aus der

Bernoullische Ungleichung (Aufgabe 2 (a), 2. Übungsblatt): Ist $x \in \mathbb{R}, x \geq -1$ und $n \geq 1$, so gilt $(1+x)^n \geq 1 + nx$.

Ausserdem ist $\{c_n\}$ beschränkt, denn

$$c_n = \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k}$$

und für die einzelnen Summanden $\binom{n}{k}\frac{1}{n^k}$ für $k \geq 1$ gilt

$$\binom{n}{k} \frac{1}{n^k} = \frac{n(n-1)\dots(n-k+1)}{n^k} \frac{1}{k!} = \frac{n}{n} \frac{(n-1)}{n} \cdots \frac{n}{(n-k+1)} \frac{1}{k!}$$

$$\le \frac{1}{k!} = \frac{1}{2}\frac{1}{3}\cdots\frac{1}{k} \le \frac{1}{2^{k-1}}.$$

Es folgt

$$c_n \le \sum_{k=0}^{n} \frac{1}{k!} \le 3.$$

Der Grenzwert $e\colon = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = 2,718281\ldots$ heisst *Eulersche Zahl.*

3.1.18 Definition: Sei $\{a_n\}_{n\ge 0}$ eine Folge. Eine Teilfolge von $\{a_n\}_{n\ge 0}$ ist eine Folge $\{b_k\}_{k\ge 0}$ für die es eine streng monoton wachsende Folge $\{n_k\}_{k\ge 0}$ natürlicher Zahlen gibt, so dass

$$b_k = a_{n_k} \text{ für alle } k \in \mathbb{N}.$$

Man sagt oft einfach: Sei $\{a_{n_k}\}$ eine Teilfolge von $\{a_n\}$. Eine Teilfolge entsteht aus der Originalfolge durch weglassen von Gliedern, so dass aber noch eine unendliche Folge übrig bleibt.

3.1.19 Beispiel: Die Folge $\{\frac{1}{2n}\}$ ist eine Teilfolge der Folge $\{\frac{(-1)^n}{n}\}$.

3.1.20 Bemerkung: Konvergiert die Folge $\{a_n\}_{n\ge 0}$ gegen a, so konvergiert auch jede Teilfolge $\{a_{n_k}\}$ von $\{a_n\}$ gegen a.

3.1.21 Satz (Existenz monotoner Teilfolgen): Jede reelle Zahlenfolge besitzt eine monotone Teilfolge.

Beweis: Sei $\{x_n\}_n$ eine reelle Zahlenfolge. Wir nennen ein Folgenglied x_n dominant, wenn es mindestens so gross ist wie alle folgenden Glieder, d.h. wenn

$$x_n \ge x_m \text{ für alle } m \ge n.$$

1. Fall: Es gibt unendlich viele dominante x_n: Dann lassen wir alle anderen Folgenglieder weg und erhalten eine monoton fallende Teilfolge.

2. Fall: Es gibt nur endlich viele dominante x_n: Dann lassen wir zunächst den Anfang der Folge bis zum letzten dominanten Glied weg. Die verbleibende Folge enthält dann kein dominantes Glied mehr, d.h. zu jedem Glied gibt es ein größeres nachfolgendes. Also können wir eine (streng) monoton wachsende unendliche Teilfolge auswählen. □

Als nächstes wollen wir ein hinreichendes und notwendiges Kriterium für die Konvergenz einer reellen Folge diskutieren.

3.1.22 Definition: Eine Folge reeller Zahlen $\{x_n\}_n$ heißt Cauchyfolge, wenn es zu jedem $\varepsilon > 0$ eine $N \in \mathbb{N}_0$ gibt, so daß

$$|x_n - x_m| < \varepsilon \qquad \text{für alle } n, m \geq N.$$

3.1.23 Bemerkung: Eine konvergente Folge ist eine Cauchyfolge. Ist nämlich $\{x_n\}_n$ konvergent mit Grenzwert a und $\varepsilon > 0$, so gibt es ein $N \in \mathbb{N}$, so daß

$$|x_n - a| < \frac{\varepsilon}{2} \text{ für alle } n \geq N.$$

Dann ist aber

$$|x_n - x_m| = |(x_n - a) + (a - x_m)| \leq |x_n - a| + |x_m - a| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

für alle $m, n \geq N$.

Umgekehrt gilt aber auch:

3.1.24 Satz (Cauchysches Konvergenzkriterium:) Jede Cauchyfolge ist konvergent.

Beweis: Sei $\{x_n\}_n$ eine Cauchyfolge. Zunächst ist $\{x_n\}_n$ beschränkt. Es gibt nämlich ein $N \in \mathbb{N}$ so dass

$$|x_n - x_m| < 1 \qquad \forall\ n \geq N$$

also

$$x_N - 1 < x_n < x_N + 1 \text{ für alle } n \geq N$$

Also ist die Folge $\{x_n\}_{n \geq N}$ beschränkt und die endlich vielen Glieder $x_0, \ldots, x_{N-1}$ ändern daran nichts.

Nach Satz 3.1.21 gibt es eine monotone Teilfolge $\{x_{n_k}\}$ von $\{x_n\}$. Diese ist natürlich wieder beschränkt, also konvergent. Sei a ihr Grenzwert. Wir wollen zeigen, daß die Originalfolge $\{x_n\}_n$ auch gegen a konvergiert. Sei $\varepsilon > 0$ vorgegeben. Dann gibt es ein $K \in \mathbb{N}_0$ so daß

$$|a - x_{n_k}| < \frac{\varepsilon}{2} \qquad \text{für alle } k \geq K$$

und ein $N_1 \in \mathbb{N}_0$ so daß

$$|x_n - x_m| < \frac{\varepsilon}{2} \qquad \text{für alle } m, n \geq N_1.$$

Für $m \geq N = \max(N_1, K)$ gilt dann

$$|x_m - a| \leq |x_m - x_{n_m}| + |x_{n_m} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

3.1.25 Beispiel: Sei $\{x_n\}_{n\geq 0} = \frac{a_{n+1}}{a_n}$ die Folge der Quotienten aufeinander folgender Fibonacci Zahlen

$$\frac{1}{1}, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \frac{13}{8}, \dots,$$

d.h. $x_n = \frac{a_{n+1}}{a_n}$ mit $a_0 = 1 = a_1$ und $a_{n+1} = a_n + a_{n-1}$ für $n \geq 1$. Es gilt also:

$$x_0 = 1, \quad x_{n+1} = 1 + \frac{1}{x_n} \quad \text{für } n \geq 0.$$

Wir wollen zeigen, daß $\{x_n\}_n$ konvergent ist und den Grenzwert berechnen. Wir zeigen zunächst durch Induktion, dass $\frac{3}{2} \leq x_n \leq 2$ für alle $n \geq 1$.

