Grundlagen der Mathematik I - II

Kaiserslautern WS 2009/2010-SS 2010

Prof. Dr. Wolfram Decker

19. Februar 2010

Dieses Skript entstand aus von Studenten erstellten Skripten zu meinen Vorlesungen Lineare Algebra I, II (WS 2000/01, SS 2001) bzw. Mathematik für Informatiker I,II,III (WS 2004/05, SS 2005, WS 2005/06) an der Universität Saarbrücken. Ich danke Oliver Barth, Martin Kaiser, Marko Kurz, Sebastian Kirsch bzw. Martin Grochulla, Pascal Gwosdek, Christian Doczkal, Sebastian Meiser für die Erstellung dieser Skripte.

Wolfram Decker `decker@mathematik.uni-kl.de`

Inhaltsverzeichnis

Kapitel 0

Einführung

Auf den Vorlesungen Grundlagen der Mathematik I und II bauen alle weiteren Mathematikvorlesungen auf. Das Beherrschen der vermittelten Beweismethoden und Rechentechniken ist die unabdingbare Voraussetzung für das Verständnis der Mathematik in den höheren Semestern. Im Einzelnen eingeführt werden Grundlagen der Analysis und der Linearen Algebra.

Gegenstand der Analysis ist die von Newton und Leibniz im 17. Jahrhundert begründete Differential- und Integralrechnung. Diese liefert Methoden, die es erlauben, das Änderungsverhalten von Funktionen zu studieren. Hauptinhalt der Vorlesung im ersten Semester ist die Differential- und Integralrechnung einer reellen Variablen. Ausgangspunkt ist die Axiomatik der reellen Zahlen – alle mathematischen Aussagen werden von einigen wenigen Grundeigenschaften der reellen Zahlen abgeleitet. Fundamentale Begriffe sind Konvergenz, Stetigkeit, Differenzierbarkeit und Integrierbarkeit. Der Fall mehrerer Variabler wird im zweiten Semester behandelt.

In der linearen Algebra geht es um das Lösen *linearer* Gleichungssysteme. Allgemein geht das Wort Algebra und das Bemühen um Lösungen von Gleichungen zurück auf die arabischen Mathematiker des 9. Jahrhunderts. Ein zentraler Algorithmus zur Lösung *linearer* Gleichungssysteme ist das von Gauss um 1800 gefundene Eliminationsverfahren. Fundamentale Begriffe wie Vektorräume, lineare Abbildungen, Matrizen und Determinanten werden im ersten Semester behandelt. Im zweiten Semester spielen Eigenwerte, Eigenvektoren und Normalformen für Matrizen sowie Skalarprodukte eine zentrale Rolle.

Beginnen wollen wir mit der Festlegung mathematischer Sprechweisen. Dabei setzen wir zunächst eine gewisse Vertrautheit im Umgang mit den reellen Zahlen voraus. In Kapitel 6 werden wir dann auf die Axiomatik der rellen Zahlen eingehen.

Literatur

ANALYSIS	LINEARE ALGEBRA
Forster, Analysis 1, 2	Fischer, Lineare Algebra
Heuser, Analysis 1, 2	Kowalsky-Michler, Lineare Algebra
Barner-Flohr, Analysis 1, 2	Bosch, Lineare Algebra
Hildebrandt, Analysis 1, 2	Jänich, Lineare Algebra
Königsberger, Analysis 1, 2	

Kapitel 1

Logik und Beweismethoden

1.1 Motivation

Die Regeln der Logik bilden die Grundlagen des mathematischen Argumentierens und damit die Grundlage für mathematische Beweise. Letztere dienen dazu, die Wahrheit mathematischer Aussagen zu verifizieren. □

1.2 Aussagen

Eine **Aussage** ist ein Satz, dem genau einer der Wahrheitswerte wahr (w) oder falsch (f) zugeordnet werden kann.

1.2.1 Beispiel

Aussagen sind:

(i) Kaiserslautern liegt in der Pfalz.

(ii) $2 + 2 = 7$.

Keine Aussagen sind:

(i) Saumagen schmeckt gut.

(ii) $x + 2 = 5$. □

1.3 Verknüpfungen von Aussagen

Mit Hilfe **logischer Operatoren** erhält man neue Aussagen aus bereits vorhandenen Aussagen. Sind etwa A und B Aussagen, so hat man auch die Aussagen:

- $\neg A$ „nicht A“ (**Negation**),
- $A \wedge B$ „A und B“ (**Konjunktion**),
- $A \vee B$ „A oder B“ (**Disjunktion**).

Die Wahrheitswerte dieser verknüpften Aussagen werden durch die folgenden Wahrheitstafeln festgelegt:

A	$\neg A$
w	f
f	w

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

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

Weitere wichtige Verknüpfungen sind:

- $A \Rightarrow B$ „aus A folgt B“ (**Implikation**),
- $A \Leftrightarrow B$ „A ist äquivalent zu B“ (**Aquivalenz**).

Die entsprechenden Wahrheitswerte sind wie folgt erklärt:

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

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

Die Reihenfolge der Ausführungen logischer Operatoren in zusammengesetzten Aussagen kann mit Hilfe von Klammern festgelegt werden.

1.3.1 Beispiel

A	B	$\neg A$	$A \wedge B$	$(\neg A) \vee B$	$(A \wedge B) \Rightarrow ((\neg A) \vee B)$
w	w	f	w	w	w
w	f	f	f	f	w
f	w	w	f	w	w
f	f	w	f	w	w

□

Um Klammern zu sparen, legen wir fest, dass $\neg$ eine höhere Priorität hat als $\vee$, $\wedge$, $\Rightarrow$, $\Leftrightarrow$, und dass $\Rightarrow$ und $\Leftrightarrow$ eine niedrigere Priorität haben als $\neg$, $\vee$, $\wedge$. Dann können wir statt $(A \wedge B) \Rightarrow ((\neg A) \vee B)$ auch $A \wedge B \Rightarrow \neg A \vee B$ schreiben.

Im folgenden bezeichnen wir mit $A, B, C, D, \ldots$ auch **Aussagenvariable**, für die man konkrete Aussagen einsetzen kann. Die Verknüpfung der Variablen und der **logischen Konstanten** w und f mit Hilfe von $\neg$, $\wedge$, $\vee$, $\Rightarrow$ und $\Leftrightarrow$ führt zu einer **logischen Formel**, aus der durch Einsetzen konkreter Aussagen eine zusammengesetzte Aussage wird.

1.3.2 Definition

(i) Eine logische Formel, die unabhängig von den Wahrheitswerten eingesetzter Aussagen immer wahr ist, heißt **Tautologie**. Ist sie immer falsch, so sprechen wir von einem **Widerspruch**.

(ii) Zwei logische Formeln heißen **logisch äquivalent**, wenn ihre Verknüpfung durch $\Leftrightarrow$ eine Tautologie ist. □

1.3.3 Satz

Die folgenden logischen Formeln sind Tautologien:

(i) Assoziativgesetze:

$$(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)\,,$$
$$(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)\,.$$

(ii) Kommutativgesetze:

$$A \wedge B \Leftrightarrow B \wedge A,$$
$$A \vee B \Leftrightarrow B \vee A.$$

(iii) Distributivgesetze:

$$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)\,,$$
$$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)\,.$$

(iv) Identitätsgesetze:

$$A \vee f \Leftrightarrow A,$$
$$A \wedge w \Leftrightarrow A.$$

(v) Komplementgesetze:

$$A \vee \neg A \Leftrightarrow w,$$
$$A \wedge \neg A \Leftrightarrow f.$$

Beweis

Die Gesetze zeigt man durch Aufstellen von Wahrheitstafeln. Zum Beispiel:

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

□

Die zu Beginn von 1.3 angegebenen Wahrheitstafeln zeigen, dass die logischen Formeln $A \Rightarrow B$ und $\neg A \vee B$ äquivalent sind. Weitere äquivalente logische Formeln liefert der folgende Satz:

1.3.4 Satz

Die folgenden logischen Formeln sind Tautologien:

(i) De Morgan'sche Gesetze:

$$\neg\,(A \vee B) \Leftrightarrow \neg A \wedge \neg B,$$
$$\neg\,(A \wedge B) \Leftrightarrow \neg A \vee \neg B.$$

(ii) Idempotenz:

$$A \vee A \Leftrightarrow A,$$
$$A \wedge A \Leftrightarrow A.$$

(iii) Doppelte Verneinung:
$$\neg(\neg A) \Leftrightarrow A.$$

(iv) Kontraposition:
$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A).$$

(v) Kettenschluss:
$$(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C).$$

(vi) indirekter Schluss:
$$(\neg A \Rightarrow f) \Leftrightarrow A.$$

Beweis

Wieder durch Wahrheitstafeln. □

1.4 Quantoren

Abkürzend schreiben wir:

$\forall$: „für alle“,

$\exists$: „es existiert ein“,

$\exists!$: „es existiert genau ein“,

$\nexists$: „es existiert kein“.

1.5 Beweismethoden

Beweise verifizieren in der Regel, dass in mathematischen Sätzen Aussagen wie $A \Rightarrow B$ oder $A \Leftrightarrow B$ wahr sind. Dabei sind A und B konkrete mathematische Aussagen.

Ist eine mathematische Aussage A wahr, so sagen wir auch „es gilt A“. Schreiben wir in einem mathematischen Satz oder Beweis $A \Rightarrow B$, so meinen wir damit, dass $A \Rightarrow B$ gilt. In diesem Fall sagen wir „aus A folgt B“, „A ist hinreichend für B“, oder „B ist notwendig für A“. Ebenso bedeutet $A \Leftrightarrow B$, dass $A \Leftrightarrow B$ gilt. Wir sagen dann auch „A und B sind äquivalent“, oder „A gilt genau dann, wenn B gilt“.

1.5.1 Beweisprinzip der vollständigen Induktion

Die Zahlen 1, 2, 3, ... heißen **natürliche Zahlen**. Sie lassen sich durch die Peano-Axiome charakterisieren. Diese beinhalten das Prinzip der vollständigen Induktion, das man wie folgt formulieren kann:

Für jede natürliche Zahl n sei eine Aussagen $A(n)$ gegeben. Man weise nach,

(1) dass $A(1)$ wahr ist und

(2) dass $A(n) \Rightarrow A(n+1)$ wahr ist für jede natürliche Zahl n .

Dann ist bewiesen, dass $A(n)$ wahr ist für jede natürliche Zahl n.

Der Beweisschritt (1) heißt **Induktionsanfang** und (2) **Induktionsschluss**.

Das Prinzip lässt sich wie folgt abwandeln:

Ersetzt man (1) durch

(1') $A(k)$ ist wahr für eine beliebige aber feste ganze Zahl k,

und (2) durch

(2') $A(n) \Rightarrow A(n+1)$ ist wahr für jede ganze Zahl $n \geq k$,

so beweist man, dass $A(n)$ wahr ist für jede ganze Zahl $n \geqslant k$.
Dabei heißen $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$ **ganze Zahlen**. □

1.5.2 Beispiel

Wir zeigen: Für jede natürlich Zahl n gilt $1 + 2 + \ldots + n = \frac{n \cdot (n+1)}{2}$.

Beweis

Induktion nach n.

- **Induktionsanfang**: $n = 1$
  - Es gilt $1 = \frac{1 \cdot 2}{2} = \frac{1 \cdot (1+1)}{2}$.
- **Induktionsschluss**: $n \Rightarrow n + 1$
  - **Induktionsvoraussetzung**:
    Sei n eine natürliche Zahl, für die gilt $1 + 2 + \ldots + n = \frac{n \cdot (n+1)}{2}$.
  - **Induktionsbehauptung**:
    Dann gilt $1 + 2 + \ldots + n + (n+1) = \frac{(n+1) \cdot (n+2)}{2}$.

  Wir zeigen die Behauptung:

$$\begin{aligned} 1 + 2 + \ldots + (n-1) + n + (n+1) &\overset{IV}{=} \frac{n \cdot (n+1)}{2} + (n+1) \\ &= \frac{n \cdot (n+1) + 2 \cdot (n+1)}{2} \\ &= \frac{(n+1) \cdot (n+2)}{2} \ . \end{aligned}$$

□

1.5.3 Der direkte Beweis

Hier nutzt man aus:

- Gilt jede Implikation einer Kette von Implikationen

$$A \Rightarrow C_1 \Rightarrow C_2 \Rightarrow \ldots \Rightarrow C_n \Rightarrow B,$$

  so gilt auch $A \Rightarrow B$ (dies folgt aus 1.3.4, (v) durch Induktion).

1.5.4 Beispiel

Wir zeigen: Sind a, b reelle Zahlen ≥ 0 mit $a \neq b$, so gilt

$$\overbrace{\frac{a+b}{2}}^{\text{arithmetisches Mittel}} > \overbrace{\sqrt{a \cdot b.}}^{\text{geometrisches Mittel}}$$

Beweis

$$\begin{array}{rrcl} & a, b \text{ reell}, a \neq b,\, a, b & \geqslant & 0 \\ \Rightarrow & (a-b)^2 & > & 0 \\ \Rightarrow & a^2 - 2 \cdot a \cdot b + b^2 & > & 0 \\ \Rightarrow & a^2 + 2 \cdot a \cdot b + b^2 & > & 4 \cdot a \cdot b \\ \Rightarrow & (a+b)^2 & > & 4 \cdot a \cdot b \\ \Rightarrow & \frac{(a+b)^2}{4} & > & a \cdot b \\ \Rightarrow & \frac{a+b}{2} & > & \sqrt{a \cdot b} \end{array}$$

□

1.5.5 Warnung

Im direkten Beweis zeigt man also die Gültigkeit von $A \Rightarrow B$ mit Hilfe einer Kette von Implikationen. Gilt dann A, so gilt auch B. Gilt hingegen B, so kann A durchaus falsch sein:

$$2 = 1 \Rightarrow 2 \cdot 0 = 1 \cdot 0 \Rightarrow 0 = 0.$$

Vergleichen Sie dazu noch einmal die Wahrheitstafel von $A \Rightarrow B$. □

1.5.6 Beweis durch Kontraposition

Hier nutzt man aus:

- $A \Rightarrow B$ ist äquivalent zu $\neg B \Rightarrow \neg A$

(siehe 1.3.4, (iv)). Um die Gültigkeit von $A \Rightarrow B$ zu zeigen, weist man die Gültigkeit von $\neg B \Rightarrow \neg A$ nach.

1.5.7 Der indirekte Beweis

Hier nützt man aus:

- $(\neg A \Rightarrow f)$ ist äquivalent zu A

(siehe 1.3.4, (vi)). Um die Gültigkeit von A zu zeigen, nimmt man an, dass A falsch ist und leitet daraus eine bereits als falsch bekannte Aussage ab. Mit anderen Worten, wir erhalten einen Widerspruch (↯) der zeigt, dass A doch richtig sein muss.

1.5.8 Beispiel

Wir zeigen, dass es unendlich viele Primzahlen gibt.

Beweis

- **Annahme**: Die Behauptung ist falsch.

  $\Rightarrow$ es gibt nur endlich viele Primzahlen, etwa $p_1, \ldots, p_r$,

  $\Rightarrow n = p_1 \cdot \ldots \cdot p_r + 1$ ist keine Primzahl,

  $\Rightarrow$ eine der Primzahlen muss n teilen, etwa p_i,

  $\Rightarrow p_i | 1$ ↯ □

1.6 Summen- und Produktschreibweisen

Um Summen und Produkte zu schreiben, benutzen wir das Summenzeichen Σ und das Produktzeichen Π. Sind etwa $m \leqslant n$ ganzen Zahlen und sind $a_m, \ldots, a_n$ reelle Zahlen, so schreiben wir:

$$\sum_{i=m}^{n} a_i = a_m + \ldots + a_n,$$

$$\prod_{i=m}^{n} a_i = a_m \cdot \ldots \cdot a_n.$$

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

$$\sum_{i=m}^{m-1} a_i = 0 \qquad \text{„leere Summe“},$$

$$\prod_{i=m}^{m-1} a_i = 1 \qquad \text{„leeres Produkt“}.$$

Kapitel 2

Mengen

2.1 Motivation

Die Mengenlehre ist das fundamentale Hilfsmittel zur Spezifizierung mathematischer Objekte und zur Bildung mathematischer Strukturen. Sie wird insbesondere dort gebraucht, wo Strukturen wie Halbgruppen, Monoide, Gruppen, Ringe, Körper, Algebren oder Verbände eine Rolle spielen. □

2.2 Mengentheoretische Sprechweisen

Wir wollen hier nicht auf die schwierige Frage eingehen, wie der Begriff „Menge“ erklärt werden kann. Dies ist Gegenstand der mathematischen Grundlagenforschung. Naiv könnte man sagen: Unter einer **Menge** verstehen wir eine Zusammenfassung verschiedener wohlbestimmter Objekte zu einem Ganzen. Zwei Mengen heißen gleich, wenn sie dieselben Objekte enthalten.

2.2.1 Beispiel

Für uns wichtige Mengen sind:

$\mathbb{N}$ = Menge der natürlichen Zahlen,

$\mathbb{N}_0$ = Menge der natürlichen Zahlen mit der Null,

$\mathbb{Z}$ = Menge der ganzen Zahlen,

$\mathbb{Q}$ = Menge der rationalen Zahlen,

$\mathbb{R}$ = Menge der reellen Zahlen,

$\mathbb{C}$ = Menge der komplexen Zahlen. □

Gehört ein Objekt x zu einer Menge X, so nennen wir x ein **Element** von X und schreiben $x \in X$. Andernfalls schreiben wir $x \notin X$.

2.2.2 Beispiel

$$2 \in \mathbb{N}, \quad -4 \notin \mathbb{N}.$$ □

Endliche Mengen kann man im Prinzip durch eine vollständige Liste ihrer Elemente angeben. Wir schreiben $X = \{x_1, \ldots, x_n\}$, wenn die Menge X aus den Elementen $x_1, \ldots, x_n$ besteht. Dabei müssen die Elemente nicht notwendig verschieden sein und die Reihenfolge ist gleichgültig.

2.2.3 Beispiel

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

$x_1, \ldots, x_n$ heißen **paarweise verschieden**, wenn $x_i \neq x_j$ für $i \neq j$. In diesem Fall ist $X = \{x_1, \ldots, x_n\}$ eine **n-elementige Menge**, wir schreiben $|X| = n$.

Beliebige Mengen lassen sich beschreiben, indem man charakteristische Eigenschaften ihrer Elemente angibt.

2.2.4 Beispiel

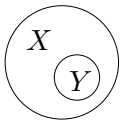
(i) $X_1 = \{\text{Einwohner von Kaiserslautern}\}$,

(ii) $X_2 = \{n \in \mathbb{N} \mid n \leqslant 5\} = \{1, 2, 3, 4, 5\}$,

(iii) $X_3 = \{x \in \mathbb{R} \mid x^2 + 1 = 0\}$. □

Die Menge X_3 enthält kein Element. Die Menge ohne Elemente nennt man auch die **leere Menge** und bezeichnet sie mit $\emptyset$. Die Menge X_2 ist eine Teilmenge von $\mathbb{N}$. Sind dabei X und Y beliebige Mengen, so heißt Y eine **Teilmenge** von X, in Zeichen $Y \subset X$ oder $X \supset Y$, wenn jedes Element von Y auch Element von X ist:

$$x \in Y \Rightarrow x \in X.$$



2.2.5 Beispiel

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$ □

Falls $Y \subset X$ und $X \neq Y$, so schreiben wir $Y \subsetneq X$ oder $X \supsetneq Y$ und nennen Y eine **echte Teilmenge** von X. Zwei Mengen X und Y sind genau dann gleich, wenn gilt $X \subset Y$ und $Y \subset X$:

$$X = Y \Leftrightarrow (X \subset Y) \wedge (Y \subset X).$$

2.3 Mengentheoretische Operationen

Wir definieren einige wichtige Operationen für Mengen.

2.3.1 Definition

Sind A, B Mengen, so definieren wir

(i) den **Durchschnitt** von A und B durch

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

(ii) die **Vereinigung** von A und B durch

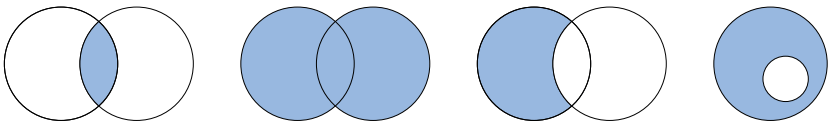
$$A \cup B := \{x \mid x \in A \text{ oder } x \in B\},$$

(iii) die **Differenz** $A \setminus B$ durch

$$A \setminus B := \{x \mid x \in A \text{ und } x \notin B\}.$$

Gilt $A \cap B = \emptyset$, so nennen wir A und B **disjunkt**. □

Wenden wir diese Operationen auf Teilmengen einer Menge X an, so erhalten wir wieder Teilmengen von X. Ist A eine Teilmenge von X, so schreiben wir auch $\complement_X A := X \setminus A$ und nennen $\complement_X A$ das **Komplement** von A in X.



$A \cap B$ $A \cup B$ $A \setminus B$ $\complement_X A$

2.3.2 Satz

Sind A, B, C, X Mengen, so gilt:

(i) Assoziativgesetze:

$$(A \cup B) \cup C = A \cup (B \cup C),$$
$$(A \cap B) \cap C = A \cap (B \cap C).$$

(ii) Kommutativgesetze:

$$A \cup B = B \cup A,$$
$$A \cap B = B \cap A.$$

(iii) Distributivgesetze:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

(iv) Identitätsgesetze:

$$A \cup \emptyset = A,$$
$$A \subset X \Rightarrow A \cap X = A.$$

(v) Komplementgesetze:

$$A \subset X \Rightarrow A \cup \complement_X A = X,$$
$$A \subset X \Rightarrow A \cap \complement_X A = \emptyset.$$

Beweis

Wir zeigen jeweils, dass jedes Element der einen Menge in der anderen Menge liegt und umgekehrt. Zum Beispiel:

$$\begin{aligned} x \in A \cap (B \cup C) &\Leftrightarrow x \in A \text{ und } x \in B \cup C \\ &\Leftrightarrow x \in A \text{ und } (x \in B \text{ oder } x \in C) \\ &\overset{1.3.3,\,(iii)}{\Leftrightarrow} (x \in A \text{ und } x \in B) \text{ oder } (x \in A \text{ und } x \in C) \\ &\Leftrightarrow (x \in A \cap B) \text{ oder } (x \in A \cap C) \\ &\Leftrightarrow x \in (A \cap B) \cup (A \cap C)\,. \end{aligned}$$

□

2.3.3 Warnung

Bevor man einen Implikationspfeil in einem Beweis umkehrt, überprüfe man sorgfältig, ob die Implikation in die andere Richtung auch gilt. □

Man beachte die formale Analogie von Satz 2.3.2 und Satz 1.3.3.

2.3.4 Satz

Sind A, B Teilmengen einer Menge X, so gilt:

(i) De Morgansche Gesetze:

$$\begin{aligned} \complement_X (A \cup B) &= \complement_X A \cap \complement_X B, \\ \complement_X (A \cap B) &= \complement_X A \cup \complement_X B. \end{aligned}$$

(ii) Doppelte Komplementbildung:

$$\complement_X \left(\complement_X A\right) = A.$$

Beweis

Wie im Beweis von Satz 2.3.2. □

2.4 Potenzmengen

2.4.1 Definition

Ist X eine Menge, so heißt die Menge $\mathcal{P}(X)$ aller Teilmengen von X die Potenzmenge von X. □

2.4.2 Beispiel

(i) $\mathcal{P}(\emptyset) = \{\emptyset\}$

(ii) $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$

(iii) $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$

(iv) $\mathcal{P}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$. □

2.4.3 Satz

Sei X eine n-elementige Menge, wobei $n \in \mathbb{N}_0$. Dann gilt $|\mathcal{P}(X)| = 2^n$.

Beweis

Induktion nach n:

- **Induktionsanfang**: n = 0

$$\mathcal{P}(\emptyset) = \{\emptyset\} \Rightarrow |\mathcal{P}(\emptyset)| = 1 = 2^0.$$

- **Induktionsschluss**: $n \Rightarrow n+1$
  - **Induktionsvoraussetzung**: Sei $n \in \mathbb{N}_0$ eine Zahl, sodass jede n-elementige Menge genau 2^n Teilmengen hat.
  - **Induktionsbehauptung**: Dann hat jede $(n+1)$-elementige Menge genau 2^{n+1} Teilmengen.
- Wir zeigen die Behauptung:
  Sei $X = \{x_1, \ldots, x_{n+1}\}$ eine $(n+1)$-elementige Menge. Wir betrachten die Teilmenge $Y = \{x_1, \ldots, x_n\}$ von X. Dann gilt $|\mathcal{P}(Y)| = 2^n$ nach der Induktionsvoraussetzung. Ist $A \in \mathcal{P}(Y)$, so sind A und $A \cup \{x_{n+1}\}$ Elemente von $\mathcal{P}(X)$. Auf diese Weise erhält man genau $2 \cdot 2^n = 2^{n+1}$ paarweise verschiedene Elemente von $\mathcal{P}(X)$. Weitere Elemente von $\mathcal{P}(X)$ gibt es nicht. □

2.5 Binomialkoeffizienten

Wir besprechen eine wichtige Aussage der Kombinatorik.

2.5.1 Definition

Ist $n \in \mathbb{N}$, so setzen wir

$$n! = \prod_{i=1}^{n} i$$

(„n **Fakultät**“). Wir schreiben auch $0! := 1$. □

2.5.2 Definition

Sind $n, k \in \mathbb{N}_0$, so setzen wir

$$\binom{n}{k} := \prod_{i=1}^{k} \frac{n-i+1}{i} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k}$$

(„**Binomialkoeffizient** n **über** k“). □

Unmittelbar aus der Definition folgt:

$$\binom{n}{k} = 0 \quad \text{für } k > n \text{ und}$$

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{n-k} \quad \text{für } 0 \leqslant k \leqslant n.$$

2.5.3 Lemma

Für $1 \leqslant k \leqslant n$ gilt:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Beweis

Gilt $k = n$, so ist die Formel sicher richtig. Gilt $1 \leqslant k < n$, so hat man

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

□

2.5.4 Satz

Seien $n, k \in \mathbb{N}_0$. Dann ist die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge gleich $\binom{n}{k}$.

Beweis

Induktion nach n unter Benutzung von 2.5.3. □

Der Satz zeigt insbesondere, dass die Zahlen $\binom{n}{k}$ natürliche Zahlen sind (oder 0). Nach Lemma 2.5.3 erhält man die Binomialkoeffizienten rekursiv über das **Pascalsche Dreieck**:

					1					
				1		1				
			1		2		1			
		1		3		3		1		
	1		4		6		4		1	
1		5		10		10		5		1

2.5.5 Beispiel

Es gilt:

$$\binom{49}{6} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 13\,983\,816.$$

□

2.5.6 Binomischer Lehrsatz

Sind $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$, so gilt:

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

Beweis

Aufgaben. □

2.6 Das kartesische Produkt

2.6.1 Definition

Ist $n \in \mathbb{N}$ und sind $X_1, \ldots, X_n$ nichtleere Mengen, so heißt die Menge

$$\prod_{i=1}^{n} X_i := X_1 \times \ldots \times X_n := \{(x_1, \ldots, x_n) \mid x_i \in X_i \text{ für } 1 \leqslant i \leqslant n\}$$

das **kartesische Produkt** der X_i. □

Wir nennen die Elemente $(x_1, \ldots, x_n)$ von $\prod_{i=1}^{n} X_i$ auch **geordnete n-Tupel** und betonen damit, dass eine Reihenfolge festgelegt ist.

2.6.2 Beispiel

$$\{1,2\} \times \{a,b,c\} = \{(1,a), (1,b), (1,c), (2,a), (2,b), (2,c)\}. \qquad \square$$

Ist X eine Menge, so schreiben wir abkürzend

$$X^n := \underbrace{X \times \ldots \times X}_{n-\text{mal}}.$$

Kapitel 3

Abbildungen

3.1 Motivation

Abbildungen zwischen Mengen spielen genauso wie Mengen selbst eine zentrale Rolle innerhalb der Mathematik. Sie erlauben es insbesondere, mathematische Strukturen auf möglicherweise verschiedenen Mengen miteinander zu vergleichen. □

3.2 Grundlegende Definitionen

3.2.1 Definition

Seien X und Y Mengen. Eine Vorschrift, die jedem Element $x \in X$ genau ein Element $y = f(x) \in Y$ zuordnet, heißt **Abbildung** von X nach Y. Man schreibt dafür auch

$$f : X \to Y \text{ oder } X \xrightarrow{f} Y \text{ oder } f : X \to Y,\ x \mapsto f(x).$$

Dabei heißen X **Definitionsmenge** und Y **Wertebereich** oder **Bildbereich** von $f : X \to Y$. □

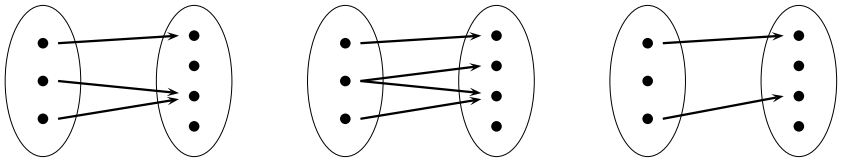


Abbildung keine Abbildung keine Abbildung

3.2.2 Bemerkung und Definition

(i) Ist X eine Menge, so ist

$$\mathrm{id}_X : X \to X,\ x \mapsto x,$$

eine Abbildung, die **identische Abbildung** von X.

(ii) Sind $f : X \to Y$ und $g : Y \to Z$ zwei Abbildungen, so erhält man eine neue Abbildung durch die Vorschrift

$$g \circ f : X \to Z,\ x \mapsto g(f(x)).$$

Diese Abbildung heißt die **Komposition** oder **Hintereinanderausführung** von f und g .

(iii) Sind $X_1, \ldots, X_n$ Mengen, so ist

$$\pi_i : X_1 \times \ldots \times X_n \to X_i,\ (x_1, \ldots, x_n) \mapsto x_i$$

eine Abbildung, die **i-te kanonische Projektion**. □

3.2.3 Definition

Ist $f : X \to Y$ eine Abbildung zwischen zwei Mengen, so heißt

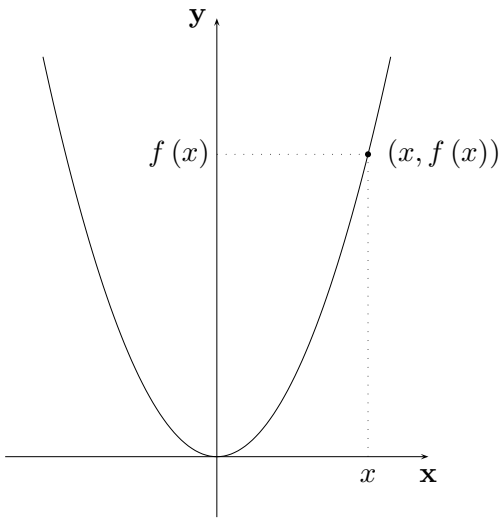
$$\Gamma_f := \{(x, y) \in X \times Y \mid f(x) = y\}$$

der **Graph** von f. □

Den Graph benutzt man oft zur Veranschaulichung von Abbildungen.

3.2.4 Beispiel

Wir skizzieren den Graphen der Abbildung $\mathbb{R} \to \mathbb{R},\ x \mapsto x^2$:



□

3.2.5 Bemerkung und Definition

Sei $f : X \to Y$ eine Abbildung

(i) Ist $A \subset X$, so heißt

$$f(A) := \{y \in Y \mid \exists x \in A \text{ mit } f(x) = y\} \subset Y$$

das **Bild** von A unter f. Insbesondere heißt

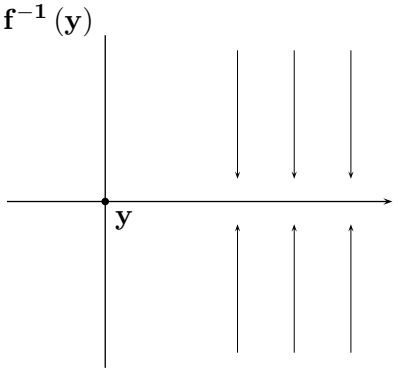
$$\text{Bild } f := f(X)$$

das **Bild** von f.

(ii) Ist $B \subset Y$, so heißt

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

das **Urbild** von B unter f. Hat insbesondere B nur ein Element y, so heißt $f^{-1}(\{y\})$ die **Faser** von f über y.



(iii) Ist $A \subset X$, so ist

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

eine Abbildung, die **Einschränkung** von f auf A. □

Ist $f : X \to Y$ eine Abbildung, so sind die Fasern von f aufgrund der Definition einer Abbildung paarweise disjunkt: Sind $y, y' \in Y$ mit $y \neq y'$, so gilt $f^{-1}(\{y\}) \cap f^{-1}(\{y'\}) = \emptyset$.

3.2.6 Definition

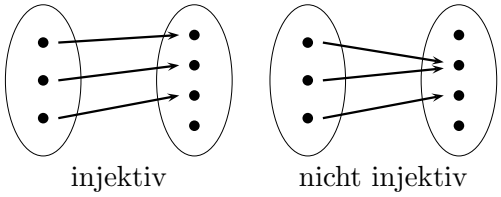
Eine Abbildung $f : X \to Y$ heißt

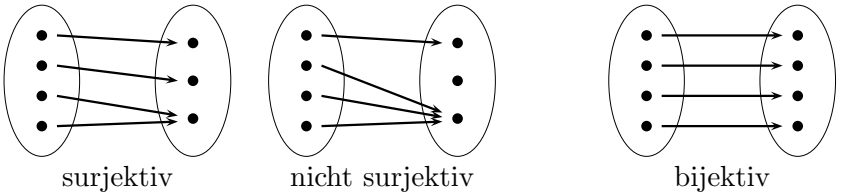
(i) **injektiv**, falls $\forall x, x' \in X$ gilt: $f(x) = f(x') \Rightarrow x = x'$;

(ii) **surjektiv**, falls $\forall y \in Y\ \exists x \in X$ mit $f(x) = y$;

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

Eine bijektive Abbildung nennt man auch eine **Bijektion**. □





3.3 Mächtigkeit von Mengen

3.3.1 Satz

Ist $f : X \to Y$ eine Abbildung zwischen endlichen Mengen mit $|X| = |Y|$, so sind äquivalent:

(i) f ist injektiv,

(ii) f ist surjektiv,

(iii) f ist bijektiv.

Beweis

Per Definition gilt (i) $\wedge$ (ii) $\Leftrightarrow$ (iii) . Es genügt also, (i) $\Leftrightarrow$ (ii) zu zeigen.

(i) $\Rightarrow$ (ii): Sei f injektiv.

$$\begin{aligned}
&\Rightarrow && \left|f^{-1}(\{y\})\right| \leqslant 1 \qquad \forall y \in Y \\
&\Rightarrow && |X| = \left|f^{-1}(Y)\right| = \sum_{y\in Y} \left|f^{-1}(\{y\})\right| \leqslant \sum_{y\in Y} 1 = |Y| \\
&\overset{|X|=|Y|}{\Rightarrow} && \left|f^{-1}(\{y\})\right| = 1 \qquad \forall y \in Y \\
&\Rightarrow && f \text{ ist surjektiv.}
\end{aligned}$$

(ii) $\Rightarrow$ (i): Sei f surjektiv.

$$\begin{aligned}
&\Rightarrow && \left|f^{-1}(\{y\})\right| \geqslant 1 \qquad \forall y \in Y \\
&\Rightarrow && |Y| = \sum_{y\in Y} 1 \leqslant \sum_{y\in Y} \left|f^{-1}(\{y\})\right| = |X| \\
&\overset{|X|=|Y|}{\Rightarrow} && \left|f^{-1}(\{y\})\right| = 1 \qquad \forall y \in Y \\
&\Rightarrow && f \text{ ist injektiv.}
\end{aligned}$$

□

3.3.2 Korollar

Ist $f : X \to Y$ eine Abbildung zwischen endlichen Mengen, so gilt:

(i) f injektiv $\Rightarrow |X| \leqslant |Y|$,

(ii) f surjektiv $\Rightarrow |X| \geqslant |Y|$.

Beweis

Dies haben wir im Beweis von 3.3.1 gezeigt. □

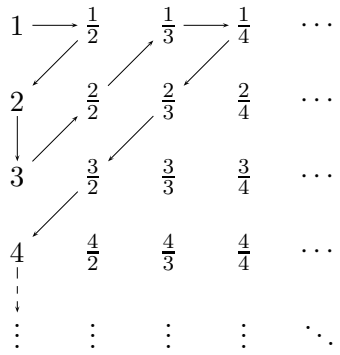
Wir nennen zwei Mengen **gleichmächtig**, wenn es eine bijektive Abbildung zwischen ihnen gibt. Wie wir gerade gesehen haben, sind zwei *endliche* Mengen X und Y genau dann gleichmächtig, wenn gilt $|X| = |Y|$. Wir nennen eine Menge X **abzählbar**, wenn es eine surjektive Abbildung $\mathbb{N} \to X$ gibt.

3.3.3 Beispiel

(i) Jede endliche Menge $X = \{x_1, \ldots, x_n\}$ ist abzählbar: Durch die Vorschrift $k \mapsto x_k$ für $k \leq n$ bzw. $k \mapsto x_n$ für $k > n$ erhält man eine surjektive Abbildung $\mathbb{N} \to X$.

(ii) Ist k eine beliebige aber feste ganze Zahl, so ist $X = \{n \in \mathbb{Z} \mid n \geq k\}$ abzählbar: Die Abbildung $\mathbb{N} \to X,\ n \mapsto n + k - 1$ ist surjektiv.

(iii) Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar. Wir illustrieren das **Cantorsche Diagonalverfahren** zur Abzählung der strikt positiven rationalen Zahlen:



Lässt man doppelte Elemente weg, so ergibt sich die Anordnung

1	$\frac{1}{2}$	2	3	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{3}{2}$	4	$\ldots$
1	2	3	4	5	6	7	8	9	$\ldots$

und somit eine Bijektion $\mathbb{N} \to \mathbb{Q}^+ := \{x \in \mathbb{Q} \mid x > 0\}$. Wandelt man das Verfahren leicht ab, so erhält man eine Bijektion $\mathbb{N} \to \mathbb{Q}$. □

Sind X, Y Mengen, sodass es eine injektive aber keine bijektive Abbildung $X \to Y$ gibt, so sagen wir, Y hat eine **höhere Mächtigkeit** als X. Mit einem weiteren Diagonalverfahren von Cantor lässt sich etwa zeigen, dass $\mathbb{R}$ eine höhere Mächtigkeit als $\mathbb{N}$ hat. Wir sagen auch, $\mathbb{R}$ ist **überabzählbar**.

3.4 Charakterisierung von injektiv, surjektiv und bijektiv

3.4.1 Satz

Ist $f : X \to Y$ eine Abbildung, so gilt:

(i) f injektiv $\Leftrightarrow \exists$ Abbildung $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$.

(ii) f surjektiv $\Leftrightarrow \exists$ Abbildung $g : Y \to X$ mit $f \circ g = \mathrm{id}_Y$.

(iii) f bijektiv $\Leftrightarrow \exists$ Abbildung $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$.

Beweis

(i) "⇒" Sei f injektiv. Dann gibt es zu jedem $y \in f(X)$ *genau* ein $x \in X$ mit $f(x) = y$ und wir setzen $g(y) := x$. Weiter wählen wir ein beliebiges $x_0 \in X$ und setzen $g(y) = x_0 \ \forall y \in Y \setminus f(X)$. So erhalten wir eine Abbildung $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$.

"⇐" Ist umgekehrt $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$ gegeben und gilt $f(x) = f(x')$ für $x, x' \in X$, so folgt $x = \mathrm{id}_X(x) = g(f(x)) = g(f(x')) = \mathrm{id}_X(x') = x'$. Also ist f injektiv.

(ii) "⇒" Sei f surjektiv. Dann können wir zu jedem $y \in Y$ ein festes $x \in X$ mit $f(x) = y$ wählen und setzen $g(y) := x$. So erhalten wir eine Abbildung $g : Y \to X$ mit $f \circ g = \mathrm{id}_Y$.

"⇐" Ist umgekehrt $g : Y \to X$ mit $f \circ g = \mathrm{id}_Y$ gegeben und $y \in Y$, so ist $y = f(g(y))$, also y Bild von $g(y)$ unter f. Also ist f surjektiv.

(iii) "⇒" Sei f bijektiv. Wir definieren g wie folgt: Sei $y \in Y$. Da f surjektiv ist, $\exists x \in X$ mit $f(x) = y$ und dieses x ist durch y und f *eindeutig bestimmt*, da f auch injektiv ist. Wir setzen $g(y) := x$ und erhalten dadurch eine Abbildung $g : Y \to X$ mit den gewünschten Eigenschaften.

"⇐" ergibt sich aus (i) und (ii) . □

3.4.2 Bemerkung und Definition

Ist $f : X \to Y$ bijektiv, so heißt die in Teil (iii), "⇒" des Beweises definierte Abbildung $f^{-1} := g$ auch **Umkehrabbildung** von f. Sie ist durch die Eigenschaften

$$g \circ f = \mathrm{id}_X \text{ und } f \circ g = \mathrm{id}_Y$$

eindeutig bestimmt. In der Tat, ist $\widetilde{g} : Y \to X$ eine weitere solche Abbildung, so gilt $\forall y \in Y$:

$$f(g(y)) = y = f(\widetilde{g}(y));$$

da f injektiv ist, folgt $g(y) = \widetilde{g}(y)$. □

Wie in obiger Bemerkung zeigt man die Gleichheit zweier Abbildungen oft dadurch, dass man für jedes Element des Definitionsbereichs nachweist, dass die entsprechenden Werte übereinstimmen.

3.4.3 Beispiel

Die Abbildung

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2,$$

ist

- nicht injektiv, da zum Beispiel $f(1) = 1 = f(-1)$,
- nicht surjektiv, da zum Beispiel -1 kein Urbild hat.

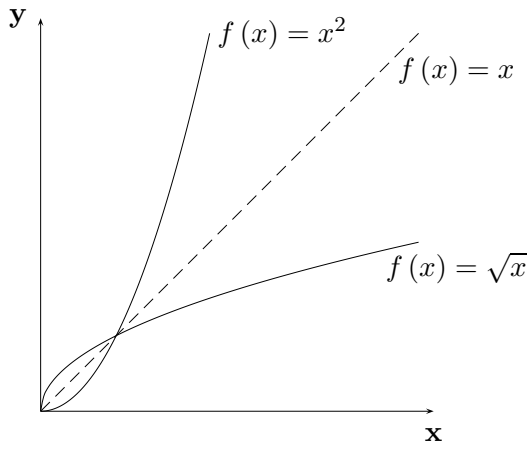
Setzen wir

$$\mathbb{R}_0^+ := \{x \in \mathbb{R} \mid x \geqslant 0\},$$

so ist

- $\mathbb{R} \to \mathbb{R}_0^+$, $x \mapsto x^2$, surjektiv,
- $\mathbb{R}_0^+ \to \mathbb{R}$, $x \mapsto x^2$, injektiv und
- $\mathbb{R}_0^+ \to \mathbb{R}_0^+$, $x \mapsto x^2$, bijektiv mit Umkehrabbildung

$$\mathbb{R}_0^+ \to \mathbb{R}_0^+,\ y \mapsto \sqrt{y}.$$



□

3.4.4 Bemerkung

Sind $f : X \to Y$ und $g : Y \to Z$ bijektive Abbildungen, so ist auch $g \circ f : X \to Z$ bijektiv und es gilt:

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

□

3.5 Familien von Mengen

Jedes n-Tupel $(x_1, \ldots, x_n)$ von Elementen einer Menge X können wir als Abbildung

$$\{1, \ldots, n\} \to X,\ i \mapsto x_i,$$

auffassen. Allgemeiner definieren wir:

3.5.1 Definition

Ist

$$\varphi : I \to X$$

eine Abbildung zwischen zwei Mengen, so schreiben wir auch

$$(x_i)_{i \in I} \text{ mit } x_i = \varphi(i)$$

für φ und fassen φ als eine **Familie** (oder ein I-**Tupel**) von Elementen in X mit Indexmenge I auf. Ist speziell $I = \mathbb{N}$, so sprechen wir von einer **Folge** $(x_n)_{n \in \mathbb{N}}$ in X. Sind allgemeiner k eine beliebige aber feste ganze Zahl und $I = \{n \in \mathbb{Z} \mid n \geq k\}$, so sprechen wir ebenfalls von einer **Folge** in X und schreiben dafür $(x_n)_{n \geq k}$. Die x_n heißen auch **Glieder** der Folge. □

3.5.2 Bemerkung und Definition

Betrachten wir in 3.5.1 Abbildungen von einer Indexmenge in eine Menge von Mengen, so ergibt sich der Begriff der **Familie von Mengen**. Ist $(X_i)_{i\in I}$ eine solche, so erhalten wir analog zu 2.3.1 die **Vereinigung** der $X_i,\ i \in I$, durch

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

und den **Durchschnitt** der $X_i,\ i \in I$, durch

$$\bigcap_{i\in I} X_i := \{x \mid x \in X_i\ \forall i \in I\}\,.$$

Ist insbesondere $I = \{1, \ldots, n\}$, so schreiben wir

$$\bigcup_{i=1}^{n} X_i := \bigcup_{i\in I} X_i \quad \text{sowie} \quad \bigcap_{i=1}^{n} X_i := \bigcap_{i\in I} X_i.$$

□

Kapitel 4

Relationen

4.1 Motivation

Relationen beschreiben Beziehungen zwischen Objekten. Beispiele sind Adressdatenbanken, die die Beziehungen zwischen Vornamen, Nachnamen, Postleitzahl, Straße, Hausnummer und Wohnort erfassen. Ein weiteres Beispiel sind Äquivalenzrelationen, die es ermöglichen, Objekte, die bestimmte Eigenschaften teilen, als gleich anzusehen. Beispiele wie die lexikographische Ordnung der Wörter einer Sprache oder die Ordnung von Objekten hinsichtlich ihrer Größe unterstreichen die Bedeutung von Ordnungsrelationen. □

4.2 Grundlegende Definitionen

4.2.1 Definition

Eine **Relation** zwischen zwei nichtleeren Mengen X und Y ist eine Teilmenge $R \subset X \times Y$. Wir schreiben auch xRy für $(x, y) \in R$ und sagen „x steht in Relation R zu y“. □

4.2.2 Beispiel

(i) Ist X die Menge der Studierenden einer Universität und ist Y die Menge der Studiengänge, so ist

$$R := \{(x, y) \in X \times Y \mid x \text{ belegt den Studiengang} y\}$$

eine Relation zwischen X und Y.

(ii) Der Graph einer Abbildung $f : X \to Y$ ist eine Relation zwischen X und Y. □

Gilt $X = Y$ in 4.2.1, so sprechen wir auch von einer **Relation auf** X.

4.2.3 Definition

Eine Relation R auf X heißt

- **reflexiv**, falls xRx $\forall x \in X$,
- **symmetrisch**, falls $xRy \Rightarrow yRx$ $\forall x, y \in X$,
- **antisymmetrisch**, falls $xRy \wedge yRx \Rightarrow x = y$ $\forall x, y \in X$,
- **transitiv**, falls $xRy \wedge yRz \Rightarrow xRz$ $\forall x, y, z \in X$. □

4.3 Äquivalenzrelationen

4.3.1 Definition

Eine Relation R auf einer Menge X heißt **Äquivalenzrelation**, wenn sie reflexiv, symmetrisch und transitiv ist. In diesem Fall schreibt man oft $\sim$ für R und somit $x \sim y$ für xRy. □

4.3.2 Beispiel

Sei X eine Menge von nichtleeren Mengen. Dann ist durch

$$A \sim B \Leftrightarrow A \text{ und } B \text{ sind gleichmächtig}$$

eine Äquivalenzrelation auf X erklärt. In der Tat ist $\sim$

- reflexiv, da $\mathrm{id}_A : A \to A$ bijektiv,
- symmetrisch, da mit $f : A \to B$ auch $f^{-1} : B \to A$ bijektiv,
- transitiv, da mit $f : A \to B$ und $g : B \to C$ auch $g \circ f : A \to C$ bijektiv ist nach 3.4.4. □

4.3.3 Beispiel

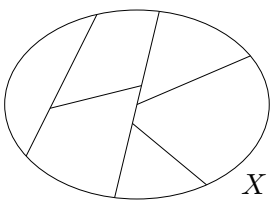
Sind endlich viele Aussagevariablen gegeben und ist X die Menge der logischen Formeln in diesen Variablen, so ist durch logische Äquivalenz eine Äquivalenzrelation auf X erklärt. □

4.3.4 Definition

Ist X eine Menge und $(X_i)_{i \in I}$ eine Familie von Teilmengen von X mit

- $X_i \cap X_j = \emptyset$ für $i \neq j$,
- $\bigcup_{i \in I} X_i = X$,

so heißt $(X_i)_{i \in I}$ eine **Partition** von X.



□

4.3.5 Definition

Sei $\sim$ eine Äquivalenzrelation auf einer Menge X. Für $x \in X$ heißt

$$[x] := [x]_\sim := \{y \in X \mid y \sim x\}$$

die **Äquivalenzklasse** von x. Jedes Element $y \in [x]$ nennen wir einen **Repräsentanten** der Äquivalenzklasse $[x]$. □

4.3.6 Satz

Ist $\sim$ eine Äquivalenzrelation auf einer Menge X, so bilden die Äquivalenzklassen bezüglich $\sim$ eine Partition von X.

Beweis

(i) Seien $x, y \in X$ mit $[x] \cap [y] \neq \emptyset$.
Dann $\exists z \in [x] \cap [y]$, das heißt es gilt $z \sim x$ und $z \sim y$. Wegen der Symmetrie und der Transitivität folgt $x \sim y$. Daraus ergibt sich $[x] = [y]$: Ist nämlich $u \in [x]$, so ist $u \sim x$ und wegen $x \sim y$ auch $u \sim y$, das heißt $u \in [y]$. Es folgt $[x] \subset [y]$. Analog ergibt sich $[y] \subset [x]$.

(ii) Sei $x \in X$. Dann ist $x \in [x]$ wegen $x \sim x$. Also ist X die Vereinigung der Äquivalenzklassen. □

4.3.7 Bemerkung und Definition

Ist $\sim$ eine Äquivalenzrelation auf der Menge X, so schreiben wir auch $X/\!\sim$ für die Menge aller Äquivalenzklassen. Die Vorschrift

$$\pi : X \to X/\!\sim,\ x \mapsto [x]\,,$$

definiert eine surjektive Abbildung, die **kanonische Projektion** von X auf $X/\!\sim$. □

4.3.8 Beispiel

Sei $n \in \mathbb{N}$. Dann heißen zwei Zahlen $a, b \in \mathbb{Z}$ **kongruent modulo** n, in Zeichen

$$a \equiv b \mod n\,,$$

wenn $a - b$ durch n teilbar ist. Es gilt etwa $17 \equiv 5 \bmod 3$ oder $38 \equiv 20 \bmod 6$.
Durch kongruent mod n ist eine Äquivalenzrelation auf $\mathbb{Z}$ definiert:

- Reflexivität: $a - a = 0$ ist durch n teilbar,
- Symmetrie: $a - b = t \cdot n \Rightarrow b - a = (-t) \cdot n$,
- Transitivität: $a - b = t_1 \cdot n$ `und` $b - c = t_2 \cdot n \Rightarrow a - c = (t_1 + t_2) \cdot n$.

Wir schreiben $\overline{a} := [a]$ für die Äquivalenzklasse von $a \bmod n$ sowie $\mathbb{Z}_n$ für die Menge der Äquivalenzklassen. Es gibt genau n Äquivalenzklassen, denn es gilt

$$\mathbb{Z}_n = \left\{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\right\} .$$

Ist etwa $n = 3$, so gibt es genau die Äquivalenzklassen

$$\begin{aligned}\overline{0} &= \{\ldots, -6, -3, 0, 3, 6, \ldots\},\\ \overline{1} &= \{\ldots, -5, -2, 1, 4, 7, \ldots\},\\ \overline{2} &= \{\ldots, -4, -1, 2, 5, 8, \ldots\}.\end{aligned}$$

□

Wir erwähnen, dass man beim Ablesen einer Uhr modulo 12 rechnet.

4.3.9 Beispiel

Konstruiert man die rationalen Zahlen aus den ganzen Zahlen, so unterscheidet man zum Beispiel nicht zwischen $\frac{2}{3}$ und $\frac{4}{6}$. Formal korrekt geht man dabei wie folgt vor. Man überlegt sich, dass durch

$$(p_1, q_1) \sim (p_2, q_2) \Leftrightarrow p_1 \cdot q_2 = p_2 \cdot q_1$$

eine Äquivalenzrelation auf $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definiert ist. Dann schreibt man $\frac{p}{q}$ für die Äquivalenzklasse von $(p, q) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ sowie

$$\mathbb{Q} := \left\{ \frac{p}{q} \;\middle|\; (p, q) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \right\}.$$

Man fasst $\mathbb{Z}$ als Teilmenge von $\mathbb{Q}$ auf vermöge der Abbildung

$$\mathbb{Z} \to \mathbb{Q},\ p \mapsto \frac{p}{1}.$$

Die Addition und Multiplikation in $\mathbb{Q}$ führt man auf die Addition und Multiplikation in $\mathbb{Z}$ zurück, indem man die üblichen Gesetze des Bruchrechnens fordert. Darauf gehen wir hier nicht näher ein. □

4.4 Ordnungsrelationen

4.4.1 Definition

Eine Relation auf einer Menge X heißt **partielle Ordnung**, wenn sie reflexiv, antisymmetrisch und transitiv ist. In diesem Fall schreibt man oft $\leqslant$ für R und somit $x \leqslant y$ für xRy. Falls zusätzlich $\forall x, y \in X$ gilt $x \leqslant y$ oder $y \leqslant x$ („je zwei Elemente von X sind bezüglich $\leqslant$ vergleichbar“), so sprechen wir von einer **totalen Ordnung**. □

Wir schreiben $x < y$ falls $x \leqslant y$ und $x \neq y$. Weiter verwenden wir die Schreibweisen $y \geqslant x$ für $x \leqslant y$ sowie $y > x$ falls $x < y$.

4.4.2 Beispiel

Sind k eine beliebige aber feste ganze Zahl und $X = \{n \in \mathbb{Z} \mid n \geq k\}$, so ist durch

$$m \leq n \Leftrightarrow \exists \ell \in \mathbb{N}_0 \text{ mit } m + \ell = n$$

eine totale Ordnung auf X definiert. □

4.4.3 Wohlordnungssatz

Seien k eine beliebige aber feste ganze Zahl, $X = \{n \in \mathbb{Z} \mid n \geq k\}$ und $A \subset X$ eine nichtleere Teilmenge. Dann besitzt A ein **kleinstes Element**:

$$\exists a_0 \in A \text{ mit } (a \in A \Rightarrow a \geqslant a_0)\,.$$

Insbesondere besitzt jede nichtleere Teilmenge von $\mathbb{N}$ ein kleinstes Element.

Beweis

- **Annahme**: Die Behauptung ist falsch.

Dann $\exists$ nichtleere Teilmenge $A \subset X$ ohne ein kleinstes Element. Wir setzen

$$B := \{n \in X \mid \{k, \ldots, n\} \cap A = \emptyset\}\,.$$

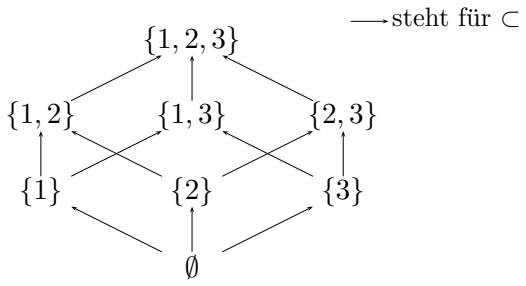
Durch Induktion nach n, $n \geq k$, zeigt man $B = X$. Dies ist ein Widerspruch zu $A \neq \emptyset$. □

4.4.4 Beispiel

Ist X eine Menge, so ist durch $\subset$ eine partielle Ordnung auf der Potenzmenge $\mathcal{P}(X)$ erklärt. In der Tat gilt: $\forall A, B, C \in \mathcal{P}(X)$:

- Reflexivität: $A \subset A$ $\qquad \forall A \in \mathcal{P}(X)\,,$
- Antisymmetrie: $A \subset B$ und $B \subset A \Rightarrow A = B$ $\qquad \forall A, B \in \mathcal{P}(X)\,,$
- Transitivität: $A \subset B$ und $B \subset C \Rightarrow A \subset C$ $\qquad \forall A, B, C \in \mathcal{P}(X)\,.$

Enthält X mindestens 2 Elemente. so handelt sich nicht um eine totale Ordnung. Dies veranschaulichen wir am Beispiel $X = \{1, 2, 3\}$:



Zum Beispiel sind $\{2\}$ und $\{1, 3\}$ nicht vergleichbar. □

Kapitel 5

Gruppen, Ringe, Körper

5.1 Motivation

Die Addition in $\mathbb{Z}$ und $\mathbb{Q}$ sowie die Multiplikation in $\mathbb{Q} \setminus \{0\}$ sind Rechenoperationen, die weitgehend übereinstimmenden Rechengesetzen unterliegen. So gilt zum Beispiel

$$(a+b)+c = a+(b+c) \text{ bzw. } (a \cdot b) \cdot c = a \cdot (b \cdot c)\,.$$

Weiter gibt es ausgezeichnete Zahlen, nämlich 0 bzw. 1, die sich bei diesen Operationen neutral verhalten:

$$0+a = a = a+0 \text{ bzw. } 1 \cdot a = a = a \cdot 1.$$

Schließlich gibt es zu jedem a ein a', nämlich $a' = -a$ bzw. $a' = \frac{1}{a}$, sodass die Summe bzw. das Produkt von a und a' gerade die jeweils neutrale Zahl ergibt:

$$(-a)+a = 0 = a+(-a) \text{ bzw. } \frac{1}{a} \cdot a = 1 = a \cdot \frac{1}{a}.$$

Da diese Rechenregeln das Zahlenrechnen weitgehend beherrschen und auch in vielen anderen Fällen auftreten, ist es naheliegend, sie unabhängig von der speziellen Natur der Rechengrößen und der jeweiligen Operationen zu untersuchen. Das hat insbesondere den Vorteil, dass mathematische Aussagen, die auf den grundlegenden Rechenregeln beruhen, nicht in jeder Situation wieder neu bewiesen werden müssen.

Bei dieser abstrakten Betrachtungsweise stehen also die Rechengesetze im Vordergrund: Nicht womit man rechnet ist wesentlich, sondern wie man rechnet. Man setzt lediglich voraus, dass für die Elemente einer gegebenen Menge eine Operation definiert ist, die jedem geordneten Paar (a, b) von Elementen der Menge wieder ein Element der Menge zuordnet, sodass die oben genannten Rechenregeln gelten. Oft betrachtet man auch mehrere "miteinander verträgliche" Operationen gleichzeitig (wie etwa $+$, $\cdot$). □

5.2 Gruppen

5.2.1 Definition

Eine **Verknüpfung** $\circ$ auf einer Menge X ist eine Abbildung

$$X \times X \to X, (x, y) \mapsto x \circ y.$$

Ist $A \subset X$ eine Teilmenge und gilt

$$a, b \in A \Rightarrow a \circ b \in A,$$

so heißt A bezüglich $\circ$ **abgeschlossen**. Ist dies der Fall, so ist die Einschränkung von $\circ$ auf $A \times A$ eine Verknüpfung auf A, die **induzierte Verknüpfung**. □

5.2.2 Bemerkung und Definition

Sei $\circ$ eine Verknüpfung auf der Menge X.

(i) $\circ$ heißt **assoziativ**, wenn gilt

$$(a \circ b) \circ c = a \circ (b \circ c) \quad \forall a, b, c \in X.$$

In diesem Fall nennen wir das Paar $(X, \circ)$ auch eine **Halbgruppe**.

(ii) Ein Element $e \in X$ heißt **neutrales Element** bezüglich $\circ$, wenn gilt

$$e \circ a = a = a \circ e \quad \forall a \in X.$$

Ein neutrales Element ist im Falle der Existenz eindeutig bestimmt: Sind e, f zwei neutrale Elemente, so gilt:

$$e \underset{f \text{ neutral}}{=} e \circ f \underset{e \text{ neutral}}{=} f.$$

Jede Halbgruppe, die ein neutrales Element besitzt, heißt ein **Monoid**.

(iii) Ist $(X, \circ)$ ein Monoid mit neutralem Element e, so heißt $a \in X$ **invertierbar** in X, wenn es ein zu a **inverses Element** in X gibt, das heißt wenn $\exists a' \in X$ mit

$$a' \circ a = e = a \circ a'.$$

Ein zu a inverses Element ist im Falle der Existenz eindeutig bestimmt und wird mit a^{-1} bezeichnet: Sind a', a'' zu a invers, so gilt

$$a' = e \circ a' = \big(a'' \circ a\big) \circ a' = a'' \circ \big(a \circ a'\big) = a'' \circ e = a''.$$

(iv) Sind $(X, \circ)$ ein Monoid mit neutralem Element e und $a \in X$ ein invertierbares Element, so gelten für alle $x, y \in X$ die folgenden **Kürzungsregeln**:

$$a \circ x_1 = a \circ x_2 \Rightarrow x_1 = x_2 \;\text{ und }\; y_1 \circ a = y_2 \circ a \Rightarrow y_1 = y_2.$$

Die erste Regel ergibt sich durch Verknüpfung mit a^{-1} von links:

$$\begin{aligned} a \circ x_1 = a \circ x_2 &\Rightarrow a^{-1} \circ (a \circ x_1) = a^{-1} \circ (a \circ x_2) \\ &\Rightarrow (a^{-1} \circ a) \circ x_1 = (a^{-1} \circ a) \circ x_2 \\ &\Rightarrow e \circ x_1 = e \circ x_2 \\ &\Rightarrow x_1 = x_2. \end{aligned}$$

Die zweite Regel folgt analog durch Verknüpfung mit a^{-1} von rechts.

(v) $\circ$ heißt **kommutativ**, wenn gilt:

$$a \circ b = b \circ a \quad \forall a, b \in X.$$ □

Steht $\circ$ für eine additiv geschriebene Verknüpfung $+$, so wird das neutrale Element mit 0 („**Nullelement**“) sowie das zu a inverse Element mit $-a$ bezeichnet (falls existent). Ist $\cdot$ eine multiplikativ geschriebene Verknüpfung, so schreibt man analog 1 für das neutrale Element („**Einselement**“) sowie $\frac{1}{a}$ für das zu a inverse Element (falls existent).

5.2.3 Definition

Eine **Gruppe** ist ein Monoid, in dem jedes Element invertierbar ist. Mit anderen Worten: Ein Paar $(G, \circ)$ bestehend aus einer Menge G und einer Verknüpfung $\circ$ auf G heißt **Gruppe**, falls gilt:

(i) $\circ$ ist assoziativ,

(ii) $\exists$ neutrales Element bezüglich $\circ$ in G,

(iii) jedes Element in G besitzt ein inverses Element in G.

Ist zusätzlich $\circ$ kommutativ, so heißt $(G, \circ)$ eine **kommutative** oder **abelsche Gruppe**. □

5.2.4 Satz

Ist $(G, \circ)$ eine Gruppe, so sind für alle $a, b \in G$ die Gleichungen

$$a \circ x = b \text{ und } y \circ a = b$$

in G lösbar.

Beweis

Das Element $x := a^{-1} \circ b \in G$ ist eine Lösung von $a \circ x = b$: Ist e das neutrale Element, so gilt

$$a \circ x = a \circ \left(a^{-1} \circ b\right) = \left(a \circ a^{-1}\right) \circ b = e \circ b = b.$$

Analog behandelt man die zweite Gleichung. □

5.2.5 Bemerkung

Man kann zeigen, dass die Gruppen gerade diejenigen Halbgruppen sind, in denen alle Gleichungen wie in 5.2.4 lösbar sind. Ist $(G, \circ)$ eine Gruppe, so sind die Gleichungen sogar eindeutig in G lösbar: Sind etwa $x_1, x_2 \in G$ zwei Lösungen von $a \circ x = b$, das heißt gilt

$$a \circ x_1 = b = a \circ x_2,$$

so gilt $x_1 = x_2$ nach der ersten Kürzungsregel in 5.2.2, (iv). Analog für die zweite Gleichung. □

5.2.6 Beispiel

(i) $(\mathbb{N}, +)$ ist eine kommutative Halbgruppe, aber kein Monoid.
$(\mathbb{N}_0, +)$ ist ein kommutatives Monoid, aber keine Gruppe.

