Mathematik I für Studierende der Physik

Vicente Cortés
Department Mathematik
Universität Hamburg

Hamburg, Wintersemester 2006-2007[1]

[1] last update: April 19, 2007

Inhaltsverzeichnis I

Inhaltsverzeichnis II

Inhaltsverzeichnis III

Inhaltsverzeichnis IV

Inhaltsverzeichnis V

Inhaltsverzeichnis VI

Reelle Zahlen: Körperaxiome

Addition und Multiplikation

- Das Symbol $\mathbb{R}$ steht für die Menge der reellen Zahlen.
- Auf dieser Menge sind zwei Verknüpfungen gegeben:
- Die Addition

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$(x, y) \mapsto x + y$$

- und die Multiplikation

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$(x, y) \mapsto x \cdot y = xy.$$

- Diese erfüllen folgende Rechenregeln:

A) Axiome der Addition

- Kommutativgesetz:

$$x + y = y + x \quad \text{für alle} \quad x, y \in \mathbb{R}.$$

- Assoziativgesetz:

$$(x + y) + z = x + (y + z) \quad \text{für alle} \quad x, y, z \in \mathbb{R}.$$

- Es gibt ein neutrales Element $0 \in \mathbb{R}$, so dass

$$x + 0 = x \quad \text{für alle} \quad x \in \mathbb{R}.$$

  (Man folgert leicht, dass es genau ein neutrales Element gibt: 0=0+0'=0'+0=0'.)

- Es gibt zu jedem $x \in \mathbb{R}$ ein additives Inverses $-x \in \mathbb{R}$, so dass

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

  (Man zeigt leicht, dass das es zu jedem x genau ein additives Inverses gibt.)

M) Axiome der Multiplikation

- Kommutativgesetz:

$$xy = yx \quad \text{für alle} \quad x, y \in \mathbb{R}.$$

- Assoziativgesetz:

$$(xy)z = x(yz) \quad \text{für alle} \quad x, y, z \in \mathbb{R}.$$

- Es gibt ein neutrales Element $1 \in \mathbb{R}$, so dass

$$x \cdot 1 = x \quad \text{für alle} \quad x \in \mathbb{R}.$$

- Es gibt zu jeder reellen Zahl $x \neq 0$ ein Inverses $x^{-1} \in \mathbb{R}$, so dass

$$x \cdot x^{-1} = 1.$$

  (Wie im Fall der Addition folgert man leicht die Eindeutigkeit von 1 und x^{-1}.)

D) Distributivgesetz

- Das Distributivgesetz drückt die Verträglichkeit von Addition und Multiplikation aus:

$$x(y+z) = xy + xz \quad \text{für alle} \quad x, y, z \in \mathbb{R}.$$

- Die Axiome A), M) und D) faßt man folgendermaßen zusammen:

$$\boxed{(\mathbb{R}, +, \cdot) \text{ ist ein Körper.}}$$

Folgerungen aus den Körperaxiomen

Satz

Für alle $a, b \in \mathbb{R}$ ist die Gleichung

$$a + x = b$$

eindeutig lösbar.

Beweis.

Das folgt aus A). Die Lösung ist $x = b - a := b + (-a)$. $\square$

Satz

Für alle $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$ ist die Gleichung

$$ax = b$$

eindeutig lösbar.

Beweis.

Das folgt aus M). Die Lösung ist $x = \frac{b}{a} := a^{-1}b$. $\square$

Weitere Folgerungen aus den Körperaxiomen

Satz

Sei $\mathbb{K}$ ein Körper (z.B. $\mathbb{K} = \mathbb{R}$) und $x, y \in \mathbb{K}$.
Es gilt $xy = 0$ genau dann wenn $x = 0$ oder $y = 0$.

Beweis.

"$\Leftarrow$" Es genügt zu zeigen $x \cdot 0 = 0$ (wg. Kommutativität). Es gilt

$$x \cdot 0 \overset{\text{A)}}{=} x(0+0) \overset{\text{D)}}{=} x \cdot 0 + x \cdot 0$$

Abziehen von $x \cdot 0$ auf beiden Seiten liefert $0 = x \cdot 0$.

"$\Rightarrow$" Es genügt zu zeigen: $x \neq 0$ und $xy = 0 \Rightarrow y = 0$. Da $x \neq 0$, hat die Gleichung $xy = 0$ die eindeutige Lsg. $y = x^{-1} \cdot 0 = 0$.

$\square$

- Weitere Folgerungen (ÜA): $-0 = 0$, $-(-x) = x$, $(-x)y = -(xy)$, $(-x)(-y) = xy$, $1^{-1} = 1$, für alle $x \neq 0$ ist $x^{-1} \neq 0$ und $(x^{-1})^{-1} = x$, etc.

Natürliche, ganze und rationale Zahlen

- Die natürlichen Zahlen $\mathbb{N} := \{1, 2, 3, \ldots\}$ fassen wir als Teilmenge von $\mathbb{R}$ auf, indem wir $n \in \mathbb{N}$ mit der reellen Zahl $1 + \cdots + 1$ (n Summanden) identifizieren.
- Durch Hinzunahme der additive Inversen erhält man die ganzen Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} \subset \mathbb{R}$. Sie bilden keinen Körper (multiplikative Inverse).
- Der kleinste Körper, der $\mathbb{Z}$ enthält ist der Körper der rationalen Zahlen $\mathbb{Q} = \{\frac{p}{q} | p \in \mathbb{Z}, \quad q \in \mathbb{N}\}$.

Reelle Zahlen: O) Ordnungsaxiome

In $\mathbb{R}$ ist die Teilmenge $\mathbb{R}_+ \subset \mathbb{R}$ der positiven Zahlen ausgezeichnet. Wir schreiben $x > 0$, wenn $x \in \mathbb{R}_+$. Es gelten folgende Axiome:

- O1 Für jede reelle Zahl x gilt genau eine der folgenden drei Relationen:

$$x > 0 \quad \text{oder} \quad x = 0 \quad \text{oder} -x > 0.$$

- O2 Aus $x > 0$ und $y > 0$ folgt $x + y > 0$ und $xy > 0$ (Verträglichkeit mit der Addition und Multiplikation).
- O3 Archimedisches Axiom: Für alle $x \in \mathbb{R}$ gibt es $n \in \mathbb{N}$ mit $n - x > 0$.

Notation

Es werden folgende Vereinbarungen getroffen:

- $x > y :\iff x - y > 0$,
- $x \geq y :\iff x > y$ oder $x = y$,
- $x < y :\iff y > x$,
- $x \leq y :\iff y \geq x$.

Folgerungen aus den Ordnungsaxiomen

Satz

Für $a, x, y, z \in \mathbb{R}$ *gilt:*

(i) *Transitivität:* $x < y$ *und* $y < z \Longrightarrow x < z$,

(ii) $x < y$ *und* $a \in \mathbb{R} \Longrightarrow x + a < y + a$,

(iii) $x < y$ *und* $a > 0 \Longrightarrow ax < ay$,

(iv) $x < y$ *und* $a < 0 \Longrightarrow ax > ay$.

Beweis.

Wir beweisen z.B. (i):

$$z - x = \underbrace{(z - y)}_{>0} + \underbrace{(y - x)}_{>0} > 0$$

□

Weitere Folgerungen aus den Ordnungsaxiomen

Satz

(i) $x^2 > 0$ *für alle* $x \neq 0$,

(ii) $x > 0 \Longrightarrow x^{-1} > 0$,

(iii) $0 < x < y \Longrightarrow x^{-1} > y^{-1} > 0$.

Beweis.

(i) Aus $x > 0$ folgt $x^2 = x \cdot x > 0$, wg. O2). Aus $x < 0$ folgt $-x > 0$ und somit ebenfalls $x^2 = (-x)^2 > 0$.

(ii) Für $x > 0$ gilt wegen (i):

$$x^{-1} = (xx^{-1})x^{-1} = \underbrace{x}_{>0} \cdot \underbrace{(x^{-1})^2}_{>0} > 0.$$

(iii) Für $0 < x < y$ gilt, wegen (ii), $x^{-1} > 0$ und $y^{-1} > 0$ und somit $x^{-1}y^{-1} > 0$. Multiplikation der Ungleichung $x < y$ mit der positiven Zahl $x^{-1}y^{-1}$ liefert schließlich $y^{-1} < x^{-1}$. $\square$

Zusammenfassung:

$\boxed{(\mathbb{R}, +, \cdot) \text{ ist ein archimedisch angeordneter Körper.}}$

- $(\mathbb{Q}, +, \cdot)$ ist ebenfalls ein archimedisch angeordneter Körper.
- $\mathbb{R}$ erfüllt, im Gegensatz zu $\mathbb{Q}$, das Vollständigkeitsaxiom, welches die Liste der Axiome der reellen Zahlen beendet und später besprochen wird.

Komplexe Zahlen

Addition und Multiplikation

- Wir betrachten die kartesische Ebene $\mathbb{R}^2 = \{(x, y) | x, y \in \mathbb{R}\}$, die Menge aller Paare reeller Zahlen,
- und definieren zwei Verküpfungen darauf:
- die Addition
$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$$
$$(x, y) + (x', y') := (x + x', y + y')$$
- und die Multiplikation
$$\cdot : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$$
$$(x, y) \cdot (x', y') := (xx' - yy', xy' + yx').$$

Komplexe Zahlen

Theorem

$(\mathbb{R}^2, +, \cdot)$ *ist ein Körper.*

Beweis.

Es genügt die Axiome A), M) und D) nachzuweisen.

A) Kommutativ- und Assoziativgesetz für $(\mathbb{R}^2, +)$ folgen direkt aus den entsprechenden Axiomen für $(\mathbb{R}, +)$, da die Addition komponentenweise erklärt ist. Das neutrale Element in $(\mathbb{R}^2, +)$ ist $(0, 0)$ und $-(x, y) = (-x, -y)$.

M) Wir überprüfen zuerst das Assoziativgesetz:

$$(x, y) \cdot \big((x', y') \cdot (x'', y'')\big) = (x, y) \cdot (x'x'' - y'y'', x'y'' + y'x'')$$

$$= (x(x'x''-y'y'')-y(x'y''+y'x''), x(x'y''+y'x'')+y(x'x''-y'y''))$$

$$= ((xx' - yy')x'' - (xy' + yx')y'', (xx' - yy')y'' + (xy' + yx')x'')$$

$$= (xx' - yy', xy' + yx') \cdot (x'', y'') = \big((x, y) \cdot (x', y')\big) \cdot (x'', y'')$$

Weiter im Beweis:

M) Das Kommutativgesetz für die Multiplikation in $\mathbb{R}^2$ folgt aus der Invarianz des Ausdrucks $(xx' - yy', xy' + yx')$ unter Vertauschung der Paare (x, y) und (x', y').

Das neutrale Element der Multiplikation ist $(1, 0)$.

In der Tat: $(x, y) \cdot (1, 0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + y \cdot 1) = (x, y)$.

Das multiplikative Inverse zu $(x, y) \neq (0, 0)$ ist

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

In der Tat:
$(x, y) \cdot (\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}) = (\frac{x^2}{x^2+y^2} - \frac{y(-y)}{x^2+y^2}, \frac{x(-y)}{x^2+y^2} + \frac{yx}{x^2+y^2})$
$= (1, 0)$

D) Distributivgesetz: ÜA. $\square$

Reelle Zahlen als Teilmenge der komplexen Zahlen

- Wir identifizieren reelle Zahlen x mit Paaren $(x, 0) \in \mathbb{R}^2$.
- Diese Identifizierung ist mit den Verknüpfungen in $\mathbb{R}$ und $\mathbb{R}^2$ verträglich.
- Jedes Paar $(x, y) \in \mathbb{R}^2$ läßt sich dann in der Form $(x, y) = x + iy$ schreiben, wobei $i := (0, 1) \in \mathbb{R}^2$.
- Die reellen Zahlen $x = \operatorname{Re} z$ und $y = \operatorname{Im} z$ heißen Real- bzw. Imaginärteil der komplexen Zahl $z = x + iy$.
- Die Menge (genauer der Körper) der komplexen Zahlen wird mit dem Symbol $\mathbb{C}$ bezeichnet:
$$\mathbb{C} := \{x + iy | x, y \in \mathbb{R}\} = \mathbb{R}^2 .$$
- Die komplexe Zahl $\bar{z} := x - iy$ heißt die zu $z = x + iy \in \mathbb{C}$ komplex konjugierte Zahl.
- Es gilt
$$\begin{aligned} x &= \operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \\ y &= \operatorname{Im} z = \frac{1}{2i}(z - \bar{z}) \end{aligned}$$

Komplexe Zahlen

Die Gleichung $z^2 = -1$

- Die rein imaginäre Zahl $i = (0, 1)$ ist eine Lösung der Gleichung $z^2 = -1$:

$$i^2 = (0,1) \cdot (0,1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1.$$

- Die einzigen komplexen Lösungen der Gleichung $z^2 = -1$ sind i und $-i$ (ÜA). Man vereinbart $\sqrt{-1} := i$.
- Allgemeiner gilt:

Satz (Fundamentalsatz der Algebra)

Jede polynomiale Gleichung der Form $z^n = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}$ *mit komplexen Koeffizienten* c_k *hat mindestens eine komplexe Lösung.*

- (Dieser Satz läßt sich mit Methoden der Analysis beweisen, die Teil der Vorlesung sind.)

Definition

Der Absolutbetrag einer reellen Zahl x ist definiert durch

$$|x| := \begin{cases} x, & \text{wenn} \quad x \geq 0 \\ -x, & \text{wenn} \quad x < 0 \end{cases}$$

Satz

Für alle $x, y \in \mathbb{R}$ gilt:

(i) $|x| \geq 0$ *und* $|x| = 0$, *genau dann wenn* $x = 0$,

(ii) $|x + y| \leq |x| + |y|$ *(Dreiecksungleichung),*

(iii) $|xy| = |x||y|$ *und*

(iv) $|x + y| \geq ||x| - |y||$.

Beweis.

(i)-(iii) sind einfache ÜA. (iv) folgt aus (ii). □

Konvergenz von Folgen

Definition

Eine Folge reeller Zahlen ist eine Abbildung $\mathbb{N} \to \mathbb{R}$, $n \mapsto x_n$. *Man schreibt dafür auch* $(x_n)_{n\in\mathbb{N}} = (x_1, x_2, \ldots)$ *oder einfach* (x_n).

Wir lassen allgemeiner auch Folgen $(x_n)_{n\geq n_0} = (x_{n_0}, x_{n_0+1}, \ldots)$ zu, die erst ab einem gewissen Rang $n_0 \in \mathbb{N}$ definiert sind.

Definition

Eine Folge (x_n) *heißt konvergent mit dem Grenzwert (=Limes)* ℓ *(symbolisch* $\lim_{n\to\infty} x_n = \ell$*), wenn es für alle* $\epsilon > 0$ *einen Rang* $N = N(\epsilon) \in \mathbb{N}$ *gibt, so dass*

$$|x_n - \ell| < \epsilon \quad \text{für alle} \quad n \geq N.$$

Konvergenz von Folgen

Beispiel

Die Folge $x_n = \frac{1}{n}$ konvergiert gegen 0.

Beweis.

Sei $\epsilon > 0$. Nach dem Archimedischen Axiom gibt es $N \in \mathbb{N}$ mit $N > \frac{1}{\epsilon}$. Für alle $n \geq N$ gilt dann $|x_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \epsilon$. $\square$

Konvergenz von Folgen

Satz

Eine konvergente Folge hat genau einen Grenzwert.

Beweis.

- Wir nehmen an, dass (x_n) zwei Grenzwerte ℓ_1, ℓ_2 hat. Falls $\ell_1 \neq \ell_2$, so können wir Œ annehmen, dass $\ell_1 < \ell_2$.
- Dann gibt es zu $\epsilon = \frac{\ell_2 - \ell_1}{2}$ natürliche Zahlen N_1 und N_2, so dass

$$\begin{array}{rcll} |x_n - \ell_1| & < & \epsilon & \text{für alle} \quad n \geq N_1, \\ |x_n - \ell_2| & < & \epsilon & \text{für alle} \quad n \geq N_2. \end{array}$$

- Daraus folgt nach der Dreiecksungleichung für alle $n \geq \max\{N_1, N_2\}$:
$2\epsilon = \ell_2 - \ell_1 = \ell_2 - x_n + x_n - \ell_1 \leq |\ell_2 - x_n| + |x_n - \ell_1| < 2\epsilon.$
- Ein Widerspruch! Also $\ell_1 = \ell_2$. □

Beschränkte Folgen

Definition

- *Eine Folge* (x_n) *heißt* *nach oben beschränkt*, *wenn es* $K \in \mathbb{R}$ *gibt derart, dass*

$$x_n \leq K \quad \text{für alle} \quad n.$$

- *Sie heißt* *nach unten beschränkt*, *wenn es* $K \in \mathbb{R}$ *gibt derart, dass*

$$K \leq x_n \quad \text{für alle} \quad n.$$

- *Eine Folge* (x_n) *heißt* *beschränkt*, *wenn sie nach oben und nach unten beschränkt ist.*

Einfache ÜA:

Eine Folge (x_n) *ist beschränkt, genau dann wenn es* $K \in \mathbb{R}$ *gibt derart, dass*

$$|x_n| \leq K \quad \text{für alle} \quad n.$$

Konvergenz und Beschränktheit

Satz

Jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ *ist beschränkt.*

Beweis.

- Sei $\ell = \lim_{n\to\infty} x_n$. Dann existiert (zu $\epsilon = 1$) eine natürliche Zahl $N \in \mathbb{N}$, so dass

$$|x_n - \ell| < 1 \quad \text{für alle} \quad n \geq N.$$

- Daraus folgt für alle $n \geq N$:

$$|x_n| = |x_n - \ell + \ell| \leq |x_n - \ell| + |\ell| \leq |\ell| + 1.$$

- und somit für alle $n \in \mathbb{N}$:

$$|x_n| \leq \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, |\ell| + 1\}. \quad \square$$

Beispiele

1) Die Umkehrung des letzten Satzes gilt nicht:

Die Folge $x_n = (-1)^n$ ist beschränkt aber nicht konvergent (ÜA).

2)

Die Folge $x_n = n$ ist unbeschränkt und somit nicht konvergent.

Konvergenz und algebraische Operationen

Satz

(x_n) und (y_n) seien konvergente Folgen. Dann gilt:

(i) $\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n$ *und*

(ii) $\lim_{n\to\infty}(x_n y_n) = (\lim_{n\to\infty} x_n) \cdot (\lim_{n\to\infty} y_n)$.

Beweis.

Sei $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$.

(i) Es gibt $N_1, N_2 \in \mathbb{N}$, so dass

$$\begin{array}{lcll} |x_n - x| & < & \epsilon/2 & \text{für alle} \quad n \geq N_1, \\ |y_n - y| & < & \epsilon/2 & \text{für alle} \quad n \geq N_2. \end{array}$$

Dann gilt für alle $n \geq \max\{N_1, N_2\}$:

$$|x_n + y_n - (x + y)| = |x_n - x + y_n - y| \leq |x_n - x| + |y_n - y| < \epsilon.$$

Weiter im Beweis:

(ii) (x_n) ist als konvergente Folge beschränkt. Daher können wir $K > |y| \geq 0$ wählen, so dass $|x_n| \leq K$ für alle n.

Desweiteren existieren $N_1, N_2 \in \mathbb{N}$ mit

$$\begin{aligned} |x_n - x| &< \frac{\epsilon}{2K} \quad \text{für alle} \quad n \geq N_1, \\ |y_n - y| &< \frac{\epsilon}{2K} \quad \text{für alle} \quad n \geq N_2. \end{aligned}$$

Somit gilt nun für alle $n \geq \max\{N_1, N_2\}$:

$$|x_n y_n - xy| = |x_n y_n \underbrace{- x_n y + x_n y}_{=0} - xy| \leq$$

$$\underbrace{|x_n|}_{\leq K} \underbrace{|y_n - y|}_{< \frac{\epsilon}{2K}} + \underbrace{|x_n - x|}_{< \frac{\epsilon}{2K}} \underbrace{|y|}_{< K} < \epsilon.$$

$\square$

Konvergenz und algebraische Operationen

Satz

Sei (x_n) eine konvergente Folge mit Grenzwert $x \neq 0$. Dann gibt es $n_0 \in \mathbb{N}$, so dass $x_n \neq 0$ für alle $n \geq n_0$ und

$$\lim_{n\to\infty} \left(\frac{1}{x_n}\right)_{n\geq n_0} = \frac{1}{x}.$$

Beweis.

ÜA $\square$

Konvergenz und Ordnungsrelation

Satz

(x_n), (y_n) seien konvergente Folgen mit $x_n \leq y_n$. Dann ist

$$\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n.$$

Beweis.

ÜA □

Beispiel

- Für die Folgen $x_n := 0 < y_n := \frac{1}{n}$ gilt in der Tat $\lim_{n\to\infty} x_n = 0 \leq \lim_{n\to\infty} y_n = 0$,
- aber Achtung: $\lim_{n\to\infty} x_n \not< \lim_{n\to\infty} y_n$.

Cauchy-Folgen

Definition

Eine Folge (x_n) *heißt* *Cauchy-Folge*, *wenn es für alle* $\epsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass*

$$|x_n - x_m| < \epsilon \quad \textit{für alle} \quad n, m \geq N.$$

Satz

Jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ *ist eine Cauchy-Folge.*

Beweis.

Sei $\ell = \lim_{n\to\infty} x_n$. Für alle $n, m \geq N(\frac{\epsilon}{2})$ gilt:

$$|x_n - x_m| = |x_n - \ell + \ell - x_m| \leq \underbrace{|x_n - \ell|}_{<\frac{\epsilon}{2}} + \underbrace{|\ell - x_m|}_{<\frac{\epsilon}{2}} < \epsilon.$$

□

Das Vollständigkeitsaxiom

Die Umkehrung des letzten Satzes ist:

Vollständigkeitsaxiom

$\mathbb{R}$ ist vollständig, d.h. jede Cauchy-Folge reeller Zahlen konvergiert gegen eine reelle Zahl.

Insbesondere konvergiert jede Cauchy-Folge rationaler Zahlen gegen eine reelle Zahl, i.a. aber nicht gegen eine rationale Zahl – wie wir noch sehen werden. $\mathbb{Q}$ ist also nicht vollständig.

Definition

- *Eine Folge (x_n) heißt monoton wachsend (bzw. streng monoton wachsend), wenn*
- $x_n \leq x_{n+1}$ *(bzw.* $x_n < x_{n+1}$*) für alle* $n \in \mathbb{N}$*.*
- *Analog definiert man 'monoton fallend' und 'streng monoton fallend'.*

Bemerkung (einfache ÜA)

Sei (x_n) eine monoton wachsende (bzw. fallende) Folge mit dem Grenzwert ℓ. Dann gilt $x_n \leq \ell$ (bzw. $\ell \leq x_n$) für alle $n \in \mathbb{N}$.

Folgerungen aus dem Vollständigkeitsaxiom

Teilfolgen

Definition

- *Eine Folge* (y_n) *heißt* *Teilfolge* *einer Folge* (x_n),
- *wenn es eine streng monoton wachsende Folge* $\varphi : \mathbb{N} \to \mathbb{N}$ *gibt, so dass*
- $y_n = x_{\varphi(n)} \quad$ *für alle* $\quad n \in \mathbb{N}$.

Bemerkung (einfache ÜA)

Sei (x_n) eine konvergente Folge mit dem Grenzwert ℓ. Dann konvergiert jede Teilfolge von (x_n) ebenfalls gegen ℓ.

Theorem

Jede beschränkte Folge reeller Zahlen besitzt eine konvergente Teilfolge.

Beispiel

Die beschränkte Folge $x_n = (-1)^n$ ist nicht konvergent. Die Teilfolgen $y_n := x_{2n}$ und $y_n' := x_{2n-1}$ sind konstant und daher konvergent.

Beweis des Satzes von Bolzano-Weierstraß

Beweis.

- Sei (x_n) eine beschränkte reelle Folge, d.h. es gibt $A, B \in \mathbb{R}$, so dass
$$A \leq x_n \leq B \quad \text{für alle} \quad n \in \mathbb{N}.$$
- Wir konstruieren rekursiv eine Folge von Intervallen $[A_k, B_k] \subset \mathbb{R}$ $(k \in \mathbb{N})$, derart dass
  - (i) das Invervall $[A_k, B_k]$ unendlich viele Glieder der Folge (x_n) enthält,
  - (ii) $[A_k, B_k] \subset [A_{k-1}, B_{k-1}]$, falls $k \geq 2$ und
  - (iii) $B_k - A_k = 2^{-(k-1)}(B - A)$.
- Dazu setzen wir $A_1 := A$, $B_1 := B$ und nehmen an, die Intervalle $[A_k, B_k]$ mit den obigen Eigenschaften seien für $k \in \{1, \dots \ell\}$ bereits konstruiert.
- Wir definieren dann $[A_{\ell+1}, B_{\ell+1}]$ wie folgt:

Weiter im Beweis des Satzes von Bolzano-Weierstraß

- Sei $M := \frac{1}{2}(A_\ell + B_\ell)$ der Mittelpunkt des ℓ-ten Intervalls.
- Wir setzen $[A_{\ell+1}, B_{\ell+1}] := [A_\ell, M]$, falls $[A_\ell, M]$ unendlich viele Glieder der Folge (x_n) enthält und $[A_{\ell+1}, B_{\ell+1}] := [M, B_\ell]$ sonst.
- Dann erfüllen die Intervalle $[A_k, B_k]$, $k = 1, 2, \ldots, \ell + 1$, die Bedingungen (i)-(iii).
- Als nächstes konstruieren wir, wieder rekursiv, eine Teilfolge $(y_k) = (x_{n_k})$ von (x_n), so dass $y_k \in [A_k, B_k]$.
- Wir setzen $y_1 := x_1$ und nehmen an, $y_1, \ldots, y_k$ seien bereits konstruiert.
- Da $[A_{k+1}, B_{k+1}]$ unendlich viele Glieder der Folge (x_n) enthält, gibt es eine natürliche Zahl $n_{k+1} > n_k$ mit $x_{n_{k+1}} \in [A_{k+1}, B_{k+1}]$.
- Wir setzen $y_{k+1} := x_{n_{k+1}}$.