Beweis durch Induktion: $n = 1 \quad \surd$.

Induktionsschritt $n \to n+1$: Es ist $x_{n+1} = 1 + \frac{1}{x_n} \geq 1 + \frac{1}{2} = \frac{3}{2}$ und $x_{n+1} = 1 + \frac{1}{x_n} \leq 1 + \frac{2}{3} < 2$. □

Für die Differenz zweier aufeinander folgender Glieder gilt daher

$$x_{n+1} - x_n = \frac{1}{x_n} - \frac{1}{x_{n-1}} = \frac{x_{n-1} - x_n}{x_{n-1}x_n}$$

und somit

$$|x_{n+1} - x_n| = \frac{|x_n - x_{n-1}|}{x_{n-1}x_n} \leq \frac{2}{3}|x_n - x_{n-1}| \qquad \forall\, n \geq 1.$$

Wieder mit vollständiger Induktion erhalten wir

$$|x_{n+1} - x_n| \leq \left(\frac{2}{3}\right)^n |x_1 - x_0| = \left(\frac{2}{3}\right)^n \qquad \forall\, n \geq 1.$$

Für $n > m$ folgt:

$$\begin{aligned}
|x_m - x_n| &\leq |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \ldots + |x_{m+1} - x_m| \\
&\leq \left(\frac{2}{3}\right)^{n-1} + \left(\frac{2}{3}\right)^{n-2} + \ldots + \left(\frac{2}{3}\right)^m \\
&= \left(\frac{2}{3}\right)^m \left(1 + \frac{2}{3} + \ldots + \left(\frac{2}{3}\right)^{n-m-1}\right) \\
&= \left(\frac{2}{3}\right)^m \frac{1 - \left(\frac{2}{3}\right)^{n-m}}{1 - \frac{2}{3}} \leq 3\left(\frac{2}{3}\right)^m.
\end{aligned}$$

Sei $\varepsilon > 0$ und $N \in \mathbb{N}$, so daß $3\left(\frac{2}{3}\right)^N < \varepsilon$. Dann gilt:

$$|x_n - x_m| < \varepsilon \qquad \forall\, n, m \geq$$

d.h. $\{x_n\}_n$ ist eine Cauchyfolge. Sei $a = \lim\limits_{n\to\infty} x_n$ der Grenzwert. Wegen $x_n \geq 1$ für alle n ist $a \geq 1$. Es gilt

$$a = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} 1 + \frac{1}{x_n} = 1 + \frac{1}{a}$$

ist $a^2 = 1 + a$, also $a = \frac{1+\sqrt{5}}{2}$.

3.1.26 Definition: Ein Punkt $a \in \mathbb{R}$ heißt Häufungspunkt der reellen Zahlenfolge $\{x_n\}_{n\in\mathbb{N}_0}$, wenn diese eine gegen a konvergente Teilfolge besitzt.

3.1.27 Beispiele: (a) Ist $\lim\limits_{n\to\infty} x_n = a$, so ist a ein Häufungspunkt und zwar der einzige.

(b) Die Folge $\{(-1)^n\}_{n\geq 0}$ hat die beiden Häufungspunkte 1 und -1.

(c) Die Folge $\{n^2\}_{n\geq 0}$ hat keinen Häufungspunkt.

3.1.28 Satz (Bolzano-Weierstraß): Jede beschränkte Folge von reellen Zahlen hat mindestens einen Häufungspunkt.

Beweis: Eine beschränkte Folge enthält eine monotone Teilfolge. Diese ist wieder beschränkt, also konvergent. $\square$

3.1.29 Limes superior und Limes inferior: Sei $\{a_n\}_{n\geq 0}$ eine Folge von reellen Zahlen und $s_n := \sup\{a_k \mid k \in \mathbb{N}_0, k \geq n\} \in \mathbb{R} \cup \{+\infty\}$. Ist $\{a_k \mid k \in \mathbb{N}_0\}$ nach oben beschränkt, so ist $\{s_n\}_{n\geq 0}$ eine monoton fallende Folge reeller Zahlen, denn mit wachsendem n wird die Menge $\{a_k \mid k \geq n\}$ kleiner. Ist $\{s_n \mid n \geq 0\}$ überdies nach unten beschränkt (das ist insbesondere dann der Fall, wenn $\{a_n\}_{n\geq 0}$ auch nach unten beschränkt ist), so ist die Folge $\{s_n\}_{n\geq 0}$ konvergent. Andernfalls konvergiert sie uneigentlich gegen $-\infty$.

Ist $\{a_k \mid k \in \mathbb{N}_0\}$ nach oben beschränkt, so definiert man den *Limes superior* von $\{x_n\}_{n\geq 0}$ durch

$$\limsup a_n = \begin{cases} \lim\limits_{n\to\infty} s_n & \text{falls } \{s_n \mid n \geq 0\} \text{ nach unten beschränkt ist,} \\ -\infty & \text{sonst.} \end{cases}$$

Ist $\{a_n \mid n \in \mathbb{N}_0\}$ nicht nach oben beschränkt, so setzt man

$$\limsup a_n = +\infty.$$

Analog definiert man den *Limes inferior*:

$$\liminf a_n = \lim_{n\to\infty} \inf\{a_k \mid k \geq n\} \in \mathbb{R} \cup \{+\infty\}$$

Offenbar gilt: $\liminf a_n = -\limsup(-a_n)$.