(ii) $(\mathbb{Z}, +), (\mathbb{Q}, +)$ bzw. $(\mathbb{Q} \setminus \{0\}, \cdot)$ sind kommutative Gruppen. □

Hier ist ein weiteres Beispiel:

5.2.7 Bemerkung und Definition

(i) Ist X eine Menge, so heißt jedes Element von

$$S(X) := \{f : X \to X \mid f \text{ ist bijektiv}\}$$

eine **Permutation** von X. Nach Bemerkung 3.4.4 ist die Komposition von zwei Permutationen wieder eine solche. Also ist $\circ$ eine Verknüpfung auf $S(X)$. Diese ist assoziativ:

$$\begin{aligned}((f \circ g) \circ h)(x) &= (f \circ g)(h(x)) = f(g(h(x))) = f((g \circ h)(x)) \\ &= (f \circ (g \circ h))(x) \qquad \forall f, g, h \in S(X),\ \forall x \in X.\end{aligned}$$

Weiter ist id_X ein neutrales Element:

$$\begin{aligned}(\mathrm{id}_X \circ f)(x) &= \mathrm{id}_X(f(x)) = f(x) = f(\mathrm{id}_X(x)) \\ &= (f \circ \mathrm{id}_X)(x) \qquad \forall f \in S(X),\ \forall x \in X.\end{aligned}$$

Ist schließlich $f \in S(X)$, so ist die Umkehrabbildung f^{-1} ein zu f inverses Element:

$$f^{-1} \circ f = \mathrm{id}_X = f \circ f^{-1}.$$

Also ist $(S(X), \circ)$ eine Gruppe, die **Permutationsgruppe** von X.

(ii) Im Falle der endlichen Menge $\{1, \ldots, n\} \subset \mathbb{N}$ schreibt man auch

$$S_n := S(\{1, \ldots, n\}).$$

Für $\sigma \in S_n$ schreiben wir auch

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma_1 & \sigma_2 & \ldots & \sigma_n \end{pmatrix} \text{ mit } \sigma_i = \sigma(i).$$

(iii) $\tau \in S_n$ heißt **Transposition**, wenn $\exists k, l \in \{1, \ldots, n\}, k \neq l$, mit

$$\tau(i) = \begin{cases} l & \text{falls } \ i = k, \\ k & \text{falls } \ i = l, \\ i & \text{sonst,} \end{cases} \qquad \text{für } \ 1 \leqslant i \leqslant n.$$

Wir sprechen auch von der **Vertauschung** von k und l. Für jedes solche τ gilt $\tau \circ \tau = \mathrm{id}_{\{1,\ldots,n\}}$, das heißt $\tau^{-1} = \tau$.

(iv) Man beachte, dass S_n genau $n!$ Elemente hat (Aufgaben). So besteht etwa S_3 aus den 6 Elementen

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Es gilt zum Beispiel

$$\sigma_2 \circ \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \sigma_5.$$

(v) Wie man leicht nachrechnet, ist S_n kommutativ für $n = 1, 2$. Für $n = 3$ gilt dies nicht mehr: Sind τ_{12} die Vertauschung von 1 und 2 sowie τ_{23} die Vertauschung von 2 und 3, so gilt:

$$\tau_{23} \circ \tau_{12} = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 3 & 1 & 2 & 4 & \dots & n \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 3 & 1 & 4 & \dots & n \end{pmatrix} = \tau_{12} \circ \tau_{23}.$$

□

Ist $(H, \circ)$ eine Halbgruppe, so ergibt sich durch Induktion ein Assoziativgesetz für mehr als drei Elemente von H. Das bedeutet, dass

$$a_1 \circ \ldots \circ a_n := (\ldots((a_1 \circ a_2) \circ a_3) \circ \ldots \circ a_{n-1}) \circ a_n$$

für $a_1, \ldots, a_n \in H$ unabhängig von der Klammerung definiert ist, sodass wir die Klammern auch einfach weglassen können. Ist $(H, \circ)$ kommutativ, so gilt entsprechend ein allgemeines Kommutativgesetz, das besagt, dass wir Elemente beim Verknüpfen beliebig vertauschen dürfen: Sind $a_1, \ldots, a_n \in H$ und ist $\sigma \in S_n$ eine Permutation von $\{1, \ldots, n\}$, so gilt

$$a_1 \circ \ldots \circ a_n = a_{\sigma(1)} \circ \ldots \circ a_{\sigma(n)}.$$

5.2.8 Definition

Seien $(H, \circ)$ eine Halbgruppe mit neutralem Element e, $a \in H$ und $n \in \mathbb{Z}$. Dann setzen wir

$$a^n = \underbrace{a \circ \ldots \circ a}_{n-\text{mal}}, \qquad \text{falls } n > 0,$$
$$a^0 = e,$$

sowie, falls a invertierbar,

$$a^n = \left(a^{-1}\right)^{-n}, \qquad \text{falls } n < 0.$$

Wir nennen a^n die n-**te Potenz** von a. □

Steht $\circ$ für eine additiv geschriebene Verknüpfung $+$, so schreiben wir na an Stelle von a^n.

5.2.9 Satz

Sei $(G, \circ)$ eine Gruppe mit neutralem Element e. Dann gelten die folgenden Rechenregeln:

(i) $e^{-1} = e$,

(ii) $(a^{-1})^{-1} = a \quad \forall a \in G$,

(iii) $(a_1 \circ \cdots \circ a_n)^{-1} = a_n^{-1} \circ \cdots \circ a_1^{-1} \quad \forall a_1, \ldots, a_n \in G$,

(iv) $a^m \circ a^n = a^{m+n}$, $(a^m)^n = a^{m \cdot n} \quad \forall m, n \in \mathbb{Z} \ \forall a \in G$.

Beweis

Siehe Vorlesung Algebraische Strukturen. □

Streng genommen haben wir bei der Definition von $a_1 \circ \ldots \circ a_n$ und damit auch bei der von a^n ein mathematisches Prinzip benutzt, das **Rekursionsprinzip** heißt, und das man durch vollständige Induktion begründen kann. Grob gesprochen erlaubt uns dieses Prinzip die Definition (Konstruktion) einer Folge $(a_n)_{n\in\mathbb{N}}$ in einer Menge X durch Angabe von a_1 zusammen mit einer Vorschrift wie man a_{n+1} aus a_n gewinnt (oder wie man a_{n+1} aus $a_1, \ldots, a_n$ gewinnt). Man spricht dann von einer **rekursiven Definition (Konstruktion)** oder auch von einer **induktiven Definition (Konstruktion)**. Will man etwa die Potenz a^n in 5.2.8 für $n > 0$ rekursiv definieren, so setzt man $a^1 := a$ sowie $a^{n+1} := a^n \circ a$. Für eine Diskussion des Rekursionsprinzips verweisen wir auf Barner-Flohr, Analysis 1.

5.3 Ringe und Körper

5.3.1 Definition

Sei R eine Menge mit zwei Verknüpfungen $+, \cdot$. Dann heißt das Tripel $(R, +, \cdot)$ ein **Ring**, falls

(i) $(R, +)$ eine abelsche Gruppe ist,

(ii) $(R, \cdot\,)$ eine Halbgruppe ist und

(iii) die Distributivgesetze gelten:

$$\begin{aligned} &a \cdot (b + c) = a \cdot b + a \cdot c, \\ &(b + c) \cdot a = b \cdot a + c \cdot a \qquad \forall a, b, c \in R\ . \end{aligned}$$

Ein Ring heißt

- **kommutativ**, falls $(R, \cdot\,)$ kommutativ ist,
- **Ring mit 1**, falls $(R, \cdot\,)$ ein (mit 1 bezeichnetes) neutrales Element besitzt. □

5.3.2 Rechenregeln

Sei $(R, +, \cdot\,)$ ein Ring. Dann gilt

(i) $a \cdot 0 = 0 \cdot a = 0 \quad \forall a \in R$,

(ii) $a \cdot (-b) = (-a) \cdot b = -a \cdot b$ und $(-a) \cdot (-b) = a \cdot b \qquad \forall a, b \in R$,

(iii) ist $(R, +, \cdot\,)$ ein Ring mit 1 und gilt $R \neq \{0\}$, so gilt $1 \neq 0$.

Beweis

Siehe Vorlesung Algebraische Strukturen. □

5.3.3 Definition

Ein kommutativer Ring $(K, +, \cdot)$ mit 1 heißt ein **Körper**, wenn $K \setminus \{0\}$ bezüglich $\cdot$ eine abelsche Gruppe ist. □

In einem Körper gilt stets $1 \neq 0$. Jeder Körper hat also mindestens zwei Elemente.

5.3.4 Beispiel

(i) $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit 1.

(ii) Seien $X \neq \emptyset$ und $R = \{f : X \to \mathbb{Z} \mid f \text{ Abbildung}\}$. Definieren wir für $f, g \in R$ die Summe $f + g$ und das Produkt $f \cdot g$ elementweise,

$$(f + g)(x) = f(x) + g(x), \quad (f \cdot g)(x) = f(x) \cdot g(x) \quad \forall x \in X ,$$

so ist $(R, +, \cdot)$ ein kommutativer Ring mit 1 (das Einselement ist die Abbildung, die jedes $x \in X$ auf die Zahl $1 \in \mathbb{Z}$ abbildet).

(iii) $(\mathbb{Q}, +, \cdot)$ ist ein Körper. □

5.3.5 Bemerkung

(i) Ist $(R, +, \cdot)$ ein Ring, so erhalten wir induktiv das **allgemeine Distributivgesetz**: Sind $a_1, \ldots, a_n, b_1, \ldots, b_m \in R$, so gilt

$$\left(\sum_{i=1}^{n} a_i\right) \cdot \left(\sum_{j=1}^{m} b_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \cdot b_j .$$

(ii) Betrachet man in der Mathematik Strukturen auf Mengen, so stellt sich immer auch die Frage, wie diese Strukturen mit mengentheoretischen Operationen und Abbildungen verträglich sind. Darauf kommen wir in dem Linearen-Algebra-Teil der Vorlesung zurück, verweisen aber vor allem auf die Vorlesung Algebraische Strukturen. □

Zum Schluss gehen wir kurz auf ein weiteres wichtiges Beispiel für einen Ring ein:

5.3.6 Satz und Definition

Seien $(R, +, \cdot)$ ein kommutativer Ring mit 1 und t eine Unbestimmte. Dann heißt ein Ausdruck der Form

$$p = \sum_{i=0}^{n} a_i t^i = a_0 + a_1 t + \cdots + a_n t^n \quad \text{mit } a_0, a_1, \ldots, a_n \in R$$

ein **Polynom** in t mit **Koeffizienten** $a_i \in R$. Ist dabei $a_n \neq 0$, so heißt $\deg(p) := n$ der **Grad** von p. Sind alle $a_i = 0$, so sprechen wir vom **Nullpolynom** und schreiben $p = 0$. Der Vollständigkeit halber setzen wir $\deg(0) = -\infty$. Die Polynome vom Grad 0 zusammen mit dem Nullpolynom werden als **konstante Polynome** bezeichnet. Mit anderen Worten, wir fassen die Elemente von R als konstante Polynome auf. Für die Menge aller Polynome in t mit Koeffizienten in R schreiben wir $R[t]$.

Zwei Polynome

$$p = a_0 + a_1 t + \cdots + a_n t^n, \;\; q = b_0 + b_1 t + \cdots + b_m t^m \in R\,[t]$$

können miteinander addiert und multipliziert werden. Zur Definition der Addition können wir $m = n$ annehmen (ist etwa $m < n$, so setzen wir $b_i = 0$ für $i = m+1, \ldots, n$). Dann definieren wir

$$p + q := \sum_{i=0}^{n} (a_i + b_i) t^i.$$

Die Multiplikation wird so definiert, dass man formal entsprechend dem allgemeinen Distributivgesetz ausmultipliziert und dann bezüglich den Potenzen von t zusammenfasst:

$$p \cdot q = (\sum_{i=0}^{n} a_i t^i) \cdot (\sum_{j=0}^{m} b_j t^j) := \sum_{\ell=0}^{n+m} \left(\sum_{k=0}^{\ell} (a_{\ell-k} \cdot b_k) \right) t^\ell.$$

Dann ist $(R\,[t]\,, +, \cdot)$ ein kommutativer Ring mit 1.

Beweis

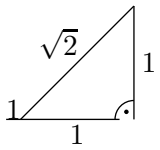
Die Ringgesetze von $(R\,[x]\,, +, \cdot)$ ergeben sich aus denen von $(R, +, \cdot)$ durch Nachrechnen. Dabei ist das Nullpoynom neutral bezüglich $+$ und zu $\sum_{i=0} a_i t^i$ ist $\sum_{i=0}^{n} (-a_i)\, t^i$ invers bezüglich $+$. Das konstante Polynom 1 ist das Einselement. □

Kapitel 6

Axiomatik der reellen Zahlen

6.1 Motivation

Die reellen Zahlen sind die Grundlage allen Messens. Wie wir in den Aufgaben gesehen haben, kommen wir dazu mit den rationalen Zahlen nicht aus: Man kann Strecken konstruieren, deren Länge keine rationale Zahl ist:



Wir haben uns bisher auf den Standpunkt gestellt, dass wir mit den Zahlen und ihren Rechengesetzen vertraut sind, und sind nur gelegentlich auf den Aufbau des Zahlsystems eingegangen (Erwähnung der Peano-Axiome zur Charakterisierung der natürlichen Zahlen in Abschnitt 1.5.1; Konstruktion der ganzen aus den natürlichen Zahlen in den Aufgaben; Konstruktion der rationalen aus den ganzen Zahlen in Beispiel 4.3.9). Nicht eingehen wollen wir auf die Konstruktion der reellen aus den rationalen Zahlen. Um auf sicherem Boden zu stehen, werden wir stattdessen als Grundlage für alles Weitere einige Axiome (Grundgesetze) formulieren, die die reellen Zahlen charakterisieren und aus denen sich die weiteren Eigenschaften und Gesetze dieser Zahlen ableiten lassen. Dabei bemühen wir uns darum, mit möglichst wenig Axiomen auszukommen. Das Prinzip, an den Anfang einer Theorie Axiome zu stellen, und alles weitere der Theorie daraus abzuleiten, nennt man auch die **axiomatische Methode**. Diese ist ein Grundprinzip der modernen Mathematik.

Mehr zum Aufbau des Zahlsystems finden Sie in dem Buch *Zahlen* von Ebbinghaus et al. □

6.2 Die Körperaxiome

Die hier vorgestellten Grundgesetze begründen das Rechnen mit den reellen Zahlen. Wir gehen davon aus, dass es Verknüpfungen

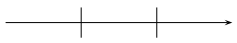
$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ (a, b) \mapsto a + b,\ \text{ und}$$

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ (a, b) \mapsto a \cdot b,$$

gibt, so dass $(\mathbb{R}, +, \cdot)$ ein Körper ist. Die in Kapitel 5 formulierten Gesetze für einen Körper nennen wir in diesem Zusammenhang dann die **Körperaxiome** der reellen Zahlen. Alle Gesetze, die wir in Kapitel 5 aus den Körpergesetzen abgeleitet haben, gelten auch für $\mathbb{R}$. **Subtraktion** bzw. **Division** sind erklärt durch $a - b := a + (-b)$ bzw. $\frac{a}{b} := a \cdot \frac{1}{b} = a \cdot b^{-1}$.

6.3 Die Anordnungsaxiome

Wir sind daran gewöhnt, dass wir zwei reelle Zahlen entsprechend ihrer Größe vergleichen können:



Axiomatisch fordern wir die Existenz einer Relation $\leqslant$ auf $\mathbb{R}$, die folgenden Grundgesetzen genügt:

6.3.1 Anordnungsaxiome

(i) $\leqslant$ ist eine totale Ordnung auf $\mathbb{R}$,

(ii) **Verträglichkeit mit der Addition:** $a \leqslant b \Rightarrow a + c \leqslant b + c \quad \forall a, b, c \in \mathbb{R}$,

(iii) **Verträglichkeit mit der Multiplikation:** $a \leqslant b,\ 0 \leqslant c \Rightarrow a \cdot c \leqslant b \cdot c \quad \forall a, b, c \in \mathbb{R}$. □

Eine Zahl $x \in \mathbb{R}$ heißt **positiv**, wenn gilt $x > 0$. Sie heißt **negativ**, wenn gilt $x < 0$.

Aus den Anordnungsaxiomen lassen sich weitere Gesetze ableiten:

6.3.2 Satz

Für alle $a, b, c, d \in \mathbb{R}$ gilt:

(i) $a \leqslant b \quad \Leftrightarrow\ 0 \leqslant b - a \Leftrightarrow -b \leqslant -a$,

(ii) $a \leqslant b,\ c \leqslant d \ \Rightarrow\ a + c \leqslant b + d$,

(iii) $a \leqslant b,\ c \leqslant 0 \ \Rightarrow\ a \cdot c \geqslant b \cdot c$,

(iv) $0 < a \leqslant b \quad \Rightarrow\ 0 < \frac{1}{b} \leqslant \frac{1}{a}$,

(v) $a \neq 0 \quad \Leftrightarrow\ 0 < a^2$.

Beweis

Wir zeigen (v). Zunächst ergibt sich aus den Körpergesetzen, dass gilt $a \neq 0 \Leftrightarrow a^2 \neq 0$. Ist dann also $a > 0$, so folgt aus 6.3.1, (iii) und den Körpergesetzen auch $0 = 0 \cdot a < a \cdot a = a^2$. Ist $a < 0$, so ist $0 < -a$ nach (i), also gilt $a^2 = (-a)^2 > 0$ nach dem gerade bewiesenen. □

Wir wollen nun erklären, wie wir bei unserem axiomatischen Ansatz die Menge $\mathbb{N}$ der natürlichen Zahlen als Teilmenge von $\mathbb{R}$ wiederfinden können. Dazu schreiben wir vorübergehend $0_{\mathbb{R}}$ bzw. $1_{\mathbb{R}}$ für das Null- bzw. Einselement von $\mathbb{R}$, um diese von den natürlichen Zahlen 0 und

1 zu unterscheiden. Aus 6.3.2, (v) folgt dann $0_{\mathbb{R}} < 1_{\mathbb{R}}$. Ist allgemein $n \in \mathbb{N}$, so schreiben wir vorübergehend

$$n_{\mathbb{R}} = \underbrace{1_{\mathbb{R}} + \ldots + 1_{\mathbb{R}}}_{n-\text{mal}} \in \mathbb{R}.$$

Induktion liefert dann $0_{\mathbb{R}} < n_{\mathbb{R}}$. Sind nun $m \neq n$ zwei natürliche Zahlen, so gilt auch $m_{\mathbb{R}} \neq n_{\mathbb{R}}$. Denn ist etwa $m < n$, so $\exists k \in \mathbb{N}$ mit $m + k = n$. Dann gilt auch $m_{\mathbb{R}} + k_{\mathbb{R}} = n_{\mathbb{R}}$. Wäre also $m_{\mathbb{R}} = n_{\mathbb{R}}$, so folgte aus der Kürzungsregel $k_{\mathbb{R}} = 0_{\mathbb{R}}$ im Widerspruch zu $0_{\mathbb{R}} < k_{\mathbb{R}}$.

Wir haben somit gezeigt, dass die Abbildung

$$\mathbb{N} \to \{n_{\mathbb{R}} \mid n \in \mathbb{N}\}\,,\ n \mapsto n_{\mathbb{R}},$$

eine Bijektion ist. Addition und Multiplikation bei den natürlichen Zahlen und den ihnen zugeordneten reellen Zahlen laufen auf dasselbe hinaus: Induktion liefert $(m+n)_{\mathbb{R}} = m_{\mathbb{R}} + n_{\mathbb{R}}$ und $(m \cdot n)_{\mathbb{R}} = m_{\mathbb{R}} \cdot n_{\mathbb{R}}$. Gleiches gilt für die Ordnung. Wir können somit die natürlichen Zahlen mit ihrem Bild in $\mathbb{R}$ identifizieren und n für $n_{\mathbb{R}}$ schreiben. Nehmen wir nun die Null und die negativen Zahlen hinzu und betrachten dann Brüche, so haben wir auch $\mathbb{Z}$ und $\mathbb{Q}$ wiedergefunden. Wir haben wie gewohnt

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

6.3.3 Definition

Für $a \in \mathbb{R}$ setzt man:

$$|a| := \begin{cases} a, & \text{falls } a \geqslant 0 \\ -a, & \text{falls } a < 0. \end{cases}$$

□

6.3.4 Satz (Eigenschaften des Absolutbetrags)

Für alle $a, b \in \mathbb{R}$ gilt:

(i) $|a| \geqslant 0 \quad$ und $\quad (|a| = 0 \Leftrightarrow a = 0)$,

(ii) $|-a| = |a|$,

(iii) $|a| \cdot |b| = |a \cdot b|$,

(iv) $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$, falls $b \neq 0$,

(v) $|a + b| \leqslant |a| + |b|$, (**Dreiecksungleichung**)

(vi) $||a| - |b|| \leqslant |a - b|$ und $||a| - |b|| \leqslant |a + b|$.

Beweis

Wir zeigen (vi). Mit Hilfe der Dreiecksungleichung ergibt sich

$$|a| = |a - b + b| \leq |a - b| + |b|$$

und somit $|a| - |b| \leq |a - b|$. Vertauscht man a und b, so erhält man analog $-(|a| - |b|) \leq |a - b|$. Aus der Definition des Betrages folgt $||a| - |b|| \leq |a - b|$ und somit die erste Ungleichung in (vi). Die zweite Ungleichung ergibt sich aus der ersten mit Hilfe von (ii):

$$||a| - |b|| = ||a| - |-b|| \leq |a - (-b)| = |a + b|$$

□

6.3.5 Notation (Intervalle)

Sind $a, b \in \mathbb{R}$ mit $a \leqslant b$, so setzen wir

- $[a,b] \quad := \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$. (**abgeschlossenes Intervall**)

Sind $a, b \in \mathbb{R}$ mit $a < b$, so setzen wir

- $]a,b[\quad := \{x \in \mathbb{R} \mid a < x < b\}$, (**offenes Intervall**)
- $[a,b[\quad := \{x \in \mathbb{R} \mid a \leqslant x < b\}$, (**halboffenes Intervall**)
- $]a,b] \quad := \{x \in \mathbb{R} \mid a < x \leqslant b\}$. (**halboffenes Intervall**)

Ist $a \in \mathbb{R}$, so setzen wir:

- $[a,\infty[\quad := \{x \in \mathbb{R} \mid x \geqslant a\}$,
- $]a,\infty[\quad := \{x \in \mathbb{R} \mid x > a\}$,
- $]-\infty,a] \quad := \{x \in \mathbb{R} \mid x \leqslant a\}$, (**uneigentliche Intervalle**)
- $]-\infty,a[\quad := \{x \in \mathbb{R} \mid x < a\}$,
- $]-\infty,\infty[:= \mathbb{R}$. □

Wir fordern ein weiteres Axiom:

6.3.6 Das Archimedische Axiom

Seien $a, b \in \mathbb{R}, a > 0, b > 0$. Dann $\exists n \in \mathbb{N}$ mit $n \cdot a > b$. □

Wir notieren einige Folgerungen:

6.3.7 Satz

(i) $\forall a \in \mathbb{R}\ \exists! n \in \mathbb{Z}$ mit $n \leqslant a < n+1$,

(ii) $\forall \varepsilon \in \mathbb{R},\ \varepsilon > 0,\ \exists n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$.

(iii) **Bernoullische Ungleichung**: Sei $a \in \mathbb{R},\ a \geqslant -1$. Dann gilt

$$(1+a)^n \geqslant 1 + n \cdot a \quad \forall n \in \mathbb{N}_0.$$

(iv) Sei $b \in \mathbb{R},\ b > 1$. Dann gilt $\forall K \in \mathbb{R},\ K > 0$: $\exists n \in \mathbb{N}$ mit $b^n > K$.

(v) Sei $q \in \mathbb{R},\ 0 < q < 1$. Dann gilt $\forall \varepsilon \in \mathbb{R},\ \varepsilon > 0$: $\exists n \in \mathbb{N}$ mit $q^n < \varepsilon$.

(vi) **Dichtheit der rationalen Zahlen**: Jedes offene Intervall $]a,b[$ enthält mindestens eine rationale Zahl.

Beweis

Die Gültigkeit von (i) - (v) zeigen wir in den Aufgaben.

(vi) Wegen $a < b$ existiert nach (ii) ein $n \in \mathbb{N}$ mit $\frac{1}{n} < b - a$. Dann können wir nach (i) ein $m \in \mathbb{Z}$ finden mit $m \leq na < m + 1$. Daraus ergibt sich mit Hilfe der Anordnungsaxiome:

$$a < \frac{m+1}{n} = \frac{m}{n} + \frac{1}{n} \leq a + \frac{1}{n} < a + (b - a) = b.$$

□

Wir fassen die bisherigen Axiome 6.2, 6.3.1 und 6.3.3 zusammen, indem wir sagen, dass $\mathbb{R}$ ein **archimedisch angeordneter Körper** ist. Neben $\mathbb{R}$ ist auch $\mathbb{Q}$ ein solcher. Also ist $\mathbb{R}$ durch die bisherigen Axiome noch nicht eindeutig bestimmt. Wir benötigen noch zusätzlich das im nächsten Kapitel formulierte **Vollständigkeitsaxiom**.

6.3.8 Notation

Wenn wir im folgenden Ungleichungen wie $\varepsilon > 0$ bzw. $M \geq 0$ verwenden, so ist stets ε bzw. M eine reelle Zahl.

Kapitel 7

Folgen, Grenzwerte und das Vollständigkeitsaxiom

7.1 Motivation

Der Begriff des Grenzwertes ist zentral für die Analysis. Seine Bedeutung beruht darauf, dass viele Größen nicht durch einen in endlich vielen Schritten exakt berechenbaren Ausdruck gegeben, sondern nur mit beliebiger Genauigkeit approximiert werden können. Eine Zahl mit beliebiger Genauigkeit zu approximieren heißt, sie als Grenzwert einer Folge darzustellen. □

7.2 Konvergente Folgen

Folgen in $\mathbb{R}$ heißen auch **Folgen reeller Zahlen**.

7.2.1 Beispiel:

(i) Ist $a \in \mathbb{R}$, so ist durch $a_n := a \ \forall n \in \mathbb{N}$ die konstante Folge $a, a, \ldots$ erklärt.

(ii) Setzen wir $a_n = \frac{1}{n} \ \forall n \in \mathbb{N}$, dann ergibt sich $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$.

(iii) Mit $a_n = (-1)^n \ \forall n \in \mathbb{N}$ erhalten wir $-1, 1, -1, 1, \ldots$.

(iv) Definieren wir rekursiv $a_1 = 1, a_2 = 1$ und $a_n = a_{n-1} + a_{n-2}$ für $n \geqslant 3$, so ergeben sich die **Fibonacci-Zahlen** $1, 1, 2, 3, 5, 8, 13, 21, \ldots$.

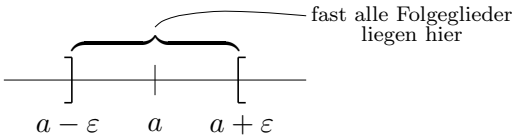
(v) Ist $a \in \mathbb{R}$, so ist durch $a_n := a^n \ \forall n \in \mathbb{N}$ eine Folge erklärt. □

Wir formulieren Definitionen und Sätze für Folgen der Form $(a_n)_{n\in\mathbb{N}}$. Für Folgen der Form $(a_n)_{n\geq k}$ (siehe 3.5.1) gelten die analogen Definitionen und Sätze.

7.2.2 Definition

Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt **konvergent gegen eine Zahl** $a \in \mathbb{R}$, falls gilt:

$$\forall \varepsilon > 0 \ \exists n_0(\varepsilon) \in \mathbb{N} \text{ mit } |a_n - a| < \varepsilon \quad \forall n \geqslant n_0(\varepsilon).$$



In diesem Fall heißt a **Grenzwert** oder **Limes** der Folge $(a_n)_{n\in\mathbb{N}}$ und wir schreiben

$$\lim a_n := \lim_{n\to\infty} a_n := a \quad \text{oder} \quad a_n \to a \ \text{ für } \ n \to \infty \ .$$

Ist speziell $a = 0$, so sprechen wir von einer **Nullfolge**. Eine Folge, die gegen einen Grenzwert konvergiert, heißt **konvergent**. Andernfalls heißt sie **divergent**. □

7.2.3 Satz

Jede konvergente Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen besitzt *genau einen* Grenzwert.

Beweis

- **Annahme:** $\exists$ zwei Grenzwerte $a \neq b$.

Dann ist $\varepsilon := \frac{|a-b|}{2} > 0$. Also

$$\exists n_a, n_b \in \mathbb{N} \ \text{ mit } \ |a_n - a| < \varepsilon \ \forall n \geqslant n_a \ \text{ und } \ |a_n - b| < \varepsilon \ \forall n \geqslant n_b.$$

Ist n_0 die größere der beiden Zahlen n_a und n_b, so gilt sowohl $|a_n - a| < \varepsilon$ als auch $|a_n - b| < \varepsilon \ \forall n \geqslant n_0$. Wählen wir ein $n \geqslant n_0$, so ergibt sich mit Hilfe der Dreiecksungleichung der folgende Widerspruch:

$$\begin{aligned} |a-b| &= |(a-a_n)+(a_n-b)| \leqslant |a_n-a| + |a_n-b| \\ &< 2\cdot\varepsilon = |a-b| \,. \quad \lightning \end{aligned}$$

□

7.2.4 Definition

Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt **beschränkt**, wenn

$$\exists M \geqslant 0 \ \text{ mit } \ |a_n| \leqslant M \ \forall n \in \mathbb{N}.$$ □

7.2.5 Satz

Jede konvergente Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen ist beschränkt.

Beweis

Sei $\lim a_n = a$. Dann $\exists n_0 \in \mathbb{N}$ mit

$$|a_n - a| < 1 \quad \forall n \geqslant n_0.$$

Es folgt:

$$|a_n| = |a + a_n - a| \leqslant |a| + |a_n - a| \leqslant |a| + 1 \ \forall n \geqslant n_0.$$

Ist also M die größte der Zahlen $|a_1|, |a_2|, \ldots, |a_{n_0-1}|, |a|+1$, so gilt $|a_n| \leqslant M \ \forall n \in \mathbb{N}$. □

7.2.6 Beispiel

(i) Ist $a_n = a \ \forall n \in \mathbb{N}$, so gilt $\lim a_n = a$. Denn ist $\varepsilon > 0$ beliebig vorgegeben, so gilt

$$|a_n - a| = 0 < \varepsilon \quad \forall n \geqslant 1.$$

(ii) Es gilt $\lim \frac{1}{n} = 0$. Denn ist $\varepsilon > 0$ beliebig vorgegeben, so $\exists$ nach 6.3.7, (ii) ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon$. Also folgt mit 6.3.2, (iv):

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leqslant \frac{1}{n_0} < \varepsilon \quad \forall n \geqslant n_0.$$

(iii) Die durch $a_n = (-1)^n \ \forall n \in \mathbb{N}$ definierte Folge ist divergent:

- **Annahme**: $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen $a \in \mathbb{R}$.

Dann $\exists n_0 \in \mathbb{N}$ mit $|a_n - a| < 1 \ \forall n \geqslant n_0$. Wählen wir ein $n \geqslant n_0$, so ergibt sich mit Hilfe der Dreiecksungleichung der folgende Widerspruch:

$$\begin{aligned} 2 = |a_{n+1} - a_n| &= |(a_{n+1} - a) + (a - a_n)| \\ &\leqslant |a_{n+1} - a| + |a - a_n| < 1 + 1 = 2. \ \lightning \end{aligned}$$

(iv) Die Folge der Fibonacci-Zahlen divergiert nach 7.2.5, denn man zeigt leicht durch Induktion, dass $a_n \geqslant n - 1 \ \forall n \in \mathbb{N}$. Also ist die Folge unbeschränkt.

(v) Ist $a_n = a^n \ \forall n \in \mathbb{N}$, so machen wir eine Fallunterscheidung. Gilt $|a| < 1$, so gilt $\lim a_n = 0$. Denn ist $\varepsilon > 0$ beliebig vorgegeben, so $\exists$ nach 6.3.7, (v) ein $n_0 \in \mathbb{N}$ mit $|a|^{n_0} < \varepsilon$. Also folgt mit dem dritten Anordnungsaxiom und 6.3.4 , (iii)

$$|a^n - 0| = |a^n| = |a|^n \leq |a|^{n_0} < \varepsilon \quad \forall n \geqslant n_0.$$

Ist $a = 1$, so gilt $\lim a_n = 1$ nach (i). Ist $a = -1$, so liegt Divergenz vor nach (iii). Ist schließlich $|a| > 1$, so ist die Folge wegen 6.3.7, (iv) unbeschränkt und somit divergent nach 7.2.5,. □

Wir sagen, eine Aussage für die Glieder einer Folge **gilt fast immer**, wenn es ein $N \in \mathbb{N}$ gibt, sodass die Aussage $\forall n \geqslant N$ gilt.

7.2.7 Satz (Rechenregeln für Grenzwerte)

Sind $(a_n)_{n\in\mathbb{N}}$ bzw. $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen reeller Zahlen mit Grenzwerten a bzw. b, und ist $(c_n)_{n\in\mathbb{N}}$ eine weitere Folge reeller Zahlen, so gilt:

(i) **Vergleichssatz**:
Gilt fast immer $a_n \leqslant b_n$, so ist auch $a \leqslant b$.

(ii) **Sandwichkriterium**:
Gilt $a = b$ und gilt fast immer $a_n \leqslant c_n \leqslant b_n$, so konvergiert auch $c_n \to a = b$.

(iii) **Betragssatz**:
Es konvergiert auch $|a_n| \to |a|$.

(iv) **Limessätze**:
Es konvergieren auch

$$\begin{aligned} a_n + b_n &\to a + b && \text{(\textbf{Summensatz})} \\ a_n - b_n &\to a - b \\ a_n \cdot b_n &\to a \cdot b && \text{(\textbf{Produktsatz})} \\ \lambda \cdot a_n &\to \lambda \cdot a \qquad \forall \lambda \in \mathbb{R} \end{aligned}$$

Ist überdies $b \neq 0$, so gilt fast immer $b_n \neq 0$ und es konvergiert auch

$$\left(\frac{a_n}{b_n}\right)_{n \geq k} \to \frac{a}{b} \qquad \text{(\textbf{Quotientensatz})}$$

(dabei sei k so gewählt dass gilt $b_n \neq 0 \; \forall n \geq k$).

Beweis

Wir zeigen den Summensatz. Sei $\varepsilon > 0$ beliebig. Dann $\exists n_a, n_b \in \mathbb{N}$ mit $|a_n - a| < \frac{\varepsilon}{2}$ für $n \geqslant n_a$ und $|b_n - b| < \frac{\varepsilon}{2}$ für $n \geqslant n_b$. Ist n_0 die größere der beiden Zahlen n_a, n_b, so folgt mit der Dreiecksungleichung

$$\begin{aligned} &|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \\ \leqslant\ &|a_n - a| + |b_n - b| \ < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \; \forall n \geqslant n_0. \end{aligned}$$

□

7.2.8 Beispiel

(i) Sei $a_n = \frac{4 \cdot n^3 - 6}{6 \cdot n^3 - 2 \cdot n^2} \; \forall n \in \mathbb{N}$. Dann gilt $a_n = \frac{4 - \frac{6}{n^3}}{6 - \frac{2}{n}}$. Beispiel 7.2.6 (i), (ii) und die Limessätze zeigen, dass die Zähler- bzw. die Nennerfolge konvergiert und zwar gegen 4 bzw. 6. Wegen $6 \neq 0$ können wir den Quotientensatz anwenden und erhalten, dass $(a_n)_{n \in \mathbb{N}}$ konvergiert mit Grenzwert $\frac{4}{6} = \frac{2}{3}$.

(ii) Sei $a_n = \frac{3^{n+1} + 2^n}{3^n + 1}$. Einsetzen großer n (Computer) liefert $a_n \approx 3$. Wir raten $a = 3$ und formen um:

$$0 \leq |a_n - 3| = \left|\frac{3^{n+1} + 2^n - 3^{n+1} - 3}{3^n + 1}\right| = \left|\frac{2^n - 3}{3^n + 1}\right| < \left|\frac{2^n}{3^n}\right| = \left|\frac{2}{3}\right|^n$$

Da die rechte Seite nach 7.2.6, (v) eine Nullfolge darstellt, folgt mit dem Sandwichkriterium $a_n \to 3$. □

7.3 Vollständigkeit

Das Vollständigkeitsaxiom eröffnet eine Möglichkeit, die Konvergenz einer Folge nachzuweisen, ohne dass man den Grenzwert kennt.

7.3.1 Definition

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen heißt **Cauchy-Folge**, falls gilt:

$$\forall \varepsilon > 0 \; \exists n_0(\varepsilon) \in \mathbb{N} \text{ mit } |a_n - a_m| < \varepsilon \quad \forall n, m \geqslant n_0(\varepsilon).$$

□

7.3.2 Satz

Jede konvergente Folge reeller Zahlen ist eine Cauchy-Folge.

Beweis

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen mit $a_n \to a$ und sei $\epsilon > 0$. Dann $\exists\, n_0 \in \mathbb{N}$, sodass $|a_n - a| < \frac{\epsilon}{2}\ \forall\ n \geq n_0$. Somit gilt $\forall\ n, m \geq n_0$:

$$\begin{aligned} &|a_n - a_m| = |(a_n - a) + (a - a_m)| \leq |a_n - a| + |a - a_m| \\ &< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon\ . \end{aligned}$$

□

Die Umkehrung des Satzes fordern wir als Axiom:

7.3.3 Vollständigkeitsaxiom

Jede Cauchy-Folge reeller Zahlen konvergiert gegen einen reellen Grenzwert. □

Ein Körper, in dem die Anordnungsaxiome, das Archimedische Axiom und das Vollständigkeitsaxiom erfüllt sind, heißt **vollständig archimedisch angeordneter Körper**. Man kann beweisen, dass ein solcher Körper eindeutig bestimmt ist (siehe Ebbinghaus et al., *Zahlen*).

7.4 Der Satz von Bolzano-Weierstraß

7.4.1 Definition

Seien $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen und

$$n_1 < n_2 < n_3 < ...$$

eine aufsteigende Folge natürlicher Zahlen. Dann heißt die Folge

$$(a_{n_k})_{k\in\mathbb{N}}$$

eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$. □

Es folgt unmittelbar aus den Definitionen, dass jede Teilfolge einer gegen $a \in \mathbb{R}$ konvergenten Folge reeller Zahlen ebenfalls gegen a konvergiert.

7.4.2 Satz (Bolzano-Weierstraß)

Jede beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen besitzt eine konvergente Teilfolge.

Notation

Ist $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall, so setzen wir diam $(I) := (b - a)$.

Beweis

(a) Wegen der Beschränktheit von $(a_n)_{n\in\mathbb{N}}$ existieren $A, B \in \mathbb{R}$ mit $A \leq a_n \leq B \quad \forall\, n \in \mathbb{N}$. Wir setzen $I_1 := [A, B]$ und konstruieren rekursiv eine Folge abgeschlossener Intervalle $I_k \subset \mathbb{R}$, $k \in \mathbb{N}$, so dass $\forall k \in \mathbb{N}$ gilt:

(i) I_k enthält unendlich viele Glieder von $(a_n)_{n\in\mathbb{N}}$.

(ii) $I_k \subset I_{k-1} \subset \cdots \subset I_1$.

(iii) diam $I_k = 2^{-k+1}$ diam $I_1 = (\frac{1}{2})^{k-1} \cdot (B - A)$.

Ist $k \geq 1$ und sind die Intervalle $I_\ell = [A_\ell, B_\ell]$ mit den Eigenschaften (i), (ii) und (iii) bereits konstruiert für $\ell \leq k$, so sei $M_k := \frac{A_k+B_k}{2}$ der Mittelpunkt des Intervalls $[A_k, B_k]$. Wir betrachten die beiden Teilintervalle $[A_k, M_k]$ und $[M_k, B_k]$. Wegen (i) für I_k muss eines dieser Teilintervalle unendlich viele Folgenglieder enthalten. Wir setzen $I_{k+1} := [A_k, M_k]$, falls dieses Intervall unendlich viele Folgenglieder enthält, und $I_{k+1} := [M_k, B_k]$ andererseits. Dann erfüllt I_{k+1} offensichtlich (i), (ii) und (iii).

(b) Ausgehend von $n_1 := 1$ definieren wir nun rekursiv eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ von $(a_n)_{n\in\mathbb{N}}$ mit $a_{n_k} \in I_k \;\forall\, k \in \mathbb{N}$: Sind $n_1 < \ldots < n_k$ bereits definiert, so $\exists\, m > n_k$ mit $a_m \in I_{k+1}$ da I_{k+1} unendlich viele Folgeglieder enthält. Wir setzen $n_{k+1} := m$.

(c) Wir beweisen mit Hilfe des Vollständigkeitsaxioms, dass $(a_{n_k})_{k\in\mathbb{N}}$ konvergiert: Ist $\epsilon > 0$ beliebig vorgegeben, so können wir nach (iii), 7.2.6, (v) und den Limesätzen ein $k_0 \in \mathbb{N}$ wählen, sodass diam $I_{k_0} < \epsilon$. Dann gilt $\forall\; k, \ell \geq k_0$: $a_{n_k}, a_{n_\ell} \in I_{k_0}$ und somit

$$|a_{n_k} - a_{n_\ell}| \leq \text{diam } I_{k_0} < \epsilon \ . \qquad \square$$

7.4.3 Definition

Eine Zahl $a \in \mathbb{R}$ heißt **Häufungspunkt** einer Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen, wenn gilt:

$$\forall\, \epsilon > 0\; \forall\, n_1 \in \mathbb{N}\; \exists n_2 \in \mathbb{N},\; n_2 \geq n_1,\; \text{mit } |a_{n_2} - a| < \epsilon \ . \qquad \square$$

7.4.4 Bemerkung

Man zeigt sofort, dass a genau dann Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ ist, wenn a Grenzwert einer konvergenten Teilfolge von $(a_n)_{n\in\mathbb{N}}$ ist. Der Satz von Bolzano-Weierstraß besagt also: Jede beschränkte Folge reeller Zahlen besitzt einen Häufungspunkt. $\square$

7.4.5 Beispiel

(i) Die durch $a_n = (-1)^n$ definierte Folge hat die Häufungspunkte ± 1, denn es gilt $\lim\limits_{k\to\infty} a_{2k} = 1$ und $\lim\limits_{k\to\infty} a_{2k-1} = -1$. Gleiches gilt für die durch $a_n = (-1)^n + \frac{1}{n}$ definierte Folge.

(ii) Die durch

$$a_n := \begin{cases} n & \text{falls } n \text{ gerade,} \\ \frac{1}{n} & \text{falls } n \text{ ungerade,} \end{cases}$$

definierte Folge ist unbeschränkt, hat aber den Häufungspunkt 0, denn $\lim\limits_{k\to\infty} a_{2k-1} = 0$. Es gibt also Folgen, die einen Häufungspunkt besitzen aber nicht beschränkt sind. $\square$

7.5 Das Monotoniekriterium

Auch hier kommen wir ohne Kenntnis des Grenzwertes aus.

7.5.1 Definition (monotone Folgen)

Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt

- **monoton wachsend**, falls $a_n \leqslant a_{n+1} \quad \forall n \in \mathbb{N}$,
- **streng monoton wachsend**, falls $a_n < a_{n+1} \quad \forall n \in \mathbb{N}$,
- **monoton fallend**, falls $a_n \geqslant a_{n+1} \quad \forall n \in \mathbb{N}$,
- **streng monoton fallend**, falls $a_n > a_{n+1} \quad \forall n \in \mathbb{N}$.

□

7.5.2 Satz (Monotoniekriterium)

Jede beschränkte monotone Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen konvergiert.

Beweis

Nach dem Satz von Bolzano-Weierstraß besitzt $(a_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$. Sei $a = \lim\limits_{k\to\infty} a_{n_k}$. Wir zeigen: Dann gilt auch $\lim\limits_{n\to\infty} a_n = a$.

Sei $\epsilon > 0$ beliebig vorgegeben. Dann $\exists\, k_0 \in \mathbb{N}$ mit $|a_{n_k} - a| < \epsilon \quad \forall\ k \geq k_0$. Wir setzen $n_0 = n_{k_0}$. Ist dann $n \geq n_0$ beliebig gegeben, so $\exists k \geq k_0$ mit $n_k \leq n < n_{k+1}$. Da $(a_n)_{n\in\mathbb{N}}$ monoton wachsend (bzw. fallend) ist, gilt dann auch $a_{n_k} \leq a_n \leq a_{n_{k+1}}$ (bzw. $a_{n_k} \geq a_n \geq a_{n_{k+1}}$). In jedem Fall gilt: Ist b die größere der beiden Zahlen $|a_{n_k} - a|$, $|a_{n_{k+1}} - a|$, so ist

$$|a_n - a| \leq b < \epsilon\ .$$

□

Wir behandeln eine wichtige Formel, die in dem nachfolgenden Beispiel zum Monotoniekriterium benötigt wird:

7.5.3 Bemerkung (Geometrische Summenformel)

Sind $x \neq 1$ eine reelle Zahl und $n \in \mathbb{N}_0$, so gilt

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x}\ .$$

Dies zeigt man sofort durch Induktion, wobei sich der Induktionsschritt $n \to n+1$ wie folgt ergibt:

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

□

7.5.4 Beispiel

Die Folge $s_n := \sum_{k=0}^{n} \frac{1}{k!}$ ist sicher monoton wachsend. Wegen

$$\frac{1}{k!} = \frac{1}{1 \cdot 2 \cdot 3 \cdot \ldots \cdot k} < \frac{1}{1 \cdot 2 \cdot 2 \cdot \ldots \cdot 2} = \frac{1}{2^{k-1}} \text{ für } k \geqslant 3$$

folgt mit der geometrischen Summenformel:

$$\begin{aligned} 0 \leqslant s_n &\leqslant 1 + \left(1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}}\right) \\ &= 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} < 1 + \frac{1}{\frac{1}{2}} = 3 \qquad \forall\, n \in \mathbb{N}_0. \end{aligned}$$

Die Folge ist also auch beschränkt. Nach 7.5.2 ist die Folge konvergent, der Grenzwert wird mit e bezeichnet (**Eulersche Zahl**). Im Tutorium werden wir zeigen, dass gilt

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n. \qquad \square$$

7.6 Bestimmte Divergenz

Bisher haben wir uns hauptsächlich um Eigenschaften konvergenter Folgen gekümmert. Jetzt gehen wir kurz auf Folgen ein, die ein besonders übersichtliches Divergenzverhalten zeigen. Für diese führen wir eine zweckdienliche Notation ein.

7.6.1 Definition

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen heißt **bestimmt divergent** (oder **uneigentlich konvergent**) gegen ∞, falls gilt:

$$\forall K \in \mathbb{R}\ \exists n_0 \in \mathbb{N} \ \text{ mit } \ a_n > K\ \forall n \geqslant n_0.$$

In diesem Fall schreiben wir $a_n \to \infty$ sowie $\lim a_n = \lim_{n \to \infty} a_n = \infty$.
Analog definiert man bestimmte Divergenz gegen $-\infty$. $\square$

Man beachte, dass ∞ bzw. $-\infty$ Symbole sind und keine Zahlen, mit denen man wie üblich rechnen kann.

7.6.2 Beispiel

(i) Ist $a_n = n\ \forall n \in \mathbb{N}$, so gilt $a_n \to \infty$. Gleiches gilt für die Folge der Fibonacci-Zahlen.

(ii) Ist $a_n = -2^n\ \forall n \in \mathbb{N}$, so gilt $a_n \to -\infty$.

(iii) Die durch $a_n = (-1)^n$ definierte Folge ist divergent aber nicht bestimmt divergent. $\square$

7.7 Schranken

Sei $X \subset \mathbb{R}$.

7.7.1 Definition

(i) X heißt **nach oben beschränkt**, wenn $\exists b \in \mathbb{R}$ mit $x \leqslant b \ \forall x \in X$. Jedes solche b heißt auch eine **obere Schranke** von X.

(ii) X heißt **nach unten beschränkt**, wenn $\exists a \in \mathbb{R}$ mit $x \geqslant a \ \forall x \in X$. Jedes solche a heißt auch eine **untere Schranke** von X.

(iii) X heißt **beschränkt**, wenn X nach oben und unten beschränkt ist. □

Für Folgen stimmt diese Definition mit der in 7.2.4 überein.

7.7.2 Definition

(i) Eine Zahl $s \in \mathbb{R}$ heißt **Supremum** von X, wenn s kleinste obere Schranke von X ist, das heißt wenn gilt

(a) $x \leqslant s \ \forall x \in X$,

(b) ist $s' \in \mathbb{R}$ und gilt $x \leqslant s' \ \forall x \in X$, so folgt $s \leqslant s'$.

(ii) Eine Zahl $r \in \mathbb{R}$ heißt **Infimum** von X, wenn r größte untere Schranke von X ist, das heißt wenn gilt:

(a) $x \geqslant r \ \forall x \in X$,

(b) ist $r' \in \mathbb{R}$ und gilt $x \geqslant r' \ \forall x \in X$, so folgt $r \geqslant r'$. □

7.7.3 Satz

(i) Ist $X \neq \emptyset$ und nach oben beschränkt, so besitzt X ein Supremum und es gibt eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ mit $x_n \to \sup X$.

(ii) Ist $X \neq \emptyset$ und nach unten beschränkt, so besitzt X ein Infimum und es gibt eine Folge $(y_n)_{n\in\mathbb{N}} \subset X$ mit $y_n \to \inf X$.

Beweis

Wir zeigen (i). Die Gültigkeit von (ii) ergibt sich analog.

Sei also $X \neq \emptyset$ und nach oben beschränkt. Dann existieren ein Element $x_0 \in X$ sowie eine obere Schranke b_0 von X. Es gilt $x_0 \leq b_0$ und somit $r := b_0 - x_0 \geq 0$. Ausgehend von x_0 und b_0 definieren wir rekursiv Folgen

- $x_0 \leq x_1 \leq x_2 \ldots$ von Elementen von X und
- $b_0 \geq b_1 \geq b_2 \ldots$ von oberen Schranken von X,

sodass gilt

$$b_n - x_n \leq 2^{-n} \cdot r \quad \forall n \in \mathbb{N}_0.$$

Ist $n \geq 1$ und sind $x_0, \ldots, x_n$ sowie $b_0, \ldots, b_n$ mit den genannten Eigenschaften bereits definiert, so sei $M_n = \frac{x_n + b_n}{2}$ der Mittelpunkt des Intervalls $[x_n, b_n]$. Dann gibt es genau zwei Möglichkeiten:

(i) M_n ist eine obere Schranke von X. Dann setzen wir $x_{n+1} = x_n$ und $b_{n+1} = M_n$.

(ii) M_n ist keine obere Schranke von X. Dann $\exists x_{n+1} \in X \cap]M_n, b_n]$ und wir setzen $b_{n+1} = b_n$.

In jedem Fall gilt $x_n \leq x_{n+1}$, $b_n \geq b_{n+1}$ und $b_{n+1} - x_{n+1} \leq 2^{-n-1} \cdot r$. Da die Folge $(b_n)_{n\in\mathbb{N}_0}$ monoton falled und nach unten durch x_0 beschränkt ist, konvergiert sie nach dem Monotoniekriterium gegen eine Zahl $b \in \mathbb{R}$. Wir zeigen, dass b kleinste obere Schranke von X ist:

(a) b ist obere Schranke von X: Ist $x \in X$, so gilt $x \leq b_n \ \forall n \in \mathbb{N}_0$, also auch $x \leq \lim b_n = b$ nach dem Vergleichssatz.

(b) b ist *kleinste* obere Schranke von X: Ist $b' < b$ eine weitere reelle Zahl, so $\exists n \in \mathbb{N}$ mit $2^{-n} \cdot r < b - b'$. Dann gilt

$$x_n \geq b_n - 2^{-n} \cdot r \geq b - 2^{-n} \cdot r > b'.$$

Also ist b' keine obere Schranke von X.

Schließlich gilt wegen $0 \leq b_n - x_n \leq 2^{-n} \cdot r \ \forall n \in \mathbb{N}_0$ und dem Sandwichkriterium, dass $(b_n - x_n)_{n\geq 0}$ eine Nullfolge ist. Mit $(b_n)_{n\geq 0}$ konvergiert dann auch $(x_n)_{n\geq 0}$ und zwar gegen den gleichen Grenzwert b. □

7.7.4 Bemerkung und Notation

(i) Supremum bzw. Infimum sind im Falle der Existenz eindeutig bestimmt und werden dann mit $\sup X$ bzw. $\inf X$ bezeichnet.

(ii) Ist X nach oben bzw. unten unbeschränkt, so existiert eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ bzw. $(y_n)_{n\in\mathbb{N}} \subset X$ die bestimmt gegen ∞ bzw. $-\infty$ divergiert. Wir schreiben dann $\sup X = \infty$ bzw. $\inf X = -\infty$. □

7.7.5 Definition

(i) Eine reelle Zahl s heißt **Maximum** von X, wenn gilt $s = \sup X$ und $s \in X$.

(ii) Eine reelle Zahl r heißt **Minimum** von X, wenn gilt $r = \inf X$ und $r \in X$. □

7.7.6 Bemerkung und Notation

Maximum bzw. Minimum sind im Falle der Existenz eindeutig bestimmt und werden dann mit $\max X$ bzw. $\min X$ bezeichnet. □

7.7.7 Beispiel

(i) Sicher gilt $\max [a, b] = b$ und $\min [a, b] = a$.

(ii) Wir zeigen $\sup]a,b[= b$. Sicher ist b eine obere Schranke von $]a,b[$. Ist $b' < b$ eine weitere reelle Zahl, so setzen wir

$$x = \max\left(\frac{a+b}{2}, \frac{b'+b}{2}\right).$$

Dann gilt $x \in]a,b[$ und $b' < x$. Also ist b' keine obere Schranke von $]a,b[$.
Analog erhält man $\inf]a,b[= a$. □

7.7.8 Satz und Definition

Seien $a \geq 0$ eine reelle Zahl und $n \in \mathbb{N}$. Dann hat die Menge

$$X := \{x \in \mathbb{R} \mid x \geqslant 0 \text{ und } x^n \leq a\}$$

ein Supremum s und für dieses gilt $s^n = a$. Weiter ist die Zahl s durch die Bedingungen $s \geqslant 0$ und $s^n = a$ eindeutig bestimmt. Wir nennen $\sqrt[n]{a} := s$ die n-**te Wurzel aus a**.

Beweis

Wir können annehmen, dass gilt $a > 0$. Sind x, y beliebige reelle Zahlen > 0, so folgt aus dem dritten Anordnungsaxiom $x < y \Leftrightarrow x^n < y^n$. Daraus ergibt sich die *Eindeutigkeit* der Wurzel im Falle der Existenz. Um die *Existenz* zu zeigen, bemerken wir zunächst, dass X wegen $0 \in X$ nicht leer ist. Weiter ist X nach oben beschränkt: Wegen der Bernoullischen Ungleichung gilt $\forall x \in X$ stets $x^n \leq a < 1 + na \leq (1+a)^n$ und somit auch $x < 1 + a$. Also $\exists s := \sup X$ nach 7.7.3. In den Aufgaben werden wir jede der beiden Annahmen $s^n < a$ und $s^n > a$ zu einem Widerspruch führen. Daraus folgt dann $s^n = a$. □

7.8 Limes superior und Limes inferior

7.8.1 Bemerkung und Definition

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen.

(i) Ist $(a_n)_{n\in\mathbb{N}}$ nach oben beschränkt, so ist $(\sup\{a_k \mid k \geq n\})_{n\in\mathbb{N}}$ eine monoton fallende Folge reeller Zahlen. Diese konvergiert also eigentlich gegen eine reelle Zahl oder uneigentlich gegen $-\infty$. Wir nennen dann

$$\overline{\lim}\, a_n := \quad \overline{\lim_{n\to\infty}}\, a_n := \quad \lim_{n\to\infty}(\sup\{a_k \mid k \geq n\})$$

den **Limes superior** von $(a_n)_{n\in\mathbb{N}}$. Ist $(a_n)_{n\in\mathbb{N}}$ nach oben unbeschränkt, so schreiben wir $\overline{\lim}\, a_n = \infty$.

(ii) Ist $(a_n)_{n\in\mathbb{N}}$ nach unten beschränkt, so ist $(\inf\{a_k \mid k \geq n\})_{n\in\mathbb{N}}$ eine monoton wachsende Folge reeller Zahlen. Diese konvergiert also eigentlich gegen eine reelle Zahl oder uneigentlich gegen ∞. Wir nennen dann

$$\underline{\lim}\, a_n := \quad \underline{\lim_{n\to\infty}}\, a_n := \quad \lim_{n\to\infty}(\inf\{a_k \mid k \geq n\})$$

den **Limes inferior** von $(a_n)_{n\in\mathbb{N}}$. Ist $(a_n)_{n\in\mathbb{N}}$ nach unten unbeschränkt, so schreiben wir $\underline{\lim}\, a_n = -\infty$. □

7.8.2 Beispiel

Für die durch $a_n = (-1)^n(1 + \frac{1}{n})$ definierte Folge gilt

$$\sup\{a_k \mid k \geq n\} \quad = \quad \begin{cases} 1 + \frac{1}{n}, & \text{falls} \quad n \quad \text{gerade}, \\ 1 + \frac{1}{n+1}, & \text{falls} \quad n \quad \text{ungerade}, \end{cases}$$

und somit $\overline{\lim\limits_{n\to\infty}}\ a_n = 1$. Entsprechend folgt $\underline{\lim\limits_{n\to\infty}}\ a_n = -1$. □

7.8.3 Satz

Seien $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge reeller Zahlen und H die Menge ihrer Häufungspunkte. Dann ist H nicht leer und besitzt sowohl ein Minimum als auch ein Maximum. Es gilt

$$\underline{\lim}\ a_n = \min H \quad \text{bzw.} \quad \overline{\lim}\ a_n = \max H.$$

Beweis

Aufgaben. □

7.8.4 Satz

Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen ist genau dann konvergent, wenn sie beschränkt ist und nur einen Häufungspunkt besitzt. In diesem Fall gilt

$$\underline{\lim}\ a_n = \lim a_n = \overline{\lim}\ a_n.$$

Beweis

Aufgaben. □

Kapitel 8

Reihen

8.1 Motivation

Reihen sind spezielle Folgen, die z. B. in der Form der Potenzreihen wichtige Anwendungen in der Approximation klassischer Funktionen wie sin, cos, exp und log haben. □

8.2 Konvergente Reihen

8.2.1 Definition

Ist $(a_k)_{k\in\mathbb{N}}$ eine Folge reeller Zahlen, so erhält man eine neue Folge $(s_n)_{n\in\mathbb{N}}$ durch Aufaddieren:

$$s_n := \sum_{k=1}^{n} a_k, \qquad n \in \mathbb{N}.$$

Die Folge $(s_n)_{n\in\mathbb{N}}$ heißt die (unendliche) **Reihe** mit den **Gliedern** $a_1, a_2, a_3, \ldots$ und den **Partialsummen** $s_1, s_2, s_3, \ldots$ und wird mit $\sum_{k=1}^{\infty} a_k$ bezeichnet. □

8.2.2 Definition

Die Reihe $\sum_{k=1}^{\infty} a_k$ heißt **konvergent**, wenn die Folge der Partialsummen konvergiert. Wir schreiben dann auch

$$s := \sum_{k=1}^{\infty} a_k := \lim_{n\to\infty} s_n$$

und nennen s den **Grenzwert** oder die **Summe** der Reihe. Eine Reihe heißt **divergent**, wenn sie nicht konvergiert. □

Zunächst ist also $\sum_{k=1}^{\infty} a_k$ nur ein Symbol. Erst im Falle der Konvergenz steht $\sum_{k=1}^{\infty} a_k$ für eine reelle Zahl. Analog zu unserem Vorgehen bei Folgen betrachten wir auch Reihen der Form

$$\sum_{k=k_0}^{\infty} a_k \text{ mit } k_0 \in \mathbb{Z}.$$

Im folgenden wählen wir oft $k_0 = 0$.

8.2.3 Beispiel

(i) **geometrische Reihe**: Ist $x \in \mathbb{R}$ mit $|x| < 1$, so ist $\sum_{k=0}^{\infty} x^k$ konvergent und es gilt

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

In der Tat liefert die geometrische Summenformel $s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$. Mit den Limessätzen und 7.2.6, (v) folgt $s_n \to \frac{1}{1-x}$.

(ii) **harmonische Reihe**: Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ ist divergent. Denn nach 6.2.5 ist jede konvergente Folge beschränkt. Wir zeigen aber: Für die harmonische Reihe ist die Folge der Partialsummen $s_n = \sum_{k=1}^{n} \frac{1}{k}$ unbeschränkt. In der Tat gilt für $n = 2^m, m \in \mathbb{N}_0$:

$$\begin{aligned} s_{2^m} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \ldots + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{m-1}+1} + \ldots + \frac{1}{2^m}\right) \\ &\geqslant 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \ldots + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^m} + \ldots + \frac{1}{2^m}\right) \\ &= 1 + \frac{m}{2} \end{aligned}$$

und diese Folge ist unbeschränkt. □

8.2.4 Satz (Rechenregeln für konvergente Reihen)

Sind $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ konvergent, so konvergieren auch $\sum_{k=0}^{\infty} (a_k \pm b_k)$ und $\sum_{k=0}^{\infty} \lambda \cdot a_k$ für $\lambda \in \mathbb{R}$ und es gilt

$$\sum_{k=0}^{\infty} (a_k \pm b_k) = \sum_{k=0}^{\infty} a_k \pm \sum_{k=0}^{\infty} b_k,$$

$$\sum_{k=0}^{\infty} \lambda \cdot a_k = \lambda \cdot \sum_{k=0}^{\infty} a_k.$$

Beweis

Dies folgt aus den Limessätzen 7.2.7, (iv). □

8.3 Konvergenzkriterien für Reihen

8.3.1 Satz (Cauchy-Kriterium)

Eine Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert genau dann, wenn gilt:

$$\forall \varepsilon > 0 \; \exists n_0(\varepsilon) \text{ mit } \left| \sum_{k=m}^{n} a_k \right| < \varepsilon \quad \forall n \geqslant m \geqslant n_0(\varepsilon).$$

Beweis

Dies folgt aus 7.3.2 und 7.3.3. □

8.3.2 Korollar

Ist $\sum_{k=0}^{\infty} a_k$ konvergent, so ist $(a_k)_{k\in\mathbb{N}_0}$ eine Nullfolge.

Beweis

Dies folgt mit $n = m$ aus 8.3.1. □

Das Beispiel der harmonischen Reihe zeigt, dass die Umkehrung im Allgemeinen falsch ist.

8.3.3 Satz (Monotoniekriterium für Reihen)

Eine Reihe mit nicht-negativen Gliedern konvergiert genau dann, wenn die Folge ihrer Partialsummen beschränkt ist.

Beweis

Dies folgt aus 7.5.2 und 7.2.5 . □

8.3.4 Satz (Leibniz-Kriterium für alternierende Reihen)

Sei $(a_k)_{k\in\mathbb{N}}$ eine monoton fallende Folge nicht-negativer reeller Zahlen mit $a_k \to 0$. Dann konvergiert

$$\sum_{k=0}^{\infty} (-1)^k \cdot a_k.$$

Beweis

Wir setzen $s_n := \sum_{k=0}^{n} (-1)^k \cdot a_k \ \forall n \in \mathbb{N}_0$.