Ende des Beweises des Satzes von B-W

Behauptung

Die Teilfolge (y_k) ist eine Cauchy-Folge und somit, nach dem Vollständigkeitsaxiom, konvergent.

- Die Behauptung beendet den Beweis des Satzes von B-W.

Beweis der Behauptung.

- Zum Beweis der Behauptung benutzen wir folgende Tatsache:

$$\lim_{n\to\infty} 2^{-n} = 0 \quad (\text{ÜA}).$$

- Daher gibt es für alle $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass

$$2^{-(N-1)}(B - A) < \epsilon.$$

- Für alle $k, \ell \geq N$ gilt $y_k, y_\ell \in [A_N, B_N]$ und somit

$$|y_k - y_\ell| \leq B_N - A_N = 2^{-(N-1)}(B - A) < \epsilon. \quad \square$$

Theorem

Jede monoton wachsende nach oben beschränkte reelle Zahlenfolge (x_n) konvergiert.

(Ebenso konvergiert jede monoton fallende nach unten beschränkte Folge.)

Beweis.

- Die Folge (x_n) ist beschränkt.
- Nach dem Satz von Bolzano-Weierstraß existiert also eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$.
- Wir zeigen, dass (x_n) gegen $\ell := \lim_{k\to\infty} x_{n_k}$ konvergiert.

Weiter im Beweis:

- Wegen $\lim_{k\to\infty} x_{n_k} = \ell$ gibt es zu jedem $\epsilon > 0$ ein $k_0 \in \mathbb{N}$, so dass
$$|x_{n_k} - \ell| < \epsilon \quad \text{für alle} \quad k \geq k_0.$$
- Wir setzen $N := n_{k_0}$.
- Für alle $n \geq N = n_{k_0}$ existiert $k \geq k_0$, so dass
$$n_k \leq n < n_{k+1}.$$
- Da (x_n) monoton wachsend ist, folgt daraus
$$x_{n_k} \leq x_n \leq x_{n_{k+1}} \leq \ell$$
- und somit $|x_n - \ell| \leq |x_{n_k} - \ell| < \epsilon$.

$\square$

Die Quadratwurzel einer positiven Zahl

Theorem

Sei $a > 0$*.*

(i) *Dann hat die Gleichung* $x^2 = a$ *genau eine positive Lösung. (Diese wird mit* $\sqrt{a}$ *bezeichnet.)*

(ii) *Sei* $b > 0$ *und* (x_n) *die durch* $x_1 := b$ *und*

$$x_{n+1} := \frac{1}{2}(x_n + \frac{a}{x_n}) \quad (n \in \mathbb{N})$$

rekursiv definierte Folge.

Dann konvergiert (x_n) *gegen* $\sqrt{a}$*.*

Beweis.

- Wir zeigen zuerst, dass die Folge (x_n) konvergiert.
- Ein einfaches Induktionsargument zeigt, dass $x_n > 0$ für alle $n \in \mathbb{N}$.

Die Quadratwurzel einer positiven Zahl

Weiter im Beweis:

1) Wir zeigen, dass $x_n^2 \geq a$ für alle $n \geq 2$:

$$\begin{aligned} x_n^2 - a &= \frac{1}{4}(x_{n-1} + \frac{a}{x_{n-1}})^2 - a \\ &= \frac{1}{4}(x_{n-1}^2 + 2a + \frac{a^2}{x_{n-1}^2}) - a \\ &= \frac{1}{4}(x_{n-1}^2 - 2a + \frac{a^2}{x_{n-1}^2}) \\ &= \frac{1}{4}(x_{n-1} - \frac{a}{x_{n-1}})^2 \geq 0 \end{aligned}$$

Die Quadratwurzel einer positiven Zahl

Weiter im Beweis:

2) Die Folge $(x_n)_{n\geq 2}$ ist monoton fallend:

$$\begin{aligned} x_n - x_{n+1} &= x_n - \frac{1}{2}(x_n + \frac{a}{x_n}) = \frac{1}{2}(x_n - \frac{a}{x_n}) \\ &= \underbrace{\frac{1}{2x_n}}_{>0}\underbrace{(x_n^2 - a)}_{\geq 0} \geq 0 \end{aligned}$$

- Die Folge $(x_n)_{n\geq 2}$ ist also monoton fallend und durch 0 nach unten beschränkt.
- Somit konvergiert (x_n) gegen eine Zahl $\ell \geq 0$.

Die Quadratwurzel einer positiven Zahl

Weiter im Beweis:

3) Behauptung.

$\ell = \lim_{n\to\infty} x_n > 0$.

Beweis der Behauptung.

- Betrachte die monoton wachsende Folge $y_n := \frac{a}{x_n} > 0$, $n \geq 2$.
- Es gilt

$$(*) \quad y_n^2 = \frac{a^2}{x_n^2} \underset{1)}{\leq} \frac{a^2}{a} = a \leq x_n^2$$

- und somit $y_2 \leq y_n \leq x_n$ für alle $n \geq 2$,
  denn aus $y_n > x_n > 0$ würde $y_n^2 > x_n^2$ folgen, im Widerspruch zu (*).
- Daraus folgt $\ell = \lim_{n\to\infty} x_n \geq y_2 > 0$. $\square$

Die Quadratwurzel einer positiven Zahl

Weiter im Beweis:

- Wir haben gezeigt, dass die Folge (x_n) gegen eine positive Zahl ℓ konvergiert.
- Grenzübergang in der Rekursionsformel $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$ liefert $\ell = \frac{1}{2}(\ell + \frac{a}{\ell})$ und somit $\ell^2 = a$.
- Also ist ℓ eine Lösung der Gleichung $x^2 = a$.
- Sei $\ell' \neq \ell$ eine zweite Lösung der Gleichung $x^2 = a$.
- Dann gilt
$$0 = \ell^2 - \ell'^2 = \underbrace{(\ell - \ell')}_{\neq 0}(\ell + \ell'),$$
woraus $\ell' = -\ell$ folgt.
- M.a.W. ℓ und $-\ell$ sind die einzigen Lösungen der Gleichung $x^2 = a$. Die eine ist positiv die andere negativ. $\square$

Irrationalität von $\sqrt{2}$

Satz

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

Beweis (durch Widerspruch).

- Wir nehmen an $\sqrt{2} = \frac{p}{q}$, $p, q \in \mathbb{N}$ teilerfremd.
- Die Gleichung $p^2 = 2q^2$ zeigt dann, dass p gerade ist.
- Somit ist p^2 durch 4 teilbar und q gerade. Widerspruch! $\square$.

Unvollständigkeit von $\mathbb{Q}$

Folgerung

$\mathbb{Q}$ *ist nicht vollständig.*

Beweis.

- Wie wir gesehen haben, konvergiert die durch $x_1 := b > 0$ und $x_{n+1} := \frac{1}{2}(x_n + \frac{2}{x_n})$, $n \in \mathbb{N}$, rekursiv definierte Folge gegen $\sqrt{2}$.
- Insbesondere ist (x_n) eine Cauchyfolge.
- Wenn wir b rational wählen, so sind alle Folgenglieder rational.
- Dann ist (x_n) eine rationale Cauchyfolge, die nicht in $\mathbb{Q}$ konvergiert.

  Der eindeutig bestimmt Grenzwert ist nämlich die irrationale Zahl $\sqrt{2}$.

$\square$

Der Betrag einer komplexen Zahl

- Sei $z = x + iy \in \mathbb{C}$ (wie immer $x, y \in \mathbb{R}$).
- Dann ist $z\bar{z} = x^2 + y^2$ eine positive reelle Zahl oder Null
- und wir definieren den Betrag der komplexen Zahl z durch

$$|z| := \sqrt{z\bar{z}} \in \mathbb{R}_+ \cup \{0\} \quad (\text{wobei} \quad \sqrt{0} = 0).$$

- Für eine reelle Zahl $x \in \mathbb{R} \subset \mathbb{C}$ ergibt sich als Betrag

$$|x| = \sqrt{x^2} = \begin{cases} x, & \text{wenn} \quad x \geq 0 \\ -x, & \text{wenn} \quad x < 0 \end{cases}$$

genau der Absolutbetrag von x.

Satz

Für alle $z, w \in \mathbb{C}$ gilt:

(i) $|z| \geq 0$ *und* $|z| = 0$*, genau dann wenn* $z = 0$*,*

(ii) $|z + w| \leq |z| + |w|$ *(Dreiecksungleichung),*

(iii) $|zw| = |z||w|$ *und*

(iv) $|z + w| \geq ||z| - |w||$*.*

Beweis.

(i) ist klar und (iv) folgt aus (ii).

(iii) folgt aus der Gleichung $|zw|^2 = zw\overline{zw} = zw\bar{z}\bar{w} = |z|^2|w|^2$ durch Wurzelziehen.

(Hierbei haben wir benutzt, dass $\overline{zw} = \bar{z}\bar{w}$, ÜA)

Der Betrag einer komplexen Zahl

Beweis der Dreiecksungleichung

- Für alle $z = x + iy \in \mathbb{C}$ gilt

$$\mathrm{Re}\, z = x \leq \sqrt{x^2 + y^2} = |z|.$$

- Somit

$$\begin{aligned} |z+w|^2 &= |z|^2 + |w|^2 + z\bar{w} + \bar{z}w \\ &= |z|^2 + |w|^2 + 2\mathrm{Re}\,(z\bar{w}) \\ &\leq |z|^2 + |w|^2 + 2|z\bar{w}| \\ &\overset{(iii)}{=} |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2. \quad \square \end{aligned}$$

Vollständigkeit von $\mathbb{C}$

- Die Begriffe 'konvergente Folge', 'Grenzwert' und 'Cauchyfolge' übertragen sich von reellen auf komplexe Zahlenfolgen, indem man den Absolutbetrag reeller Zahlen durch den Betrag komplexer Zahlen ersetzt.
- Es gilt dann:

Satz

$\mathbb{C}$ ist vollständig, d.h. jede Cauchyfolge komplexer Zahlen konvergiert gegen eine komplexe Zahl.

Beweis.

- Sei $(z_n = x_n + iy_n)$ eine komplexe Cauchyfolge.
- Dann sind (x_n) und (y_n) reelle Cauchyfolgen, denn
- $|x_n - x_m| \leq |z_n - z_m|$ und $|y_n - y_m| \leq |z_n - z_m|$.

Vollständigkeit von $\mathbb{C}$

Weiter im Beweis:

- Da $\mathbb{R}$ vollständig ist, konvergieren (x_n), (y_n) gegen reelle Zahlen x bzw. y.
- Wir setzen $z := x + iy$.
- Die Folge (z_n) konvergiert gegen z,
  denn $|z_n - z| = |x_n - x + i(y_n - y)| \leq |x_n - x| + |i(y_n - y)| = |x_n - x| + |y_n - y|$. $\square$

- *Bemerkung:* Für jede konvergente Folge (z_n) konvergiert die Folge $(|z_n|)$ und es gilt

$$|\lim_{n\to\infty} z_n| = \lim_{n\to\infty} |z_n|.$$

(Das folgt aus $||z_n| - |z|| \leq |z_n - z|$ mit $z = \lim_{n\to\infty} z_n$.)

Konvergenz von Reihen

Der Begriff der Reihe

Definition

- *Eine Reihe (komplexer Zahlen) ist eine Folge* $(s_n)_{n\in\mathbb{N}}$*, der Gestalt*

$$s_n = \sum_{k=1}^{n} z_k,$$

wobei $(z_k)_{k\in\mathbb{N}}$ *eine Folge (komplexer Zahlen) ist.*
- *Die Zahlen* z_k *heißen Glieder der Reihe.*
- *Die Zahlen* $s_n \in \mathbb{C}$ *heißen Partialsummen der Reihe.*
- *Ist die Reihe konvergent, so wird der Grenzwert mit*

$$\sum_{k=1}^{\infty} z_k := \lim_{n\to\infty} s_n \quad \textit{bezeichnet.}$$

(Bemerkung: Jede Folge kann als Reihe aufgefasst werden.)

Definition

- *Eine Reihe*

$$\left(\sum_{k=1}^{n} z_k\right)_{n\in\mathbb{N}}$$

heißt absolut konvergent,

- *wenn die Reihe*

$$\left(\sum_{k=1}^{n} |z_k|\right)_{n\in\mathbb{N}}$$

konvergent ist.

Konvergenz von Reihen

Cauchysches Konvergenzkriterium

Satz

- *Eine Reihe* $(\sum_{k=1}^{n} z_k)_{n\in\mathbb{N}}$ *konvergiert, genau dann wenn*
- *es zu jedem* $\epsilon > 0$ *ein* $N \in \mathbb{N}$ *gibt, so dass*

$$\left|\sum_{k=m}^{n} z_k\right| < \epsilon$$

für alle $n \geq m \geq N$.

Beweis.

Die Bedingung des Satzes besagt einfach, dass die Folge der Partialsummen (s_n) eine Cauchyfolge ist, denn

$$\sum_{k=m}^{n} z_k = s_n - s_{m-1}. \quad \square$$

Konvergenz von Reihen

Absolute Konvergenz impliziert Konvergenz

Satz

Jede absolut konvergente Reihe $(\sum_{k=1}^{n} z_k)$ *ist konvergent.*

Beweis.

- Es genügt zu überprüfen, dass die Reihe $(\sum_{k=1}^{n} z_k)$ das Cauchysche Konvergenzkriterium erfüllt.
- Wg. der absoluten Konvergenz erfüllt die Reihe $(\sum_{k=1}^{n} |z_k|)$ das Cauchysche Konvergenzkriterium.

  Daher gibt es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\sum_{k=m}^{n} |z_k| < \epsilon \quad \text{für alle} \quad n \geq m \geq N.$$

- Daraus folgt für alle $n \geq m \geq N$:
  $|\sum_{k=m}^{n} z_k| \leq \sum_{k=m}^{n} |z_k| < \epsilon$.
- Das zeigt, die Konvergenz der Reihe $(\sum_{k=1}^{n} z_k)$. $\square$

Konvergenz von Reihen

Majorantenkriterium

Satz

- *Sei* $(\sum_{k=1}^n z_k)$ *eine Reihe, so dass* $|z_k| \leq a_k \in \mathbb{R}_+ \cup \{0\}$ *für alle* $k \in \mathbb{N}$*, wobei* $(\sum_{k=1}^n a_k)$ *eine konvergente Reihe ist.*
- *Dann ist die Reihe* $(\sum_{k=1}^n z_k)$ *absolut konvergent und* $|\sum_{k=1}^\infty z_k| \leq \sum_{k=1}^\infty |z_k| \leq \sum_{k=1}^\infty a_k$.

Beweis.

- Die absolute Konvergenz folgt z. B. aus der Abschätzung $\sum_{k=m}^n |z_k| \leq \sum_{k=m}^n a_k$.
- Die Ungleichungskette folgt dann aus

$$|\sum_{k=1}^n z_k| \leq \sum_{k=1}^n |z_k| \leq \sum_{k=1}^n a_k.$$

durch Grenzübergang. $\square$

Konvergenz von Reihen

Die geometrische Reihe

- Sei $z \in \mathbb{C}$. Die Reihe $(\sum_{k=0}^{n} z^k)$ heißt geometrische Reihe.
- Für alle $z \neq 1$ gilt

$$(*) \quad \sum_{k=0}^{n} z^k = 1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}.$$

  (Um das zu verifizieren genügt es die Gleichung mit $(1 - z)$ zu multiplizieren.)
- Sei nun $|z| < 1$. Dann gilt $\lim_{n\to\infty} z^{n+1} = 0$, siehe Übung.
- Wegen (*) konvergiert dann die geometrische Reihe:

$$\sum_{k=0}^{\infty} z^k = 1 + z + z^2 + \cdots = \lim_{n\to\infty} \frac{1 - z^{n+1}}{1 - z} = \frac{1}{1 - z}.$$

- Die Konvergenz ist absolut: $\sum_{k=0}^{\infty} |z^k| = \sum_{k=0}^{\infty} |z|^k = \frac{1}{1-|z|}$.

Satz

- *Sei* $(a_k)_{k\geq 0}$ *eine Folge komplexer Zahlen. Es gebe* $0 \leq \theta < 1$, *so dass*

$$(*) \quad |a_{k+1}| \leq \theta |a_k| \quad \textit{für alle} \quad k.$$

- *Dann ist die Reihe* $(\sum_{k=0}^{n} a_k)$ *absolut konvergent und*

$$\sum_{k=0}^{\infty} |a_k| \leq \frac{|a_0|}{1-\theta}.$$

Beweis.

- Aus (*) erhält man durch Induktion $|a_k| \leq \theta^k |a_0|$.
- Die absolute Konvergenz folgt nun aus dem Majorantenkrit. durch Vergleich mit der geom. Reihe $\sum_{k=0}^{\infty} \theta^k = \frac{1}{1-\theta}$. □

Definition

- *Sei $z \in \mathbb{C}$. Die Reihe*

$$\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)_{n \geq 0}$$

heißt Exponentialreihe.

- Die n-te Partialsumme der Exponentialreihe ist also

$$s_n = \sum_{k=0}^{n} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2} + \cdots + \frac{z^n}{n!}.$$

Satz

Die Exponentialreihe ist absolut konvergent.

Beweis.

- Für alle $z \in \mathbb{C}$ existiert $N \in \mathbb{N}$, so dass $\frac{|z|}{N+1} \leq \frac{1}{2}$.
- Für alle $k \geq N$ gilt dann

$$\left|\frac{z^{k+1}}{(k+1)!}\right| = \frac{|z|}{k+1} \cdot \left|\frac{z^k}{k!}\right| \leq \frac{1}{2}\left|\frac{z^k}{k!}\right|$$

- Nach dem Quotientenkriterium ist also die Reihe $(\sum_{k=N}^{n} \frac{z^k}{k!})_{n \geq N}$ absolut konvergent und daher auch die Reihe $\sum_{k=0}^{n} \frac{z^k}{k!} = \sum_{k=0}^{N-1} \frac{z^k}{k!} + \sum_{k=N}^{n} \frac{z^k}{k!}$.

□

Definition

Die Exponentialfunktion ist die durch die Exponentialreihe definierte Abbildung

$$\exp : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Definition

Das Produkt *zweier Reihen* $(\sum_{k=0}^{n} a_k)$ *und* $(\sum_{k=0}^{n} b_k)$ *ist die Reihe* $(\sum_{k=0}^{n} c_k)$ *mit den Gliedern*

$$c_k := \sum_{j=0}^{k} a_j b_{k-j}.$$

Theorem

- *Das Produkt* $(\sum_{k=0}^{n} c_k)$ *zweier absolut konvergenter Reihen* $(\sum_{k=0}^{n} a_k)$ *und* $(\sum_{k=0}^{n} b_k)$ *ist absolut konvergent*
- *mit dem Grenzwert*

$$\sum_{k=0}^{\infty} c_k = (\sum_{k=0}^{\infty} a_k) \cdot (\sum_{k=0}^{\infty} b_k).$$

Das Produkt absolut konvergenter Reihen ist absolut konvergent

Beweis.

- Wir setzen $a'_k := |a_k|$, $b'_k := |b_k|$ und $c'_k := |c_k|$.
- Die absolute Konvergenz der Reihe $(\sum_{k=0}^{n} c_k)$ folgt aus:

$$
\begin{aligned}
\sum_{k=0}^{n} c'_k &= \sum_{0\le i,j\le n,\ i+j\le n} a'_i b'_j \le \sum_{0\le i,j\le n} a'_i b'_j \\
&= (\sum_{i=0}^{n} a'_i)\cdot(\sum_{j=0}^{n} b'_j) \le (\sum_{i=0}^{\infty} a'_i)\cdot(\sum_{j=0}^{\infty} b'_j).
\end{aligned}
$$

- Bleibt zu zeigen, dass $\sum_{k=0}^{\infty} c_k = (\sum_{k=0}^{\infty} a_k)\cdot(\sum_{k=0}^{\infty} b_k)$.
- Es genügt zu zeigen, dass $\sum_{k=0}^{n} c_k - (\sum_{k=0}^{n} a_k)\cdot(\sum_{k=0}^{n} b_k)$ gegen Null konvergiert.

Das Produkt absolut konvergenter Reihen ist absolut konvergent

Weiter im Beweis:

Wir schätzen ab:

$$\left|\sum_{k=0}^{n} c_k - \Big(\sum_{k=0}^{n} a_k\Big) \cdot \Big(\sum_{k=0}^{n} b_k\Big)\right| = \left|\sum_{0\le i,j\le n,\ i+j\le n} a_i b_j - \sum_{0\le i,j\le n} a_i b_j\right|$$

$$\begin{aligned} &= \left|\sum_{0\le i,j\le n,\ i+j>n} a_i b_j\right| \le \sum_{0\le i,j\le n,\ i+j>n} a_i' b_j' \\ &\le \sum_{k=n+1}^{\infty} c_k' = \sum_{k=0}^{\infty} c_k' - \sum_{k=0}^{n} c_k' \longrightarrow 0 \quad \square \end{aligned}$$

Die Funktionalgleichung der Exponentialfunktion

Folgerung

Für alle $z, w \in \mathbb{C}$ gilt

$$\exp(z+w) = \exp(z)\exp(w).$$

Beweis.

- Wir benutzen die binomische Formel (siehe Übung): $(z+w)^k = \sum_{j=0}^{k} \binom{k}{j} z^{k-j} w^j$, wobei $\binom{k}{j} = \frac{k!}{j!(k-j)!}$.
- Daraus ergibt sich:

$$\exp(z+w) = \sum_{k=0}^{\infty} \frac{(z+w)^k}{k!} = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{z^{k-j}}{(k-j)!} \frac{w^j}{j!} \stackrel{Thm}{=} \exp(z)\exp(w). \quad \square$$

Weitere Folgerungen

Folgerung

Für alle $z \in \mathbb{C}$ gilt

$$\exp(z) \neq 0.$$

Beweis.

Das folgt aus $\exp(z)\exp(-z) = \exp(z-z) = \exp(0) = 1 \neq 0.$ $\square$

- Man definiert die Eulersche Zahl

$$e := \exp(1).$$

- Dann gilt $\exp(n) = \exp(1 + \cdots + 1) = \exp(1)\cdots\exp(1) = e^n$.
- *Bemerkung:* Für jede konvergente Folge (z_n) konvergiert die Folge $(\overline{z_n})$ und es gilt $\lim_{n\to\infty} \overline{z_n} = \overline{\lim_{n\to\infty} z_n}$.
- Insbesondere gilt $\exp(\bar{z}) = \overline{\exp(z)}$.

Sinus und Cosinus

Definition

Für $x \in \mathbb{R}$ definiert man

$$\begin{aligned} \cos x &:= \operatorname{Re} \exp(ix) \quad \textit{und} \\ \sin x &:= \operatorname{Im} \exp(ix), \quad x \in \mathbb{R}. \end{aligned}$$

Satz

Für alle $x \in \mathbb{R}$ gilt $(\cos x)^2 + (\sin x)^2 = 1$, d.h.

$$\exp(ix) = \cos x + i \sin x \in S^1 = \{z \in \mathbb{C} \,|\, |z| = 1\}.$$

Beweis.

$(\cos x)^2 + (\sin x)^2 = |\exp(ix)|^2 = \exp(ix)\overline{\exp(ix)} =$
$\exp(ix)\exp(-ix) = \exp(ix - ix) = \exp(0) = 1.$ $\square$

Einige Eigenschaften von Sinus und Cosinus

Einfache Übungsaufgaben:

Für alle $x \in \mathbb{R}$ gilt

- $\cos(-x) = \cos x$, $\sin(-x) = -\sin x$,
- $$\begin{aligned} \cos(x+y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x+y) &= \sin x \cos y + \cos x \sin y \quad \text{(Additionstheoreme)}, \end{aligned}$$
- $$\begin{aligned} \cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} \cdots, \\ \sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} \cdots \end{aligned}$$

  (mit absoluter Konvergenz).

Ausblick: Ebene Polarkoordinaten

- Wir werden später noch sehen, dass die Funktion $\mathbb{R} \ni \varphi \mapsto \exp(i\varphi) \in S^1 \subset \mathbb{C}$, die Kreislinie periodisch (gegen den Uhrzeigersinn) durchläuft.
- Die Periode ist 2π, wobei die reelle Zahl $\pi = 3,14159\ldots$ noch zu definieren ist.
- Der Parameter φ läßt sich als Bogenlänge des entsprechenden Kreisbogens interpretieren ($\rightarrow$ Zeichnung).
- Insbesondere ist die Periode 2π genau die Bogenlänge der Einheitskreislinie $S^1 = \{z \in \mathbb{C} | |z| = 1\}$.
- Demensprechend hat jede komplexe Zahl $z \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$ eine eindeutige Darstellung der Form

$$z = r \exp(i\varphi), \quad \text{wobei} \quad r > 0 \quad \text{und} \quad \varphi \in [0, 2\pi).$$

- Man nennt $r = |z|$ und φ die Polarkoordinaten von z ($\rightarrow$ Zeichnung). Die reelle Zahl $\arg z := \varphi$ heißt Argument von z.

Multiplikation in Polarkoordinaten

- Die Polarkoordinatendarstellung erleichtert die *Multiplikation* komplexer Zahlen:
$$r \exp(i\varphi) r' \exp(i\varphi') = rr' \exp(i(\varphi + \varphi')).$$

Beispiel

Berechne z^{20} für $z = (1 + i)$.

- Die Polarkoordinaten von z sind $r = \sqrt{2}$ und $\varphi = \frac{\pi}{4}$,
- Also
$$z^{20} = \sqrt{2}^{20} \exp(i20\frac{\pi}{4}) = 2^{10} \exp(i5\pi) = 1024 \exp(i\pi) = -1024.$$
- Hierbei haben wir benutzt, dass $\cos\frac{\pi}{4} = \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}$ und $\exp(i\pi) = -1$. Das folgt unmittelbar aus der (nur mitgeteilten) geometrischen Interpretation am Kreis und wird später noch bewiesen.