3.1.30 Besipiele: (a) Wir betrachten die Folge $a_n = (-1)^n(1 + \frac{1}{n})$. Dann gilt

$$\limsup a_n = 1, \qquad \liminf a_n = -1$$

(b) Für die Folge

$$a_n = \begin{cases} \frac{1}{n} & \text{für } n \text{ ungerade;} \\ -7n & \text{für } n \text{ gerade.} \end{cases}$$

gilt $\limsup a_n = 0$ und $\liminf a_n = -\infty$.

3.1.31 Lemma: Sei $\{a_n\}_{n\geq 0}$ eine beschränkte Folge reeller Zahlen. Dann ist $\limsup a_n$ der grösste und $\liminf a_n$ der kleinste Häufungspunkt von $\{a_n\}_{n\geq 0}$.

Beweis: Sei a ein Häufungspunkt und sei $\{a_{n_k}\}$ eine Teilfolge von $\{a_n\}$ die gegen a konvergiert. Da $n_m \geq m$ ist, ist $a_{n_m} \in \{a_k \mid k \geq m\}$, also $a_{n_m} \leq s_m :\ = \sup\{a_k \mid k \in k \geq m\}$ und damit $a = \lim a_{n_m} \leq \lim s_m = \limsup a_n$.

Um zu zeigen, dass $\limsup a_n$ ein Häufungspunkt von $\{a_n\}_{n\geq 0}$ ist, bemerken wir zunächst, dass es nach Definition der kleinsten oberen Schranke für jedes $n \in \mathbb{N}_0$ einen Index $i(n) \geq n$ gibt mit

$$s_n - \frac{1}{n} < a_{i(n)} \leq s_n.$$

Wir definieren rekursiv eine streng monoton wachsende Folge von Indizes $n_0 < n_1 < n_2 < \ldots$ durch $n_0 :\ = i(0)$ und $n_{k+1} = i(n_k + 1)$ für $k \geq 0$. Es gilt $n_{k+1} = i(n_k + 1) \geq n_k + 1 > n_k$ und

$$s_{n_k+1} - \frac{1}{n_k + 1} < a_{n_k} \leq s_{n_k+1} \qquad \text{für alle } k \in \mathbb{N}_0.$$

Durch Übergang zum Limes folgt mit Lemma 3.1.10

$$\lim_{k\to\infty} a_{n_k} = \lim_{k\to\infty} s_{n_k+1} = \lim_{n\to\infty} s_n = \limsup a_n,$$

d.h. $\limsup a_n$ ist ein Häufungspunkt.

Das der Limes inferior der kleinste Häufungspunkt ist kann, man analog zeigen, oder aus dem schon Bewiesenen ableiten, indem man es auf die Folge $\{-a_n\}_{n\geq 0}$ anwendet. □

3.1.32 Folgerung: Eine beschränkte Folge $\{a_n\}_{n\geq 0}$ ist genau dann konvergent gegen $a \in \mathbb{R}$, wenn $\limsup a_n = a = \liminf a_n$.

Beweis: Das folgt aus 3.1.31 und Aufgabe 4 des 10. Übungsblatt.

3.1.33 Konvergenz komplexer Zahlenfolgen: Analog zur Konvergenz einer reellen Folge definiert man die Konvergenz von komplexen Folgen.

3.1.34 Definition: Eine Folge von komplexen Zahlen $\{z_n\}_{n\geq 0}$ heisst konvergent, wenn es eine Zahl $a \in \mathbb{C}$ mit der folgenden Eigenschaft gibt: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}_0$, so dass

$$|z_n - a| < \varepsilon$$

für alle $n \geq N$.

Dabei bezeichnet $|\ \ |$ die komplexe Betragsfunktion: für eine komplexe Zahl $z = x + iy$ ist $|z| = \sqrt{x^2 + y^2}$.

Die Zahl a heisst Grenzwert oder Limes der Folge und man schreibt

$$\lim_{n\to\infty} z_n = a \qquad \text{oder} \qquad z_n \to a \text{ für } n \to \infty.$$

Die Rechenregeln (i)–(iv) (Satz 3.1.8) gelten genauso für komplexe Zahlenfolgen (die Beweise verlaufen ganz analog).

3.1.35 Satz: Sei $\{z_n\}_{n\geq 0}$ komplexe Folge und $a \in \mathbb{C}$. Dann gilt: die Folge $\{z_n\}_{n\geq 0}$ konvergiert genau dann gegen a, wenn sowohl die Folge der Realteile $\{\mathrm{Re}(z_n)\}_{n\geq 0}$ gegen $\mathrm{Re}(a)$ als auch die Folge der Imaginärteile $\{\mathrm{Im}(z_n)\}_{n\geq 0}$ gegen $\mathrm{Im}(a)$ konvergieren.

Beweis: Für jede komplexe Zahl z gilt:

$$|\mathrm{Re}(z)|, |\mathrm{Im}(z)| \leq |z| \leq |\mathrm{Re}(z) + i\mathrm{Im}(z)| \leq |\mathrm{Re}(z)| + |i||\mathrm{Im}(z)| = |\mathrm{Re}(z)| + |\mathrm{Im}(z)|$$

(siehe Ungleichungen (8) und (9) auf Seite 24).

Konvergiert $\{z_n\}_{n\geq 0}$ gegen a, so erhalten wir damit

$$|\mathrm{Re}(z_n) - \mathrm{Re}(a)| \leq |z_n - a| \to 0,$$

also $\mathrm{Re}(z_n) \to \mathrm{Re}(a)$. Entsprechend folgt $\mathrm{Im}(z_n) \to \mathrm{Im}(a)$ aus $|\mathrm{Im}(z_n) - \mathrm{Im}(a)| \leq |z_n - a|$.