Wegen $s_{2k+2} - s_{2k} = -a_{2k+1} + a_{2k+2} \leqslant 0$ ist die Folge $(s_{2k})_{k\in\mathbb{N}}$ monoton fallend. Entsprechend ist wegen $s_{2k+3} - s_{2k+1} = a_{2k+2} - a_{2k+3} \geqslant 0$ die Folge $(s_{2k+1})_{k\in\mathbb{N}}$ monoton wachsend. Weiter gilt

$$s_{2k+1} = s_{2k} - a_{2k+1} \leqslant s_{2k} \quad \forall k \in \mathbb{N}_0.$$

Also ist $(s_{2k})_{k\in\mathbb{N}_0}$ nach unten durch s_1 und $(s_{2k+1})_{k\in\mathbb{N}_0}$ nach oben durch s_0 beschränkt. Nach dem Monotoniekriterium sind beide Folgen konvergent, etwa

$$\lim_{k\to\infty} (s_{2k}) = s \quad \text{und} \quad \lim_{k\to\infty} (s_{2k+1}) = s'.$$

Wegen $s_{2k} - s_{2k+1} = a_{2k+1} \to 0$ folgt $s = s'$.

Ist nun $\varepsilon > 0$ beliebig, so $\exists n_1, n_2 \in \mathbb{N}_0$ mit

$$|s_{2k} - s| < \varepsilon \quad \forall k \geqslant n_1 \quad \text{und} \quad |s_{2k+1} - s| < \varepsilon \quad \forall k \geqslant n_2.$$

Setzen wir $n_0 := \max(2 \cdot n_1, 2 \cdot n_2 + 1)$, so gilt

$$|s_n - s| < \varepsilon \quad \forall n \geqslant n_0.$$

□

8.3.5 Beispiel

Die alternierende harmonische Reihe $\sum_{k=1}^{\infty} (-1)^{k-1} \cdot \frac{1}{k}$ konvergiert nach dem Leibniz-Kriterium. Gleiches gilt für die Reihe $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$. Mit dem jeweiligen Grenzwert werden wir uns später befassen. □

8.4 Absolut konvergente Reihen

Nach dem allgemeinen Kommutativegesetz darf man bei endlichen Summen die Reihenfolge der Summation vertauschen, ohne dass sich die Summe ändert (man darf beliebig umordnen). Um bei Reihen umordnen zu können, braucht man absolute Konvergenz.

8.4.1 Definition

Eine Reihe $\sum_{k=0}^{\infty} a_k$ heißt **absolut konvergent**, wenn die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergiert. □

8.4.2 Bemerkung

(i) Mit Hilfe der Dreiecksungleichung ergibt sich aus dem Cauchy-Kriterium, dass jede absolut konvergente Reihe auch konvergent ist. Das Beispiel der alternierenden harmonischen Reihe zeigt, dass die Umkehrung im Allgemeinen nicht gilt.

(ii) Das Monotoniekriterium für Reihen zeigt, dass eine Reihe $\sum_{k=0}^{\infty} a_k$ genau dann absolut konvergiert, wenn die Folge $\left(\sum_{k=0}^{n} |a_k|\right)_{n \in \mathbb{N}_0}$ beschränkt ist. □

8.5 Kriterien für absolute Konvergenz

8.5.1 Satz (Majorantenkriterium)

Sind $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ Reihen mit $b_k \geqslant 0 \ \forall k \in \mathbb{N}_0$, so gilt:

(i) Ist $\sum_{k=0}^{\infty} b_k$ konvergent und gilt fast immer $|a_k| \leqslant b_k$, so ist $\sum_{k=0}^{\infty} a_k$ absolut konvergent.

(ii) Ist $\sum_{k=0}^{\infty} b_k$ divergent und gilt fast immer $a_k \geqslant b_k$, so ist $\sum_{k=0}^{\infty} a_k$ divergent.

Im Falle (i) heißt $\sum_{k=0}^{\infty} b_k$ eine **konvergente Majorante**, im Falle (ii) eine **divergente Minorante**.

Beweis

(i) Nach Voraussetzung $\exists k_0 \in \mathbb{N}_0$ mit

$$\sum_{k=k_0}^{n} |a_k| \leqslant \sum_{k=k_0}^{n} b_k \leqslant \sum_{k=0}^{\infty} b_k \quad \forall n \geqslant k_0.$$

Also ist die Folge der Partialsummen beschränkt. Das Monotoniekriterium liefert die Behauptung.

(ii) folgt aus (i). □

8.5.2 Satz (Wurzelkriterium)

Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe.

(i) Gibt es eine reelle Zahl $0 < q < 1$, sodass fast immer gilt $\sqrt[k]{|a_k|} \leqslant q$, so ist $\sum_{k=0}^{\infty} a_k$ absolut konvergent.

(ii) Gilt hingegen fast immer $\sqrt[k]{|a_k|} \geqslant 1$, so ist $\sum_{k=0}^{\infty} a_k$ divergent.

Beweis

(i) Die Voraussetzung impliziert, dass fast immer gilt $|a_k| \leq q^k$. Also ist die geometrische Reihe $\sum_{k=0}^{\infty} q^k$ eine konvergente Majorante für $\sum_{k=0}^{\infty} a_k$.

(ii) Die Voraussetzung impliziert, dass $(a_k)_{k\in\mathbb{N}_0}$ keine Nullfolge ist. □

8.5.3 Satz (Quotientenkriterium)

Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe.

(i) Gibt es eine reelle Zahl $0 < q < 1$, sodass fast immer gilt $a_k \neq 0$ und $\left|\frac{a_{k+1}}{a_k}\right| \leqslant q$, so ist $\sum_{k=0}^{\infty} a_k$ absolut konvergent.

(ii) Gilt hingegen fast immer $a_k \neq 0$ und $\left|\frac{a_{k+1}}{a_k}\right| \geqslant 1$, so ist $\sum_{k=0}^{\infty} a_k$ divergent.

Beweis

(i) Der erste Teil der Voraussetzung bedeutet, dass es ein $m \in \mathbb{N}_0$ gibt mit $a_k \neq 0$ für alle $k \geq m$. Aus dem zweiten Teil der Voraussetzung ergibt sich dann durch vollständige Induktion: $|a_k| \leq |a_m|\, q^k$. Also ist $|a_m| \sum_{k=0}^{\infty} q^k$ eine konvergente Majorante.

(ii) Wie beim Wurzelkriterium. □

8.5.4 Korollar

Ist $\sum_{k=0}^{\infty} a_k$ eine Reihe und existiert der Grenzwert $\lim_{k\to\infty} \sqrt[k]{|a_k|} = c$ bzw. $\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = c$, so gilt:

(i) $c < 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ absolut konvergent,

(ii) $c > 1 \Rightarrow \sum\limits_{k=0}^{\infty} a_k$ divergent.

Beweis

Dies folgt unmittelbar aus 8.5.3 und 8.5.2. □

8.5.5 Beispiel

(i) **Warnung:** Im Falle der harmonischen Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ gilt $\left|\frac{a_{k+1}}{a_k}\right| = \frac{k}{k+1} < 1 \ \forall k \geqslant 1$. Wegen $\lim_{k\to\infty} \frac{k}{k+1} = 1$ ist die Bedingung in 8.5.3, (i) aber nicht erfüllt.

(ii) Für $x \in \mathbb{R}$ betrachten wir die Reihe $\sum_{k=1}^{\infty} \frac{x^{k-1}}{k}$. Ist $x = 0$, so liegt absolute Konvergenz vor. Andernfalls gilt

$$\left|\frac{\frac{x^k}{k+1}}{\frac{x^{k-1}}{k}}\right| = \frac{k}{k+1} \cdot |x| \to |x| \,.$$

Also ist die Reihe nach 8.5.4 absolut konvergent, falls $|x| < 1$ und divergent, falls $|x| > 1$. Für $x = 1$ erhalten wir die harmonische Reihe und somit Divergenz, für $x = -1$ erhalten wir die alternierende harmonische Reihe und somit Konvergenz.

(iii) Für $n \in \mathbb{N}$ gilt: $\sum_{k=1}^{n} \frac{1}{k\cdot(k+1)} = \frac{n}{n+1}$, wie man leicht durch Induktion zeigt. Also konvergiert $\sum_{k=1}^{\infty} \frac{1}{k\cdot(k+1)}$ und es gilt $\sum_{k=1}^{\infty} \frac{1}{k\cdot(k+1)} = 1$. Nach dem Majorantenkriterium konvergieren dann auch die Reihen $\sum_{k=1}^{\infty} \frac{1}{k^n}, n \geqslant 2$. In der Tat gilt $\forall k \in \mathbb{N}$ und $n \geqslant 2$:

$$\frac{1}{k^n} \leqslant \frac{1}{k^2} \leqslant \frac{2}{k \cdot (k+1)}.$$

□

8.6 Umordnung von Reihen

8.6.1 Definition

Ist $\sum_{k=0}^{\infty} a_k$ eine Reihe und $\sigma : \mathbb{N}_0 \to \mathbb{N}_0$ eine Permutation, so heißt $\sum_{k=0}^{\infty} a_{\sigma(k)}$ eine **Umordnung** von $\sum_{k=0}^{\infty} a_k$. □

8.6.2 Umordnungssatz

Ist $\sum_{k=0}^{\infty} a_k$ eine absolut konvergente Reihe, so ist auch jede Umordnung $\sum_{k=0}^{\infty} a_{\sigma(k)}$ absolut konvergent und es gilt

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_{\sigma(k)}.$$

Beweis

(i) Ist $N \in \mathbb{N}_0$, so gibt es ein $n_1 \in \mathbb{N}_0$ mit

$$\{\sigma(0), \ldots, \sigma(N)\} \subset \{0, \ldots, n_1\}.$$

Also gilt

$$\sum_{k=0}^{N} \left|a_{\sigma(k)}\right| \leqslant \sum_{k=0}^{n_1} |a_k| \leqslant \sum_{k=0}^{\infty} |a_k| \,.$$

Nach dem Monotoniekriterium ist $\sum_{k=0}^{\infty} a_{\sigma(k)}$ absolut konvergent.

(ii) Wir schreiben

$$s_n = \sum_{k=0}^{n} a_k, \quad t_n = \sum_{k=0}^{n} a_{\sigma(k)}$$

und zeigen $\lim_{n\to\infty}(s_n - t_n) = 0$ (woraus sich dann $\lim_{n\to\infty} t_n = \lim_{n\to\infty} s_n$ ergibt): Sei $\varepsilon > 0$ beliebig. Nach dem Cauchy-Kriterium $\exists n_0 \in \mathbb{N}_0$ mit

$$\sum_{k=n_0+1}^{n} |a_k| < \varepsilon \quad \forall n \geqslant n_0 + 1.$$

Wir wählen ein $N_0 \in \mathbb{N}$ mit

$$\{0, \ldots, n_0\} \subset \{\sigma(0), \ldots, \sigma(N_0)\}$$

und betrachten ein beliebiges $N \geqslant N_0$. Zu N wählen wir ein $n_1 \geq n_0 + 1$ wie in (i). Dann gilt

$$|s_N - t_N| \leqslant \sum_{k=n_0+1}^{n_1} |a_k| < \varepsilon.$$

Dies zeigt die Behauptung. □

Ohne die Voraussetzung der absoluten Konvergenz ist Satz 8.6.2 falsch. Bevor wir ein Beispiel geben, stellen wir fest:

8.6.3 Bemerkung

Ist $\sum_{k=0}^{\infty} a_k$ eine konvergente Reihe, so dürfen ihre Glieder durch Klammern beliebig zusammengefasst werden. Genauer gilt: Ist $0 \leqslant k_1 < k_2 < \ldots$ und setzt man

$$A_1 := a_0 + \ldots + a_{k_1}, \quad A_2 := a_{k_1+1}, \ldots, a_{k_2}, \quad \ldots,$$

so ist auch $\sum_{\nu=1}^{\infty} A_\nu$ konvergent und es gilt:

$$\sum_{\nu=1}^{\infty} A_\nu = \sum_{k=0}^{\infty} a_k.$$

In der Tat ist die Folge der Partialsummen $(A_1 + \ldots + A_m)_{m\in\mathbb{N}}$ eine Teilfolge der Folge der Partialsummen $(a_0 + \ldots + a_n)_{n\in\mathbb{N}}$. □

8.6.4 Beispiel

Wir betrachten die alternierende harmonische Reihe mit Grenzwert

$$s := \sum_{k=1}^{\infty} (-1)^{k-1} \cdot \frac{1}{k}.$$

Nach obiger Bemerkung dürfen wir schreiben

$$\begin{aligned}\frac{1}{2}\cdot s &= \sum_{\nu=1}^{\infty}\frac{1}{2}\cdot\left(\frac{1}{2\nu-1}-\frac{1}{2\nu}\right)\\ &= \sum_{\nu=1}^{\infty}\left(\frac{1}{2\nu-1}-\frac{1}{4\nu-2}-\frac{1}{4\nu}\right)\\ &= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\dots\ .\end{aligned}$$

Letzteres ist aber eine Umordnung von

$$\sum_{k=1}^{\infty}(-1)^{k-1}\cdot\frac{1}{k}\ .$$

□

8.7 Produkt von Reihen

8.7.1 Satz (Cauchy-Produkt von Reihen)

Es seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergente Reihen. Für $n \in \mathbb{N}_0$ setzen wir

$$c_n := \sum_{k=0}^{n} a_{n-k}\cdot b_k.$$

Dann ist auch die Reihe $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{k=0}^{\infty} a_k\right)\cdot\left(\sum_{k=0}^{\infty} b_k\right).$$

Beweis

Siehe Heuser, 32.6. □

8.7.2 Beispiel

Für jedes $x \in \mathbb{R}$ ist die Reihe

$$\sum_{k=0}^{\infty}\frac{x^k}{k!}$$

absolut konvergent. In der Tat folgt dies aus dem Quotientenkriterium, denn ist $x \neq 0$ und $k \geqslant 2\cdot|x|$ -1, so gilt

$$\left|\frac{\frac{x^{k+1}}{(k+1)!}}{\frac{x^k}{k!}}\right| = \frac{|x|}{k+1} \leqslant \frac{1}{2}.$$

Sind $x, y \in \mathbb{R}$, so gilt nach dem binomischen Lehrsatz

$$c_n := \sum_{k=0}^{n}\frac{x^{n-k}}{(n-k)!}\cdot\frac{y^k}{k!} = \frac{1}{n!}\cdot\sum_{k=0}^{n}\binom{n}{k}\cdot x^{n-k}\cdot y^k = \frac{1}{n!}\cdot(x+y)^n\,.$$

Mit 7.7.1 folgt

$$\left(\sum_{k=0}^{\infty}\frac{x^k}{k!}\right)\cdot\left(\sum_{k=0}^{\infty}\frac{y^k}{k!}\right) = \sum_{n=0}^{\infty}\frac{(x+y)^n}{n!}.$$

□

8.8 Die Exponentialfunktion

8.8.1 Bemerkung und Definition

Bereits in 7.5.4 hatten wir die Konvergenz von $\sum_{k=0}^{\infty} \frac{1}{k!}$ gezeigt und den Grenzwert e genannt (**Eulersche Zahl**). Nun setzen wir allgemein

$$\exp(x) := \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \forall x \in \mathbb{R},$$

und nennen die Abbildung

$$\exp : \mathbb{R} \to \mathbb{R},\ x \mapsto \exp(x),$$

die **Exponentialfunktion** (reellwertige Abbildungen werden oft als Funktionen bezeichnet). Nach 8.7.2 gilt die **Funktionalgleichung**

$$\exp(x+y) = \exp(x) \cdot \exp(y).$$

□

8.8.2 Korollar

(i) $\forall x \in \mathbb{R}$ gilt $\exp(x) > 0$.

(ii) $\forall x \in \mathbb{R}$ gilt $\exp(-x) = (\exp(x))^{-1}$.

(iii) $\forall n \in \mathbb{Z}$ ist $\exp(n) = e^n$.

Beweis

Dies folgt aus der Reihendarstellung bzw. der Funktionalgleichung. Siehe Aufgaben. □

8.9 Potenzreihen

Wie das Beispiel der Exponentialfunktion zeigt, können wir Reihen benutzen, um wichtige Funktionen zu definieren. Später werden wir sehen, dass man sie oft auch benutzen kann, um bereits definierte Funktionen zu approximieren. In diesem Zusammenhang definieren wir:

8.9.1 Definition

Sind $(a_k)_{k\in\mathbb{N}_0}$ eine Folge reeller Zahlen und $x_0 \in \mathbb{R}$, so nennen wir eine Reihe der Form

$$\sum_{k=0}^{\infty} a_k \cdot (x - x_0)^k$$

Potenzreihe mit **Entwicklungspunkt** x_0 und **Koeffizienten** a_k. □

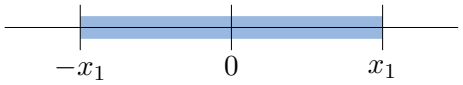
Der Einfachheit halber betrachten wir im Folgenden $x_0 = 0$ (durch die Substitution $\widetilde{x} = x - x_0$ kann man jede Potenzreihe zu einer Potenzreihe mit Entwicklungspunkt 0 machen).

8.9.2 Satz

Konvergiert die Potenzreihe

$$\sum_{k=0}^{\infty} a_k \cdot x^k$$

für ein $0 \neq x_1 \in \mathbb{R}$ (das man für x einsetzt), so konvergiert sie absolut $\forall x \in \mathbb{R}$ mit $|x| < |x_1|$.



Beweis

Da $\sum_{k=0}^{\infty} a_k \cdot x_1^k$ konvergiert, ist $\left(a_k \cdot x_1^k\right)_{k\in\mathbb{N}_0}$ eine Nullfolge. Insbesondere ist die Folge beschränkt. Also $\exists M > 0$ mit $\left|a_k \cdot x_1^k\right| \leqslant M \ \forall k \in \mathbb{N}_0$. Somit gilt $\forall x \in \mathbb{R}$:

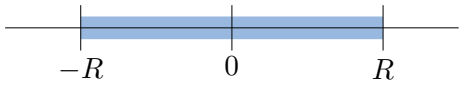
$$\left|a_k \cdot x^k\right| = \left|a_k \cdot x_1^k \cdot \frac{x^k}{x_1^k}\right| \leqslant M \cdot \left|\frac{x}{x_1}\right|^k \quad \forall k \in \mathbb{N}_0.$$

Gilt nun $|x| < |x_1|$, so ist $\left|\frac{x}{x_1}\right| < 1$ und somit $M \cdot \sum_{k=0}^{\infty} \left|\frac{x}{x_1}\right|^k$ eine konvergente Majorante für $\sum_{k=0}^{\infty} a_k \cdot x^k$ (geometrische Reihe). □

8.9.3 Korollar und Definition

Sei $\sum_{k=0}^{\infty} a_k \cdot x^k$ eine Potenzreihe. Dann gilt genau eine der beiden folgenden Aussagen:

(i) $\exists$ reelle Zahl $R \geqslant 0$, sodass die Potenzreihe $\forall x \in \mathbb{R}$ mit $|x| < R$ absolut konvergiert und $\forall x \in \mathbb{R}$ mit $|x| > R$ divergiert.



(ii) Die Potenzreihe konvergiert $\forall x \in \mathbb{R}$.

Im Falle (ii) schreiben wir auch $R = \infty$. In jedem Fall nennen wir R den **Konvergenzradius** der Potenzreihe.

Beweis

Dies folgt mit

$$R := \sup\left\{x \in \mathbb{R} \;\middle|\; \sum_{k=0}^{\infty} a_k \cdot x^k \text{ konvergiert}\right\}$$

aus 8.9.2. □

8.9.4 Satz (Cauchy-Hadamard)

Sei $\sum_{k=0}^{\infty} a_k \cdot x^k$ eine Potenzreihe mit Konvergenzradius R. Dann gilt

$$R = \frac{1}{\overline{\lim}\sqrt[k]{|a_k|}} \ .$$

Dabei setzt man $\frac{1}{\infty} = 0$ und $\frac{1}{0} = \infty$.

Beweis

Dies folgt aus dem Wurzelkriterium. □

8.9.5 Bemerkung

Gilt fast immer $a_k \neq 0$ und existiert der Grenzwert

$$c = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| \in \mathbb{R} \cup \{\pm\infty\},$$

so liefert das Quotientenkriterium entsprechend

$$R = \frac{1}{c}.$$ □

8.9.6 Beispiel

(i) Für die Exponentialreihe erhalten wir, wie erwartet,

$$\frac{k!}{(k+1)!} = \frac{1}{k+1} \to 0,$$

also $R = \infty$.

(ii) Für die geometrische Reihe $\sum_{k=0}^{\infty} x^k$ sind alle Koeffizienten und somit auch R gleich 1.

(iii) Für die Reihe $\sum_{k=1}^{\infty} \frac{x^{k-1}}{k}$ ist $R = 1$ nach 8.5.5, (ii). Dieses Beispiel zeigt, dass wir für $|x| = R$ keine allgemeinen Konvergenzaussagen treffen können. □

Kapitel 9

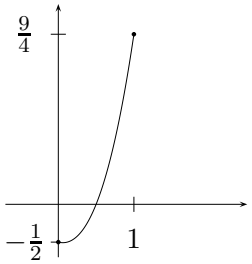
Stetige Funktionen

9.1 Motivation

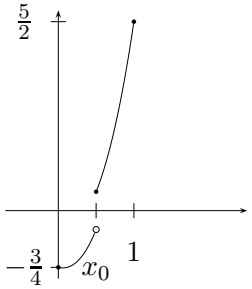
Das mathematische Modellieren von Anwendungsproblemen führt oft zu Gleichungen oder Gleichungssystemen, um deren Lösung man sich kümmern muss. Wir betrachten den Fall einer reellen Gleichung mit einer reellen Unbekannten: Gegeben sei eine Funktion $f : I \to \mathbb{R}$, die auf einem Intervall $I \subset \mathbb{R}$ definiert ist und gesucht sind alle $x \in I$ mit $f(x) = 0$. Diese x heißen auch **Nullstellen** von f. Als Beispiel fragen wir uns, ob die Funktion

$$f(x) = x^3 + 2 \cdot x^2 - \frac{1}{4} \cdot x - \frac{1}{2}$$

eine Nullstelle im Intervall $I = [0, 1]$ hat. Einsetzen der Randpunkte von I liefert $f(0) = -\frac{1}{2}$, $f(1) = \frac{9}{4}$. Wegen des Vorzeichenwechsels erwarten wir bei kontinuierlichen Verlauf des Graphen von f in der Tat eine Nullstelle, das heißt einen Punkt, bei dem der Graph die x-Achse schneidet:



„Springt“ die Funktion hingegen, so braucht keine Nullstelle vorzuliegen:



Im ersten Beispiel sprechen wir von einer stetigen Funktion, im zweiten Beispiel ist die Funktion unstetig im Punkt x_0. Wir wollen den Begriff der Stetigkeit formal einführen. □

9.2 Definition der Stetigkeit

9.2.1 Definition

Seien $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ eine Funktion. Dann heißt f **stetig** in einem Punkt $x_0 \in X$, wenn für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset X$ gilt:

$$x_n \to x_0 \;\Rightarrow\; f(x_n) \to f(x_0).$$

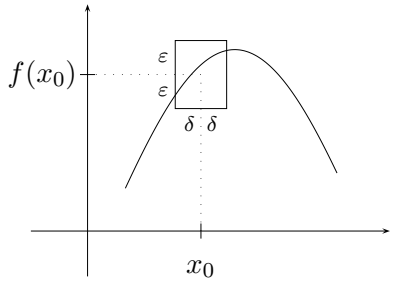
Wir nennen f **stetig** (in X), falls f stetig ist in jedem Punkt von X. □

9.2.2 Satz (ε-δ–Kriterium der Stetigkeit)

Seien $X \subset \mathbb{R}$, $f : X \to \mathbb{R}$ eine Funktion und $x_0 \in X$. Dann ist f genau dann stetig in x_0, wenn gilt:

$$\forall \varepsilon > 0 \; \exists \delta > 0, \text{ sodass } \forall x \in X \text{ gilt :}$$

$$|x - x_0| < \delta \;\Rightarrow\; |f(x) - f(x_0)| < \varepsilon.$$



Beweis

"⇒" Sei f stetig in x_0.

- **Annahme**: Die ε-δ–Bedingung ist nicht erfüllt.

Dann $\exists\, \epsilon > 0$, so dass kein $\delta > 0$ wie gewünscht existiert. Mit anderen Worten:

$$\forall\ \delta > 0 \quad \exists\ x \in X \ \text{ mit } \ |x - x_0| < \delta \ \text{ aber } \ |f(x) - f(x_0)| \geq \epsilon.$$

Insbesondere gibt es dann $\forall\ n \in \mathbb{N}$ ein $x_n \in X$ mit

$$|x_n - x_0| < \frac{1}{n} \quad \text{aber} \quad |f(x_n) - f(x_0)| \geq \epsilon\ .$$

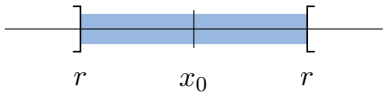
Dann haben wir im Widerspruch zur Voraussetzung eine Folge $(x_n)_{n\in\mathbb{N}} \subset X$ gefunden mit $x_n \to x_0$ aber $f(x_n) \nrightarrow f(x_0)$.

"⇐" Sei nun die ε-δ–Bedingung erfüllt. Ist dann $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge mit $x_n \to x_0$, so müssen wir zeigen: $f(x_n) \to f(x_0)$. Sei dazu $\epsilon > 0$. Nach Voraussetzung $\exists\ \delta > 0$ mit $x \in X,\ |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$. Da $x_n \to x_0$ existiert zu diesem δ ein $n_0 \in \mathbb{N}$ mit $|x_n - x_0| < \delta \quad \forall\ n \geq n_0$. Es folgt $|f(x_n) - f(x_0)| < \epsilon\ \forall\ n \geq n_0$ und somit $f(x_n) \to f(x_0)$. □

9.2.3 Bemerkung und Notation

Sind $x_0 \in \mathbb{R}$ und $r > 0$, so setzt man auch

$$U_r(x_0) :=]x_0 - r, x_0 + r[= \{x \in \mathbb{R} \mid |x - x_0| < r\}.$$



Mit dieser Notation lautet die ε-δ–Bedingung:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \;\text{ mit }\; x \in X \cap U_\delta(x_0) \;\Rightarrow\; f(x) \in U_\varepsilon(f(x_0)).$$ □

9.2.4 Definition

Sind $f, g : X \to \mathbb{R}$ Funktionen und $\lambda \in \mathbb{R}$, so definiert man Funktionen

$$\begin{aligned} f \pm g &: X \to \mathbb{R},\ x \mapsto f(x) \pm g(x), \\ f \cdot g &: X \to \mathbb{R},\ x \mapsto f(x) \cdot g(x), \\ \lambda \cdot f &: X \to \mathbb{R},\ x \mapsto \lambda \cdot f(x). \end{aligned}$$

Setzt man $X' := \{x \in X \mid g(x) \neq 0\}$, so definiert man die Funktion

$$\frac{f}{g} : X' \to \mathbb{R},\ x \mapsto \frac{f(x)}{g(x)}.$$ □

9.2.5 Satz

Seien $f, g : X \to \mathbb{R}$ Funktionen, $\lambda \in \mathbb{R}$ und $x_0 \in X$. Dann gilt: Sind f und g stetig in x_0, so auch $f \pm g$, $f \cdot g$ und $\lambda \cdot f$. Gleiches gilt für $\frac{f}{g}$, falls $g(x_0) \neq 0$.

Beweis

Dies folgt direkt aus den Limessätzen für konvergente Folgen. □

9.2.6 Satz (Komposition stetiger Funktionen)

Seien $f : X \to \mathbb{R}$ und $g : Y \to \mathbb{R}$ Funktionen mit $f(X) \subset Y$ und sei $x_0 \in X$. Dann gilt:

$$f \text{ stetig in } x_0,\ g \text{ stetig in } f(x_0) \Rightarrow g \circ f \text{ stetig in } x_0.$$

Beweis

Sei $(x_n)_{n\in\mathbb{N}} \subset X$ eine Folge mit $x_n \to x_0$. Wegen der Stetigkeit von f in x_0 gilt dann $f(x_n) \to f(x_0)$. Nach Voraussetzung gilt aber $(f(x_n))_{n\in\mathbb{N}} \subset Y$. Also liefert die Stetigkeit von g in $f(x_0)$ die Behauptung: $(g \circ f)(x_n) = g(f(x_n)) \to g(f(x_0)) = (g \circ f)(x_0)$. □

9.2.7 Bemerkung

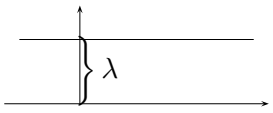
Unmittelbar aus der Definition der Stetigkeit ergibt sich:

(i) Einschränkungen stetiger Funktionen sind stetig: Sind $Y \subset X \subset \mathbb{R}$ Teilmengen, ist $x_0 \in Y$ und ist die Funktion $f : X \to \mathbb{R}$ stetig in x_0, so ist auch $f|_Y$ stetig in x_0.

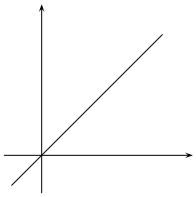
(ii) Die Stetigkeit ist eine lokale Eigenschaft in folgendem Sinn: Sind $f : X \to \mathbb{R}$ und $g : Y \to \mathbb{R}$ Funktionen auf Teilmengen $X, Y \subset \mathbb{R}$, ist $x_0 \in X \cap Y$ und gibt es ein $r > 0$ mit $U_r(x_0) \cap X = U_r(x_0) \cap Y$ und $f|_{U_r(x_0)\cap X} = g|_{U_r(x_0)\cap Y}$, so ist f genau dann stetig in x_0, wenn g stetig ist in x_0. □

9.2.8 Beispiel

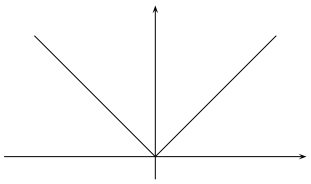
(i) **Konstante Funktionen**: Ist $\lambda \in \mathbb{R}$, so ist $\mathbb{R} \to \mathbb{R},\ x \mapsto \lambda$, stetig.



(ii) Die identische Abbildung $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R},\ x \mapsto x$, ist stetig.



(iii) Die **Betragsfunktion** $|\ | : \mathbb{R} \to \mathbb{R},\ x \mapsto |x|$, ist stetig.



In der Tat gilt für $x_0 \in \mathbb{R}$: Ist $x_0 > 0$, so stimmt $|\ |$ auf $U_{|x_0|}(x_0) =]0, 2x_0[$ mit $\mathrm{id}_{\mathbb{R}}$ überein, ist also nach (ii) und 9.2.7 in x_0 stetig. Ist $x_0 < 0$, so stimmt $|\ |$ auf $U_{|x_0|}(x_0) =]2x_0, 0[$ mit $-\mathrm{id}_{\mathbb{R}}$ überein und die Behauptung folgt aus (ii), 9.2.4 und 9.2.7. Ist $x_0 = 0$, so folgt aus $x_n \to 0 = x_0$ auch $|x_n| \to 0 = f(x_0)$.

(iv) Ist $f : X \to \mathbb{R}$ stetig, so auch $|f| : X \to \mathbb{R},\ x \mapsto |f(x)|$, nach (iii) und 9.2.6.

(v) Polynome mit Koeffizienten in $\mathbb{R}$ definieren Funktionen der Form

$$p : \mathbb{R} \to \mathbb{R},\ x \mapsto p(x) = \sum_{i=0}^{n} a_i x^i, \qquad a_0, \ldots, a_n \in \mathbb{R}.$$

Diese heißen auch **Polynomfunktionen** und sind nach (i), (ii) und 9.2.4 stetig.

(vi) Funktionen der Form

$$X \to \mathbb{R},\ x \mapsto \frac{p(x)}{q(x)}, \quad p, q \text{ Polynomfunktionen},\ X := \{x \in \mathbb{R} \mid q(x) \neq 0\},$$

heißen **rationale Funktionen**. Sie sind stetig nach (v) und 9.2.5. □

9.3 Grenzwerte von Funktionen

Sei $X \subset \mathbb{R}$. Wir schreiben

$$\overline{X} := \{x \in \mathbb{R} \mid \exists \text{ Folge } (x_n)_{n\in\mathbb{N}} \subset X \text{ mit } x_n \to x\}.$$

Dann gilt $X \subset \overline{X}$. Ist z.B. $X =]a, b[$ ein offenes Intervall, so ist $\overline{X} = [a, b]$ (siehe 7.7.7, (ii)).

9.3.1 Definition

Seien $f : X \to \mathbb{R}$ eine Funktion und $x_0 \in \overline{X}$. Dann schreiben wir

$$\lim_{x\to x_0} f(x) = a, \quad \text{mit } a \in \mathbb{R},$$

falls für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset X$ gilt:

$$x_n \to x_0 \Rightarrow f(x_n) \to a.$$

In diesem Fall sagen wir, f besitzt den **Grenzwert** a für $x \to x_0$ sowie f **konvergiert** (oder **strebt**) **gegen a für** $x \to x_0$. □

Die **Limessätze** für konvergente Folgen übertragen sich auf Funktionen. Zum Beispiel gilt:

$$\lim_{x\to x_0} (f(x) + g(x)) = \left(\lim_{x\to x_0} f(x)\right) + \left(\lim_{x\to x_0} g(x)\right).$$

Diese Aussage bedeutet wie üblich: Existieren die Grenzwerte auf der rechten Seite, so auch der auf der linken Seite und es gilt die angegebene Gleichheit.

9.3.2 Bemerkung

Seien $f : X \to \mathbb{R}$ eine Funktion und $x_0 \in \overline{X}$.

(i) Ist $x_0 \in X$, so ist f genau dann stetig in x_0, wenn $\lim_{x\to x_0} f(x) =: a \in \mathbb{R}$ existiert. In diesem Fall gilt $a = f(x_0)$ da man insbesondere die konstante Folge mit $x_n = x_0\ \forall n \in \mathbb{N}$ betrachten kann.

(ii) Ist $x_0 \notin X$ und existiert $\lim_{x\to x_0} f(x) =: a \in \mathbb{R}$, so ist die Funktion

$$\widetilde{f} : X \cup \{x_0\} \to \mathbb{R},\ x \to \begin{cases} f(x), & \text{falls } x \neq x_0, \\ a, & \text{falls } x = x_0, \end{cases}$$

stetig in x_0. Wir sprechen dann von der **stetigen Fortsetzung** von f in x_0. □

9.3.3 Definition (Uneigentliche Grenzwerte, Grenzwerte im Unendlichen)

(i) Seien $f : X \to \mathbb{R}$ eine Funktion und $x_0 \in \overline{X}$. Dann schreiben wir

$$\lim_{x\to x_0} f(x) = \infty,$$

falls für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset X$ gilt:

$$x_n \to x_0 \Rightarrow f(x_n) \to \infty.$$

Entsprechend definiert man $\lim_{x\to x_0} f(x) = -\infty$.

(ii) Ist X nach oben unbeschränkt und $f : X \to \mathbb{R}$ eine Funktion, so schreiben wir

$$\lim_{x\to\infty} f(x) = a, \quad \text{mit } a \in \mathbb{R} \text{ oder } a = \pm\infty,$$

wenn $\forall$ Folge $(x_n)_{n\in\mathbb{N}} \subset X$ gilt:

$$x_n \to \infty \Rightarrow f(x_n) \to a.$$

Analog definiert man $\lim_{x\to-\infty} f(x) = a$.

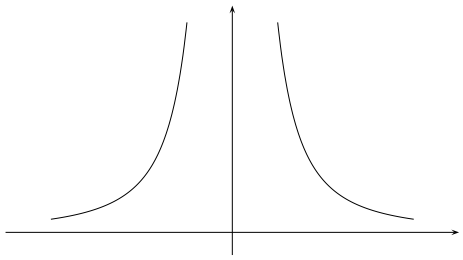
Wir verwenden die zu 9.3.1 analogen Sprechweisen (in (i) ist der Grenzwert uneigentlich; gleiches gilt in (ii) falls $a = \pm\infty$). □

9.3.4 Beispiel

(i) Für die Funktion

$$f : X := \mathbb{R} \setminus \{0\} \to \mathbb{R},\ x \mapsto \frac{1}{x^2},$$

gilt $\lim_{x\to 0} f(x) = \infty$.



(ii) Für die rationale Funktion

$$f : [0,\infty[\to \mathbb{R},\ x \mapsto \frac{3x^3 + 2x - 1}{2x^3 + 6},$$

gilt

$$\lim_{x\to\infty} f(x) = \frac{3}{2},$$

denn für jede Folge $(x_n)_{n\in\mathbb{N}} \subset]0,\infty[$ mit $x_n \to \infty$ gilt:

$$f(x_n) = \frac{3x_n^3 + 2x_n - 1}{2x_n^3 + 6} = \frac{3 + \frac{2}{x_n^2} - \frac{1}{x_n^3}}{2 + \frac{6}{x_n^3}} \to \frac{3}{2}.$$

(iii) Das Wachstum einer Polynomfunktion für $x \to \pm\infty$ wird durch den Term höchsten Grades bestimmt: Ist $p : \mathbb{R} \to \mathbb{R},\ x \mapsto p(x) = \sum_{i=0}^{n} a_i x^i$, eine Polynomfunktion vom Grad n und ist $a_n > 0$, so haben wir

$$\lim_{x\to\infty} p(x) = \infty$$

sowie

$$\lim_{x\to-\infty} p(x) = \begin{cases} \infty, & \text{falls } n \text{ gerade,} \\ -\infty, & \text{falls } n \text{ ungerade.} \end{cases}$$

Ist $a_n < 0$, so gilt die analoge Aussage. Siehe Aufgaben. □

9.3.5 Definition (Einseitige Grenzwerte)

Sind $X \subset \mathbb{R}$ und $x_0 \in \mathbb{R}$, so setzen wir

$$X_{x_0}^{+} := \{x \in X \mid x > x_0\} \quad \text{und} \quad X_{x_0}^{-} := \{x \in X \mid x < x_0\}.$$

Ist $f : X \to \mathbb{R}$ eine Funktion und gilt $x_0 \in \overline{X_{x_0}^{+}}$, so setzen wir

$$\lim_{x \to x_0^{+}} f(x) = a, \text{ mit } a \in \mathbb{R} \text{ oder } a = \pm\infty,$$

falls $\forall$ Folge $(x_n)_{n\in\mathbb{N}} \subset X_{x_0}^{+}$ gilt:

$$x_n \to x_0 \Rightarrow f(x_n) \to a.$$

Entsprechend definiert man

$$\lim_{x \to x_0^{-}} f(x) = a.$$

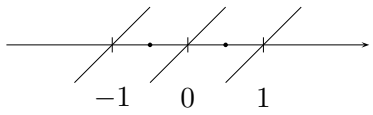
Falls existent heißt a **rechts-** bzw. **linksseitiger Grenzwert** von f in x_0 (im Falle $a = \pm\infty$ sind die Grenzwerte uneigentlich). □

9.3.6 Bemerkung

Ist $x_0 \in X$ und gilt sowohl $x_0 \in \overline{X_{x_0}^{+}}$ als auch $x_0 \in \overline{X_{x_0}^{-}}$, so ist f genau dann stetig in x_0, wenn rechts- und linksseitiger Grenzwert von f in x_0 existieren und mit $f(x_0)$ übereinstimmen. □

9.3.7 Beispiel

Die **Sägezahnkurve**



ist der Graph der Funktion

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x - n & \text{für } n - \frac{1}{2} < x < n + \frac{1}{2},\ n \in \mathbb{Z}, \\ 0 & \text{für } x = n + \frac{1}{2}, \quad n \in \mathbb{Z}. \end{cases}$$

Diese Funktion ist unstetig in den Punkten $n + \frac{1}{2}, n \in \mathbb{Z}$ und stetig sonst. Es gilt

$$\lim_{x \to n+\frac{1}{2}^{-}} f(x) = \frac{1}{2}, \qquad f\left(n + \frac{1}{2}\right) = 0, \qquad \lim_{x \to n+\frac{1}{2}^{+}} f(x) = -\frac{1}{2}.$$

□

9.4 Der Zwischenwertsatz

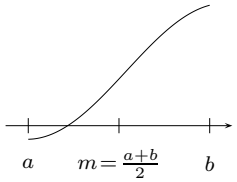
Wir zeigen, dass stetige Funktionen die in der Motivation 9.1 geforderte Eigenschaft haben.

9.4.1 Nullstellensatz

Ist $f : [a, b] \to \mathbb{R}$ eine stetige Funktion auf dem abgeschlossenen Intervall $[a, b]$ und gilt $f(a) \cdot f(b) < 0$, so hat f in dem offenen Teil $]a, b[$ mindestens eine Nullstelle.

Beweis

Die Voraussetzung bedeutet, dass $f(a)$ und $f(b)$ verschiedene Vorzeichen haben. Wir betrachten den Fall $f(a) < 0$ und $f(b) > 0$ (im anderen Fall ersetzen wir f durch $-f$). Wir verwenden das **Intervallhalbierungsverfahren** (oder **Bisektionsverfahren**), das wir im Prinzip schon aus den Beweisen der Sätze 7.4.2 und 7.7.3 kennen.



Zu Beginn setzen wir $a_0 = a$, $b_0 = b$ und $m := \frac{a+b}{2}$. Ist $f(m) = 0$, so sind wir fertig. Ist $f(m) > 0$ (bzw. $f(m) < 0$), so betrachten wir als nächstes $[a_1, b_1] = [a, m]$ (bzw. $[a_1, b_1] = [m, b]$). Fahren wir so fort, so erhalten wir entweder nach endlich vielen Schritten eine Nullstelle, oder wir bekommen rekursiv eine **Intervallschachtelung**

$$[a, b] = [a_0, b_0] \supset [a_1, b_1] \supset \ldots \supset [a_n, b_n] \supset \ldots$$

mit folgenden Eigenschaften:

(i) $f(a_n) < 0 < f(b_n) \quad \forall n \in \mathbb{N}_0$.

(ii) $b_n - a_n = \frac{b-a}{2^n} \quad \forall n \in \mathbb{N}_0$.

(iii) Die Folgen $(a_n)_{n\in\mathbb{N}_0}$ bzw. $(b_n)_{n\in\mathbb{N}_0}$ sind monoton wachsend bzw. monoton fallend und beschränkt.

Wegen (iii) sind $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ konvergent. Mit (ii) folgt dann:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n =: x_0.$$

Da f stetig ist, folgt mit (i):

$$f(x_0) = \lim_{n\to\infty} f(a_n) \leqslant 0 \leqslant \lim_{n\to\infty} f(b_n) = f(x_0).$$

Also ist $f(x_0) = 0$. □

9.4.2 Bemerkung

(i) Der Beweis liefert ein konstruktives Verfahren zur näherungsweisen Berechnung der Nullstellen stetiger Funktionen.

(ii) Der Nullstellensatz liefert einen neuen Beweis für die Existenz der n-ten Wurzel (vergleiche 7.7.8, wo auch die Eindeutigkeit gezeigt wurde): Ist $a > 0$ eine reelle Zahl, so betrachten wir die Funktion $f : [0, a+1] \to \mathbb{R}$, $x \mapsto x^n - a$. Dann gilt $f(0) = -a < 0$ sowie $f(a+1) = (a+1)^n - a \geq 1 + na - a = 1 + (n-1)a > 0$ (Bernoullische Ungleichung). Also existiert ein $s \in]0, a+1[$ mit $s^n - a = 0$. □

9.4.3 Beispiel

Polynomfunktionen $p : \mathbb{R} \to \mathbb{R},\ x \mapsto p(x) = \sum_{i=0}^{n} a_i x^i$, sind stetig nach 9.2.8, (v).

(i) Ist der Grad n ungerade, so gilt nach 9.3.4, (iii) entweder $\lim_{x\to\infty} p(x) = \infty$ und $\lim_{x\to-\infty} p(x) = -\infty$ (oder $\lim_{x\to\infty} p(x) = -\infty$ und $\lim_{x\to-\infty} p(x) = \infty$). Insbesondere existieren reelle Zahlen a bzw. b mit $f(a) \cdot f(b) < 0$. Der Nullstellensatz zeigt, dass p dann mindestens eine reelle Nullstelle hat.

(ii) Ist der Grad hingegen gerade, so muss keine reelle Nullstelle vorliegen. Man betrachte etwa die Polynome $x^{2m} + 1,\ m \geq 1$. □

9.4.4 Zwischenwertsatz

Eine stetige Funktion $f : [a, b] \to \mathbb{R}$ auf dem abgeschlossenen Intervall $[a, b]$ nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an.

Beweis

Gilt $f(a) = f(b)$, so ist nichts zu zeigen. Gilt $f(a) < f(b)$ und ist $y_0 \in]f(a), f(b)[$, so wende man den Nullstellensatz auf die durch $g(x) := f(x) - y_0$ definierte Funktion an. Analog für den Fall $f(a) > f(b)$. □

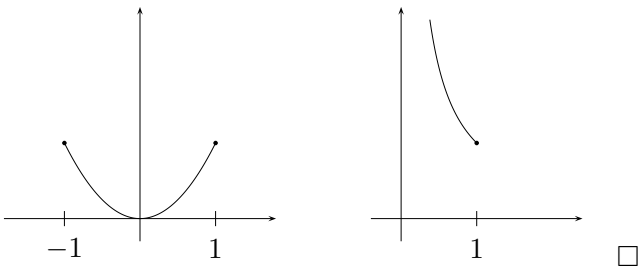
9.5 Maximum - Minimum - Eigenschaft stetiger Funktionen

Die in 7.7 eingeführten Begriffe für Teilmengen von $\mathbb{R}$ übertragen sich auf Funktionen $f : X \to \mathbb{R}$ durch Betrachten von $f(X)$. Wir sagen etwa, f ist **nach oben beschränkt**, wenn dies auf $f(X)$ zutrifft. In jedem Fall schreiben wir $\sup_{x\in X} f(x) = \sup f(X) \in \mathbb{R} \cup \{\infty\}$. Man beachte, dass $\sup_{x\in X} f(x)$ nicht notwendig zu $f(X)$ gehört: Eine Funktion braucht ihr Supremum nicht anzunehmen. Tut sie es doch, das heißt $\exists x_0 \in X$ mit $f(x_0) = \sup_{x\in X} f(x)$, so sagen wir, $\sup_{x\in X} f(x)$ ist das **Maximum** von f auf X, in Zeichen $\max_{x\in X} f(x) := f(x_0)$, und nennen x_0 eine **Maximalstelle**. Entsprechend führen wir etwa die Begriffe **Minimum** und **Minimalstelle** ein.

9.5.1 Beispiel

(i) Die Funktion $f : [-1, 1] \to \mathbb{R},\ x \mapsto x^2$, hat eine Minimalstelle in 0 und Maximalstellen in ± 1. Betrachten wir aber f auf dem offenen Intervall $]-1, 1[$, so gilt $\sup_{x\in]-1,1[} f(x) = 1$, aber ein Maximum liegt nicht vor.

(ii) Die Funktion $f :]0, 1] \to \mathbb{R},\ x \mapsto \frac{1}{x}$, hat eine Minimalstelle in 1, ist aber nach oben unbeschränkt: $\sup_{x\in]0,1]} f(x) = \infty$.



□

9.5.2 Satz

Eine stetige Funktion $f : [a, b] \to \mathbb{R}$ auf dem abgeschlossenen Intervall $[a, b]$ ist beschränkt und nimmt Maximum und Minimum an:

$$\exists\, \xi, \eta \in [a, b] \quad \text{mit} \quad f(\xi) \leqslant f(x) \leqslant f(\eta) \;\; \forall x \in [a, b]\,.$$

Beweis

Wir zeigen die Aussage für das Maximum (die Aussage für das Minimum folgt durch Betrachten von $-f$). Sei

$$s := \sup_{x \in [a,b]} f(x) \in \mathbb{R} \cup \{\infty\} \;\; .$$

Dann $\exists$ Folge $(x_n)_{n\in\mathbb{N}} \subset [a, b]$ mit

$$f(x_n) \underset{n\to\infty}{\longrightarrow} s$$

(vergleiche 7.7.3 und 7.7.4). Da $(x_n)_{n\in\mathbb{N}}$ beschränkt ist, existiert nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$. Sei η der Grenzwert:

$$x_{n_k} \underset{k\to\infty}{\longrightarrow} \eta \in [a, b] \;\; .$$

Die Stetigkeit von f liefert dann

$$f(\eta) = \lim_{k\to\infty} f(x_{n_k}) = s \;\; .$$

Es folgt $s \in \mathbb{R}$, also ist f nach oben beschränkt und η ist eine Maximalstelle. □

In der Situation von 9.5.2 zeigt der Zwischenwertsatz, dass das Bild $f([a, b]) = [f(\xi), f(\eta)]$ ebenfalls ein abgeschlossenes Intervall ist. Dieses Argument werden wir im nächsten Abschnitt wieder aufgreifen.

9.6 Der Umkehrsatz für streng monotone Funktionen

In Verallgemeinerung von 7.5.1 definieren wir:

9.6.1 Definition

Sei $X \subset \mathbb{R}$. Eine Funktion $f : X \to \mathbb{R}$ heißt **monoton wachsend** (bzw. **streng monoton wachsend, monoton fallend, streng monoton fallend**), wenn aus $x, x' \in X, x < x'$, stets folgt $f(x) \leqslant f(x')$ (bzw. $f(x) < f(x'), f(x) \geqslant f(x'), f(x) > f(x')$). □

Im folgenden betrachten wir ein beliebiges Intervall $I = [a, b]$ bzw. $I =]a, b[$ bzw. $I =]a, b]$ bzw. $I = [a, b[$, wobei wir, wenn sinnvoll, auch die Fälle $a = -\infty$ und $b = \infty$ zulassen. In jedem Fall nennen wir a und b die **Randpunkte** von I und sagen, I ist **links** bzw. **rechts abgeschlossen**, wenn a bzw. b eine reelle Zahl ist, die zu I gehört.

9.6.2 Satz (Umkehrsatz für streng monotone Funktionen)

Seien $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine streng monoton wachsende Funktion. Dann gilt:

(i) Die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ existiert. Sie ist streng monoton wachsend und stetig.

(ii) Ist f selbst stetig, so ist $f(I)$ ein Intervall mit den Randpunkten

$$A := \inf f(I) \quad \text{bzw.} \quad B := \sup f(I).$$

In diesem Fall ist $f(I)$ genau dann links bzw. rechts abgeschlossen, wenn dies für I gilt.

Die analoge Aussage gilt für streng monoton fallende Funktionen.

Beweis

Wir betrachten den Fall f streng monoton wachsend (andernfalls geht man zu $-f$ über).

(i) Da f streng monoton wächst, gilt $x < y \Leftrightarrow f(x) < f(y)$. Es folgt die Injektivität von f und damit die Existenz von $f^{-1} : f(I) \to \mathbb{R}$. Ebenso folgt, dass f^{-1} streng monoton wächst. Wir zeigen nun, dass f^{-1} in einem beliebig vorgegebenen Punkt $y_0 \in f(I)$ stetig ist.

Sei dazu $f^{-1}(y_0) = x_0$, also $y_0 = f(x_0)$. Wir behandeln den Fall, dass x_0 kein Randpunkt von I ist (und lassen die Fälle x_0 linker bzw. rechter Randpunkt als Übung). Dann $\exists r > 0$ mit $[x_0 - r, x_0 + r] \subset I$. Wir wenden das ε-δ–Kriterium der Stetigkeit an.

Sei $\varepsilon > 0$ vorgegeben. Wir können annehmen, dass $\varepsilon \leq r$. Dann gilt $[x_0 - \varepsilon, x_0 + \varepsilon] \subset I$. Da f streng monoton wächst, gilt $f(x_0 - \varepsilon) < f(x_0) = y_0 < f(x_0 + \varepsilon)$. Dann $\exists \delta > 0$ mit

$$f(x_0 - \varepsilon) < y_0 - \delta < y_0 + \delta < f(x_0 + \varepsilon).$$

Ist also $y \in f(I) \cap U_\delta(y_0)$, so gilt $f(x_0 - \varepsilon) < y < f(x_0 + \varepsilon)$. Da f^{-1} streng monoton wächst, folgt daraus $x_0 - \varepsilon < f^{-1}(y) < x_0 + \varepsilon$, d.h. $f^{-1}(y) \in U_\varepsilon(x_0)$. Also ist f^{-1} in y_0 stetig.

(ii) Sei nun f stetig. Da f streng monoton wächst, gilt $A = \inf f(I) < B = \sup f(I)$. Also gibt es zu jedem Punkt $y_0 \in]A, B[$ Punkte $x_1, x_2 \in I$ mit $f(x_1) < y_0 < f(x_2)$. Der Zwischenwertsatz angewandt auf die (stetige) Funktion $f|_{[x_1,x_2]}$ ergibt die Existenz eines Punktes $x_0 \in [x_1, x_2]$ mit $f(x_0) = y_0$. Also ist $]A, B[\subset f(I)$. Aufgrund der Definition von Infimum und Supremum folgt, dass $f(I)$ ein Intervall mit den Randpunkten A und B ist. Die zweite Aussage in (ii) lasse ich als Übung. □

Eine Funktion wie im zweiten Bild der Motivation 9.1 zeigt, dass die Voraussetzung stetig in (ii) notwendig ist.

9.6.3 Beispiel

(i) Ist $n \in \mathbb{N}$, so ist die **Potenzfunktion** $[0,\infty[\to \mathbb{R},\ x \mapsto x^n$, stetig nach 9.2.8, (v). Wie im Beweis von 7.7.8 gesehen, ist sie auch streng monoton wachsend (mit Bild $[0,\infty[$). Somit gilt gleiches für ihre Umkehrfunktion, die **Wurzelfunktion** $[0,\infty[\to \mathbb{R},\ x \mapsto \sqrt[n]{x}$.

(ii) Ist $f : X \to \mathbb{R}$ eine stetige Funktion mit $f(X) \subset [0,\infty[$, so ist nach (i) und 9.2.6 auch die Funktion

$$\sqrt[n]{f} : X \to [0,\infty[\,,\ x \mapsto \sqrt[n]{f(x)},$$

stetig. Zum Beispiel ist

$$\mathbb{R} \to [0,\infty[\,,\ x \mapsto \sqrt{x^2+1}$$

stetig. □

9.7 Exponentialfunktion und Logarithmus

Wir wollen zunächst zeigen, dass die in 8.8 eingeführte Exponentialfunktion stetig ist. Dazu benötigen wir:

9.7.1 Lemma

Es gilt:

$$|\exp(x) - 1| \leqslant 2 \cdot |x| \quad \forall x \in \mathbb{R} \text{ mit } |x| \leqslant 1.$$

Beweis

Es gilt $\frac{1}{k!} \leqslant \frac{1}{2^{k-1}}\ \forall k \in \mathbb{N}$ (vergleiche 7.5.4). Daraus folgt:

$$\begin{aligned} |\exp(x) - 1| &= \left|\sum_{k=0}^{\infty} \frac{x^k}{k!} - 1\right| = \left|\sum_{k=1}^{\infty} \frac{x^k}{k!}\right| \leqslant |x| \cdot \sum_{k=1}^{\infty} \frac{|x|^{k-1}}{k!} \\ &\leqslant |x| \cdot \sum_{k=1}^{\infty} \frac{|x|^{k-1}}{2^{k-1}} = |x| \cdot \sum_{k=0}^{\infty} \left(\frac{|x|}{2}\right)^k \underset{|x|\leqslant 1}{\leqslant} |x| \cdot \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \\ &= |x| \cdot \frac{1}{1-\frac{1}{2}} = 2 \cdot |x|\,. \end{aligned}$$

□

9.7.2 Satz

Die Funktion $\exp : \mathbb{R} \to \mathbb{R}$ ist stetig.

Beweis

(i) Die Stetigkeit im Nullpunkt ergibt sich aus dem Lemma:

$$x_n \to 0 \Rightarrow |\exp(x_n) - \exp(0)| = |\exp(x_n) - 1| \leqslant 2 \cdot |x_n| \to 0$$

(ii) Sei nun $x_0 \in \mathbb{R}$ beliebig und $(x_n)_{n\in\mathbb{N}}$ eine Folge mit $x_n \to x_0$. Dann gilt: $x_n - x_0 \to 0$ und somit $\exp(x_n - x_0) \to 1$ nach (i). Mit der Funktionalgleichung von exp folgt (siehe 8.8.1 und 8.8.2, (ii)):

$$\exp(x_n) = \frac{\exp(x_0)}{\exp(x_0)} \cdot \exp(x_n) = \exp(x_0) \cdot \exp(x_n - x_0) \to \exp(x_0)\,.$$

□

9.7.3 Satz

Die Funktion $\exp : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend und hat den Wertebereich $]0, \infty[$.

Beweis

(i) Ist $x > 0$, so gilt:

$$\exp(x) = 1 + x + \frac{x^2}{2} + \ldots > 1.$$

Seien nun $x_1, x_2 \in \mathbb{R}$ mit $x_1 < x_2$. Dann ist $x_2 - x_1 > 0$ und es folgt:

$$\exp(x_2) = \exp(x_1 + (x_2 - x_1)) = \exp(x_1) \cdot \exp(x_2 - x_1) > \exp(x_1).$$

Also ist exp streng monoton wachsend.

(ii) Für alle $n \in \mathbb{N}$ gilt:

$$\exp(n) = 1 + n + \frac{n^2}{2} + \ldots \geqslant 1 + n$$

und

$$\exp(-n) = \frac{1}{\exp(n)} \leqslant \frac{1}{1+n}.$$

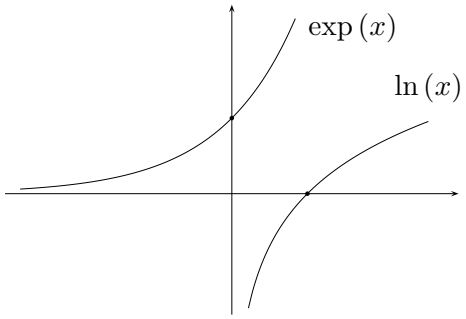
Daraus folgt

$$\lim_{n\to\infty} \exp(n) = \infty \quad \text{und} \quad \lim_{n\to\infty} \exp(-n) = 0.$$

Andererseits gilt $\exp(x) > 0 \; \forall x \in \mathbb{R}$ (siehe 8.8.2, (i)) und nach (i) und dem Umkehrsatz 9.6.2 ist der Wertebereich $\exp(\mathbb{R})$ ein Intervall. Insgesamt folgt $\exp(\mathbb{R}) =]0, \infty[$. □

9.7.4 Bemerkung und Definition

Nach 9.7.3 und dem Umkehrsatz 9.6.2 hat $\exp : \mathbb{R} \to]0, \infty[$ eine Umkehrfunktion $]0, \infty[\to \mathbb{R}$, die ebenfalls streng monoton wachsend und stetig ist.



Die Umkehrfunktion heißt der **natürliche Logarithmus** und wird mit ln bezeichnet. Wegen $\exp(0) = 1$ gilt $\ln(1) = 0$. Weitere Rechenregeln finden sich in 9.7.8. □

Wir führen jetzt allgemein die Exponentialfunktion und die Logarithmusfunktion zur Basis a ein (bisher haben wir den Fall $a = \exp(1) = e$ betrachtet).

9.7.5 Definition

Für $a \in \mathbb{R}, a > 0$, ist die **Exponentialfunktion zur Basis** a definiert durch

$$\exp_a : \mathbb{R} \to \mathbb{R}, \; x \mapsto \exp(x \cdot \ln(a)).$$

□

9.7.6 Satz

Die Funktion $\exp_a : \mathbb{R} \to \mathbb{R}$ ist stetig und es gilt:

(i) $\exp_a(x+y) = \exp_a(x) \cdot \exp_a(y) \quad \forall x, y \in \mathbb{R}.$

(ii) $\exp_a(n) = a^n \quad \forall n \in \mathbb{Z}.$

(iii) $\exp_a\left(\frac{p}{q}\right) = \sqrt[q]{a^p} \quad \forall p \in \mathbb{Z} \text{ und } \forall\, q \in \mathbb{N}, q \geqslant 2.$

Beweis

Als Komposition stetiger Funktionen ist $\exp_a$ stetig. Die Funktionalgleichung (i) ergibt sich unmittelbar aus der Funktionalgleichung von exp. Insbesondere gilt

$$\exp_a(-x) = \frac{1}{\exp_a(x)}$$

(setze $y = -x$ in (i)). Weiter folgt aus (i) durch vollständige Induktion:

$$\exp_a(nx) = \exp_a(x)^n \quad \forall n \in \mathbb{N},\ \forall x \in \mathbb{R}.$$

Wegen $\exp_a(1) = \exp(\ln(a)) = a$ und $\exp_a(-1) = \frac{1}{a}$ folgt daraus mit $x = \pm 1$:

$$\exp_a(n) = a^n \ \text{ und } \ \exp_a(-n) = \frac{1}{a^n} \quad \forall n \in \mathbb{N}.$$

Dies zeigt (ii). Schließlich gilt (iii) wegen

$$a^p = \exp_a(p) = \exp_a(q \cdot \frac{p}{q}) = \left(\exp_a(\frac{p}{q})\right)^q.$$

□

Satz 9.7.6 rechtfertigt die Schreibweise

$$a^x := \exp_a(x) = \exp(x \cdot \ln(a)) \ .$$

Die Funktionalgleichung aus 9.7.6, (i) schreibt sich jetzt so:

$$a^{x+y} = a^x \cdot a^y \quad \forall\ a > 0 \text{ und } \forall\ x, y \in \mathbb{R}.$$

9.7.7 Satz (Rechenregeln)

Für alle $a, b \in]0, \infty[$ und $x, y \in \mathbb{R}$ gilt:

(i) $a^x \cdot a^y = a^{x+y}, \quad (a^x)^y = a^{x \cdot y}.$

(ii) $a^x \cdot b^x = (a \cdot b)^x.$

(iii) $\left(\frac{1}{a}\right)^x = a^{-x}.$

Beweis

Wir zeigen die zweite Formel in (i): Wegen $a^x = \exp(x \cdot \ln(a))$ gilt $\ln(a^x) = x \cdot \ln(a)$ und somit

$$(a^x)^y = \exp(y \cdot \ln(a^x)) = \exp(y \cdot x \cdot \ln(a)) = a^{x \cdot y}.$$

□

9.7.8 Satz und Definition

Die Funktion

$$\mathbb{R} \to \mathbb{R},\ x \mapsto a^x,$$

ist für $a > 1$ streng monoton wachsend und für $0 < a < 1$ streng monoton fallend. In beiden Fällen wird $\mathbb{R}$ bijektiv auf $]0, \infty[$ abgebildet. Die Umkehrfunktion

$${}^a\log :]0, \infty[\to \mathbb{R}$$

ist stetig und es gilt:

$${}^a\log(x) = \frac{\ln(x)}{\ln(a)} \quad \forall x \in]0, \infty[\,.$$

Die Funktion ${}^a\log$ heißt **Logarithmus zur Basis** a. Es gelten die folgenden **Rechenregeln**:

(i) ${}^a\log(x \cdot y) = {}^a\log(x) + {}^a\log(y) \quad \forall x, y \in]0, \infty[.$

(ii) ${}^a\log\left(\frac{x}{y}\right) = {}^a\log(x) - {}^a\log(y) \quad \forall x, y \in]0, \infty[.$

(iii) $\lambda \cdot {}^a\log(x) = {}^a\log\left(x^\lambda\right) \quad \forall \lambda \in \mathbb{R},\ \forall x \in]0, \infty[\,.$

Beweis

Die Monotonie von $\exp_a$ und das Wachstumsverhalten für $x \to \pm\infty$ ergeben sich aus den entsprechenden Eigenschaften von exp (man unterscheide die Fälle $a > 1$ und $0 < a < 1$, d.h. die Fälle $\ln(a) > 0$ und $\ln(a) < 0$). Der Umkehrsatz 9.6.2 liefert dann wieder die Existenz und Stetigkeit von ${}^a\log$. Die Formel ${}^a\log(x) = \frac{\ln(x)}{\ln(a)}$ folgt durch Vergleich der Exponenten in den Gleichungen

$$x = e^{\ln(x)} \ \text{ und } \ x = a^{{}^a\log(x)} = e^{{}^a\log(x)\cdot\ln(a)}.$$

Die Funktionalgleichung (i) ergibt sich aus der Funktionalgleichung von $\exp_a$:

$$x \cdot y = \exp_a({}^a\log(x)) \cdot \exp_a({}^a\log(y)) = \exp_a({}^a\log(x) + {}^a\log(y))\,.$$

(ii) und (iii) lasse ich als Übung. □

Wir berechnen einige wichtige Grenzwerte:

9.7.9 Satz

(i) $\forall n \in \mathbb{N}_0$ gilt:

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty.$$

Man sagt auch, e^x wächst für $x \to \infty$ schneller gegen ∞ als jede Potenz von x.