Weitere Konvergenzkriterien für Reihen

Satz

Die Glieder z_k einer konvergenten Reihe $\sum z_k$ bilden eine Nullfolge.

Beweis.

Nach dem Cauchyschen Konvergenzkriterium gibt es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass insbesondere $|z_n| = |\sum_{k=n}^{n} z_k| < \epsilon$ für alle $n \geq N$. $\square$

Bemerkung

Das im Satz formulierte Konvergenzkriterium ist notwendig aber nicht hinreichend. Beispielsweise ist die sogenannte harmonische Reihe

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots$$

divergent (ÜA), obwohl die Folge $(\frac{1}{n})$ eine Nullfolge ist.

Leibnizkriterium für alternierende Reihen

Beispiel

- Die alternierende harmonische Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \cdots$$

  konvergiert
- (und zwar gegen $\ln 2$, wie wir zu gegebener Zeit sehen werden).
- Die Konvergenz ergibt sich aus dem Leibnizkriterium:

Satz

Sei $a_k \geq 0$ ($k = 0, 1, 2, \ldots$) eine monoton fallende Nullfolge. Dann konvergiert die Reihe $\sum_{k=0}^{\infty} (-1)^k a_k$.

Leibnizkriterium für alternierende Reihen

Beweis.

- Wir betrachten die Partialsummen $s_n = \sum_{k=0}^{n}(-1)^k a_k$ und setzen $x_n := s_{2n}$ und $y_n := s_{2n+1}$.
- Die Folge (x_n) ist monoton fallend, denn $x_{n+1} - x_n = a_{2n+2} - a_{2n+1} \leq 0$.
- Die Folge (y_n) ist monoton wachsend, denn $y_{n+1} - y_n = -a_{2n+3} + a_{2n+2} \geq 0$.
- Desweiteren gilt $y_0 \leq y_n \leq x_n \leq x_0$, denn $y_n - x_n = -a_{2n+1} \leq 0$.
- Also existieren $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$.
- Wegen $\lim_{n\to\infty}(y_n - x_n) = \lim_{n\to\infty}(-a_{2n+1}) = 0$, gilt $x = y$.
- Daraus folgt leicht $\lim_{n\to\infty} s_n = \lim_{n\to\infty} s_{2n} = \lim_{n\to\infty} s_{2n+1}$. (ÜA)

$\square$

Abzählbare Mengen

Definition

Eine Menge A heißt abzählbar, wenn es eine surjektive Abbildung $\varphi : \mathbb{N} \to A$ gibt.

Beispiele

- Jede endliche Menge $A = \{a_1, a_2, \ldots, a_N\}$ ist abzählbar. Eine Surjektion $\varphi : \mathbb{N} \to A$ ist $(\varphi(n))_{n\in\mathbb{N}} = (a_1, a_2, \ldots, a_N, a_N, \ldots)$.
- Die Menge $\mathbb{Z}$ ist abzählbar. Eine Bijektion $\varphi : \mathbb{N} \to \mathbb{Z}$ ist $(\varphi(n))_{n\in\mathbb{N}} = (0, 1, -1, 2, -2, 3, -3, \ldots)$.

Satz

Die Vereinigung von abzählbar vielen abzählbaren Mengen ist abzählbar.

Beweis.

Das beweist man mit dem Cantorschen Diagonalverfahren ($\to$ Zeichnung). $\square$

Abzählbarkeit von $\mathbb{Q}$

Folgerung

$\mathbb{Q}$ *ist abzählbar.*

Beweis.

$\mathbb{Q} = \mathbb{Z} \cup \{\frac{n}{2} | n \in \mathbb{Z}\} \cup \{\frac{n}{3} | n \in \mathbb{Z}\} \cup \cdots.$ □

Folgerung

$\mathbb{Q}^n = \{x = (x_1, x_2, \ldots, x_n) | x_1, x_2, \ldots, x_n \in \mathbb{Q}\}$ *ist abzählbar für alle* $n \in \mathbb{N}$*.*

Beweis durch Induktion nach *n*.

- Der Induktionsanfang ist die Abzählbarkeit von $\mathbb{Q}$.
- Aus der Abzählbarkeit von $\mathbb{Q}^n$ folgt die von $\mathbb{Q}^{n+1}$, denn

$$\mathbb{Q}^{n+1} = \bigcup_{y \in \mathbb{Q}} \{(x, y) | x \in \mathbb{Q}^n\}. \quad \square$$

Überabzählbarkeit von $\mathbb{R}$

Satz

Die reelle Zahlengerade ist nicht abzählbar.

Beweis.

- Wir nehmen an, die reellen Zahlen seien abzählbar, d.h. $\mathbb{R} = \{x_1, x_2, x_3, \ldots\}$ und leiten einen Widerspruch her.
- I_1 sei ein Intervall der Länge 1 mit $x_1 \notin I_1$.
- Wir unterteilen I_1 in drei gleich große Intervalle und wählen eines der drei aus, das x_2 nicht enthält, und nennen es I_2.
- Durch Fortsetzung dieses rekursiven Verfahrens erhalten wir eine Intervallschachtelung durch Intervalle I_n der Länge $1/3^{n-1}$ mit $x_n \notin I_n$.
- Sei x die eindeutig bestimmt reelle Zahl mit $x \in \cap_{n\in\mathbb{N}} I_n$. Es folgt $x \neq x_n$ für alle $n \in \mathbb{N}$, im Widerspruch zur Annahme.

$\square$

Approximation reeller Zahlen durch rationale Zahlen

Satz

Für jede reelle Zahl x gibt es eine monoton wachsenden (und ebenso eine fallende) Folge rationaler Zahlen, die gegen x konvergiert.

Beweis.

- Definiere $x_n := max\{\frac{k}{2^n} | k \in \mathbb{Z},\ k \leq 2^n x\}$.
- Dann ist x_n monoton wachsend und $|x - x_n| < 1/2^n \to 0$.

$\square$

- ÜA: Zeigen Sie, dass man die Folgen auch *streng* monoton wählen kann.

Infimum und Supremum

Definition

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

- *Eine Zahl s heißt obere (bzw. untere) Schranke für D, falls $x \leq s$ (bzw. $s \leq x$) für alle $x \in D$.*

Satz

- *Sei $\emptyset \neq D \subset \mathbb{R}$ nach oben beschränkt.*
- *Dann hat D ein Supremum, das heißt eine kleinste obere Schranke.*

  (Diese wird mit $\sup D$ bezeichnet.)
- *Ebenso hat jede nach unten beschränkte nicht leere Teilmenge $D \subset \mathbb{R}$ ein Infimum, eine größte untere Schranke $\inf D$.*

Infimum und Supremum

Beweis.

- Wir beweisen z.B. die Existenz des Supremums.
- Dazu konstruieren wir rekursiv $a_n \leq b_n$, $n = 0, 1, 2, \ldots$, so dass
  (i) $a_n \in D$,
  (ii) b_n is eine obere Schranke für D,
  (iii) $a_{n+1} \geq a_n$,
  (iv) $b_{n+1} \leq b_n$ und
  (v) $b_n - a_n \leq 2^{-n}(b_0 - a_0)$.
- Induktionsanfang: Wir beginnen mit einem Element $a_0 \in D$ und einer oberen Schranke b_0 von D.
- Induktionsschritt: Ausgehend von $a_0 \leq b_0, a_1 \leq b_1, \ldots, a_n \leq b_n$ mit (i-v) konstruiren wir $a_{n+1} \leq b_{n+1}$.

Infimum und Supremum

Weiter im Beweis:

- Sei $m = \frac{1}{2}(a_n + b_n)$.
- Wir setzen $[a_{n+1}, b_{n+1}] := [a_n, m]$, falls m eine obere Schranke von D ist.
- Sonst gibt es $d \in D$ mit $d > m$ und wir setzen dann $[a_{n+1}, b_{n+1}] := [d, b_n] \subset [m, b_n]$.
- In beiden Fällen gilt $a_{n+1} \leq b_{n+1}$ und die Eigenschaften (i-v) sind erfüllt.
- Die monoton wachsende Folge (a_n) ist durch b_0 nach oben beschränkt.
- Die monoton fallende Folge (b_n) ist durch a_0 nach unten beschränkt.
- Wegen (v) können wir schließen, dass (a_n) und (b_n) gegen *denselben* Grenzwert c konvergieren.

Infimum und Supremum

Weiter im Beweis:

Behauptung

$c = \sup D$.

Beweis.

- Wir zeigen zuerst, dass c eine obere Schranke von D ist.
  - Sonst gäbe es $a \in D$ mit $a > c$.
  - Wegen $c = \lim_{n\to\infty} b_n$ gäbe es dann $N \in \mathbb{N}$, so dass $a > b_n$ für alle $n \geq N$.
  - Das widerspricht der Tatsache, dass b_n eine obere Schranke für D ist. Also ist c eine obere Schranke von D.
- Wir zeigen als Nächstes, dass c die *kleinste* obere Schranke von D ist.
  - Sonst gäbe es eine kleinere obere Schranke $b < c$ von D.
  - Wegen $c = \lim_{n\to\infty} a_n$ gäbe es dann $N \in \mathbb{N}$, so dass $b < a_n$ für alle $n \geq N$.
  - Das widerspricht der Tatsache, dass $a_n \in D$. Also ist c die kleinste obere Schranke von D. $\square$

Maximum und Minimum

Definition

- *Sei $D \subset \mathbb{R}$ eine nach unten (bzw. oben) beschränkte Menge.*
- *Falls* $\inf D \in D$ *(bzw.* $\sup D \in D$*) so heißt* $\inf D$ *(bzw.* $\sup D$*) das Minimum (bzw. Maximum) der Menge D.*
- *Wir schreiben dann auch* $\min D$ *bzw.* $\max D$ *statt* $\inf D$ *bzw.* $\sup D$*.*

Notation

Falls $\emptyset \neq D \subset \mathbb{R}$ nicht nach unten (bzw. nicht nach oben) beschränkt ist, so setzen wir $\inf D := -\infty$ (bzw. $\sup D := \infty$).

Stetigkeit

Definition

- *Sei $D \subset \mathbb{C}$ und $p \in D$. Eine Funktion $f : D \to \mathbb{C}$ heißt stetig in p, wenn*

$$\lim_{n\to\infty} f(z_n) = f(p)$$

für jede gegen p konvergierende Folge (z_n) in D.

- *$f : D \to \mathbb{C}$ heißt stetig, wenn f in allen Punkten $p \in D$ stetig ist.*

Stetigkeit

Beispiele

(i) Jede konstante Funktion ist stetig.

(ii) Die Funktionen $z \mapsto z, \bar{z}, |z|, \mathrm{Re}\, z, \mathrm{Im}\, z$ sind stetig.

(iii) Die Summe $f + g$ und das Produkt $f \cdot g$ in $p \in D$ stetiger Funktionen $f, g : D \to \mathbb{C}$ sind stetig in p.

(iv) $\frac{1}{f}$ ist stetig in $p \in D$, falls $f : D \to \mathbb{C}^* \subset \mathbb{C}$ stetig in p ist.

(v) Jede rationale Funktion

$$f(z) = \frac{a_n z^n + \cdots + a_0}{b_m z^m + \cdots + b_0}$$

ist stetig auf $D := \{z \in \mathbb{C} | b_m z^m + \cdots + b_0 \neq 0\} \subset \mathbb{C}$.

(vi) Die Funktionen $\bar{f}, |f|, \mathrm{Re}\, f, \mathrm{Im}\, f$ sind stetig in $p \in D$, wenn $f : D \to \mathbb{C}$ stetig in p ist.

Stetigkeit der Exponentialfunktion

Satz

Die Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ *ist stetig.*

Beweis.

- Es genügt zu zeigen, dass exp stetig in 0 ist. In der Tat:
- Sei (z_n) eine konvergente Folge und $p = \lim_{n\to\infty} z_n$. Dann konvergiert $\exp(z_n) = \exp(z_n - p)\exp(p)$ gegen $\exp(0)\exp(p) = \exp(p)$, falls exp stetig in 0 ist.
- Wir zeigen die Stetigkeit im Nullpunkt. Für $|z| < 1$ gilt:

$$\begin{aligned} |\exp(z) - 1| &= |z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots| \leq |z|(1 + \frac{|z|}{2!} + \frac{|z|^2}{3!} + \cdots) \\ &\leq |z|(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots) = |z| \cdot (e - 1) \end{aligned}$$

- Also erfüllt jede Nullfolge z_n ab einem gewissen Folgenglied die Ungleichung $|\exp(z_n) - 1| \leq |z_n| \cdot (e - 1)$, d.h. $\lim_{n\to\infty} \exp(z_n) = 1$. $\square$

Stetigkeit von Sinus und Cosinus

Folgerung

Die Funktionen $\sin, \cos : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

Beweis.

- Aus der Stetigkeit der Exponentialfunktion folgt die Stetigkeit der Funktionen $\mathbb{R} \ni x \mapsto \exp(ix)$ und $x \mapsto \exp(-ix)$.
- Die Stetigkeit von Sinus und Cosinus folgt nun aus der Darstellung

$$\begin{aligned} \cos(x) &= \frac{1}{2}(\exp(ix) + \exp(-ix)) \\ \sin(x) &= \frac{1}{2i}(\exp(ix) - \exp(-ix)). \end{aligned}$$

$\square$

Hyperbolische Funktionen

Definition

Die durch

$$\begin{aligned}\cosh(x) &:= \frac{1}{2}(\exp(x) + \exp(-x)) \\ \sinh(x) &= \frac{1}{2}(\exp(x) - \exp(-x))\end{aligned}$$

definierten Funktionen $\sinh : \mathbb{R} \to \mathbb{R}$ *und* $\cosh : \mathbb{R} \to \mathbb{R}$ *heißen* *Sinus bzw. Cosinus hyperbolicus.*

Folgerung

Die Funktionen $\sinh, \cosh : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

Beweis.

Das folgt aus der Stetigkeit der Exponentialfunktion. $\square$

Hyperbolische Funktionen

Einfache Übungsaufgaben:

Für alle $x \in \mathbb{R}$ gilt $(\cosh x)^2 - (\sinh x)^2 = 1$,

- $\cosh(-x) = \cosh x$, $\sinh(-x) = -\sinh x$,
-
$$\begin{aligned} \cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y, \\ \sinh(x+y) &= \sinh x \cosh y + \cosh x \sinh y \quad \text{(Additionstheoreme)} \end{aligned}$$
-
$$\begin{aligned} \cosh x &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2} + \frac{x^4}{4!} + \cdots, \\ \sinh x &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \end{aligned}$$

(mit absoluter Konvergenz).

Verkettung stetiger Funktionen

Satz
Sei $f : D \to \mathbb{C}$ stetig in $p \in D \subset \mathbb{C}$, $f(D) \subset E \subset \mathbb{C}$ und $g : E \to \mathbb{C}$ stetig in $f(p)$. Dann ist $g \circ f : D \to \mathbb{C}$ stetig in p.

Beweis.

- Sei $z_n \in D$ mit $\lim_{n\to\infty} z_n = p$.
- Wir betrachten $f(z_n) \in E$. Wegen der Stetigkeit von f in p folgt $\lim_{n\to\infty} f(z_n) = f(p)$.
- Aus der Stetigkeit von g in $f(p)$ folgt schließlich $\lim_{n\to\infty} g(f(z_n)) = g(f(p))$.

$\square$

Grenzwert einer Funktion

Notation

- Sei $f : D \to \mathbb{C}$ eine Funktion.
- Wir schreiben
$$\lim_{z \to p} f(z) = q,$$
falls es
  (i) erstens eine Folge $z_n \in D$ gibt, die gegen p konvergiert und
  (ii) zweitens $\lim_{n \to \infty} f(z_n) = q$ für jede solche Folge (z_n) gilt.
  (Wir setzen hier nicht voraus, dass $p \in D$.)
- Mit dieser Notation gilt dann:
$$f \quad \text{ist stetig in} \quad p \in D \quad \Longleftrightarrow \lim_{z \to p} f(z) = f(p).$$

Zwischenwerteigenschaft stetiger Funktionen

- Im Folgenden sei $a < b$. Wir betrachten reellwertige stetige Funktionen auf dem abgeschlossenen Intervall $[a, b] := \{x \in \mathbb{R} | a \leq x \leq b\}$.

Satz (Nullstellensatz von Bolzano)

Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, so dass $f(a)f(b) < 0$. Dann existiert $c \in [a, b]$ mit $f(c) = 0$.

Folgerung (Zwischenwertsatz)

Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion und d eine reelle Zahl zwischen $f(a)$ und $f(b)$. Dann gibt es $c \in [a, b]$ mit $f(c) = d$.

Beweis

- Falls $f(a) < d < f(b)$ oder $f(b) < d < f(a)$, so erfüllt $g(x) := f(x) - d$ die Voraussetzungen des Nullstellensatzes.
- Daher existiert $c \in [a, b]$ mit $0 = g(c) = f(c) - d$ und somit $f(c) = d$. □

Beweis des Nullstellensatzes

- Durch Ersetzen von f durch $-f$, falls notwendig, können wir Œ annehmen, dass $f(a) < f(b)$.
- Wir konstruieren rekursiv Intervalle $[a_n, b_n]$, so dass
  (i) $[a_n, b_n] \subset [a_{n-1}, b_{n-1}]$ für alle $n \in \mathbb{N}$,
  (ii) $b_n - a_n = 2^{-n}(b-a)$,
  (iii) $f(a_n) \leq 0$ und $f(b_n) \geq 0$.
- Wir beginnen mit $[a_0, b_0] := [a, b]$.
- Seien $[a_0, b_0], [a_1, b_1], \ldots, [a_n, b_n]$ mit (i-iii) bereits konstruiert. Wir konstruieren $[a_{n+1}, b_{n+1}]$.
- Sei $m = \frac{1}{2}(a_n + b_n)$. Wir setzen:

$$[a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m], & \text{falls} \quad f(m) \geq 0 \\ [m, b_n], & \text{falls} \quad f(m) < 0. \end{cases}$$

- Die Eigenschaften (i-iii) sind dann erfüllt.

Weiter im Beweis:

- (a_n) ist monoton wachsend, (b_n) fallend und beide Folgen sind durch a nach unten und durch b nach oben beschränkt.
- Wegen (ii) können wir schließen, dass $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n =: c$.
- Die Stetigkeit von f ermöglicht den Grenzübergang in den Ungleichungen $f(a_n) \leq 0$ und $f(b_n) \geq 0$ und liefert somit $f(c) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = 0$. $\square$

Beispiel

- Die durch
$$f(x) := \begin{cases} 0, & \text{falls} \quad x < \sqrt{2} \\ 1, & \text{falls} \quad x > \sqrt{2} \end{cases}$$
definierte Funktion $f : \mathbb{Q} \to \mathbb{R}$ ist stetig,
- besitzt aber keine Fortsetzung zu einer stetigen Funktion $\tilde{f} : \mathbb{R} \to \mathbb{R}$. In der Tat:
- Für jede monoton wachsende rationale Folge (x_n) mit Grenzwert $\sqrt{2}$ gilt $\lim_{n\to\infty} f(x_n) = 0$ und somit müsste $\tilde{f}(\sqrt{2}) = 0$ gelten.
- Andererseits gilt für jede monoton fallende rationale Folge (x_n) mit Grenzwert $\sqrt{2}$ hingegen $\lim_{n\to\infty} f(x_n) = 1$ und somit müsste $\tilde{f}(\sqrt{2}) = 1$ gelten.

Infimum und Supremum einer reellwertigen Funktion

Definition

- *Sei $f : D \to \mathbb{R}$ eine reellwertige Funktion auf einer Menge D (z.B. $D \subset \mathbb{C}$).*
- *f heißt **beschränkt**, falls $f(D) \subset \mathbb{R}$ beschränkt ist.*

  *(f heißt **nach unten** (bzw. **nach oben**) **beschränkt**, falls $f(D)$ nach unten (bzw. nach oben) beschränkt ist.)*
- *Wir setzen*

$$\begin{aligned} \inf f &:= \inf f(D) \in \mathbb{R} \cup \{-\infty\} \\ \sup f &:= \sup f(D) \in \mathbb{R} \cup \{\infty\}. \end{aligned}$$

- *Falls $\inf f \in f(D)$ bzw. $\sup f \in f(D)$, so schreibt man stattdessen auch **$\min f$** bzw. **$\max f$**.*
- *Man sagt dann, dass die Funktion ihr **Minimum** bzw. **Maximum annimmt**.*

Stetige Funktionen auf beschränkten abgeschlossenen Intervallen

Theorem

Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ *ist beschränkt und nimmt ihr Minimum und Maximum an,*

> *d.h. es gibt* x_{min} *und* $x_{max} \in [a, b]$ *mit* $f(x_{min}) = \min f$ *und* $f(x_{max}) = \max f$*.*

Beweis

- Wir zeigen z.B. dass f nach oben beschränkt ist und ihr Maximum annimmt.

1. Fall: Falls das Bild $B := f([a, b]) = \{f(x) | x \in [a, b]\}$ nach oben beschränkt ist, so existiert eine Folge $x_n \in [a, b]$, so dass $f(x_n) \in B$ gegen $\sup B = \sup f$ konvergiert

   (vgl. Existenzbeweis für das Supremum).

- Nach Bolzano-Weierstraß gibt es eine konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$. Wir setzen $\ell := \lim_{k \to \infty} x_{n_k} \in [a, b]$.

Stetige Funktionen auf beschränkten abgeschlossenen Intervallen

Weiter im Beweis:

- Stetigkeit liefert $f(\ell) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = \sup f$, d.h. f nimmt an der Stelle ℓ ihr Maximum an.

2. Fall: Falls $B = f([a,b])$ nach oben unbeschränkt ist, so existiert eine Folge $x_n \in [a,b]$, so dass die Folge $f(x_n) \in B$ monoton wachsend und unbeschränkt ist.

- Nach Bolzano-Weierstraß gibt es eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$; $\ell := \lim_{k\to\infty} x_{n_k} \in [a,b]$.
- Stetigkeit liefert nun $f(\ell) = \lim_{k\to\infty} f(x_{n_k})$. Das ist unmöglich, denn jede Teilfolge einer monoton wachsenden unbeschränkten Folge ist monoton wachsend und unbeschränkt. $\square$

Stetige Funktionen auf beschränkten abgeschlossenen Intervallen

Folgerung

Sei $f : [a, b] \to \mathbb{R}$ *eine stetige Funktion. Dann gilt* $f([a, b]) = [\min f, \max f]$*.*

Beweis.

Das folgt aus dem vorherigen Satz und der Zwischenwerteigenschaft. $\square$

Streng monotone Funktionen

Definition

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

- *Eine Funktion $f : D \to \mathbb{R}$ heißt monoton wachsend (bzw. streng monoton wachsend), falls*

$$f(x) \leq f(x') \quad (bzw. \quad f(x) < f(x'))$$

für alle $x, x' \in D$ mit $x < x'$.

(Die Begriffe 'monoton fallend' und 'streng monoton fallend' werden analog definiert.)

Umkehrfunktionen streng monotoner Funktionen

Satz

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

- *Jede streng monotone Funktion $f : D \to \mathbb{R}$ besitzt eine Umkehrfunktion $g : f(D) \to D$, d.h. eine Funktion so dass $g(f(x)) = x$ für alle $x \in D$.*
- *Die Umkehrfunktion g ist wieder streng monoton, und zwar wachsend, wenn f wachsend ist und fallend, wenn f fallend ist.*

Umkehrfunktionen streng monotoner Funktionen

Beweis.

- Da f streng monoton ist, gibt es zu jedem $y \in f(D)$ *genau* ein $x \in D$ mit $f(x) = y$. Wir definieren $g(y) := x$.
- Um zu zeigen, dass g streng monoton ist, nehmen wir z.B. an, dass f streng monoton wachsend ist, d.h.

  $x < x' \Longrightarrow f(x) < f(x')$.
- Es gilt sogar $x < x' \Longleftrightarrow f(x) < f(x')$, denn $x \geq x' \Longrightarrow f(x) \geq f(x')$.
- Die Substitution $y = f(x)$ und $y' = f(x')$ liefert

$$g(y) < g(y') \Longleftrightarrow y < y'.$$

- Also ist g streng monoton wachsend.

$\square$

Umkehrfunktionen streng monotoner Funktionen

Bemerkungen:

(i) Für beliebige Funktionen $f : D \to \mathbb{R}$ und $g : f(D) \to \mathbb{R}$ impliziert die Gleichung $g(f(x)) = x$ für alle $x \in D$ die Gleichung $f(g(y)) = y$ für alle $y \in f(D)$.

- Um das einzusehen genügt es $y = f(x)$ zu setzen und f auf die Gleichung $g(y) = x$ anzuwenden.

(ii) Wenn $f : D \to \mathbb{R}$ eine Umkehrfunktion $g : f(D) \to D \subset \mathbb{R}$ besitzt, dann schreiben wir diese als $g = f^{-1}$.

- Vorsicht: Verwechseln sie nicht $f^{-1}(x)$, $x \in f(D)$, mit dem multiplikativen Inversen $f(x)^{-1} = \frac{1}{f(x)}$

  (das im Übrigen nur für $f(x) \neq 0$ existiert).

Umkehrfunktionen stetiger streng monotoner Funktionen

Satz

Sei $f : [a, b] \to \mathbb{R}$ *eine stetige streng monoton wachsende (bzw. streng monoton fallende) Funktion und* $[c, d] := [f(a), f(b)]$ *(bzw.* $[c, d] := [f(b), f(a)]$*).*
Dann gilt $f([a, b]) = [c, d]$ *und die streng monotone Umkehrfunktion* $f^{-1} : [c, d] \to [a, b]$ *ist stetig.*

Beweis

- Wenn f z.B. monoton wachsend ist, dann gilt $\min f = f(a)$ und $\max f = f(b)$.
- Wegen der Stetigkeit von f folgt daraus $f([a, b]) = [\min f, \max f] = [f(a), f(b)] = [c, d]$.
- Wenn f fallend ist, erhalten wir $\min f = f(b)$ und $\max f = f(a)$ und somit $f([a, b]) = [\min f, \max f] = [f(b), f(a)] = [c, d]$.