Gilt umgekehrt $\mathrm{Re}(z_n) \to \mathrm{Re}(a)$ und $\mathrm{Im}(z_n) \to \mathrm{Im}(a)$ so folgt

$$|z_n - a| \leq |\mathrm{Re}(z_n) - \mathrm{Re}(a)| + |\mathrm{Im}(z_n) - \mathrm{Im}(a)| \to 0,$$

also $z_n \to a$. $\square$

Mit Hilfe des Satzes kann man z.B. zeigen, dass eine Folge in $\mathbb{C}$ genau dann konvergiert, wenn sie eine Cauchyfolge ist; dass jede konvergente Folge beschränkt ist und dass jede beschränkte Folge eine konvergente Teilfolge, also einen Häufungspunkt besitzt. Dabei sind die Begriffe "Cauchyfolge" und "Beschränktheit" wie für reelle Folgen definiert (mit der komplexen Betragsfunktion an Stelle des reellen Betrags):

3.1.36 Definition: (a) Eine komplexe Folge $\{z_n\}_{n\geq 0}$ heisst Cauchyfolge, wenn es zu jedem, wenn es zu jedem $\varepsilon > 0$ eine $N \in \mathbb{N}_0$ gibt, so daß

$$|z_n - z_m| < \varepsilon \qquad \text{für alle } n, m \geq N.$$

(b) $\{z_n\}_{n\geq 0}$ heisst beschränkt, wenn es ein $K \in \mathbb{R}, K > 0$ gibt mit

$$|z_n| < \varepsilon \qquad \text{für alle } n \in \mathbb{N}_0.$$

Monotonie hingegen macht für komplexe Folgen keinen Sinn.

3.2 Unendliche Reihen

3.2.1 Definition: Sei $\{a_n\}_{n\geq 0}$ eine Folge von reellen oder komplexen Zahlen.

(a) Durch

$$s_n = a_0 + a_1 + \ldots + a_n = \sum_{k=0}^{n} a_k$$

definieren wir eine neue Folge $\{s_n\}_{n\geq 0}$ die wir mit

$$\sum_{k=0}^{\infty} a_k$$

bezeichnen. Folgen dieser Form heissen unendliche Reihen.

Man nennt s_n die n-te Partialsumme von $\sum_{k=0}^{\infty} a_k$.

(b) Die unendliche Reihe $\sum_{k=0}^{\infty} a_k$ heisst konvergent, wenn die Folge $\{s_n\}_{n\geq 0}$ konvergiert. In diesem Fall schreibt man auch $\sum_{k=0}^{\infty} a_k$ statt $\lim_{n\to\infty} s_n$ und nennt diesen Grenzwert die Summe der unendlichen Reihe. Eine unendliche Reihe die nicht konvergiert, heisst divergent.

Das Symbol $\sum_{k=0}^{\infty} a_k$ hat also zwei Bedeutungen: es bezeichnet die Folge der Partialsummen und ggf. deren Grenzwert.

Ist $\sum_{k=0}^{\infty} a_k$ konvergent und A die Summe, dann konvergiert auch die Folge $\{s_{n+1}\}_{n\geq 0}$ gegen A (sie hat ja dieselben Glieder wie $\{s_n\}_{n\geq 0}$ nur in der Nummerierung um eins verschoben). Damit ergibt sich

$$0 = A - A = \lim_{n\to\infty} s_{n+1} - \lim_{n\to\infty} s_n = \lim_{n\to\infty} (s_{n+1} - s_n) = \lim_{n\to\infty} a_n.$$

Wir erhalten damit die folgende notwendige Bedingung für die Konvergenz einer unendlichen Reihe:

3.2.2 Lemma: Sei $\sum_{k=0}^{\infty} a_k$ eine unendliche Reihe. Dann gilt:

$$\sum_{k=0}^{\infty} a_k \text{ konvergiert} \quad \Longrightarrow \quad \{a_k\}_{k\geq 0} \text{ ist eine Nullfolge}$$

3.2.3 Beipiel: (a) Das wichtigste Beispiel einer unendlichen Reihe ist die *geometrische Reihe*:

$$1 + q + q^2 + \ldots = \sum_{k=0}^{\infty} q^k$$

wobei q eine reelle oder komplexe Zahl ist. Für die Partialsummen

$$s_n = \sum_{k=0}^{n} q^k$$

gilt

$$(1-q)s_n = (\sum_{k=0}^{n} q^k) - (\sum_{k=0}^{n} q^{k+1}) = (1+q+\ldots+q^n) - (q+\ldots+q^n+q^{n+1}) = 1-q^{n+1}$$

also

$$s_n = \begin{cases} \frac{1-q^{n+1}}{1-q} & \text{falls } q \neq 0, \\ n+1 & \text{falls } q = 1. \end{cases}$$

Für $|q| < 1$ gilt $q^n \to 0$ für $n \to \infty$. In diesem Fall ist $\sum_{k=0}^{\infty} q^k$ konvergent mit der Summe

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

Für $|q| \geq 1$ ist $\sum_{k=0}^{\infty} q^k$ dagegen divergent (ist q reell und ≥ 1 so konvergiert $\sum_{k=0}^{\infty} q^k$ uneigentlich gegen $+\infty$).

(b) Die Reihe

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots = \sum_{k=1}^{\infty} \frac{1}{k}$$

heisst *harmonische Reihe*. Sie ist divergent. Dazu betrachten wir für $m \in \mathbb{N}$ die folgende Summe

$$\frac{1}{2^m+1} + \frac{1}{2^m+2} + \ldots + \frac{1}{2^{m+1}-1} + \frac{1}{2^{m+1}} = \sum_{k=2^m+1}^{2^{m+1}} \frac{1}{k}.$$

Jeder einzelne Summand ist $\geq \frac{1}{2^{m+1}}$. Also ist die gesammte Summe

$$\geq \frac{1}{2^{m+1}} + \ldots + \frac{1}{2^{m+1}} \quad (2^m\text{-Summanden}) \quad = \frac{1}{2}.$$

Für ein fest gewähltes $n \in \mathbb{N}$ sei m die grösste ganze Zahl mit $2^m \leq n$. Für die n-te Partialsumme der harmonischen Reihe erhalten wir

$$s_n = \sum_{k=1}^{n} \frac{1}{k} \geq \sum_{k=1}^{2^m} \frac{1}{k} = 1 + \sum_{l=0}^{m-1} \sum_{k=2^l+1}^{2^{l+1}} \frac{1}{2^{m+1}} \geq 1 + \frac{m}{2}.$$

Da mit n auch m gegen unendlich strebt, folgt $s_n \to \infty$ für $n \to \infty$.

Das Beispiel der harmonischen Reihe zeigt, dass es für die Konvergenz einer Reihe $\sum_{k=0}^{\infty} a_k$ nicht ausreicht, dass die Glieder a_n eine Nullfolge bilden.