(ii) $\forall n \in \mathbb{N}_0$ gilt:

$$\lim_{x\to\infty} x^n \cdot e^{-x} = 0 \quad \text{und} \quad \lim_{x\to 0^+} x^n \cdot e^{\frac{1}{x}} = \infty.$$

(iii)

$$\lim_{x\to\infty} \ln(x) = \infty \quad \text{und} \quad \lim_{x\to 0^+} \ln(x) = -\infty.$$

(iv) $\forall \lambda \in \mathbb{R}, \lambda > 0$, gilt:

$$\lim_{x \to 0^+} x^\lambda = 0 \quad \text{und} \quad \lim_{x \to 0^+} x^{-\lambda} = \infty.$$

(v) $\forall \lambda \in \mathbb{R}, \lambda > 0$, gilt:

$$\lim_{x \to \infty} \frac{\ln(x)}{x^\lambda} = 0.$$

Beweis

(i) $\forall x > 0$ und $\forall n \in \mathbb{N}_0$ gilt

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} > \frac{x^{n+1}}{(n+1)!},$$

also auch

$$\frac{e^x}{x^n} > \frac{x}{(n+1)!}.$$

Daraus folgt die Behauptung.

(ii) folgt aus (i) wegen $x^n \cdot e^{-x} = \left(\frac{e^x}{x^n}\right)^{-1}$ bzw. mit $y = \frac{1}{x}$ wegen $\lim_{x \to 0^+} x^n \cdot e^{\frac{1}{x}} = \lim_{y \to \infty} \frac{e^y}{y^n} = \infty$.

(iii) Sei $K \in \mathbb{R}$ beliebig vorgegeben. Da ln streng monoton wachsend ist, gilt: $\ln(x) > K \ \forall x > e^K$. Dies zeigt $\lim_{x \to \infty} \ln(x) = \infty$. Die zweite Behauptung folgt daraus mit $y = \frac{1}{x}$ wegen $\lim_{x \to 0^+} \ln(x) = \lim_{y \to \infty} \ln\left(\frac{1}{y}\right) = -\lim_{y \to \infty} \ln(y) = -\infty$.

(iv) und (v) ergeben sich analog. □

9.7.10 Bemerkung

Die Potenzfunktion

$$]0, \infty[\to \mathbb{R}, \ x \mapsto x^\lambda,$$

ist jetzt also für jedes beliebige aber feste $\lambda \in \mathbb{R}$ durch die Formel

$$x^\lambda = \exp(\lambda \cdot \ln(x)) \qquad \text{für } x > 0$$

erklärt. Sie ist streng monoton wachsend falls $\lambda > 0$ bzw. streng monoton fallend falls $\lambda < 0$. Wegen 9.7.9, (iv) lässt sie sich im Falle $\lambda > 0$ durch $0^\lambda := 0$ stetig in 0 fortsetzen. □

Kapitel 10

Komplexe Zahlen und trigonometrische Funktionen

10.1 Motivation

Wir wollen die trigonometrischen Funktionen Cosinus und Sinus mit Hilfe der komplexen Exponentialfunktion einführen. Deswegen behandeln wir zunächst die komplexen Zahlen. □

10.2 Der Körper der komplexen Zahlen

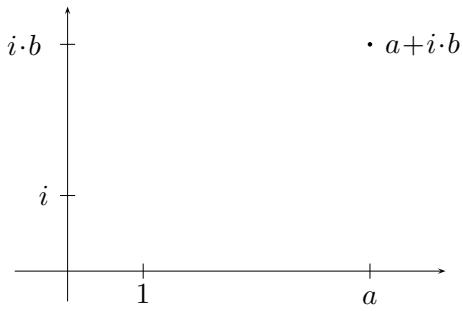
Beim Aufbau des Zahlsystems beginnt man mit den natürlichen Zahlen, die man zum Abzählen braucht. Die Subtraktion von natürlichen Zahlen führt aus diesen heraus zu den ganzen Zahlen. Die Division von ganzen Zahlen führt aus diesen heraus zu den rationalen Zahlen. Nun hat man das Problem, dass nicht jede Cauchy-Folge rationaler Zahlen gegen eine rationale Zahl konvergiert. Nimmt man die Grenzwerte hinzu, so gelangt man zu den reellen Zahlen. Zu den komplexen Zahlen wird man schließlich geleitet, wenn man nach Nullstellen von Polynom(funktion)en fragt. So hat etwa das Polynom x^2+1 keine reelle Nullstelle. Erweitert man die reellen Zahlen zu den komplexen Zahlen, so wollen wir mit den neuen Zahlen wie üblich rechnen können. Nehmen wir also eine Zahl i mit $i^2+1=0$ hinzu, so sollte für $b \in \mathbb{R}$ auch die Zahl $i \cdot b$ erklärt sein. Und ist $a \in \mathbb{R}$ eine weitere Zahl, so sollte auch $a + i \cdot b$ erklärt sein. Legt man die üblichen Rechengesetze zu Grunde, so sollte gelten

$$(a + i \cdot b) + (c + i \cdot d) = (a + c) + i \cdot (b + d)$$

und

$$\begin{aligned}(a + i \cdot b) \cdot (c + i \cdot d) &= a \cdot c + i \cdot a \cdot d + i \cdot b \cdot c + i^2 \cdot b \cdot d \\ &= (a \cdot c - b \cdot d) + i \cdot (a \cdot d + b \cdot c)\,.\end{aligned}$$

Zahlen der Form $i \cdot b$ heißen **imaginäre Zahlen**, wir können sie uns auf der **imaginären Achse** angeordnet vorstellen. Identifizieren wir i mit dem Punkt $(0,1) \in \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, so kann jede komplexe Zahl als Punkt in der Ebene gedeutet werden. Wir sprechen auch von der **komplexen Zahlebene**.



Wir identifizieren also die reelle Zahl a mit dem Paar $(a, 0)$, die imaginäre Zahl $i \cdot b$ mit dem Punkt $(0, b)$ und die komplexe Zahl $a + i \cdot b$ mit dem Paar (a, b). Dies führt zu den folgenden Definitionen.

10.2.1 Satz und Definition

Auf der Menge

$$\mathbb{C} := \{z = (a, b) \mid a, b \in \mathbb{R}\}$$

erklären wir eine Addition und Multiplikation durch

$$(a, b) + (c, d) = (a + c, b + d)$$

und

$$(a, b) \cdot (c, d) = (a \cdot c - b \cdot d, a \cdot d + b \cdot c)\,.$$

Dann ist $(\mathbb{C}, +, \cdot)$ ein Körper, der **Körper der komplexen Zahlen**. Dabei ist $0 = (0, 0)$ das Nullelement und $1 = (1, 0)$ das Einselement. Das zu $z = (a, b)$ bzgl. $+$ inverse Element ist $-z = (-a, -b)$, das zu $z = (a, b) \neq 0$ bzgl. $\cdot$ inverse Element ist $z^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$.

Beweis

Die Körpergesetze von $\mathbb{C}$ ergeben sich aus denen von $\mathbb{R}$ durch Nachrechnen. Zum Beispiel gilt für $z = (a, b) \neq 0$:

$$\begin{aligned}&(a, b) \cdot \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right)\\ =&\left(a \cdot \frac{a}{a^2 + b^2} - b \cdot \frac{-b}{a^2 + b^2}, a \cdot \frac{-b}{a^2 + b^2} + b \cdot \frac{a}{a^2 + b^2}\right) = (1, 0)\,.\end{aligned}$$

□

10.2.2 Bemerkung und Definition

(i) Wir fassen $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auf, indem wir $a \in \mathbb{R}$ mit $(a, 0) \in \mathbb{C}$ identifizieren. Dies ist verträglich mit Addition und Multiplikation.

$$(a, 0) + (b, 0) = (a + b, 0)\,, \quad (a, 0) \cdot (b, 0) = (a \cdot b, 0)\,.$$

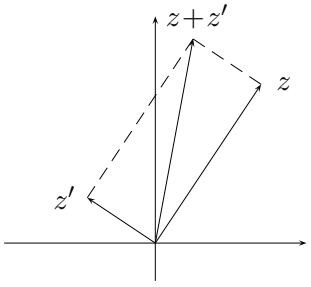
(ii) Setzt man $i = (0, 1)$, so erhält man die gebräuchliche Schreibweise für die komplexen Zahlen:

$$(a, b) = (a, 0) + (0, 1) \cdot (b, 0) = a + i \cdot b.$$

Es gilt

$$i^2 = (0,1) \cdot (0,1) = (-1,0) = -1.$$

(iii) Veranschaulicht man die komplexen Zahlen in der komplexen Zahlebene, so ist die Addition gerade die übliche Vektoraddition:



(iv) **Subtraktion** bzw. **Division** sind erklärt durch:

$$(a + i \cdot b) - (c + i \cdot d) := (a + i \cdot b) + (-(c + i \cdot d)) = (a - c) + i \cdot (b - d)$$

bzw.

$$\frac{a + i \cdot b}{c + i \cdot d} := (a + i \cdot b) \cdot (c + i \cdot d)^{-1} = \frac{ac + bd}{c^2 + d^2} + i \cdot \frac{bc - ad}{c^2 + d^2}, \quad \text{falls } (c,d) \neq 0.$$

(v) Für $z = a + i \cdot b \in \mathbb{C}$ setzen wir

$$\operatorname{Re} z := a \qquad \textbf{(Realteil)}$$

und

$$\operatorname{Im} z := b. \qquad \textbf{(Imaginärteil)}$$

Zwei komplexe Zahlen z und z' sind genau dann gleich, wenn gilt $\operatorname{Re} z = \operatorname{Re} z'$ und $\operatorname{Im} z = \operatorname{Im} z'$. □

10.3 Konjugiert komplexe Zahlen und Betrag

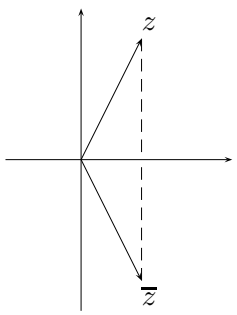
10.3.1 Bemerkung und Definition

Ist $z = a + ib \in \mathbb{C}$, so heißt $\overline{z} := a - ib \in \mathbb{C}$ die zu z **konjugiert komplexe Zahl**. Es gilt

$$\operatorname{Re} z = \frac{1}{2}(z + \overline{z}), \quad \operatorname{Im} z = \frac{1}{2i}(z - \overline{z}).$$

□

Anschaulich erhält man $\overline{z}$ aus z durch Spiegelung an der reellen Achse:



10.3.2 Satz (Rechenregeln)

Für alle $z, z' \in \mathbb{C}$ gilt:

(i) $\overline{\overline{z}} = z$,

(ii) $\overline{z \pm z'} = \overline{z} \pm \overline{z'}$,

(iii) $\overline{z \cdot z'} = \overline{z} \cdot \overline{z'}$,

(iv) $\overline{\left(\frac{z}{z'}\right)} = \frac{\overline{z}}{\overline{z'}}$, falls $z' \neq 0$,

(v) $\overline{z^n} = (\overline{z})^n \quad \forall n \in \mathbb{N}_0$.

Beweis

Wir zeigen (iii): Für $z = a + ib$ und $z' = a' + ib'$ gilt:

$$\begin{aligned}
\overline{z \cdot z'} &= \overline{(a+ib)\cdot(a'+ib')} \\
&= \overline{(aa' - bb') + i\cdot(ab' + ba')} \\
&= \big(aa' - bb'\big) - i\cdot\big(ab' + ba'\big) \\
&= (a - ib)\cdot\big(a' - ib'\big) \\
&= \overline{z}\cdot\overline{z'}.
\end{aligned}$$

□

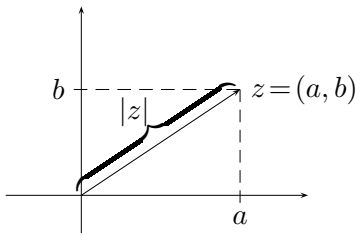
10.3.3 Bemerkung und Definition

Ist $z = a + i \cdot b \in \mathbb{C}$, so heißt $|z| = \sqrt{a^2 + b^2} \in \mathbb{R}$ der **Betrag** von z. Es gilt:

$$|z| = \sqrt{z \cdot \overline{z}}.$$

□

Anschaulich ist $|z|$ der Abstand von (a, b) zum Nullpunkt:



10.3.4 Satz (Rechenregeln)

Es gelten wortwörtlich die gleichen Rechenregeln wie in 6.3.4.

Beweis

Nachrechnen. □

10.4 Folgen und Reihen, Stetigkeit

Unendliche Folgen und Reihen von komplexen Zahlen werden analog zu Folgen und Reihen reeller Zahlen erklärt. Gleiches gilt für die Konvergenz und die dazu eingeführten Bezeichnungen. So gilt etwa für Folgen in $\mathbb{C}$:

$$c_n \to c \Leftrightarrow \forall\, \varepsilon > 0\ \exists\, n_0(\varepsilon) \in \mathbb{N} : |c_n - c| < \varepsilon \quad \forall n \geqslant n_0(\varepsilon).$$

Das Cauchy-Kriterium und die Rechenregeln für Folgen und Reihen übertragen sich wortwörtlich. Gleiches gilt für die absolute Konvergenz von Reihen, das Majoranten-, Wurzel- und Quotientenkriterium. Lediglich Definitionen und Sätze, die Ordnungseigenschaften enthalten (wie zum Beispiel das Monotoniekriterium) lassen sich nicht ins Komplexe übertragen, da für komplexe Zahlen keine Beziehungen $>$ oder $<$ eingeführt sind.

Die Definition der Stetigkeit überträgt sich: Ist $Z \subset \mathbb{C}$, so heißt eine Funktion $f : Z \to \mathbb{C}$ stetig in $z_0 \in Z$, wenn gilt:

$$(z_n)_{n\in\mathbb{N}} \subset Z,\ z_n \to z_0 \Rightarrow f(z_n) \to f(z_0).$$

Gleiches gilt für das Rechnen mit stetigen Funktionen, außer für den Fall, dass Ordnungseigenschaften im Spiel sind (Zwischenwertsatz). Wir notieren noch:

10.4.1 Satz

Sei $c_n = a_n + ib_n, n \in \mathbb{N}$, eine Folge komplexer Zahlen. Dann konvergiert $(c_n)_{n\in\mathbb{N}}$ genau dann, wenn $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergieren. In diesem Fall gilt:

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} a_n + i \lim_{n\to\infty} b_n.$$

Insbesondere konvergiert dann auch die Folge $(\overline{c_n})_{n\in\mathbb{N}}$ und es gilt:

$$\lim_{n\to\infty} \overline{c_n} = \overline{\lim_{n\to\infty} c_n}.$$

Beweis

Gilt $c_n \to c = a + ib$ und ist $\varepsilon > 0$ vorgegeben, so $\exists n_0 \in \mathbb{N}$ mit $|c_n - c| < \varepsilon\ \forall n \geqslant n_0$. Dann gilt aber auch:

$$|a_n - a| = |\mathrm{Re}\,(c_n - c)| \leq |c_n - c| < \varepsilon,$$

$$|b_n - b| = |\mathrm{Im}\,(c_n - c)| \leq |c_n - c| < \varepsilon.$$

Es folgt $a_n \to a$ und $b_n \to b$.

Ist umgekehrt $a_n \to a$ und $b_n \to b$ vorausgesetzt und ist $\varepsilon > 0$ vorgegeben, so $\exists n_a, n_b \in \mathbb{N}$ mit $|a_n - a| < \frac{\varepsilon}{2}\ \forall n \geqslant n_a$ und $|b_n - b| < \frac{\varepsilon}{2}\ \forall n \geqslant n_b$. Mit $c := a + ib$ und $n_0 := \max(n_a, n_b)$ gilt dann $\forall n \geq n_0$:

$$|c_n - c| = |(a_n - a) + i(b_n - b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Es folgt $c_n \to c$.

Die zweite Aussage des Satzes folgt wegen $\mathrm{Re}\,(\overline{c_n}) = \mathrm{Re}\,(c_n)$ und $\mathrm{Im}\,(\overline{c_n}) = -\mathrm{Im}\,(c_n)$. $\square$

10.5 Die komplexe Exponentialfunktion

Wie im Reellen erhält man die Absolutkonvergenz der komplexen Exponentialreihe

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

Auch hier gilt die Funktionalgleichung

$$\exp(w+z) = \exp(w) \cdot \exp(z).$$

Wie im Reellen schreiben wir $e^z := \exp(z)$.

10.5.1 Satz

Für alle $z \in \mathbb{C}$ gilt:

(i) $\exp(z) \neq 0$.

(ii) $\exp(\overline{z}) = \overline{\exp(z)}$.

Für alle $x \in \mathbb{R}$ gilt:

(iii) $|\exp(ix)| = 1$.

Beweis

(i) Die Funktionalgleichung liefert

$$\exp(z) \cdot \exp(-z) = \exp(z-z) = \exp(0) = 1.$$

(ii) folgt aus den Rechenregeln 10.3.2 und 10.4.1:

$$\sum_{k=0}^{n} \frac{\overline{z}^k}{k!} = \overline{\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)} \to \overline{\left(\sum_{k=0}^{\infty} \frac{z^k}{k!}\right)}.$$

(iii) folgt aus 10.3.3 und (ii) mit der Funktionalgleichung:

$$\begin{aligned} |\exp(ix)|^2 &= \exp(ix) \cdot \overline{\exp(ix)} \\ &= \exp(ix) \cdot \exp(-ix) = \exp(0) = 1. \end{aligned}$$

□

Wie im Reellen folgt die Stetigkeit der Exponentialfunktion aus der Abschätzung

$$|\exp(z) - 1| \leqslant 2 \cdot |z| \quad \forall z \in \mathbb{C} \text{ mit } |z| \leqslant 1.$$

Allgemein schreiben wir

$$r_{n+1}(z) := \exp(z) - \sum_{k=0}^{n} \frac{z^k}{k!} \qquad \textbf{(Restglied)}$$

und erhalten mit dem zu 9.7.1 analogen Beweis:

10.5.2 Satz (Restgliedabschätzung)

Es gilt

$$|r_{n+1}(z)| \leqslant 2 \cdot \frac{|z|^{n+1}}{(n+1)!} \quad \forall z \in \mathbb{C} \text{ mit } |z| \leqslant 1 + \frac{n}{2}.$$

□

Restgliedabschätzungen spielen eine wichtige Rolle bei der näherungsweisen Berechnung von Summen von Reihen.

10.5.3 Beispiel

Es gilt $|r_{16}(1)| \leqslant \frac{2}{16!} < 10^{-13}$. Also liefert die Berechnung von $\sum_{k=0}^{15} \frac{1}{k!}$ die Eulersche Zahl $e = \sum_{k=0}^{\infty} \frac{1}{k!}$ auf 12 Stellen nach dem Komma genau:

$$e = 2,718\,281\,828\,459\ldots \ .$$

□

10.6 Das Horner-Schema

Wie wertet man polynomiale Ausdrücke wie $\sum_{k=0}^{15} \frac{1}{k!}$ in 10.5.3 aus?

10.6.1 Beispiel

Wir werten

$$p(z) = 3z^5 - 4z^4 + 2z^2 - 7z - 16$$

in $z_0 = 2$ aus. Berechnen wir $p(2)$ direkt, so müssten wir 12 Multiplikationen und 4 Additionen durchführen. Schreiben wir dagegen

$$p(z) = ((((3z - 4) \cdot z + 0) \cdot z + 2) \cdot z - 7) \ \ z - 16,$$

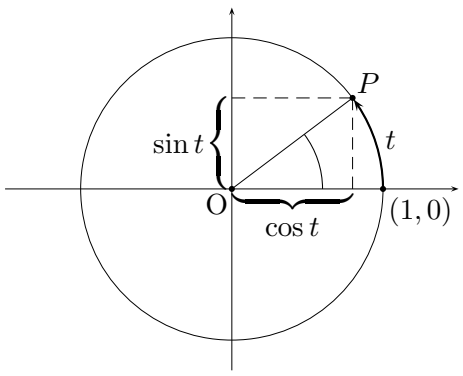
so kommen wir mit 5 Multiplikationen und 5 Additionen aus. □

Obiges Verfahren funktioniert allgemein und ist unter dem Namen **Horner-Schema** bekannt.

10.7 Die trigonometrischen Funktionen Sinus und Cosinus

10.7.1 Motivation

Aus der Schule sind uns die Funktionen Cosinus und Sinus über die folgende geometrische Beziehung bekannt:



Dabei bezeichnet t das Bogenmaß des Winkels zwischen der reellen Achse und der Geraden $\overline{OP}$. Gibt man sich die Mühe, die Bogenlänge mathematisch exakt zu definieren, so kann man zeigen, dass gilt $P = e^{it}$ (vergleiche auch 10.5.1, (iii)). Dies führt zur folgenden Definition. □

10.7.2 Definition

Die reellen Funktionen **Cosinus** und **Sinus** sind definiert durch

$$\cos : \mathbb{R} \to \mathbb{R},\ x \mapsto \cos x := \operatorname{Re}\left(e^{ix}\right)$$

bzw.

$$\sin : \mathbb{R} \to \mathbb{R},\ x \mapsto \sin x := \operatorname{Im}\left(e^{ix}\right). \qquad \square$$

10.7.3 Bemerkung

Per Definition gilt die **Eulersche Formel**

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

Es folgt

(i) $\cos 0 = 1, \qquad \sin 0 = 0,$

(ii) $\cos(-x) = \cos x, \qquad \sin(-x) = -\sin x \qquad \forall x \in \mathbb{R},$

(iii) $\cos^2 x + \sin^2 x = 1 \qquad \forall x \in \mathbb{R},$

(iv) $|\cos x| \leqslant 1, \qquad |\sin x| \leqslant 1 \qquad \forall x \in \mathbb{R}.$

Beweis

(i) gilt wegen $e^{i0} = e^0 = 1$,

(ii) gilt wegen $\overline{e^{ix}} = e^{-ix}$,

(iii) bedeutet gerade $\left|e^{ix}\right|^2 = 1$ und

(iv) folgt aus (iii). □

10.7.4 Satz

Die Funktionen cos und sin sind auf ganz $\mathbb{R}$ stetig.

Beweis

Sind $x_0 \in \mathbb{R}$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen mit $x_n \to x_0$, so gilt auch $ix_n \to ix_0$ nach 10.4.1. Da exp auf ganz $\mathbb{C}$ stetig ist, folgt $e^{ix_n} \to e^{ix_0}$. Daraus ergibt sich wieder mit 10.4.1:

$$\cos x_n = \operatorname{Re}\left(e^{ix_n}\right) \to \operatorname{Re}\left(e^{ix_0}\right) = \cos x_0$$

$$\sin x_n = \operatorname{Im}\left(e^{ix_n}\right) \to \operatorname{Im}\left(e^{ix_0}\right) = \sin x_0. \qquad \square$$

10.7.5 Satz (Additionstheoreme)

Für alle $x, y \in \mathbb{R}$ gilt:

$$\cos(x+y) = \cos x \cdot \cos y - \sin x \cdot \sin y,$$
$$\sin(x+y) = \sin x \cdot \cos y + \cos x \cdot \sin y.$$

Beweis

Die Funktionalgleichung von exp liefert

$$e^{i(x+y)} = e^{ix+iy} = e^{ix} \cdot e^{iy}.$$

Daraus ergibt sich mit der Eulerschen Formel:

$$\begin{aligned}\cos(x+y) + i\sin(x+y) &= (\cos x + i\sin x) \cdot (\cos y + i \sin y) \\ &= (\cos x \cdot \cos y - \sin x \cdot \sin y) + i(\cos x \cdot \sin y + \sin x \cdot \cos y).\end{aligned}$$

Vergleicht man Real- und Imaginärteil, so folgt die Behauptung. □

10.7.6 Korollar (Additionstheoreme)

Für alle $x, y \in \mathbb{R}$ gilt:

$$\begin{aligned}\sin x - \sin y &= 2\cos\left(\tfrac{x+y}{2}\right)\sin\left(\tfrac{x-y}{2}\right), \\ \cos x - \cos y &= -2\sin\left(\tfrac{x+y}{2}\right)\sin\left(\tfrac{x-y}{2}\right).\end{aligned}$$

Beweis

Wir zeigen die zweite Formel in den Aufgaben. □

10.7.7 Satz

Für alle $x \in \mathbb{R}$ gilt

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \dots,$$
$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \dots$$

und diese Reihen konvergieren absolut.

Beweis

Für die Potenzen von i gilt:

$$i^n = \begin{cases} 1, & \text{falls } n = 4m \\ i, & \text{falls } n = 4m+1 \\ -1, & \text{falls } n = 4m+2 \\ -i, & \text{falls } n = 4m+3 \end{cases} \qquad (m \in \mathbb{N}_0)$$

(siehe Aufgaben). Also folgt mit der Exponentialreihe

$$e^{i\cdot x} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{k=0}^{\infty} i^k \frac{x^k}{k!} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

□

10.7.8 Satz (Restgliedabschätzung)

Schreiben wir

$$r_{2n+2}(x) := \cos x - \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!},$$
$$r_{2n+3}(x) := \sin x - \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!},$$

so gilt

$$|r_{2n+2}(x)| \leqslant \frac{|x|^{2n+2}}{(2n+2)!} \qquad \forall x \in \mathbb{R} \text{ mit } |x| \leqslant 2n+3,$$
$$|r_{2n+3}(x)| \leqslant \frac{|x|^{2n+3}}{(2n+3)!} \qquad \forall x \in \mathbb{R} \text{ mit } |x| \leqslant 2n+4,$$

Beweis

Dies mit den jetzigen Mitteln der Vorlesung zu zeigen, ist eine nützliche Übung. Später werden wir sehen, dass die Abschätzungen sogar $\forall x \in \mathbb{R}$ gelten. □

10.8 Die Zahl π

10.8.1 Satz und Definition

Die Funktion cos hat im Intervall $[0, 2]$ genau eine Nullstelle. Diese nennen wir $\frac{\pi}{2}$.

Beweis

Es gilt $\cos 0 = 1$ und mit Hilfe der Restgliedabschätzung zeigt man $\cos 2 \leqslant -\frac{1}{3}$. Nach dem Zwischenwertsatz gibt es eine Nullstelle. Aus dem Additionstheorem

$$\cos x_2 - \cos x_1 = -2 \sin \frac{x_1 + x_2}{2} \sin \frac{x_2 - x_1}{2}$$

in 10.7.6 und wegen

$$\sin x > 0 \ \forall x \in]0, 2]$$

(Restgliedabschätzung) folgt: cos ist auf $[0, 2]$ streng monoton fallend. Also gibt es genau eine Nullstelle. □

10.8.2 Bemerkung

Mit Hilfe des Intervallhalbierungsverfahrens und der Restgliedabschätzung lässt sich π näherungsweise berechnen:

$$\pi = 3,141\ldots \ .$$

□

Aus 10.8.1 ergeben sich weitere Eigenschaften von Cosinus und Sinus:

10.8.3 Satz

(i) Es gilt die folgende Wertetabelle:

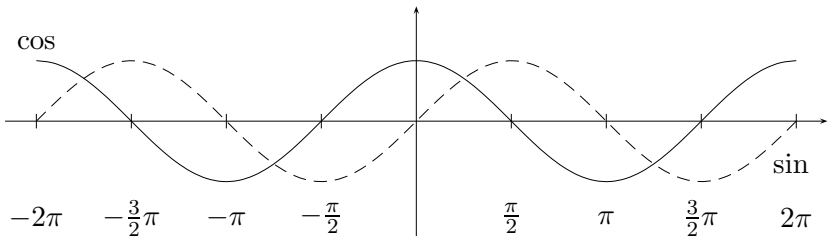
x	0	$\frac{\pi}{2}$	π	$\frac{3}{2}\pi$	2π
$\sin x$	0	1	0	-1	0
$\cos x$	1	0	-1	0	1

(ii) $\forall x \in \mathbb{R}$ gilt:

$$\cos(x + 2\pi) = \cos x, \qquad \sin(x + 2\pi) = \sin x$$

(cos und sin sind (2π)-**periodisch**),

$$\cos(x + \pi) = -\cos x, \qquad \sin(x + \pi) = -\sin x,$$
$$\cos x = \sin\left(\frac{\pi}{2} - x\right), \qquad \sin x = \cos\left(\frac{\pi}{2} - x\right).$$



(iii) Nullstellen von cos und sin:

$$\sin x = 0 \Leftrightarrow \exists k \in \mathbb{Z} \text{ mit } x = k\pi,$$
$$\cos x = 0 \Leftrightarrow \exists k \in \mathbb{Z} \text{ mit } x = \frac{\pi}{2} + k\pi.$$

Insbesondere gilt:

$$e^{ix} = 1 \Leftrightarrow \exists k \in \mathbb{Z} \text{ mit } x = 2k\pi.$$

Beweis

Aufgaben. □

10.9 Tangens und Cotangens

Wir führen Tangens und Cotangens mit Hilfe von Cosinus und Sinus ein. Die Eigenschaften von Tangens und Cotangens ergeben sich aus den entsprechenden Eigenschaften von Sinus und Cosinus. Die entsprechenden Beweise wollen wir nicht näher ausführen.

10.9.1 Definition

Die Funktionen **Tangens** bzw. **Cotangens** sind definiert durch

$$\tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \;\middle|\; k \in \mathbb{Z} \right\} \to \mathbb{R},\ x \mapsto \frac{\sin x}{\cos x},$$

bzw.

$$\cot : \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\} \to \mathbb{R},\ x \mapsto \frac{\cos x}{\sin x}.$$

10.9.2 Satz

(i) tan und cot sind stetig, wo definiert.

(ii) Die Nullstellen von tan bzw. cot sind die Nullstellen von sin bzw. cos.

(iii) tan und cot sind π-**periodisch**:

$$\tan(x+\pi) = \tan x, \qquad \cot(x+\pi) = \cot x$$

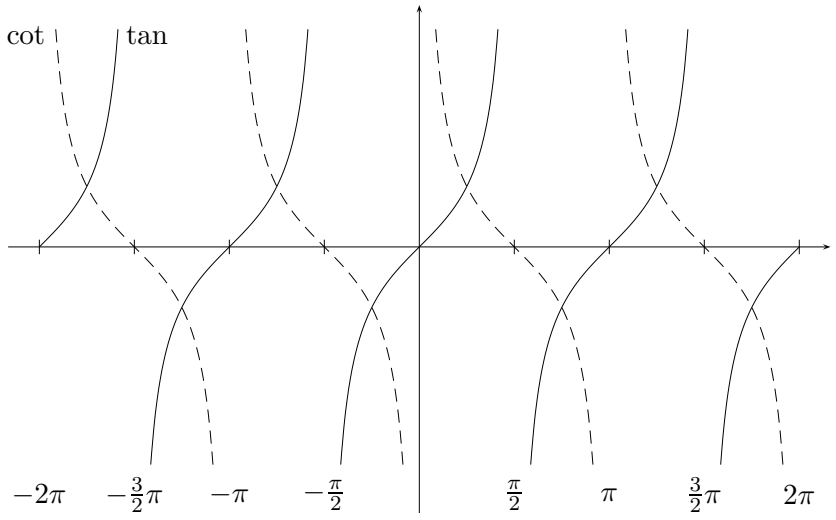
(wo definiert).

(iv) $\forall k \in \mathbb{Z}$ gilt:

$$\lim_{x \to \left(\frac{\pi}{2}+k\pi\right)^+} \tan x = -\infty, \qquad \lim_{x \to \left(\frac{\pi}{2}+k\pi\right)^-} \tan x = \infty$$

bzw.

$$\lim_{x \to (k\pi)^+} \cot x = \infty, \qquad \lim_{x \to (k\pi)^-} \cot x = -\infty\ .$$



(v) **Additionstheoreme**:
Es gilt

$$\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \cdot \tan y},$$
$$\cot(x \pm y) = \frac{\cot x \cdot \cot y \mp 1}{\cot y \pm \cot x},$$

(wo definiert). □

10.10 Arcusfunktionen

10.10.1 Satz und Definition

(i) Die Funktion cos ist im Intervall $[0, \pi]$ streng monoton fallend und bildet dieses Intervall bijektiv auf $[-1, 1]$ ab. Die Umkehrfunktion

$$\arccos : [-1, 1] \to \mathbb{R}$$

heißt **Arcus-Cosinus**.

(ii) Die Funktion sin ist im Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ streng monoton wachsend und bildet dieses Intervall bijektiv auf $[-1, 1]$ ab. Die Umkehrfunktion

$$\arcsin : [-1, 1] \to \mathbb{R}$$

heißt **Arcus-Sinus**.

(iii) Die Funktion tan ist im Intervall $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ streng monoton wachsend und bildet dieses Intervall bijektiv auf $\mathbb{R}$ ab. Die Umkehrfunktion

$$\arctan : \mathbb{R} \to \mathbb{R}$$

heißt **Arcus-Tangens**.

(iv) Die Funktion cot ist im Intervall $]0, \pi[$ streng monoton fallend und bildet dieses Intervall bijektiv auf $\mathbb{R}$ ab. Die Umkehrfunktion

$$\operatorname{arccot} : \mathbb{R} \to \mathbb{R}$$

heißt **Arcus-Cotangens**.

Beweis

(i) Wie im Beweis zu 10.8.1 gesehen, ist cos in $[0, 2]$ und somit insbesondere in $\left[0, \frac{\pi}{2}\right]$ streng monoton fallend. Wegen

$$\cos x = -\cos(\pi - x)$$

ist cos auch in $\left[\frac{\pi}{2}, \pi\right]$ streng monoton fallend. Aus dem Umkehrsatz 9.6.2 folgt

$$\cos[0, \pi] = [\cos \pi, \cos 0] = [-1, 1]$$

und insgesamt die Behauptung.

(ii) folgt aus (i) wegen

$$\sin x = \cos\left(\frac{\pi}{2} - x\right).$$

(iii) und (iv) folgen aus (i) und (ii) zusammen mit 10.9.2, (iv) . □

10.10.2 Bemerkung

Die in 10.10.1 definierten Funktionen nennt man auch **Hauptzweige** der Arcus-Funktionen.

Für beliebige $k \in \mathbb{Z}$ gilt:

(i) cos bildet $[k\pi, (k+1)\pi]$ bijektiv auf $[-1, 1]$ ab.

(ii) sin bildet $[-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi]$ bijektiv auf $[-1, 1]$ ab.

(iii) tan bildet $]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[$ bijektiv auf $\mathbb{R}$ ab.

(iv) cot bildet $]k\pi, (k+1)\pi[$ bijektiv auf $\mathbb{R}$ ab.

Die zugehörigen Umkehrfunktionen

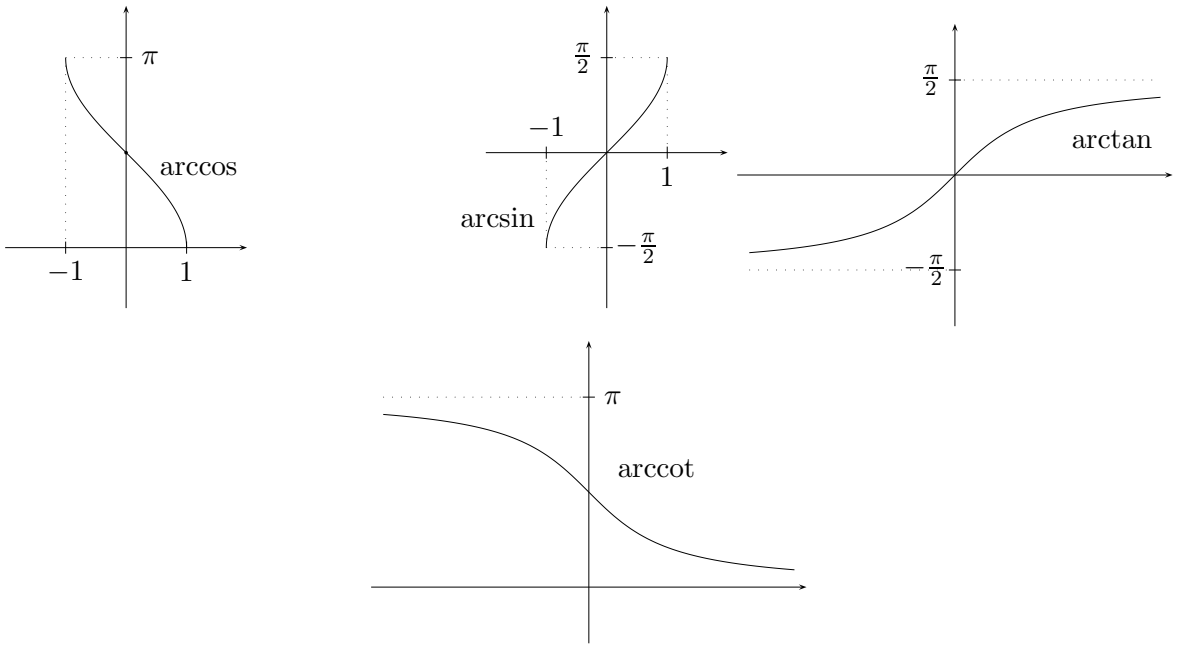
(i) $\arccos_k : [-1, 1] \to \mathbb{R}$,

(ii) $\arcsin_k : [-1, 1] \to \mathbb{R}$,

(iii) $\arctan_k : \mathbb{R} \to \mathbb{R}$,

(iv) $\operatorname{arccot}_k : \mathbb{R} \to \mathbb{R}$

heißen für $k \neq 0$ **Nebenzweige** der Arcus-Funktionen.



□

Zum Schluss kommen wir noch einmal zurück zu den komplexen Zahlen.

10.11 Polarkoordinaten

10.11.1 Satz

Jedes $z \in \mathbb{C}$ lässt sich in der Form

$$z = re^{i\varphi}$$

schreiben mit

$$\varphi \in \mathbb{R} \text{ und } r = |z| \in [0, \infty[\,.$$

Für $z \neq 0$ ist φ bis auf ein ganzzahliges Vielfaches von 2π eindeutig bestimmt.

Beweis

Das ist klar für $z = 0$. Ist $z \neq 0$, so seien $r = |z|$ und $\zeta := \frac{z}{r}$. Ist dann $\zeta = \xi + i\eta$ die Zerlegung von ζ in Real- und Imaginärteil, so gilt $|\zeta| = 1$ und somit auch $\xi^2 + \eta^2 = 1$ sowie $|\xi| \leq 1$. Also ist $\alpha := \arccos \xi$ definiert und wegen $\cos \alpha = \xi$ folgt $\sin \alpha = \pm\sqrt{1 - \xi^2} = \pm\eta$. Wir setzen

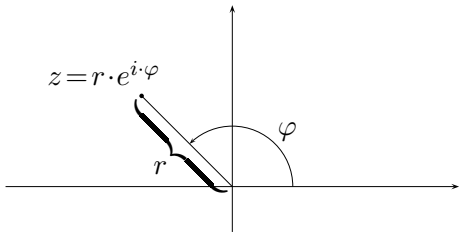
$$\varphi := \begin{cases} \alpha, & \text{falls} \quad \sin\alpha = \eta, \\ -\alpha, & \text{falls} \quad \sin\alpha = -\eta. \end{cases}$$

Dann gilt $z = re^{i\varphi}$. Die Eindeutigkeit von φ bis auf ganzzahlige Vielfache von 2π ergibt sich wegen $e^{i\varphi_1} = e^{i\varphi_2} \Leftrightarrow e^{i(\varphi_1 - \varphi_2)} = 1$ aus 10.8.3, (iii). □

10.11.2 Bemerkung

(i) Die Zahl φ gibt den Winkel (im Bogenmaß) zwischen der positiven reellen Achse und dem Ortsvektor von z an. Wir schreiben auch

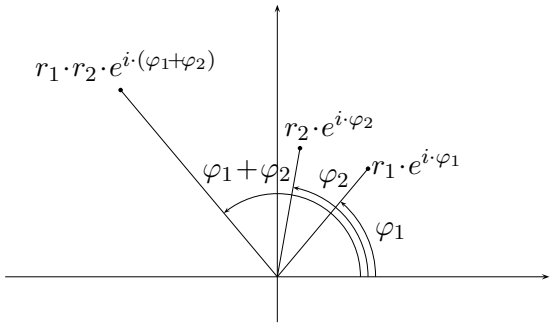
$$\varphi = \arg z. \qquad (\textbf{Argument} \text{ von } z)$$



(ii) Dies erlaubt nun auch eine anschauliche Interpretation der Multiplikation komplexer Zahlen:

$$z_1 = r_1 e^{i\varphi_1}, z_2 = r_2 e^{i\varphi_2} \quad \Rightarrow \quad z_1 \cdot z_2 = r_1 \cdot r_2 \cdot e^{i(\varphi_1 + \varphi_2)}.$$

Man erhält also das Produkt zweier komplexer Zahlen, indem man ihre Beträge multipliziert und ihre Argumente addiert.


$r_1 \cdot r_2 \cdot e^{i\cdot(\varphi_1+\varphi_2)}$
$r_2 \cdot e^{i\cdot\varphi_2}$
$r_1 \cdot e^{i\cdot\varphi_1}$
$\varphi_1+\varphi_2$
φ_2
φ_1

□

Kapitel 11

Differenzierbare Funktionen

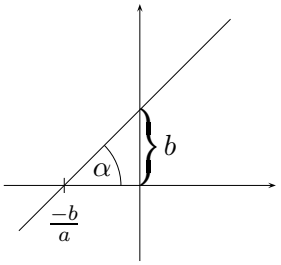
11.1 Motivation

Die Differentialrechnung ist die Lehre von den Veränderungen. Mit ihr werden Wachstumsraten, Verlustquoten, Geschwindigkeiten, Beschleunigungen usw. beschrieben. Sie hilft beim Lösen von Gleichungen, beim Maximieren und Minimieren, bei der Berechnung komplizierter Funktionen, von Bewegungen, Kräften, Impulsen, Energien usw. Die Differentialrechnung ist das wesentliche Hilfsmittel in den Anwendungen der Mathematik in Naturwissenschaften und Technik.

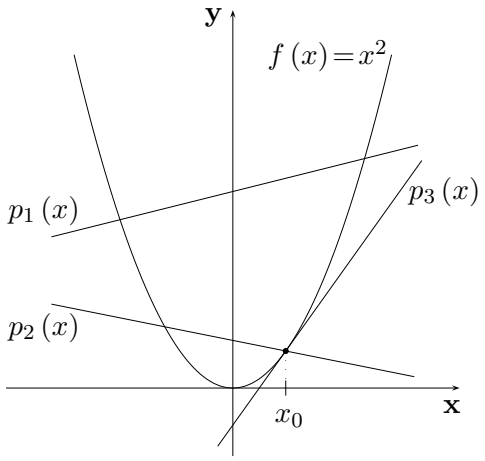
Um zur Definition der Differenzierbarkeit zu gelangen, stellen wir uns die Aufgabe, eine gegebene Funktion $f : X \to \mathbb{R}$ „in der Nähe“ eines gegebenen Punktes $x_0 \in X$ durch eine „Funktion besonders einfacher Gestalt“ anzunähern (dies kann in der Praxis das Berechnen von Funktionswerten oder das Auffinden von Nullstellen erleichtern). Tatsächlich wollen wir f in der Nähe von x_0 *linear approximieren*, das heißt durch eine Polynomfunktion

$$p : \mathbb{R} \to \mathbb{R},\ x \mapsto ax + b, \quad a, b \in \mathbb{R},\ a \neq 0,$$

vom Grad 1 annähern. Der Graph von p ist dann eine Gerade mit der Steigung $a = \tan\alpha$:



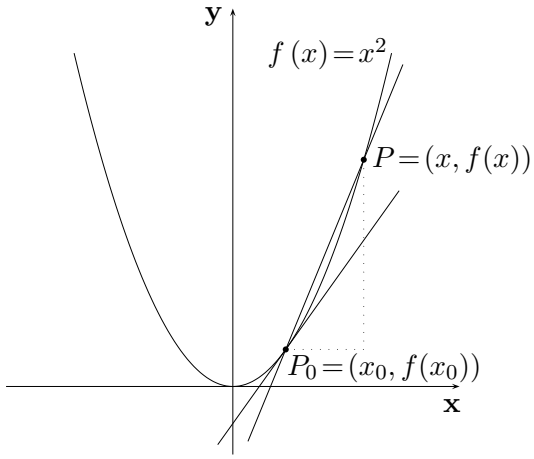
In dem folgenden Bild approximieren p_1 und p_2 die Funktion $f(x) = x^2$ schlecht in der Nähe des Punktes x_0 (für p_1 gilt nicht einmal $p_1(x_0) = f(x_0)$):



Die Funktion p_3 stellt die beste lineare Approximation von f in der Nähe von x_0 dar: Es gilt

$$p_3(x_0) = f(x_0)$$

und unter allen Geraden durch den Punkt $P_0 = (x_0, f(x_0))$ zeichnet sich der Graph von p_3 dadurch aus, dass er sich in der Nähe von P_0 an die Kurve „anschmiegt“. Geometrisch reden wir von der *Tangente* an den Graphen von f in P_0. Diese erhalten wir als Grenzlage der *Sekanten* des Graphen durch P_0:



Die Steigung der Sekante durch P_0 und $P = (x, f(x))$ ist gerade

$$\frac{f(x) - f(x_0)}{x - x_0}. \qquad \textit{(Differenzenquotient)}$$

Die Steigung der Tangente erhält man daraus durch unbegrenzte Annäherung von P an P_0:

$$a = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \qquad \textit{(Differentialquotient)}$$

Die Tangente hat dann die Gleichung

$$y = f(x_0) + a(x - x_0).$$

Dies führt uns zu den folgenden Definitionen. □

11.2 Definitionen

Sei $X \subset \mathbb{R}$. Sind $f : X \to \mathbb{R}$ eine Funktion und $x_0 \in X$ ein Punkt, so fassen wir

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$$

als eine Funktion mit Definitionsbereich $X \setminus \{x_0\}$ auf. Zur Bildung von

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

lassen wir also alle Folgen $(x_n)_{n \in \mathbb{N}}$ mit $x_n \to x_0$ zu, für die gilt $x_n \in X \setminus \{x_0\}\ \forall n \in \mathbb{N}$. In der Notation drückt man dies manchmal so aus, dass man schreibt

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

Dies wollen wir hier aber der Einfachheit halber nicht tun.

Damit es Sinn macht, über die Existenz des Grenzwertes zu reden, setzen wir im folgenden immer voraus, dass es mindestens eine Folge $(x_n)_{n \in \mathbb{N}} \subset X \setminus \{x_0\}$ mit $x_n \to x_0$ gibt. Entsprechend der Notation aus Kapitel 9 fordern wir also, das der Punkt x_0 in $\overline{X \setminus \{x_0\}}$ liegt (man sagt dann auch, x_0 ist kein **isolierter Punkt** von X).

11.2.1 Definition

Seien $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ eine Funktion. Dann heißt f **differenzierbar in einem Punkt** $x_0 \in X$, falls x_0 in $\overline{X \setminus \{x_0\}}$ liegt und der Grenzwert

$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. In diesem Fall heißt $f'(x_0)$ auch **Differentialquotient** oder **Ableitung** von f in x_0. Wir nennen f **differenzierbar**, falls f differenzierbar ist in jedem Punkt von X. In diesem Fall heißt die Funktion $f' : X \to \mathbb{R},\ x \mapsto f'(x)$, die **Ableitung** von f.

11.2.2 Bemerkung

(i) Ist X ein Intervall oder die Vereinigung von Intervallen, so gilt stets

$$x_0 \in X \Rightarrow x_0 \in \overline{X \setminus \{x_0\}}.$$

(ii) Die Bemerkungen in 9.2.7 über Stetigkeit (Einschränkung von Funktionen, lokale Eigenschaft) gelten analog (!) auch für die Differenzierbarkeit.

(iii) Existiert $f'(x_0)$, so kann man mit $h = x - x_0$ den Differentialquotienten auch schreiben als

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Hier sind zur Limesbildung nur solche Folgen $(h_n)_{n \in \mathbb{N}}$ mit $h_n \to 0$ zugelassen, für die gilt $h_n \neq 0$ und $x_0 + h_n \in X\ \forall n \in \mathbb{N}$. □

11.2.3 Beispiel

(i) Konstante Funktionen

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto \lambda,$$

sind differenzierbar mit Ableitung

$$f' : \mathbb{R} \to \mathbb{R},\ x \mapsto 0 :$$

$\forall x_0 \in \mathbb{R}$ gilt:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{\lambda - \lambda}{x - x_0} = 0.$$

(ii) $\mathrm{id}_{\mathbb{R}}$ ist differenzierbar mit Ableitung

$$\mathrm{id}'_{\mathbb{R}} : \mathbb{R} \to \mathbb{R},\ x \mapsto 1 :$$

$\forall x_0 \in \mathbb{R}$ gilt:

$$\lim_{x \to x_0} \frac{\mathrm{id}_{\mathbb{R}}(x) - \mathrm{id}_{\mathbb{R}}(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{x - x_0}{x - x_0} = 1.$$

(iii) $\exp : \mathbb{R} \to \mathbb{R}$ ist differenzierbar mit

$$\exp' = \exp :$$

$\forall x_0 \in \mathbb{R}$ gilt:

$$\begin{aligned} \lim_{h \to 0} \frac{\exp(x_0 + h) - \exp(x_0)}{h} &\overset{\text{Funktional-gleichung}}{=} \lim_{h \to 0} \exp(x_0) \cdot \frac{\exp(h) - 1}{h} \\ &\overset{(!)}{=} \exp(x_0) \cdot 1 \qquad = \exp(x_0). \end{aligned}$$

In der Tat liefert die Restgliedabschätzung in 10.5.2 für $n = 1$:

$$|\exp(h) - (1 + h)| \leq \left|h^2\right| \quad \text{für} \quad |h| \leq \frac{3}{2}.$$

Division durch $|h|$ ergibt für $0 < |h| \leq \frac{3}{2}$:

$$\left| \frac{\exp(h) - 1}{h} - 1 \right| = \left| \frac{\exp(h) - (1 + h)}{h} \right| \leq |h|.$$

Also folgt (!).

(iv) $\sin : \mathbb{R} \to \mathbb{R}$ ist differenzierbar mit

$$\sin' = \cos :$$

$\forall x_0 \in \mathbb{R}$ gilt:

$$\begin{aligned} \lim_{h \to 0} \frac{\sin(x_0 + h) - \sin(x_0)}{h} &\overset{10.7.6}{=} \lim_{h \to 0} \frac{2 \cdot \cos\left(\frac{2x_0 + h}{2}\right) \cdot \sin\frac{h}{2}}{h} \\ &= \left(\lim_{h \to 0} \cos\left(x_0 + \frac{h}{2}\right)\right) \cdot \left(\lim_{h \to 0} \frac{\sin\frac{h}{2}}{\frac{h}{2}}\right) \\ &\overset{\substack{\cos \text{ ist stetig} \\ \text{Aufgaben}}}{=} \cos(x_0) \cdot 1 \qquad = \cos(x_0). \end{aligned}$$

(v) $\cos : \mathbb{R} \to \mathbb{R}$ ist differenzierbar mit

$$\cos' = -\sin :$$

$\forall x_0 \in \mathbb{R}$ gilt:

$$\begin{aligned}\lim_{h\to 0} \frac{\cos(x_0+h)-\cos(x_0)}{h} &\overset{10.7.6}{=} \lim_{h\to 0} \frac{-2\cdot \sin\left(\frac{2x_0+h}{2}\right)\cdot \sin\frac{h}{2}}{h} \\ &= -\left(\lim_{h\to 0}\sin\left(x_0+\frac{h}{2}\right)\right)\cdot\left(\lim_{h\to 0}\frac{\sin\frac{h}{2}}{\frac{h}{2}}\right) \\ &= -\sin x_0. \end{aligned}$$

□

11.2.4 Satz

Für eine Funktion $f : X \to \mathbb{R}$ und $x_0 \in X$ gilt:

$$f \text{ differenzierbar in } x_0 \Rightarrow f \text{ stetig in } x_0.$$

Beweis

Für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset X \setminus \{x_0\}$ gilt:

$$\begin{aligned} f(x_n) &= (f(x_n)-f(x_0))+f(x_0) \\ &= \frac{f(x_n)-f(x_0)}{x_n-x_0}\cdot(x_n-x_0)+f(x_0) \\ &\overset{x_n\to x_0}{\to} f'(x_0)\cdot 0+f(x_0) \qquad = f(x_0). \end{aligned}$$

□

11.2.5 Beispiel

Die Umkehrung von 11.2.4 gilt nicht: Die Funktion

$$|\,| : \mathbb{R} \to \mathbb{R},\ x \mapsto |x|,$$

ist überall stetig, aber nicht differenzierbar in 0: Für

$$h_n := (-1)^n \cdot \frac{1}{n}$$

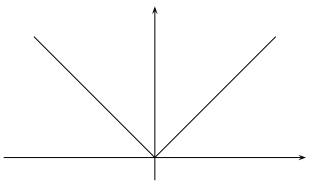
gilt

$$h_n \to 0,$$

aber

$$\frac{|0+h_n|-|0|}{h_n} = \frac{\frac{1}{n}-0}{(-1)^n\cdot\frac{1}{n}} = (-1)^n$$

ist divergent. Anschaulich gibt es im Nullpunkt keine eindeutige Tangente an den Graphen von $|\,|$:



□

11.2.6 Satz

Seien $f : X \to \mathbb{R}$ eine Funktion und $x_0 \in X$ ein Punkt mit $x_0 \in \overline{X \setminus \{x_0\}}$. Dann gilt:

$$f \text{ differenzierbar in } x_0 \Leftrightarrow \exists a \in \mathbb{R} \text{ mit } f(x) = f(x_0) + a \cdot (x - x_0) + \varphi(x) \quad \forall x \in X,$$

wobei $\varphi : X \to \mathbb{R}$ eine Funktion ist mit $\lim_{x \to x_0} \frac{\varphi(x)}{x - x_0} = 0$. In diesem Fall ist $a = f'(x_0)$.

Beweis

i) Sei zunächst f differenzierbar in x_0 und sei $a := f'(x_0)$. Wir definieren die Funktion $\varphi : X \to \mathbb{R}$ durch $\varphi(x) := f(x) - f(x_0) - a(x - x_0)$. Dann gilt

$$\lim_{x \to x_0} \frac{\varphi(x)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - a = f'(x_0) - a = 0.$$

ii) Nun gelte, dass $\exists\, a \in \mathbb{R}$ mit $f(x) = f(x_0) + a(x - x_0) + \varphi(x)$, wobei φ die angegebenen Eigenschaften hat. Dann gilt

$$\lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - a\right) = \lim_{x \to x_0} \frac{\varphi(x)}{x - x_0} = 0.$$

Also ist f differenzierbar in x_0 mit $f'(x_0) = a$. □

11.2.7 Bemerkung

Der Satz drückt aus, dass in der Tat die Differenzierbarkeit in x_0 gleichbedeutend mit der lokalen Approximierbarkeit durch eine Polynomfunktion vom Grad 1 ist. Der Graph von

$$\mathbb{R} \to \mathbb{R},\ x \mapsto f(x_0) + f'(x_0) \cdot (x - x_0),$$

ist die Tangente an den Graphen von f in $(x_0, f(x_0))$. □

11.3 Regeln

11.3.1 Rechenregeln für differenzierbare Funktionen

Sind $f, g : X \to \mathbb{R}$ in $x_0 \in X$ differenzierbare Funktionen und ist $\lambda \in \mathbb{R}$, so sind auch

$$f \pm g,\ f \cdot g \ \text{ und } \ \lambda \cdot f$$

differenzierbar in x_0 und es gilt:

$$(f \pm g)'(x_0) = f'(x_0) \pm g'(x_0),$$

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0), \qquad \textbf{(Produktregel)}$$

$$(\lambda \cdot f)'(x_0) = \lambda \cdot f'(x_0).$$

Ist $g(x_0) \neq 0$, so ist auch

$$\frac{f}{g} : X \setminus \{x \in X \mid g(x) = 0\} \to \mathbb{R}$$

differenzierbar in x_0 mit

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0) \cdot g(x_0) - f(x_0) \cdot g'(x_0)}{(g(x_0))^2}. \qquad \textbf{(Quotientenregel)}$$

Beweis

Wir zeigen die Produktregel:

$$\begin{aligned}
&\lim_{h\to 0}\frac{f(x_0+h)\cdot g(x_0+h)-f(x_0)\cdot g(x_0)}{h}\\
=\;&\lim_{h\to 0}\frac{1}{h}\cdot\left(f(x_0+h)\cdot(g(x_0+h)-g(x_0))+(f(x_0+h)-f(x_0))\cdot g(x_0)\right)\\
=\;&\lim_{h\to 0}f(x_0+h)\cdot\frac{g(x_0+h)-g(x_0)}{h}+\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}\cdot g(x_0)\\
\overset{\substack{f\text{ stetig}\\ \text{in } x_0}}{=}\;&f(x_0)\cdot g'(x_0)+f'(x_0)\cdot g(x_0).
\end{aligned}$$

□

11.3.2 Beispiel

(i) Für alle $n\in\mathbb{N}$ ist die Funktion

$$f_n:\mathbb{R}\to\mathbb{R},\ x\mapsto x^n,$$

differenzierbar mit

$$f_n'(x)=n\cdot x^{n-1}\quad\forall x\in\mathbb{R}.$$

Dies zeigen wir durch vollständige Induktion:

- **Induktionsanfang**: $n=1$

  Wir wissen bereits, dass $f_1=\mathrm{id}_{\mathbb{R}}$ differenzierbar ist mit $f_1'(x)=1\ \forall x\in\mathbb{R}$.

- **Induktionsschluss**: $n\Rightarrow n+1$

  Da $f_{n+1}=f_1\cdot f_n$, folgt aus der Produktregel:

  $$f_{n+1}\text{ ist differenzierbar}$$

  und es gilt

  $$\begin{aligned}f_{n+1}'(x)&=f_1'(x)\cdot f_n(x)+f_1(x)\cdot f_n'(x)\\&=1\cdot x^n+x\cdot\left(n\cdot x^{n-1}\right)=(n+1)\cdot x^n.\end{aligned}$$

(ii) Polynomfunktionen

$$p:\mathbb{R}\to\mathbb{R},\ x\mapsto\sum_{k=0}^{n}a_k\cdot x^k,\quad a_0,\ldots,a_n\in\mathbb{R}$$

sind nach (i) und 11.3.1 differenzierbar mit

$$p'(x)=\sum_{k=1}^{n}k\cdot a_k\cdot x^{k-1}.$$

(iii) Für alle $n\in\mathbb{N}$ ist die Funktion $f:\mathbb{R}\setminus\{0\}\to\mathbb{R},\ x\mapsto\frac{1}{x^n}$, nach (i) und 11.3.1 differenzierbar mit

$$f'(x)=\frac{-n\cdot x^{n-1}}{\left(x^n\right)^2}=\frac{-n}{x^{n+1}}.$$

(iv) $\tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\} \to \mathbb{R}$ ist nach 11.2.3 und 11.3.1 differenzierbar mit

$$\begin{aligned} \tan' x &= \left(\frac{\sin x}{\cos x} \right)' &&= \frac{\sin' x \cdot \cos x - \sin x \cdot \cos' x}{\cos^2 x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} &&= \frac{1}{\cos^2 x}. \end{aligned}$$

(v) $\cot : \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\} \to \mathbb{R}$ ist nach 11.2.3 und 11.3.1 differenzierbar mit

$$\begin{aligned} \cot' x &= \left(\frac{\cos x}{\sin x} \right)' &&= \frac{\cos' x \cdot \sin x - \cos x \cdot \sin' x}{\sin^2 x} \\ &= \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} &&= -\frac{1}{\sin^2 x}. \end{aligned}$$

□

11.3.3 Satz (Ableitung der Umkehrfunktion)

Sei $f : I \to \mathbb{R}$ eine stetige, streng monotone Funktion auf einem Intervall $I \subset \mathbb{R}$. Ist f differenzierbar in $x_0 \in I$ und gilt $f'(x_0) \neq 0$, so ist die Umkehrfunktion

$$f^{-1} : f(I) \to \mathbb{R}$$

differenzierbar in $y_0 := f(x_0)$ und es gilt

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'\left(f^{-1}(y_0)\right)}.$$

Beweis

Der Umkehrsatz 9.6.2 liefert Existenz (und Bijektivität) der Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ auf dem Intervall $f(I)$. Ist nun $(y_n)_{n\in\mathbb{N}} \subset f(I) \setminus \{y_0\}$ eine Folge mit $y_n \to y_0$, so gilt

$$x_n := f^{-1}(y_n) \in I \setminus \{x_0\} \quad \forall n \in \mathbb{N}.$$

Es folgt

$$\begin{aligned} \lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} &= \lim_{n\to\infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} \\ &= \lim_{n\to\infty} \frac{1}{\frac{f(x_n)-f(x_0)}{x_n - x_0}} \qquad = \frac{1}{f'(x_0)}. \end{aligned}$$

□

11.3.4 Beispiele

(i) $\ln :]0, \infty[\to \mathbb{R}$ ist die Umkehrfunktion von $\exp : \mathbb{R} \to \mathbb{R}$. Da

$$\exp'(x) = \exp(x) \neq 0 \quad \forall x \in \mathbb{R}$$

ist ln differenzierbar mit Ableitung

$$\ln' :]0, \infty[\to \mathbb{R},\ y \mapsto \frac{1}{\exp'(\ln(y))} = \frac{1}{y}.$$

(ii) $\arccos : [-1,1] \to \mathbb{R}$ ist die Umkehrfunktion von $\cos : [0,\pi] \to \mathbb{R}$. Es gilt

$$\cos'(x) = -\sin(x) \neq 0 \quad \forall x \in]0,\pi[\,.$$

Also ist arccos auf $]-1,1[$ differenzierbar. Wir berechnen die Ableitung im Punkt

$$y_0 = \cos x_0 \in]-1,1[\,:$$

$$\arccos'(y_0) = \frac{1}{\cos'(x_0)} = \frac{-1}{\sin(x_0)} \underset{x_0 \in]0,\pi[}{=} -\frac{1}{\sqrt{1-\cos^2 x_0}} = -\frac{1}{\sqrt{1-y_0^2}}.$$

(iii) Analog erhält man:

$$\arcsin :]-1,1[\to \mathbb{R}$$

ist differenzierbar mit Ableitung

$$\arcsin' :]-1,1[\to \mathbb{R},\ y \mapsto \frac{1}{\sqrt{1-y^2}}.$$

(iv) $\arctan : \mathbb{R} \to \mathbb{R}$ ist die Umkehrfunktion von $\tan : \left]-\frac{\pi}{2},\frac{\pi}{2}\right[\to \mathbb{R}$. Es gilt

$$\tan' x = \frac{1}{\cos^2 x} \neq 0 \quad \forall x \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[.$$

Also ist arctan auf $\mathbb{R}$ differenzierbar. Wir berechnen die Ableitung im Punkt

$$y_0 = \tan x_0 \in \mathbb{R}\,:$$

$$\begin{aligned}\arctan'(y_0) = \frac{1}{\tan'(x_0)} &= \cos^2 x_0 \\ &= \frac{\cos^2 x_0}{\cos^2 x_0 + \sin^2 x_0} = \frac{1}{1+\tan^2 x_0} = \frac{1}{1+y_0^2}.\end{aligned}$$

(v) Analog erhält man:

$$\operatorname{arccot} : \mathbb{R} \to \mathbb{R}$$

ist differenzierbar mit Ableitung

$$\operatorname{arccot}' : \mathbb{R} \to \mathbb{R},\ y \mapsto -\frac{1}{1+y^2}.$$

□

11.3.5 Satz (Kettenregel)

Seien $f : X \to \mathbb{R},\ g : Y \to \mathbb{R}$ Funktionen mit $f(X) \subset Y$ und $x_0 \in X$. Dann gilt:

$$f \text{ differenzierbar in } x_0, \quad g \text{ differenzierbar in } f(x_0) = y_0$$

$$\Rightarrow g \circ f \text{ differenzierbar in } x_0$$

und in diesem Fall ist

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)\,.$$

Beweis

Wir betrachten die Funktion

$$h : Y \to \mathbb{R},\ y \mapsto \begin{cases} \dfrac{g(y) - g(y_0)}{y - y_0}, & \text{falls } y \neq y_0 \\ g'(y_0), & \text{falls } y = y_0. \end{cases}$$

Dass g differenzierbar ist in y_0 bedeutet gerade

$$\lim_{y \to y_0} h(y) = g'(y_0) = h(y_0).$$

Außerdem gilt:

$$g(y) - g(y_0) = h(y) \cdot (y - y_0) \quad \forall y \in Y.$$

Es folgt

$$\begin{aligned} \lim_{x \to x_0} \frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} &= \lim_{x \to x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} \\ = \lim_{x \to x_0} \frac{h(f(x)) \cdot (f(x) - f(x_0))}{x - x_0} &= \lim_{x \to x_0} h(f(x)) \cdot \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ = g'(f(x_0)) \cdot f'(x_0). & \end{aligned}$$

□

11.3.6 Beispiele

(i) Für $\lambda \in \mathbb{R}$ betrachten wir die allgemeine Potenzfunktion

$$f :]0, \infty[\to \mathbb{R},\ x \mapsto x^\lambda = \exp(\lambda \cdot \ln x).$$

Diese ist nach der Kettenregel differenzierbar und es gilt

$$\begin{aligned} f'(x) = \exp'(\lambda \cdot \ln x) \cdot (\lambda \cdot \ln x)' &= \exp(\lambda \cdot \ln x) \cdot \frac{\lambda}{x} \\ = x^\lambda \cdot \frac{\lambda}{x} \qquad &= \lambda \cdot x^{\lambda - 1}. \end{aligned}$$

Dies verallgemeinert 11.3.2, (i) und (iii).