Weiter im Beweis:

- Wir müssen noch die Stetigkeit der Umkehrfunktion
$$f^{-1} : [c, d] \to [a, b] \quad \text{zeigen.}$$
- Sei $y_n \in [c, d]$ eine konvergente Folge, $y = \lim_{n\to\infty} y_n$.
- Wir beweisen durch Widerspruch, dass die Folge $x_n := f^{-1}(y_n) \in [a, b]$ gegen $x := f^{-1}(y)$ konvergiert.
- Wenn (x_n) nicht gegen x konvergiert, so gibt es $\epsilon > 0$, so dass für alle $N \in \mathbb{N}$ eine natürliche Zahl $n \geq N$ existiert mit $|x_n - x| \geq \epsilon$.
- Daher kann man eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ konstruieren mit $|x_{n_k} - x| \geq \epsilon$ für alle k.
- Da $x_{n_k} \in [a, b]$, können wir durch Übergang zu einer Teilfolge Œ annehmen, dass (x_{n_k}) gegen $x' \in [a, b] \setminus \{x\}$ konvergiert.
- Aus der Stetigkeit von f erhalten wir nun
$$y_{n_k} = f(x_{n_k}) \overset{(k\to\infty)}{\longrightarrow} f(x') \underset{(f \text{ str. mon.})}{\neq} f(x) = y$$
- Im Widerspruch zu $\lim_{n\to\infty} y_n = y$. $\square$

Beispiel 1: Potenz- und Wurzelfunktionen

- Sei $k \in \mathbb{N}$. Die Potenzfunktion $x \mapsto f(x) = x^k$ definiert eine stetige und streng monoton wachsende Funktion $f : [0, \infty) \to \mathbb{R}$.
- Die Umkehrfunktion $f^{-1} : [0, \infty) \to \mathbb{R}$, $f^{-1}(x) =: \sqrt[k]{x} =: x^{\frac{1}{k}}$, ist ebenfalls stetig und streng monoton wachsend.
- Das folgt durch Anwendung des vorhergehenden Satzes auf die Einschränkung $f|_{[0,n]} : [0, n] \to [0, n^k] \subset \mathbb{R}$, $n = 1, 2, \ldots$.
- Für ungerades k ist die stetige Funktion $x \mapsto x^k$ streng monoton wachsend auf ganz $\mathbb{R}$, ebenso wie die Umkehrfunktion $f^{-1} : \mathbb{R} \to \mathbb{R}$, $f^{-1}(x) =: \sqrt[k]{x} =: x^{\frac{1}{k}}$.
- ($\rightarrow$ Zeichnung)

Beispiel 2: Exponential- und Logarithmusfunktion

Satz

- *Die reelle Exponentialfunktion* $\exp : \mathbb{R} \to \mathbb{R}$ *ist stetig, streng monoton wachsend und erfüllt* $\exp(\mathbb{R}) = \mathbb{R}_+$.
- *Die Umkehrfunktion*

$$\ln := \exp^{-1} : \mathbb{R}_+ \to \mathbb{R} \quad \textit{(natürlicher Logarithmus)}$$

- *erfüllt die Gleichung*

$$\ln(xx') = \ln(x) + \ln(x')$$

  für alle $x, x' \in \mathbb{R}_+$.
- *(→ Zeichnung)*

Exponential- und Logarithmusfunktion

Beweis

- Die Stetigkeit der Exponentialfunktion haben wir bereits bewiesen.
- Sei $x \geq 0$. Dann ist
$$\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots \geq 1 > 0.$$
- Daraus folgt $\exp(-x) = \exp(x)^{-1} > 0$. Also $\exp(\mathbb{R}) \subset \mathbb{R}_+$.
- Wir zeigen als Nächstes, dass $\exp : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend ist.
- Für $x < x'$ haben wir $\exp(x' - x) > 1$ und somit:
$$\exp(x') = \exp(x' - x + x) = \exp(x' - x)\exp(x) > \exp(x).$$

Exponential- und Logarithmusfunktion

Weiter im Beweis:

- Für alle $n \in \mathbb{N}$ gilt
$$\exp(n) = 1 + n + \frac{n^2}{2!} + \cdots \geq 1 + n$$
- und somit
$$\exp(-n) = (\exp n)^{-1} \leq \frac{1}{1+n}.$$
- Es folgt
$$\mathbb{R}_+ \supset \exp(\mathbb{R}) = \bigcup_{n\in\mathbb{N}} \exp([-n, n]) \supset \bigcup_{n\in\mathbb{N}} [\frac{1}{1+n}, 1+n] = \mathbb{R}_+.$$
- und daher $\exp(\mathbb{R}) = \mathbb{R}_+$.

Exponential- und Logarithmusfunktion

Weiter im Beweis:

- Wir haben gezeigt, dass $\exp : \mathbb{R} \to \mathbb{R}_+$ stetig, streng monoton wachsend und surjektiv ist.
- Daher existiert eine stetige, streng monoton wachsende Umkehrfunktion $\ln = \exp^{-1} : \mathbb{R}_+ \to \mathbb{R}$.
- Es bleibt noch die Funktionalgleichung für den Logarithmus zu beweisen.
- Seien $x, x' \in \mathbb{R}_+$. Wir schreiben $x = \exp(y)$ und $x' = \exp(y')$, wobei $y = \ln(x)$, $y' = \ln(x') \in \mathbb{R}$.
- Die Funktionalgleichung der Exponentialfkt. liefert dann

$$xx' = \exp(y)\exp(y') = \exp(y + y').$$

- Anwendung des Logarithmus ergibt:

$$\ln(xx') = \ln(\exp(y + y')) = y + y' = \ln(x) + \ln(x').\square$$

Exponentialfunktion zur Basis a

Definition

Sei $a > 0$. Wir setzen

$$\exp_a(x) := \exp(x \ln(a)).$$

Bemerkung

Es gilt $\exp_e(x) = \exp(x)$, denn $\ln(e) = \ln \exp(1) = 1$.

Exponentialfunktion zur Basis a

Satz

Die Funktion $\exp_a : \mathbb{R} \to \mathbb{R}$ *ist stetig und erfüllt:*

(i) $\exp_a(x+y) = \exp_a(x)\exp_a(y)$ *für alle* $x, y \in \mathbb{R}$*,*

(ii) $\exp_a(n) = a^n$, $\exp_a(-n) = \frac{1}{a^n}$ *für alle* $n \in \mathbb{N}$ *und*

(iii) $\exp_a(\frac{1}{n}) = a^{\frac{1}{n}} := \sqrt[n]{a}$ *für alle* $n \in \mathbb{N}$*.*

Beweis.

- $\exp_a$ ist als Verkettung $g \circ f$ der stetigen Funktionen $g = \exp$ und $x \mapsto f(x) = x\ln(a)$ stetig.
- (i-iii) sind einfache Folgerungen aus der Funktionalgleichung der Exponentialfunktion (ÜA).

□

Exponentialfunktion zur Basis a

- Wir haben festgestellt, dass $\exp_a(n) = a^n$ für alle $n \in \mathbb{N}$.

Definition

Für $x \in \mathbb{R}$ setzen wir

$$a^x := \exp_a(x).$$

- Dann gilt insbesondere $e^x = \exp(x)$.

Exponentialfunktion zur Basis a

Satz

Für alle $a, b \in \mathbb{R}_+$, $x, y \in \mathbb{R}$ *gilt:*

(i) $a^{x+y} = a^x a^y$,

(ii) $(a^x)^y = a^{xy}$,

(iii) $a^x b^x = (ab)^x$,

(iv) $(\frac{1}{a})^x = a^{-x}$.

Beweis.

(i) ist die Funktionalgleichung von $\exp_a$.

(ii) $(a^x)^y = \exp(y \ln(a^x)) = \exp(y \ln(\exp(x \ln a))) = \exp(yx \ln a) = a^{yx} = a^{xy}$.

(iii) $a^x b^x = \exp(x \ln a) \exp(x \ln b) = \exp(x \ln a + x \ln b) = \exp(x(\ln a + \ln b)) = \exp(x(\ln(ab)) = (ab)^x$.

(iv) $(\frac{1}{a})^x = (a^{-1})^x \underset{(ii)}{=} a^{-x}$. $\square$

Exponentialfunktion zur Basis a

Satz

(i) *Falls $a > 1$, so ist die Funktion $\exp_a : \mathbb{R} \to \mathbb{R}$ streng monoton wachsend.*

(ii) *Falls $0 < a < 1$, so ist sie streng monoton fallend.*

(iii) *In beiden Fällen gilt $\exp_a(\mathbb{R}) = \mathbb{R}_+$.*

Beweis.

(i-ii) Es gilt $\ln(a) > 0$ für $a > 1$ und $\ln(a) < 0$ für $0 < a < 1$, denn

- die Funktion $\ln : \mathbb{R}_+ \to \mathbb{R}$ ist streng monoton wachsend und $\ln 1 = \ln \exp(0) = 0$.
- Entsprechend ist $x \mapsto \exp_a(x) = \exp(x \ln a)$ streng monoton wachsend bzw. fallend.

(iii) Da $\ln a \neq 0$, haben wir $\exp_a(\mathbb{R}) = \exp(\mathbb{R}) = \mathbb{R}_+$.

$\square$

Logarithmus zur Basis a

Definition

Sei $0 < a \neq 1$. *Die Umkehrfunktion* $\log_a := \exp_a^{-1} : \mathbb{R}_+ \to \mathbb{R}$ *der Exponentialunktion* $\exp_a : \mathbb{R} \to \mathbb{R}_+$ *zur Basis a heißt* *Logarithmusfunktion zur Basis a.*

Satz

Es gilt

$$\log_a = \frac{\ln}{\ln a}$$

(insbesondere also $\log_e = \ln$*).*

Beweis.

Für alle $x \in \mathbb{R}_+$ gilt $\exp_a(\frac{\ln x}{\ln a}) = \exp(\frac{\ln x}{\ln a} \ln a) = \exp(\ln x) = x$. $\square$

Charakterisierung der Exponentialfunktionen

Theorem

- *Sei* $f : \mathbb{R} \to \mathbb{R}$ *eine stetige Funktion, die nicht konstant gleich Null ist und die Funktionalgleichung*

$$(*) \quad f(x+y) = f(x)f(y) \quad \textit{für alle} \quad x, y \in \mathbb{R}$$

erfüllt.

- *Dann gilt* $f(1) =: a > 0$ *und* $f = \exp_a$.

Beweis.

- $a = f(1) \overset{(*)}{=} f(\frac{1}{2})^2 \geq 0$.
- Es gilt $a > 0$, denn sonst wäre
  $f(x) \overset{(*)}{=} f(x-1)f(1) = f(x-1)a = 0$ für alle $x \in \mathbb{R}$.
  Das ist unmöglich, denn f ist nach Voraussetzung nicht die Nullfunktion. Also ist $a > 0$.

Charakterisierung der Exponentialfunktionen

Weiter im Beweis:

- Es ist zu zeigen, dass $f(x) = a^x$ für alle $x \in \mathbb{R}$.
- Wir überprüfen zuerst, dass $f(0) = a^0 = 1$:

$$a = f(1) = f(1+0) \overset{(*)}{=} af(0) \Longrightarrow f(0) = 1.$$

- Für $n \in \mathbb{N}$ haben wir $f(n) \overset{(*)}{=} f(1) \cdots f(1) = a^n$ und
- aus $1 = f(n-n) \overset{(*)}{=} a^n f(-n)$ folgt $f(-n) = a^{-n}$.
- Wir haben also bereits gezeigt, dass $f(n) = a^n$ für alle $n \in \mathbb{Z}$.
- Als Nächstes betrachten wir $\frac{p}{q} \in \mathbb{Q}$, wobei $p \in \mathbb{Z}$, $q \in \mathbb{N}$, $q \geq 2$:

$$f\left(\frac{p}{q}\right)^q = \underbrace{f\left(\frac{p}{q}\right) \cdots f\left(\frac{p}{q}\right)}_{\text{q Faktoren}} \overset{(*)}{=} f(p) = a^p > 0.$$

- Diese Rechnung zeigt, dass $f(\frac{p}{q}) = \sqrt[q]{a^p} = a^{\frac{p}{q}}$.

Charakterisierung der Exponentialfunktionen

Weiter im Beweis:

- Wir haben gezeigt, dass $f(x) = a^x$ für alle $x \in \mathbb{Q}$.
- Sei nun $x \in \mathbb{R}$ und $x_n \in \mathbb{Q}$ eine Folge mit Grenzwert x.
- Aus der Stetigkeit von f und $\exp_a$ folgt nun

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} a^{x_n} = a^x.$$

$\square$

Bemerkung/ÜA:

Es genügt die Stetigkeit von $f : \mathbb{R} \to \mathbb{R}$ im Nullpunkt vorauszusetzen.

Uneigentliche Grenzwerte

Definition

- *Sei* (x_n) *eine Folge reeller Zahlen. Man schreibt* $\lim_{n\to\infty} x_n = +\infty$, *falls*

(i) $x_n > 0$ *ab einem gewissen Folgenglied und*

(ii) $\lim_{n\to\infty} \frac{1}{x_n} = 0$.

- *Man schreibt* $\lim_{n\to\infty} x_n = -\infty$, *falls* $\lim_{n\to\infty}(-x_n) = +\infty$.
- *Sei* $f : D \to \mathbb{R}$ *eine Funktion,* $D \subset \mathbb{R}$ *und* $p, \ell \in \mathbb{R} \cup \{-\infty, +\infty\}$. *Man schreibt* $\lim_{x\to p} f(x) = \ell$, *falls es*

(i) *eine Folge* $x_n \in D$ *gibt, mit* $\lim_{n\to\infty} x_n = p$ *und*

(ii) $\lim_{n\to\infty} f(x_n) = \ell$ *für jede solche Folge.*

Uneigentliche Grenzwerte

Beispiele

- $\lim_{x\to-\infty} e^x = 0$, $\lim_{x\to+\infty} e^x = +\infty$,
- $\lim_{x\to 0} \ln x = -\infty$, $\lim_{x\to+\infty} \ln x = +\infty$.

Exponentielles Wachstum

Satz

Für alle $k \in \mathbb{N}$ gilt

$$\lim_{x \to +\infty} \frac{\exp x}{x^k} = +\infty.$$

Beweis.

Für alle $x \geq 0$ gilt $\exp x \geq \frac{x^{k+1}}{(k+1)!}$ und somit

$$\frac{\exp x}{x^k} \geq \frac{x}{(k+1)!} \to +\infty.$$

$\square$

- Die Exponentialfunktion wächst also schneller als jede Potenzfunktion.

Logarithmisches Wachstum

Satz

Für alle $k \in \mathbb{N}$ gilt

$$\lim_{x \to +\infty} \frac{\sqrt[k]{x}}{\ln x} = +\infty.$$

Beweis.

- Für alle $x > 1$ gilt $x = \exp(ky)$, mit $y = \frac{\ln x}{k} > 0$.
- Also $\frac{\sqrt[k]{x}}{\ln x} = \frac{\exp y}{ky}$ und somit
- $$\lim_{x \to +\infty} \frac{\sqrt[k]{x}}{\ln x} = \lim_{y \to +\infty} \frac{\exp y}{ky} = +\infty.$$

□

- Der Logarithmus wächst also langsamer als jede Wurzelfunktion.

Die Zahl π

Theorem

Die Cosinusfunktion hat im Intervall $[0, 2]$ *genau eine Nullstelle* c.

Definition

$\pi := 2c$.

Beweis.

- Wir wissen, dass die Cosinusfunktion stetig ist und dass $\cos 0 = 1 > 0$.
- Wir zeigen, dass $\cos 2 < 0$. Die Existenz der Nullstelle c folgt dann aus dem Nullstellensatz von Bolzano.

Weiter im Beweis:

- Die folgende Abschätzung gilt für $|x| \leq 3$:

$$\begin{aligned}
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \\
&= 1 - \left(\frac{x^2}{2!} - \frac{x^4}{4!}\right) - \left(\frac{x^6}{6!} - \frac{x^8}{8!}\right) - \cdots \\
&= 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3 \cdot 4}\right) - \frac{x^6}{6!}\left(1 - \frac{x^2}{7 \cdot 8}\right) - \cdots \\
&< 1 - \frac{x^2}{2!}\left(1 - \frac{x^2}{3 \cdot 4}\right).
\end{aligned}$$

- Also $\cos 2 < 1 - 2(1 - \frac{1}{3}) = 1 - \frac{4}{3} = -\frac{1}{3} < 0$.
- Als Nächstes zeigen wir, dass die Cosinusfunktion auf dem Intervall $[0, 2]$ streng monoton fallend ist. Somit ist die Nullstelle $c \in (0, 2)$ eindeutig.

Weiter im Beweis:

- Wir haben zu zeigen
$$0 \le x < y \le 2 \Longrightarrow \cos x - \cos y > 0.$$
- Wir schreiben $x = \alpha - \beta$ und $y = \alpha + \beta$, wobei $\alpha = \frac{x+y}{2} \in (0,2)$ und $\beta = \frac{y-x}{2} \in (0,2]$.
- Dann folgt aus den Additionstheoremen: $\cos x - \cos y = \cos(\alpha - \beta) - \cos(\alpha + \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta - (\cos\alpha\cos\beta - \sin\alpha\sin\beta) = 2\sin\alpha\sin\beta$.
- Es genügt daher zu zeigen, dass $\sin x > 0$ für alle $x \in (0,2]$.
- Das folgt aus:
$$\begin{aligned}\sin x &= \left(x - \frac{x^3}{3!}\right) + \left(\frac{x^5}{5!} - \frac{x^7}{7!}\right) + \cdots \\ &= x\left(1 - \frac{x^2}{2\cdot 3}\right) + \frac{x^5}{5!}\left(1 - \frac{x^2}{6\cdot 7}\right) + \cdots\end{aligned}$$

$\square$

Trigonometrische Funktionen

Satz

$$\exp(i\frac{\pi}{2}) = i$$

Beweis.

- Nach Definition von π gilt $\cos\frac{\pi}{2} = 0$.
- Es folgt $(\sin\frac{\pi}{2})^2 = 1 - (\cos\frac{\pi}{2})^2 = 1$
- und somit $\sin\frac{\pi}{2} = 1$, denn $\sin > 0$ auf $(0, 2)$. $\square$

Folgerung

$$\exp(i\pi) = -1, \quad \exp(i\frac{3\pi}{2}) = -i \quad \textit{und} \quad \exp(i2\pi) = 1.$$

Beweis.

Das folgt aus $\exp(in\frac{\pi}{2}) = (\exp(i\frac{\pi}{2}))^n = i^n$ für $n = 2, 3$ und 4. $\square$

Trigonometrische Funktionen

Folgerung

(i) $\exp(z + i2\pi) = \exp z$ *für alle* $z \in \mathbb{C}$,

(ii) $\cos(x + 2\pi) = \cos x$, $\sin(x + 2\pi) = \sin x$,

(iii) $\cos(x + \pi) = -\cos x$, $\sin(x + \pi) = -\sin x$ *und*

(iv) $\cos(x + \frac{\pi}{2}) = -\sin x$, $\sin(x + \frac{\pi}{2}) = \cos x$ *für alle* $x \in \mathbb{R}$.

Beweis.

$\exp(z + in\frac{\pi}{2}) = \exp(z)i^n,\ n = 4, 2, 1.$ □

Trigonometrische Funktionen

Folgerung

Die Funktionen Sinus und Cosinus sind vollständig durch die Einschränkung $\cos\big|_{[0,\frac{\pi}{2}]}$ *bestimmt. (→ Zeichnung)*

Beweis.

- Wegen $\cos x = \sin(x + \frac{\pi}{2})$ erhält man den Graphen der Sinusfunktion durch Verschiebung des Graphen der Cosinusfunktion um $\pi/2$ nach rechts.
- Die Cosinusfunktion ist vollständig durch die Einschränkung $\cos\big|_{[0,\frac{\pi}{2}]}$ bestimmt:
  - Wegen (iii) genügt es cos auf einem Intervall der Länge π zu kennen, z.B. auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$.
  - Wegen der Symmetrie $\cos(x) = \cos(-x)$, genügt $[0, \frac{\pi}{2}]$.

$\square$

Trigonometrische Funktionen

Folgerung

- *Die Funktionen* sin, cos *und* $\mathbb{R} \ni x \mapsto \exp(ix)$ *sind periodisch mit Periode* 2π.
- $\cos x = 0 \iff x = \frac{\pi}{2} + n\pi,\ n \in \mathbb{Z}$.
- $\sin x = 0 \iff x = n\pi,\ n \in \mathbb{Z}$.
- $\exp(ix) = 1 \iff x = n2\pi,\ n \in \mathbb{Z}$.
- *Die Cosinusfunktion ist auf dem Intervall* $[0, \pi]$ *streng monoton fallend.*
- *Die Sinusfunktion ist auf dem Intervall* $[-\frac{\pi}{2}, \frac{\pi}{2}]$ *streng monoton wachsend.*
- *Die Tangensfunktion* $\tan := \frac{\sin}{\cos}$ *(definiert dort wo* $\cos \neq 0$*) ist auf dem Intervall* $(-\frac{\pi}{2}, \frac{\pi}{2})$ *streng monoton wachsend und* $\tan(-\frac{\pi}{2}, \frac{\pi}{2}) = \mathbb{R}$.

Arcusfunktionen

Definition

- *Die zugehörigen (streng montonen und stetigen) Umkehrfunktionen*
  - $\arccos := (\cos|_{[0,\pi]})^{-1} : [-1,1] \to \mathbb{R}$,
  - $\arcsin := (\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]})^{-1} : [-1,1] \to \mathbb{R}$ *und*
  - $\arctan := (\tan|_{(-\frac{\pi}{2},\frac{\pi}{2})})^{-1} : \mathbb{R} \to \mathbb{R}$
- *heißen Arcus-Cosinus, Arcus-Sinus und Arcus-Tangens. ($\rightarrow$ Zeichnung)*

Definition

- *Sei* $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ *eine Funktion und* $x \in D$.
- *Die auf* $D \setminus \{x\}$ *definierte Funktion*

$$\xi \mapsto \frac{f(\xi) - f(x)}{\xi - x}$$

  heißt *Differenzenquotient* *von* f *an der Stelle* x.
- *Ein Punkt* $x \in \mathbb{R}$ *heißt* *Häufungspunkt* *von* D,
- *wenn es eine Folge* $x_n \in D \setminus \{x\} = \{\xi \in D | \xi \neq x\}$ *gibt, die gegen* x *konvergiert.*

Ableitung

Definition

Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x \in D$ ein Häufungspunkt von D.

- *Man sagt, dass f in x differenzierbar ist,*
- *wenn der Grenzwert*

$$f'(x) := \lim_{\xi \to x} \frac{f(\xi) - f(x)}{\xi - x}$$

  existiert.
- *Der Grenzwert $f'(x)$ heißt Ableitung der Funktion f an der Stelle x.*
- *Die Funktion f heißt differenzierbar, wenn sie in allen Punkten $x \in D$ differenzierbar ist.*
- *Die Funktion $f' : x \mapsto f'(x)$ heißt Ableitung von f.*

Differentialquotient und zeitliche Ableitung

Bemerkung

- Nach Leibniz (1646-1716) schreibt man auch

$$\frac{df}{dx}$$

statt f'.

- Nach Newton (1643-1727) schreibt man auch

$$\dot{f}$$

statt f', wenn f eine Funktion der Zeit ist. Die Zeitvariable heißt dann in der Regel t.

Geometrische Interpretation

- Der Differenzenquotient

$$\frac{f(\xi) - f(x)}{\xi - x}$$

  ist genau die Steigung der Geraden durch die Punkte $(x, f(x))$ und $(\xi, f(\xi))$. ($\rightarrow$ Zeichnung)
- Diese Gerade nennt man die Sekante (="die Schneidende").
- Wenn f in x differenzierbar ist, dann strebt die Sekante für $\xi \rightarrow x$ gegen eine Grenzgerade, die sogenannte Tangente:
- Die Tangente (="die Berührende") an den Graphen von f im Punkt $p = (x, f(x))$ ist die Gerade durch p mit Steigung $f'(x)$. ($\rightarrow$ Zeichnung)

Kinematische Interpretation

- Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion.
- Wir können $t \mapsto f(t)$ als die Bewegung eines Punktes im eindimensionalen Raum $\mathbb{R}$ auffassen.
- $f'(t)$ ist dann die Geschwindigkeit zum Zeitpunkt t.
- Die Bewegung eines Punktes $x(t) := (x_1(t), x_2(t), x_3(t))$ im dreidimensionalen Raum $\mathbb{R}^3$ wird entsprechend durch drei Funktionen $x_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$, beschrieben.
- Die Geschwindigkeit $v(t)$ zum Zeitpunkt t hat dementsprechend drei Komponenten:

$$v(t) := \dot{x}(t) := (\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t)).$$

- Nochmaliges Ableiten liefert die Beschleunigung

$$a(t) := \dot{v}(t).$$

Beispiele

(i) Für jede konstante Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, gilt

$$f'(x) = \lim_{\xi \to x} \frac{c - c}{\xi - x} = 0.$$

(ii) Für jede lineare Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = cx$, gilt

$$f'(x) = \lim_{\xi \to x} \frac{c\xi - cx}{\xi - x} = c.$$

(iii) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, gilt

$$f'(x) = \lim_{\xi \to x} \frac{\xi^2 - x^2}{\xi - x} = \lim_{\xi \to x} (\xi + x) = 2x.$$

Weitere Beispiele

(iv) Für $f : \mathbb{R}^* \to \mathbb{R}$, $f(x) = \frac{1}{x}$, gilt

$$f'(x) = \lim_{\xi \to x} \frac{\frac{1}{\xi} - \frac{1}{x}}{\xi - x} = -\lim_{\xi \to x} \frac{1}{\xi x} = -\frac{1}{x^2}.$$

(v) Es gilt $\exp' = \exp$, denn mit $h := \xi - x$ haben wir

- $$\frac{\exp \xi - \exp x}{\xi - x} = \frac{\exp(x+h) - \exp x}{h} = \exp x \frac{\exp h - 1}{h}$$
- und $\lim_{h\to 0} \frac{\exp(h)-1}{h} = 1$.
- In der Tat

$$\left| \frac{\exp h - 1}{h} - 1 \right| = \left| \frac{\frac{h^2}{2!} + \frac{h^3}{3!} + \cdots}{h} \right| \le \frac{|h|}{2!} + \frac{|h|^2}{3!} + \cdots$$

$$\le |h| + \frac{|h|^2}{2!} + \cdots = \exp(|h|) - 1 \to 0 \quad (h \to 0).$$

Weitere Beispiele

(vi) Es gilt $\sin' = \cos$, denn

- $$\begin{aligned} \frac{\sin(x+h)-\sin x}{h} &= \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h}, \end{aligned}$$
- wobei $\lim_{h\to 0} \frac{\cos h - 1}{h} = 0$ und $\lim_{h \to 0} \frac{\sin h}{h} = 1$.
- Das folgt aus:

$$\begin{aligned} \left| \frac{\cos h - 1}{h} \right| &\le \frac{|h|}{2!} + \frac{|h|^3}{4!} + \cdots \le |h| + \frac{|h|^3}{3!} + \cdots \le e^{|h|} - 1 \\ \left| \frac{\sin h}{h} - 1 \right| &\le \frac{|h|^2}{3!} + \frac{|h|^4}{5!} + \cdots \le \frac{|h|^2}{2!} + \frac{|h|^4}{4!} + \cdots \le e^{|h|} - 1. \end{aligned}$$

(vii) $\cos' = -\sin$ (ÜA).