Wir wollen nun einige hinreichende Kriterien für die Konvergenz von unendlichen Reihen geben.

3.2.4 Satz (Majorantenkriterium): Sei $\sum\limits_{k=0}^{\infty} a_k$ eine (reelle oder komplexe) unendliche Reihe. Sei $\sum\limits_{k=0}^{\infty} b_k$ eine konvergente reelle Reihe, so daß gilt

$$|a_k| \leq b_k \qquad \text{für fast alle } k$$

(d.h. es gibt ein $N \in \mathbb{N}_0$ so dass $|a_k| \leq b_k$ für alle $k \geq N$). Dann ist $\sum_{k=0}^{\infty} a_k$ konvergent.

Beweis: Es genügt zu zeigen, dass die Folge der Partialsummen $s_n = \sum\limits_{k=0}^{n} a_k$ eine Cauchyfolge ist. Wir setzen $t_n := \sum_{k=0}^{n} b_k$. Dann gilt:

$$(13) \qquad |s_n - s_m| = |\sum_{k=m+1}^{n} a_k| \leq \sum_{k=m+1}^{n} |a_k| =\leq \sum_{k=m+1}^{n} b_k = t_n - t_m$$

für alle $m, n \in \mathbb{N}_0$ mit $N \leq m \leq n$. Nach Voraussetzung ist $\{t_n\}_{n\in\mathbb{N}_0}$ konvergent, also eine Cauchyfolge. Die Ungleichung(13) impliziert daher, dass auch $\{s_n\}_{n\in\mathbb{N}_0}$ eine Cauchyfolge ist. $\square$

Sind $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ unendliche Reihen und ist $|a_k| \leq b_k$, so heisst $\sum_{k=0}^{\infty} b_k$ *Majorante* für die Reihe $\sum_{k=0}^{\infty} a_k$.

3.2.5 Beispiele: (a) Die unendliche Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist konvergent, da $\frac{1}{k^2} \leq \frac{1}{k(k-1)}$ für alle $k \geq 2$ ist, und da

$$\sum_{k=2}^{\infty} \frac{1}{k(k-1)} = \lim_{n\to\infty} \sum_{k=2}^{n} \frac{1}{k(k-1)} = \lim_{n\to\infty} \sum_{k=2}^{n} (\frac{1}{k-1} - \frac{1}{k}) =$$

$$\lim_{n\to\infty} (\frac{1}{1} - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + \ldots + (\frac{1}{n-1} - \frac{1}{n}) = \lim_{n\to\infty} (1 - \frac{1}{n}) = 1.$$

Also ist $\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \sum_{k=2}^{\infty} \frac{1}{k^2} \leq 1 + \sum_{k=2}^{\infty} \frac{1}{k(k-1)} = 2$. Mit Hilfe der Theorie der *Fourierreihen* kann man zeigen, dass

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

(b) Die Reihe $\sum_{k=0}^{\infty} \frac{\sin(k^3+1)}{2k^3+4k+1}$ ist konvergent, denn $|\frac{\sin(k^3+1)}{2k^3+4k+1}| \leq \frac{1}{2k^3+4k+1} \leq \frac{1}{k^2}$ und $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergiert.

3.2.6 Definition: Eine unendliche Reihe $\sum_{k=0}^{\infty} a_k$ heisst absolut konvergent, wenn die Reihe $\sum\limits_{k=0}^{\infty} |a_k|$ konvergiert.

Nach dem Majorantenkriterium gilt:

3.2.7 Folgerung: Eine absolut konvergente Reihe ist konvergent.

Wir werden weiter unten sehen, dass absolut konvergente Reihen besonders günstige Eigenschaften habe (z.B. kann man die Glieder beliebig "umordnen" ohne die Summe zu verändern; siehe Satz 3.2.18). Eine Reihe die konvergent, aber nicht absolut konvergent ist, heisst *bedingt konvergent.*

Durch Vergleich mit der geometrischen Reihe erhält man folgendes hinreichende Kriterium.

3.2.8 Satz (Quotientenkriterium): Sei $\sum_{k=0}^{\infty} a_k$ eine unendliche Reihe mit $a_k \neq 0$ für fast alle k. Wenn es eine reele Zahl $0 < q < 1$ gibt, so dass gilt

$$\frac{|a_{k+1}|}{|a_k|} \leq q \qquad \text{für fast alle } k \in \mathbb{N}_0,$$

so ist $\sum_{k=0}^{\infty} a_k$ absolut konvergent.

Beweis: Sei $N \in \mathbb{N}_0$ so gewählt, dass $a_k \neq 0$ für alle $k \geq N$ und

$$\frac{|a_{k+1}|}{|a_k|} \leq q \qquad \forall\ k \geq N. \tag{14}$$

Sei

$$b_k = \begin{cases} 0 & \text{falls } n < N \\ |a_N| q^{k-N} & \text{falls } k \geq N. \end{cases}$$

Durch vollständige Induktion erhält man aus (14):

$$|a_k| \leq b_k \qquad \text{für alle } k \geq N.$$

Da $\sum_{k=0}^{\infty} b_k = |a_N| \sum_{n=0}^{\infty} q^n$ konvergent ist, folgt die Behauptung aus dem Majorandenkriterium. □

Beispiel: Die Reihe $\sum_{k=0}^{\infty} \frac{k^2}{2^k}$ ist konvergent. Es gilt nämlich:

$$\frac{\frac{(k+1)^2}{2^{k+1}}}{\frac{k^2}{2^k}} = \frac{(k+1)^2}{k^2} \frac{1}{2}$$

Da $\frac{k+1}{k} \to 1$ für $k \to \infty$ gibt es ein $N \in \mathbb{N}_0$ mit $\frac{(k+1)^2}{k^2} \leq \frac{3}{2}$ für alle $k \geq N$ (man kann z.B. $N = 5$ wählen).

Wir können im Quotientenkriterium also $q = \frac{3}{2}\frac{1}{2} = \frac{3}{4}$ wählen (oder jedes $\frac{1}{2} < q < 1$).