(ii) Insbesondere sind die Wurzelfunktionen

$$]0, \infty[\to \mathbb{R},\ x \mapsto \sqrt[n]{x},$$

differenzierbar mit Ableitung

$$]0, \infty[\to \mathbb{R},\ x \mapsto \frac{1}{n} \cdot x^{\frac{1}{n} - 1} = \frac{1}{n \cdot \sqrt[n]{x^{n-1}}}.$$

(iii) Wieder nach der Kettenregel ist die Funktion

$$f : \mathbb{R} \to [0, \infty[,\ x \mapsto \sqrt{x^2 + 1},$$

differenzierbar mit Ableitung

$$f' : \mathbb{R} \to [0, \infty[,\ x \mapsto \frac{1}{2 \cdot \sqrt{x^2 + 1}} \cdot 2 \cdot x = \frac{x}{\sqrt{x^2 + 1}}.$$

□

11.4 Höhere Ableitungen

11.4.1 Definition

Seien $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ differenzierbar. Falls die Ableitung $f' : X \to \mathbb{R}$ im Punkt $x_0 \in X$ differenzierbar ist, so heißt

$$f''(x_0) := \left(f'\right)'(x_0)$$

die **zweite Ableitung** von f in x_0. Allgemein definieren wir induktiv: Die Funktion $f : X \to \mathbb{R}$ heißt k**-mal differenzierbar in** $x_0 \in X$, falls $\exists r > 0$, sodass die Funktion

$$f|_{U_r(x_0)\cap X} : U_r(x_0) \cap X \to \mathbb{R}$$

$(k-1)$-mal differenzierbar ist und ihre $(k-1)$-te Ableitung in x_0 differenzierbar ist. In diesem Fall schreiben wir

$$f^{(k)}(x_0) := \left(f^{(k-1)}\right)'(x_0)$$

und nennen diese Zahl die k**-te Ableitung von** $\boldsymbol{f}$ **an der Stelle** $\boldsymbol{x_0}$. Wir verwenden auch die Schreibweisen

$$f^{(0)} := f$$

sowie

$$\frac{df}{dx}(x_0) := f'(x_0)$$

und allgemeiner

$$\frac{d^k f}{dx^k}(x_0) := f^{(k)}(x_0).$$

Wir sagen, f ist k**-mal differenzierbar** (in X), falls f in jedem Punkt von X k-mal differenzierbar ist. Ist zusätzlich die k-te Ableitung

$$f^{(k)} : X \to \mathbb{R}$$

stetig, so sagen wir, f ist k**-mal stetig differenzierbar** (in X). Schließlich heißt f **beliebig oft differenzierbar** (in X), wenn f ist k-mal differenzierbar ist $\forall k \in \mathbb{N}$. □

11.4.2 Bemerkung

Ist $f : X \to \mathbb{R}$ k-mal differenzierbar in X, so sind nach 11.2.4 alle Ableitungen

$$f^{(i)} : X \to \mathbb{R}, \quad i = 0, \ldots, k-1,$$

stetig. □

11.4.3 Beispiele

(i) Induktion zeigt: $\exp : \mathbb{R} \to \mathbb{R}$ ist beliebig oft differenzierbar und es gilt

$$\exp^{(k)} = \exp \quad \forall k.$$

(ii) Induktion zeigt: sin und cos sind beliebig oft differenzierbar und es gilt

$$\sin^{(k)} = \begin{cases} (-1)^m \cdot \cos, & \text{falls } k = 2m+1, \\ (-1)^m \cdot \sin, & \text{falls } k = 2m, \end{cases} \qquad \forall k,$$

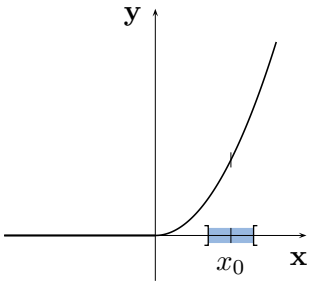
sowie

$$\cos^{(k)} = \begin{cases} (-1)^{m+1} \cdot \sin, & \text{falls } k = 2m+1, \\ (-1)^m \cdot \cos, & \text{falls } k = 2m, \end{cases} \qquad \forall k.$$

(iii) Die Funktion

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x^2, & \text{falls } x \geqslant 0, \\ 0, & \text{falls } x < 0, \end{cases}$$

ist stetig differenzierbar, aber die Ableitung ist nicht differenzierbar in 0.



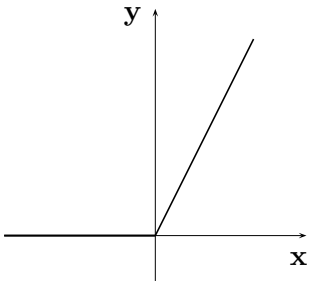
Zunächst ist f differenzierbar in jedem Punkt $x_0 \in \mathbb{R}$. Ist nämlich $x_0 > 0$, so ist f etwa in $U_{\frac{x_0}{2}}(x_0)$ durch $x \to x^2$ gegeben und es folgt: f ist differenzierbar in x_0 mit $f'(x_0) = 2x_0$. Analog erhalten wir für $x_0 < 0$, dass f differenzierbar ist in x_0 mit $f'(x_0) = 0$. Ist $x_0 = 0$, so folgt aus $x_n \to x_0 = 0,\ x_n \neq x_0 = 0\ \forall n$:

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} = \begin{cases} x_n, & \text{falls } x_n > 0, \\ 0, & \text{falls } x_n < 0. \end{cases} \quad \to\ 0 = f(x_0).$$

Die Ableitung

$$f' : \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} 2x, & \text{falls } x \geqslant 0, \\ 0, & \text{falls } x < 0, \end{cases}$$

ist zwar stetig, aber nicht differenzierbar in $x_0 = 0$:



Es gilt

$$\lim_{x \to x_0^+} \frac{f'(x) - f'(x_0)}{x - x_0} = 2,$$

aber

$$\lim_{x \to x_0^-} \frac{f'(x) - f'(x_0)}{x - x_0} = 0.$$

□

Kapitel 12

Lokale Extrema, Mittelwertsätze und erste Anwendungen

12.1 Motivation

Die Mittelwertsätze sind die Grundlage für zahlreiche Anwendungen, wie etwa die Berechnung von Grenzwerten und die Kurvendiskussion. □

12.2 Lokale Extrema

12.2.1 Definition

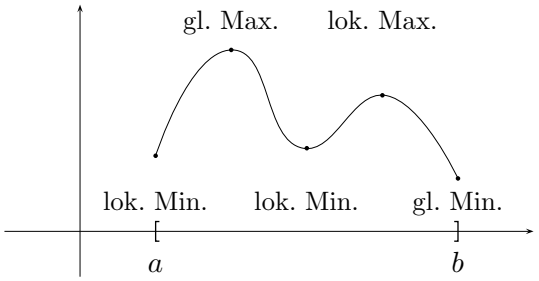
Seien $I \subset \mathbb{R}$ ein Intervall oder eine Vereinigung von Intervallen und $f : I \to \mathbb{R}$ eine Funktion. Wir sagen, f hat in $x_0 \in I$ ein **lokales Maximum** (bzw. **lokales Minimum**), wenn gilt:

$$\exists \varepsilon > 0 \text{ mit } f(x) \leqslant f(x_0) \quad (\text{bzw. } f(x) \geqslant f(x_0)) \quad \forall x \in U_\varepsilon(x_0) \cap I.$$

Wir nennen dann $f(x_0)$ ein **lokales Maximum** (bzw. **lokales Minimum**) und x_0 eine **lokale Maximalstelle** (bzw. **lokale Minimalstelle**). Im Unterschied hierzu nennen wir

$$\max_{x \in I} f(x) \quad \left(\text{bzw. } \min_{x \in I} f(x)\right)$$

im Falle der Existenz auch **globales Maximum** (bzw. **globales Minimum**). Der Sammelbegriff für Maximum und Minimum ist **Extremum**. □



In obigem Bild liegen in den Extremalstellen, die keine Randpunkte sind, waagerechte Tangenten vor. Dies ist immer so:

12.2.2 Satz

Seien $f : I \to \mathbb{R}$ eine Funktion auf einem Intervall I und $x_0 \in I$ ein Punkt, der kein Randpunkt von I ist. Ist dann x_0 eine lokale Extremalstelle von f und ist f differenzierbar in x_0, so gilt

$$f'(x_0) = 0.$$

Beweis

Ist x_0 kein Randpunkt von I und liegt in x_0 etwa eine lokale Maximalstelle von f vor, so

$$\exists \varepsilon > 0 \text{ mit } U_\varepsilon(x_0) \subset I \text{ und } f(x) \leqslant f(x_0) \quad \forall x \in U_\varepsilon(x_0).$$

Es folgt

$$f'(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geqslant 0, \quad f'(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leqslant 0$$

und somit

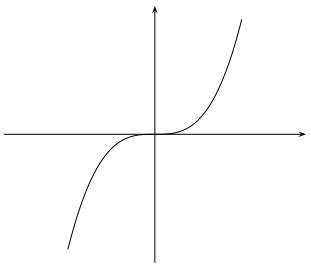
$$f'(x_0) = 0.$$

□

12.2.3 Beispiel

Wir betrachten die Funktion

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^3 :$$



Hier gilt $f'(0) = 0$ obwohl 0 keine lokale Extremalstelle ist. Die Umkehrung von 12.2.2 gilt also nicht. □

12.3 Mittelwertsatz

12.3.1 Satz von Rolle

Ist die Funktion $f : [a, b] \to \mathbb{R}$ stetig in dem abgeschlossenen Intervall $[a, b]$ und differenzierbar in dem offenen Teil $]a, b[$, so gilt:

$$f(a) = f(b) \Rightarrow \exists x_0 \in]a, b[\text{ mit } f'(x_0) = 0.$$

Beweis

Die Aussage ist sicher richtig, falls f konstant ist. Im anderen Fall

$$\exists x_1 \in]a, b[\text{ mit } f(x_1) \neq f(a) = f(b).$$

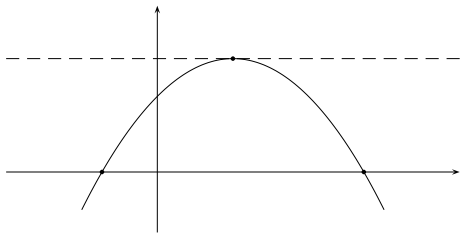
Sei etwa

$$f(x_1) > f(a) = f(b) \qquad (*)$$

(sonst ersetze f durch $-f$). Nach dem Satz vom Maximum 9.5.2 existiert eine (globale) Maximalstelle $x_0 \in [a, b]$ von f. Wegen $(*)$ gilt $x_0 \in]a, b[$ und somit $f'(x_0) = 0$ nach 12.2.2. □

12.3.2 Bemerkung

Der Satz von Rolle besagt insbesondere, dass zwischen zwei Nullstellen einer differenzierbaren Funktion stets eine Nullstelle der Ableitung liegt:



□

12.3.3 Mittelwertsatz

Ist $f : [a, b] \to \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in $]a, b[$, so

$$\exists x_0 \in]a, b[\text{ mit } f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

Beweis

Die Hilfsfunktion

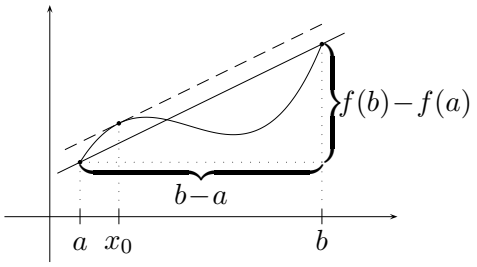
$$F : [a, b] \to \mathbb{R},\ x \mapsto f(x) - \frac{f(b) - f(a)}{b - a}(x - a),$$

erfüllt die Voraussetzungen das Satzes von Rolle. Also $\exists x_0 \in]a, b[$ mit

$$0 = F'(x_0) = f'(x_0) - \frac{f(b) - f(a)}{b - a}.$$

□

Anschaulich bedeutet der Mittelwertsatz, dass die Steigung der Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ gleich der Steigung der Tangente an den Graphen von f an einer gewissen Zwischenstelle $(x_0, f(x_0))$ ist:



12.4 Erste Anwendungen des Mittelwertsatzes

12.4.1 Satz

Ist $f : [a, b] \to \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in $]a, b[$, so gilt:

$$f'(x) = 0 \quad \forall x \in]a, b[\;\Rightarrow\; f \text{ ist konstant.}$$

Beweis

Wir zeigen: Es gilt $f(x) = f(a) \;\forall x \in [a, b]$. Sei $x \in]a, b]$. Betrachten wir f nur auf dem Intervall $[a, x]$, so erfüllt die eingeschränkte Funktion die Voraussetzungen des Mittelwertsatzes. Also

$$\exists x_0 \in]a, x[\text{ mit } 0 = f'(x_0) = \frac{f(x) - f(a)}{x - a}.$$

Es folgt

$$f(x) = f(a).$$

□

Im folgenden schreiben wir $\overset{\circ}{I}$ für das **Innere** eines Intervalls I, d.h. für I ohne die eventuell zu I gehörenden Randpunkte.

12.4.2 Satz (Ableitung und Monotonie)

Seien $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion, die im Innern $\overset{\circ}{I}$ differenzierbar ist. Dann gilt:

(i) $f'(x) \geqslant 0 \;\forall x \in \overset{\circ}{I} \;\Leftrightarrow\; f$ ist monoton wachsend,

(ii) $f'(x) > 0 \;\forall x \in \overset{\circ}{I} \;\Rightarrow\; f$ ist streng monoton wachsend,

(iii) $f'(x) \leqslant 0 \;\forall x \in \overset{\circ}{I} \;\Leftrightarrow\; f$ ist monoton fallend,

(iv) $f'(x) < 0 \;\forall x \in \overset{\circ}{I} \;\Rightarrow\; f$ ist streng monoton fallend.

Beweis

(i) "⇒" Sei $f'(x) \geq 0 \quad \forall x \in \overset{\circ}{I}$. Seien $x_1, x_2 \in I$ mit $x_1 < x_2$. Betrachten wir f nur auf dem Intervall $[x_1, x_2]$, so erfüllt die eingeschränkte Funktion die Voraussetzungen des Mittelwertsatzes. Also

$$\exists x_0 \in]x_1, x_2[\text{ mit } 0 \leq f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Aus $0 < x_2 - x_1$ folgt dann $0 \leq f(x_2) - f(x_1)$.

"⇐" Nun sei f monoton wachsend. Dann gilt für $x_1, x_2 \in \overset{\circ}{I}$ mit $x_1 < x_2$:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0 \text{ und somit } f'(x_1) = \lim_{x_2 \to x_1} \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0.$$

(ii) Im Beweis von (i) ersetze man im "⇒"-Teil an den entsprechenden Stellen $\leq$ durch $<$ (man beachte, dass dies im "⇐"-Teil nicht funktioniert).

(iii) und (iv) ergeben sich analog. □

12.4.3 Satz

Sei $f :]a, b[\to \mathbb{R}$ eine differenzierbare Funktion. In $x_0 \in]a, b[$ sei f zweimal differenzierbar und es gelte

$$f'(x_0) = 0 \quad \text{und} \quad f''(x_0) > 0 \quad \big(\text{bzw. } f''(x_0) < 0\big).$$

Dann besitzt f in x_0 ein lokales Minimum (bzw. Maximum).

Beweis

Wir betrachten den Fall $f''(x_0) > 0$. Da

$$f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} > 0,$$

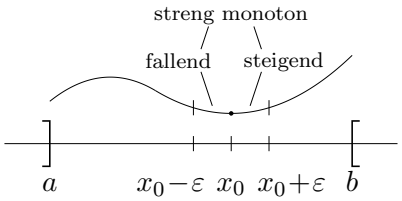
gibt es ein $\varepsilon > 0$ mit

$$\frac{f'(x) - f'(x_0)}{x - x_0} > 0 \quad \forall x \in U_\varepsilon(x_0).$$

Da $f'(x_0) = 0$ folgt daraus

$$f'(x) < 0 \text{ für } x_0 - \varepsilon < x < x_0,$$
$$f'(x) > 0 \text{ für } x_0 < x < x_0 + \varepsilon.$$

Die Behauptung ergibt sich also aus 12.4.2:



□

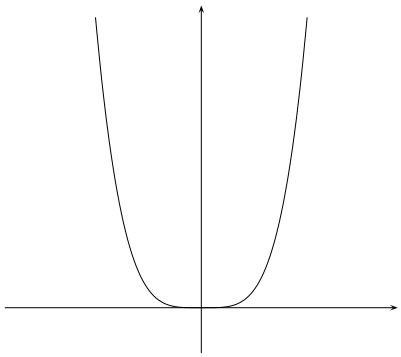
Der Beweis zeigt sogar, dass unter den genannten Voraussetzungen ein **isoliertes lokales Minimum** (bzw. ein **isoliertes lokales Maximum**) vorliegt:

$$\exists \varepsilon > 0 \text{ mit } f(x) < f(x_0) \quad (\text{bzw. } f(x) > f(x_0)) \quad \forall x \in U_\varepsilon(x_0) \text{ mit } x \neq x_0.$$

12.4.4 Beispiel

(i) Die Funktion

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^4,$$



besitzt in $x_0 = 0$ ein lokales Minimum, es gilt jedoch $f''(0) = 0$.

(ii) Für die Funktion

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto e^x - 2x + 1$$

gilt

$$f'(x) = e^x - 2.$$

Es folgt

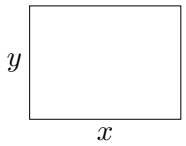
$$f'(x_0) = 0 \Leftrightarrow e^{x_0} = 2 \Leftrightarrow x_0 = \ln 2.$$

Wegen

$$f''(x_0) = e^{x_0} = 2 > 0$$

ist x_0 eine lokale Minimalstelle. Nach 12.2.2 gibt es keine weiteren Extremalstellen (es sind keine Randpunkte zu betrachten).

(iii) Es soll unter allen Rechtecken mit gleichem Flächeninhalt F dasjenige mit kleinstem Umfang gesucht werden. Welche Form hat es?



$$F = xy$$
$$U = 2(x + y)$$

Einsetzen von $y = \frac{F}{x}$ liefert den Umfang als Funktion von x:

$$U :]0, \infty[\to \mathbb{R},\ x \mapsto 2\left(x + \frac{F}{x}\right).$$

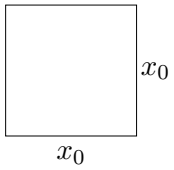
Es gilt

$$U'(x) = 2\left(1 - \frac{F}{x^2}\right)$$

und somit

$$U'(x_0) = 0 \Leftrightarrow 2\left(1 - \frac{F}{x_0^2}\right) = 0 \Leftrightarrow x_0 = \sqrt{F}.$$

Die zweite Ableitung $U'' = \frac{4F}{x^3}$ ist in x_0 positiv, also ist $x_0 = \sqrt{F}$ eine Minimalstelle. Nach 12.2.2 gibt es keine weiteren Extremalstellen. Das gesuchte Rechteck ist also ein Quadrat mit Seitenlänge $x_0 = \sqrt{F}$:



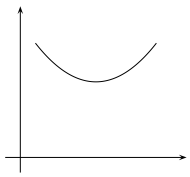
□

Während die erste Ableitung einer Funktion etwas über die Monotonie der Funktion aussagt, hängt die zweite Ableitung mit der Konvexität zusammen.

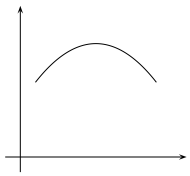
12.5 Konvexität, geometrische Bedeutung der zweiten Ableitung

12.5.1 Motivation

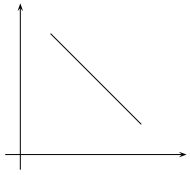
Eine Funktion, deren Graph von der Form



ist, soll **konvex** heißen, eine solche, deren Graph von der Form

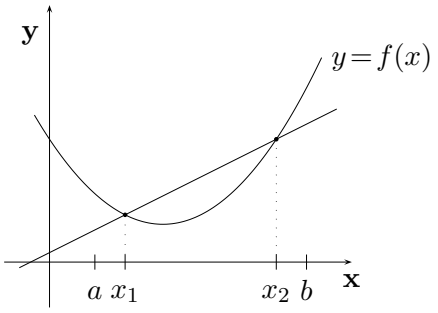


ist, **konkav**. Dabei ist auch der Grenzfall eines Geradenstücks zugelassen:



Hier ist die Funktion sowohl konvex als auch konkav. Liegt eine konvexe (konkave) Funktion vor, deren Graph kein Geradenstück enthält, so sprechen wir auch von einer **streng konvexen** (bzw. **streng konkaven**) Funktion.

Wir fragen uns, wie wir diese anschauliche Idee mathematisch exakt fassen können. Dazu betrachten wir eine konvexe Funktion f wie in folgendem Bild:



Wählen wir $x_1 < x_2$ im Definitionsbereich von f, so gilt $\forall x$ mit $x_1 < x < x_2$ stets

$$f(x) \leqslant g(x), \qquad (*)$$

wobei g die Funktion ist, deren Graph die Gerade durch $(x_1, y_1 = f(x_1))$ und $(x_2, y_2 = f(x_2))$ ist: g hat die Gleichung

$$g(x) = (y_2 - y_1)\frac{x - x_1}{x_2 - x_1} + y_1.$$

Mit der Abkürzung

$$\lambda = \frac{x - x_1}{x_2 - x_1}$$

folgt

$$g(x) = (1 - \lambda) y_1 + \lambda y_2.$$

λ ist das Verhältnis der Strecken $x - x_1$ zu $x_2 - x_1$, also gilt $0 < \lambda < 1$. Löst man umgekehrt nach x auf,

$$x = (1 - \lambda) x_1 + \lambda x_2,$$

so erhält man durch Einsetzen in $(*)$:

$$f((1 - \lambda) x_1 + \lambda x_2) \leqslant (1 - \lambda) y_1 + \lambda y_2 \quad \forall \lambda \in]0, 1[.$$

Dies führt zur folgenden Definition. □

12.5.2 Definition

Eine Funktion $f : I \to \mathbb{R}$ auf einem Intervall I heißt **konvex**, wenn $\forall x_1, x_2 \in I$ und $\forall \lambda \in]0, 1[$ gilt:

$$f((1 - \lambda) x_1 + \lambda x_2) \leqslant (1 - \lambda) f(x_1) + \lambda f(x_2).$$

Steht in der Ungleichung sogar $<$, so heißt f **streng konvex**. Im Falle $\geqslant$ bzw. $>$ heißt f **konkav** bzw. **streng konkav**. □

12.5.3 Satz (2. Ableitung und Konvexität)

Seien $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion, die im Innern $\overset{\circ}{I}$ zweimal differenzierbar ist. Dann gilt:

(i) $f''(x) \geqslant 0 \ \forall x \in \overset{\circ}{I} \Leftrightarrow f$ ist konvex,

(ii) $f''(x) > 0 \ \forall x \in \overset{\circ}{I} \Rightarrow f$ ist streng konvex,

(iii) $f''(x) \leqslant 0 \ \forall x \in \overset{\circ}{I} \Leftrightarrow f$ ist konkav,

(iv) $f''(x) < 0 \ \forall x \in \overset{\circ}{I} \Rightarrow f$ ist streng konkav auf I.

Beweis

(i) Nach 12.4.2 bedeutet die Bedingung auf der linken Seite dass f' in $\overset{\circ}{I}$ monoton wächst.

"$\Rightarrow$" Sei f' monoton wachsend in $\overset{\circ}{I}$. Seien $x_1, x_2 \in I$ und $\lambda \in]0,1[$ sowie $x := (1-\lambda)\, x_1 + \lambda x_2$. Es gelte etwa $x_1 < x_2$. Nach dem Mittelwertsatz gibt es dann $\xi_1 \in]x_1, x[$ und $\xi_2 \in]x, x_2[$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Wegen $x - x_1 = \lambda(x_2 - x_1)$ und $x_2 - x = (1-\lambda)(x_2 - x_1)$ folgt

$$\frac{f(x) - f(x_1)}{\lambda} \leq \frac{f(x_2) - f(x)}{1-\lambda}$$

und somit auch $f(x) \leq (1-\lambda)\, f(x_1) + \lambda f(x_2)$. Also ist f konvex.

"$\Leftarrow$" Nun setzen wir voraus, dass f konvex ist. Seien $x_1, x_2 \in \overset{\circ}{I}$ und es gelte etwa $x_1 < x_2$. Wir wählen ein $\lambda \in]0,1[$ und setzen $x := (1-\lambda)\, x_1 + \lambda x_2$. Dann gilt $f(x) \leq (1-\lambda)\, f(x_1) + \lambda f(x_2)$ und somit auch

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x) - f(x_2)}{x - x_2}.$$

Wegen

$$f'(x_1) = \lim_{x \to x_1} \frac{f(x) - f(x_1)}{x - x_1} \ \text{ und } \ f'(x_2) = \lim_{x \to x_2} \frac{f(x) - f(x_2)}{x - x_2}$$

folgt $f'(x_1) \leq f'(x_2)$. Also ist f' monoton wachsend in $\overset{\circ}{I}$.

(ii) Nach 12.4.2 impliziert die Bedingung auf der linken Seite dass f' in $\overset{\circ}{I}$ monoton wächst. Nun ersetze man im'$\Rightarrow$"-Teil des Beweises von (i) an den entsprechenden Stellen $\leq$ durch $<$.

(iii) und (iv) ergeben sich analog. □

12.5.4 Beispiel

(i) Die folgenden Funktionen sind streng konvex auf $\mathbb{R}$, da ihre zweiten Ableitungen immer > 0 sind:

$$f(x) = x^2, \quad f(x) = e^x.$$

(ii) Die Funktion

$$f(x) = x^3$$

ist streng konvex auf $]0, \infty[$ und streng konkav auf $]-\infty, 0[$. □

12.6 Der verallgemeinerte Mittelwertsatz

12.6.1 Verallgemeinerter Mittelwertsatz

Sind $f, g : [a, b] \to \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in $]a, b[$, so $\exists x_0 \in]a, b[$ mit:

$$f'(x_0)(g(b) - g(a)) = g'(x_0)(f(b) - f(a)).$$

Gilt darüberhinaus $g'(x) \neq 0$ für alle $x \in]a, b[$, so können wir diese Gleichung auch so schreiben:

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Beweis

Die erste Aussage ergibt sich, indem man den Satz von Rolle auf die Hilfsfunktion

$$F : [a, b] \to \mathbb{R},\ x \mapsto f(x)(g(b) - g(a)) - g(x)(f(b) - f(a))$$

anwendet. Die zweite Aussage folgt: gilt $g'(x) \neq 0 \quad \forall x \in]a, b[$, so ist auch $g(b) - g(a) \neq 0$ nach dem Mittelwertsatz 12.3.3. □

12.6.2 Bemerkung

Setzt man in 12.6.1 $g(x) = x \quad \forall x \in [a, b]$, so erhält man den Mittelwertsatz 12.3.3. □

12.7 Die Regeln von de l'Hospital

Seien $f, g : I \to \mathbb{R}$ Funktionen auf einem Intervall I und $x_0 \in I$. Die Bestimmung des Grenzwertes

$$\lim_{x \to x_0} \frac{f(x)}{g(x)}$$

kann schwierig sein, wenn $f(x_0) = g(x_0) = 0$ ist. Sind allerdings f und g differenzierbar in x_0 und ist $g'(x_0) \neq 0$, so ist die Grenzwertbildung einfach:

12.7.1 Regel von de l'Hospital, elementarer Fall

Seien $f, g : I \to \mathbb{R}$ differenzierbar in $x_0 \in I, I$ ein Intervall. Es gelte

$$f(x_0) = g(x_0) = 0$$

sowie

$$g'(x_0) \neq 0 \text{ und } g(x) \neq 0 \quad \forall x \in I \setminus \{x_0\}.$$

Dann folgt

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}.$$

Beweis

Für $x \in I \setminus \{x_0\}$ gilt:

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{\frac{f(x)-f(x_0)}{x-x_0}}{\frac{g(x)-g(x_0)}{x-x_0}} \underset{x \to x_0}{\to} \frac{f'(x_0)}{g'(x_0)}.$$

□

12.7.2 Beispiel

(i)

$$\lim_{x \to 0} \frac{\sin x}{e^x - 1} = \frac{\cos 0}{e^0} = 1.$$

(ii)

$$\lim_{x \to 1} \frac{\ln x}{x^2 - 1} = \frac{\frac{1}{1}}{2 \cdot 1} = \frac{1}{2}.$$

□

Manchmal ergibt sich ein Ergebnis durch mehrmaliges Anwenden der Regel von de l'Hospital:

12.7.3 Beispiel

$$\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \lim_{x \to 0} \frac{\sin x}{2x} = \lim_{x \to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

□

Gibt man sich etwas mehr Mühe, so erhält man mit dem verallgemeinerten Mittelwertsatz:

12.7.4 Regel von de l'Hospital, allgemeiner Fall

Seien $f, g :]a, b[\to \mathbb{R}$ differenzierbare Funktionen (hierbei sind auch $a = -\infty$ und $b = \infty$ zugelassen). Es gelte entweder

(i) $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$

oder

(ii) $\lim_{x \to a} g(x) = \infty$ oder $= -\infty$.

Weiter sei

$$g'(x) \neq 0 \quad \forall x \in]a, b[.$$

Dann folgt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

sofern der rechtsstehende Grenzwert existiert (eigentlich oder uneigentlich). Die analoge Aussage gilt für $x \to b$.

Beweis

Siehe Heuser, Analysis 1, 50.1. □

12.7.5 Beispiel

Für $\lambda \in \mathbb{R}, \lambda > 0$ betrachten wir die Funktion

$$]0,\infty[\to \mathbb{R},\ x \mapsto x^\lambda \ln x = \frac{\ln x}{x^{-\lambda}} =: \frac{f(x)}{g(x)}.$$

Dann sind f, g differenzierbar in $]0,\infty[$, es gilt $\lim_{x\to 0} g(x) = \infty$ (siehe 9.7.9, (iv)) sowie $g'(x) = -\lambda \cdot x^{-\lambda-1} \neq 0\ \forall x \in]0,\infty[$. Mit 12.7.4 folgt

$$\lim_{x\to 0} x^\lambda \ln x = \lim_{x\to 0} \frac{\ln x}{x^{-\lambda}} = \lim_{x\to 0} \frac{\frac{1}{x}}{-\lambda \cdot x^{-\lambda-1}} = \lim_{x\to 0} \frac{-x^\lambda}{\lambda} = 0.$$

Für die Funktion

$$]0,\infty[\to \mathbb{R},\ x \mapsto x^x,$$

ergibt sich dann

$$\lim_{x\to 0} x^x = \lim_{x\to 0} e^{x \ln x} = e^0 = 1.$$

Betrachtet man speziell die Folge $x_n = \frac{1}{n}$, so erhält man $\lim_{n\to\infty} \frac{1}{n^{\frac{1}{n}}} = 1$ und somit auch

$$\lim_{n\to\infty} \sqrt[n]{n} = 1.$$

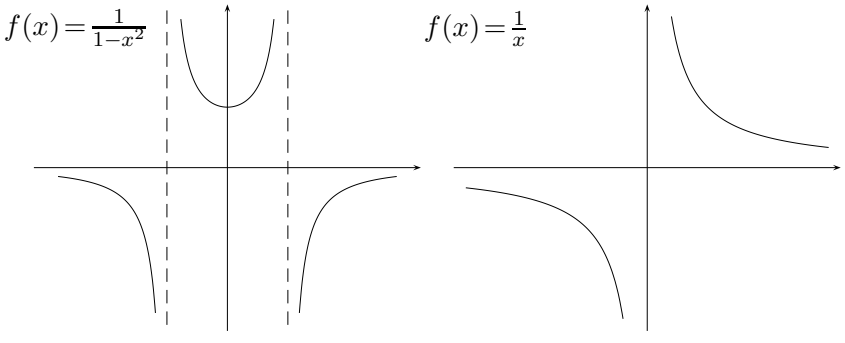
Letzteres kann man natürlich auch anders zeigen. □

12.8 Kurvendiskussion

In Anwendungen benutzt man oft Schaubilder von Funktionen. Um den wesentlichen Verlauf einer reellen Funktion f zu überblicken, geht man zweckmäßig die folgenden Gesichtspunkte der Reihe nach durch (wir setzen voraus, dass f hinreichend oft differenzierbar ist.)

(i) **Definitionsbereich**.
Da die vorgelegte Funktion oft formelmäßig gegeben ist (zum Beispiel $f(x) = \frac{1}{x}$), muss geprüft werden, für welche reellen x der Ausdruck sinnvoll ist (zum Beispiel $\forall x \neq 0$). Definitionsbereiche sind in Anwendungsbeispielen normalerweise Intervalle oder Vereinigungen von endlich vielen Intervallen.

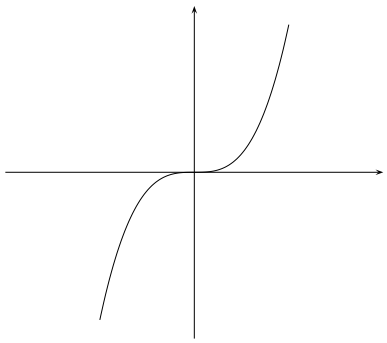
(ii) **Symmetrie**.
Man prüfe, ob f eine **gerade Funktion** ist (das heißt $f(-x) = f(x)\ \forall x$) oder vielleicht eine **ungerade Funktion** ist (das heißt $f(-x) = -f(x)\ \forall x$) ist. Der Graph einer geraden Funktion liegt spiegelbildlich zur y-Achse (zum Beispiel $f(x) = \frac{1}{1-x^2}$), derjenige einer ungeraden Funktion spiegelbildlich zum Nullpunkt (zum Beispiel $f(x) = \frac{1}{x}$).



(iii) **Nullstellen von** f, f', f''.
Bestimme die Nullstellen und so die Intervalle, in denen f, f', f'' positiv bzw. negativ sind. Damit ist insbesondere klar, wo f streng monoton wachsend ($f'(x) > 0$) bzw. streng monoton fallend ($f'(x) < 0$) ist und wo f streng konvex ($f''(x) > 0$) bzw streng konkav ($f''(x) < 0$) ist.

(iv) **Extremalstellen**.
Die Nullstellen von f', zusammen mit den Vorzeichen von f'' ausgewertet in diesen Nullstellen, liefern lokale Maxima und Minima (in Punkten mit $f'(x_0) = f''(x_0) = 0$ sind Sonderuntersuchungen durchzuführen). Man vergesse nicht die Randpunkte des Definitionsbereichs.

(v) **Wendepunkte**.
x_0 heißt **Wendepunkt** von f, wenn die zweite Ableitung beim Durchgang durch x_0 ihr Vorzeichen wechselt. Das heißt, $f''(x_0) = 0$ und $\exists \varepsilon > 0$ so dass in $U_\varepsilon(x_0)$ gilt: links von x_0 ist $f''(x) < 0$ und rechts von x_0 ist $f''(x) > 0$ (oder umgekehrt). Beim Durchgang durch einen Wendepunkt wechselt also die Funktion von streng konkaver Krümmung zu streng konvexer Krümmung (oder umgekehrt). Da dies bedeutet, dass f' in x_0 ein lokales Extremum hat, gilt: Ist $f''(x_0) = 0$ und $f'''(x_0) \neq 0$, so ist x_0 ein Wendepunkt (in Punkten mit $f''(x_0) = f'''(x_0) = 0$ sind Sonderuntersuchungen durchzuführen). Zum Beispiel ist der Nullpunkt ein Wendepunkt von $f(x) = x^3$:



(vi) **Einseitige Grenzwerte, Pole**.
Liegt $x_0 \in \mathbb{R}$ nicht im Definitionsbereich X von f und erfüllt x_0 die Voraussetzungen an $X_{x_0}^+$ bzw. $X_{x_0}^-$ wie in 9.3.5 so bestimme man

$$\lim_{x \to x_0^+} f(x) \quad \text{bzw.} \quad \lim_{x \to x_0^-} f(x)$$

(falls existent). Gilt

$$\lim_{x \to x_0} |f(x)| = \infty,$$

so heißt x_0 ein **Pol** von f (siehe Beispiel 9.3.4, (i)).

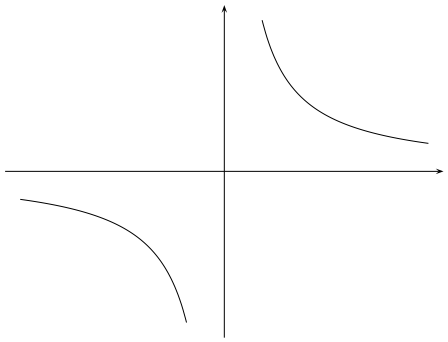
(vii) **Verhalten für große** $|x|$**, Asymptoten**.
Falls sinnvoll, bestimme man

$$\lim_{x \to \infty} f(x) \quad \text{und} \quad \lim_{x \to -\infty} f(x).$$

Allgemeiner suche man, falls sinnvoll, nach „einfachen“ Funktionen h mit

$$|f(x) - h(x)| \to 0 \text{ für } x \to \infty \text{ bzw. } x \to -\infty.$$

Jede solche Funktion h heißt eine **Asymptote** von f. Besonders interessant sind Asymptoten, deren Graph eine Geraden ist. Zum Beispiel:



(viii) **Zeichnung**.
Skizzieren Sie den Graph von f.

12.8.1 Beispiel

Wir diskutieren die rationale Funktion

$$f(x) = \frac{x^4 - 5x^2 + 2}{2x^3}.$$

(i) Der **Definitionsbereich** von f ist $\mathbb{R} \setminus \{0\}$.

(ii) Es gilt $f(-x) = -f(x)\ \forall x \neq 0$, das heißt f ist ungerade. Aus diesem Grunde diskutieren wir im folgenden nur $x > 0$, da für $x < 0$ sich alle Eigenschaften aus der **Symmetrie** zum Nullpunkt ergeben.

(iii) **Nullstellen** von f:

$$f(x) = 0 \Leftrightarrow x^4 - 5x^2 + 2 = 0.$$

Wir setzen $z = x^2$ und lösen $z^2 - 5z + 2 = 0$: $z_{1,2} = \frac{(5 \mp \sqrt{17})}{2}$. Daraus ergeben sich die positiven Nullstellen von f:

$$x_1 = \sqrt{\frac{5 - \sqrt{17}}{2}}, \quad x_2 = \sqrt{\frac{5 + \sqrt{17}}{2}}.$$

Die einzige positive **Nullstelle** von

$$f'(x) = \frac{x^4 + 5x^2 - 6}{2x^4} \quad \text{bzw.} \quad f''(x) = \frac{12 - 5x^2}{x^5}$$

ist

$$x_3 = 1 \quad \text{bzw.} \quad x_4 = \sqrt{\frac{12}{5}}.$$

Durch Berechnung einiger weiterer Werte von f, f', f'' erkennt man:

In $]0, x_1[$ und $]x_2, \infty[$ ist f positiv,
in $]x_1, x_2[$ ist f negativ,
in $]0, x_3[$ ist $f' < 0$, also f streng monoton fallend,
in $]x_3, \infty[$ ist $f' > 0$, also f streng monoton wachsend,
in $]0, x_4[$ ist $f'' > 0$, also f streng konvex,
in $]x_4, \infty[$ ist $f'' < 0$, also f streng konkav.

(iv) Es gilt $f''(x_3) > 0$ (und $f'(x_3) = 0$), d. h. x_3 ist eine **lokale Minimalstelle**. Da keine Randpunkte zu betrachten sind, ist $x_3 = 1$ nach (ii) die einzige Extremalstelle in $]0, \infty[$.

(v) Einziger **Wendepunkt** in $]0, \infty[$ ist x_4. In der Tat ist $f'''(x_4) \neq 0$ (und $f''(x_4) = 0$).

(vi) Es gilt

$$\lim_{x \to 0^+} f(x) = \infty, \quad \lim_{x \to 0^-} f(x) = -\infty,$$

in 0 liegt also ein **Pol** von f vor.

(vii) Wir schreiben $f(x)$ um als

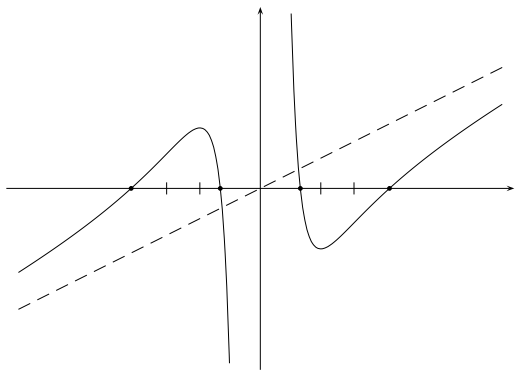
$$f(x) = \frac{x}{2} + \frac{-5 \cdot x^2 + 2}{2 \cdot x^3}.$$

Der rechte Summand strebt für $x \to \pm\infty$ gegen 0, das heißt die Gerade

$$h(x) = \frac{x}{2}$$

ist eine Asymptote für f.

(viii) Der Graph von f sieht wie folgt aus:



□

12.9 Das Newtonsche Verfahren

Wie erwähnt, ist es wichtig, Gleichungen der Form

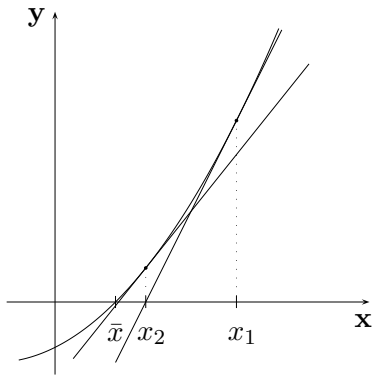
$$f(x) = 0,$$

wobei

$$f : I \to \mathbb{R}$$

eine auf einem Intervall I definierte Funktion ist, lösen zu können. Nicht immer kann man die Lösungen – wie etwa bei quadratischen Polynomfunktionen – durch eine explizite Formel angeben. Es sind Näherungsmethoden notwendig, bei denen die Lösungen als Grenzwerte von Folgen dargestellt werden, deren einzelne Glieder berechnet werden können. Für die Brauchbarkeit eines Näherungsverfahrens ist es wichtig, Fehlerabschätzungen zu haben, damit man weiß, wann man bei einer gegebenen Fehlerschranke das Verfahren abbrechen kann.

Das **Newtonsche Verfahren** beruht auf einer einfachen geometrischen Idee:



Gesucht ist die Schnittstelle $\overline{x}$ des Graphen von f mit der x-Achse. Wir nehmen an, dass ein Punkt x_0 in der Nähe von $\overline{x}$ bekannt ist (den man etwa durch Probieren oder Skizzieren des Graphen gewonnen hat). Man betrachtet die Tangente an den Graphen von f in dem Punkt $(x_0, f(x_0))$ und sucht deren Schnittstelle x_1 mit der x-Achse. Da die Gleichung der Tangente

$$t(x) = f(x_0) + f'(x_0)(x - x_0)$$

lautet, gewinnt man x_1 aus $t(x_1) = 0$:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Wir setzen voraus, dass f stetig differenzierbar ist und dass gilt $f'(x) \neq 0 \; \forall x \in I$. Liegt x_1 in I, so kann man die gleiche Überlegung noch mal anwenden. Sukzessive erhält man die Zahlen

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0,\, 1,\, 2, \ldots \;,$$

von denen wir annehmen wollen, dass sie alle in I liegen (sonst bricht das Verfahren ab). Falls nun diese Folge (**Newtonfolge**) gegen ein $\overline{x} \in I$ konvergiert, so folgt aus Stetigkeitsgründen

$$\overline{x} = \overline{x} - \frac{f(\overline{x})}{f'(\overline{x})}, \text{ also } f(\overline{x}) = 0.$$

Im allgemeinen braucht das Verfahren jedoch nicht zu konvergieren. Wir beschreiben einen wichtigen Fall, in dem Konvergenz auftritt:

12.9.1 Satz

Seien $I = [x_0 - r, x_0 + r]$ ein Intervall und $f : I \to \mathbb{R}$ eine dreimal stetig differenzierbare Funktion mit $f'(x) \neq 0 \ \forall x \in I$. Ferner existiere eine positive Zahl $q < 1$ mit

$$\left| \frac{f(x) f''(x)}{f'(x)^2} \right| \leqslant q \quad \forall x \in I$$

und

$$\left| \frac{f(x_0)}{f'(x_0)} \right| \leqslant (1 - q) r.$$

Dann hat f genau eine Nullstelle $\overline{x}$ in I und die wie oben definierte Newtonfolge $x_0, x_1, x_2, \ldots$ **konvergiert quadratisch** gegen $\overline{x}$, das heißt, es gilt

$$|x_{n+1} - \overline{x}| \leqslant K (x_n - \overline{x})^2 \quad \forall n \in \mathbb{N}_0$$

mit einer geeigneten Konstanten K. Schließlich haben wir die **Fehlerabschätzung**

$$|x_n - \overline{x}| \leqslant \frac{f(x_n)}{M} \text{ mit } M := \min_{x \in I} \left| f'(x) \right|. \qquad \square$$

Es bleibt die Frage: Wie findet man geeignete Anfangsnäherungen x_0? Im Falle konvexer Funktionen ist man mit jeder Anfangsnäherung x_0 erfolgreich, für die gilt $f(x_0) \geq 0$:

12.9.2 Satz

Sei $f : [a, b] \to \mathbb{R}$ zweimal stetig differenzierbar und konvex. Es gelte

$$f'(x) \neq 0 \quad \forall x \in [a, b]$$

und außerdem seien die Vorzeichen von $f(a)$ und $f(b)$ verschieden. Dann gilt: Ist $x_0 \in [a, b]$ beliebig mit $f(x_0) \geqslant 0$, so konvergiert die Newtonfolge, und zwar gegen die einzige Nullstelle von f in $[a, b]$. Es gilt dieselbe Fehlerabschätzung wie in 12.9.1 und ist f sogar dreimal stetig differenzierbar, so ist die Konvergenz quadratisch. $\square$

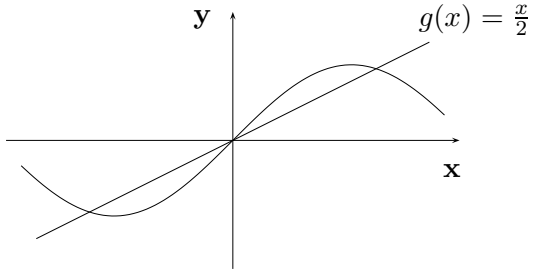
Die analoge Aussage gilt für konkave Funktionen.

12.9.3 Beispiel

Wir suchen die Lösungen von

$$f(x) = \frac{x}{2} - \sin x = 0,$$

das heißt wir suchen die Schnittstellen der Graphen der Funktionen $\sin x$ und $g(x) = \frac{x}{2}$:



Der Nullpunkt ist eine Schnittstelle und wir erwarten zwei weitere, eine etwa bei $x = 2$ und die dritte etwa bei $x = -2$. Wegen der Symmetrie von f brauchen wir nur die Lösung $\overline{x}$ in der Nähe von 2 zu berechnen. Nun ist f (mindestens) dreimal steig differenzierbar und es gilt

$$f'(x) = \frac{1}{2} - \cos x, \quad f''(x) = \sin x, \quad f'''(x) = \cos x.$$

Betrachten wir etwa das Intervall $\left[\frac{\pi}{2}, \pi\right] =: I$, so hat f' in I keine Nullstelle. Weiter ist

$$x_0 = 2 \in I$$

und es gilt

$$f(x_0) = 1 - \sin 2 \geqslant 0.$$

Wir erhalten rekursiv die Newtonfolge

$$x_{n+1} = x_n - \frac{\frac{x_n}{2} - \sin x_n}{\frac{1}{2} - \cos x_n}, \quad n = 0,\, 1,\, 2, \ldots \ .$$

Auf 10 Dezimalstellen gerundet ergibt sich:

$$\begin{aligned} x_0 &= 2,000\,000\,000 \\ x_1 &= 1,900\,995\,594 \\ x_2 &= 1,895\,511\,645 \\ x_3 &= 1,895\,494\,267 \\ x_4 &= 1,895\,494\,267 \end{aligned}$$

In der Tat zeigt die Fehlerabschätzung, dass der Fehler kleiner als $5 \cdot 10^{-10}$ ist. Man beachte die schnelle Konvergenz des Verfahrens. □

Noch ein Wort zu den Beweisen: Die Fehlerabschätzung ergibt sich direkt aus dem Mittelwertsatz. Dass die Konvergenz unter den angegebenen Bedingungen quadratisch ist, folgt mit Hilfe der Taylorformel (dazu später mehr). Dass die Newtonfolge in 12.9.2 konvergiert, ergibt sich wie folgt: Ist etwa $f'(x_0) > 0\ \forall x \in [a, b]$ (sonst ersetze man f durch $-f$), so ist f streng monoton wachsend und hat wegen des Vorzeichenwechsels genau eine Nullstelle $\overline{x} \in [a, b]$. Wegen der Konvexität fällt die Newtonfolge monoton und ist durch $\overline{x}$ nach unten beschränkt. Also ist die konvergent Newtonfolge. Die Konvergenz in 12.9.1 ergibt sich aus dem in den Aufgaben behandelten Banachschen Fixpunktsatz angewandt auf die Funktion

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

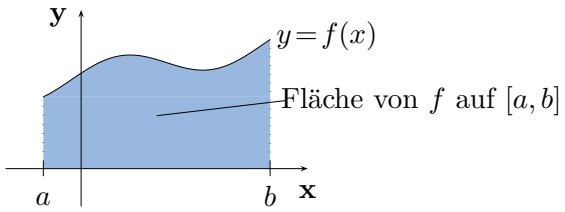
Für Einzelheiten siehe Heuser, Analysis 1, § 70 und Forster, Analysis 1, § 17.

Kapitel 13

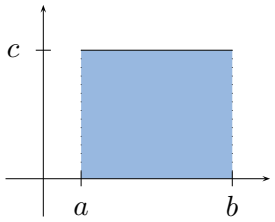
Das Riemannsche Integral

13.1 Motivation

Wie groß ist der Flächeninhalt einer Ellipse, die Länge einer Freileitung, die Energie einer Gasmenge oder die Fluchtgeschwindigkeit einer Rakete? Wie berechnet man Satellitenbahnen, den Schwerpunkt einer Halbkugel, das Trägheitsmoment eines Kegels oder die Wahrscheinlichkeit für den Ausfall eines Bauteils? Dieses vielfältige Spektrum von Fragen kann mit Mitteln der Integralrechnung beantwortet werden.
Dabei geht man von einer elementaren Grundaufgabe aus, nämlich der Bestimmung der Flächeninhalte krummlinig berandeter Flächen. Insbesondere beschäftigt man sich mit Flächen, die zwischen der x-Achse und dem Funktionsgraphen einer reellen Funktion mit positiven Werten liegen:



Handelt es sich um eine konstante Funktion $f(x) = c$, so ist die entsprechende Fläche ein Rechteck:



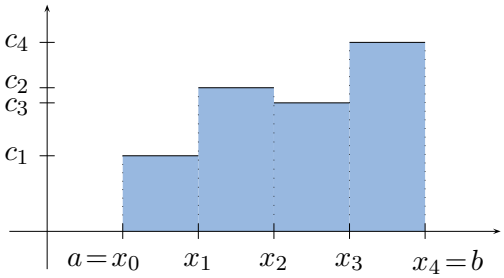
In diesem Fall gilt für die Fläche

$$F = c \cdot (b - a) .$$

Wir schreiben auch

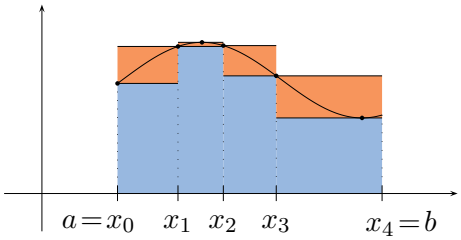
$$\int_a^b f(x)\, dx = c \cdot (b - a) .$$

Durch Aufsummieren kann man nun auch *Treppenfunktionen* integrieren:



$$\int_a^b f(x)\, dx = \sum_{k=1}^{4} c_k \cdot (x_k - x_{k-1})\,.$$

Um den Flächeninhalt für eine „allgemeine“ Funktion zu bestimmen, ja, ihn überhaupt mathematisch exakt zu definieren, wollen wir die Funktion durch Treppenfunktionen annähern:



Wir *zerlegen* das Grundintervall $[a, b]$ und wählen in jedem der entstehenden Teilintervalle den größten bzw. kleinsten Funktionswert aus. Die durch die zugehörigen Treppenfunktionen definierten Flächen (*Obersummen* bzw. *Untersummen*) nähern den gesuchten Flächeninhalt an: Macht man die Zerlegung immer *feiner*, so erhält man immer genauere Annäherungen: Die Obersummen werden immer kleiner (jedenfalls nicht größer), die Untersummen immer größer (jedenfalls nicht kleiner). Infimum bzw. Supremum über alle Ober- bzw. Untersummen nennt man, falls existent, auch *Ober-* bzw. *Unterintegral* von f. Stimmen diese überein, so nennen wir f *integrierbar* auf $[a, b]$ und den gemeinsamen Wert von Ober- und Unterintegral

$$\int_a^b f(x)\, dx$$

das *Integral* von f auf $[a, b]$. Wir werden sehen, dass jede auf $[a, b]$ stetige Funktion integrierbar ist und dass die Integralrechnung die Umkehrung der Differentialrechnung ist: Während man in der Differentialrechnung von gegebenen Funktion die Ableitung berechnet, versucht man umgekehrt in der Integralrechnung aus gegebenen Ableitungen die ursprünglichen Funktionen zu gewinnen. □

13.2 Grundlegende Definitionen

Seien $I = [a, b]$ ein abgeschlossenes Intervall und $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion.

13.2.1 Definition

Eine **Zerlegung** Z von $[a, b]$ ist eine Menge von Punkten

$$Z = \{x_0, x_1, \ldots, x_n\}$$

mit

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

Dabei heißt x_k der ***k*-te Teilpunkt** von Z und $I_k = [x_{k-1}, x_k]$ das k**-te Teilintervall** von Z. Wir nennen $|I_k| = x_k - x_{k-1}$ die **Länge** von I_k und

$$|Z| := \max_{k\in\{1,\ldots,n\}} |I_k|$$

die **Feinheit** von Z. Weiter heißt Z **äquidistant**, wenn alle I_k die gleiche Länge haben. Die Zerlegung Z' heißt **feiner** als Z, falls gilt $Z \subset Z'$. □

13.2.2 Definition

Ist Z eine Zerlegung von I wie oben, so setzen wir

$$M_k := \sup_{x\in I_k} f(x), \quad m_k := \inf_{x\in I_k} f(x)$$

sowie

$$O(Z,f) := \sum_{k=1}^{n} M_k \cdot (x_k - x_{k-1}) \qquad (\textbf{Obersumme von } f \textbf{ bzgl. } Z)$$

und

$$U(Z,f) := \sum_{k=1}^{n} m_k \cdot (x_k - x_{k-1}). \qquad (\textbf{Untersumme von } f \textbf{ bzgl. } Z)$$

□

13.2.3 Satz

(i) Für jede Zerlegung Z von $[a,b]$ gilt

$$U(Z,f) \leqslant O(Z,f).$$

(ii) Sind Z, Z' Zerlegungen von $[a,b]$ und ist Z' feiner als Z, so gilt

$$U(Z,f) \leqslant U\left(Z',f\right)$$

und

$$O(Z,f) \geqslant O\left(Z',f\right).$$

(iii) Für beliebige Zerlegungen Z und Z' von $[a,b]$ gilt

$$U(Z,f) \leqslant O\left(Z',f\right).$$

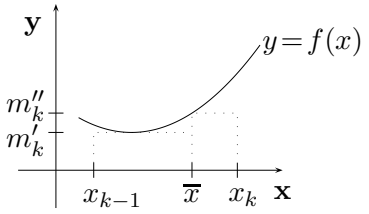
Beweis

(i) Es gilt

$$O(Z,f) - U(Z,f) = \sum_{k=1}^{n} (M_k - m_k) \cdot (x_k - x_{k-1}) \geqslant 0.$$

(ii) Z' enthält die Teilpunkte von Z sowie weitere Teilpunkte. In dem wir das folgende Argument sukzessive anwenden, können wir annehmen, dass Z' genau einen Teilpunkt mehr als Z enthält. Seien etwa

$$Z = \{x_0, \ldots, x_n\}\,,\ Z' = \{x_0, \cdots, x_{k-1}, \overline{x}, x_k, \ldots, x_n\}\,.$$



Wir zeigen die Aussage für die Untersummen. Seien $m_1, \ldots, m_n$ bzgl. Z wie üblich definiert und

$$m'_k := \inf\{f(x) \mid x_{k-1} \leqslant x \leqslant \overline{x}\}\,,$$
$$m''_k := \inf\{f(x) \mid \overline{x} \leqslant x \leqslant x_k\}\,.$$

Dann gilt

$$m_k \leqslant m'_k, \quad m_k \leqslant m''_k$$

und somit

$$\begin{aligned} & U(Z,f) - U\left(Z',f\right) \\ &= m_k \cdot (x_k - x_{k-1}) - m'_k \cdot (\overline{x} - x_{k-1}) - m''_k \cdot (x_k - \overline{x}) \\ &= \left(m_k - m''_k\right) \cdot (x_k - \overline{x}) + \left(m_k - m'_k\right) \cdot (\overline{x} - x_{k-1}) \quad \leqslant 0. \end{aligned}$$

(iii) Es ist

$$Z'' = Z \cup Z'$$

feiner als Z und Z'. Also folgt

$$U(Z,f) \overset{\text{(ii)}}{\leqslant} U\left(Z'',f\right) \overset{\text{(i)}}{\leqslant} O\left(Z'',f\right) \overset{\text{(ii)}}{\leqslant} O\left(Z',f\right). \qquad \square$$

Insbesondere ist die Menge der Obersummen von f nach unten bzw. die Menge der Untersummen von f nach oben beschränkt. Deswegen ist die folgende Definition sinnvoll:

13.2.4 Definition

Wir nennen

$$\overline{I}_f := \inf_Z O(Z,f)$$

Oberintegral von f auf $[a,b]$ und

$$\underline{I}_f := \sup_Z U(Z,f)$$

Unterintegral von f auf $[a,b]$. Stimmen Ober- und Unterintegral von f auf $[a,b]$ überein, so heißt f (**Riemann-**) **integrierbar** auf $[a,b]$. Der gemeinsame Wert

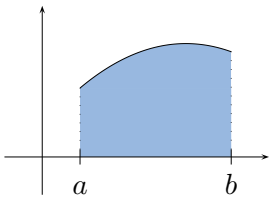
$$\int_a^b f(x)\,dx := \overline{I}_f = \underline{I}_f$$

heißt dann (**Riemann-**) **Integral** von f auf $[a,b]$. $\square$

13.2.5 Definition

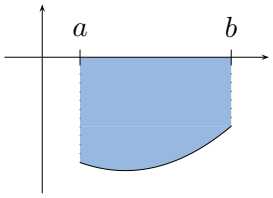
Sei f auf $[a,b]$ integrierbar. Gilt $f(x) \geqslant 0$ auf $[a,b]$, so ist der **Flächeninhalt** der Fläche von f auf $[a,b]$ gleich

$$\int_a^b f(x)\, dx.$$

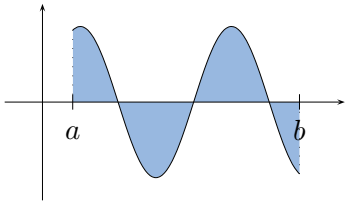


Ist hingegen $f(x) \leqslant 0$ auf $[a,b]$, so ist der **Flächeninhalt** gleich

$$-\int_a^b f(x)\, dx.$$



Ist f sowohl positiv als auch negativ auf $[a,b]$, so muss man entsprechende Teilflächen aufsummieren:

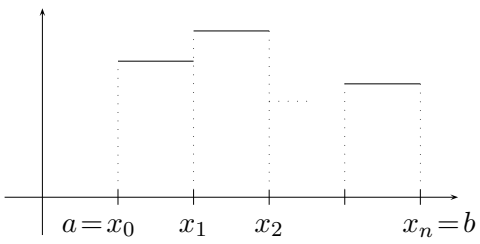


□

13.2.6 Beispiele

(i) Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt **Treppenfunktion**, wenn eine Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a,b]$ existiert, sodass f auf jedem offenen Teilintervall $]x_{k-1}, x_k[$ konstant ist:

$$f(x) = c_k \quad \forall x \in]x_{k-1}, x_k[\,.$$



Jede beschränkte Treppenfunktion ist integrierbar mit

$$\int_a^b f(x)\,dx = \sum_{k=1}^{n} c_k \cdot (x_k - x_{k-1})\,.$$

(ii) Die **Dirichlet-Funktion**

$$f : [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} 1, & \text{falls } x \in \mathbb{Q}, \\ 0, & \text{falls } x \notin \mathbb{Q}, \end{cases}$$

ist nicht integrierbar. Es gilt

$$\overline{I}_f = 1 \text{ und } \underline{I}_f = 0.$$

□

13.3 Rechenregeln

13.3.1 Definition

(i) Für jede in $a \in \mathbb{R}$ definierte Funktion setzen wir

$$\int_a^a f(x)\,dx = 0.$$

(ii) Ist f auf $[a,b]$ integrierbar, so sei

$$\int_b^a f(x)\,dx := -\int_a^b f(x)\,dx.$$

□

Ist $f : [a,b] \to \mathbb{R}$ eine Funktion, so haben wir auch die Funktionen

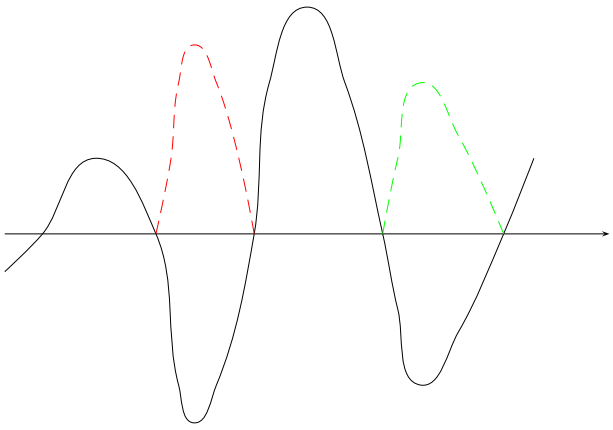
$$f^+ : [a,b] \to \mathbb{R},\ x \mapsto \max\{f(x), 0\}\,,$$

und

$$f^- : [a,b] \to \mathbb{R},\ x \mapsto -\min\{f(x), 0\}\,.$$

Es gilt

$$|f| = f^+ + f^-, \quad f = f^+ - f^-\,.$$



13.3.2 Rechenregeln zur Integration

Seien $[a,b] \subset \mathbb{R}$ ein abgeschlossenes Intervall und f, g integrierbar auf $[a,b]$. Dann sind auch

$$f^+, \quad f^-, \quad |f|, \quad f \pm g, \quad \lambda \cdot f \ (\lambda \in \mathbb{R}), \quad f \cdot g, \quad \frac{f}{g} \ (\text{falls } g(x) \neq 0 \ \forall x \in [a,b])$$

integrierbar auf $[a,b]$ und es gilt:

(i)
$$\int_a^b (f(x) \pm g(x)) \, dx = \int_a^b f(x) \, dx \pm \int_a^b g(x) \, dx.$$

(ii)
$$\int_a^b \lambda \cdot f(x) \, dx = \lambda \cdot \int_a^b f(x) \, dx.$$

(iii)
$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \quad \forall c \in [a,b]$$

(iv) Gilt $m \leqslant f(x) \leqslant M \ \forall x \in [a,b]$, so auch
$$m \cdot (b-a) \leqslant \int_a^b f(x) \, dx \ \leqslant M \cdot (b-a).$$

(v) Gilt $f(x) \leqslant g(x) \ \forall x \in [a,b]$, so auch
$$\int_a^b f(x) \, dx \leqslant \int_a^b g(x) \, dx.$$
Sind f und g darüber hinaus stetig auf $[a,b]$ und $\exists x_0 \in [a,b]$ mit $f(x_0) < g(x_0)$, so folgt sogar
$$\int_a^b f(x) \, dx < \int_a^b g(x) \, dx.$$

(vi) Mit $C := \sup_{x \in [a,b]} |f(x)|$ erhält man
$$\left| \int_a^b f(x) \, dx \right| \leqslant \int_a^b |f(x)| \, dx \leqslant C \cdot (b-a).$$

Beweis

Dies folgt mehr oder weniger direkt aus den Definitionen. Siehe Aufgaben zur GdM2. □

13.4 Das Riemannsche Kriterium

Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion auf einem abgeschlossenen Intervall. Ist Z eine Zerlegung von $[a,b]$, so ist das Supremum bzw. Infimum der Funktion über den zugehörigen Teilintervallen oft nicht einfach zu berechnen. Somit sind die entsprechenden Ober- bzw. Untersummen nicht einfach zu bestimmen. Eine einfachere Methode zur Bestimmung des Integrals ist das **Riemannsche Summenkriterium**. Hier werden statt der Suprema bzw. Infima beliebige Funktionswerte in den Teilintervallen benutzt.