Die Betragsfunktion

Beispiel

- Die stetige Funktion $f(x) := |x|$, $x \in \mathbb{R}$, ist im Nullpunkt nicht differenzierbar.
- Um das einzusehen, betrachten wir die beiden Nullfolgen $x_n^{\pm} := \pm\frac{1}{n}$.
- Einerseits gilt

$$\lim_{n\to\infty} \frac{f(x_n^+) - f(0)}{x_n^+ - 0} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\frac{1}{n}} = 1$$

- und andererseits

$$\lim_{n\to\infty} \frac{f(x_n^-) - f(0)}{x_n^- - 0} = \lim_{n\to\infty} \frac{\frac{1}{n}}{-\frac{1}{n}} = -1.$$

- Also ist f in 0 nicht differenzierbar.

Differenzierbare Funktionen sind stetig

Satz

- *Sei $f : D \rightarrow \mathbb{R}$ in $a \in D \subset \mathbb{R}$ differenzierbar.*
- *Dann ist f in a stetig.*

Beweis.

- Aus
$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a)$$
- folgt
$$\lim_{x \rightarrow a} f(x) = f(a).$$

□

Affine Approximation

Satz

- *Sei $f : D \to \mathbb{R}$ eine in $a \in D \subset \mathbb{R}$ differenzierbare Funktion.*
- *Wir betrachten die affine Funktion $h : D \to \mathbb{R}$,*
  $h(x) := f(a) + f'(a)(x - a)$.

  (Funktionen $h : \mathbb{R} \to \mathbb{R}$ von der Form $h(x) = cx + d$, wobei $c, d \in \mathbb{R}$, heissen affin.)
- *Es gilt*

$$\lim_{x \to a} \frac{f(x) - h(x)}{x - a} = 0.$$

Beweis.

$$\frac{f(x) - h(x)}{x - a} = \frac{f(x) - f(a)}{x - a} - f'(a) \to 0 \quad (x \to a).$$

□

Umgekehrt gilt:

Satz

- *Sei $f : D \to \mathbb{R}$ eine Funktion und $a \in D \subset \mathbb{R}$ ein Häufungspunkt von D.*
- *Desweiteren sei $h : \mathbb{R} \to \mathbb{R}$ eine affine Funktion, derart dass*

$$(*) \quad \lim_{x \to a} \frac{f(x) - h(x)}{x - a} = 0.$$

- *Dann ist f in a differenzierbar, $f(a) = h(a)$ und $f'(a) = h'(a)$.*

Beweis.

- Aus (*) folgt $f(a) = h(a)$ und somit $h(x) = f(a) + c(x - a)$.
- Daraus folgt
  $0 = \lim_{x \to a} \frac{f(x) - h(x)}{x - a} = \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} - c \right) = f'(a) - c$,
  d.h. $f'(a) = c = h'(a)$.

$\square$

Affine Approximation

Definition

- *Sei $f : D \to \mathbb{R}$ in $a \in D$ differenzierbar, $D \subset \mathbb{R}$.*
- *Die affine Funktion $h(x) = f(a) + f'(a)(x - a)$ heisst affine Approximation von f in a.*

  (Manchmal spricht man auch von linearer Approximation.)

Ableitungsregeln

Satz

- *Die Funktionen $f, g : D \to \mathbb{R}$ seien in $a \in D \subset \mathbb{R}$ differenzierbar. Dann gilt:*

(i) *Die Funktion $f + g : D \to \mathbb{R}$ ist in a differenzierbar und*

$$(f+g)'(a) = f'(a) + g'(a).$$

(ii) *Die Funktion $f \cdot g : D \to \mathbb{R}$ ist in a differenzierbar und*

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

(iii) *Wenn $g(D) \subset \mathbb{R}^*$, dann ist die Funktion $\frac{f}{g} : D \to \mathbb{R}$ in a differenzierbar und*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Ableitungsregeln

Beweis.

(i)

$$
\begin{aligned}
& \frac{(f+g)(a+h)-(f+g)(a)}{h} \\
= \quad & \frac{f(a+h)-f(a)}{h} + \frac{g(a+h)-g(a)}{h} \xrightarrow[h\to 0]{} f'(a)+g'(a)
\end{aligned}
$$

(ii)

$$
\begin{aligned}
& \frac{(fg)(a+h)-(fg)(a)}{h} \\
= \quad & \frac{f(a+h)g(a+h)-f(a)g(a+h)}{h} + \frac{f(a)g(a+h)-f(a)g(a)}{h} \\
= \quad & \frac{f(a+h)-f(a)}{h} g(a+h) + f(a)\frac{g(a+h)-g(a)}{h} \\
\to \quad & f'(a)g(a)+f(a)g'(a) \quad (h\to 0)
\end{aligned}
$$

Ableitungsregeln

Weiter im Beweis:

(iii) Wir betrachten zunächst den Spezialfall $f = 1$:

$$\frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} = \frac{g(a) - g(a+h)}{hg(a+h)g(a)}$$

$$= -\underbrace{\frac{g(a+h) - g(a)}{h}}_{\to g'(a)} \cdot \underbrace{\frac{1}{g(a+h)g(a)}}_{\to 1/g(a)^2} \to -\frac{g'(a)}{g(a)^2} \quad (h \to 0),$$

- d.h. $(\frac{1}{g})'(a) = -\frac{g'(a)}{g(a)^2}$. Mit (ii) erhalten wir dann für beliebiges f:

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a) = \frac{f'(a)}{g(a)} + f(a)\left(\frac{1}{g}\right)'(a)$$

$$= \frac{f'(a)g(a)}{g(a)^2} - \frac{f(a)g'(a)}{g(a)^2}.$$

$\square$

Beispiele

- Eine einfache Induktion mit Hilfe der Produktregel (ii) ergibt

$$(x^n)' = nx^{n-1} \quad \text{für alle} \quad n \in \mathbb{N}.$$

- Die Regel $(1/g)' = -g'/g^2$ mit $g(x) = x^n$ liefert dann

$$(x^{-n})' = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1} \quad (x \neq 0) \quad \text{für alle} \quad n \in \mathbb{N}.$$

- Die Quotientenregel liefert

$$\tan' = \frac{\sin' \cos - \sin \cos'}{\cos^2} = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2} = 1 + \tan^2 .$$

Ableitung der Umkehrfunktion

Satz

- *Sei* $f : [a, b] \to \mathbb{R}$ *eine stetige streng monotone Funktion und* $g = f^{-1} : [c, d] \to \mathbb{R}$ *sei die zugehörige Umkehrfunktion, wobei* $[c, d] = f([a, b])$.
- *Wenn* f *in* $x \in [a, b]$ *differenzierbar ist und* $f'(x) \neq 0$, *dann ist* g *in* $y = f(x)$ *differenzierbar und*
- $$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'(g(y))}.$$

Ableitung der Umkehrfunktion

Bemerkung

- Wenn im obigen Satz f überall differenzierbar ist und $f'(x) \neq 0$ für alle $x \in [a, b]$,
- so erhalten wir für die Ableitung der Umkehrfunktion $g = f^{-1} : [c, d] \to \mathbb{R}$ die Regel

$$g' = \frac{1}{f' \circ g}.$$

- Mit der Leibnizschen Notation schreibt sich diese Regel einprägsamer als:

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}},$$

wobei die Funktion $\frac{dy}{dx} = \frac{df}{dx} = f'$ an der Stelle x auszuwerten ist und $\frac{dx}{dy} = \frac{dg}{dy}$ an der Stelle $y = f(x)$.

Ableitung der Umkehrfunktion

Beweis.

- Sei $y_n \in [c, d] \setminus \{y\}$ eine Folge, die gegen y konvergiert.
- Da g stetig ist, konvergiert die Folge $x_n := g(y_n) \in [a, b] \setminus \{x\}$ gegen $x := g(y)$.
- Somit

$$\lim_{n\to\infty} \frac{g(y_n) - g(y)}{y_n - y} = \lim_{n\to\infty} \frac{x_n - x}{f(x_n) - f(x)} = \frac{1}{f'(x)},$$

- d.h. $g'(y) = 1/f'(x)$.

□

Beispiele

(i)

$$\ln'(x) = \frac{1}{\exp'(\ln x)} = \frac{1}{\exp(\ln x)} = \frac{1}{x} \quad (x > 0).$$

(ii) $\arcsin = (\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]})^{-1} : [-1,1] \to \mathbb{R}$ hat für $x \in (-1,1)$ die Ableitung

$$\arcsin'(x) = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}},$$

denn $\cos = \sqrt{1-\sin^2}$ auf $[-\frac{\pi}{2},\frac{\pi}{2}]$.

Beispiele

(iii) Ebenso hat $\arccos = (\cos|_{[0,\pi]})^{-1} : [-1,1] \to \mathbb{R}$ für $x \in (-1,1)$ die Ableitung

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}.$$

(iv)

$$\arctan'(x) = \frac{1}{\tan'(\arctan x)} = \frac{1}{1+(\tan(\arctan x))^2} = \frac{1}{1+x^2},$$

denn $\tan' = 1 + \tan^2$.

Die Kettenregel

Satz

- *Sei* $f : D \to \mathbb{R}$ *an der Stelle* $x \in D \subset \mathbb{R}$ *differenzierbar,*
- $f(D) \subset E \subset \mathbb{R}$ *und*
- $g : E \to \mathbb{R}$ *an der Stelle* $y := f(x)$ *differenzierbar.*
- *Dann ist* $g \circ f : D \to \mathbb{R}$ *an der Stelle* x *differenzierbar und*

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

Die Kettenregel

Beweis.

- Sei $h : E \to \mathbb{R}$ die stetige Fortsetzung des Differenzenquotienten von g in $y = f(x)$, d.h.

$$h(\eta) := \begin{cases} \frac{g(\eta)-g(y)}{\eta-y}, & \text{falls} \quad \eta \in E \setminus \{y\} \\ g'(y), & \text{falls} \quad \eta = y. \end{cases}$$

- Für alle $\eta \in E$ gilt dann

$$g(\eta) - g(y) = h(\eta)(\eta - y).$$

- Daraus folgt (mit $\eta = f(\xi)$) für alle $\xi \in D \setminus \{x\}$:

$$\frac{(g \circ f)(\xi) - (g \circ f)(x)}{\xi - x} = h(f(\xi))\frac{f(\xi) - f(x)}{\xi - x}$$

$$\to h(f(x))f'(x) = g'(f(x))f'(x) \quad (\xi \to x).$$

□

Beispiele

(i) Sei $f : \mathbb{R} \to \mathbb{R}$ differenzierbar, $a, b \in \mathbb{R}$ und $\tilde{f} : \mathbb{R} \to \mathbb{R}$ definiert durch $\tilde{f}(x) := f(ax + b)$.

- Dann ist $\tilde{f}$ differenzierbar und

$$\tilde{f}'(x) = af'(ax + b).$$

(ii) Sei $f : \mathbb{R}_+ \to \mathbb{R}$, $f(x) = x^\alpha$, wobei $\alpha \in \mathbb{R}$.

- Dann ist

$$f'(x) = \alpha x^{\alpha-1},$$

- denn aus $f(x) = \exp(\alpha \ln x)$ folgt mit der Kettenregel

$$f'(x) = \exp'(\alpha \ln x)(\alpha \ln x)' = \underbrace{\exp(\alpha \ln x)}_{x^\alpha}\frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

Lokale Extrema

Definition

- *Sei $f : D \to \mathbb{R}$ eine Funktion und $z \in D \subset \mathbb{C}$.*
- *Man sagt, dass f in z ein lokales Maximum (bzw. ein lokales Minimum) annimmt, falls es $\epsilon > 0$ gibt, derart dass*

  $$f(z) \geq f(\zeta) \quad (\textit{bzw.} \quad f(z) \leq f(\zeta))$$

  für alle $\zeta \in D$ mit $|z - \zeta| < \epsilon$.
- *Statt von (lokalen bzw. globalen) Minima und Maxima spricht man auch von (lokalen bzw. globalen) Extrema.*
- *Man spricht von einem isolierten lokalen Extremum, falls zusätzlich $f(\zeta) \neq f(z)$ für alle $\zeta \in D \setminus \{z\}$ mit $|z - \zeta| < \epsilon$.*

Lokale Extrema

Satz

- *Die Funktion* $f : (a,b) \to \mathbb{R}$ *sei an der Stelle* $x \in (a,b)$ *differenzierbar und nehme dort ein lokales Extremum an.*
- *Dann gilt* $f'(x) = 0$.

Beweis.

- Wir nehmen z.B. an, dass ein lokales Maximum vorliegt.
- Dann gibt es $\epsilon > 0$, so dass $f(x) - f(\xi) \geq 0$ für alle ξ mit $|x - \xi| < \epsilon$.
- Somit
$$f'(x) = \lim_{\xi \nearrow x} \frac{f(\xi) - f(x)}{\xi - x} \geq 0$$
- und ebenso $f'(x) = \lim_{\xi \searrow x} \frac{f(\xi)-f(x)}{\xi - x} \leq 0$, d.h. $f'(x) = 0$.

□

Beispiele

(i) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = (x+2)(x+1)x^2$ (siehe Zeichnung), ist nach oben unbeschränkt ($\sup f = +\infty$) und hat daher kein (globales) Maximum.

- Sie nimmt ihr (globales) Minimum $\min f$ an einer Stelle $a \in (-2, -1)$ an.
- Sie hat zwei weitere lokale Extrema: ein lokales Maximum an einer Stelle $b \in (-1, 0)$ und ein lokales Minimum bei 0.
- ÜA: Berechnen Sie a, b und $\min f = f(a)$.

(ii) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, erfüllt $f'(0) = 0$, hat aber an der Stelle 0 kein lokales Extremum (siehe Zeichnung).

(iii) Die Funktion $f : [-1, 2] \to \mathbb{R}$, $f(x) = x^2$, nimmt an der Stelle 0 ihr Minimum $\min f = 0$, an der Stelle 2 ihr Maximum $\max f = 4$ und an der Stelle -1 ein lokales Maximum an.

Der Satz von Rolle

Satz

- *Sei* $f : [a, b] \to \mathbb{R}$ *(weiterhin* $a < b$*) eine stetige Funktion mit* $f(a) = f(b)$.
- *Falls* f *auf* (a, b) *differenzierbar ist, so existiert* $\xi \in (a, b)$ *mit* $f'(\xi) = 0$ *(*$\to$ *Zeichnung).*

Beweis.

- Falls f konstant ist, so gilt $f' = 0$ und der Satz ist erfüllt.
- Falls f nicht konstant ist, so existiert $x \in (a, b)$ mit $f(x) > f(a) = f(b)$ oder $f(x) < f(a) = f(b)$.
- Im ersten Fall existiert $\xi \in (a, b)$ mit $f(\xi) = \max f$, im zweiten Fall $\xi \in (a, b)$ mit $f(\xi) = \min f$.
- In beiden Fällen ist $f'(\xi) = 0$.

□

Mittelwertsatz

Satz

Sei $f : [a, b] \to \mathbb{R}$ *(*$a < b$*) eine stetige, auf* (a, b) *differenzierbare Funktion. Dann existiert* $\xi \in (a, b)$ *mit*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \quad \textit{(siehe Zeichnung).}$$

Beweis.

- Wir betrachten die Hilfsfunktion $g : [a, b] \to \mathbb{R}$,

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

- g erfüllt die Voraussetzungen des Satzes von Rolle, insbesondere $g(a) = g(b) = f(a)$.
- Somit existiert $\xi \in (a, b)$ mit $0 = g'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}$. □

Folgerungen aus dem Mittelwertsatz

Der Schrankensatz

Folgerung

- *Unter den Voraussetzungen des MWS gelte zusätzlich*

$$(*) \quad m \leq f'(\xi) \leq M \quad \textit{für alle} \quad \xi \in (a,b).$$

- *Dann gilt*

$$(**) \quad m(y-x) \leq f(y) - f(x) \leq M(y-x)$$

für alle $x, y \in [a,b]$ *mit* $x \leq y$.

Beweis.

- Für $x = y$ ist nichts zu zeigen. Für $x < y$ gibt es nach dem MWS ein $\xi \in (x,y)$ mit $f'(\xi) = \frac{f(y)-f(x)}{y-x}$.
- Multiplikation der Ungleichung (*) mit $y - x > 0$ ergibt dann (**). $\square$

Folgerung

Sei $f : [a, b] \to \mathbb{R}$ *stetig, auf* (a, b) *differenzierbar und* $f'(x) = 0$ *für alle* $x \in (a, b)$. *Dann ist* f *konstant.*

Beweis.

- Die Funktion erfüllt die Voraussetzungen des Schrankensatzes mit $m = M = 0$.
- Also $f(y) = f(x)$ für alle $x, y \in [a, b]$ mit $x \leq y$, d.h. $f = const$.

□

Bemerkung

- Dieser Satz ist sehr hilfreich beim Studium der Eindeutigkeit von Lösungen von Differentialgleichungen bei gegebenen Anfangsbedingungen.
- Er besagt, dass die Dgl. $f' = 0$ genau eine Lösung mit der Anfangsbedingung $f(x_0) = c$ hat, nämlich die konstante Funktion $f = c$. (Hierbei ist $x_0 \in [a, b]$ und $c \in \mathbb{R}$.)

Charakterisierung der Exponentialfunktion durch eine Dgl.

Satz

- *Sei $c \in \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ eine Lösung der Dgl.*

$$(*) \quad f' = cf.$$

- *Dann gilt $f(x) = f(0)e^{cx}$ (für alle $x \in \mathbb{R}$).*

Beweis.

- Wir betrachten die differenzierbare Funktion $g : \mathbb{R} \to \mathbb{R}$, $g(x) := f(x)e^{-cx}$.
- Ableiten liefert:

$$g'(x) = f'(x)e^{-cx} - cf(x)e^{-cx} \overset{(*)}{=} 0.$$

- Somit $g = const = g(0) = f(0)$, d.h. $f(x) = g(x)e^{cx} = f(0)e^{cx}$. $\square$

Monotonie und Ableitung

Satz

Sei $f : [a, b] \to \mathbb{R}$ *stetig und auf* (a, b) *differenzierbar.*

(i) *Falls* $f'(x) \geq 0$ *(bzw.* $f'(x) > 0$*) für alle* $x \in (a, b)$*, so ist* f *auf* $[a, b]$ *monoton wachsend (bzw. streng monoton wachsend).*

(ii) *Falls* $f'(x) \leq 0$ *(bzw.* $f'(x) < 0$*) für alle* $x \in (a, b)$*, so ist* f *auf* $[a, b]$ *monoton fallend (bzw. streng monoton fallend).*

(iii) *Umgekehrt gilt* f *monoton wachsend (bzw. monoton fallend)* $\Longrightarrow f'(x) \geq 0$ *(bzw.* $f'(x) \leq 0$*) für alle* $x \in (a, b)$*.*

Monotonie und Ableitung

Beweis.

(i-ii) Wir betrachten z.B. den Fall $f' > 0$ auf (a, b) und zeigen, dass f streng monoton wachsend ist.

- Wäre f nicht streng monoton wachsend, so gäbe es $a \leq x < y \leq b$ mit $f(x) \geq f(y)$.
- Wg. des MWS gibt es dann $\xi \in (x, y)$ mit

$$f'(\xi) = \frac{f(y) - f(x)}{y - x} \leq 0,$$

im Widerspruch zur Annahme $f' > 0$ auf (a, b).

(iii) Für die Umkehrung nehmen wir z.B. an, dass f monoton wachsend ist.

- Dann gilt für alle $x \in (a, b)$:

$$f'(x) = \lim_{\xi \searrow x} \frac{f(\xi) - f(x)}{\xi - x} \geq 0. \square$$

Monotonie und Ableitung

Bemerkung (zum vorhergehenden Satz)

- Aus dem streng monotonen Wachstum von f folgt nur $f' \geq 0$ und nicht die strikte Ungleichung.
- Beispielsweise ist die Funktion $f(x) = x^3$ streng monoton wachsend, aber $f'(0) = 0$.

Monotonie und Ableitung

Folgerung

- *Sei $f : (a, b) \to \mathbb{R}$ differenzierbar, $x \in (a, b)$ und $\epsilon > 0$ derart, dass $a < x - \epsilon < x + \epsilon < b$.*
- *Es gelte weiterhin*

$$(1) \quad f'(\xi) \leq 0 \quad (bzw. \quad \geq 0)$$

für alle ξ mit $x - \epsilon < \xi < x$ und

$$(2) \quad f'(\xi) \geq 0 \quad (bzw. \quad \leq 0)$$

für alle ξ mit $x < \xi < x + \epsilon$.

- *Dann hat f ein lokales Minimum (bzw. Maximum) an der Stelle x.*
- *Ersetzt man die Ungleichungen (1) und (2) durch strikte Ungleichungen, $f'(\xi) < 0$ (bzw. > 0) usw., so folgt, dass das lokale Extremum isoliert ist.*

Monotonie und Ableitung

Beweis.

- Aus (1) und (2) folgt, dass f auf $[x - \epsilon, x]$ monoton fallend (bzw. wachsend) und auf $[x, x + \epsilon]$ monoton wachsend (bzw. fallend) ist.
- Also hat f an der Stelle x ein lokales Minimum (bzw. Maximum).

$\square$

Höhere Ableitungen

- Eine differenzierbare Funktion f heißt zweimal differenzierbar, wenn f' differenzierbar ist.
- Die Ableitung $f'' := (f')'$ von f' heißt zweite Ableitung von f.
- Allgemein definiert man rekursiv $f^{(0)} := f$ und für alle $k \in \mathbb{N}$ die k-te Ableitung $f^{(k)}$ von f als

  $$f^{(k)} := (f^{(k-1)})',$$

  falls $f^{(k-1)}$ existiert und differenzierbar ist. Man sagt dann, dass f k-mal differenzierbar ist. Falls zusätzlich $f^{(k)}$ stetig ist, so heißt f k-mal stetig differenzierbar.
- *Bemerkung:* Jede k-mal differenzierbare Funktion ist (k-1)-mal stetig differenzierbar ($k \in \mathbb{N}$).

Zweite Ableitung und Extrema

Satz

- *Sei $f : (a,b) \to \mathbb{R}$ zweimal differenzierbar und $x \in (a,b)$, derart dass $f'(x) = 0$ und $f''(x) > 0$ (bzw. $f''(x) < 0$).*
- *Dann hat f ein isoliertes lokales Minimum (bzw. Maximum) an der Stelle x.*

Beweis.

- Wir betrachten z.B. den Fall $f''(x) > 0$.
- Wegen
$$0 < f''(x) = \lim_{\xi \to x} \frac{f'(\xi) - f'(x)}{\xi - x}$$
gibt es $\epsilon > 0$, so dass $\frac{f'(\xi)-f'(x)}{\xi - x} > 0$ für alle ξ mit $|x - \xi| < \epsilon$.
- Mit anderen Worten gilt also $f'(\xi) < f'(x) = 0$ für alle $\xi \in (x - \epsilon, x)$ und $f'(\xi) > f'(x) = 0$ für alle $\xi \in (x, x + \epsilon)$.
- $\Longrightarrow$ f hat an der Stelle x ein isol. lok. Min. (vgl. S. 171). $\square$

Beispiele

(i) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, gilt $f'(0) = 0$ und $f''(0) = 2 > 0$. Also hat f an der Stelle 0 ein isoliertes lokales Minimum. Das ist auch das globale Minimum der Funktion.

(ii) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$, hat ebenfalls an der Stelle 0 ihr (isoliertes) globales Minimum. In diesem Fall gilt jedoch $f''(0) = 0$.

Konvexität

Definition

- *Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt konvex, wenn für alle $x_1, x_2 \in I$ mit $x_1 < x_2$ gilt*

$$(*) \quad f(x) \leq s(x) \quad \textit{für alle} \quad x \in (x_1, x_2),$$

- *wobei $s(x) = f(x_1)\frac{x_2 - x}{x_2 - x_1} + f(x_2)\frac{x - x_1}{x_2 - x_1}$ die Sekante durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$ ist.*

Bemerkung

- Die Bedingung $(*)$ läßt sich folgendermaßen umformulieren

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2) \quad \text{für alle} \quad t \in (0,1).$$

- Um das einzusehen genügt es $t := \frac{x - x_1}{x_2 - x_1}$ zu setzen. Daraus folgt $1 - t = \frac{x_2 - x}{x_2 - x_1}$ und $(1-t)x_1 + tx_2 = x$.

Konvexität

Satz

Sei $f : I \to \mathbb{R}$ eine zweimal differenzierbare Funktion auf einem Intervall $I \subset \mathbb{R}$. f ist genau dann konvex, wenn $f'' \geq 0$.

Beweis.

"$\Longrightarrow$" Wir nehmen an, es gäbe $x_0 \in I$ mit $f''(x_0) < 0$ und führen diese Annahme zum Widerspruch.