3.2.9 Satz (Wurzelkriterium): Sei $\sum\limits_{k=0}^{\infty} a_k$ eine (reelle oder komplexe) unendliche Reihe. Wenn es ein $0 < q < 1$ gibt, so dass gilt

$$\sqrt[k]{|a_k|} \leq q \qquad \text{für fast alle } k \in \mathbb{N}_0,$$

so ist die Reihe $\sum\limits_{k=0}^{\infty} a_k$ absolut konvergent.

Beweis: Sei $N \in \mathbb{N}_0$, so dass

$$\sqrt[k]{|a_k|} \leq q \qquad \forall\ k \geq N.$$

Wegen $|a_k| \leq q^k$ für alle $k \geq N$ folgt die Behauptung aus der Konvergenz der geometrischen Reihe $1 + q + q^2 + \ldots$ und dem Majorantenkriterium. $\square$

3.2.10 Bemerkung: Man kann das Wurzelkriteriums auch wie folgt formulieren:

$$\limsup \sqrt[k]{|a_k|} < 1 \quad \Longrightarrow \quad \textstyle\sum_{k=0}^{\infty} a_k \text{ ist absolut konvergent}$$

Ist nämlich $\limsup \sqrt[k]{|a_k|} < 1$ und q eine reelle Zahl mit $\limsup \sqrt[k]{|a_k|} < q < 1$ so gilt

$$\sqrt[k]{|a_k|} \leq q \qquad \text{für fast alle } k \in \mathbb{N}_0,$$

denn andernfalls gäbe es unendlich viele Glieder der Folge $\sqrt[k]{|a_k|}$ mit $\sqrt[k]{|a_k|} \geq q$, aus denen man dann eine konvergente Teilfolge $\sqrt[k_m]{|a_{k_m}|}$ auswählen könnte. Dann wäre $\lim_{m\to\infty} \sqrt[k_m]{|a_{k_m}|}$ ein Häufungspunkt von $\{\sqrt[k]{|a_k|}\}_{k\geq 0}$ der grösser als $\limsup \sqrt[k]{|a_k|}$ ist; ein Widerspruch!

Gibt es umgekehrt ein $q < 1$ mit

$$\sqrt[k]{|a_k|} \leq q \qquad \text{für fast alle } k \in \mathbb{N}_0,$$

so gilt offenbar auch $\limsup \sqrt[k]{|a_k|} \leq q$, also $\limsup \sqrt[k]{|a_k|} < 1$.

3.2.11 Beispiel: Sei $\{a_k\}_{k\geq 0}$ eine Folge mit $a_0 \in \mathbb{N}_0$ und $a_k \in \{0, 1, \ldots, 9\}$ für alle $k \geq 1$. Dann ist die Dezilmaldarstellung

$$a_0, a_1 a_2 a_3 \ldots a_n \ldots :\ = a_0 + \sum_{k=0}^{\infty} \frac{a_k}{10^k}$$

nach dem Wurzelkriterium eine konvergente Reihe, da

$$\sqrt[k]{\frac{a_k}{10^k}} \leq \frac{\sqrt[k]{9}}{10} \leq \frac{3}{10} < 1 \qquad \text{für alle } k \geq 2.$$

3.2.12 Bemerkung: Obwohl die Reihe $\sum\limits_{k=1}^{\infty} \frac{1}{k^2}$ konvergiert, kann man weder das Quotienten- noch durch das Wurzelkriterium anwenden. Denn wegen

$$\frac{\frac{1}{(k+1)^2}}{\frac{1}{k^2}} = \frac{k^2}{(k+1)^2} \to \infty$$

und

$$\sqrt[k]{\frac{1}{k^2}} = \frac{1}{\sqrt[k]{k}^2} \to \infty$$

gibt es kein $q < 1$ mit $\frac{\frac{1}{(k+1)^2}}{\frac{1}{k^2}} \leq q$ (bzw. $\sqrt[k]{\frac{1}{k^2}} \leq q$) für fast alle k. Das bedeutet, dass das Quotienten- und das Wurzelkriterium nur hinrechende aber nicht notwendige Bedingungen für die Konvergenz sind.

3.2.13 Definition: Eine unendliche Reihe der Form $\sum_{k=0}^{\infty}(-1)^k b_k$ mit $b_k \geq 0$ heißt alternierende Reihe.

Für alternierende Reihen gibt es das folgende Konvergenzkriterium:

3.2.14 Satz (Leibniz-Kriterium): Sei $\{a_k\}_{k\geq 0}$ eine monoton fallende Folge nicht-negativer reeller Zahlen mit $\lim\limits_{k\to\infty} a_k = 0$. Dann ist die Reihe

$$\sum_{k=0}^{\infty}(-1)^k a_k$$

konvergent.

Beweis: Für die Folge der Partialsummen $s_n = \sum_{k=0}^{n}(-1)^k a_k$ und $n > m$ gilt:

$$|s_n - s_m| = \begin{cases} |(a_{m+1} - a_{m+2}) + \ldots + (a_{n-1} - a_n)| & \text{falls } n-m \text{ gerade ist,} \\ |(a_{m+1} - a_{m+2}) + \ldots + (a_{n-2} - a_{n-1}) + a_n| & \text{falls } n-m \text{ ungerade ist.} \end{cases}$$

Da alle Terme $(a_{k-1} - a_k)$ und $a_k \geq 0$ sind, können wir die Betragsstriche weglassen. Es folgt:

$$|s_n - s_m| = a_{m+1} - (a_{m+1} - a_{m+2}) - \ldots - \begin{cases} a_n & \text{falls } n-m \text{ gerade ist,} \\ (a_{n-1} - a_n) & \text{falls } n-m \text{ ungerade ist.} \end{cases}$$

also $|s_n - s_m| \leq a_{m+1}$ für alle $m, n \in \mathbb{N}_0$ mit $m \leq n$. Wegen $a_m \to 0$ folgt, dass $\{s_n\}_{n\geq 0}$ eine Cauchyfolge ist. $\square$

3.2.15 Beispiele: (a) (Alternierende harmonische Reihe) Die Reihe $\sum\limits_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$ konvergiert (das ist also ein Beispiel einer bedingt konvergenten Reihe). Man kann zeigen, dass