13.4.1 Definition

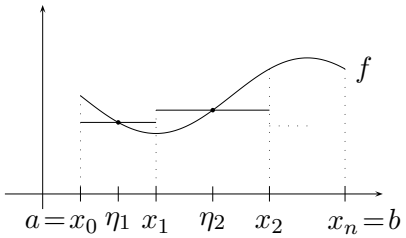
Sei f und Z wie oben, $Z = \{x_0, \ldots, x_n\}$. Wählt man in jedem Teilintervall $[x_{k-1}, x_k]$ einen Zwischenpunkt η_k (**Stützstelle**) und schreibt

$$\eta = (\eta_1, \ldots, \eta_n),$$

so heißt

$$R(Z, \eta, f) := \sum_{k=1}^{n} f(\eta_k) \cdot (x_k - x_{k-1})$$

eine **Riemannsche Summe** von f.



□

13.4.2 Satz

Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion. Dann ist f genau dann integrierbar, wenn jede Folge $(R_\ell)_{\ell \in \mathbb{N}}$ Riemannscher Summen von f, bei denen die Folge der Feinheiten $|Z_\ell|$ der zugehörigen Zerlegungen gegen Null streben, konvergiert. In diesem Fall gilt für jede solche Folge $(R_\ell)_{\ell \in \mathbb{N}}$:

$$\lim_{\ell \to \infty} R_\ell = \int_a^b f(x)\, dx.$$

Beweis

Wird nachgeliefert. □

13.4.3 Beispiel

Wir wollen für $a > 0$

$$\int_0^a \cos x\, dx$$

mittels Riemannscher Summen berechnen. Dazu zeigen wir zunächst:

Sei $t \in \mathbb{R} \setminus \{2\pi k \mid k \in \mathbb{Z}\}$. Dann gilt $\forall n \in \mathbb{N}$:

$$\frac{1}{2} + \sum_{k=1}^{n} \cos(kt) = \frac{\sin\left(n + \frac{1}{2}\right) t}{2 \sin \frac{1}{2} t}. \qquad (*)$$

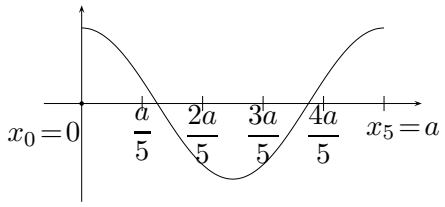
Es gilt $\cos(kt) = \frac{1}{2}\left(e^{ikt} + e^{-ikt}\right)$, also

$$\begin{aligned}
\frac{1}{2} + \sum_{k=1}^{n} \cos(kt) &= \frac{1}{2} \sum_{k=-n}^{n} e^{ikt} \\
= \frac{1}{2} e^{-int} \sum_{k=0}^{2n} e^{ikt} &= \frac{1}{2} e^{-int} \frac{1 - e^{(2n+1)it}}{1 - e^{it}} \\
= \frac{1}{2} \frac{e^{i\left(n+\frac{1}{2}\right)t} - e^{-i\left(n+\frac{1}{2}\right)t}}{e^{i\frac{t}{2}} - e^{-i\frac{t}{2}}} &= \frac{1}{2} \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}\,t}
\end{aligned}$$

Nun zum Integral. Für alle $n \geqslant 1$ erhält man durch

$$\frac{ka}{n}, \quad k = 0, 1, \ldots, n,$$

eine äquidistante Zerlegung Z_n von $[0, a]$ der Feinheit $\frac{a}{n}$. Als Stützstellen wählen wir gerade die rechten Randpunkte der Teilintervalle: Sei $\eta^{(n)} := (\frac{a}{n}, \ldots, a)$.



Für die zugehörigen Riemannschen Summen gilt dann:

$$\begin{aligned}
R\left(Z_n, \eta^{(n)}, \cos x\right) &= \sum_{k=1}^{n} \left(\cos\frac{ka}{n}\right)\left(\frac{ka}{n} - \frac{(k-1)\,a}{n}\right) \\
= \frac{a}{n} \sum_{k=1}^{n} \cos\frac{ka}{n} &\overset{(*)}{=} \frac{a}{n}\left(\frac{\sin\left(n+\frac{1}{2}\right)\frac{a}{n}}{2\sin\frac{a}{2n}} - \frac{1}{2}\right) \\
= \frac{\frac{a}{2n}}{\sin\frac{a}{2n}} \sin\left(a + \frac{a}{2n}\right) - \frac{a}{2n} &\overset{\text{Aufgabe 31, (iii)}}{\longrightarrow} \sin a.
\end{aligned}$$

Andererseits ist die Funktion cos als stetige Funktion auf $[0, a]$ integrierbar (dies werden wir gleich zeigen). Mit obigem Satz folgt

$$\int_0^a \cos x\, dx = \sin a.$$

□

13.5 Gleichmäßige Stetigkeit

Wir zeigen einen Satz, den wir zum Beweis der Tatsache benötigen, dass stetige Funktionen integrierbar sind. Wir betrachten die Teilmenge $X \subset \mathbb{R}$ und eine Funktion $f : X \to \mathbb{R}$. Nach dem ε-δ–Kriterium ist f stetig in X, wenn $\forall\, x_0 \in X$ gilt:

$$\begin{aligned}
&\forall\, \varepsilon > 0\ \ \exists\, \delta > 0, \quad \text{so dass} \quad \forall\, x \in X \ \text{ gilt :} \\
&\qquad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon\ .
\end{aligned}$$

Dabei hängt δ bei festem ε von x_0 ab.

13.5.1 Definition

Eine Funktion $f : X \to \mathbb{R}$ heißt **gleichmäßig stetig** in X, wenn gilt:

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta > 0, \quad \text{so dass} \quad \forall\, x_1, x_2 \in X \;\; \text{gilt}: \\ |x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon \; .$$

□

Sicher ist jede in X gleichmäßig stetige Funktion auch stetig in X. Die Umkehrung gilt aber nicht:

13.5.2 Beispiel

Die Funktion

$$f :]0,1] \to \mathbb{R}, \; x \mapsto \frac{1}{x} \; ,$$

ist sicher stetig in $]0,1]$. Sie ist aber nicht gleichmäßig stetig, denn sonst gäbe es insbesondere zu $\varepsilon = 1$ ein $\delta > 0$, sodass $\forall\, x_1, x_2 \in]0,1]$ gilt:

$$|x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon.$$

Da wir aber ein $n \in \mathbb{N}$ wählen können mit $\left|\frac{1}{n} - \frac{1}{2n}\right| = \frac{1}{2n} < \delta$ ergibt sich wegen $\frac{1}{n},\ \frac{1}{2n} \in]0,1]$ der Widerspruch

$$|f\Big(\frac{1}{n}\Big) - f\Big(\frac{1}{2n}\Big)| = n \geq 1 = \varepsilon \; .$$

□

13.5.3 Satz

Ist $f : [a,b] \to \mathbb{R}$ eine stetige Funktion auf einem abgeschlossenen Intervall $[a,b]$, so ist f in $[a,b]$ gleichmäßig stetig.

Beweis

Wir nehmen an, dass f nicht gleichmäßig stetig ist. Dann $\exists\, \varepsilon > 0$, so dass $\forall\, n \in \mathbb{N}$ Punkte $x_1^{(n)}, x_2^{(n)} \in [a,b]$ existieren mit

$$|x_1^{(n)} - x_2^{(n)}| < \frac{1}{n} \quad \text{und} \quad |f(x_1^{(n)}) - f(x_2^{(n)})| \geq \varepsilon \; .$$

Da alle Glieder der Folge $(x_1^{(n)})_{n\in\mathbb{N}}$ in $[a,b]$ liegen, ist die Folge beschränkt. Sie besitzt also nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge $(x_1^{(n_k)})_{k\in\mathbb{N}}$. Diese muss nach dem Sandwichkriterium gegen einen Punkt $x_0 \in [a,b]$ konvergieren. Wegen $|x_1^{(n_k)} - x_2^{(n_k)}| < \frac{1}{n_k}$ gilt auch $\lim\limits_{k\to\infty} x_2^{(n_k)} = x_0$. Die Stetigkeit von f impliziert

$$\lim_{k\to\infty} \big(f(x_1^{(n_k)}) - f(x_2^{(n_k)})\big) = f(x_0) - f(x_0) = 0$$

im Widerspruch zu $|f(x_1^{(n_k)}) - f(x_2^{(n_k)})| \geq \varepsilon\ \forall\ k$. □

13.6 Integrierbarkeit stetiger und monotoner Funktionen

Welche Funktionen sind integrierbar?

13.6.1 Satz

Ist $f : [a, b] \to \mathbb{R}$ stetig, so ist f auf $[a, b]$ integrierbar.

Beweis

Nach 9.5.2 ist f beschränkt.Weiter ist f nach 13.5.3 sogar gleichmäßig stetig. Ist also $\varepsilon > 0$ beliebig, so $\exists\ \delta > 0$, sodass $\forall\ x_1, x_2 \in [a, b]$ gilt:

$$|x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon. \qquad (*)$$

Für jede Zerlegung $Z = \{x_0, \ldots, x_n\}$ der Feinheit $|Z| < \delta$ gilt dann

$$\begin{aligned} O(Z, f) - U(Z, f) &= \sum_{k=1}^{n} (M_k - m_k)(x_k - x_{k-1}) \\ &\overset{(*)}{\leq} \varepsilon \sum_{k=1}^{n} (x_k - x_{k-1}) \\ &= \varepsilon(b - a)\ . \end{aligned}$$

Da ε beliebig klein gewählt werden kann, wird bei entsprechender Wahl von Z auch $O(Z, f) - U(Z, f)$ beliebig klein. Daraus ergibt sich die Behauptung. □

13.6.2 Bemerkung

(i) Wir nennen eine Funktion $f : [a, b] \to \mathbb{R}$ **stückweise stetig**, wenn sie bis auf endlich viele Stellen stetig ist. Betrachtet man im obigen Beweis nur Zerlegungen Z, für die die endlich vielen Sprungstellen einer solchen Funktion Teilpunkte sind, so sieht man, dass jede beschränkte, stückweise stetige Funktion auf $[a, b]$ auch dort integrierbar ist.

(ii) Mit einem ähnlichen Argument lässt sich zeigen, dass jede beschränkte, stückweise monotone Funktion $f : [a, b] \to \mathbb{R}$ auf $[a, b]$ integrierbar ist (**stückweise monoton** bedeutet, dass es eine Zerlegung Z_0 von $[a, b]$ gibt, sodass f zwischen je zwei benachbarten Teilpunkten von Z_0 monoton ist). Siehe Aufgaben GdM2. □

Eine vollständige Charakterisierung integrierbarer Funktionen werden wir in GdM2 herleiten.

13.7 Mittelwertsatz der Integralrechnung

13.7.1 Mittelwertsatz der Integralrechnung

Ist $f : [a, b] \to \mathbb{R}$ stetig, so

$$\exists x_0 \in [a, b] \ \text{ mit } \int_a^b f(x)\ dx = f(x_0) \cdot (b - a)\,.$$

Beweis

Nach 9.5.2 nimmt f sein Maximum M und sein Minimum m auf $[a,b]$ an, es gilt

$$m \leqslant f(x) \leqslant M \quad \forall x \in [a,b].$$

Mit 13.3.2, (iv) folgt

$$m \cdot (b-a) \leqslant \int_a^b f(x)\, dx \leqslant M \cdot (b-a).$$

Also gilt

$$m \leqslant \frac{1}{b-a} \cdot \int_a^b f(x)\, dx \leqslant M.$$

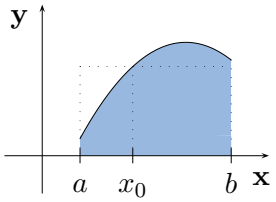
Nach dem Zwischenwertsatz $\exists x_0 \in [a,b]$ mit

$$f(x_0) = \frac{1}{b-a} \cdot \int_a^b f(x)\, dx.$$

□

13.7.2 Bemerkung

Anschaulich bedeutet der Mittelwertsatz, dass es ein Rechteck mit der Höhe $f(x_0)$ gibt, dessen Fläche gleich der Fläche von f ist:



□

Der gleiche Beweis wie eben liefert:

13.7.3 Verallgemeinerter Mittelwertsatz der Integralrechnung

Sind $f, g : [a,b] \to \mathbb{R}$ stetig und gilt $g(x) \geqslant 0\ \forall x \in [a,b]$, so

$$\exists x_0 \in [a,b] \text{ mit } \int_a^b f(x) \cdot g(x)\, dx = f(x_0) \cdot \int_a^b g(x)\, dx.$$

□

Kapitel 16

Vektorräume, Basis, Dimension

16.1 Motivation

Das Wort "Algebra" entstand aus dem Titel eines arabischen Werkes von Muhammed ibn Mûsâ Alchwarizmî aus dem ersten Viertel des 9. Jahrhunderts nach Christus. Naiv können wir sagen: Algebra ist die mathematische Wissenschaft, die sich mit dem Lösen algebraischer Gleichungen beschäftigt. In der linearen Algebra, mit der wir uns hier beschäftigen, geht es um das Lösen linearer Gleichungen über einem Körper K (für die anschauliche Interpretation der folgenden Diskussion betrachten wir den Fall $K = \mathbb{R}$). Eine Gleichung in zwei Unbekannten x und y sieht so aus:

$$a \cdot x + b \cdot y = \alpha, \text{ mit } a, b, \alpha \in K.$$

Gilt hier $(a, b) \neq (0, 0)$, so handelt es sich um die Gleichung einer Geraden (lateinisch: linea). Betrachten wir zwei solche Gleichungen, so erhalten wir ein System von zwei linearen Gleichungen in zwei Unbekannten x und y:

$$\left.\begin{array}{l} a \cdot x + b \cdot y = \alpha \\ c \cdot x + d \cdot y = \beta \end{array}\right\}, \text{ mit } a, b, \alpha, c, d, \beta \in K. \qquad (*)$$

Setzen wir voraus, dass gilt $(a, b) \neq (0, 0)$ und $(c, d) \neq (0, 0)$, so sind

$$L = \{(x, y) \mid ax + by = \alpha\} \text{ und}$$
$$L' = \{(x, y) \mid cx + dy = \beta\}$$

Geraden in der (x, y)-Ebene und die Lösungsmenge von $(*)$ besteht aus dem Durchschnitt $L \cap L'$. Es sind drei Fälle möglich:

(i) L und L' sind nicht parallel, das heißt $L \cap L'$ besteht aus genau einem Punkt.

(ii) L und L' sind parallel und verschieden, das heißt $L \cap L' = \emptyset$.

(iii) L und L' sind gleich, das heißt $L \cap L' = L = L'$ ist eine Gerade.

Für $(*)$ bedeutet dies:

(i) Es gibt genau eine Lösung.

(ii) Es gibt keine Lösung.

(iii) Es gibt unendlich viele Lösungen.

Um herauszufinden, unter welchen algebraischen Bedingungen die Fälle (i), (ii) und (iii) eintreten, formen wir $(*)$ um: Multiplikation mit d, b, c, a liefert

$$\begin{aligned} adx + bdy &= d\alpha \\ bcx + bdy &= b\beta \\ acx + bcy &= c\alpha \\ acx + ady &= a\beta. \end{aligned}$$

Durch Subtraktion ergibt sich:

$$\begin{aligned} (ad - bc)\, x &= (\alpha d - b\beta) \\ (ad - bc)\, y &= (a\beta - \alpha c)\,. \end{aligned} \qquad (**)$$

Die Zahl $ad - bc$ ist eine wichtige Größe, die dem Schema der Koeffizienten a, b, c, d des Gleichungssystems zugeordnet ist. Wir schreiben das Koeffizientenschema auch in Form einer *Matrix*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Die Zahl $ad - bc$ heißt dann die *Determinante* der Matrix, wir schreiben

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Im Hinblick auf Lösungen ergeben sich drei verschiedene Fälle.

Fall A

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0.$$

Dann hat $(**)$ und damit auch $(*)$ genau eine Lösung, nämlich

$$x = \frac{\begin{vmatrix} \alpha & b \\ \beta & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}},\ y = \frac{\begin{vmatrix} a & \alpha \\ c & \beta \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

(*Cramersche Regel*, 1750). Es liegt also der Fall (i) vor. □

Wenn die Determinante verschwindet, gibt es zwei mögliche Fälle:

Fall B

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0 \text{ und } \left(\begin{vmatrix} \alpha & b \\ \beta & d \end{vmatrix} \neq 0 \text{ oder } \begin{vmatrix} a & \alpha \\ c & \beta \end{vmatrix} \neq 0 \right).$$

Hier hat $(**)$ und somit auch $(*)$ keine Lösung, es liegt also der Fall (ii) vor. □

Fall C

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} \alpha & b \\ \beta & d \end{vmatrix} = \begin{vmatrix} a & \alpha \\ c & \beta \end{vmatrix} = 0.$$

Nach Voraussetzung gilt $(a, b) \neq (0, 0)$ und $(c, d) \neq (0, 0)$. Wir betrachen den Fall $a \neq 0$. Dann ist auch $c \neq 0$, denn sonst wäre wegen $ad - bc = 0$ auch $d = 0$ im $\lightning$ zu $(c, d) \neq (0, 0)$. Also sind a und c ungleich 0. Deswegen hat $(*)$ dieselben Lösungen wie

$$acx + bcy = c\alpha$$
$$acx + ady = a\beta.$$

Diese Gleichungen sind aber im Falle C identisch, wir erhalten als Lösungen alle Punkte auf der Geraden

$$ax + by = \alpha.$$

Es liegt also der Fall (iii) vor. □

Wir haben somit die verschiedenen Möglichkeiten für die Lösungen von $(*)$ eindeutig durch Determinantenbedingungen charakterisiert und damit die Theorie von zwei linearen Gleichungen mit zwei Unbekannten im Wesentlichen erledigt.

In der Praxis hat man es mit sehr viel größeren Gleichungssystemen zu tun. Es ist deswegen zunächst wichtig, systematische Schreibweisen einzuführen. Für die Unbekannten benutzen wir indizierte Buchstaben, etwa $x_1, \ldots, x_n$, für die Koeffizienten doppelt indizierte Buchstaben, etwa $a_{11}, \ldots, a_{1n}, \ldots, a_{mn}$. Ein System von m linearen Gleichungen in n Unbekannten $x_1, \ldots, x_n$ geben wir meist in der folgenden Form an:

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

Dabei sind die a_{ij} und die b_i Elemente von K. Eine *Lösung* des Gleichungssystems besteht aus n Elementen $x_1, \ldots, x_n \in K$, die die Gleichungen des Systems erfüllen. Wir fassen die a_{ij} zu einer **Matrix** (a_{ij}) zusammen. Diese stellt man als rechteckige Anordnung der a_{ij} dar:

$$A = (a_{ij}) = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

Die a_{ij} heißen **Koeffizienten** von A, wir sprechen auch von einer Matrix mit Koeffizienten in K. Die Koeffizienten $a_{i1}, \ldots, a_{ij}, \ldots, a_{in}$ bilden die ***i*-te Zeile** von A, die Koeffizienten $a_{1j}, \ldots, a_{ij}, \ldots a_{mj}$ die ***j*-te Spalte** von A:

$$\begin{array}{c} \\ \\ \begin{pmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \ldots & a_{mj} & \ldots & a_{mn} \end{pmatrix} \end{array} \begin{array}{l} j\text{−te Spalte} \\ \downarrow \\ \\ \\ \leftarrow i\text{−te Zeile} \end{array}$$

Eine Matrix der Form $(a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ heißt auch eine $m \times n$–**Matrix**. Die Menge aller $m \times n$–Matrizen mit Koeffizienten in K bezeichnen wir mit $M(m \times n, K)$.

Zusammenfassung

Ein System von m linearen Gleichungen in n Unbekannten über K wird durch eine Koeffizientenmatrix $A = (a_{ij}) \in M\,(m \times n, K)$ sowie ein senkrecht als Spalte geschriebenes m-Tupel $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$ gegeben (dass wir in diesem Kontext Tupel als Spalten schreiben, wird sich später als zweckdienlich erweisen). In Kurzform:

$$A \cdot x = b, \qquad \text{mit } \; x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Ist ein solches System gegeben, so stellen wir uns folgende Fragen:

- Hat das System Lösungen?
- Wenn ja, wie viele Lösungen hat das System?
- Wie findet man die Lösungen?

Die lineare Algebra stellt eine systematische Theorie zur Beantwortung dieser Fragen zur Verfügung. Grundlegend ist die Struktur des *Vektorraums*, die sich automatisch ergibt, wenn man ein **homogenes lineares Gleichungssystem** betrachtet. Dabei bedeutet homogen, dass die rechte Seite des Systems verschwindet:

$$\begin{aligned} a_{11}x_1 \; + \; \ldots \; + \; a_{1n}x_n \; &= \; 0 \\ &\vdots \\ a_{m1}x_1 \; + \; \ldots \; + \; a_{mn}x_n \; &= \; 0 \; . \end{aligned} \tag{$***$}$$

Schreiben wir die Lösungen senkrecht als Spalten, so ist mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\text{ und }\; y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \;\text{ auch }\; x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

eine Lösung von $(***)$. Ebenso ist für jedes $\lambda \in K$ mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\text{ auch }\; \lambda \cdot x := \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

eine Lösung von $(***)$. Die so definierten Operationen $+$ und $\cdot$ unterliegen Rechenregeln, die man dann in der Definition des *Vektorraums* aufgreift. Wie sich herausstellen wird, ist die Lösungsmenge eines Gleichungssystems $A \cdot x = b$ entweder leer, oder sie ergibt sich durch „Parallelverschiebung“ aus der Lösungsmenge des *zugehörigen homogenen Gleichungssystems* $A \cdot x = 0$, wobei die Null auf der rechten Seite für den entsprechenden *Nullvektor* steht. Die Lösungsmenge des homogenen Systems wiederum bildet einen *Untervektorraum* des K^n. $\square$

16.2 Grundlegende Definitionen

Sei K ein Körper.

16.2.1 Definition

Eine Menge V zusammen mit einer Verknüpfung

$$+ : V \times V \to V,\ (v, w) \mapsto v + w, \qquad \textbf{(Addition)}$$

und einer Abbildung

$$\cdot : K \times V \to V,\ (\lambda, v) \mapsto \lambda \cdot v, \qquad \textbf{(Skalarmultiplikation)}$$

heißt ***K*-Vektorraum** (**Vektorraum über *K***), wenn gilt:

(i) $(V, +)$ ist eine abelsche Gruppe.

(ii) Es gelten die **Distributivgesetze**

$$\begin{array}{l}(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v, \\ \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w,\end{array} \qquad \forall \lambda, \mu \in K, \forall v, w \in V,$$

das **Assoziativgesetz**

$$\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v \qquad \forall \lambda, \mu \in K, \forall v \in V$$

sowie

$$1 \cdot v = v \qquad \forall v \in V.$$

In diesem Fall nennen wir $+$ eine **innere Verknüpfung** und $\cdot$ eine **äußere Verknüpfung**. Die Elemente von V heißen **Vektoren**, das neutrale Element 0 der Addition heißt **Nullvektor**, die Elemente von K werden als **Skalare** bezeichnet. Wir schreiben auch

$$\lambda v := \lambda \cdot v \text{ für } \lambda \in K, v \in V.$$

Ist $K = \mathbb{R}$ bzw. $K = \mathbb{C}$, so sprechen wir auch von einem **reellen** bzw. **komplexen Vektorraum**. □

16.2.2 Satz (Rechenregeln)

Sei V ein K-Vektorraum. Dann gilt

(i) $0 \cdot v = 0$ $\qquad \forall v \in V.$

(ii) $\lambda \cdot 0 = 0$ $\qquad \forall \lambda \in K.$

(iii) $\lambda \cdot v = 0 \Rightarrow \lambda = 0$ oder $v = 0$ $\qquad \forall \lambda \in K, v \in V.$

(iv) $(-1) \cdot v = -v$ $\qquad \forall v \in V.$

Beweis

(i) $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$ und wir können die Kürzungsregel in der Gruppe $(V, +)$ anwenden.

(ii) $\lambda \cdot 0 = \lambda \cdot (0+0) = \lambda \cdot 0 + \lambda \cdot 0$ sowie Kürzungsregel.

(iii) Sei $\lambda \cdot v = 0$ aber $\lambda \neq 0$. Dann gilt:

$$v = 1 \cdot v = \left(\lambda^{-1} \cdot \lambda\right) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0 \overset{\text{(ii)}}{=} 0.$$

(iv) $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1-1) \cdot v = 0 \cdot v \overset{\text{(i)}}{=} 0.$ □

16.2.3 Beispiel

(i) Sei $n \in \mathbb{N}$. Wir betrachten die Menge

$$K^n = \{x = (x_1, \ldots, x_n) \mid x_i \in K\}$$

der n-Tupel von Elementen in K. Entsprechend unseren Überlegungen in Motivation 16.1 erklären wir eine Addition

$$+ : K^n \times K^n \to K^n,\ (x, y) \mapsto x + y,$$

durch

$$x + y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

sowie eine Skalarmultiplikation

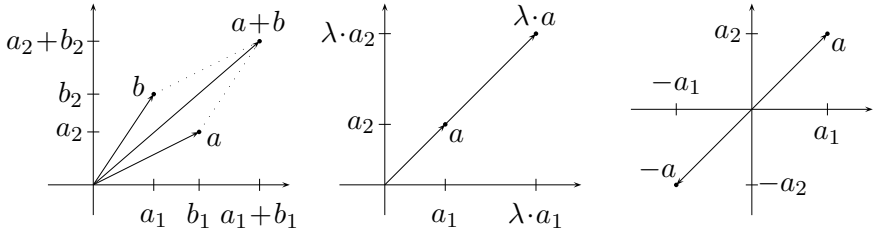
$$\cdot : K \times K^n \to K^n,\ (\lambda, x) \mapsto \lambda \cdot x,$$

durch

$$\lambda \cdot x = \lambda \cdot (x_1, \ldots, x_n) = (\lambda \cdot x_1, \ldots, \lambda \cdot x_n)\,.$$

Dadurch wird K^n zu einem K-Vektorraum mit Nullvektor $0 = (0, \ldots, 0)$.

Im Falle $K = \mathbb{R}$ können wir Addition und Skalarmultiplikation wie üblich geometrisch deuten. Etwa für $n = 2$:



(ii) Seien $m, n \in \mathbb{N}$. Sind $A = (a_{ij})$ und $B = (b_{ij})$ Matrizen in $M\,(m \times n, K)$ und ist $\lambda \in K$, so sind auch

$$A + B := (a_{ij} + b_{ij}) \text{ und } \lambda \cdot A = (\lambda \cdot a_{ij})$$

Matrizen in $M(m \times n, K)$. Mit den so definierten Operationen wird $M(m \times n, K)$ zu einem K-Vektorraum, dessen Nullvektor die Nullmatrix mit m Zeilen und n Spalten ist:

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

Schreibt man die Zeilen einer Matrix (a_{ij}) hintereinander als

$$(a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{m1}, \dots, a_{mn}),$$

so sieht man, dass $M(m \times n, K)$ als K-Vektorraum letztendlich nichts anderes ist als $K^{m \cdot n}$.

(iii) Für den Polynomring $K[t]$, in dem ja insbesondere schon eine Addition erklärt ist, definieren wir eine Skalarmultiplikation durch

$$K \times K[t] \to K[t], \ (\lambda, p) \mapsto \lambda \cdot p,$$

durch

$$\lambda \cdot \sum_{i=0}^{n} a_i t^i = \sum_{i=0}^{n} (\lambda \cdot a_i) t^i$$

Dadurch wird $K[t]$ zu einem K-Vektorraum. Der Nullvektor ist das Nullpolynom.

(iv) Ist $\emptyset \neq X$ eine Menge und K ein Körper, so macht man den Ring aller Abbildungen $f : X \to K$ aus Beispiel 5.3.4 , (ii) zu einem K-Vektorraum, indem man $\lambda \cdot f$ elementweise definiert:

$$(\lambda \cdot f)(x) := \lambda \cdot f(x) \quad \forall x \in X.$$

□

In Motivation 16.1 haben wir im Hinblick auf die Lösungen eines homogenen linearen Gleichungssystems mit n Unbekannten bereits angesprochen, dass diese einen Untervektorraum des K^n bilden. Hier ist die entsprechende Definition:

16.2.4 Satz und Definition

Seien V ein K-Vektorraum und $U \subset V$ eine Teilmenge. Dann heißt U ein **Untervektorraum** von V, wenn gilt:

(i) $U \neq \emptyset$.

(ii) U ist abgeschlossen bezüglich der Addition: $u, v \in U \Rightarrow u + v \in U$.

(iii) U ist abgeschlossen bezüglich der Skalarmultiplikation: $\lambda \in K, v \in U \Rightarrow \lambda v \in U$.

In diesem Fall wird U zusammen mit der durch V induzierten Addition und Skalarmultiplikation zu einem K-Vektorraum.

Beweis

Wegen (iii) und 16.2.2, (iv) gilt: Mit $v \in U$ ist auch $-v = (-1) \cdot v \in U$. Andererseits gibt es nach (i) wenigstens ein solches Element $v \in U$. Also ist nach (ii) auch das Nullelement $0 = v - v \in U$. Daraus folgt die Behauptung. □

Ist $(U_i)_{i \in I}$ eine beliebige Familie von Untervektorräumen von V, so ist offensichtlich auch $U := \bigcap_{i \in I} U_i$ ein Untervektorraum von V.

16.2.5 Beispiel

(i) $U := \{(a, b, 0) \mid a, b \in \mathbb{R}\}$ ist ein Untervektorraum von $\mathbb{R}^3$.

(ii) $U := \{(a, b, 1) \mid a, b \in \mathbb{R}\}$ ist kein Untervektorraum von $\mathbb{R}^3$, denn $(0, 0, 0) \notin U$. □

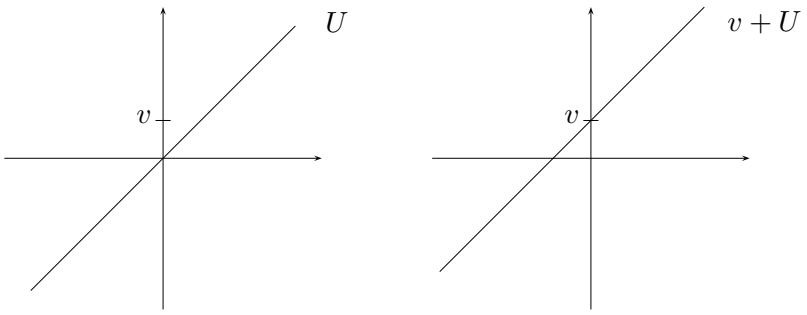
In Motivation 16.1 hatten wir im Hinblick auf die Lösungen eines linearen Gleichungssystems auch Teilmengen des K^n angesprochen, die durch „Parallelverschiebung“ eines Untervektorraums entstehen. Hier ist die entsprechende Definition:

16.2.6 Definition

Seien V ein K-Vektorraum und $X \subset V$ eine Teilmenge. Dann heißt X ein **affiner Unterraum** von V, falls es einen Vektor $v \in V$ und einen Untervektorraum U von V gibt mit

$$X = v + U := \{v + u \mid u \in U\}.$$

Auch die leere Teilmenge wird affiner Unterraum genannt. □



16.2.7 Lemma

Sei $X = v + U$ ein affiner Unterraum von V. Dann gilt:

(i) $v' \in X \;\Rightarrow\; X = v' + U$.

(ii) Sind $v' \in V$ ein Vektor und $U' \subset V$ ein Untervektorraum mit

$$v + U = v' + U',$$

so folgt

$$U = U' \text{ und } v' - v \in U.$$

Beweis

Einfache Übung. □

Zu einem affinen Unterraum $X = v + U$ von V ist also der Untervektorraum U eindeutig bestimmt, während der '„Aufhängepunkt“ v beliebig in X gewählt werden kann.

16.3 Linearkombinationen, lineare Unabhängigkeit

Jede Teilmenge eines Vektorraums kann zu einem Untervektorraum „abgeschlossen“ werden:

16.3.1 Bemerkung und Definition

Sei V ein K-Vektorraum.

(i) Sei $(v_i)_{i\in I}$ eine Familie von Vektoren in V. Ein Vektor $v \in V$ heißt **Linearkombination** der $(v_i)_{i\in I}$, wenn es Indizes $i_1, \ldots, i_n$ und Skalare $\lambda_1, \ldots, \lambda_n \in K$ gibt mit

$$v = \sum_{\mu=1}^{n} \lambda_\mu v_{i_\mu}.$$

Wir bezeichnen mit

$$\operatorname{span}(v_i)_{i\in I} = \operatorname{span}_K (v_i)_{i\in I}$$

die Menge aller Linearkombinationen der $(v_i)_{i\in I}$. Tatsächlich erhalten wir so den kleinsten Untervektorraum von V, der alle v_i, $i \in I$, enthält (nämlich den Durchschnitt aller Untervektorräume von V, die diese Vektoren enthalten). Dieser heißt der von den $(v_i)_{i\in I}$ **aufgespannte Vektorraum**. Wir setzen noch $\operatorname{span}(\emptyset) := \{0\}$.

Ist speziell $I = \{1, \ldots, n\}$ und sind $v_1, \ldots, v_n \in V$, so schreiben wir auch

$$\begin{aligned}\operatorname{span}(v_1, \ldots, v_n) &:= \operatorname{span}_K (v_1, \ldots, v_n) \\ &:= \left\{ v \in V \;\middle|\; \exists \lambda_1, \ldots, \lambda_n \in K \text{ mit } v = \sum_{i=1}^{n} \lambda_i v_i \right\}\end{aligned}$$

für den von $v_1, \ldots, v_n$ aufgespannten Untervektorraum.

(ii) Ist $(v_i)_{i\in I}$ eine Familie von Vektoren in V mit

$$\operatorname{span}(v_i)_{i\in I} = V,$$

so nennen wir die Familie ein **Erzeugendensystem** von V. Wir sagen, V ist **endlich erzeugt**, falls es ein endliches Erzeugendensystem gibt. □

16.3.2 Beispiel

(i) Ist K ein Körper, so betrachten wir in K^n die Vektoren

$$e_i := (\ 0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \ldots, 0\ ), \quad i = 1, \ldots, n.$$

Dann gilt

$$\operatorname{span}(e_1, \ldots, e_n) = K^n.$$

(ii) Der Polynomring $K[t]$ über einem Körper K wird erzeugt von $1, t, t^2, \ldots$:

$$K[t] = \operatorname{span}\left(t^i\right)_{i \in \mathbb{N}_0}.$$

(iii) Sei $0 \neq v_1 \in \mathbb{R}^3$. Dann ist $\operatorname{span}(v_1)$ die Gerade durch 0 und v_1. Ist $v_2 \in \mathbb{R}^3$ ein weiterer Vektor, $v_2 \notin \operatorname{span}(v_1)$, so ist $\operatorname{span}(v_1, v_2)$ die Ebene durch $0, v_1, v_2$.

(iv) Seien

$$\begin{aligned} v_1 &= (1,2,3,4), & v_2 &= (6,6,10,10), \\ v_3 &= (1,0,1,0), & v_4 &= (1,1,1,1) \in \mathbb{R}^4. \end{aligned}$$

Dann ergibt sich zum Beispiel für den Vektor $v = (4,2,4,2) \in \mathbb{R}^4$:

$$v = v_2 - 2v_1 = 2(v_3 + v_4).$$

Die Darstellung von v als Linearkombination von $v_1, \ldots, v_4$ ist also nicht eindeutig. □

Dass im obigen Beispiel die Darstellung nicht eindeutig ist, liegt daran, dass zum Beispiel v_2 bereits im span der anderen drei Vektoren enthalten ist:

$$v_2 = 2v_1 + 2v_3 + 2v_4.$$

Daraus folgt dann

$$\operatorname{span}(v_1, \ldots, v_4) = \operatorname{span}(v_1, v_3, v_4).$$

Natürlich ist es sinnvoll, mit der geringstmöglichen Anzahl von Vektoren zu arbeiten, die einen Untervektorraum aufspannen. Wie kann man aber überprüfen, ob ein gegebenes Erzeugendensystem verkleinert werden kann? Was man leicht testen kann, ist, ob der Nullvektor nur eine Darstellung aus den gegebenen Vektoren zulässt. Dies führt zur folgenden Definition:

16.3.3 Definition

Sei V ein K-Vektorraum. Eine endliche Familie $(v_1, \ldots, v_n)$ von Vektoren aus V heißt **linear unabhängig**, wenn gilt:

$$\lambda_1, \ldots, \lambda_n \in K, \ \sum_{i=1}^{n} \lambda_i v_i = 0 \ \Rightarrow \ \lambda_1 = \ldots = \lambda_n = 0.$$

Eine beliebige Familie $(v_i)_{i \in I}$ von Vektoren heißt **linear unabhängig**, wenn je endlich viele Vektoren aus der Familie linear unabhängig sind. Andernfalls heißt die Familie **linear abhängig**. □

Tatsächlich reicht es im Hinblick auf die oben gestellte Frage, den Nullvektor auf eindeutige Darstellbarkeit zu testen:

16.3.4 Satz

Für eine Familie $(v_i)_{i\in I}$ von Vektoren eines K-Vektorraums V sind äquivalent:

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

(ii) Jeder Vektor $v \in \operatorname{span}(v_i)_{i\in I}$ lässt sich in eindeutiger Weise aus Vektoren der Familie $(v_i)_{i\in I}$ linear kombinieren.

Beweis

"(ii) $\Rightarrow$ (i)" ist klar.

"(i) $\Leftrightarrow$ (ii)" Sei

$$v = \sum_{i\in I} \lambda_i v_i = \sum_{i\in I} \mu_i v_i \in \operatorname{span}(v_i)_{i\in I}$$

auf zwei Arten linear kombiniert (dabei sind die Summen so zu verstehen, dass nur endlich viele λ_i bzw. μ_i von Null verschieden sind und tatsächlich nur über die entsprechenden Indizes summiert wird). Dann gilt

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

(wobei jetzt tatsächlich nur über die endlich vielen Indizes mit $\lambda_i \neq 0$ oder $\mu_i \neq 0$ summiert wird). Es folgt $\lambda_i = \mu_i \ \forall i \in I$ nach (i). □

Ist $(v_1, \ldots, v_n)$ eine linear unabhängige (abhängige) Familie von Vektoren, so sagen wir auch einfach: $v_1, \ldots, v_n$ sind linear unabhängig (bzw. linear abhängig).

16.3.5 Beispiel

(i) Die Vektoren $e_1, \ldots, e_n$ des K^n sind linear unabhängig.

(ii) Die Polynome $1, t, t^2, t^3, \ldots$ des Polynomrings $K[t]$ sind linear unabhängig. □

16.3.6 Lemma

Sei V ein K-Vektorraum. Dann gilt:

(i) $v \in V$ linear unabhängig $\Leftrightarrow v \neq 0$.

(ii) Ist $n \geqslant 2$, so sind $v_1, \ldots, v_n \in V$ genau dann linear abhängig, wenn einer der Vektoren Linearkombination der anderen ist.

Beweis

(i) "$\Leftarrow$" Ist v linear abhängig , so $\exists \lambda \neq 0$ mit $\lambda v = 0$. Mit 16.2.2, (iii) folgt $v = 0$.

"$\Rightarrow$" Der Nullvektor 0 ist linear abhängig wegen $1 \cdot 0 = 0$.

(ii) "⇒" Sind $v_1, \ldots, v_n$ linear abhängig, so $\exists \lambda_1, \ldots, \lambda_n \in K$ mit $\sum_{i=1}^{n} \lambda_i v_i = 0$ und für mindestens ein k gilt $\lambda_k \neq 0$. Es folgt

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

"⇐" Ist umgekehrt etwa

$$v_k = \mu_1 v_1 + \ldots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \ldots + \mu_n v_n,$$

so folgt

$$\mu_1 v_1 + \ldots + \mu_{k-1} v_{k-1} + (-1) \cdot v_k + \mu_{k+1} v_{k+1} + \ldots + \mu_n v_n = 0. \qquad \square$$

16.3.7 Beispiel

Die Vektoren $v_1, \ldots, v_4 \in \mathbb{R}^4$ in Beispiel 16.3.2, (iv) sind linear abhängig. $\square$

16.4 Basen und Dimensionen

16.4.1 Definition

Sei V ein K-Vektorraum. Eine Familie $\mathcal{B} = (v_i)_{i \in I}$ von Vektoren in V heißt **Basis** von V, wenn gilt:

(i) $\mathcal{B}$ ist ein Erzeugendensystem von V.

(ii) $\mathcal{B}$ ist linear unabhängig.

Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine endliche Basis, so heißt n **Länge der Basis**. Ist I unendlich, so sprechen wir von einer **Basis unendlicher Länge**. $\square$

Im Falle des Nullvektorraums $V = \{0\}$ ist die Definition so zu verstehen, dass die leere Familie eine Basis der Länge 0 ist.

16.4.2 Beispiel

(i) $\mathcal{K} := (e_1, \ldots, e_n)$ ist eine Basis des K^n, sie heißt die **kanonische Basis** oder **Standardbasis**.

(ii) Die Matritzen

$$E_i^j = \begin{pmatrix} & 0 & \\ & \vdots & \\ & 0 & \\ 0 \ldots 0 & 1 & 0 \ldots 0 \\ & 0 & \\ & \vdots & \\ & 0 & \end{pmatrix} \leftarrow i\text{−te Zeile}$$

$$\uparrow$$

$$j\text{−te Spalte}$$

$1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$, bilden eine Basis von $M(m \times n, K)$. Dies ist wieder nur eine Variante von (i).

(iii) $1, t, t^2, t^3, \ldots$ ist eine Basis unendlicher Länge von $K[t]$. □

16.4.3 Satz

Für eine Familie $\mathcal{B} = (v_1, \ldots, v_n)$ von Vektoren eines K-Vektorraums $V \neq \{0\}$ sind äquivalent:

(i) $\mathcal{B}$ ist eine Basis.

(ii) $\mathcal{B}$ ist ein unverkürzbares Erzeugendensystem, das heißt $\mathcal{B}$ ist ein Erzeugendensystem, aber

$$(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n)$$

ist für jedes $k \in \{1, \ldots, n\}$ kein Erzeugendensystem mehr.

(iii) $\forall v \in V \ \exists! \lambda_1, \ldots, \lambda_n \in K$ mit

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

das heißt $\mathcal{B}$ ist ein Erzeugensystem mit der zugehörigen Eindeutigkeitseigenschaft.

(iv) $\mathcal{B}$ ist linear unabhängig und $\forall v \in V$ ist $(v_1, \ldots, v_n, v)$ linear abhängig.

Beweis

"(i) ⇒ (ii)" Ist $\mathcal{B}$ eine Basis, so ist $\mathcal{B}$ ein Erzeugendensystem. Wäre $\mathcal{B}$ verkürzbar, so könnte man ein v_r als Linearkombination der anderen v_i schreiben und $\mathcal{B}$ wäre linear abhängig.

"(ii) ⇒ (iii)" Sei $\mathcal{B}$ ein unverkürzbares Erzeugendensystem. Gäbe es ein $v \in V$ mit

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

und $\lambda_k \neq \mu_k$ für ein k, so folgte

$$v_k = \frac{\mu_1 - \lambda_1}{\lambda_k - \mu_k} v_1 + \cdots + \frac{\mu_{k-1} - \lambda_{k-1}}{\lambda_k - \mu_k} v_{k-1} + \frac{\mu_{k+1} - \lambda_{k+1}}{\lambda_k - \mu_k} v_{k+1} + \cdots + \frac{\mu_n - \lambda_n}{\lambda_k - \mu_k} v_n,$$

und $\mathcal{B}$ wäre verkürzbar.

"(iii) ⇒ (iv)" Gilt (iii), so folgt aus 16.3.4: $\mathcal{B}$ ist linear unabhängig. Weiter gilt: Ist $v \in V$, so ist v von der Form $v = \sum_{i=1}^{n} \lambda_i v_i$, es folgt

$$\lambda_1 v_1 + \cdots + \lambda_n v_n + (-1)v = 0,$$

d. h. $(v_1, \ldots, v_n, v)$ ist linear abhängig.

"(iv) ⇒ (i)" Nun gelte (iv). Dann ist $\mathcal{B}$ linear unabhängig. Sei $v \in V$. Da $(v_1, \ldots, v_n, v)$ linear abhängig ist, $\exists \lambda_1, \ldots, \lambda_n, \lambda \in K, \lambda \neq 0$, mit $\lambda_1 v_1 + \cdots + \lambda_n v_n + \lambda v = 0$, es folgt

$$v = -\frac{\lambda_1}{\lambda} v_1 - \cdots - \frac{\lambda_n}{\lambda} v_n.$$

Also ist $\mathcal{B}$ auch ein Erzeugendensystem von V. □

16.4.4 Korollar

Ist V ein K-Vektorraum, der nicht endlich erzeugt ist, so gibt es eine linear unabhängige Familie von Vektoren in V mit unendlich vielen Elementen.

Beweis

Wir zeigen durch vollständige Induktion nach n: $\forall n \geq 1$ gibt es linear unabhängige Vektoren $v_1, \ldots, v_n \in V$.

- **Induktionsanfang**: $n = 1$
  - Es gibt einen Vektor $0 \neq v_1 \in V$, denn sonst wäre $V = \{0\}$ endlich erzeugt.
- **Induktionsschluss**: $n \Rightarrow n + 1$
  - **Induktionsvoraussetzung**:
    Sei n eine natürliche Zahl, für die gilt: $\exists v_1, \ldots, v_n \in V$ linear unabhängig.
  - **Induktionsbehauptung**:
    Dann gilt: $\exists v_{n+1} \in V$, so daß $v_1, \ldots, v_{n+1}$ linear unabhängig sind.

  Wir zeigen die Behauptung: Wären $v_1, \ldots, v_n, v$ für jedes $v \in V$ linear abhängig, dann wäre $v_1, \ldots, v_n$ nach 16.4.3 auch ein Erzeugendensystem von V im Widerspruch zur Voraussetzung. □

16.4.5 Basisauswahlsatz

Aus jedem endlichen Erzeugendensystem eines K-Vektorraums kann man Vektoren auswählen, die eine Basis bilden. Insbesondere hat jeder endlich erzeugte K-Vektorraum eine Basis endlicher Länge.

Beweis

Aus dem endlichen Erzeugendensystem nehme man so lange Vektoren weg, bis es unverkürzbar geworden ist. □

16.4.6 Bemerkung

Allgemeiner kann man zeigen, dass jeder K-Vektorraum eine Basis hat. Der Beweis ist wesentlich komplizierter und benutzt das sogenannte Lemma von Zorn. □

Wir wollen nun die Länge verschiedener Basen eines endlich erzeugten K-Vektorraums vergleichen. Dazu muss man systematisch Vektoren austauschen.

16.4.7 Austauschlemma von Steinitz

Seien V ein K-Vektorraum, $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und $w = \sum_{i=1}^{n} \lambda_i v_i \in V$. Dann ist für jedes $k \in \{1, \ldots, n\}$ mit $\lambda_k \neq 0$ auch

$$\mathcal{B}' = (v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$$

eine Basis von V. Man kann also v_k gegen w austauschen.

Beweis

Wir nehmen $k = 1$ und somit $\lambda_1 \neq 0$ an (sonst nummerieren wir um). Es ist also zu zeigen, dass $\mathcal{B}' = (w, v_2, \ldots, v_n)$ unter den genannten Voraussetzungen eine Basis von V ist.

$\mathcal{B}'$ erzeugt V: Sei $v \in V$. Dann $\exists \mu_1, \ldots, \mu_n \in K$ mit $v = \sum_{i=1}^{n} \mu_i v_i$. Andererseits gilt:

$$v_1 = \frac{1}{\lambda_1} w - \sum_{i=2}^{n} \frac{\lambda_i}{\lambda_1} v_i.$$

Also gilt:

$$v = \frac{\mu_1}{\lambda_1} w + \left(\mu_2 - \frac{\mu_1 \lambda_2}{\lambda_1} \right) v_2 + \ldots + \left(\mu_n - \frac{\mu_1 \lambda_n}{\lambda_1} \right) v_n.$$

$\mathcal{B}'$ ist linear unabhängig: Seien $\mu, \mu_2, \ldots, \mu_n \in K$ mit

$$\mu w + \mu_2 v_2 + \ldots + \mu_n v_n = 0.$$

Dann folgt durch Einsetzen:

$$\mu \lambda_1 v_1 + (\mu \lambda_2 + \mu_2) v_2 + \ldots + (\mu \lambda_n + \mu_n) v_n = 0$$

und somit

$$\mu \lambda_1 = \mu \lambda_2 + \mu_2 = \ldots = \mu \lambda_n + \mu_n = 0,$$

da $\mathcal{B}$ linear unabhängig ist. Da $\lambda_1 \neq 0$, folgt $\mu = 0$ und somit auch $\mu_2 = \ldots = \mu_n = 0$. □

16.4.8 Austauschsatz von Steinitz

Seien V ein K-Vektorraum, $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und $(w_1, \ldots, w_r)$ eine Familie linear unabhängiger Vektoren in V. Dann gilt:

(i) $r \leqslant n$.

(ii) Es gibt Indizes $i_1, \ldots, i_r \in \{1, \ldots, n\}$, sodass man nach Austausch von v_{i_1} durch $w_1, \ldots, v_{i_r}$ durch w_r wieder eine Basis von V erhält.

Beweis

Wir führen vollständige Induktion nach r.

- **Induktionsanfang**: $r = 1$
  - Der Fall $r = 1$ ist das gerade bewiesene Austauschlemma.
- **Induktionsschluss**: $r \Rightarrow r + 1$
  - **Induktionsvoraussetzung**:
    Sei r eine natürliche Zahl, für die die Aussage richtig ist.
  - **Induktionsbehauptung**:
    Dann ist die Aussage richtig für $r + 1$.

Wir zeigen die Behauptung: Seien $(w_1, \ldots, w_{r+1})$ eine Familie linear unabhängiger Vektoren in V. Dann ist auch $(w_1, \ldots, w_r)$ linear unabhängig. Also ergibt sich aus der Induktionsvoraussetzung, dass (evtl. nach Umnumerierung)

$$(w_1, \ldots, w_r, v_{r+1}, \ldots, v_n)$$

eine Basis von V ist. Ebenfalls gilt $r \leq n$ nach Induktionsvoraussetzung. Wäre $r = n$, so wäre $(w_1, \ldots, w_r)$ bereits eine Basis von V im Widerspruch zu 16.4.3 , (iv). Also gilt $r + 1 \leq n$. Nun schreiben wir

$$w_{r+1} = \lambda_1 w_1 + \cdots + \lambda_r w_r + \lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n$$

mit $\lambda_1, \ldots, \lambda_n \in K$. Wäre $\lambda_{r+1} = \cdots = \lambda_n = 0$, so hätte man einen Widerspruch zur linearen Unabhängigkeit von $(w_1, \ldots, w_{r+1})$. Bei geeigneter Numerierung kann man also $\lambda_{r+1} \neq 0$ annehmen und nach dem Austauschlemma v_{r+1} gegen w_{r+1} austauschen. □

16.4.9 Korollar

Hat ein K-Vektorraum V eine endliche Basis, so ist jede Basis von V endlich.

Beweis

Seien $(v_1, \ldots, v_n)$ eine endliche Basis und $(w_i)_{i \in I}$ eine beliebige Basis von V. Wäre I unendlich, so gäbe es Indizes $i_1, \ldots, i_{n+1} \in I$, sodass $w_{i_1}, \ldots, w_{i_{n+1}}$ linear unabhängig sind. Dies widerspricht 16.4.8. □

16.4.10 Korollar

Je zwei endliche Basen eines K-Vektorraums V haben gleiche Länge.

Beweis

Sind $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ zwei Basen von V, so kann man 16.4.8 zweimal anwenden und erhält $m \leqslant n$, aber auch $n \leqslant m$. □

16.4.11 Definition

Ist V ein K-Vektorraum, so heißt

$$\dim V := \dim_K V := \begin{cases} \infty, & \text{falls } V \text{ keine endliche Basis hat,} \\ n, & \text{falls } V \text{ eine Basis der Länge } n \text{ hat,} \end{cases}$$

die **Dimension** von V über K. □

16.4.12 Korollar

Seien V ein endlich erzeugter K-Vektorraum und $U \subset V$ ein Untervektorraum. Dann gilt:

(i) U ist endlich erzeugt und $\dim U \leqslant \dim V$.

(ii) Ist $\dim U = \dim V$, so folgt $U = V$.

Beweis

Folgt direkt aus 16.4.4 und 16.4.8. □

16.4.13 Basisergänzungssatz

Seien V ein endlich erzeugter K-Vektorraum und $w_1, \ldots, w_r \in V$ linear unabhängig. Dann gibt es $v_{r+1}, \ldots, v_n \in V$, sodass

$$\mathcal{B} = (w_1, \ldots, w_r, v_{r+1}, \ldots, v_n)$$

eine Basis von V ist.

Beweis

Wähle eine Basis $(v_1, \ldots, v_n)$ von V und wende 16.4.8 an. □

16.4.14 Beispiel

(i) $\dim_K K^n = n$, denn K^n hat die kanonische Basis $(e_1, \ldots, e_n)$.

(ii) $\dim_K \operatorname{Mat}(m \times n, K) = mn$.

(iii) $\dim_K K[t] = \infty$. □

16.4.15 Definition

Ist $X = v + U$ ein affiner Unterraum des K-Vektorraums V, so heißt

$$\dim X := \dim U$$

die **Dimension** von X. □

Diese Definition macht Sinn, da U eindeutig durch X bestimmt ist.

16.5 Das Eliminationsverfahren von Gauss

In 16.4.5 haben wir bewiesen, dass man aus jedem endlichen Erzeugendensystem eines K-Vektorraums eine Basis auswählen kann. Für die Praxis ist das Verfahren des Weglassens (und die Kontrolle, ob ein Erzeugendensystem übrig bleibt) nicht ökonomisch. Wir geben nun einen Algorithmus an, wie man aus einem Erzeugendensystem eine Basis linear kombinieren kann. Dabei behandeln wir den Spezialfall eines Untervektorraumes U von K^n (der allgemeine Fall lässt sich darauf zurückführen). Seien also Vektoren $u_1, \ldots, u_m \in K^n$ gegeben und sei $U = \operatorname{span}(u_1, \ldots, u_m)$. Sind $a_{i1}, \ldots, a_{in}$ die Komponenten von u_i, das heißt ist $u_i = (a_{i1}, \ldots, a_{in})$, so ergeben die Vektoren als Zeilen untereinander geschrieben eine Matrix

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

16.5.1 Beispiel

Aus der kanonischen Basis $(e_1, \ldots, e_n)$ des K^n erhält man die n-**reihige Einheitsmatrix**

$$E = (\delta_{ij}) = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \ddots & \\ 0 & \ldots & 0 & 1 \end{pmatrix} \in M(n \times n, K),$$

wobei

$$\delta_{ij} = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j. \end{cases} \qquad (\textbf{Kronecker-Symbol}).$$

□

16.5.2 Definition

Ist $A = (a_{ij}) \in M(m \times n, K)$, so heißen $a_1 = (a_{11}, \ldots, a_{1n}), \ldots, a_m = (a_{m1}, \ \ldots, a_{mn})$ auch die **Zeilen(vektoren)** von A und

$$ZR(A) := \operatorname{span}(a_1, \ldots, a_m) \subset K^n$$

der **Zeilenraum** von A. Wir setzen

$$\text{Zeilenrang } A := \dim_K ZR(A).$$

□

Unser obiges Problem können wir also so umformulieren: Bestimme eine Basis von $ZR(A)$!

Ist die Matrix in einer bestimmten Form, so kann man die Basis direkt ablesen.

16.5.3 Definition

Sei $A \in M(m \times n, K)$. Dann hat A **Zeilenstufenform**, wenn für jede Zeile von A gilt: Sind die ersten $s-1$ Elemente der Zeile Null, so sind für die folgenden Zeilen mindestens die ersten s Elemente Null (soweit vorhanden). Zum Beispiel:

$$\begin{pmatrix} \bullet & & & & & & & & & \\ 0 & \bullet & * & & & & & & & \\ 0 & 0 & 0 & \bullet & * & * & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & & & \\ & \ddots & & & & & & \ddots & & \\ & & & & & & & & \bullet & * \\ 0 & \ldots & & & & & & & 0 & 0 \\ 0 & \ldots & & & & & & & 0 & 0 \end{pmatrix}$$

Jeder mit • gekennzeichnete Eintrag ist von Null verschieden und heißt **Pivot** (der entsprechende Zeile). □

16.5.4 Bemerkung

Ist $A \in M(m \times n, K)$ in **Zeilenstufenform**, so gibt es eine Zahl r, $0 \leqslant r \leqslant m$, sodass in jeder der ersten r Zeilen mindestens ein von Null verschiedener Eintrag steht und in den anderen Zeilen nur Nullen stehen. Ist $r \geqslant 1$, so sind die ersten r Zeilen linear unabhängig, bilden also eine Basis von $ZR(A)$. Es gilt $\dim_K ZR(A) = r$. □

16.5.5 Beispiel

Für

$$A = \begin{pmatrix} 0 & 2 & 0 & 4 & 6 & 0 & 5 \\ 0 & 0 & 1 & 3 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ist $\dim_K ZR(A) = 3$. □

Ist nun A beliebig, so können wir A durch elementare Zeilenumformungen auf Zeilenstufenform bringen, ohne den Zeilenraum zu verändern.

16.5.6 Definition

Unter einer **elementaren Zeilenumformung** von $A \in M(m \times n, K)$ verstehen wir eine der folgenden Operationen:

I. Multiplikation der i-ten Zeile mit einem Skalar $\lambda \in K \setminus \{0\}$:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} \to \begin{pmatrix} \vdots \\ \lambda \cdot a_i \\ \vdots \end{pmatrix}$$

II. Addition der j-ten Zeile zur i-ten Zeile:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \to \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}$$

III. Addition der λ-fachen j-ten Zeile zur i-ten Zeile, $\lambda \in K$:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \to \begin{pmatrix} \vdots \\ a_i + \lambda \cdot a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}$$

IV. Vertauschung der i-ten Zeile mit der j-ten Zeile:

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \to \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix}$$

Dabei bezeichnen $a_1, \ldots, a_m$ die Zeilen von A, es ist stets $i \neq j$ vorausgesetzt und an den mit Punkten gekennzeichneten Zeilen ändert sich nichts. □

16.5.7 Bemerkung

Elementare Zeilenumformungen vom Typ III bzw. IV kann man aus denen vom Typ I bzw. II kombinieren:

$$\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{\text{I}}{\to} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ \lambda \cdot a_j \\ \vdots \end{pmatrix} \overset{\text{II}}{\to} \begin{pmatrix} \vdots \\ a_i + \lambda \cdot a_j \\ \vdots \\ \lambda \cdot a_j \\ \vdots \end{pmatrix} \overset{\text{I}}{\to} \begin{pmatrix} \vdots \\ a_i + \lambda \cdot a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}, \qquad \lambda \neq 0,$$

bzw.

$$\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{\text{I}}{\to} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ -a_j \\ \vdots \end{pmatrix} \overset{\text{II}}{\to} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i - a_j \\ \vdots \end{pmatrix} \overset{\text{I+II}}{\to} \begin{pmatrix} \vdots \\ a_i - (a_i - a_j) \\ \vdots \\ a_i - a_j \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i - a_j \\ \vdots \end{pmatrix} \overset{\text{II}}{\to} \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix}.$$

□

16.5.8 Lemma

Ist B aus A durch elementare Zeilenumformungen entstanden, so gilt $ZR(A) = ZR(B)$.

Beweis

Nach 16.5.7 genügt es, den Fall zu betrachten, dass B aus A durch eine einzige Zeilenumformung vom Typ I bzw. II entstanden ist. Dafür ist aber alles klar. □

16.5.9 Satz

Jede Matrix $A \in M(m \times n, K)$ kann durch elementare Zeilenumformungen auf Zeilenstufenform gebracht werden.

Beweis

Wir formen A so lange um, bis wir eine Matrix in Zeilenstufenform erhalten.