- Sei $g(x) = f(x_0) + f'(x_0)(x - x_0)$ die affine Approximation von f an der Stelle x_0.
- Die Hilfsfunktion $h := f - g$ nimmt an der Stelle x_0 ein isoliertes lokales Maximum an, denn
  $h'(x_0) = f'(x_0) - g'(x_0) = f'(x_0) - f'(x_0) = 0$ und
  $h''(x_0) = f''(x_0) - g''(x_0) = f''(x_0) - 0 = f''(x_0) < 0$.
- Also gibt es $x_1, x_2 \in I$, so dass $x_1 < x_0 < x_2$, $f(x_0) > f(x_1)$ und $f(x_0) > f(x_2)$.
- Daraus folgt $f(x_0) > (1 - t)f(x_1) + tf(x_2)$ für alle $t \in (0, 1)$, im Widerspruch zur Konvexität von f.

Konvexität

Weiter im Beweis:

"$\Longleftarrow$" Wir zeigen nun umgekehrt, dass aus $f'' > 0$ die Konvexität von f folgt.

- Seien $x_1, x_2 \in I$, $x_1 < x_2$, $x = (1-t)x_1 + tx_2$, $t \in (0,1)$.
- Wegen des Mittelwertsatzes gibt es $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$, so dass

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \underset{(f''>0)}{\leq} f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}$$

- Also
$$f(x) \leq f(x_1)\frac{x_2 - x}{x_2 - x_1} + f(x_2)\frac{x - x_1}{x_2 - x_1}.$$
- Das ist genau die Konvexität von f. $\square$

Im Folgenden sei immer $a, b \in \mathbb{R}$, $a < b$.

Definition

Eine Funktion $\varphi : [a, b] \to \mathbb{R}$ *heißt* *Treppenfunktion*, *wenn es eine Unterteilung* $Z : a = x_0 < x_1 < \cdots < x_n = b$ *des Intervalls* $[a, b]$ *gibt, so dass* φ *auf jedem der Teilintervalle* (x_{i-1}, x_i) *konstant ist,* $i = 1, 2, \ldots, n$. *(→ Zeichnung)*

Übungsaufgabe

Seien $\varphi, \psi : [a, b] \to \mathbb{R}$ Treppenfunktionen. Dann sind $\varphi + \psi$ und $\varphi\psi : [a, b] \to \mathbb{R}$ Treppenfunktionen.

Integral einer Treppenfunktion

Definition

- *Sei $\varphi : [a, b] \to \mathbb{R}$ eine Treppenfunktion und $Z : a = x_0 < x_1 < \cdots < x_n = b$ eine Unterteilung des Intervalls $[a, b]$, so dass $\varphi|_{(x_{i-1}, x_i)} = c_i = const$, $i = 1, 2, \ldots, n$.*
- *Man definiert dann*

$$\int_a^b \varphi(x)dx := \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

Übungsaufgabe

Überlegen Sie sich, dass diese Definition nicht von der Wahl der Unterteilung abhängt und dass für alle $c \in (a, b)$

$$\int_a^b \varphi(x)dx = \int_a^c \varphi(x)dx + \int_c^b \varphi(x)dx.$$

Geometrische Interpretation

- Sei F das zwischen der x-Achse und dem Graphen der Funktion φ liegende Gebiet.
- F ist eine endliche Vereinigung von Rechtecken ($\rightarrow$ Zeichnung). Sei $A(F)$ der Flächeninhalt von F.
- Wenn $\varphi \geq 0$, so ist $\int_a^b \varphi(x)dx = A(F) \geq 0$.
- Wenn $\varphi \leq 0$, so ist $\int_a^b \varphi(x)dx = -A(F) \leq 0$.
- D.h. die Fläche oberhalb der x-Achse trägt positiv, die unterhalb der x-Achse negativ zum Integral bei.

Linearität und Monotonie des Integrals von Treppenfunktionen

Satz

- *$\varphi, \psi : [a, b] \to \mathbb{R}$ seien Treppenfunktionen und $\lambda \in \mathbb{R}$.*

(i) $\int_a^b (\varphi + \psi)(x)dx = \int_a^b \varphi(x)dx + \int_a^b \psi(x)dx.$

(ii) $\int_a^b (\lambda\varphi)(x)dx = \lambda \int_a^b \varphi(x)dx.$

(iii) $\varphi \le \psi \Longrightarrow \int_a^b \varphi(x)dx \le \int_a^b \psi(x)dx.$

Beweis.

- Sei $a = x_0 < x_1 < \cdots < x_n = b$ eine Unterteilung, so dass $\varphi|_{(x_{i-1},x_i)} = c_i$ und $\psi|_{(x_{i-1},x_i)} = d_i$.

(i) $\int_a^b (\varphi + \psi)(x)dx = \sum_{i=1}^n (c_i + d_i)(x_i - x_{i-1})$
$= \sum_{i=1}^n c_i(x_i - x_{i-1}) + \sum_{i=1}^n d_i(x_i - x_{i-1})$
$= \int_a^b \varphi(x)dx + \int_a^b \psi(x)dx.$

Linearität und Monotonie des Integrals von Treppenfunktionen

Weiter im Beweis:

(ii) $\int_a^b (\lambda\varphi)(x)dx = \sum_{i=1}^n (\lambda c_i)(x_i - x_{i-1}) = \lambda \sum_{i=1}^n c_i(x_i - x_{i-1})$
$= \lambda \int_a^b \varphi(x)dx.$

(iii) $\varphi \le \psi \Longrightarrow c_i \le d_i$ für alle $i \in \{1, 2, \dots, n\}$.
$\Longrightarrow \int_a^b \varphi(x)dx = \sum_{i=1}^n c_i(x_i - x_{i-1}) \le \sum_{i=1}^n d_i(x_i - x_{i-1}) =$
$\int_a^b \psi(x)dx.$ $\square$

Ober- und Unterintegral

Definition

- *Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion.*
- *Das Oberintegral von f ist die Zahl*

$$\int_a^{*b} f(x)dx := \inf \left\{ \int_a^b \varphi(x)dx \middle| \varphi \quad \textit{Treppenfkt. mit} \quad \varphi \geq f \right\}.$$

- *Das Unterintegral von f ist die Zahl*

$$\int_{*a}^{b} f(x)dx := \sup \left\{ \int_a^b \varphi(x)dx \middle| \varphi \quad \textit{Treppenfkt. mit} \quad \varphi \leq f \right\}.$$

Bemerkungen/Übungsaufgaben

- $\int_{*a}^{b} f(x)dx \leq \int_a^{*b} f(x)dx$.
- $\int_{*a}^{b} \varphi(x)dx = \int_a^{*b} \varphi(x)dx = \int_a^b \varphi(x)dx$ für jede Treppenfkt.

Ober- und Unterintegral

Beispiel

- Sei $f : [0,1] \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 1, & \text{wenn} \quad x \in \mathbb{Q} \\ 0, & \text{sonst.} \end{cases}$$

- Dann gilt

$$\int_{*a}^{b} f(x)dx = 0 \quad \text{und} \quad \int_{a}^{*b} f(x)dx = 1.$$

Eigenschaften des Ober- und Unterintegrals

Satz

- *$f, g : [a.b] \to \mathbb{R}$ seien beschränkte Funktionen. Dann gilt:*

(i) $\int^*(f+g) \leq \int^* f + \int^* g$,

(ii) $\int^*(\lambda f) = \lambda \int^* f$ *für alle* $\lambda \geq 0$,

(iii) $\int_*(f+g) \geq \int_* f + \int_* g$,

(iv) $\int_*(\lambda f) = \lambda \int_* f$ *für alle* $\lambda \geq 0$,

(v) $\int^*(\lambda f) = \lambda \int_* f$ *und* $\int_*(\lambda f) = \lambda \int^* f$ *für alle* $\lambda \leq 0$.

(Hierbei ist $\int^* f = \int_a^{*b} f(x)dx$*, usw.)*

Beweis.

- Für jede Teilmenge $D \subset \mathbb{R}$ gilt $\sup D = -\inf(-D)$, wobei $-D := \{-x | x \in D\}$.
- Insbesondere gilt $\int_* f = -\int^*(-f)$.
- (iii-v) folgt damit aus (i-ii).

Eigenschaften des Ober- und Unterintegrals

Weiter im Beweis:

(i) Wir zeigen $\int^*(f+g) \le \int^* f + \int^* g$:
$\int^*(f+g) = \inf\{\int \xi | \xi \ge f+g \text{ Treppenfkt.}\}$
$\le \inf\{\int \xi | \xi = \varphi + \psi, \varphi \ge f, \psi \ge g \text{ Treppenfktn.}\}$
$= \inf\{\int \varphi + \int \psi | \varphi \ge f, \psi \ge g \text{ Treppenfktn.}\}$
$= \inf\{\int \varphi | \varphi \ge f \text{ Treppenfkt.}\} + \inf\{\int \psi | \psi \ge g \text{ Treppenfkt.}\}$
$= \int^* f + \int^* g.$

(ii) Wir zeigen $\int^*(\lambda f) = \lambda \int^* f$ für $\lambda > 0$:
$\int^*(\lambda f) = \inf\{\int \varphi | \varphi \ge \lambda f \text{ Treppenfkt.}\}$
$= \inf\{\int \varphi | \frac{1}{\lambda}\varphi \ge f \text{ Treppenfkt.}\}$
$= \inf\{\lambda \int \psi | \psi \ge f \text{ Treppenfkt.}\}$
$= \lambda \inf\{\int \psi | \psi \ge f \text{ Treppenfkt.}\} = \lambda \int^* f.$ $\square$

Integrierbare Funktionen

Definition

- *Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ heißt integrierbar (genauer: Riemann-integrierbar), wenn*

$$\int_a^{*b} f(x)dx = \int_{*a}^{b} f(x)dx.$$

- *Man definiert dann das Integral von f als*

$$\int_a^b f(x)dx := \int_a^{*b} f(x)dx = \int_{*a}^{b} f(x)dx.$$

Linearität und Monotonie des Integrals

Satz

- *$f, g : [a, b] \to \mathbb{R}$ seien integrierbar und $\lambda \in \mathbb{R}$. Dann sind $f + g$ und λf integrierbar und es gilt:*

(i) $\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx,$

(ii) $\int_a^b (\lambda f)(x)dx = \lambda \int_a^b f(x)dx$ *und*

(iii) $f \leq g \Longrightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx.$

Linearität und Monotonie des Integrals

Beweis.

(i) $\int^*(f+g) \leq \int^* f + \int^* g = \int f + \int g = \int_* f + \int_* g \leq$
$\int_*(f+g) \leq \int^*(f+g)$
$\Longrightarrow \int^*(f+g) = \int_*(f+g) = \int f + \int g$.

(ii) Für $\lambda > 0$ gilt $\int^*(\lambda f) = \lambda \int^* f = \lambda \int f = \lambda \int_* f = \int_*(\lambda f)$.

- Für $\lambda < 0$ gilt $\int^*(\lambda f) = \lambda \int_* f = \lambda \int f = \lambda \int^* f = \int_*(\lambda f)$.
- In beiden Fälllen ist also λf integrierbar und $\int(\lambda f) = \lambda \int f$.

(iii) $f \leq g \Longrightarrow \int f = \int_* f = \sup\{\int \varphi | \varphi \leq f \text{ Treppenfkt.}\}$
$\leq \sup\{\int \varphi | \varphi \leq g \text{ Treppenfkt.}\} = \int_* g = \int g$.

□

Bemerkungen/Übungsaufgaben:

(i) Sei $a < b < c$. Dann ist $f : [a, c] \to \mathbb{R}$ integrierbar, genau dann wenn $f|_{[a,b]}$ und $f|_{[b,c]}$ integrierbar sind.

- Es gilt dann
$$\int_a^c f = \int_a^b f + \int_b^c f.$$
- Wir setzen $\int_b^a = -\int_a^b$ und $\int_a^a = 0$.

(ii) Positiv- und Negativteil $f_\pm : [a, b] \to \mathbb{R}_+ \cup \{0\}$ einer Fkt. $f : [a, b] \to \mathbb{R}$ sind definiert als $f_+(x) := \max\{f(x), 0\}$ bzw. $f_-(x) := -\min\{f(x), 0\}$, $x \in [a, b]$.

- Es gilt: f integrierbar $\iff$ f_+ und f_- integrierbar $\implies$ $|f|$ integrierbar.

(iii) Aus der Monotonie des Integrals folgt
$$\left| \int_a^b f \right| \leq \int_a^b |f|$$
für jede integrierbare Funktion f.

Integrierbarkeit stetiger Funktionen

Theorem

Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ ist integrierbar.

Beweis.

- Es genügt zu zeigen, dass es für alle $\epsilon > 0$ Treppenfunktionen $\varphi, \psi : [a,b] \to \mathbb{R}$ gibt mit $\varphi \leq f \leq \psi$ und $\int \psi - \int \varphi < \epsilon$.
- Dazu benutzen wir folgenden Hilfssatz:

Lemma

- *Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ ist gleichmäßig stetig, d.h.*
- *Für alle $\epsilon > 0$ existiert $\delta > 0$, so dass für alle $x, y \in [a,b]$:*

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon.$$

Beweis des Lemmas

- Wir nehmen an, f sei nicht gleichmäßig stetig und führen das zum Widerspruch.
- Wir nehmen also an, es gibt $\epsilon > 0$ und Folgen $x_n, y_n \in [a, b]$ mit
$$|x_n - y_n| < \frac{1}{n} \quad \text{und} \quad |f(x_n) - f(y_n)| \geq \epsilon.$$
- Nach Bolzano-Weierstraß existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, die gegen $\ell \in [a, b]$ konvergiert.
- Wegen $\lim_{n\to\infty} |x_n - y_n| = 0$ gilt auch $\lim_{k\to\infty} y_{n_k} = \ell$.
- Die Stetigkeit von f liefert dann
$$\lim_{k\to\infty} f(x_{n_k}) = f(\ell) = \lim_{k\to\infty} f(y_{n_k}),$$
- im Widerspruch zu $|f(x_n) - f(y_n)| \geq \epsilon$ für alle $n \in \mathbb{N}$. $\square$

Beweis des Theorems

- Wir betrachten die äquidistante Unterteilung $x_i = a + i\frac{(b-a)}{n}, \quad i = 0, 1, \ldots, n.$
- und setzen $c_i := \min f|_{[x_{i-1}, x_i]}$ und $d_i := \max f|_{[x_{i-1}, x_i]}$.
- Nach dem Lemma gibt es $N \in \mathbb{N}$, so dass

$$d_i - c_i < \frac{\epsilon}{b-a} \quad \text{für alle} \quad n \geq N.$$

- Wir definieren

$$\varphi|_{[x_{i-1}, x_i)} := c_i \quad \text{und} \quad \psi|_{[x_{i-1}, x_i)} := d_i, \quad i = 1, \ldots, n$$

  und $\varphi(b) := \psi(b) := f(b)$.
- Offenbar gilt dann $\varphi \leq f \leq \psi$ und

$$\int_a^b \psi(x)dx - \int_a^b \varphi(x)dx = \sum_{i=1}^n \underbrace{(d_i - c_i)}_{<\epsilon/(b-a)} \frac{b-a}{n} < \epsilon. \square$$

Übungsaufgabe

Jede monotone Funktion $f : [a, b] \to \mathbb{R}$ ist integrierbar.

Riemannsche Summen

Definition

- *Sei* $f : [a, b] \to \mathbb{R}$ *eine Funktion,* $a = x_0 < x_1 < \cdots < x_n = b$ *eine Unterteilung des Intervalls* $[a, b]$ *und* $\xi_i \in [x_{i-1}, x_i]$*.*
- *Die zugehörige* ***Riemannsche Summe*** *ist die Zahl*

$$\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

Approximation des Integrals durch Riemannsche Summen

Satz

- *Für jede stetige Funktion* $f : [a,b] \to \mathbb{R}$ *approximiert die Riemannsche Summe in folgendem Sinne das Integral:*
- *Für alle* $\epsilon > 0$ *gibt es* $\delta > 0$*, so dass*

$$\left| \int_a^b f(x)dx - \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \right| < \epsilon$$

für jede Riemannsche Summe mit $\max_{1 \leq i \leq n} (x_i - x_{i-1}) < \delta$*.*

Bemerkung

Der Satz gilt sogar für beliebige (Riemann-) integrierbare Funktionen $f : [a,b] \to \mathbb{R}$, vgl. Forster.

Approximation des Integrals durch Riemannsche Summen

Beweis.

- Wegen der (gleichmäßigen) Stetigkeit von $f : [a, b] \to \mathbb{R}$ gibt es $\delta > 0$, so dass $|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon/(b - a)$.
- Insbesondere folgt aus $\max\limits_{1 \le i \le n}(x_i - x_{i-1}) < \delta$:
  $|f(x) - f(\xi_i)| < \epsilon/(b - a)$ für alle $x \in [x_{i-1}, x_i]$.
- Damit erhalten wir

$$\left| \int_a^b f(x)dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| = \left| \sum_{i=1}^n \int_{x_{i-1}}^{x_i} (f(x) - f(\xi_i))dx \right|$$

$$\le \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \underbrace{|f(x) - f(\xi_i)|}_{<\epsilon/(b-a)} dx < \sum_{i=1}^n (x_i - x_{i-1}) \frac{\epsilon}{b - a} = \epsilon.$$

□

Fundamentalsatz der Differenzial- und Integralrechnung

Im Folgenden sei $I \subset \mathbb{R}$ immer ein Intervall.

Definition

Sei $f : I \to \mathbb{R}$ eine Funktion. Eine **Stammfunktion** *von f ist eine differenzierbare Funktion $F : I \to \mathbb{R}$, so dass*

$$F' = f.$$

Theorem

- *Sei $f : I \to \mathbb{R}$ eine stetige Funktion und $x_0 \in I$.*
- *Die durch*
$$F(x) := \int_{x_0}^{x} f(t)dt$$
definierte Funktion $F : I \to \mathbb{R}$ ist eine Stammfunktion von f.

Beweis des Fundamentalsatzes

Wir benutzen folgendes Lemma:

Lemma (Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann existiert $\xi \in [a, b]$, so dass

$$\int_a^b f(x)dx = f(\xi)(b - a).$$

Beweis.

- Aus $\min f \leq f \leq \max f$ folgt wg. der Monotonie des Integrals

$$(b - a) \min f \leq \int_a^b f(x)dx \leq (b - a) \max f.$$

- Der Zwischenwertsatz liefert dann die Existenz von $\xi \in [a, b]$ mit $f(\xi) = \int_a^b f(x)dx/(b - a)$.

$\square$

Beweis des Fundamentalsatzes

- Sei $0 \neq h \in \mathbb{R}$ derart, dass $x, x+h \in I$. Dank des Lemmas gilt

  $$\int_x^{x+h} f(t)dt = f(\xi)h,$$

  wobei $\xi \in [x, x+h]$, falls $h > 0$ und $\xi \in [x+h, x]$, falls $h < 0$.
- Damit erhalten wir

  $$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= \frac{\int_{x_0}^{x+h} f(t)dt - \int_{x_0}^{x} f(t)dt}{h} \\ &= \frac{\int_x^{x+h} f(t)dt}{h} = f(\xi) \to f(x) \quad (h \to 0), \end{aligned}$$

  d.h. $F'(x) = f(x)$. $\square$

Fundamentalsatz der Differenzial- und Integralrechnung

Folgerung

- *Sei $f : I \to \mathbb{R}$ eine stetige Funktion und F eine Stammfunktion von f.*
- *Dann gilt für alle $a, b \in I$:*

$$\int_a^b f(x)dx = F(b) - F(a).$$

Beweis.

- Aus
$$(F(x) - \int_{x_0}^x f(t)dt)' = F'(x) - f(x) = 0$$
folgt $F(x) = \int_{x_0}^x f(t)dt + const$ und somit $F(b) - F(a) = \int_{x_0}^b f(t)dt + const - (\int_{x_0}^a f(t)dt + const) = \int_a^b f(t)dt$. $\square$

Notation

$$F|_a^b := F(b) - F(a).$$

Theorem (Substitutionsregel)

- Sei $f : I \to \mathbb{R}$ eine stetige Funktion und $\varphi : [a, b] \to \mathbb{R}$ eine stetig differenzierbare Funktion mit $\varphi([a, b]) \subset I$. Dann gilt
-
$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Beweis.

- Sei $F : I \to \mathbb{R}$ eine Stammfunktion von f.
- Nach der Kettenregel gilt für alle $t \in [a, b]$
$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$
- Somit $\int_a^b f(\varphi(t))\varphi'(t)dt = \int_a^b (F \circ \varphi)'(t)dt = F \circ \varphi|_a^b = F|_{\varphi(a)}^{\varphi(b)} = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$ □

Integration durch Substitution

Beispiel

- $$\begin{aligned} \int_0^2 \frac{2t}{t^2+1} dt &\underset{(x=t^2)}{=} \int_0^4 \frac{dx}{x+1} = \ln(x+1)\Big|_0^4 \\ &= \ln 5 - \ln 1 = \ln 5 \end{aligned}$$
- Hier ist $f(x) = \frac{1}{x+1}$ und $\varphi(t) = t^2$.

Partielle Integration

Theorem

- $f, g : [a, b] \to \mathbb{R}$ *seien stetig differenzierbar.*
- *Dann gilt*

$$\int_a^b f'(x)g(x)dx = fg|_a^b - \int_a^b f(x)g'(x)dx.$$

Beweis.

- Das folgt aus der Produktregel $(fg)' = f'g + fg'$ durch Integration:
- $$fg|_a^b = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx.$$

□

Partielle Integration

Beispiel

- $$\begin{aligned} \int_a^b \ln x \, dx &= \int_a^b 1 \cdot \ln x \, dx \\ &= x \ln x|_a^b - \int_a^b \underbrace{x \cdot \frac{1}{x}}_{=1} dx = x(\ln x - 1)|_a^b \end{aligned}$$
- Hier ist $f(x) = x$ und $g(x) = \ln x$.

Vektorräume

Definition

- *Sei $\mathbb{K}$ ein Körper, z.B. $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$.*
- *Ein Vektorraum über $\mathbb{K}$ ist eine Menge V auf der zwei Verknüpfungen gegeben sind,*
- *die Addition*

$$+ : V \times V \to V, \quad (v, w) \mapsto v + w$$

- *und die skalare Multiplikation*

$$\mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v \ (=: \lambda v),$$

so dass

1) $(V, +)$ eine kommutative Gruppe ist, d.h. es gilt:

   (i) $v + w = w + v$ für alle $v, w \in V$ (Kommutativgesetz),

   (ii) $(u + v) + w = u + (v + w)$ für alle $u, v, w \in V$ (Assoziativgesetz),

   (iii) es gibt $0 \in V$, so dass $v + 0 = v$ für alle v und

   (iv) zu jedem $v \in V$ existiert $-v$, so dass $v + (-v) = 0$. (Wie im Fall von $(\mathbb{R}, +)$ zeigt man, dass die Eindeutigkeit des neutralen Elements 0 und des additiven Inversen $-v$ von v.)

2) Für alle $\lambda, \mu \in \mathbb{K}$, $v, w \in V$ gilt:

   (i) $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$,

   (ii) $1 \cdot v = v$,

   (iii) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,

   (iv) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

Definition

Die Elemente von $\mathbb{K}$ heißen Skalare, die von V Vektoren.

Beispiele von Vektorräumen

1) Der kartesische Raum

- $\mathbb{K}^n = \{(x_1, \dots, x_n) | x_1, \dots, x_n \in \mathbb{K}\}$
- mit der Addition

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

- und der skalaren Multiplikation

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n)$$

  ist ein Vektorraum über $\mathbb{K}$.
- Das neutrale Element 0 der kommutativen Gruppe $(\mathbb{K}^n, +)$ ist der Vektor $(0, \dots, 0)$ und das additive Inverse $-(x_1, \dots, x_n)$ des Vektors $(x_1, \dots, x_n)$ ist $(-x_1, \dots, -x_n)$.

Beispiele von Vektorräumen

2) Funktionenräume

- Sei X eine Menge und $F(X, \mathbb{K})$ die Menge der Funktionen $f : X \to \mathbb{K}$.
- $F(X, \mathbb{K})$ ist mit der üblichen, punktweise definierten, Addition und skalaren Multiplikation

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ (\lambda f)(x) &= \lambda f(x) \quad (f, g \in F(X, \mathbb{K}), \quad \lambda \in \mathbb{K}, \quad x \in X) \end{aligned}$$

ein Vektorraum über $\mathbb{K}$.

Übungsaufgabe

- Sei V ein Vektorraum über $\mathbb{K}$. Für alle $v \in V$, $\lambda \in \mathbb{K}$ gilt
  (i) $\lambda v = 0 \iff \lambda = 0$ oder $v = 0$,
  (ii) $-v = (-1) \cdot v$.

Unterräume

Definition

- *Ein Unterraum (genauer: ein Untervektorraum) eines Vektorraumes V ist eine nicht leere Teilmenge $U \subset V$, so dass*

(i) $v + w \in U$ *für alle* $v, w \in U$ *und*

(ii) $\lambda v \in U$ *für alle* $\lambda \in \mathbb{K}, v \in U$.

Bemerkung

Wegen (i) und (ii) induzieren die Addition und die skalare Multiplikation in V eine Addition

$$+ : U \times U \to U$$

und eine skalare Multiplikation

$$\cdot : \mathbb{K} \times U \to U,$$

die U zu einem Vektorraum machen.

Beispiele von Unterräumen

- Sei V ein Vektorraum.

1) $\{0\}$ und V sind Untervektorräume von V.
2) Für jeden Vektor $v \in V$ ist $\mathbb{K}v = \{\lambda v | \lambda \in \mathbb{K}\} \subset V$ der kleinste Unterraum, der v enthält.

- Wenn $v \neq 0$, so ist $\mathbb{K}v \neq \{0\}$ und heißt die von v erzeugte (lineare) Gerade.

3) Für jedes Paar von Vektoren v, w ist

$$\mathbb{K}v + \mathbb{K}w = \{\lambda v + \mu w | \lambda, \mu \in \mathbb{K}\} \subset V$$

der kleinste Unterraum, der v und w enthält.