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}\pm\ldots=\log 2.$$

(b) Die Reihe $\sum_{k=0}^{\infty}(-1)^k\frac{1}{2k+1}$ konvergiert (ebenfalls nur bedingt). Es gilt:

$$1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}\mp\ldots=\frac{\pi}{4}.$$

3.2.16 Umordnungen von Reihen: Wir wolll zunächst am Beispiel der alternierenden harmonischen Reihe zeigen, dass sich bei einer bedingt konvergenten Reihe das Konvergenzverhalten ändert, wenn man die Reihenfolge der Glieder abändert (für unendliche Summen gilt also kein "Kommutativgesetzt"). Die Reihen der positiven bzw. negativen Glieder der alternierenden harmonischen Reihe

$$\begin{array}{l}1+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\ldots\\ \frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\frac{1}{8}+\ldots\end{array}$$

sind beide divergent (die zweite ist einfach die halbe harmonische Reihe und deshalb divergent; die Glieder und deshalb auch die Partialsummen der ersten sind grösser als die der zweiten, weshalb auch die erste Reihe divergent ist). Deshalb kann man von der alternierenden harmonischen Reihe zunächst so viele positive Glieder (der Reihe nach) addieren bis man z.B. über 1959 ist. Dann addiert man solange negative Glieder bis man unter 1959 kommt. Dann addiert man positive Glieder, bis man wieder über 1959 ist. Auf diese Weise summiert man schliesslich alle Glieder der Reihe und hat sie so umgeordnet, dass die neue Reihe nun gegen 1959 konvergiert. Es ist offensichtlich, dass man auf diese Weise statt 1959 auch jeden anderen Grenzwert (oder auch bestimmte Divergenz gegen $+\infty$ oder $-\infty$) durch geeignete Umordnung erzielen kann.

Dieses Phänomen tritt offenbar nicht bei einer Reihe auf, deren positive und negative Glieder jeweils eine konvergente Reihe bilden, d.h. bei einer absolut konvergenten Reihe.

3.2.17 Definition: Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ zwei (reelle oder komplexe) unendliche Reihen. Die Reihe $\sum_{k=0}^{\infty} b_k$ heisst Umordnung von $\sum_{k=0}^{\infty} a_k$, wenn es zu jedem $k \in \mathbb{N}_0$ genau ein $m \in \mathbb{N}_0$ gibt mit $b_k = a_m$, d.h. wenn es eine bijektive Abbildung $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ gibt mit $b_k = a_{\sigma(k)}$ für alle $k \in \mathbb{N}_0$.

3.2.18 Satz (Umordnungssatz): Sei $\sum_{k=0}^{\infty} a_k$ eine absolut konvergente Reihe und $\sum_{k=0}^{\infty} b_k$ eine Umordnung von $\sum_{k=0}^{\infty} a_k$. Dann konvergiert auch $\sum_{k=0}^{\infty} b_k$ absolut und beide Reihen haben dieselbe Summe:

$$\sum_{k=0}^{\infty} b_k = \sum_{k=0}^{\infty} a_k.$$

Beweis: Wir bezeichnen die Partialsummen von $\sum_{k=0}^{\infty} a_k$ mit s_n und die von $\sum_{k=0}^{\infty} b_k$ mit t_n. Ferner sei $A = \lim s_n = \sum_{k=0}^{\infty} a_k$. Es genügt zu zeigen, dass $t_n - s_n$ eine Nullfolge ist, denn dann gilt $t_n = s_n + (t_n - s_n) \to \infty$.

Sei $\varepsilon > 0$. Aufgrund der Konvergenz von $\sum_{k=0}^{\infty} |a_k|$ gibt es ein $n_0 \in \mathbb{N}_0$ mit

$$\sum_{k=n_0}^{\infty} |a_k| < \frac{\varepsilon}{2}$$

Daraus folgt

$$|s_n - s_{n_0}| = |\sum_{k=n_0+1}^{n} a_k| \leq \sum_{k=n_0}^{\infty} |a_k| < \frac{\varepsilon}{2} \qquad \forall n \geq n_0.$$

Wählen wir N gross genug, so kommen irgendwann die Glieder $a_0, a_1, \ldots, a_{n_0}$ alle in t_N vor. Für $n \geq N$ heben sich in der Differenz

$$t_n - s_n = (b_0 + b_1 + \ldots + b_n) - (a_0 + a_1 + \ldots + a_n)$$

also mindestens alle $a_0, a_1, a_2, \ldots, a_{n_0}$ weg. Die Terme b_k die übrig bleiben sind von der Form $b_k = a_m$ mit $m > n$.

Für $n \geq N$ (also wegen $N \geq n_0$ auch $n \geq n_0$) folgt daher

$$|t_n - s_n| \leq |\sum_{k=n+1}^{\infty} |a_k| + \sum_{k=n_0+1}^{\infty} |a_k| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Das zeigt, dass $t_n - s_n$ eine Nullfolge ist.

Da $\sum_{k=0}^{\infty} |b_k|$ eine Umordnung von $\sum_{k=0}^{\infty} |a_k|$ ist, folgt aus dem bereits Bewiesenen auch die Konvergenz von $\sum_{k=0}^{\infty} |b_k|$, d.h. die absolute Konvergenz von $\sum_{k=0}^{\infty} b_k$. $\square$

3.2.19 Die Exponentialfunktion:

3.2.20 Satz und Definition: Für jedes $z \in \mathbb{C}$ ist die Reihe $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ absolut konvergent. Die Funktion

$$\exp : \mathbb{C} \to \mathbb{C}, z \mapsto \exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

heisst Exponentialfunktion.