- Ist $A = 0$, so sind wir fertig.
- Ist $A \neq 0$, so gibt es mindestens eine Spalte von A, deren Einträge nicht alle Null sind. Wir wählen diejenige Spalte mit dem kleinsten Index j_1:

$$j_1 := \min \{ j \mid \exists i \text{ mit } a_{ij} \neq 0 \}.$$

Ist $a_{1j_1} \neq 0$, so können wir a_{1j_1} als Pivot wählen. Andernfalls suchen wir uns ein $a_{i_1 j_1} \neq 0$ und vertauschen die Zeile 1 mit der Zeile i_1. Die neue erste Zeile ist schon die erste Zeile von B. Also gilt für den ersten Pivot

$$b_{1j_1} = a_{i_1 j_1}.$$

Durch Umformungen vom Typ III kann man alle unterhalb von b_{1j_1} stehenden Einträge zu Null machen: Ist a einer dieser Einträge, so soll

$$a + \lambda \cdot b_{1j_1} = 0$$

werden, also wählt man

$$\lambda = \frac{-a}{b_{1j_1}}.$$

Als Ergebnis dieser Umformungen erhalten wir eine Matrix der Gestalt

$$\widetilde{A}_1 = \left(\begin{array}{ccccc|ccc} 0 & \dots & 0 & \dots & b_{1j_1} & * & \dots & * \\ \vdots & & \vdots & & 0 & & & \\ \vdots & & \vdots & & \vdots & & A_2 & \\ 0 & & 0 & & 0 & & & \end{array} \right).$$

Im nächsten Schritt macht man mit A_2 das gleiche wie im ersten Schritt mit $A = A_1$. Die dazu nötigen Zeilenumformungen kann man auf $\widetilde{A}_1$ ausdehnen, ohne dass sich in Spalten 1 bis j_1 etwas ändert, denn „links von A_2“ stehen nur Nullen. Fährt man so fort, so erhält man nach endlich vielen Schritten eine Matrix B in Zeilenstufenform. □

16.5.10 Beispiel

Wir bestimmen eine Basis des von

$$\begin{aligned} a_1 &= (0, \quad 0, \quad 0, \quad 2, -1)\,, \\ a_2 &= (0, \quad 1, -2, \quad 1, \quad 0)\,, \\ a_3 &= (0, -1, \quad 2, \quad 1, -1)\,, \\ a_4 &= (0, \quad 0, \quad 0, \quad 1, \quad 2) \end{aligned}$$

aufgespannten Untervektorraums U des $\mathbb{R}^5$:

$$A = \begin{pmatrix} 0 & 0 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \overset{IV}{\to} \begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \overset{III}{\to} \begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\overset{IV}{\to} \begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 \end{pmatrix} \overset{III}{\to} \begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix} \overset{III}{\to} \begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} =: B.$$

□

16.5.11 Bemerkung

Der Beweis von 16.5.9 liefert also einen Algorithmus zur Überführung einer Matrix in Zeilenstufenform. Diesen nennt man das **Eliminationsverfahren von Gauss**. □

Ob man Vektoren als Zeilen- oder als Spaltenvektoren schreibt, ist eine Frage der Konvention. Für eine Matrix A ist der Spaltenraum gleich dem Zeilenraum der zu A transponierten Matrix:

16.5.12 Definition

Sei $A = (a_{ij}) \in M(m \times n, K)$. Dann heißt die Matrix

$$^tA := \left(a'_{ij}\right) \in M(n \times m, K) \text{ mit } a'_{ji} := a_{ij}$$

die zu A **transponierte Matrix**. □

16.5.13 Beispiel

Es gilt

$$^t\begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{pmatrix}.$$

□

Unmittelbar aus der Definition ergibt sich:

16.5.14 Rechenregeln

Für $A, B \in M(m \times n, K)$ gilt:

(i) $^t(A + B) = {}^tA + {}^tB$.

(ii) $^t\left({}^tA\right) = A$.

(iii) $^t(\lambda A) = \lambda\, {}^tA \qquad \forall \lambda \in K$.

□

16.6 Quotienten und Summen

Zum Abschluss dieses Abschnitts wollen wir weitere Möglichkeiten studieren, wie man aus bereits bekannten Vektorräumen andere Beispiele ableiten kann.

16.6.1 Satz und Definition

Seien V ein K-Vektorraum und U ein Untervektorraum. Dann ist durch $v_1 \underset{U}{\sim} v_2 \Leftrightarrow v_1 - v_2 \in U$ eine Äquivalenzrelation auf V definiert, sodass die Äquivalenzklasse von $v \in V$ gerade der affine Unterraum $v + U$ von V ist.

Wir bezeichnen mit V/U die Menge aller Äquivalenzklassen. Dann ist durch

$$V/U \times V/U \to V/U,\ (v_1 + U, v_2 + U) \mapsto (v_1 + v_2) + U,$$

eine Addition und durch

$$K \times V/U \to V/U,\ (\lambda, v+U) \mapsto (\lambda \cdot v) + U,$$

eine Skalarmultiplikation erklärt, sodass V/U zusammen mit dieser Addition und Skalarmultiplikation ein K-Vektorraum ist. Dieser heißt **Quotientenvektorraum** oder **Faktorraum** von V nach U. Das Nullelement von V/U ist die durch das Nullelement von V definierte Äquivalenzklasse: $0_{V/U} = U$.

Beweis

Dass es sich um eine Äquivalenzrelation auf V handelt, ergibt sich direkt aus den Eigenschaften eines Untervektorraums:

- Reflexivität: $v - v = 0 \in U$,
- Symmetrie: $v_1 - v_2 \in U \Rightarrow v_2 - v_1 = -(v_1 - v_2) \in U$,
- Transitivität: $v_1 - v_2 \in U$ und $v_2 - v_3 \in U \Rightarrow v_1 - v_3 = (v_1 - v_2) + (v_2 - v_3) \in U$.

Sind $v_1, v_2 \in V$, so gilt $v_2 \underset{U}{\sim} v_1 \Leftrightarrow v_2 - v_1 \in U \Leftrightarrow v_2 \in v_1 + U$. Also sind die Äquivalenzklassen von der angegebenen Gestalt.

Als nächstes müssen wir zeigen, dass die Addition und Skalarmultiplikation wohldefiniert, d.h. vertreterunabhängig definiert sind. Wir behandeln beispielhaft die Skalarmultiplikation: Seien $v_1, v_2 \in V$ mit $v_1 + U = v_2 + U$ und $\lambda \in K$. Dann $\exists u \in U$ mit $v_1 - v_2 = u$. Es folgt

$$\lambda \cdot v_1 - \lambda \cdot v_2 = \lambda \cdot (v_1 - v_2) = \lambda \cdot u \in U$$

und somit

$$\lambda \cdot v_1 + U = \lambda \cdot v_2 + U.$$

Die Vektorraumgesetze schließlich übertragen sich von V auf V/U. □

16.6.2 Satz

Ist V endlich erzeugt, so auch V/U und es ist

$$\dim V = \dim U + \dim V/U.$$

Genauer gilt: Ist $(u_1, \ldots, u_r)$ eine Basis von U und $(u_1, \ldots, u_r, v_{r+1}, \ldots, v_n)$ eine Basis von V, so ist

$$(v_{r+1} + U, \ldots, v_n + U) =: \mathcal{B}$$

eine Basis von V/U.

Beweis

Ist V endlich erzeugt, so gilt dies auch für U nach 16.4.12 , (i). Sei $(u_1, \ldots, u_r)$ eine Basis von U. Nach dem Basisergänzungssatz 16.4.13 können wir diese zu einer Basis von V ergänzen. Sei

$$(u_1, \ldots, u_r, v_{r+1}, \ldots, v_n)$$

eine solche und $\mathcal{B}$ wie im Satz definiert. Es gilt

$$u_i + U = U = 0_{V/U}, \quad i = 1, \ldots, r.$$

$\mathcal{B}$ erzeugt V/U: Sei $v + U \in V/U$, $v \in V$. Dann gibt es $\lambda_1, \ldots, \lambda_n \in K$ mit

$$v = \sum_{i=1}^{r} \lambda_i \cdot u_i + \sum_{i=r+1}^{n} \lambda_i \cdot v_i.$$

Es folgt

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

$\mathcal{B}$ ist linear unabhängig: Seien $\lambda_{r+1}, \ldots, \lambda_n \in K$ mit

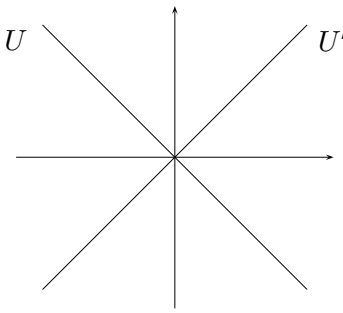
$$0_{V/U} = \sum_{i=r+1}^{n} \lambda_i \cdot (v_i + U) = \left(\sum_{i=r+1}^{n} \lambda_i \cdot v_i \right) + U.$$

Dann ist $\sum_{i=r+1}^{n} \lambda_i \cdot v_i \in U$, das heißt es gibt $\lambda_1, \ldots, \lambda_r \in K$ mit

$$\sum_{i=r+1}^{n} \lambda_i \cdot v_i = \sum_{i=1}^{r} \lambda_i \cdot u_i.$$

Da aber $(u_1, \ldots, u_r, v_{r+1}, \ldots, v_n)$ linear unabhängig sind, folgt $\lambda_i = 0 \; \forall i$. □

Im Gegensatz zu Durchschnitten ist die Vereinigung von Untervektorräumen im Allgemeinen kein Untervektorraum, man betrachte etwa zwei verschiedene Nullpunktsgeraden im $\mathbb{R}^2$:



Nach Abschnitt 16.3 kann man aber den kleinsten Untervektorraum betrachten, der alle gegebenen Untervektorräume enthält (in unserem Beispiel ist dies der ganze $\mathbb{R}^2$). Wir schreiben das für den Fall zweier Untervektorräume auf. Induktiv ergibt sich der Fall endlich vieler Untervektorräume.

16.6.3 Bemerkung und Definition

Seien V ein K-Vektorraum und $U, U' \subset V$ Untervektorräume. Dann ist

$$U + U' := \operatorname{span}\left(U \cup U'\right) = \left\{ v \in V \mid \exists u \in U, u' \in U' \text{ mit } v = u + u' \right\}$$

ein Untervektorraum von V. Dieser heißt die **Summe** von U und U'. □

16.6.4 Dimensionsformel für Summen

Sind V ein K-Vektorraum und U, U' endlichdimensionale Untervektorräume, so gilt

$$\dim\left(U + U'\right) = \dim U + \dim U' - \dim\left(U \cap U'\right).$$

Beweis

Nach 16.4.12 , (i) ist auch $U \cap U'$ endlichdimensional. Sei $u_1, \ldots, u_r$ eine Basis von $U \cap U'$. Diese können wir nach 16.4.13 zu Basen

$$(u_1, \ldots, u_r, v_{r+1}, \ldots v_n) \text{ bzw. } \left(u_1, \ldots, u_r, v'_{r+1}, \ldots v'_{n'}\right)$$

von U bzw. U' ergänzen. Die Behauptung folgt, indem wir zeigen, dass

$$\mathcal{B} := \left\{u_1, \ldots, u_r, v_{r+1}, \ldots v_n, v'_{r+1}, \ldots v'_{n'}\right\}$$

eine Basis von $U + U'$ ist. Dass $U + U'$ von $\mathcal{B}$ erzeugt wird, ist klar. Die lineare Unabhängigkeit lasse ich als Übung. □

Will man den „Korrekturterm“ dim $(U \cap U')$ loswerden, so muss man $U \cap U' = \{0\}$ fordern.

16.6.5 Definition

Ein K-Vektorraum V heißt **direkte Summe** von zwei Untervektorräumen U und U', in Zeichen

$$V = U \oplus U',$$

wenn gilt:

(i) $V = U + U'$,

(ii) $U \cap U' = \{0\}$. □

Kapitel 17

Lineare Abbildungen und Matrizen, lineare Gleichungssysteme

17.1 Motivation

Für das Studium von Mengen mit algebraischen Strukturen eines gewissen Typs ist es von großer Bedeutung, auch die strukturverträglichen Abbildungen zwischen den Mengen zu betrachten. Dies machen wir nun für Vektorräume, indem wir *lineare Abbildungen* einführen. Wir beschreiben, wie man lineare Abbildungen zwischen endlichdimensionalen K-Vektorräumen durch Matrizen darstellen kann, und stellen den Zusammenhang zu linearen Gleichungssystemen her. Gleichzeitig erklären wir, wie man solche Systeme mit Hilfe des Eliminationsverfahrens von Gauss lösen kann. □

17.2 Grundlegende Definitionen

17.2.1 Bemerkung und Definition

Eine Abbildung

$$\varphi : V \to W$$

zwischen zwei K-Vektorräumen V und W heißt **linear** (genauer ***K*-linear** oder **Homomorphismus von *K*-Vektorräumen**), wenn gilt:

(i) $\varphi(v+w) = \varphi(v) + \varphi(w) \qquad \forall v, w \in V.$

(ii) $\varphi(\lambda \cdot v) = \lambda \cdot \varphi(v) \qquad \forall \lambda \in K, v \in V.$

Äquivalent dazu ist:

(iii) $\varphi(\lambda \cdot v + \mu \cdot w) = \lambda \cdot \varphi(v) + \mu \cdot \varphi(w) \quad \forall \lambda, \mu \in K,\ \forall v, w \in V.$

Es gilt:

(iv) Die Komposition von zwei linearen Abbildungen ist linear (sofort).

Ist $\varphi : V \to W$ eine lineare Abbildung, so heißt

$$\text{Kern } \varphi = \{v \in V \mid \varphi(v) = 0\}$$

der **Kern** von φ. Weiter heißt dann φ ein **Monomorphismus** bzw. **Epimorphismus** bzw. **Isomorphismus**, wenn φ injektiv bzw. surjektiv bzw. bijektiv ist.

Im Falle $V = W$ sprechen wir von einem **Endomorphismus** von V sowie – im bijektiven Fall – von einem **Automorphismus** von V.

Es gilt:

(v) Ist $\varphi : V \to W$ ein Isomorphismus, so ist auch $\varphi^{-1} : W \to V$ ein solcher:

In der Tat ist φ^{-1} bijektiv, und sind $w, w' \in W$, so existieren $v, v' \in V$ mit $\varphi(v) = w$ bzw. $\varphi(v') = w'$, da φ surjektiv ist. Es folgt

$$\begin{aligned}\varphi^{-1}\left(w + w'\right) &= \varphi^{-1}\left(\varphi(v) + \varphi\left(v'\right)\right) \\ &= \varphi^{-1}\left(\varphi\left(v + v'\right)\right) = v + v' = \varphi^{-1}(w) + \varphi^{-1}\left(w'\right).\end{aligned}$$

Entsprechend verfährt man mit der Skalarmultiplikation.

Zwei K-Vektorräume V und W heißen **isomorph**, in Zeichen $V \cong W$, wenn es einen Isomorphismus zwischen V und W gibt. □

Wir notieren einige einfache Folgerungen aus der Definition:

17.2.2 Satz

Sei $\varphi : V \to W$ eine lineare Abbildung zwischen zwei K-Vektorräumen. Dann gilt:

(i) $\varphi(0) = 0$.

(ii) $\varphi\left(\sum_{i=1}^{n} \lambda_i \cdot v_i\right) = \sum_{i=1}^{n} \lambda_i \cdot \varphi(v_i)$.

(iii) $(v_i)_{i \in I}$ linear abhängig in $V \Rightarrow (\varphi(v_i))_{i \in I}$ linear abhängig in W.

(iv) Sind $V' \subset V$ bzw. $W' \subset W$ Untervektorräume, so auch $\varphi(V') \subset W$ und $\varphi^{-1}(W') \subset V$. Insbesondere sind

$$\text{Bild } \varphi \subset W \text{ und Kern } \varphi \subset V$$

Untervektorräume.

(v) φ injektiv $\Leftrightarrow$ Kern $\varphi = \{0\}$.

(vi) φ injektiv, $v_1, \ldots, v_n \in V$ linear unabhängig $\Rightarrow \varphi(v_1), \ldots, \varphi(v_n) \in W$ linear unabhängig.

Beweis

Aufgaben. □

17.2.3 Beispiel

(i) Die Abbildung

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^3,\ (x, y) \mapsto (x, y, x + y + 1),$$

ist nicht linear, denn

$$\varphi(0, 0) = (0, 0, 1) \neq (0, 0, 0).$$

(ii) Die Abbildung

$$\varphi : \mathbb{R}^4 \to \mathbb{R}^2,\ (w,x,y,z) \mapsto (w+2x, 3y-z)\,,$$

ist linear, denn es gilt:

$$\begin{aligned} & \varphi\left((w,x,y,z)+\left(w',x',y',z'\right)\right) \\ =\ & \varphi\left(\left(w+w',x+x',y+y',z+z'\right)\right) \\ =\ & \left(w+w'+2\left(x+x'\right), 3\left(y+y'\right)-\left(z+z'\right)\right) \\ =\ & \varphi\left(w,x,y,z\right)+\varphi\left(w',x',y',z'\right) \end{aligned}$$

sowie

$$\begin{aligned} & \varphi\left(\lambda\left(w,x,y,z\right)\right) \\ =\ & \varphi\left(\left(\lambda w,\lambda x,\lambda y,\lambda z\right)\right) \\ =\ & \left(\lambda w+2\lambda x, 3\lambda y-\lambda z\right) \\ =\ & \lambda\varphi\left(w,x,y,z\right). \end{aligned}$$

(iii) Allgemeiner rechnet man nach: Ist

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

so ist durch

$$\varphi_A : K^n \to K^m,\ x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1+\ldots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\ldots+a_{mn}x_n \end{pmatrix} =: A\cdot x,$$

eine lineare Abbildung definiert. Dabei schreiben wir die Elemente von K^n bzw. K^m als Spaltenvektoren. Setzt man für x den j-ten kanonischen Basisvektor des K^n ein, so erhält man als Bild die j-te Spalte von A. Die Spalten von A sind also die Bilder der kanonischen Basisvektoren. Die Vektoren im Kern φ_A sind gerade Lösungen des homogenen linearen Gleichungssystems

$$A\cdot x = 0.$$

Wir werden später sehen, dass jede lineare Abbildung $\varphi : K^n \to K^m$ von der Form φ_A für ein $A \in M\left(m\times n, K\right)$ ist. □

Die Lösungen eines beliebiges lineares Gleichungssystem

$$A\cdot x = b, \quad A\in M\left(m\times n, K\right),\ b\in K^m,$$

sind gerade die Vektoren in der Faser $\varphi_A^{-1}\left(\{b\}\right)$. Insbesondere gibt es genau dann mindestens eine Lösung, wenn gilt

$$b \in \text{Bild } \varphi_A.$$

17.2.4 Beispiel

Sei

$$A = \begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix} \in M(2 \times 2, \mathbb{R}).$$

Dann ist φ_A die Abbildung

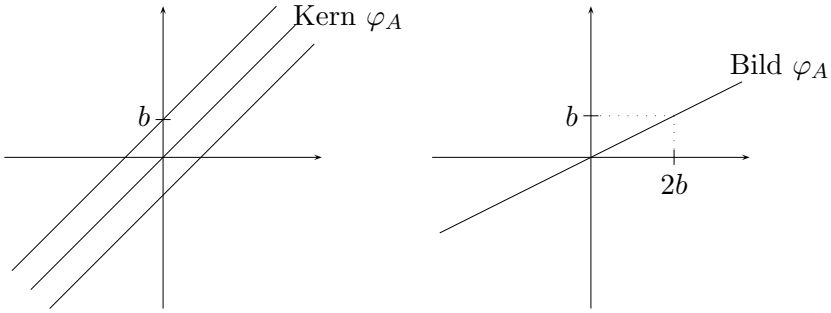
$$\varphi_A : \mathbb{R}^2 \to \mathbb{R}^2,\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -2 \cdot x_1 + 2 \cdot x_2 \\ -x_1 + x_2 \end{pmatrix}.$$

Es gilt

$$\text{Bild } \varphi_A = \text{span}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right), \quad \text{Kern } \varphi_A = \text{span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right),$$

und für $\binom{2b}{b} \in \text{Bild } \varphi_A$ ist die Faser die Gerade mit der Gleichung $x_2 = x_1 + b$, also

$$\varphi_A^{-1}\left(\left\{\begin{pmatrix} 2b \\ b \end{pmatrix}\right\}\right) = \{(\lambda, \lambda + b) \mid \lambda \in \mathbb{R}\} = \begin{pmatrix} 0 \\ b \end{pmatrix} + \text{span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right).$$



Die Fasern sind also parallele Geraden, der Kern ist die einzige Faser durch den Nullpunkt. Die Lösungen von $A \cdot x = \binom{2b}{b}$ erhält man durch Parallelverschiebung der Lösungen von $A \cdot x = \binom{0}{0}$ durch einen beliebigen Punkt der Faser

$$\varphi_A^{-1}\left(\left\{\begin{pmatrix} 2b \\ b \end{pmatrix}\right\}\right).$$

□

Allgemein gilt:

17.2.5 Lemma

Ist $\varphi : V \to W$ linear, $w \in \text{Bild } \varphi$ und $v \in \varphi^{-1}(\{w\})$ beliebig, so ist $\varphi^{-1}(\{w\})$ der affine Unterraum

$$\varphi^{-1}(\{w\}) = v + \text{Kern } \varphi = \{v + u \mid u \in \text{Kern } \varphi\}.$$

Beweis

Einfache Übung. □

17.2.6 Satz

Sei $\varphi : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W mit $\dim_K V < \infty$. Dann ist auch $\dim_K(\text{Bild}\,\varphi) < \infty$ und es gilt: Sind Basen

$$(v_1, \ldots, v_k) \text{ von } \text{Kern}\,\varphi \text{ und } \ (w_1, \ldots, w_r) \text{ von } \text{Bild}\,\varphi$$

sowie beliebige Vektoren $u_1 \in \varphi^{-1}(\{w_1\}), \ldots, u_r \in \varphi^{-1}(\{w_r\})$ gegeben, so ist

$$\mathcal{A} := (u_1, \ldots, u_r,\ v_1, \ldots, v_k)$$

eine Basis von V. Insbesondere gilt die *Dimensionsformel*

$$\dim V = \dim(\text{Bild}\,\varphi) + \dim(\text{Kern}\,\varphi)\ .$$

Beweis

Aufgaben. □

17.2.7 Korollar

Ist $\dim_K V = \dim_K W < \infty$ und $\varphi : V \to W$ K-linear, so sind äquivalent:

(i) φ injektiv.

(ii) φ surjektiv.

(iii) φ bijektiv.

Beweis

Aufgaben. □

17.3 Lineare Abbildungen und Matrizen

17.3.1 Motivation

Will man eine Abbildung $\varphi : X \to Y$ zwischen Mengen definieren, so kann man die Bilder verschiedener Elemente des Definitionsbereiches völlig unabhängig voneinander wählen. Ganz anders ist die Situation für eine lineare Abbildung $\varphi : V \to W$ zwischen Vektorräumen. Kennt man einen Wert $\varphi(v), v \neq 0$, so ist φ auf der ganzen Geraden $K \cdot v := \{\lambda \cdot v \mid \lambda \in K\}$ festgelegt. Will man für einen weiteren Vektor $v' \in V$ den Wert beliebig vorschreiben, so darf also v' nicht auf dieser Geraden $K \cdot v$ liegen. Durch $\varphi(v)$ und $\varphi(v')$ ist dann φ auf der ganzen Ebene $K \cdot v + K \cdot v'$ festgelegt und so weiter. □

Die Frage, durch wie viele Vorgaben eine lineare Abbildung festgelegt ist, hat eine einfache Antwort:

17.3.2 Satz (über die Erzeugung von linearen Abbildungen)

Gegeben seinen K-Vektorräume V und W mit $\dim_K V < \infty$ sowie Vektoren $v_1, \ldots, v_r \in V$ und $w_1, \ldots, w_r \in W$. Dann gilt:

(i) $v_1, \ldots, v_r$ linear unabhängig $\Rightarrow$ es gibt mindestens eine lineare Abbildung

$$\varphi : V \to W \text{ mit } \varphi(v_i) = w_i, \quad i = 1, \ldots, r.$$

(ii) Ist $v_1, \ldots, v_r$ sogar eine Basis, so gibt es genau eine solche Abbildung φ. Für diese gilt:

(a) Bild $\varphi = \operatorname{span}(w_1, \ldots, w_r)$.

(b) φ ist injektiv $\Leftrightarrow$ $w_1, \ldots, w_r$ sind linear unabhängig.

Beweis

(ii) Jedes $v \in V$ hat eine eindeutige Darstellung

$$v = \sum_{i=1}^{r} \lambda_i v_i, \quad \lambda_1, \ldots, \lambda_r \in K,$$

da $v_1, \ldots, v_r$ eine Basis ist. Ist φ wie gewünscht gegeben, so ist $\varphi(v)$ wegen $\varphi(v_i) = w_i$ und der Linearität von φ durch

$$\varphi(v) = \sum_{i=1}^{r} \lambda_i \varphi(v_i) = \sum_{i=1}^{r} \lambda_i w_i$$

eindeutig bestimmt. Definiert man umgekehrt φ durch diese Vorschrift, so ist φ linear. Nun zu (a) und (b):

(a) "$\subset$" Klar.
"$\supset$" Ist $w = \sum_{i=1}^{r} \mu_i w_i \in \operatorname{span}(w_1, \ldots, w_r)$, so gilt $w = \varphi(\sum_{i=1}^{r} \mu_i v_i) \in$ Bild φ.

(b) "$\Rightarrow$" Seien $w_1, \ldots, w_r$ linear abhängig. Dann

$$\exists\, (\mu_1, \ldots, \mu_r) \neq (0, \ldots, 0) \text{ mit } \sum_{i=1}^{r} \mu_i w_i = 0.$$

Also ist $\varphi(\sum_{i=1}^{r} \mu_i v_i) = 0$ und somit φ nicht injektiv, da $\sum_{i=1}^{r} \mu_i v_i \neq 0$.

"$\Leftarrow$" Seien $w_1, \ldots, w_r$ linear unabhängig. Sei $v = \sum_{i=1}^{r} \lambda_i v_i \in V$ mit $\varphi(v) = 0$. Dann ist $\sum_{i=1}^{r} \lambda_i w_i = \varphi(v) = 0$. Wegen der linearen Unabhängigkeit der w_j folgt $\lambda_1 = \ldots = \lambda_r = 0$, also $v = 0$.

(i) Wir ergänzen $(v_1, \ldots, v_r)$ zu einer Basis

$$(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$$

von V und wählen $w_{r+1}, \ldots, w_n \in W$ beliebig. Dann definieren wir φ wie im Beweis von (ii) bezüglich dieser Basis. $\square$

17.3.3 Korollar

Zwei endlichdimensionale K-Vektorräume V und W sind genau dann isomorph, wenn gilt $\dim_K V = \dim_K W$. □

17.3.4 Korollar

Ist V ein K-Vektorraum mit einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$, so gibt es genau einen Isomorphismus

$$\Phi_{\mathcal{B}} : K^n \to V \text{ mit } \Phi_{\mathcal{B}}(e_j) = v_j, \quad j = 1, \ldots, n,$$

wobei $(e_1, \ldots, e_n)$ die kanonische Basis des K^n bezeichnet. □

17.3.5 Korollar

Zu jeder linearen Abbildung $\varphi : K^n \to K^m$ gibt es genau eine Matrix $A \in M(m \times n, K)$ mit $\varphi = \varphi_A$, das heißt mit $\varphi(x) = A \cdot x$ für jeden Spaltenvektor $x \in K^n$.

Beweis

Man schreibe $\varphi(e_1), \ldots, \varphi(e_n)$ als Spaltenvektoren nebeneinander, dies ergibt A. □

Einen Zusammenhang zwischen linearen Abbildungen und Matrizen gibt es nicht nur für Vektorräume vom Typ K^n. Bevor wir das zeigen, führen wir eine Bezeichnung ein:

17.3.6 Bemerkung und Definition

(i) Sind V, W zwei K-Vektorräume, so sei

$$\mathrm{Hom}(V, W) := \mathrm{Hom}_K(V, W) := \{\varphi : V \to W \mid \varphi \text{ ist } K\text{-linear}\}.$$

Definiert man $\varphi + \psi$ bzw. $\lambda \cdot \varphi$ durch

$$(\varphi + \psi)(v) := \varphi(v) + \psi(v) \text{ bzw. } (\lambda \cdot \varphi)(v) := \lambda \cdot \varphi(v)$$

für alle $v \in V$, so wird dadurch $\mathrm{Hom}_K(V, W)$ zu einem K-Vektorraum. Der Nullvektor ist die **Nullabbildung**

$$V \to W, \ v \mapsto 0.$$

(ii) Im Spezialfall $V = W$ schreibt man

$$\mathrm{End}(V) := \mathrm{End}_K(V) := \mathrm{Hom}_K(V, V)$$

für den K-Vektorraum aller Endomorphismen von V. □

17.3.7 Satz

Gegeben seien K-Vektorräume

$$V \text{ mit Basis } \mathcal{A} = (v_1, \ldots, v_n)$$

und

$$W \text{ mit Basis } \mathcal{B} = (w_1, \ldots, w_m).$$

Dann gibt es zu jeder linearen Abbildung $\varphi : V \to W$ genau eine Matrix $A = (a_{ij}) \in M(m \times n, K)$ mit

$$\varphi(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad j = 1, \ldots, n. \tag{$*$}$$

Die durch diese Zuordnung definierte Abbildung

$$M_{\mathcal{B}}^{\mathcal{A}} : \mathrm{Hom}(V, W) \to M(m \times n, K), \ \varphi \mapsto A =: M_{\mathcal{B}}^{\mathcal{A}}(\varphi),$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis

Die a_{ij} in $(*)$ existieren, da $\mathcal{B}$ eine Basis von W ist. Aus demselben Grund sind die a_{ij} eindeutig bestimmt. Also ist $M_{\mathcal{B}}^{\mathcal{A}}$ wie angegeben definiert. Dass $M_{\mathcal{B}}^{\mathcal{A}}$ linear ist, rechnet man nach. Weiter gilt: Da auch $\mathcal{A}$ eine Basis ist, gibt es nach 17.3.2 genau ein φ, das die durch $(*)$ festgelegten Werte annimmt. Insgesamt ist $M_{\mathcal{B}}^{\mathcal{A}}$ ein Isomorphismus. □

17.3.8 Bemerkung und Definition

(i) Satz 17.3.7 bedeutet: Nach Wahl fester Basen kann man lineare Abbildungen durch Matrizen (das heißt relativ abstrakte durch konkrete Objekte) ersetzen. Man nennt $M_{\mathcal{B}}^{\mathcal{A}}(\varphi)$ auch **die Matrix, die φ bezüglich den Basen $\mathcal{A}$ und $\mathcal{B}$ darstellt**.

(ii) Im Spezialfall $V = K^n$ und $W = K^m$ mit den kanonischen Basen $\mathcal{K}$ bzw. $\mathcal{K}'$ ist der kanonische Isomorphismus

$$M_{\mathcal{K}'}^{\mathcal{K}} : \mathrm{Hom}(K^n, K^m) \to M(m \times n, K)$$

die in Korollar 17.3.4 und 17.2.3, (iii) beschriebene Beziehung.

(iii) Den in 16.4.2 beschriebenen Basisvektoren E_i^j vom $M(m \times n, K)$ entsprechen bezüglich $M_{\mathcal{B}}^{\mathcal{A}}$ die durch

$$F_i^j(v_k) := \begin{cases} w_j, & k = i \\ 0, & \text{sonst} \end{cases}$$

definierten Basisvektoren F_i^j von $\mathrm{Hom}(V, W)$. □

Als Folgerung aus 17.2.6 erhält man, dass bei Benutzung der dort konstruierten Basen die darstellende Matrix besonders einfach ist:

17.3.9 Satz

Seien $\varphi : V \to W$ linear, $n = \dim V, m = \dim W$ und $r = \dim \text{Bild } \varphi$. Dann existieren Basen $\mathcal{A}$ und $\mathcal{B}$ von V bzw. W mit

$$M_{\mathcal{B}}^{\mathcal{A}}(\varphi) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Dabei ist E_r die in 16.5.1 eingeführte r-reihige Einheitsmatrix.

Beweis

Wir wählen eine Basis $\mathcal{A}$ von V wie in 17.2.6 und ergänzen die dort gewählte Basis von Bild φ zu einer Basis $\mathcal{B}$ von W. Bezüglich dieser Basis hat $M_{\mathcal{A}}^{\mathcal{B}}(\varphi)$ die angegebene Form. □

17.4 Das Produkt von Matrizen

Eine naheliegende Frage ist, wie sich die Beziehung zwischen linearen Abbildungen und Matrizen verhält im Zusammenhang mit der Komposition von Abbildungen.

Wir betrachten zuerst den Spezialfall

$$K^r \overset{\varphi_B}{\to} K^n \overset{\varphi_A}{\to} K^m.$$

Also:

$$B \in M(n \times r, K), A \in M(m \times n, K),$$

und

$$\varphi_B : K^r \to K^n,\ y \mapsto B \cdot y,$$

sowie

$$\varphi_A : K^n \to K^m,\ x \mapsto A \cdot x.$$

Insbesondere stimmt die Spaltenzahl von A mit der Zeilenzahl von B überein. Die Komposition $\varphi_A \circ \varphi_B$ ist dann die Abbildung

$$\begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix} \underset{\varphi_B}{\to} \begin{pmatrix} b_{11} & \dots & b_{1r} \\ \vdots & & \\ b_{n1} & \dots & b_{nr} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^r b_{1k} y_k \\ \vdots \\ \sum_{k=1}^r b_{nk} y_k \end{pmatrix} =: \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}$$

$$\underset{\varphi_A}{\to} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^n a_{1j} \left(\sum_{k=1}^r b_{jk} y_k \right) \\ \vdots \\ \sum_{j=1}^n a_{mj} \left(\sum_{k=1}^r b_{jk} y_k \right) \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^r \left(\sum_{j=1}^n a_{1j} b_{jk} \right) y_k \\ \vdots \\ \sum_{k=1}^r \left(\sum_{j=1}^n a_{mj} b_{jk} \right) y_k \end{pmatrix}$$

$$=: \begin{pmatrix} c_{11} & \dots & c_{1r} \\ \vdots & & \vdots \\ c_{m1} & \dots & c_{mr} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix},$$

das heißt es gilt:

$$\varphi_A \circ \varphi_B = \varphi_C$$

mit

$$C := (c_{ik}) = \left(\sum_{j=1}^{n} a_{ij} \cdot b_{jk} \right).$$

17.4.1 Definition

Sind

$$A = (a_{ij}) \in M(m \times n, K), \quad B = (b_{jk}) \in M(n \times r, K),$$

so ist das **Produkt** von A und B definiert durch

$$A \cdot B := (c_{ik}) \in M(m \times r, K) \text{ mit } c_{ik} = \sum_{j=1}^{n} a_{ij} \cdot b_{jk}.$$

□

17.4.2 Bemerkung

Damit $A \cdot B$ gebildet werden kann, muss also die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmen. Im Ergebnis hat $A \cdot B$ hat so viele Zeilen wie A und so viele Spalten wie B. Für die Abbildungen

$$K^r \overset{\varphi_B}{\to} K^n \overset{\varphi_A}{\to} K^m$$

gilt

$$\varphi_A \circ \varphi_B = \varphi_{A \cdot B}.$$

□

17.4.3 Beispiel

(i)

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ 0 & 1 \\ -1 & -3 \end{pmatrix}.$$

(ii) Ist speziell $m = r = 1$ und n beliebig, so gilt

$$(a_1, \ldots, a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \left(\sum_{j=1}^{n} a_j \cdot b_j \right) \in M(1 \times 1, K) = K.$$

(iii) Andererseits ist

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot (b_1, \ldots, b_n) = \begin{pmatrix} a_1 \cdot b_1 & \ldots & a_1 \cdot b_n \\ \vdots & & \vdots \\ a_n \cdot b_1 & \ldots & a_n \cdot b_n \end{pmatrix} \in M(n \times n, K).$$

(iv) Ist speziell $m = n = r$, so kann man sowohl $A \cdot B$ als auch $B \cdot A$ bilden. Die Matrizenmultiplikation ist aber nicht kommutativ: Im Allgemeinen gilt $A \cdot B \neq B \cdot A$. Zum Beispiel:

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

Dieses Beispiel zeigt auch, dass das Produkt von zwei Matrizen Null sein kann, obwohl jede einzelne Matrix von Null verschieden ist. □

17.4.4 Bemerkung

Seien Matrizen $A, A' \in M(m \times n, K)$, $B, B' \in M(n \times r, K)$, $C \in M(r \times s, K)$, sowie $\lambda \in K$ gegeben. Dann gilt

(i) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (**Assoziativgesetz**)

(ii) $E_m \cdot A = A \cdot E_n = A$ (**Neutralität der Einheitsmatrix**)

(iii) $A \cdot (B + B') = A \cdot B + A \cdot B'$
$(A + A') \cdot B = A \cdot B + A' \cdot B$ (**Distributivgesetze**)

(iv) $A \cdot (\lambda \cdot B) = (\lambda \cdot A) \cdot B = \lambda \cdot (A \cdot B)$

(v) ${}^t(A \cdot B) = {}^tB \cdot {}^tA.$

Beweis

Nachrechnen. □

Im Falle $m = n$ sprechen wir auch von einer **quadratischen Matrix**. Jede solche Matrix A definiert einen Homomorphismus $\varphi_A : K^n \to K^n$. Wir klären nun, was es für die Matrix bedeutet, dass φ_A ein Isomorphismus ist:

17.4.5 Definition

Eine quadratische Matrix $A \in M(n \times n, K)$ heißt **invertierbar**, wenn $\exists A' \in M(n \times n, K)$ mit

$$A \cdot A' = A' \cdot A = E_n.$$

In diesem Fall schreiben wir auch

$$A^{-1} = A'$$

und nennen A^{-1} die zu A **inverse Matrix**. □

Aus den Rechenregeln in 17.4.4 ergibt sich, dass $M((n \times n, K)$ zusammen mt der Addition und Multiplikation von Matrizen ein Ring ist. Man beachte, dass das inverse Element A^{-1} nach 5.2.2 , (iii) im Falle der Existenz eindeutig bestimmt ist.

17.4.6 Beispiel

(i) Ist

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, K)$$

invertierbar, so gilt

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

(ii) Die Matrix

$$A = \begin{pmatrix} 5 & 10 & 20 \\ 1 & 1 & 1 \\ 12 & 12 & 20 \end{pmatrix} \in M(3 \times 3, \mathbb{Q})$$

ist invertierbar mit

$$A^{-1} = \begin{pmatrix} -\frac{1}{5} & -1 & \frac{1}{4} \\ \frac{1}{5} & \frac{7}{2} & -\frac{3}{8} \\ 0 & -\frac{3}{2} & \frac{1}{8} \end{pmatrix}.$$

□

17.4.7 Bemerkung

(i) Sind die folgenden Größen definiert, so gilt:

$$\left(A^{-1}\right)^{-1} = A, \quad (A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

(ii) Sei $A \in M(n \times n, K)$. Dass A invertierbar ist, bedeutet, dass die lineare Abbildung

$$\varphi_A : K^n \to K^n$$

ein Isomorphismus ist. In der Tat gilt

$$\varphi_A \circ \varphi_{A^{-1}} = \varphi_{A \cdot A^{-1}} = \varphi_{E_n} = id_{K^n}$$

und

$$\varphi_{A^{-1}} \circ \varphi_A = \varphi_{A^{-1} \cdot A} = \varphi_{E_n} = id_{K^n},$$

das heißt $\varphi_{A^{-1}}$ ist die Umkehrabbildung zu φ_A. □

17.4.8 Lemma

Für $A \in M(n \times n, K)$ sind äquivalent:

(i) A ist invertierbar.

(ii) ${}^tA$ ist invertierbar.

(iii) Spaltenrang $A = n$.

(iv) Zeilenrang ${}^tA = n$.

Sind diese Bedingungen erfüllt, so gilt:

$$\left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right).$$

Beweis

"(i) $\Rightarrow$ (ii)" und der Zusatz folgen aus

$$ {}^t\left(A^{-1}\right) \cdot {}^tA = {}^t\left(A \cdot A^{-1}\right) = {}^tE_n = E_n $$

und

$$ {}^tA \cdot {}^t\left(A^{-1}\right) = {}^t\left(A^{-1} \cdot A\right) = {}^tE_n = E_n. $$

"(ii) $\Rightarrow$ (i)" ergibt sich durch Transposition.

"(i) $\Leftrightarrow$ (iii)" gilt wegen

$$ \begin{aligned} A \text{ invertierbar} &\Leftrightarrow \varphi_A \text{ Isomorphismus} \\ &\Leftrightarrow \text{Spaltenrang } A = n, \end{aligned} $$

denn die Spaltenvektoren von A bilden ja eine Basis von Bild φ_A.

"(ii) $\Leftrightarrow$ (iv)" ergibt sich wieder durch Transposition. □

17.4.9 Satz und Definition

Die Menge

$$ GL(n, K) := \{A \in M(n \times n, K) \mid A \text{ ist invertierbar}\} $$

mit der Multiplikation von Matrizen als Verknüpfung ist eine Gruppe mit neutralem Element E_n. Sie heißt die **allgemeine lineare Gruppe**.

Beweis

Wir zeigen, dass $GL(n, K)$ bezüglich der Matrizenmultiplikation abgeschlossen ist:

Sind A, B invertierbar, so auch $A \cdot B$ mit inverser Matrix $B^{-1} \cdot A^{-1}$:

$$ \begin{aligned} (A \cdot B) \cdot \left(B^{-1} \cdot A^{-1}\right) &= A \cdot \left(B \cdot B^{-1}\right) \cdot A^{-1} = A \cdot E_n \cdot A^{-1} = E_n, \\ \left(B^{-1} \cdot A^{-1}\right) \cdot (A \cdot B) &= B^{-1} \cdot \left(A^{-1} \cdot A\right) \cdot B = B^{-1} \cdot E_n \cdot B = E_n. \end{aligned} $$

Dabei haben wir das Assoziativgesetz benutzt, das sogar in $M(n \times n, K)$ gilt, und wir haben benutzt, dass sich E_n neutral verhält. □

17.5 Die Berechnung der inversen Matrix

Wir wollen nun das Gauss'sche Eliminationsverfahren benutzen, um zu entscheiden, ob eine gegebene Matrix $A \in M(n \times n, K)$ invertierbar ist, und um gegebenenfalls A^{-1} zu berechnen. Zur theoretischen Begründung ist es dabei hilfreich, Matrizenumformungen zu interpretieren als Multiplikation mit speziellen invertierbaren Matrizen, den **Elementarmatrizen**. Genauer bewirkt die Multiplikation von links mit Elementarmatrizen elementare Zeilenumformungen, die Multiplikation von rechts elementare Spaltenumformungen.

17.5.1 Definition

Sind $m \in \mathbb{N}, 1 \leqslant i, j \leqslant m$ und $\lambda \in K^* = K \setminus \{0\}$, so nennt man die folgenden quadratischen Matrizen **Elementarmatrizen**:

$$S_i(\lambda) := \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ i\text{-te Spalte} \\ \downarrow \\ \\ \\ \end{matrix}$$

Im folgenden sei $i \neq j$.

$$Q_i^j := \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & 1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ j\text{-te Spalte} \\ \downarrow \\ \\ \\ \end{matrix}$$

$$Q_i^j(\lambda) := \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & \lambda & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ j\text{-te Spalte} \\ \downarrow \\ \\ \\ \end{matrix}$$

$$P_i^j := \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 0 & \dots & & \dots & 1 & & \\ & & & \vdots & 1 & & & \vdots & & \\ & & & & & \ddots & & & & \\ & & & \vdots & & & 1 & \vdots & & \\ & & & 1 & \dots & & \dots & 0 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{−te Zeile} \\ \\ \\ \\ \leftarrow j\text{−te Zeile} \\ \\ \\ \end{matrix}$$

Außer den eingetragenen oder durch Punkte angedeuteten sind dabei alle Koeffizienten gleich Null. □

17.5.2 Bemerkung

Seien $A \in M(m \times n, K)$ und $\lambda \in K^*$ gegeben.

(i) Durch elementare Zeilenumformungen wie in 16.5.6 erhält man aus A die folgenden Matrizen:

A_I	durch Multiplikation der i-ten Zeile mit λ,
A_{II}	durch Addition der j-ten Zeile zur i-ten Zeile,
A_{III}	durch Addition des λ-fachen der j-ten Zeile zur i-ten Zeile,
A_{IV}	durch Vertauschen der i-ten Zeile und der j-ten Zeile.

Dann gilt:

$$\begin{aligned} A_I &= S_i(\lambda) \cdot A, \\ A_{II} &= Q_i^j \cdot A, \\ A_{III} &= Q_i^j(\lambda) \cdot A, \\ A_{IV} &= P_i^j \cdot A. \end{aligned}$$

(ii) Für die analogen elementaren Spaltenumformungen von A gilt:

$$\begin{aligned} A^I &= A \cdot S_i(\lambda), \\ A^{II} &= A \cdot Q_i^j, \\ A^{III} &= A \cdot Q_i^j(\lambda), \\ A^{IV} &= A \cdot P_i^j. \end{aligned}$$

□

17.5.3 Bemerkung

(i) Entsprechend 16.5.7 sind die Elementarmatrizen $Q_i^j(\lambda)$ und P_i^j Produkte von Elementarmatrizen vom Typ $S_i(\lambda)$ und Q_i^j:

$$Q_i^j(\lambda) = S_j\left(\frac{1}{\lambda}\right) \cdot Q_i^j \cdot S_j(\lambda)$$

und

$$P_i^j = Q_j^i \cdot Q_i^j(-1) \cdot Q_j^i \cdot S_j(-1).$$

(ii) Die Elementarmatrizen sind invertierbar und ihre Inversen sind wieder Elementarmatrizen:

$$\begin{aligned}(S_i(\lambda))^{-1} &= S_i\left(\frac{1}{\lambda}\right),\\ \left(Q_i^j\right)^{-1} &= Q_i^j(-1),\\ \left(Q_i^j(\lambda)\right)^{-1} &= Q_i^j(-\lambda),\\ \left(P_i^j\right)^{-1} &= P_i^j.\end{aligned}$$

Dies überprüft man durch Ausmultiplizieren. □

17.5.4 Satz

Jede invertierbare Matrix $A \in M(n \times n, K)$ ist endliches Produkt von Elementarmatrizen. Wir sagen deswegen auch, dass die Gruppe $GL(n,K)$ *von den Elementarmatrizen erzeugt wird.*

Beweis

Nach 17.4.8 gilt Zeilenrang $A = n$. Überführt man also A mit Hilfe des Gauss'schen Verfahrens 16.5.9 in eine Matrix B in Zeilenstufenform, so muss B eine **obere Dreiecksmatrix** mit von Null verschiedenen Diagonalelementen sein, das heißt B ist von der Form

$$B = \begin{pmatrix} b_{11} & \dots & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & b_{nn} \end{pmatrix}$$

mit $b_{ii} \neq 0$ für alle i. Nach 17.5.2 existieren Elementarmatrizen $B_1, \dots, B_r$ mit

$$B = B_r \cdot \ldots \cdot B_1 \cdot A.$$

Man kann nun B durch weitere elementare Zeilenumformungen zur Einheitsmatrix E_n machen: Zunächst mache man $b_{1n}, \dots, b_{(n-1)n}$ mit Hilfe der letzten Zeile zu Null, dann $b_{1(n-1)}, \dots, b_{(n-2)(n-1)}$ mit Hilfe der vorletzten Zeile und so weiter. Schließlich normiere man die Koeffizienten in der Diagonalen auf 1. Es gibt also weitere Elementarmatrizen $B_{r+1}, \dots, B_s$, so dass gilt

$$B_s \cdot \ldots \cdot B_1 \cdot A = B_s \cdot \ldots \cdot B_{r+1} \cdot B = E_n.$$

Mit 17.5.3 folgt, dass

$$A = B_1^{-1} \cdot \ldots \cdot B_s^{-1}$$

ein Produkt von Elementarmatrizen ist. □

Aus dem Beweis ergibt sich auch

$$A^{-1} = B_s \cdot \ldots \cdot B_1 = B_s \cdot \ldots \cdot B_1 \cdot E_n$$

und somit ein einfaches Rechenverfahren zur Bestimmung der inversen Matrix:

17.5.5 Algorithmus (Berechnung der inversen Matrix)

Ist $A \in M(n \times n, K)$ gegeben, so schreibt man die Matrizen A und E_n nebeneinander. Alle Zeilenumformungen, die im folgenden an A vorgenommen werden, führt man gleichzeitig an E_n durch.

Zunächst bringt man A durch elementare Zeilenumformungen auf Zeilenstufenform. Dadurch stellt sich insbesondere heraus, ob gilt

$$\text{Zeilenrang } A = n,$$

das heißt, ob A überhaupt invertierbar ist. Im Falle Zeilenrang $A < n$ kann man aufhören: A ist nicht invertierbar, und die Umformungen mit E_n waren umsonst.
Andernfalls führt man weitere Zeilenumformungen durch, bis aus A die Matrix E_n entstanden ist:

A	E_n
$B_1 \cdot A$	$B_1 \cdot E_n$
$\vdots$	$\vdots$
$B_s \cdot \ldots \cdot B_1 \cdot A$	$B_s \cdot \ldots \cdot B_1 \cdot E_n$

Ist links E_n entstanden, so steht rechts A^{-1}. □

17.5.6 Beispiel

Sei $K = \mathbb{R}$.

(i)

A			E_n		
1	0	1	1	0	0
0	−1	0	0	1	0
1	1	1	0	0	1
1	0	1	1	0	0
0	−1	0	0	1	0
0	1	0	−1	0	1
1	0	1	1	0	0
0	−1	0	0	1	0
0	0	0	−1	1	1

A ist nicht invertierbar, wegen

$$\text{Zeilenrang } A = 2 < 3.$$

(ii)

A invertierbar $\Leftarrow$

A			E_n		
0	1	−4	1	0	0
1	2	−1	0	1	0
1	1	2	0	0	1
1	2	−1	0	1	0
0	1	−4	1	0	0
1	1	2	0	0	1
1	2	−1	0	1	0
0	1	−4	1	0	0
0	−1	3	0	−1	1
1	2	−1	0	1	0
0	1	−4	1	0	0
0	0	−1	1	−1	1
1	2	−1	0	1	0
0	1	−4	1	0	0
0	0	1	−1	1	−1
1	2	0	−1	2	−1
0	1	−4	1	0	0
0	0	1	−1	1	−1
1	2	0	−1	2	−1
0	1	0	−3	4	−4
0	0	1	−1	1	−1
1	0	0	5	−6	7
0	1	0	−3	4	−4
0	0	1	−1	1	−1

$= A^{-1}$

Man berechne zur Kontrolle $A \cdot A^{-1}$.

17.6 Transformationsmatrizen

Wir betrachten die Frage, wie sich die eine lineare Abbildung darstellende Matrix ändert, wenn man andere Basen betrachtet. Dies erlaubt uns dann auch, die Frage nach Matrizen und Komposition von linearen Abbildungen zwischen beliebigen Vektorräumen zu beantworten. Außerdem werden wir uns überlegen, dass Zeilen- und Spaltenrang einer Matrix übereinstimmen.

Da wir es im folgenden mit mehreren Abbildungen zwischen verschiedenen Mengen zu tun haben, ist es übersichlicher, die Abbildungen und Mengen in einem **Diagramm** anzuordnen. Stimmen dabei zu je zwei Mengen alle auftretenden Abbildungen (auch zusammengesetzte und mehrfach zusammengesetzte), die die eine Menge in die andere abbilden, überein, so nennen wir das Diagramm **kommutativ**.

17.6.1 Bemerkung und Definition

Sei V ein Vektorraum mit einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$. Dann gibt es nach 17.3.4 genau einen Isomorphismus

$$\Phi_{\mathcal{B}} : K^n \to V \text{ mit } \Phi_{\mathcal{B}}(e_j) = v_j, \quad j = 1, \ldots, n,$$

wobei $(e_1, \ldots, e_n)$ die kanonische Basis des K^n ist. Für $x = (x_1, \ldots, x_n) \in K^n$ gilt dann

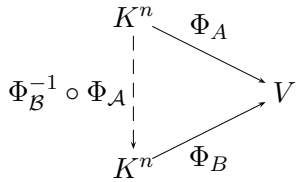
$$\Phi_{\mathcal{B}}\left((x_1, \ldots, x_n)\right) = \sum_{i=1}^{n} x_i v_i =: v.$$

Man nennt $\Phi_{\mathcal{B}}$ das durch $\mathcal{B}$ bestimmte **Koordinatensystem** in V, $x_1, \ldots, x_n$ die **Koordinaten** von v bezüglich $\mathcal{B}$ und

$$x = (x_1, \ldots, x_n) = \Phi_{\mathcal{B}}^{-1}(v) \in K^n$$

den **Koordinatenvektor** von v bezüglich $\mathcal{B}$. □

Koordinatensysteme erlauben es uns, das Rechnen mit Vektoren in V auf das Rechnen in K^n und damit letztendlich auf das Rechnen in K zurückzuführen. Für Anwendungen ist es wichtig, dass man verschiedene Koordinaten ineinander umrechnen kann. Seien dazu zwei Basen $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_n)$ von V gegeben. Dann hat man ein kommutatives Diagramm von Isomorphismen:



Da $\Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}}$ ein Isomorphismus von K^n auf K^n ist, gibt es eine Matrix $T_{\mathcal{B}}^{\mathcal{A}} \in GL(n, K)$ mit

$$\Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}} = \varphi_{T_{\mathcal{B}}^{\mathcal{A}}},$$

das heißt eine invertierbare Matrix $T_{\mathcal{B}}^{\mathcal{A}}$, die $\Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}}$ bezüglich der kanonischen Basis des K^n darstellt.

17.6.2 Definition

$T_{\mathcal{B}}^{\mathcal{A}}$ heißt die **Transformationsmatrix des Basiswechsels** von $\mathcal{A}$ und $\mathcal{B}$. □

17.6.3 Bemerkung

$T_{\mathcal{B}}^{\mathcal{A}}$ hat per Definition die folgende Eigenschaft: Ist

$$v = x_1 v_1 + \ldots + x_n v_n = y_1 w_1 + \ldots + y_n w_n,$$

so ist

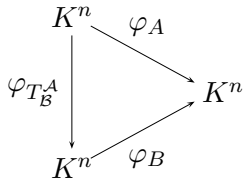
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_{\mathcal{B}}^{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

□

Kennt man also $T_{\mathcal{B}}^{\mathcal{A}}$, so kann man die neuen aus den alten Koordinaten berechnen. Wie aber bestimmt man $T_{\mathcal{B}}^{\mathcal{A}}$?

17.6.4 Bemerkung

(i) Wir betrachten zunächst den Spezialfall $V = K^n$. Sind A bzw. B die Matrizen mit den Vektoren aus $\mathcal{A}$ bzw. $\mathcal{B}$ als Spalten, so sind A und B invertierbar, und es gilt $\Phi_{\mathcal{A}} = \varphi_A$ und $\Phi_{\mathcal{B}} = \varphi_B$. Also sieht obiges kommutative Diagramm wie folgt aus:



Somit gilt:

$$\varphi_{T_{\mathcal{B}}^{\mathcal{A}}} = \varphi_B^{-1} \circ \varphi_A = \varphi_{B^{-1}} \circ \varphi_A = \varphi_{B^{-1}A},$$

das heißt $T_{\mathcal{B}}^{\mathcal{A}}$ ist das Produkt $T_{\mathcal{B}}^{\mathcal{A}} = B^{-1}A$.

(ii) Ist V beliebig, so schreibt man die Vektoren aus $\mathcal{B} = (w_1, \ldots, w_n)$ als Linearkombinationen der Vektoren aus $\mathcal{A} = (v_1, \ldots, v_n)$:

$$w_j = \sum_{i=1}^{n} s_{ij} v_i \text{ mit } s_{ij} \in K.$$

Setzt man dann $S = (s_{ij})$, so ist S invertierbar und es gilt

$$T_{\mathcal{B}}^{\mathcal{A}} = S^{-1}$$

(nachrechnen). □

17.6.5 Beispiel

Im $\mathbb{R}^2$ seien $\mathcal{A} = \mathcal{K} = (e_1, e_2)$ und

$$\mathcal{B} = (w_1, w_2) = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right).$$

Dann gilt in 17.6.4, (i):

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

Es folgt

$$B^{-1} = \frac{1}{5} \cdot \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}$$

und somit

$$T_{\mathcal{B}}^{\mathcal{A}} = B^{-1}A = B^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}.$$

Nun kann man Koordinaten ineinander umrechnen. Betrachtet man etwa den Vektor mit den Koordinaten $x = \binom{-1}{-1}$ bezüglich $\mathcal{A} = (e_1, e_2)$, das heißt den Vektor $v = -e_1 - e_2$, so hat v bezüglich $\mathcal{B}$ die Koordinaten

$$y = B^{-1}x = \begin{pmatrix} -\frac{2}{5} \\ -\frac{1}{5} \end{pmatrix}.$$

Also gilt

$$v = -\frac{2}{5}w_1 - \frac{1}{5}w_2.$$

□

Den Zusammenhang zwischen Koordinatensystem und Matrizen liefert:

17.6.6 Lemma

Seien $\varphi : V \to W$ eine lineare Abbildung, $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_m)$ Basen von V bzw. W und $A := M_{\mathcal{B}}^{\mathcal{A}}(\varphi)$. Dann ist das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_{\mathcal{A}}} & V \\ {\scriptstyle \varphi_A}\downarrow & & \downarrow{\scriptstyle \varphi} \\ K^m & \xrightarrow[\Phi_{\mathcal{B}}]{} & W \end{array}$$

kommutativ, das heißt es gilt

$$\Phi_{\mathcal{B}} \circ \varphi_A = \varphi \circ \Phi_{\mathcal{A}}.$$

Beweis

Sei $A = (a_{ij})$. Um die Gleichheit der Abbildungen zu zeigen, genügt es, die Bilder der kanonischen Basisvektoren e_j des K^n zu vergleichen (siehe Satz 17.3.2). Für diese Bilder gilt:

$$\begin{aligned} \Phi_{\mathcal{B}}(\varphi_A(e_j)) = \Phi_{\mathcal{B}}(a_{1j}, \ldots, a_{mj}) &= \sum_{i=1}^{m} a_{ij} w_i \\ \overset{17.3.7}{=} \varphi(v_j) &= \varphi(\Phi_{\mathcal{A}}(e_j)) \ . \end{aligned}$$

□

Nun zur Komposition von linearen Abbildungen:

17.6.7 Satz

Gegeben seien K-Vektorräume U, V, W mit Basen $\mathcal{A}, \mathcal{B}, \mathcal{C}$ der Längen r, n, m sowie lineare Abbildungen

$$U \overset{\psi}{\to} V \overset{\varphi}{\to} W.$$

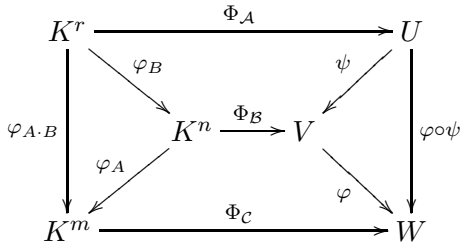
Dann gilt

$$M_{\mathcal{C}}^{\mathcal{A}}(\varphi \circ \psi) = M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot M_{\mathcal{B}}^{\mathcal{A}}(\psi).$$

Beweis

Nach 17.4.2 ist die Aussage richtig für Vektorräume vom Typ K^n zusammen mit den kanonischen Basen. Mit $A = M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$ und $B = M_{\mathcal{B}}^{\mathcal{A}}(\psi)$ folgt daraus die allgemeine Behauptung

durch Betrachten des nach 17.4.2 und 17.6.6 kommutativen Diagramms



. □

17.6.8 Beispiel

Durch

$$\psi(x_1, x_2) = (2x_1,\ x_1 + x_2, 0)$$

und

$$\varphi(y_1, y_2, y_3) = (y_1 + y_2,\ y_1 + y_3,\ y_2 + y_3,\ y_1 + y_2 + y_3)$$

sind lineare Abbildungen

$$\mathbb{R}^2 \xrightarrow{\psi} \mathbb{R}^3 \xrightarrow{\varphi} \mathbb{R}^4$$

definiert. Wir wählen die Basen

$$\mathcal{A} = (u_1, u_2) = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right) \qquad \text{von } \mathbb{R}^2,$$

$$\mathcal{B} = \mathcal{K} = (e_1, e_2, e_3) \qquad \text{von } \mathbb{R}^3,$$

$$\mathcal{C} = (w_1, w_2, w_3, w_4) = \left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix}\right) \qquad \text{von } \mathbb{R}^4.$$

Da $\mathcal{B}$ die kanonische Basis des $\mathbb{R}^3$ ist, sind die Spalten der Matrix

$$B := M_{\mathcal{B}}^{\mathcal{A}}(\psi)$$

die Bilder der Basisvektoren u_1, u_2. Wegen

$$\psi(u_1) = \begin{pmatrix}2\\1\\0\end{pmatrix} \text{ und } \psi(u_2) = \begin{pmatrix}2\\2\\0\end{pmatrix}$$

gilt also

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

Wir berechnen

$$A := M_{\mathcal{C}}^{\mathcal{B}}(\varphi):$$

Es ist

$$\varphi(e_1) = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = 0 \cdot w_1 + 1 \cdot w_2 - 1 \cdot w_3 + 1 \cdot w_4,$$

$$\varphi(e_2) = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = 1 \cdot w_1 - 1 \cdot w_2 + 0 \cdot w_3 + 1 \cdot w_4,$$

$$\varphi(e_3) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = -1 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 + 1 \cdot w_4,$$

also

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

Analog erhält man

$$C := M_{\mathcal{C}}^{\mathcal{A}}(\varphi \circ \psi) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ -2 & -2 \\ 3 & 4 \end{pmatrix},$$

denn es gilt

$$(\varphi \circ \psi)(u_1) = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 3 \end{pmatrix} = 1 \cdot w_1 + 1 \cdot w_2 - 2 \cdot w_3 + 3 \cdot w_4,$$

$$(\varphi \circ \psi)(u_2) = \begin{pmatrix} 4 \\ 2 \\ 2 \\ 4 \end{pmatrix} = 2 \cdot w_1 + 0 \cdot w_2 - 2 \cdot w_3 + 4 \cdot w_4.$$

Tatsächlich gilt

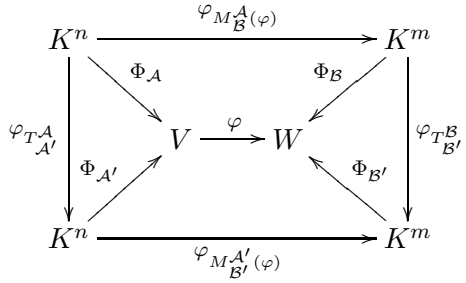
$$A \cdot B = C.$$

□

Als nächstes beantworten wir die Frage, wie sich die darstellende Matrix bei Einführung neuer Basen ändert:

17.6.9 Satz (Transformationsformel)

Ist $\varphi : V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen und sind Basen $\mathcal{A}, \mathcal{A}'$ von V bzw. $\mathcal{B}, \mathcal{B}'$ von W gegeben, so ist das Diagramm



kommutativ. Insbesondere gilt für die beteiligten Matrizen

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\varphi) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(\varphi) \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}.$$

Anders ausgedrückt: Sind

$$A = M_{\mathcal{B}}^{\mathcal{A}}(\varphi) \text{ und } B = M_{\mathcal{B}'}^{\mathcal{A}'}(\varphi)$$

die beiden Matrizen, die φ bezüglich der Basen $\mathcal{A}, \mathcal{B}$ bzw. $\mathcal{A}', \mathcal{B}'$ darstellen, und sind

$$T = T_{\mathcal{A}'}^{\mathcal{A}}, \quad S = T_{\mathcal{B}'}^{\mathcal{B}}$$

die entsprechenden Transformationsmatrizen zwischen den Basen, so gilt

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

Beweis

Das Diagramm ist aufgrund der Definition der Transformationsmatrix und nach 17.6.6 kommutativ. □

Nun können wir Zeilenrang und Spaltenrang einer Matrix vergleichen:

17.6.10 Satz

Für jede Matrix $A \in M(m \times n, K)$ gilt

$$\text{Zeilenrang } A = \text{Spaltenrang } A.$$

Beweis

Wir betrachten die lineare Abbildung

$$\varphi_A : K^n \to K^m,\ x \mapsto A \cdot x,$$

und wählen Basen $\mathcal{A}$ und $\mathcal{B}$ von K^n bzw. K^m wie in 17.3.9, also so, dass gilt

$$B := M_{\mathcal{B}}^{\mathcal{A}}(\varphi_A) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Für die Matrix B gilt offensichtlich

$$\text{Zeilenrang } B = r = \text{Spaltenrang } B.$$

Um zu zeigen, dass sich diese Gleichheit auf A überträgt, wählen wir gemäß der Transformationsformel invertierbare Matrizen S und T mit

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

Die Behauptung ergibt sich dann aus dem folgenden Lemma. □

17.6.11 Lemma

Für $A \in M(m \times n, K),\ S \in \mathrm{GL}(m, K)$ und $T \in \mathrm{GL}(n, K)$ gilt

(i) Spaltenrang SAT^{-1} = Spaltenrang A .

(ii) Zeilenrang SAT^{-1} = Zeilenrang A .