- Wenn $v \neq 0$ und $w \notin \mathbb{K}v$, so heißt $\mathbb{K}v + \mathbb{K}w \underset{\neq}{\supset} \mathbb{K}v \underset{\neq}{\supset} \{0\}$ die von v und w erzeugte (lineare) Ebene.
- Die (affine) Gerade $w + \mathbb{K}v$ ist dann *kein* Unterraum des Vektorraums V.

Beispiele von Unterräumen

4) Sei $I \subset \mathbb{R}$ ein Intervall und $k \in \mathbb{N}$. Die Folgenden Mengen sind Unterräume von $F(I, \mathbb{R})$:

$$\begin{aligned} C^k(I, \mathbb{R}) &:= \{f : I \to \mathbb{R} | f \quad \text{k-mal stetig differenzierbar}\} \\ &\subset \quad Diff(I, \mathbb{R}) := \{f : I \to \mathbb{R} | f \quad \text{differenzierbar}\} \\ &\subset \quad C(I, \mathbb{R}) := \{f : I \to \mathbb{R} | f \quad \text{stetig}\} \end{aligned}$$

5) Sei V ein Vektorraum und $U_j \subset V$ durch $j \in J$ indizierte Unterräume, wobei J eine beliebige Menge ist.

- Dann ist der Durchschnitt $U = \cap_{j \in J} U_j \subset V$ ein Unterraum (ÜA).
- Beispielsweise ist

$$C^\infty(I, \mathbb{R}) := \cap_{k \in \mathbb{N}} C^k(I, \mathbb{R})$$

ein Unterraum. Die Funktionen in $C^\infty(I, \mathbb{R})$ heißen unendlich oft differenzierbar.

Beispiele von Unterräumen

6) Ein Polynom in der Variablen x mit Koeffizienten in einem Körper $\mathbb{K}$ ist ein Ausdruck der Form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

wobei $a_0, \ldots, a_n \in \mathbb{K}$ und $n \in \mathbb{N} \cup \{0\}$.

- Die Menge aller solchen Polynome wird mit $\mathbb{K}[x]$ bezeichnet und bildet in offensichtlicher Weise (ÜA) einen Vektorraum.
- Jedes Polynom $a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{K}[x]$ definiert eine Funktion

$$P : \mathbb{K} \to \mathbb{K}, \quad \lambda \mapsto P(\lambda) = a_0 + a_1 \lambda + \cdots + a_n \lambda^n.$$

  Solche Funktionen P heißen polynomial.
- Wenn $\mathbb{K}$ unendlich ist, so ist diese Zuordnung $a_0 + \cdots + a_n x^n \mapsto P$ injektiv (ÜA) und $\mathbb{K}[x]$ wird auf diese Weise zu einem Unterraum $\mathbb{K}[x] \subset F(\mathbb{K}, \mathbb{K})$. Z.B. haben wir die Inklusion $\mathbb{R}[x] \subset C^\infty(\mathbb{R}, \mathbb{R}) \subset F(\mathbb{R}, \mathbb{R})$.

Lineare Unabhängigkeit

Definition

(i) *Sei V ein Vektorraum und $v_1, \ldots, v_r \in V$ eine endliche Familie von Vektoren.*

- *Man sagt, dass $v \in V$ eine Linearkombination der Vektoren $v_1, \ldots, v_r$ ist, wenn es $\lambda_i \in \mathbb{K}$ gibt, so dass*

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

- *Die Vektoren $v_1, \ldots, v_r$ heißen linear unabhängig, wenn*

$$\sum_{i=1}^{r} \lambda_i v_i = 0 \Longrightarrow \lambda_1 = \cdots = \lambda_r = 0.$$

Lineare Unabhängigkeit

(ii) Sei nun $(v_j)_{j \in J}$ eine durch eine unendliche Menge J indizierte Familie von Vektoren $v_j \in V$.

- Man sagt, dass $v \in V$ eine Linearkombination der Vektoren $(v_j)_{j \in J}$ ist, wenn es eine endliche Teilmenge $J_0 \subset J$ gibt, so dass v eine Linearkombination der $(v_j)_{j \in J_0}$ ist.
- Die Vektoren $(v_j)_{j \in J}$ heißen linear unabhängig, wenn die Vektoren $(v_j)_{j \in J_0}$ für jede endliche Teilmenge $J_0 \subset J$ linear unabhängig sind.

Lineare Unabhängigkeit

Beispiele

1) Wir setzen

$$e_i := (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{K}^n.$$
$$\text{(i-te Stelle)}$$

- Die Vektoren $(e_1, \dots, e_n)$ sind linear unabhängig. In der Tat:

$$0 = \sum_{i=1}^{n} \lambda_i e_i = (\lambda_1, \dots, \lambda_n) \Longrightarrow \lambda_1 = \dots = \lambda_n = 0.$$

2) Die Monome $(1, x, x^2, x^3, \dots)$ sind linear unabhängig in $\mathbb{K}[x]$.

Lineare Hülle

Definition

- *Sei V ein Vektorraum und $(v_j)_{j \in J}$ eine Familie von Vektoren.*
- *Die Lineare Hülle der $(v_j)_{j \in J}$ ist der Unterraum*

$$\mathrm{span}\{v_j | j \in J\} := \{v | v \textit{ ist eine Linearkombination der } (v_j)_{j \in J}\}.$$

Satz

- *Die Vektoren $v_j \in V$, $j \in J$, seien linear unabhängig.*
- *Dann hat jeder Vektor $v \in \mathrm{span}\{v_j | j \in J\}$ eine eindeutige Darstellung*

$$v = \sum_{j \in J} \lambda_j v_j, \quad \lambda_j \in \mathbb{K},$$

wobei $\lambda_j = 0$ für fast alle $j \in J$, d.h. für alle bis auf endlich viele $j \in J$.

Beweis.

- Nach Definition der linearen Hülle gibt es eine endliche Teilmenge $J_0 \subset J$ und eine Darstellung

$$v = \sum_{j \in J_0} \lambda_j v_j.$$

- Sei $v = \sum_{j \in J_0'} \lambda_j' v_j$ eine weitere solche Darstellung mit endlichem $J_0' \subset J$.

Weiter im Beweis:

- Dann gilt

$$\begin{aligned} 0 &= v - v = \sum_{j\in J_0} \lambda_j v_j - \sum_{j\in J_0'} \lambda_j' v_j \\ &= \sum_{j\in J_0\cap J_0'} (\lambda_j - \lambda_j')v_j + \sum_{j\in J_0\setminus J_0\cap J_0'} \lambda_j v_j - \sum_{j\in J_0'\setminus J_0\cap J_0'} \lambda_j' v_j. \end{aligned}$$

- Wegen der linearen Unabhängigkeit der v_j, $j \in J$, und der Endlichkeit der Menge $J_0 \cup J_0'$ folgt

$$\begin{aligned} \lambda_j - \lambda_j' &= 0, \quad \text{wenn} \quad j \in J_0 \cap J_0' \\ \lambda_j &= 0, \quad \text{wenn} \quad j \in J_0 \setminus J_0 \cap J_0' \\ \lambda_j' &= 0, \quad \text{wenn} \quad j \in J_0' \setminus J_0 \cap J_0'. \end{aligned}$$

- Also stimmen die beiden Darstellungen
$v = \sum_{j\in J_0} \lambda_j v_j = \sum_{j\in J_0\cap J_0'} \lambda_j v_j$ und
$v = \sum_{j\in J_0'} \lambda_j' v_j = \sum_{j\in J_0\cap J_0'} \lambda_j' v_j$ überein. $\square$

Erzeugendensysteme, Basen

Definition

- *Sei V ein Vektorraum. Eine Familie von Vektoren $(v_j)_{j\in J}$ heißt Erzeugendensystem von V, falls $V = \operatorname{span}\{v_j | j \in J\}$.*
- *Eine Basis von V ist ein linear unabhängiges Erzeugendensystem.*

Beispiele

1) $(e_1, \ldots, e_n)$ ist eine Basis von $\mathbb{K}^n$, die so genannte kanonische Basis.
2) $(1, x, x^2, \ldots)$ ist eine Basis von $\mathbb{K}[x]$.
3) $(e_1, e_2, e_1 + e_2, 2e_1)$ ist ein Erzeugendensytem von $\mathbb{K}^2$, aus dem wir folgende Basen auswählen können: (e_1, e_2), $(e_1, e_1 + e_2)$ und $(e_2, e_1 + e_2)$. Falls in $\mathbb{K}$ $0 \neq 2$ $(:= 1 + 1)$ gilt, so sind $(e_2, 2e_1)$ und $(e_1 + e_2, 2e_1)$ ebenfalls Basen.

Charakterisierung von Basen

Satz

- *Sei $V \neq 0$ ein VR. Eine Familie $(v_j)_{j\in J}$ von Vektoren von V ist genau dann eine Basis, wenn sie eine der folgenden Bedingungen erfüllt.*

(i) *$(v_j)_{j\in J}$ ist ein minimales Erzeugendensystem, d.h. $J_0 = J$ ist die einzige Teilmenge von J, für die $(v_j)_{j\in J_0}$ ein Erzeugendensystem ist.*

(ii) *$(v_j)_{j\in J}$ ist eine maximale linear unabhängige Familie, d.h. $J_0 = J$ ist die einzige Obermenge $J_0 \supset J$, für die $(v_j)_{j\in J_0}$ linear unabhängig ist.*

(iii) *Jeder Vektor $v \in V$ besitzt genau eine Darstellung*

$$v = \sum_{j\in J} \lambda_j v_j,$$

wobei $\lambda_j = 0$ für fast alle $j \in J$.

Beweis.

- Sei (0) die Aussage, dass $(v_j)_{j\in J}$ eine Basis bildet. Wir zeigen (0) $\implies$ (*iii*) $\implies$ (*i*) $\implies$ (*ii*) $\implies$ (0).

$0 \implies iii$ Wir wissen bereits, dass jeder Vektor aus $\operatorname{span}\{v_j | j \in J\}$ eine eindeutige Darstellung als (endliche) Linearkombination der linear unabhängigen Vektoren $(v_j)_{j\in J}$ hat. Da $(v_j)_{j\in J}$ ein Erzeugendensystem ist, gilt das für jeden Vektor aus $V = \operatorname{span}\{v_j | j \in J\}$.

$iii \implies i$ Aus (iii) folgt offensichtlich, dass $(v_j)_{j\in J}$ ein Erzeugendensystem ist.

- Um die Minimalität zu zeigen, nehmen wir an, dass $(v_j)_{j\in J_0}$, $J_0 \subset J$, ein Erzeugendensystem ist.
- Dann hat jeder der Vektoren v_i, $i \in J$, eine Darstellung (*) $v_i = \sum_{j\in J_0} \lambda_j v_j$. Nach (iii) ist das die eindeutige Darstellung als Linearkombination der $(v_j)_{j\in J}$.
- Daraus folgt $i \in J_0$, denn sonst wären $v_i = v_i$ und (*) zwei verschiedene Darstellungen. Also $J = J_0$.

Weiter im Beweis

$i \Longrightarrow ii$ Wir zeigen zuerst, dass aus (i) die lineare Unabhängigkeit von $(v_j)_{j\in J}$ folgt.

- Sei $\sum_{j\in J_0} \lambda_j v_j = 0$, wobei $J_0 \subset J$ endlich ist.
- Wenn für ein $j_0 \in J_0 \quad \lambda_{j_0} \neq 0$ wäre, so wäre $v_{j_0} = -\sum_{j\in J_0\setminus\{j_0\}} \frac{\lambda_j}{\lambda_{j_0}} v_j$ und somit $(v_j)_{j\in J\setminus\{j_0\}}$ ein Erzeugendensystem, im Widerspruch zur Minimalität in (i).
- Also $\lambda_{j_0} = 0$ für alle $j_0 \in J_0$. Das zeigt die lineare Unabhängigkeit von $(v_j)_{j\in J}$.
- Als Nächstes zeigen wir die Maximalität der linear unabhängigen Familie $(v_j)_{j\in J}$.
- Sei also $(v_j)_{j\in J_0}$, $J \neq J_0 \supset J$, eine Familie, die die $(v_j)_{j\in J}$ enthält und $j_0 \in J_0 \setminus J$.
- Da $(v_j)_{j\in J}$ ein Erzeugendensystem ist, gibt es eine Darst. $v_{j_0} = \sum_{j\in J} \lambda_j v_j$.
- Das zeigt, dass $(v_j)_{j\in J_0}$ linear abhängig ist. Somit ist die lin. unabh. Familie $(v_j)_{j\in J}$ maximal.

Weiter im Beweis

$ii \Longrightarrow 0$ Sei $(v_j)_{j\in J}$ eine maximale lin. unabh. Familie.

- Wir zeigen, dass $(v_j)_{j\in J}$ ein Erzeugendensystem und somit eine Basis ist.
- Sonst gäbe es $v \in V \setminus \mathrm{span}\{v_j | j \in J\}$ und $(v, (v_j)_{j\in J})$ wäre eine linear unabhängige Familie, im Widerspruch zur Maximalität von $(v_j)_{j\in J}$. □

Folgerung

Jedes endliche Erzeugendensystem eines Vektorraums V enthält eine Basis.

Beweis.

Jedes endliche Erzeugendensystem enthält ein minimales Erzeugendensystem. □

Austauschsätze von Steinitz

Austauschlemma

- Sei $(v_1, \dots, v_n)$ eine Basis von V und
$$w = \sum_{i=1}^{n} \lambda_i v_i$$
eine Linearkombination.
- Wenn $\lambda_k \neq 0$, so ist $(v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_n)$ eine Basis von V.

Beweis.
Übungsaufgabe. $\square$

Austauschsatz

- Sei V ein VR, $(v_1, \ldots, v_n)$ eine Basis und $(w_1, \ldots, w_r)$ linear unabhängige Vektoren.
- Dann gilt $n \geq r$ und es gibt eine Permutation $\varphi : \{1, \ldots, n\} \to \{1, \ldots, n\}$, so dass $(w_1, \ldots, w_r, v_{\varphi(r+1)}, \ldots, v_{\varphi(n)})$ eine Basis ist.

Beweis durch Induktion nach r

- Für $r = 0$ ist nichts zu zeigen. Wir führen den Induktionsschritt $r \to r + 1$ aus.
- Seien also $(w_1, \ldots, w_{r+1})$ lin. unabh. und $(v_1, \ldots, v_n)$ eine Basis.
- Nach Induktionsvor. gilt $r \leq n$ und es gibt eine Permutation $\varphi : \{1, \ldots, n\} \to \{1, \ldots, n\}$, so dass $(w_1, \ldots, w_r, v_{\varphi(r+1)}, \ldots, v_{\varphi(n)})$ eine Basis ist.

Austauschsatz

Weiter im Beweis:

- Wir können daher schreiben

$$w_{r+1} = \sum_{i=1}^{r} \lambda_i w_i + \sum_{j=r+1}^{n} \lambda_j v_{\varphi(j)}.$$

- Da die $(w_1, \dots, w_{r+1})$ lin. unabh. sind, existiert $j_0 \in \{r+1, \dots, n\}$ mit $\lambda_{j_0} \neq 0$. Insbesondere ist $r+1 \leq n$.
- Wir können (ggf. nach Abänderung von φ) annehmen, dass $j_0 = r+1$.
- Nach dem Austauschlemma können wir dann in der Basis $(w_1, \dots, w_r, v_{\varphi(r+1)}, \dots, v_{\varphi(n)})$ den Vektor $v_{\varphi(j_0)} = v_{\varphi(r+1)}$ durch w_{r+1} ersetzen und erhalten die Basis $(w_1, \dots, w_{r+1}, v_{\varphi(r+2)}, \dots, v_{\varphi(n)})$. $\square$

Theorem

- *Sei V ein Vektorraum.*

(i) *Wenn V eine endliche Basis besitzt, dann ist jede Basis von V endlich.*

(ii) *Alle endlichen Basen von V haben die gleiche Anzahl von Elementen.*

Beweis.

(i) Sei $(v_1, \ldots, v_n)$ eine Basis. Aus dem Austauschsatz folgt, dass jede linear unabhängige Familie höchstens n Elemente hat.

(ii) Sei $(w_1, \ldots, w_r)$ eine zweite Basis. Aus dem Austauschsatz folgt, wie gesagt, $r \leq n$ und ebenso $n \leq r$.

$\square$

Dimension

Definition

- *Sei V ein Vektorraum über $\mathbb{K}$.*
- *Die Dimension von V ist die Zahl*

$$\dim V \; (= \dim_{\mathbb{K}} V) := \begin{cases} 0, & \textit{falls} \quad V = \{0\} \\ n, & \textit{falls} \quad V \quad \textit{eine endliche Basis} \\ & \qquad (v_1, \ldots, v_n) \quad \textit{hat} \\ \infty & \textit{sonst} \end{cases}$$

Bemerkung

Die leere Familie ist die Basis des Nullvektorraums $V = \{0\}$.

Beispiele/Übungsaufgaben

1) $\dim_{\mathbb{K}} \mathbb{K}^n = n$.
2) Jeder komplexe Vektorraum V kann als reeller Vektorraum aufgefaßt werden und $\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V$. Insbesondere gilt $\dim_{\mathbb{R}} \mathbb{C}^n = 2 \dim_{\mathbb{C}} \mathbb{C}^n = 2n$.
3) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$ (Hinweis: Das folgt aus der Überabzählbarkeit von $\mathbb{R}$).
4) $\dim_{\mathbb{K}} F(X, \mathbb{K}) = \mathrm{card}(X)$, wobei

$$\mathrm{card}(X) := \begin{cases} n, & \text{falls} \quad X \quad \text{aus n Elementen besteht} \\ \infty, & \text{falls} \quad X \quad \text{unendlich ist.} \end{cases}$$

5) $\dim_{\mathbb{K}} \mathbb{K}[x] = \infty$.
6) Sei $U \subset V$ ein Unterraum. Dann gilt $\dim U \leq \dim V$. Falls V endlichdimensional ist, so gilt $\dim U = \dim V$ genau dann, wenn $U = V$.

Folgerung

- *Sei V ein Vektorraum der Dimension $n \in \mathbb{N}$.*

(i) *Jede linear unabhängige Familie von Vektoren von V hat höchstens n Elemente.*

(ii) *Eine linear unabhängige Familie von Vektoren von V ist genau dann eine Basis, wenn sie n Elemente hat.*

(iii) *Jedes Erzeugendensystem von V hat mindestens n Elemente.*

(iv) *Ein Erzeugendensystem von V ist genau dann eine Basis, wenn es n Elemente hat.*

Beweis

(i-ii) folgt aus dem Austauschsatz.

(iii) folgt daraus, dass man aus jedem endlichen Erzeugendensystem eine Basis auswählen kann.

(iv) Ein Erzeugendensystem mit n Elementen ist wegen (iii) minimal und somit eine Basis. □

Gaußscher Algorithmus

- Sei $V = \mathbb{K}^n$ und $v_1, \dots, v_r \in V$.
- Es gibt einen einfachen Algorithmus um eine Basis des Unterraums $U = \mathrm{span}\{v_1, \dots, v_r\} \subset V$ zu bestimmen. Er beruht auf folgenden Beobachtungen:

(i) $\mathrm{span}\{v_1, \dots, v_r\} = \mathrm{span}\{v_{\varphi(1)}, \dots, v_{\varphi(r)}\}$ für jede Permutation $\varphi : \{1, \dots, r\} \to \{1, \dots, r\}$,

(ii) $\mathrm{span}\{\lambda v_1, v_2, \dots, v_r\} = \mathrm{span}\{v_1, \dots, v_r\}$ für alle $\lambda \in \mathbb{K}^* := \mathbb{K} \setminus \{0\}$ und

(iii) $\mathrm{span}\{v_1, \dots, v_{i-1}, v_i + \lambda v_1, v_{i+1}, \dots, v_r\} = \mathrm{span}\{v_1, \dots, v_r\}$.

Gaußscher Algorithmus

- In der Praxis schreibt man die Vektoren

$$v_i = (v_{i1}, v_{i2}, \dots, v_{in}) \in \mathbb{K}^n$$

als Zeilen einer Matrix

$$A := \begin{pmatrix} v_{11} & \dots & v_{1n} \\ v_{21} & \dots & v_{2n} \\ \vdots & & \vdots \\ v_{r1} & \dots & v_{rn} \end{pmatrix}$$

- Die Matrix A läßt sich unter Anwendung von (i-iii) in eine Matrix

$$A' = \begin{pmatrix} w_{11} & \dots & w_{1n} \\ w_{21} & \dots & w_{2n} \\ \vdots & & \vdots \\ w_{r1} & \dots & w_{rn} \end{pmatrix}$$

überführen, derart dass

Gaußscher Algorithmus

1) die ersten $k = \dim U$ Zeilen der Matrix A' eine Basis von $U = \mathrm{span}\{v_1, \dots, v_r\}$ bilden, $w_i = 0$ für alle $k < i \leq n$ und
2) die Matrix A' Zeilenstufenform hat, d.h.

$$\begin{array}{rcl} w_1 & = & (0, \dots, 0, w_{1j_1}, \dots, w_{1n}) \\ w_2 & = & (0, \dots, 0, w_{2j_2}, \dots, w_{2n}) \\ \vdots & & \vdots \\ w_k & = & (0, \dots, 0, w_{kj_k}, \dots, w_{kn}), \end{array}$$

wobei die $w_{ij_i} \neq 0$ $(i = 1, \dots, k)$ und $1 \leq j_1 < j_2 < \dots < j_k \leq n$.

Gaußscher Algorithmus: Beispiel

- Sei z.B. $\mathbb{K} = \mathbb{Q}, = \mathbb{R}$ oder $= \mathbb{C}$ und

$$\begin{aligned} v_1 &= (0,0,2,1,0) \\ v_2 &= (0,1,0,2,1) \\ v_3 &= (0,2,1,1,1) \\ v_4 &= (0,4,4,3,2) \end{aligned}$$

- Vertauschen der ersten beiden Zeilen ändert nach (i) nicht die lineare Hülle der Zeilenvektoren:

$$A = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \xrightarrow{(i)} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

Gaußscher Algorithmus

Weiter im Beispiel:

- Die erste Spalte der Matrix

$$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$

  ist Null.
- Die zweite Spalte beginnt mit $w_{1j_1} = w_{12} = 1 \neq 0$.
- Daher kann man durch Addition von geeigneten Vielfachen der ersten Zeile zu den anderen Zeilen erreichen, dass alle Einträge unterhalb von w_{12} Null werden:

Gaußscher Algorithmus

Weiter im Beispiel:

- $$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix} \overset{-2\times I+III}{\longrightarrow} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 4 & 4 & 3 & 2 \end{pmatrix}$$
- $$\overset{-4\times I+IV}{\longrightarrow} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix}$$
- Als Nächstes produzieren wir Nullen unterhalb von $w_{2j_2} = w_{23} = 2 \neq 0$ durch Addition von Vielfachen der zweiten Zeile:

Gaußscher Algorithmus

Weiter im Beispiel:

- $$\begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 4 & -5 & -2 \end{pmatrix} \overset{-\frac{1}{2}\times II+III}{\underset{-2\times II+IV}{\longrightarrow}} \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{7}{2} & -1 \\ 0 & 0 & 0 & -7 & -2 \end{pmatrix}$$
- Wir erhalten $w_{3j_3} = w_{34} = -\frac{7}{2} \neq 0$ und $-2 \times III + IV$ liefert:

$$A' = \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{7}{2} & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Gaußscher Algorithmus

Weiter im Beispiel:

- Also $k = 3$ und

$$\begin{aligned} w_1 &= (0, 1, 0, 2, 1) \\ w_2 &= (0, 0, 2, 1, 0) \\ w_3 &= (0, 0, 0, -\frac{7}{2}, -1) \end{aligned}$$

- Somit ist (w_1, w_2, w_3) eine Basis von $U = \mathrm{span}\{v_1, v_2, v_3, v_4\} \subset \mathbb{K}^5$.
- Durch Multiplikation der Zeilen der Matrix A' mit geeigneten Faktoren $\neq 0$, können wir immer erreichen, dass $w_{1j_1} = w_{2j_2} = \cdots = w_{kj_k} = 1$.
- In dem Beispiel:

$$A' = \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & -\frac{7}{2} & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & \frac{2}{7} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Lineare Gleichungssysteme

- Der selbe Algorithmus kann auch zum lösen inhomogener linearer Gleichungssysteme

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

  verwendet werden. Hierbei sind die $a_{ij}, b_j \in \mathbb{K}$ gegeben. Gesucht sind die n Unbekannten $x_1, x_2, \ldots, x_n \in \mathbb{K}$.
- Es genügt den Gaußschen Algorithmus auf die $m \times (n+1)$-Matrix $(A|b)$ anzuwenden, wobei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Lineare Gleichungssysteme

- Der Gaußsche Algorithmus liefert eine $m \times (n+1)$-Matrix in Zeilenstufenform:

$$(A'|b') = \begin{pmatrix} \underline{|a'_{1j_1}} & & & \cdots & a'_{1n} & b'_1 \\ & \underline{|a'_{2j_2}} & & \cdots & a'_{2n} & b'_2 \\ & & & \vdots & & \\ & & \underline{|a'_{kj_k}} & \cdots & a'_{kn} & b'_k \\ 0 & & & & & \underline{|b'_{k+1}} \\ & & & & & \end{pmatrix},$$

  wobei $1 \leq j_1 < j_2 < \cdots < j_k \leq n$ und $a'_{ij_i} \neq 0$ $(i = 1, \ldots, k \leq \min\{m, n\})$.
- Das zugehörige inhomogene lineare Gleichungssystem hat die selben Lösungen wie das ursprüngliche System.
- Das Gleichungssystem ist genau dann lösbar, wenn $b'_{k+1} = 0$.
- Die Lösung hängt dann von genau $n - k$ frei wählbaren Parametern x_i, $i \in \{1, 2, \ldots, n\} \setminus \{j_1, j_2, \ldots, j_k\}$, ab.