Beweis: Wir wenden das Quotientenkriterium an (mit $q\colon = \frac{1}{2}$). Für $z \in \mathbb{C}, z \neq 0$ gilt

$$\frac{\left|\frac{z^{k+1}}{(k+1)!}\right|}{\left|\frac{z^k}{k!}\right|} = \frac{|z|}{k+1} \leq \frac{1}{2} \qquad \text{für } k \geq 2|z|.$$

□

3.2.21 Satz: (Funktionalgleichung für die Exponentialfunktion):

$$\exp(z+w) = \exp(z)\exp(w) \qquad \text{für alle } z, w \in \mathbb{C}.$$

Zum Beweis benötigen wir das folgende

3.2.22 Lemma: Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergente (reelle oder komplexe) Reihen. Für $n \in \mathbb{N}_0$ setzen wir

$$c_n\colon = \sum_{k=0}^{n} a_k b_{n-k}.$$

Dann konvergiert die Reihe $\sum_{k=0}^{\infty} c_k$ absolut und es gilt

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{k=0}^{\infty} b_k\right).$$

Beweis: Sei $A\colon = \sum_{k=0}^{\infty} a_k$, $B\colon = \sum_{k=0}^{\infty} b_k$ und $C_n\colon = \sum_{k=0}^{n} c_k$. Wir müssen zeige, dass $C_n \to AB$. Wegen $\widetilde{C}_n\colon = (\sum_{k=0}^{n} a_k)(\sum_{k=0}^{n} b_k) \to AB$ für $n \to \infty$ genügt es zu zeigen, dass $\widetilde{C}_n - C_n$ eine Nullfolge ist. Es gilt

$$\widetilde{C}_n = \sum_{0 \leq i,j \leq n} a_i b_j$$

$$C_n = \sum_{k=0}^{n} \left(\sum_{l=0}^{k} a_l b_{k-l} \right) = \sum_{k=0}^{n} \left(\sum_{i+j=k} a_i b_j \right) = \sum_{0 \le i,j \le n, i+j \le n} a_i b_j$$

Daraus folgt

$$|\widetilde{C}_n - C_n| = |\sum_{(i,j) \in \Delta_n} a_i b_j| \le \sum_{(i,j) \in \Delta_n} |a_i||b_j|$$

wobei $\Delta_n \colon = \{(i,j) \mid 0 \le i \le n, 0 \le j \le n, i+j > n\}$.

Sei $\varepsilon > 0$. Aufgrund der absoluten Konvergenz von $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ gibt es ein $N \in \mathbb{N}_0$ mit

$$\sum_{k=N}^{\infty} |a_k| < \frac{\varepsilon}{2(\sum_{k=0}^{\infty} |b_k|)}, \quad \sum_{k=N}^{\infty} |b_k| < \frac{\varepsilon}{2(\sum_{k=0}^{\infty} |a_k|)}$$

für alle $n \ge N$. Wir werden zeigen, dass

$$(15) \qquad |\widetilde{C}_n - C_n| < \varepsilon \qquad \forall\, n \ge 2N$$

Für ein gegebenes $n \in \mathbb{N}_0$ mit $n \ge 2N$ und ein Paar $(i,j) \in \Delta_n$, also $i + j > n \ge 2N$, ist entweder $i \ge N$ oder $j \ge N$. Es folgt

$$|\sum_{(i,j) \in \Delta_n} a_i b_j| \le \sum_{(i,j) \in \Delta_n} |a_i||b_j| \le \left(\sum_{k=N}^{n} |a_k| \right) \left(\sum_{k=0}^{n} |b_k| \right) + \left(\sum_{k=0}^{n} |a_k| \right) \left(\sum_{k=N}^{n} |b_k| \right)$$

$$\le \left(\sum_{k=N}^{\infty} |a_k| \right) \left(\sum_{k=0}^{\infty} |b_k| \right) + \left(\sum_{k=0}^{\infty} |a_k| \right) \left(\sum_{k=N}^{\infty} |b_k| \right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Das beweist (15) und wir erhalten $\widetilde{C}_n - C_n \to 0$.

Es bleibt noch die absolute Konvergenz von $\sum_{k=0}^{\infty} c_k$ zu zeigen. Für $n \ge 0$ sei $c'_n \colon = \sum_{k=0}^{n} |a_k||b_{n-k}|$. Aus dem bereits Bewiesenen folgt, dass die Reihe $\sum_{k=0}^{\infty} c'_k$ konvergiert (gegen $(\sum_{k=0}^{\infty} |a_k|)(\sum_{k=0}^{\infty} |b_k|)$). Da

$$|c_n| = |\sum_{k=0}^{n} a_k b_{n-k}| \le \sum_{k=0}^{n} |a_k||b_{n-k}| = c'_n$$

folgt die Konvergenz von $\sum_{k=0}^{\infty} |c_k|$ aus dem Majorantenkriterium. □

Beweis von Satz 3.2.21: Nach dem obigen Lemma gilt:

$$\exp(z)\exp(w) = \left(\sum_{k=0}^{\infty} \frac{z^k}{k!} \right) \left(\sum_{k=0}^{\infty} \frac{w^k}{k!} \right) = \sum_{m=0}^{\infty} \left(\sum_{k=0}^{m} \frac{z^k}{k!} \frac{w^{m-k}}{(m-k)!} \right) =$$

$$= \sum_{m=0}^{\infty} \frac{1}{m!} \left(\sum_{k=0}^{m} \frac{m!}{k!(m-k)!} z^k w^{m-k} \right) \overset{(*)}{=} \sum_{m=0}^{\infty} \frac{1}{m!} (z+w)^m = \exp(z+w)$$

wobei die Gleichheit $(*)$ aus der verallgemeinerten Binomischen Formel folgt.
□

3.2.23 Folgerung: (a) $\exp(-z) = (\exp(z)^{-1}$ und $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$.

(b) $\exp(r) = e^r$ für jede rationale Zahl r. Dabei ist $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ die Eulersche Zahl.

Beweis: (a) folgt aus $\exp(z)\exp(-z) = \exp(0) = 1$.

(b) Nach Aufgabe 2 des 10. Übungsblatts gilt zunächst

$$\exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Für eine beliebige positive rationale Zahl $r = \frac{m}{n}, m, n \in \mathbb{N}$ folgt dann

$$\exp(r)^n = \exp(nr) = \exp(m) = \exp(1)^m = (e^r)^n.$$

und damit $\exp(r) = e^r$. Der Fall $r = -\frac{m}{n} < 0$ mit $m, n \in \mathbb{N}$ folgt schliesslich aufgrund von $\exp(-\frac{m}{n}) = \exp(\frac{m}{n})^{-1}$. □