Beweis

Zu den gegebenen Matrizen gehört ein kommutatives Diagramm linearer Abbildungen

$$\begin{array}{ccc} K^n & \xrightarrow{\varphi_A} & K^m \\ {\scriptstyle \varphi_{T^{-1}}}\uparrow & & \downarrow{\scriptstyle \varphi_S} \\ K^n & \xrightarrow[\varphi_{SAT^{-1}}]{} & K^m \end{array} \quad .$$

Nach 17.4.7 , (ii) sind φ_S und $\varphi_{T^{-1}}$ Isomorphismen, es folgt

$$\begin{aligned} \text{Spaltenrang } A &= \dim \text{Bild}\, \varphi_A = \dim \text{Bild} \varphi_S \circ \varphi_A \circ \varphi_{T^{-1}} \\ &= \dim \text{Bild} \varphi_{SAT^{-1}} = \text{Spaltenrang } SAT^{-1} \ , \end{aligned}$$

das heißt es gilt (i). Daraus ergibt sich (ii) durch Transposition, denn

$$\text{Zeilenrang } A = \text{Spaltenrang } {}^tA$$

und

$${}^t(SAT^{-1}) = {}^t(T^{-1}) \cdot {}^tA \cdot {}^tS \ .$$

□

Wie man Transformationsmatrizen S und T wie oben aus A berechnen kann, werden wir gleich sehen.

17.6.12 Bemerkung und Definition

(i) Ist $A \in M\,(m \times n, K)$, so heißt

$$\text{rang } A := \text{Zeilenrang } A = \text{Spaltenrang } A$$

der **Rang** von A.

(ii) Ist $\varphi : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W, so heißt

$$\operatorname{rang} \varphi := \dim \operatorname{Bild} \varphi$$

der **Rang** von φ.

(iii) Sind $\mathcal{A}$ und $\mathcal{B}$ Basen der endlichdimensionalen K-Vektorräume V bzw. W, und ist $A = M_{\mathcal{B}}^{\mathcal{A}}(\varphi)$, so gilt

$$\operatorname{rang} A = \operatorname{rang} \varphi.$$

Dies folgt durch Betrachten des kommutativen Diagramms

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_{\mathcal{A}}} & V \\ {\scriptstyle \varphi_A}\downarrow & & \downarrow{\scriptstyle \varphi} \\ K^m & \xrightarrow[\Phi_{\mathcal{B}}]{} & W \end{array}$$

□

Ist $A \in M(m \times n, K)$, und ist $r = \operatorname{rang} A$, so gibt es also Transformationsmatrizen $S \in GL(m, K)$ und $T \in GL(n, K)$ mit

$$SAT^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Wir leiten nun ein Rechenverfahren zur Bestimmung von S und T ab.

17.6.13 Algorithmus (Berechnung von Transformationsmatrizen)

Sei $A \in M(m \times n, K)$. Wir betrachten folgendes Schema:

E_m	A	
$B_1 \cdot E_m$	$B_1 \cdot A$	
$\vdots$	$\vdots$	
$B_k \cdot \ldots \cdot B_1 \cdot E_m$	$B_k \cdot \ldots \cdot B_1 \cdot A$	E_n
	$B_k \cdot \ldots \cdot B_1 \cdot A \cdot C_1$	$E_n \cdot C_1$
	$\vdots$	$\vdots$
	$B_k \cdot \ldots \cdot B_1 \cdot A \cdot C_1 \cdot \ldots \cdot C_l$	$E_n \cdot C_1 \cdot \ldots \cdot C_l$

Zunächst wird A durch elementare Zeilenumformungen, die man gleichzeitig an E_m durchführt, auf Zeilenstufenform gebracht. Die Zeilenumformungen entsprechen der Multiplikation von links mit m-reihigen Elementarmatrizen $B_1, \ldots, B_k$.

Da die Matrix $B_k \cdot \ldots \cdot B_1 \cdot A$ Zeilenstufenform hat, kann man sie durch elementare Spaltenumformungen, die man gleichzeitig an E_n durchführt, auf die Form

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

mit $r = \operatorname{rang} A$ bringen. Dies entspricht der Multiplikation von rechts mit n-reihigen Elementarmatrizen $C_1, \ldots, C_l$. Wegen

$$B_k \cdot \ldots \cdot B_1 \cdot A \cdot C_1 \cdot \ldots \cdot C_l = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

sind durch

$$S = B_k \cdot \ldots \cdot B_1 = B_k \cdot \ldots \cdot B_1 \cdot E_m$$

und

$$T^{-1} = C_1 \cdot \ldots \cdot C_l = E_n \cdot C_1 \cdot \ldots \cdot C_l$$

die gewünschten Transformationsmatrizen gefunden. □

17.6.14 Beispiel

Für $K = \mathbb{R}$ und $A = \left(\begin{smallmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \end{smallmatrix}\right)$ ergibt sich:

$S =$

		A					
1	0	1	2	0			
0	1	2	2	1			
1	0	1	2	0	1	0	0
−2	1	0	−2	1	0	1	0
					0	0	1
		1	0	2	1	0	0
		0	1	−2	0	0	1
					0	1	0
		1	0	0	1	0	−2
		0	1	−2	0	0	1
					0	1	0
		1	0	0	1	0	−2
		0	1	0	0	0	1
					0	1	2

$= T^{-1}$

□

Schreibt man in der Situation von 17.6.13

$$D = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

so erhält man auch Basen $\mathcal{A}$ und $\mathcal{B}$ von K^n bzw. K^m, bezüglich derer die Abbildung

$$\varphi_A : K^n \to K^m,\ x \mapsto A \cdot x,$$

durch D beschrieben wird. Dazu betrachten wir das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{\varphi_D} & K^m \\ {\scriptstyle \varphi_{T^{-1}}}\downarrow & & \downarrow{\scriptstyle \varphi_{S^{-1}}} \\ K^n & \xrightarrow[\varphi_A]{} & K^m \end{array},$$

das wegen

$$D = SAT^{-1}$$

kommutativ ist. Nach 17.6.6 erhält man die Vektoren von $\mathcal{A}$ und $\mathcal{B}$ als Bilder der kanonischen Basisvektoren von K^n bzw. K^m unter den Isomorphismen $\varphi_{T^{-1}}$ und $\varphi_{S^{-1}}$. Also erhält man $\mathcal{A}$ und $\mathcal{B}$ als Spaltenvektoren von T^{-1} und S^{-1}. Dazu muss man S noch invertieren:

17.6.15 Beispiel

In 17.6.14 ist

$$S^{-1} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix},$$

also sind

$$\mathcal{A} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} \right), \quad \mathcal{B} = \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

die gesuchten Basen. Zur Kontrolle prüft man nach:

$$A \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A \cdot \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

□

17.7 Lineare Gleichungenssysteme

Nun wollen wir das Lösen linearer Gleichungssysteme systematisch studieren.

17.7.1 Satz und Definition

Sei

$$A \cdot x = b, \quad A = (a_{ij}) \in M(m \times n, K), \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m \qquad (*)$$

ein lineares Gleichungssystem. Dann heißt

$$A \cdot x = 0 \qquad (**)$$

das zu $(*)$ gehörige **homogene System**. Ist $b \neq 0$, so nennt man das System $(*)$ **inhomogen**. Ist $r = \operatorname{rang} A$ und setzt man

$$\text{Lös}(A, b) = \{x \in K^n \mid A \cdot x = b\} \subset K^n,$$

so gilt:

(i) $\text{Lös}(A, 0) \subset K^n$ ist ein Untervektorraum der Dimension $n - r$.

(ii) $\text{Lös}(A, b) \subset K^n$ ist entweder leer oder ein affiner Unterraum der Dimension $n - r$. Ist $v \in \text{Lös}(A, b)$ beliebig, so gilt

$$\text{Lös}(A, b) = v + \text{Lös}(A, 0).$$

Wir sprechen auch von den **Lösungsräumen** $\text{Lös}(A, b)$ bzw. $\text{Lös}(A, 0)$.

Beweis

Ist

$$\varphi_A : K^n \to K^m,\ x \mapsto A \cdot x,$$

so gilt Lös $(A, 0) =$ Kern φ_A. Aus der Dimensionsformel 17.2.6 folgt (i):

$$\dim \text{Lös}\,(A, 0) = \dim \text{Kern}\ \varphi_A = \dim K^n - \dim \text{Bild}\ \varphi_A = n - r.$$

Dann folgt (ii) wegen Lös $(A, b) = \varphi_A^{-1}(b)$ aus 17.2.5. □

Man erhält also alle Lösungen des inhomogenen Systems $(*)$, indem man zu einer speziellen Lösung von $(*)$ alle Lösungen des homogenen Systems $(**)$ addiert.

17.7.2 Definition

Das lineare Gleichungssystem

$$A \cdot x = b$$

heißt

(i) **lösbar**, wenn es mindestens eine Lösung gibt,

(ii) **eindeutig lösbar**, wenn es genau eine Lösung gibt. □

Jedes homogene System ist lösbar, denn es gibt ja mindestens die **triviale Lösung** $x = 0$. Im allgemeinen Fall erhält man ein Kriterium für die Lösbarkeit bzw. eindeutige Lösbarkeit durch Betrachten der Matrix $(A \mid b)$, die aus A durch Hinzunahme des Spaltenvektors b entsteht.

17.7.3 Satz

Für das lineare Gleichungssystem

$$A \cdot x = b,\ A = (a_{ij}) \in M(m \times n, K),\ b \in K^m \qquad (*)$$

gilt:

(i) $(*)$ lösbar $\Leftrightarrow$ rang $A =$ rang $(A \mid b)$.

(ii) $(*)$ eindeutig lösbar $\Leftrightarrow$ rang $A =$ rang $(A \mid b) = n$.

Beweis

Wir schreiben $A' = (A \mid b)$ und betrachten die lineare Abbildungen

$$\varphi_A : K^n \to K^m,\ x \mapsto A \cdot x,$$
$$\varphi'_A : K^{n+1} \to K^m,\ x' \mapsto A' \cdot x',$$

mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ x' = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \end{pmatrix}.$$

Dann gilt für die kanonischen Basen $(e_1, \ldots, e_n)$ und $\left(e'_1, \ldots, e'_n, e'_{n+1}\right)$ des K^n bzw. K^{n+1}:

$$\varphi_A(e_1) = \varphi_{A'}\left(e'_1\right), \ldots, \varphi_A(e_n) = \varphi_{A'}\left(e'_n\right) \text{ und } \varphi_{A'}\left(e'_{n+1}\right) = b.$$

Insbesondere ist stets Bild $\varphi_A \subset$ Bild $\varphi_{A'}$ und somit

$$\text{rang } A \leqslant \text{rang } A'.$$

Also gilt:

$$\text{rang } A = \text{rang } (A \mid b) \Leftrightarrow \text{Bild } \varphi_A = \text{Bild } \varphi_{A'} \Leftrightarrow b \in \text{Bild } \varphi_A \Leftrightarrow (*) \text{ lösbar.}$$

Dies ergibt (i).

Ist nun $(*)$ lösbar oder äquivalent rang $A =$ rang $(A \mid b)$, so gilt:

$(*)$ ist eindeutig lösbar

$\Leftrightarrow$ das zugehörige homogene System hat nur die triviale Lösung $x = 0$

$\Leftrightarrow \varphi_A$ ist ein Monomorphismus

$\Leftrightarrow$ rang $A = \dim \text{Bild} \varphi_A = n$.

Dies ergibt (ii). □

17.7.4 Bemerkung

Dass bei vorgegebenem $A \in M(m \times n, K)$ für jedes $b \in K^m$ mindestens eine Lösung von

$$A \cdot x = b$$

existiert, bedeutet, dass φ_A surjektiv ist, oder äquivalent, dass gilt rang $A = m$. □

17.7.5 Bemerkung

Ist

$$(A \mid b) = \left(\begin{array}{ccc|c} \bullet & & & \\ & \bullet & & \\ & & \ddots & \\ & & \bullet & \\ & & & b_{r+1} \\ & & & \vdots \\ & & & b_m \end{array}\right)$$

in Zeilenstufenform, so gilt:

$$\text{rang } A = \text{rang } (A \mid b) \Leftrightarrow b_{r+1} = \ldots = b_m = 0.$$

□

Zum expliziten Lösen eines linearen Gleichungssystems kann man in zwei Schritten vorgehen.

Im ersten Schritt bringt man A durch elementare Zeilenumformungen auf Zeilenstufenform, wobei man b entsprechend umformt. Dabei verändert sich, wie wir nun zeigen werden, der Lösungsraum nicht.

17.7.6 Lemma

Ist $S \in GL(m, K)$, so gilt

$$\text{Lös}(A, b) = \text{Lös}(SA, Sb).$$

Beweis

Es gilt:

$$x \in \text{Lös}(A, b) \Leftrightarrow A \cdot x = b \Leftrightarrow S \cdot A \cdot x = S \cdot b \Leftrightarrow x \in \text{Lös}(SA, Sb).$$

□

17.7.7 Satz

Geht (A, b) aus $(\tilde{A}, \tilde{b})$ durch elementare Zeilenumformungen hervor, so gilt

$$\text{Lös}(A, b) = \text{Lös}\left(\tilde{A}, \tilde{b}\right).$$

Beweis

Jede elementare Zeilenumformung entspricht nach 17.5.2 der Multiplikation von links mit einer Elementarmatrix. Elementarmatrizen sind aber invertierbar nach 17.5.3 , (ii). □

Im zweiten Lösungschritt kann man von einem Gleichungssystem $A \cdot x = b$ ausgehen, in dem $(A \mid b)$ bereits in Zeilenstufenform ist.

17.7.8 Algorithmus

Sei

$$(A \mid b) = \left(\begin{array}{ccccc|c} \bullet & & & & & \\ & \bullet & & & & \\ & & \bullet & & & \\ & & & \ddots & & \\ & & & & \bullet & \\ & & & & & b_{r+1} \\ & & & & & \vdots \\ & & & & & b_m \end{array}\right)$$

mit A in Zeilenstufenform. Gibt es ein $i,\ r+1 \leqslant i \leqslant m$, mit $b_i \neq 0$, so ist $A \cdot x = b$ nicht lösbar. Andernfalls sei für jedes $i,\ 1 \leqslant i \leqslant r$, mit j_i der Spaltenindex des Pivot-Elements der

i-ten Zeile bezeichnet. Das zu lösende Gleichungssystem sieht dann wie folgt aus:

$$\begin{aligned} x_{j_1}\overset{0}{\overset{\shortparallel}{a_{1j_1}}} + \quad \ldots \quad + x_n a_{1n} &= b_1 \\ \ddots & \\ x_{j_r}\overset{0}{\overset{\shortparallel}{a_{rj_r}}} + \ldots + x_n a_{rn} &= b_r \end{aligned}$$

Wir schreiben $J = \{j_1, \ldots, j_r\}$ und nennen jede Variable $x_j,\ j \in \{1, \ldots, n\} \backslash J$, **freie Variable** und jede Variable $x_j,\ j \in J$, **gebundene Variable**.
Für jede der $n-r$ freien Variablen x_j kann man ein beliebiges Element von K einsetzen. Man erhält dann einen Lösungsvektor, indem man für jede der gebundenen Variablen x_j den nach den folgenden Rekursionsformeln berechneten Wert einsetzt:

$$x_{j_k} = \frac{1}{a_{kj_k}} \left(b_k - \sum_{j>j_k} a_{kj} x_j \right), \quad k = r, \ldots, 1.$$

Durchläuft man beim Einsetzen für die $n-r$ freien Variablen den ganzen Vektorraum K^{n-r}, so durchlaufen die Lösungen den ganzen $(n-r)$-dimensionalen affinen Lösungsraum des Systems. □

17.7.9 Beispiel

Wir lösen das lineare Gleichungssystem

$$\begin{aligned} 3x_1 + 4x_2 + 5x_3 &= 9 \\ 6x_1 + 7x_2 + 8x_3 &= 9 \\ 6x_1 + 6x_2 + 6x_3 &= 0. \end{aligned}$$

Überführung in Zeilenstufenform liefert:

$$\left(\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 6 & 7 & 8 & 9 \\ 6 & 6 & 6 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 0 & -1 & -2 & -9 \\ 0 & -2 & -4 & -18 \end{array}\right) \to \left(\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 0 & -1 & -2 & -9 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Es ergibt sich also das lineare Gleichungssystem

$$\begin{aligned} 3x_1 + 4x_2 + 5x_3 &= 9 \\ -x_2 - 2x_3 &= -9\ , \end{aligned}$$

welches lösbar ist. Hier ist x_3 eine freie Variable und kann beliebig gewählt werden. Die allgemeine Lösung ergibt sich rekursiv:

$$x_2 = 9 - 2x_3$$

$$x_1 = \frac{1}{3}(9 - 4x_2 - 5x_3) = \frac{1}{3}(9 - 36 + 8x_3 - 5x_3) = -9 + x_3.$$

Diese Lösungen können wir dann in der Form

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} -9 \\ 9 \\ 0 \end{pmatrix}}_{\substack{| \\ \text{spezielle} \\ \text{Lösung}}} + \lambda \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \lambda \in K,$$

schreiben. □

Wir wollen nun noch einen Schritt weiter gehen und eine systematische Methode angeben, wie man eine Basis von Lös$(\tilde{A}, 0)$ und eine spezielle Lösung von $\tilde{A} \cdot x = \tilde{b}$ bekommt. Dazu muss man die erweiterte Matrix $(\tilde{A}, \tilde{b})$ durch elementare Zeilenumformungen auf **Zeilenstufennormalform** (A, b) bringen. Nach 17.7.7 gilt wieder

$$\text{Lös}\,(A, b) = \text{Lös}\left(\tilde{A}, \tilde{b}\right) \text{ bzw. Lös}\,(A, 0) = \text{Lös}\left(\tilde{A}, 0\right).$$

17.7.10 Definition

Sei $B = (b_{ij})$ eine Matrix in Zeilenstufenform und seien $j_1, \ldots, j_r$ die Spaltenindizes der Pivot-Elemente. Dann heißt B in **Zeilenstufennormalform**, wenn gilt:

(i) Alle Pivot-Elemente sind 1, das heißt es gilt

$$b_{ij_i} = 1, \quad i = 1, \ldots, r.$$

(ii) Oberhalb der Pivot-Elemente stehen nur Nullen, das heißt es gilt:

$$b_{kj_i} = 0, \quad i = 1, \ldots, r, \ k = 1, \ldots, i-1.$$

□

17.7.11 Bemerkung

Wir wissen bereits, dass jede Matrix durch elementare Zeilenumformungen in Zeilenstufenform gebracht werden kann. Es ist aber auch klar, dass man jede Matrix in Zeilenstufenform durch elementare Zeilenumformungen in Zeilenstufennormalform überführen kann. Man kann zeigen: Zu jeder Matrix A gibt es genau eine Matrix in Zeilenstufennormalform, in die sich A durch elementare Zeilenumformungen überführen lässt (siehe Kowalski-Michler, Lineare Algebra, Satz 4.1.31). □

17.7.12 Beispiel

$$\left(\begin{array}{ccc|c} 3 & 4 & 5 & 9 \\ 0 & -1 & -2 & -9 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & \frac{4}{3} & \frac{5}{3} & 3 \\ 0 & -1 & -2 & -9 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|c} 1 & 0 & -1 & -9 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

□

17.7.13 Satz

Seien $(A \mid b) = (a_{ij} \mid b_i) \in M(m \times (n+1), K)$ mit A in Zeilenstufennormalform, $r = \operatorname{rang} A$, j_i der Spaltenindex des Pivot-Elements in der i-ten Zeile, $i = 1, \ldots, r$, und $J = \{j_1, \ldots, j_r\}$. Dann gilt für das lineare Gleichungssystem

$$A \cdot x = b \tag{$*$}$$

und das zugehörige homogene System

$$A \cdot x = 0 : \tag{$**$}$$

(i) $(*)$ hat keine Lösung $\Leftrightarrow$ $b_i \neq 0$ für ein $i \in \{r+1, \ldots, m\}$.

(ii) Ist $b_i = 0$ für $i = r+1, \ldots, m$, und ist $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in K^n$ definiert durch

$$v_\ell := \begin{cases} b_i, & \text{falls } \ell = j_i, \\ 0, & \text{sonst,} \end{cases}$$

so ist v eine spezielle Lösung von $(*)$.

(iii) $\forall k \in \{1, \ldots, n\} \setminus J$ sei $u^{(k)} = \begin{pmatrix} u_1^{(k)} \\ \vdots \\ u_n^{(k)} \end{pmatrix} \in K^n$ definiert durch

$$u_\ell^{(k)} := \begin{cases} a_{ik}, & \text{falls } \ell = j_i, \\ -1, & \text{falls } \ell = k, \\ 0, & \text{sonst.} \end{cases} .$$

Dann ist

$$\mathcal{B} := \left(u^{(k)} \mid k \in \{1, \ldots, n\} \setminus J \right)$$

eine Basis von $\operatorname{Lös}(A, 0)$. Insgesamt gilt:

$$\operatorname{Lös}(A, b) = v + \operatorname{span}\left(u^{(k)} \mid k \in \{1, \ldots, n\} \setminus J \right).$$

Beweis

(i) ist wieder klar nach 17.7.3 , (i).

(ii) $\forall i \in \{1, \ldots, r\}$ gilt

$$\sum_{\ell=1}^{n} a_{i\ell} v_\ell = a_{ij_i} b_i = b_i \ .$$

Also ist v eine spezielle Lösung von $(*)$.

(iii) $\forall k \in \{1,\dots,n\} \setminus J,\ \forall i \in \{1,\dots,r\}$ gilt

$$\sum_{\ell=1}^{n} a_{i\ell} u_\ell^k = -a_{ik} + a_{ij_i} a_{ik} = 0.$$

Also sind alle $u^{(k)}$ Lösungen von $A \cdot x = 0$. Die $u^{(k)}$ sind auch linear unabhängig: Sei

$$\sum_{k \in \{1,\dots,n\} \setminus J} \lambda_k u^{(k)} = 0,\ \lambda_k \in K.$$

Dann gilt $\forall \ell \in \{1,\dots,n\} \setminus J$:

$$0 = \sum_{k \in \{1,\dots,n\} \setminus J} \lambda_k u_\ell^{(k)} = \lambda_\ell,$$

also sind alle $\lambda_k = 0$. Wegen $\operatorname{rang} A = r$ gibt es genau $n-r$ linear unabhängige Lösungen von $(**)$ (siehe 17.7.1, (i)) . Also bilden die $u^{(k)}$ eine Basis von $\text{Lös}(A, 0)$. □

17.7.14 Algorithmus

Seien $(A \mid b), r$ und $J = \{j_1, \dots, j_r\}$ wie in 17.7.13 .

Ist $b_i \neq 0$ für ein $i \in \{r+1, \dots, m\}$, so hat $A \cdot x = b$ keine Lösung.

Andernfalls führe man in der Matrix $(A \mid b)$ Nullzeilen ein, so dass die Pivot-Elemente in der neuen Matrix in der Diagonalen stehen. Das heißt, das Pivot-Element der i-ten Zeile steht an der Stelle (j_i, j_i).

Durch weiteres Anhängen bzw. Streichen von Nullzeilen erhält man eine $n \times (n+1)$-Matrix.

Dann ersetze man die Nullen an den Stellen $(k,k), k \in \{1,\dots,n\} \setminus J$ durch -1.

Sei $C = \left(c^{(1)}, \dots, c^{(n+1)}\right)$ die so entstandene Matrix. Dann ist $c^{(n+1)}$ eine spezielle Lösung von $A \cdot x = b$, und die Vektoren $c^{(k)},\ k \in \{1,\dots,n\} \setminus J$, bilden eine Basis von Lös $(A, 0)$.

Insgesamt gilt:

$$\text{Lös}\,(A,b) = c^{(n+1)} + \operatorname{span}\left(c^{(k)} \mid k \in \{1,\dots,n\} \setminus J\right).$$

□

17.7.15 Beispiel

(i) Noch einmal 17.7.9 bzw. 17.7.12:

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 0 & -1 & -9 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & -1 & -9 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & -1 & 0 \end{array}\right)$$

Es folgt

$$\text{Lös}\,(A,b) = \begin{pmatrix} -9 \\ 9 \\ 0 \end{pmatrix} + \operatorname{span}\left(\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}\right).$$

(ii)

$$(A \mid b) = \left(\begin{array}{ccccc|c} 1 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & 5 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

ist bereits in Zeilenstufennormalform. Unser Verfahren ergibt:

$$\begin{aligned}(A \mid b) &\to \left(\begin{array}{ccccc|c} 1 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \\ &\to \left(\begin{array}{ccccc|c} 1 & 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{array}\right) \\ &\to \left(\begin{array}{ccccc|c} 1 & 2 & 0 & -1 & 0 & 2 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 & -2 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{array}\right)\end{aligned}$$

Es folgt

$$\mathrm{Lös}\,(A, b) = \begin{pmatrix} 2 \\ 0 \\ -2 \\ 0 \\ 7 \end{pmatrix} + \operatorname{span}\left(\begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 5 \\ -1 \\ 0 \end{pmatrix}\right).$$

(iii)

$$\begin{aligned}(A \mid b) = \left(\begin{array}{cccc|c} 1 & 0 & 0 & 3 & 1 \\ 0 & 1 & 2 & 1 & 5 \end{array}\right) &\to \left(\begin{array}{cccc|c} 1 & 0 & 0 & 3 & 1 \\ 0 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \\ &\to \left(\begin{array}{cccc|c} 1 & 0 & 0 & 3 & 1 \\ 0 & 1 & 2 & 1 & 5 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right)\end{aligned}$$

Es folgt

$$\mathrm{Lös}\,(A, b) = \begin{pmatrix} 1 \\ 5 \\ 0 \\ 0 \end{pmatrix} + \operatorname{span}\left(\begin{pmatrix} 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \\ -1 \end{pmatrix}\right).$$

□

Kapitel 18

Determinanten

18.1 Motivation

In der Motivation 16.1 haben wir bereits die Determinante einer 2×2-Matrix eingeführt. Nun wollen wir erklären, wie die Determinante einer beliebigen $n \times n$-Matrix definiert ist. Dazu gibt es insbesondere die folgenden beiden Möglichkeiten:

- Wir geben eine Formel für die Determinante an, die von den Einträgen der Matrix abhängt, also etwa
$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$
- Wir charakterisieren die Determinante als Abbildung
$$\det : M(n \times n, K) \to K,\ A \mapsto \det A,$$
durch ihre grundlegenden Eigenschaften. Dann muss man zeigen, dass eine Abbildung mit diesen Eigenschaften existiert, und dass sie durch die Eigenschaften eindeutig bestimmt ist. Und man muss angeben, wie man Determinanten ausrechnet. □

Wir gehen hier nach der zweiten Möglichkeit vor. Ist dabei $A \in M(n \times n, K)$, so bezeichnen wir wie üblich mit $a_1, \ldots, a_n$ die Zeilenvektoren von A und schreiben
$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

18.2 Grundlegende Definitionen

18.2.1 Definition

Seien K ein Körper und $n \in \mathbb{N}$. Dann heißt eine Abbildung
$$\det : M(n \times n, K) \to K,\ A \mapsto \det A,$$
Determinante, falls gilt:

D1 det ist linear in jeder Zeile, das heißt $\forall i \in \{1, \ldots, n\}$ gilt:

(i) Ist $a_i = a_i' + a_i''$, so ist

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i' \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i'' \\ \vdots \end{pmatrix}.$$

(ii) Ist $\lambda \in K$, so ist

$$\det \begin{pmatrix} \vdots \\ \lambda \cdot a_i \\ \vdots \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix}.$$

Dabei stehen an den mit „…“ gekennzeichneten Stellen unverändert die Zeilenvektoren $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$.

D2 det ist **alternierend**, das heißt hat A zwei gleiche Zeilen, so ist

$$\det A = 0.$$

D3 det ist **normiert**, das heißt

$$\det E_n = 1.$$

Ist

$$A = (a_{ij}) \in M(n \times n, K),$$

so schreiben wir auch

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

und nennen diesen Skalar die **Determinante** von A. □

18.2.2 Satz (Rechenregeln für Determinanten)

Eine Determinante

$$\det : M(n \times n, K) \to K$$

hat die folgenden Eigenschaften:

D4 $\forall \lambda \in K$ ist $\det(\lambda \cdot A) = \lambda^n \cdot \det A$.

D5 Ist eine Zeile von A gleich Null, so ist $\det A = 0$.

D6 Entsteht B aus A durch Vertauschung von zwei Zeilen, so ist

$$\det B = -\det A.$$

D7 Ist $\lambda \in K$ und entsteht B aus A durch Addition der λ-fachen j-ten Zeile zur i-ten Zeile, $i \neq j$, so ist

$$\det B = \det A.$$

D8 Ist A eine obere Dreiecksmatrix, also A von der Form

$$A = \begin{pmatrix} \lambda_1 & \cdots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix},$$

so ist

$$\det A = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

D9 Ist $n \geqslant 2$ und $A \in M(n \times n, K)$ von der Gestalt

$$A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$$

mit quadratischen A_1 und A_2, so gilt

$$\det A = (\det A_1) \cdot (\det A_2).$$

D10 Es gilt

$$\det A \neq 0 \;\Leftrightarrow\; \text{rang } A = n \;\Leftrightarrow\; A \text{ ist invertierbar.}$$

D11 Es gilt

$$\det(A \cdot B) = \det A \cdot \det B \quad \forall A, B \in M(n \times n, K).$$

Insbesondere gilt:

$$\det A^{-1} = (\det A)^{-1} \quad \forall A \in GL(n, K).$$

Beweis

D4, D5 folgen sofort aus D1, (ii).

D6 Wir nehmen an, dass wir die Zeilen a_i und a_j, $i < j$, vertauschen. Dann gilt

$$\begin{aligned} \det A + \det B &= \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} \\ &\overset{\text{D2}}{=} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \\ &\overset{\text{D1, (i)}}{=} \det \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} \overset{\text{D2}}{=} 0. \end{aligned}$$

D7 gilt wegen

$$\det B = \det \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{D1}{=} \det A + \lambda \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{D2}{=} \det A \ .$$

D8 Sind alle $\lambda_i \neq 0$, so ergibt sich durch wiederholte Anwendung von D7:

$$\det A \overset{D7}{=} \det \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \overset{D1,\,(ii)}{=} \lambda_1 \cdot \ldots \cdot \lambda_n \cdot \det E_n$$
$$\overset{D3}{=} \lambda_1 \cdot \ldots \cdot \lambda_n \ .$$

Gibt es ein i mit $\lambda_i = 0$, so wählen wir i maximal mit dieser Eigenschaft. Dann gilt:

$$\lambda_i = 0, \quad \text{aber} \quad \lambda_{i+1} \cdot \ldots \cdot \lambda_n \neq 0 \ .$$

Mit Hilfe von $\lambda_{i+1}, \ldots, \lambda_n$ macht man die restlichen Einträge der i-ten Zeile auch noch zu Null. Mit D7 und D5 folgt $\det A = 0$.

D9 Durch Zeilenumformungen vom Typ III und IV an A mache man A_1 zu einer oberen Dreiecksmatrix. Dabei bleibt A_2 unverändert, aus C werde C'. Ist k die Anzahl der ausgeführten Zeilenvertauschungen, so gilt

$$\det A_1 \overset{\text{D6}}{=} (-1)^k \cdot \det B_1.$$

Geht B_2 analog aus A_2 hervor (wobei B_1 und C' unverändert bleiben), und ist ℓ die Anzahl der Zeilenvertauschungen, so gilt

$$\det A_2 \overset{\text{D6}}{=} (-1)^\ell \cdot \det B_2.$$

Schreiben wir

$$B = \begin{pmatrix} B_1 & C' \\ 0 & B_2 \end{pmatrix}$$

so ist mit B_1 und B_2 auch B eine obere Dreiecksmatrix. Aus D8 folgt

$$\det B = (\det B_1) \cdot (\det B_2) \, .$$

Wegen

$$\det A = (-1)^{k+\ell} \cdot \det B$$

folgt die Behauptung.

D10 Durch Zeilenumformungen vom Typ III und IV bringen wir A auf Zeilenstufenform B. Dann ist B insbesondere eine obere Dreiecksmatrix, also von der Form

$$B = \begin{pmatrix} \lambda_1 & \cdots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix},$$

und nach D6 und D7 ist $\det B = \pm \det A$. Weiter ist rang $A =$ rang B. Also folg aus D8 die erste zu beweisende Äquivalenz:

$$0 \neq \lambda_1, \cdot \ldots \cdot \lambda_n = \det B \;\Leftrightarrow\; \text{rang } B = n \ .$$

Die zweite Äquivalenz ergibt sich aus 17.4.8.

D11 Es gilt rang $(A \cdot B) \leq$ rang A wegen Bild$(\varphi_A \circ \varphi_B) \subset$ Bildφ_A sowie rang $(A \cdot B) =$ dim Bild$(\varphi_A \circ \varphi_B)$ und rang $A =$ dim Bildφ_A . Ist also rang $A < n$, so ist rang $(A \cdot B) < n$, und die Gleichung lautet $0 = 0$ wegen D9 .

Andernfalls ist $A \in \mathrm{GL}(n, K)$. Nach 17.5.4 gibt es Elementarmatrizen $C_1, \ldots, C_s$ mit

$$A = C_1 \cdot \ldots \cdot C_s \ .$$

Nach 17.5.3 genügt es also zu zeigen, daß für jede Elementarmatrix C vom Typ $S_i(\lambda)$ oder Q_i^j gilt

$$\det(C \cdot B) = \det C \cdot \det B \ .$$

Nach Regel D8, die analog auch für untere Dreiecksmatrizen gilt, ist

$$\det S_i(\lambda) = \lambda \quad \text{und} \quad \det Q_i^j = 1 \ .$$

Multiplizieren von links mit $S_i(\lambda)$ multipliziert die i-te Zeile von B mit λ, also ist

$$\det(S_i(\lambda) \cdot B) \overset{D1}{=} \lambda \cdot \det B \ .$$

Multiplizieren von links mit Q_i^j bewirkt die Addition der j-ten zur i-ten Zeile, also ist

$$\det(Q_i^j \cdot B) \overset{D1}{=} 1 \cdot \det B \ .$$

□

Hiermit hat man eine erste Möglichkeit zur Berechnung von Determinanten: Man bringt A durch Zeilenumformungen vom Typ III und IV auf obere Dreiecksgestalt

$$B = \begin{pmatrix} \lambda_1 & \ldots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix}.$$

Ist k die Anzahl der dabei durchgeführten Zeilenvertauschungen, so gilt

$$\det A = (-1)^k \cdot \det B = (-1)^k \lambda_1 \cdot \ldots \cdot \lambda_n \ .$$

18.2.3 Beispiel

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}$$

$$= - \begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{vmatrix} = 3.$$

□

18.2.4 Bemerkung

Wir werden später sehen, dass für jede Matrix $A \in M(n \times n, K)$ gilt

$$\det {}^tA = \det A.$$

Deswegen gelten alle definierenden Eigenschaften und Rechenregeln für Determinanten, die für Zeilen formuliert sind, analog auch für Spalten. □

18.3 Der Entwicklungssatz von Laplace

Wir benötigen einige technische Vorbereitungen, bevor wir ein weiteres Verfahren zur Determinantenberechnung angeben können.

18.3.1 Definition

Sei $A = (a_{ij}) \in M(n \times n, K)$ eine quadratische Matrix.

(i) Für feste i, j sei $A_{i,j}$ die Matrix, die aus A wie folgt entsteht:

$$A_{ij} := \begin{pmatrix} a_{1,1} & \dots & a_{1,j-1} & 0 & a_{1,j+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \dots & a_{i-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j-1} & 0 & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix}$$

Wir setzen

$$a_{ij}^\sharp := \det A_{ji}$$

und nennen

$$A^\sharp = \left(a_{ij}^\sharp\right) \in M(n \times n, K)$$

die zu A **komplementäre Matrix**.

(ii) Für feste i, j sei A'_{ij} die Matrix, die aus A wie folgt entsteht:

$$A'_{ij} := \left(\begin{array}{ccc|c|ccc} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \hline a_{i1} & \dots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \dots & a_{in} \\ \hline \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{array}\right) \in M((n-1) \times (n-1), K). \quad \square$$

18.3.2 Lemma

Für $A \in M(n \times n, K)$ gilt:

(i) $\det A_{ij} = (-1)^{i+j} \cdot \det A'_{ij}$.

(ii) $\det A_{ij} = \det\left(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n\right)$. Dabei bezeichnen $a^1, \dots, a^n$ die Spaltenvektoren von A und e^j ist der als Spalte geschriebene j-te Einheitsvektor des K^n.

Beweis

(i) Durch $i-1$ Vertauschungen benachbarter Zeilen und $j-1$ Vertauschungen benachbarter Spalten kann man A_{ij} auf die Form

$$\begin{pmatrix} 1 & 0 \\ 0 & A'_{ij} \end{pmatrix}$$

bringen. Wegen

$$(-1)^{(i-1)+(j-1)} = (-1)^{i+j} \ .$$

folgt also die Behauptung aus D6 und D9.

(ii) Durch Addition von Vielfachen der j-ten Spalte zu den anderen Spalten kann man

$$(a^1, \ldots, a^{j-1}, e^i, a^{j+1}, \ldots, a^n)$$

in A_{ij} überführen. Also folgt die Behauptung aus D7. □

18.3.3 Satz

Ist $A \in M\,(n \times n, K)$, so gilt

$$A^\sharp \cdot A = A \cdot A^\sharp = (\det A) \cdot E_n.$$

Beweis

Wir berechnen die Komponenten c_{ik} von $A^\sharp \cdot A$:

$$\begin{aligned} c_{ik} &= \sum_{j=1}^{n} a^\sharp_{ij}\ a_{jk} = \sum_{j=1}^{n} a_{jk}\ \det A_{ji} \\ &\overset{18.3.2,(ii)}{=} \sum_{j=1}^{n} a_{jk} \det\ (a^1, \ldots, a^{i-1}, e^j, a^{i+1}, \ldots, a^n) \\ &\overset{D1}{=} \det(a^1, \ldots, a^{i-1}, \sum_{j=1}^{n} a_{jk} e^j, a^{i+1}, \ldots, a^n) \\ &= \det(a^1, \ldots, a^{i-1}, a^k, a^{i+1}, \ldots, a^n) \\ &\overset{D2}{=} \delta_{ik} \cdot \det A \ . \end{aligned}$$

Also ist $A^\sharp \cdot A = (\det A) \cdot E_n$. Analog berechnet man $A \cdot A^\sharp$. □

18.3.4 Entwicklungssatz von Laplace

Ist $A \in M\,(n \times n, K)\,, n \geqslant 2$, so gilt $\forall i = 1, \ldots, n$

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det A'_{ij}$$

(**Entwicklung nach der i-ten Zeile**) und $\forall j = 1, \ldots, n$

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det A'_{ij}$$

(**Entwicklung nach der j-ten Spalte**).

Beweis

Nach 18.3.3 ist $\det A$ für jedes i gleich der i-ten Komponente in der Diagonalen der Matrix $A \cdot A^\sharp$, also

$$\det A = \sum_{j=1}^{n} a_{ij} \cdot a^\sharp_{ji} = \sum_{j=1}^{n} a_{ij} \cdot \det A_{ij} \overset{18.3.2,\ (i)}{=} \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det A'_{ij}.$$

Dies ergibt die erste Formel. Die zweite ergibt sich durch Betrachten von $A^\sharp \cdot A$. □

18.3.5 Bemerkung

Die durch den Faktor $(-1)^{i+j}$ bewirkte Vorzeichenverteilung kann man sich als „Schachbrettmuster“ vorstellen:

+	-	+	-
-	+	-	+
+	-	+	-
-	+	-	+

. □

18.3.6 Beispiel

(i) Ist

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M\,(2 \times 2, K)\,,$$

so ergibt die Entwicklung nach der 1. Zeile:

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

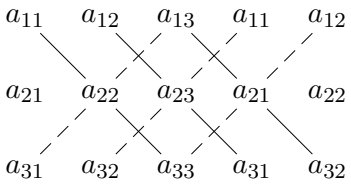
(ii) Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in M\,(3 \times 3, K)\,,$$

so ergibt die Entwicklung nach der 1. Zeile:

$$\det A = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = \ldots\ .$$

Schreibt man diese Formel aus, so erkennt man die **Regel von Sarrus**: Man schreibt den ersten und zweiten Spaltenvektor noch einmal hinter die Matrix



und erhält

$$\begin{aligned}\det A = {} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31}.\end{aligned}$$

(iii)

$$\begin{vmatrix} 1 & 2 & 5 & -1 \\ 3 & 1 & -2 & -2 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 3 \end{vmatrix} \overset{\text{3. Zeile}\downarrow}{=} 1 \cdot \begin{vmatrix} 2 & 5 & -1 \\ 1 & -2 & -2 \\ 1 & 2 & 3 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 2 & 5 \\ 3 & 1 & -2 \\ 0 & 1 & 2 \end{vmatrix}$$

$$= -33 - 7 = -40.$$

□

Eine weitere Folgerung aus 18.3.3 ist:

18.3.7 Satz

Sei $A \in GL(n, K)$. Definiert man $C = (c_{ij}) \in M(n \times n, K)$ durch

$$c_{ij} := (-1)^{i+j} \cdot \det A'_{ij},$$

so gilt

$$A^{-1} = \frac{1}{\det A} \cdot {}^tC.$$

Beweis

Nach D10 ist $\det A \neq 0$. Die Behauptung folgt aus 18.3.3 und 18.3.2, (i). □

18.3.8 Beispiel

Ist $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, K)$, so erhält man die Formel aus 17.4.6, (i):

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

□

18.4 Die Existenz und Eindeutigkeit von Determinanten

Nachdem wir nun wissen, wie man Determinanten ausrechnet, wollen wir herleiten, dass es sie tatsächlich gibt.

Was ist das Problem?

Eine Möglichkeit zur Definition der Determinante wäre es, eine gegebene Matrix A in eine obere Dreiecksmatrix

$$B = \begin{pmatrix} \lambda_1 & \dots & \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix}$$

zu überführen und $\det A$ als

$$\det A := (-1)^k \cdot \det B := (-1)^k \cdot \lambda_1 \cdot \ldots \cdot \lambda_n$$

zu definieren, wobei k die Anzahl der ausgeführten Zeilenvertauschungen ist. Da diese Zeilenumformungen nicht eindeutig bestimmt sind, müsste man dann aber zeigen, dass das Ergebnis unabhängig von allen Wahlen ist. Insbesondere muss klar sein, dass die Zahl k für ein gegebenes A immer gerade oder ungerade ist.

Dass wir Zeilenvertauschungen vornehmen, bedeutet letztendlich, dass wir eine Permutation der Menge der Zeilenindizes $\{1, \dots, n\}$ vornehmen. Eine Transposition der Zeilenindizes entspricht dabei genau einer Zeilenvertauschung. Betrachten wir ein Beispiel: Ist $\sigma \in S_n$ eine Permutation, und sind $e_1, \dots, e_n$ die kanonischen Basisvektoren des K^n, so geht die Matrix

$$E_\sigma := \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$$

aus der Einheitsmatrix durch Zeilenvertauschungen hervor. In diesem Fall ist also die Frage: Ist $\det E_\sigma = +1$ oder ist $\det E_\sigma = -1$?

Diese Frage, aber auch die Frage nach der Existenz von Determinanten, lässt sich durch ein genaueres Studium von Permutationen lösen. Wir benötigen:

18.4.1 Lemma

Ist $n \geqslant 2$, so lässt sich jedes $\sigma \in S_n$ als Komposition von endlich vielen Transpositionen schreiben.

Beweis

Ist $\sigma = \mathrm{id}_{\{1,\dots,n\}}$, so wählen wir eine beliebige Transposition $\tau \in S_n$. Dann gilt:

$$\sigma = \mathrm{id}_{\{1,\dots,n\}} = \tau \circ \tau^{-1} = \tau \circ \tau.$$

Andernfalls $\exists i_1 \in \{1, \dots, n\}$ mit

$$\sigma(i) = i \text{ für } i = 1, \dots, i_1 - 1 \text{ und } \sigma(i_1) \neq i_1.$$

Dann gilt sogar $\sigma(i_1) > i_1$. Sei τ_1 die Transposition, die i_1 mit $\sigma(i_1)$ vertauscht und $\sigma_1 := \tau_1 \circ \sigma$. Dann gilt:

$$\sigma_1(i) = i \text{ für } i = 1, \dots, i_1.$$

Entweder ist nun $\sigma_1 = \mathrm{id}$, oder $\exists i_2$ mit $i_2 > i_1$ und

$$\sigma_1(i) = i \text{ für } i = 1, \dots, i_2 - 1 \text{ und } \sigma_1(i_2) > i_2.$$

Also können wir wie oben fortfahren. □

Die Zerlegung einer Permutation in Transpositionen ist keineswegs eindeutig. Um unser Vorzeichenproblem in den Griff zu bekommen, müssen wir beweisen, dass für gegebenes σ die Anzahl der benötigten Transpositionen immer gerade oder immer ungerade ist. Zu diesem Zweck ordnen wir jeder Permutation ein Vorzeichen zu:

18.4.2 Definition

Seien $n \in \mathbb{N}$ und $\sigma \in S_n$. Dann heißt jedes Paar $i, j \in \{1, \ldots, n\}$ mit

$$i < j \text{ aber } \sigma(i) > \sigma(j)$$

ein **Fehlstand** von σ. Wir definieren das **Signum** von σ durch

$$\operatorname{sign} \sigma := \begin{cases} +1, & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen hat,} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen hat.} \end{cases}$$

Weiter nennen wir σ **gerade**, falls $\operatorname{sign} \sigma = +1$, bzw. **ungerade**, falls $\operatorname{sign} \sigma = -1$.

18.4.3 Beispiel

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in S_3$$

hat genau 2 Fehlstände, nämlich

$$1 < 3 \text{ aber } \sigma(1) = 2 > 1 = \sigma(3),$$

und

$$2 < 3 \text{ aber } \sigma(2) = 3 > 1 = \sigma(3).$$

Es folgt

$$\operatorname{sign} \sigma = +1.$$

□

18.4.4 Satz (Eigenschaften des Signums)

Sei $n \in \mathbb{N}$. Dann gilt

(i) $\forall \sigma \in S_n$ ist

$$\operatorname{sign} \sigma = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

(ii) $\forall \sigma', \sigma \in S_n$ gilt

$$\operatorname{sign}\left(\sigma' \circ \sigma\right) = \operatorname{sign} \sigma' \cdot \operatorname{sign} \sigma.$$

Insbesondere gilt $\forall \sigma \in S_n$:

$$\operatorname{sign} \sigma^{-1} = \operatorname{sign} \sigma.$$

(iii) Ist $n \geqslant 2$ und $\tau \in S_n$ eine Tranposition, so gilt:

$$\operatorname{sign} \tau = -1.$$

(iv) Ist $n \geqslant 2$ und $\sigma = \tau_1 \circ \ldots \circ \tau_k \in S_n$ mit Transpositionen $\tau_1, \ldots, \tau_k$, so gilt:

$$\operatorname{sign} \sigma = (-1)^k .$$

(v) $\forall \sigma \in S_n$ ist

$$\det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = \operatorname{sign} \sigma.$$

Beweis

(i) Sei m die Anzahl der Fehlstände von σ. Dann gilt

$$\begin{aligned} \prod_{i<j} (\sigma(j) - \sigma(i)) &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} (\sigma(j) - \sigma(i)) \cdot (-1)^m \cdot \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} |\sigma(j) - \sigma(i)| \\ &= (-1)^m \prod_{i<j} |\sigma(j) - \sigma(i)| = (-1)^m \prod_{i<j} (j - i), \end{aligned}$$

da die letzten beiden Produkte die gleichen Faktoren bis auf Reihenfolge haben.

(ii) Es gilt

$$\begin{aligned} \operatorname{sign}(\sigma' \circ \sigma) &= \prod_{i<j} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{j - i} \\ &= \prod_{i<j} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \ . \end{aligned}$$

Da das zweite Produkt gleich sign σ ist, genügt es zu zeigen, dass das erste Produkt gleich sign σ' ist. Dies ergibt sich wie folgt:

$$\begin{aligned} &\prod_{i<j} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \prod_{\substack{i>j \\ \sigma(i)<\sigma(j)}} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{\sigma(i)<\sigma(j)} \frac{\sigma'(\sigma(j)) - \sigma'(\sigma(i))}{\sigma(j) - \sigma(i)} = \prod_{i<j} \frac{\sigma'(j) - \sigma'(i)}{j - i} \\ &= \operatorname{sign} \sigma' \ . \end{aligned}$$

Dabei schließt man beim vorletzten Gleichheitszeichen wie in (i).

(iii) Sei

$$\tau_0 := \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ 2 & 1 & 3 & \dots & n \end{pmatrix}$$

die Transposition, die 1 und 2 vertauscht. Wir zeigen:

Ist $\tau \in S_n$ eine beliebige Transposition, so existiert $\sigma \in S_n$ mit $\tau = \sigma \circ \tau_0 \circ \sigma^{-1}$. $(*)$

Daraus folgt (iii): Es gilt sign $\tau_0 = -1$, denn τ_0 hat genau einen Fehlstand. Zu dem gegebenen τ wählen wir ein σ wie in $(*)$. Mit (ii) folgt

$$\begin{aligned} \text{sign } \tau &= \text{sign } \sigma \cdot \text{sign } \tau_0 \cdot (\text{sign } \sigma)^{-1} = \\ &= \text{sign } \tau_0 = -1 \ . \end{aligned}$$

Nun zeigen wir $(*)$: Seien k und ℓ die von τ vertauschten Elemente. Wir behaupten, daß jedes $\sigma \in S_n$ mit

$$\sigma(1) = k \quad \text{und} \quad \sigma(2) = \ell$$

die verlangte Eigenschaft hat. Sei $\tau' := \sigma \circ \tau_0 \circ \sigma^{-1}$. Wegen $\sigma^{-1}(k) = 1$ und $\sigma^{-1}(\ell) = 2$ ist

$$\tau'(k) = \sigma(\tau_0(1)) = \sigma(2) = \ell$$

und

$$\tau'(\ell) = \sigma(\tau_0(2)) = \sigma(1) = k \ .$$

Für $i \notin \{k, \ell\}$ ist $\sigma^{-1}(i) \notin \{1, 2\}$, also gilt

$$\tau'(i) = \sigma(\tau_0(\sigma^{-1}(i))) = \sigma(\sigma^{-1}(i)) = i \ .$$

Daraus folgt $\tau' = \tau$.

(iv) ergibt sich aus (ii) und (iii).

(v) folgt aus (iv), denn man kann E_n durch k Zeilenvertauschungen in

$$\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$$

überführen. □

Im folgenden verwenden wir einige Sprechweisen aus der Gruppentheorie, für die wir auf die Vorlesung Algebraische Strukturen oder die Aufgaben verweisen (für das Verständnis von Determinanten benötigen wir diese Sprechweisen nicht).

18.4.5 Bemerkung und Definition

Die Eigenschaft 18.4.4, (ii) besagt, dass die Abbildung

$$\operatorname{sign} : S_n \to \{+1, -1\}\,,\ \sigma \mapsto \operatorname{sign} \sigma,$$

ein *Gruppenepimorphismus* auf die multiplikative Gruppe mit 2 Elementen ist. Insbesondere ist der *Kern*

$$A_n := \{\sigma \in S_n \mid \operatorname{sign} \sigma = +1\} \subset S_n$$

ein *Normalteiler* in S_n. Wir nennen A_n die **alternierende Gruppe**. □

Für jedes $\sigma \in S_n$ können wir die *Nebenklasse*

$$A_n\sigma = \left\{\sigma' \circ \sigma \mid \sigma' \in A_n\right\}$$

betrachten.

18.4.6 Satz

Ist $\sigma \in S_n$ mit $\operatorname{sign} \sigma = -1$, so gilt

$$S_n = A_n \cup A_n\sigma \quad \text{und} \quad A_n \cap A_n\sigma = \emptyset\ .$$

Insbesondere hat A_n genau $\frac{1}{2}n!$ Elemente.

Beweis

Sei $\sigma' \in S_n$ mit $\operatorname{sign} \sigma' = -1$. Nach 18.4.4, (ii) gilt

$$\operatorname{sign}(\sigma' \circ \sigma^{-1}) = +1\ ,$$

also ist $\sigma' \in A_n\sigma$, denn

$$\sigma' = (\sigma' \circ \sigma^{-1}) \circ \sigma\ .$$

Für jedes $\sigma' \in A_n\sigma$ ist $\operatorname{sign} \sigma' = -1$, also ist die Vereinigung auch disjunkt. Aus den Gruppeneigenschaften ergibt sich, dass die Einschränkung der *Rechtsstranslation* auf A_n,

$$A_n \to A_n\sigma,\ \sigma' \mapsto \sigma' \circ \sigma\ ,$$

bijektiv ist (siehe Vorlesung algebraische Strukturen). Da S_n aus $n!$ Elementen besteht, enthalten A_n und $A_n\sigma$ je $\frac{1}{2}n!$ Elemente. □

Nach diesen Vorbereitungen können wir nun zeigen:

18.4.7 Satz (Existenz und Eindeutigkeit der Determinante)

Ist K ein Körper und $n \in \mathbb{N}$, so gibt es genau eine Determinante

$$\det : M\,(n \times n, K) \to K,$$

und zwar ist für $A = (a_{ij}) \in M\,(n \times n, K)$ die Determinante $\det A$ erklärt durch

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign} \sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}. \qquad (*)$$

Beweis

Eindeutigkeit. Wir zeigen, dass sich die Formel $(*)$ aus D1, D2, D3 und den Folgerungen daraus herleiten lässt. Seien

$$\det : M(n \times n, K) \to K$$

eine Determinante und $A \in M(n \times n, K)$. Wir bezeichnen mit a_i wieder die Zeilen von A und mit e_i die kanonischen Basisvektoren des K^n. Aus D1 ergibt sich die folgende Rechnung, die der sukzessiven Entwicklung nach Zeilen entspricht:

$$\begin{aligned}
\det A = \det \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &= \sum_{j_1=1}^{n} a_{1j_1} \det \begin{pmatrix} e_{j_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \\
&= \sum_{j_1=1}^{n} a_{1j_1} \sum_{j_2=1}^{n} a_{2j_2} \det \begin{pmatrix} e_{j_1} \\ e_{j_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \dots \\
&= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdot a_{2j_2} \cdot \ldots \cdot a_{nj_n} \cdot \det \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} .
\end{aligned}$$

Wegen D2 und 18.4.4, (v) gilt

$$\det \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} = \begin{cases} \text{sign } \sigma & \text{falls } \exists\ \sigma \in S_n \text{ mit } \sigma(\ell) = j_\ell\ \forall \ell, \\ 0 & \text{sonst .} \end{cases}$$

Also bleiben nur $n!$ der n^n Summanden übrig, und es gilt $(*)$.

Existenz. Wir definieren nun

$$\det : M(n \times n, K) \to K$$

durch die Formel $(*)$ für jedes $A \in M(n \times n, K)$, und zeigen, dass dann D1, D2 und D3 gelten.

D1: Es gilt

$$\begin{aligned}
\det \begin{pmatrix} \vdots \\ a_i' + a_i'' \\ \vdots \end{pmatrix} &= \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \cdot \ldots \cdot a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot a'_{i\sigma(i)} \cdot \ldots \cdot a_{n\sigma(n)} \\
&\quad + \sum_{\sigma \in S_n} \text{sign } \sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot a''_{i\sigma(i)} \cdot \ldots \cdot a_{n\sigma(n)} \\
&= \det \begin{pmatrix} \vdots \\ a_i' \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i'' \\ \vdots \end{pmatrix} .
\end{aligned}$$

Somit gilt D1, (i). Analog ergibt sich D1, (ii).

D2: Seien etwa die k-te und ℓ-te Zeile von A gleich, $k < \ell$, und $\tau \in S_n$ die Transposition, die k und ℓ vertauscht. Es gilt sign $\tau = -1$ nach 18.4.4, (iii), also erhält man aus 18.4.6 die disjunkte Vereinigung

$$S_n = A_n \cup A_n\tau \ .$$

Wenn σ die Gruppe A_n durchläuft, so durchläuft $\sigma \circ \tau$ die Menge $A_n\tau$. Es folgt

$$(**) \qquad \det A = \sum_{\sigma\in A_n} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} - \sum_{\sigma\in A_n} a_{1\sigma(\tau(1))} \cdot \ldots \cdot a_{n\sigma(\tau(n))} \ .$$

Da die k-te und ℓ-te Zeile von A gleich sind, gilt nach Definition von τ:

$$\begin{aligned} &a_{1\sigma(\tau(1))} \cdot \ldots \cdot a_{k\sigma(\tau(k))} \cdot \ldots \cdot a_{l\sigma(\tau(l))} \cdot \ldots \cdot a_{n\sigma(\tau(n))} \\ &\quad = a_{1\sigma(1)} \cdot \ldots \cdot a_{k\sigma(l)} \cdot \ldots \cdot a_{l\sigma(k)} \cdot \ldots \cdot a_{n\sigma(n)} \\ &\quad = a_{1\sigma(1)} \cdot \ldots \cdot a_{k\sigma(k)} \cdot \ldots \cdot a_{l\sigma(l)} \cdot \ldots \cdot a_{n\sigma(n)} \\ &\quad = a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \ . \end{aligned}$$

Also heben sich die Summanden in $(**)$ gegenseitig auf, es folgt

$$\det A = 0 \ .$$

D3: Ist δ_{ij} das Kronecker-Symbol und $\sigma \in S_n$, so ist

$$\delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)} = \begin{cases} 0 & \text{für } \sigma \neq \mathrm{id}_{\{1,\ldots,n\}} \\ 1 & \text{für } \sigma = \mathrm{id}_{\{1,\ldots,n\}} \ . \end{cases}$$

Also gilt

$$\begin{aligned} \det E_n = \det(\delta_{ij}) &= \sum_{\sigma\in S_n} \text{sign } \sigma \cdot \delta_{1\sigma(1)} \cdot \ldots \cdot \delta_{n\sigma(n)} \\ &= \text{sign}(\mathrm{id}) = 1 \ . \end{aligned}$$

□

18.4.8 Beispiel

(i) Ist $n = 2$, also etwa $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$, so gibt es genau zwei Permutationen, nämlich

$$\sigma_1 = \ \mathrm{id}_{\{1,2\}} \ = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \text{ und } \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Die Formel $(*)$ lautet also dann

$$\begin{aligned} \det A &= \text{sign } \sigma_1 \cdot a_{1\sigma_1(1)} \cdot a_{2\sigma_1(2)} + \text{sign } \sigma_2 \cdot a_{1\sigma_2(1)} \cdot a_{2\sigma_2(2)} \\ &= +a_{11} \cdot a_{22} - a_{12} \cdot a_{21}. \end{aligned}$$

(ii) Ist $n = 3$, so liefert $(*)$ gerade die aus der Regel von Sarrus abgeleitete Formel.

Unmittelbar aus $(*)$ ergibt sich die folgende Eigenschaft von Determinanten, die wir bereits früher angesprochen haben:

18.4.9 Satz

Für jede Matrix $A \in M(n \times n, K)$ gilt $\det {}^tA = \det A$.

Beweis

Ist $A = (a_{ij})$, so ist ${}^tA = (a'_{ij})$ mit $a'_{ij} = a_{ji}$. Also gilt

$$\begin{aligned}\det {}^tA &= \sum_{\sigma \in S_n} \operatorname{sign}\, \sigma \cdot a'_{1\sigma(1)} \cdot \ldots \cdot a'_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}\, \sigma \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}\, \sigma^{-1} \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \\ &= \det A\ .\end{aligned}$$

Bei der vorletzten Gleichung wurde benutzt, dass für alle $\sigma \in S_n$ gilt

$$a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)}$$

(beide Produkte enthalten bis auf die Reihenfolge die gleichen Faktoren). Außerdem wurde die Identität $\operatorname{sign} \sigma = \operatorname{sign} \sigma^{-1}$ verwendet (siehe 18.4.4, (ii)). Bei der letzten Gleichung haben wir ausgenutzt, dass mit σ auch σ^{-1} ganz S_n durchläuft. □

Wie erwähnt, bedeutet dieser Satz insbesondere, dass es egal ist, ob wir Rechenregeln für Determinanten für Zeilen oder für Spalten formulieren. Wie nützlich es ist, wenn man abwechselnd mit Zeilen und Spalten operieren kann, zeigt das folgende Beispiel:

18.4.10 Beispiel

Für $a_1, \ldots, a_n \in K$ ist die **Vandermonde-Determinante** definiert als

$$\Delta_n := \det \begin{pmatrix} 1 & a_1 & \ldots & a_1^{n-1} \\ \vdots & & & \vdots \\ 1 & a_n & \ldots & a_n^{n-1} \end{pmatrix}.$$

Offensichtlich ist $\Delta_1 = 1,\ \Delta_2 = a_2 - a_1$. Für $n = 3$ formt man zuerst Spalten um und zieht dann Faktoren aus den Zeilen:

$$\begin{aligned}\begin{vmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix} &= \begin{vmatrix} 1 & a_1 & a_1^2 - a_1^2 \\ 1 & a_2 & a_2^2 - a_1 \cdot a_2 \\ 1 & a_3 & a_3^2 - a_1 \cdot a_3 \end{vmatrix} \\ &= \begin{vmatrix} a_2 - a_1 & a_2 \cdot (a_2 - a_1) \\ a_3 - a_1 & a_3 \cdot (a_3 - a_1) \end{vmatrix} \\ &= (a_2 - a_1) \cdot (a_3 - a_1) \cdot (a_3 - a_2)\end{aligned}$$

$$\begin{aligned}&= \begin{vmatrix} 1 & a_1 - a_1 & 0 \\ 1 & a_2 - a_1 & a_2 \cdot (a_2 - a_1) \\ 1 & a_3 - a_1 & a_3 \cdot (a_3 - a_1) \end{vmatrix} \\ &= (a_2 - a_1) \cdot (a_3 - a_1) \cdot \begin{vmatrix} 1 & a_2 \\ 1 & a_3 \end{vmatrix}\end{aligned}$$

Induktiv ergibt sich mit demselben Prinzip

$$\Delta_n = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i)\,.$$

□

18.5 Die Cramersche Regel

Wie in der Motivation 16.1 angesprochen, ergibt sich als Anwendung von Determinanten ein Verfahren für die Lösung linearer Gleichungssysteme, und zwar im Fall

$$A \cdot x = b \text{ mit } A \in GL(n, K).$$

Dann gibt es nach 17.7.3, (ii) eine eindeutige Lösung, und zwar

$$x = A^{-1} \cdot b.$$

Mit Hilfe der komplementären Matrix kann man dies auch so formulieren:

18.5.1 Satz (Cramersche Regel)

Seien $A \in GL(n, K)$, $b \in K^n$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ die eindeutig bestimmte Lösung des Gleichungssystems

$$A \cdot x = b.$$

Bezeichnen $a^1, \ldots, a^n$ die Spaltenvektoren von A, so gilt $\forall i \in \{1, \ldots, n\}$:

$$x_i = \frac{\det\left(a^1, \ldots, a^{i-1}, b, a^{i+1}, \ldots, a^n\right)}{\det A}.$$

Beweis

Nach D10 ist $\det A \neq 0$. Sind $a_1, \ldots, a_n$ die Spaltenvektoren von A, so hat A^{-1} nach 18.3.2, (ii) und 18.3.7 in der i-ten Zeile und j-ten Spalte die Komponente

$$\frac{\det A_{ji}}{\det A} = \frac{\det(a^1, \ldots, a^{i-1}, e^j, a^{i+1}, \ldots, a^n)}{\det A}.$$

Für die i-te Komponente von $x = A^{-1}b$ folgt mit D1 und 18.4.9:

$$x_i = \sum_{j=1}^{n} b_j \frac{\det A_{ji}}{\det A} = \frac{\det(a^1, \ldots, a^{i-1}, b, a^{i+1}, \ldots, a^n)}{\det A}.$$

□

18.5.2 Beispiel

Seien

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} \in M(3 \times 3, \mathbb{R}) \text{ und } b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3.$$

Dann gilt etwa nach der Regel von Sarrus

$$\det A = 2$$

sowie

$$x_1 = \frac{1}{2} \cdot \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{vmatrix} = -1, \quad x_2 = \frac{1}{2} \cdot \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{vmatrix} = 2, \quad x_3 = \frac{1}{2} \cdot \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 3 & 2 & 0 \end{vmatrix} = -1.$$

□

Die Cramersche Regel ist für praktische Anwendungen im Falle großer Matrizen wegen der vielen zu bestimmenden Determinanten nicht geeignet.