Lineare Gleichungssysteme: Beispiel

$$(A|b) = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & -6 & -12 & 6 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Für den letzten Schritt setzen wir voraus, dass $3 \neq 0$ in $\mathbb{K}$ gilt.
- Das zugehörige Gleichungssystem ist:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ x_2 + 2x_3 &= -1. \end{aligned}$$

- Daraus ergibt sich $x_2 = -2x_3 - 1$,
  $x_1 = -2x_2 - 3x_3 = -2(-2x_3 - 1) - 3x_3 = x_3 + 2$.
- Die allgemeine Lösung ist also $x_1 = \lambda + 2$, $x_2 = -2\lambda - 1$,
  $x_3 = \lambda \in \mathbb{K}$ beliebig.

Lineare Abbildungen

Definition

- *V, W seien Vektorräume über $\mathbb{K}$.*
- *Eine Abbildung $F : V \to W$ heißt linear (genauer: $\mathbb{K}$-linear), wenn*

$$\begin{aligned} F(v+w) &= F(v) + F(w) \quad \textit{und} \\ F(\lambda v) &= \lambda F(v) \end{aligned}$$

für alle $v, w \in V$, $\lambda \in \mathbb{K}$.

- *Eine bijektive lineare Abbildung $F : V \to W$ heißt auch Isomorphismus (von Vektorräumen).*
- *Die Vektorräume V und W heißen isomorph, wenn es einen Isomorphismus $F : V \to W$ gibt.*

- *Übungsaufgabe:* Sei $F : V \to W$ ein Isomorphismus. Dann ist $F^{-1} : W \to V$ linear.

Lineare Abbildungen

Bemerkungen/Übungsaufgaben:

- Für jede Menge X und jeden VR V bilden die Abbildungen $X \to V$ einen Vektorraum, der mit $Abb(X, V)$ bezeichnet wird. Insbesondere ist das kartesische Produkt $V^n = V \times \cdots \times V = Abb(\{1, 2, \ldots, n\}, V)$ ein VR.
- Seien V, W Vektorräume. Die Menge aller linearen Abbildungen $V \to W$ ist ein Unterraum von $Abb(V, W)$, der mit $L(V, W)$ bezeichnet wird.
- Sei $F : V \to W$ linear. Für jeden Unterraum $V' \subset V$ ist $F(V') \subset W$ ein Unterraum.
- Ebenso ist $F^{-1}(W') \subset V$ ein Unterraum für jeden Unterraum $W' \subset W$.
- Insbesondere sind das Bild, $F(V) \subset W$, und der Kern, $\ker F := F^{-1}(0) = \{v \in V | F(v) = 0\} \subset V$, einer linearen Abbildung F Unterräume.
- F ist genau dann injektiv, wenn $\ker F = \{0\}$.

Lineare Abbildungen

Satz

- *V, W seien Vektorräume und $(v_1, \ldots, v_n)$ eine Basis von V.*

(i) *Die Zuordnung $F \mapsto (F(v_i))_i := (F(v_1), F(v_2), \ldots, F(v_n))$ definiert einen Isomorphismus $\varphi : L(V, W) \to W^n$.*

(ii) *F ist genau dann surjektiv, wenn $(F(v_i))_i$ ein Erzeugendensystem von W ist.*

(iii) *F ist genau dann injektiv, wenn $(F(v_i))_i$ linear unabhängig ist.*

(iv) *F ist genau dann ein Isomorphismus, wenn $(F(v_i))_i$ eine Basis von W ist.*

Lineare Abbildungen

Beweis.

(i) $\varphi : F \to (F(v_i))_i$ ist injektiv, denn aus $F(v_1) = F(v_2) = \cdots = F(v_n) = 0$ folgt $F(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i F(v_i) = 0$ für jeden Vektor $v = \sum_{i=1}^n \lambda_i v_i \in V$ und somit $F = 0$.

- Um die Surjektivität von φ zu zeigen, müssen wir zu jedem $(w_1, w_2, \ldots, w_n) \in W^n$ die Existenz einer linearen Abbildung $F : V \to W$ mit $F(v_i) = w_i$ zeigen, $i = 1, 2, \ldots, n$.
- Sei $v \in V$ und $v = \sum_{i=1}^n \lambda_i v_i$ die eindeutige Darstellung als Linearkombination der linear unabhängigen Vektoren $(v_i)_{i=1,2,\ldots,n}$. Wir definieren $F(v) := \sum_{i=1}^n \lambda_i w_i$.
- Dann ist $F \in L(V, W)$ und offenbar $F(v_i) = w_i$.
- (ii-iv) sind einfache Übungsaufgaben. □

Folgerung

(i) *Für jeden $\mathbb{K}$-Vektorraum W, gibt es einen kanonischen Isomorphismus*

$$L(\mathbb{K}^n, W) \to W^n, \quad F \mapsto (F(e_i))_i := (F(e_1), \dots, F(e_n)).$$

(ii) *Sei V ein endlichdimensionaler VR. Ein VR W ist genau dann isomorph zu V, wenn $\dim W = \dim V$.*

Beweis.

(i) folgt unmittelbar aus dem vorhergehenden Satz.

(ii) Sei $(v_1, \dots, v_n)$ eine Basis von V.

'$\Longrightarrow$' Wenn $F : V \to W$ ein Iso. ist, so ist $(F(v_1), \dots, F(v_n))$ eine Basis von W und somit $\dim W = n = \dim V$.

'$\Longleftarrow$' Sei umgekehrt $\dim W = \dim V = n$. Dann existiert eine Basis $(w_1, \dots, w_n)$ von W mit n Elementen. Die lineare Abbildung $F : V \to W$ mit $F(v_i) = w_i$ ist ein Isomorphismus. □

Lineare Abbildungen

- Für $W = \mathbb{K}^m$ erhalten wir aus dem letzten Satz:

Folgerung

(i) *Es gibt kanonische Isomorphismen*

$$L(\mathbb{K}^n, \mathbb{K}^m) \cong (\mathbb{K}^m)^n \cong \mathbb{K}^{mn}.$$

(ii) *Auf diese Weise können wir lineare Abbildungen $\mathbb{K}^n \to \mathbb{K}^m$ kanonisch mit $(m \times n)$-Matrizen identifizieren. Die Menge $\mathrm{Mat}(m, n, \mathbb{K})$ der $(m \times n)$-Matrizen mit Einträgen in $\mathbb{K}$ wird somit zu einem $\mathbb{K}$-Vektorraum der Dimension mn.*

Lineare Abbildungen: Beweis der Folgerung

Beweis.

- Wir erläutern die Identifizierung $(\mathbb{K}^m)^n = \mathrm{Mat}(m, n, \mathbb{K})$:
- Die Einträge w_i eines n-Tupels $(w_1, \dots, w_n) \in (\mathbb{K}^m)^n$ bilden die n Spalten einer $(m \times n)$-Matrix

$$\begin{pmatrix} w_{11} & \dots & w_{1n} \\ w_{21} & \dots & w_{2n} \\ \vdots & & \vdots \\ w_{m1} & \dots & w_{mn} \end{pmatrix}.$$

$\square$

Produkt einer $(m \times n)$-Matrix mit einem n-Vektor

- Wie wir gesehen haben entspricht einer $(m \times n)$-Matrix

$$A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ a_1^2 & \dots & a_n^2 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} \in \mathrm{Mat}(m, n, \mathbb{K}) \cong L(\mathbb{K}^n, \mathbb{K}^m)$$

  eine lineare Abbildung $A : \mathbb{K}^n \to \mathbb{K}^m$.
- Die j-te Spalte

$$a_j = \begin{pmatrix} a_j^1 \\ \vdots \\ a_j^m \end{pmatrix} = \sum_{i=1}^{m} a_j^i e_i \in \mathbb{K}^m$$

  der Matrix A ist dabei genau das Bild $A(e_j)$ des j-ten Basisvektors e_j der kanonischen Basis von $\mathbb{K}^n$.

Produkt einer $(m \times n)$-Matrix mit einem n-Vektor

- Das Produkt Ax der $(m \times n)$-Matrix A mit einem n-Vektor

$$x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \sum_{i=1}^{n} x^i e_i \in \mathbb{K}^n$$

ist nach Definition das Bild $y = A(x)$ des Vektors x unter der linearen Abbildung $A : \mathbb{K}^n \to \mathbb{K}^m$.

- Demnach gilt

$$\begin{aligned} Ax &= A(\sum_{i=1}^{n} x^i e_i) = \sum_{i=1}^{n} x^i A(e_i) = \sum_{i=1}^{n} x^i a_i \\ &= \sum_{i=1}^{n} x^i \sum_{j=1}^{m} a_i^j e_j = \sum_{j=1}^{m} (\sum_{i=1}^{n} a_i^j x^i) e_j. \end{aligned}$$

- Also $y = Ax = \sum_{j=1}^{m} y^j e_j$ mit $y^j = \sum_{i=1}^{n} a_i^j x^i$.

Matrixprodukt

- Sei nun $A = (a^i_j)_{i,j}$ eine $(m \times n)$-Matrix und $B = (b^j_k)_{j,k}$ eine $(n \times p)$-Matrix.
- Das Produkt $C = AB \in \mathrm{Mat}(m, p, \mathbb{K})$ der Matrizen $A \in L(\mathbb{K}^n, \mathbb{K}^m)$ und $B \in L(\mathbb{K}^p, \mathbb{K}^n)$ ist nach Definition die Verkettung $A \circ B \in L(\mathbb{K}^p, \mathbb{K}^m) = \mathrm{Mat}(m, p, \mathbb{K})$.
- Demnach gilt für $x = \sum_{k=1}^p x^k e_k \in \mathbb{K}^p$

$$
\begin{aligned}
(AB)x &= A(Bx) = A(\sum_{j=1}^n \sum_{k=1}^p b^j_k x^k e_j) = \sum_{k=1}^p \sum_{j=1}^n b^j_k x^k A(e_j) \\
&= \sum_{k=1}^p \sum_{j=1}^n b^j_k x^k \sum_{i=1}^m a^i_j e_i = \sum_{i=1}^m \sum_{k=1}^p (\sum_{j=1}^n a^i_j b^j_k) x^k e_i
\end{aligned}
$$

- Also $AB = C = (c^i_k)_{i,k}$ mit $c^i_k = \sum_{j=1}^n a^i_j b^j_k$.

Darstellende Matrix einer linearen Abbildung

- $B = (v_1, \dots, v_n)$ bzw. $B' = (w_1, \dots, w_m)$ seien Basen der $\mathbb{K}$-Vektorräume V bzw. W.
- $\phi_B : \mathbb{K}^n \to V$ und $\phi_{B'} : \mathbb{K}^m \to W$ seien die zugehörigen Isomorphismen.

  (Zur Erinnerung: $\phi_B(e_i) = v_i$, $\phi_{B'}(e_j) = w_j$, wobei $i = 1, \dots, n$, $j = 1, \dots, m$.)

Definition

Sei $F \in L(V, W)$. Die Matrix

$$M_B^{B'}(F) := \phi_{B'}^{-1} \circ F \circ \phi_B \in L(\mathbb{K}^n, \mathbb{K}^m) = \mathrm{Mat}(m, n, \mathbb{K})$$

heißt darstellende Matrix der linearen Abbildung F bez. der Basen B, B'.

Darstellende Matrix einer linearen Abbildung

Satz

Mit den obigen Bezeichnungen ist die darstellende Matrix $A = (a^i_j)_{i,j} = M^{B'}_B(F)$ *durch folgende Gleichung bestimmt:*

$$(*) \quad F(v_i) = \sum_{j=1}^{m} a^j_i w_j, \quad i = 1, \dots, n.$$

Beweis.

Die $A = M^{B'}_B(F)$ definierende Gleichung

$$(\phi^{-1}_{B'} \circ F \circ \phi_B)(e_i) = A(e_i) = \sum_{j=1}^{m} a^j_i e_j$$

geht durch Anwenden von $\phi_{B'}$ in die äquivalente Gleichung (*) über. $\square$

Darstellende Matrix einer linearen Abbildung

Beispiel

- Seien $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ und B bzw. B' die zugehörigen *kanonischen* Basen. Dann gilt $\phi_B = \mathrm{Id}_V$, $\phi_{B'} = \mathrm{Id}_W$ und
- $M_B^{B'}(F) = \phi_{B'}^{-1} \circ F \circ \phi_B = F$ für jede linearen Abbildung $F \in L(\mathbb{K}^n, \mathbb{K}^m) = \mathrm{Mat}(m, n, \mathbb{K})$.

Darstellende Matrix einer linearen Abbildung

Satz

- *Seien V, V', V'' endlichdimensionale Vektorräume mit Basen B, B', B'' und $F \in L(V, V')$, $G \in L(V', V'')$.*
- *Dann gilt $G \circ F \in L(V, V'')$ und*

$$M_B^{B''}(G \circ F) = M_{B'}^{B''}(G) M_B^{B'}(F).$$

Beweis.

- Die Linearität von $G \circ F$ folgt leicht aus der von F und G.
- Die darstellende Matrix $M_B^{B''}(G \circ F)$ ergibt sich aus

$$\phi_{B''}^{-1} \circ (G \circ F) \circ \phi_B = (\phi_{B''}^{-1} \circ G \circ \phi_{B'}) \circ (\phi_{B'}^{-1} \circ F \circ \phi_B). \quad \square$$

Endomorphismen

Definition

- *Sei V ein Vektorraum. Lineare Abbildungen $V \to V$ heißen auch Endomorphismen von V. Wir setzen* $\mathrm{End}(V) := L(V, V)$.
- *Für den Vektorraum der $(n \times n)$-Matrizen schreiben wir auch* $\mathrm{Mat}(n, \mathbb{K}) := \mathrm{Mat}(n, n, \mathbb{K}) = \mathrm{End}(\mathbb{K}^n)$.
- *Für die darstellende Matrix eines Endomorphismus $F \in \mathrm{End}(V)$ bez. einer Basis B von V schreiben wir auch* $M_B(F) := M_B^B(F)$.

Endomorphismen

Beispiel

- Sei
$$F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{R}) = \mathrm{End}(\mathbb{R}^2).$$
- Die darstellende Matrix bez. der Basis $B = (b_1, b_2) = (e_1 + e_2, e_1 - e_2)$ ist
$$M_B(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
- In der Tat: $F(b_1) = e_2 + e_1 = b_1$ und $F(b_2) = e_2 - e_1 = -b_2$.

Rang einer linearen Abbildung

Definition

Der Rang einer linearen Abbildung $F : V \to W$ ist

$$\mathrm{rg}(F) := \dim F(V) \leq \dim W.$$

Beispiel

- Sei $A \in \mathrm{Mat}(m, n, \mathbb{K}) = L(\mathbb{K}^n, \mathbb{K}^m)$ eine Matrix.
- Der Rang von A ist die Dimension des von den n Spalten $A(e_1), \ldots, A(e_n) \in \mathbb{K}^m$ von A erzeugten Unterraumes von $\mathbb{K}^m$.
- Man spricht daher auch vom Spaltenrang der Matrix.
- Der Zeilenrang von A ist die Dimension des von den Zeilen der Matrix A aufgespannten Unterraums von $\mathbb{K}^n$.
- Wir werden später sehen, dass Zeilenrang=Spaltenrang.

Rang einer linearen Abbildung

Satz

- *Seien V, W endlichdimensionale Vektorräume und B, B' Basen von V bzw. W.*

(i) *Dann gilt für jede lineare Abbildung $F : V \to W$*

$$\mathrm{rg}(F) = \mathrm{rg}(M_B^{B'}(F)).$$

Beweis.

(i) Das Bild von $F \circ \phi_B$ stimmt mit dem von F überein und wird durch $\phi_{B'}^{-1}$ isomorph auf das Bild von $M_B^{B'}(F)$ abgebildet. Also $\mathrm{rg}(F) = \mathrm{rg}(F \circ \phi_B) = \mathrm{rg}(M_B^{B'}(F))$.

□

Dimensionsformel

Satz

- *Seien* V, W *Vektorräume und* $\dim V < \infty$.
- *Dann gilt für jede lineare Abbildung* $F : V \to W$

$$\mathrm{rg}(F) = \dim V - \dim \ker(F) \leq \dim V.$$

Beweis.

- F bildet jede Basis von V auf ein Erzeugendensystem von $F(V)$ ab. Demnach ist $\dim F(V) \leq \dim V$.
- Sei $(u_1, \ldots, u_k)$ eine Basis von $\ker(F)$, $(w_1, \ldots, w_r)$ eine Basis von $F(V)$ und $v_1, \ldots, v_r \in V$, so dass $F(v_i) = w_i$.
- *Behauptung:* $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ ist eine Basis von V.
- Aus der Behauptung folgt die Aussage des Satzes: $\dim V = k + r$.

Beweis der Behauptung.

1) Wir zeigen zuerst, dass $(u_1, \ldots, u_k, v_1, \ldots, v_r)$ linear unabhängig ist.

- Aus $0 = \sum_{i=1}^k \lambda_i u_i + \sum_{j=1}^r \mu_j v_j$ folgt

$$0 = F(\sum_{i=1}^k \lambda_i u_i + \sum_{j=1}^r \mu_j v_j) = \sum_{j=1}^r \mu_j w_j$$

  und somit $\mu_1 = \cdots = \mu_r = 0$, wegen der linearen Unabhängigeit der w_j.

- Also $0 = \sum_{i=1}^k \lambda_i u_i$, woraus $\lambda_1 = \cdots = \lambda_k = 0$ folgt, wegen der linearen Unabhängigkeit der u_i.

2) Als Nächstes zeigen wir, dass $\mathrm{span}\{u_1, \ldots, u_k, v_1, \ldots, v_r\} = V$.

- Sei $v \in V$. Wir schreiben $F(v) = \sum_{j=1}^r \mu_j w_j$.
- Dann gilt $v - \sum_{j=1}^r \mu_j v_j \in \ker(F) = \mathrm{span}\{u_1, \ldots, u_k\}$ und somit $v \in \mathrm{span}\{u_1, \ldots, u_k, v_1, \ldots, v_r\}$. $\square$

Folgerung aus der Dimensionsformel

Folgerung

- *Seien V, W Vektorräume,* $\dim V < \infty$ *und* $F \in L(V, W)$*. Dann gilt:*

(i) *F ist genau dann surjektiv, wenn* $\mathrm{rg}(F) = \dim W$,

(ii) *F ist genau dann injektiv, wenn* $\mathrm{rg}(F) = \dim V$,

(iii) *F ist genau dann bijektiv, wenn* $\mathrm{rg}(F) = \dim V = \dim W$.

Beweis.

(i) F surjektiv $\iff F(V) = W \iff \dim F(V) = \dim W$.

(ii) F injektiv $\iff \ker(F) = \{0\} \overset{\text{(Dim.formel)}}{\iff} \mathrm{rg}(F) = \dim V$.

(iii) folgt aus (i-ii).

$\square$

Folgerung

- *Seien V, W Vektorräume und* $\dim V = \dim W < \infty$.
- *Dann gilt: Eine lineare Abbildung $F : V \to W$ ist genau dann ein Isomorphismus, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:*

(i) *F ist injektiv,*

(ii) *F ist surjektiv,*

(iii) *F ist bijektiv.*

Direktes Produkt von Vektorräumen

Definition

- *U, V seien $\mathbb{K}$-Vektorräume.*
- *Das direkte Produkt von U und V ist das kartesische Produkt $U \times V = \{(u, v) | u \in U, v \in V\}$ ausgestattet mit der durch die komponentenweise definierten Addition und skalaren Multiplikation gegebenen Vektorraumstruktur.*

Übungsaufgabe

$$\dim(U \times V) = \dim U + \dim V.$$

Summe von Unterräumen

- W, W' seien Unterräume eines Vektorraums V.
- Die Summe

$$W + W' := \{w + w' | w \in W, w' \in W'\}$$

ist der kleinste Unterraum von V, der W und W' enthält.

Satz

Falls W und W' endlichdimensional sind, so gilt folgende Dimensionsformel

$$\dim(W + W') = \dim W + \dim W' - \dim(W \cap W').$$

Beweis der Dimensionsformel für die Summe

Beweis.

- Wir betrachten die Abbildung $F : W \times W' \to W + W'$,
$$(w, w') \mapsto F(w, w') := w - w'.$$
- Offenbar ist F surjektiv und
$$\ker(F) = \{(w, w) | w \in W \cap W'\} \cong W \cap W'.$$
- Daher folgt aus der Dimensionsformel für lineare Abbildungen
$$\begin{aligned} \mathrm{rg}(F) &= \dim(W + W') = \dim(W \times W') - \dim(W \cap W') \\ &= \dim W + \dim W' - \dim(W \cap W'). \end{aligned}$$

$\square$

Direkte Summe von Unterräumen

Definition

- *W, W' seien Unterräume eines Vektorraums V.*
- *Wir sagen, dass die Summe $W + W'$ direkt ist und schreiben*

$$W \oplus W',$$

  falls $W \cap W' = \{0\}$.
- *Die Unterräume W, W' heißen komplementär, wenn $V = W \oplus W'$.*
- *Wir sagen dann, dass W' ein Komplement zu W (in V) ist.*

Direkte Summe von Unterräumen

Beispiel

- U, V seien $\mathbb{K}$-Vektorräume.
- Die Unterräume

$$U \times \{0\} = \{(u, 0) \in U \times V | u \in U\} \subset U \times V,$$

$$\{0\} \times V = \{(0, v) \in U \times V | v \in V\} \subset U \times V$$

sind komplementär:

$$(*) \quad U \times V = (U \times \{0\}) \oplus (\{0\} \times V).$$

- Durch die injektiven linearen Abbildungen $u \mapsto (u, 0)$, $U \to U \times V$, und $v \mapsto (0, v)$, $V \to U \times V$, können wir U und V mit den Unterräumen $U \times \{0\}$ bzw. $\{0\} \times V$ von $U \times V$ identifizieren und (*) auch in der Form $U \times V = U \oplus V$ schreiben.

Direkte Summe von Unterräumen

Satz

- *W, W' seien Unterräume eines Vektorraums V.*
- $V = W \oplus W' \iff$ *Jeder Vektor $v \in V$ hat eine eindeutige Darstellung* $(*)$ $\quad v = w + w', \quad w \in W, \quad w' \in W'.$

Beweis.

'$\Longrightarrow$' Aus $V = W + W'$ folgt, dass jeder Vektor $v \in V$ eine Darst. (*) hat. Sei $v = w_1 + w_1'$ eine weitere solche Darst.

- Es folgt $0 = v - v = (w - w_1) + (w' - w_1')$ und daraus $W \ni w - w_1 = -(w' - w_1') \in W'$.
- Wg. $W \cap W' = \{0\}$ folgt nun $w - w_1 = w' - w_1' = 0$ und somit die Eindeutigkeit der Darstellung (*).

'$\Longleftarrow$' Wenn umgekehrt jeder Vektor $v \in V$ eine eindeutige Darstellung (*) hat, so gilt $V = W + W'$ und aus $0 = w + (-w)$, $w \in W \cap W'$, folgt wg. der Eindeutigkeit von (*) $w = 0$, d.h. $W \cap W' = \{0\}$. □

Direkte Summe von Unterräumen

Satz

- *Sei V ein endlichdimensionaler Vektorraum. Dann gilt:*

a) *Jeder Unterraum $W \subset V$ besitzt ein Komplement.*

b) *$W, W' \subset V$ seien Unterräume. Dann sind folgende Aussagen äquivalent:*
   (i) $V = W \oplus W'$,
   (ii) $W' \cap W = \{0\}$ *und* $\dim W + \dim W' = \dim V$,
   (iii) $V = W + W'$ *und* $\dim W + \dim W' = \dim V$.

Beweis.

a) Sei $(w_1, \ldots, w_k)$ eine Basis von W. Nach dem Austauschsatz von Steinitz können wir diese zu einer Basis $(w_1, \ldots, w_k, w_1', \ldots, w_l')$ von V ergänzen.

- $W' := \mathrm{span}\{w_1', \ldots, w_l'\}$ ist dann ein Komplement zu W.

b) Die Äquivalenz der Eigenschaften (i-iii) folgt aus der Dimensionsformel. $\square$

Der Lösungsraum eines linearen Gleichungssystems

- Ein inhomogenes lineares Gleichungssystem über $\mathbb{K}$ läßt sich in der Form

$$(*) \quad Ax = b,$$

  schreiben, wobei $A \in L(\mathbb{K}^n, \mathbb{K}^m)$ und $b \in \mathbb{K}^m$ gegeben sind und $x \in \mathbb{K}^n$ gesucht ist.
- Hierbei ist m die Anzahl der Gleichungen und n die Anzahl der Unbekannten, die Komponenten des Vektors x.
- Das System (*) heißt homogen, falls $b = 0$.

Der Lösungsraum eines linearen Gleichungssystems

Satz

(i) *Die Lösungsmenge eines homogenen linearen Gleichungssytems* $Ax = 0$, $A \in L(\mathbb{K}^n, \mathbb{K}^m)$, *ist genau der Unterraum* $U = \ker(A) \subset \mathbb{K}^n$.

(ii) *Die Dimension von* U *beträgt* $n - r$, *wobei* $r = \mathrm{rg}(A)$.

(iii) *Der Spaltenrang* r *stimmt mit dem Zeilenrang von* A *überein.*

Beweis.

- (i) ist klar und (ii) folgt aus der Dimensionsformel.

(iii) Wie wir gesehen haben, läßt sich die Matrix A durch den Gaußalgorithmus auf Zeilenstufenform A' bringen und somit das lineare Gleichungssytem $Ax = 0$ lösen.

- Dabei ergab sich $\dim U = n - k$, wobei k die Anzahl der nichtrivialen Zeilen der Matrix A' ist, also genau der Zeilenrang von A. Wg. (ii) gilt nun $k = r$, d.h. Zeilenrang=Spaltenrang. □

Der Lösungsraum eines linearen Gleichungssystems

Satz

- *Wir betrachten ein inhomogenes lineares Gleichungssytem*
  $(*) \quad Ax = b,\ A \in L(\mathbb{K}^n, \mathbb{K}^m).$

(i) *Falls $b \notin A(\mathbb{K}^n)$, so hat (*) keine Lösung.*

(ii) *Falls $b \in A(\mathbb{K}^n)$, so ist die Lösungsmenge U' von (*) von der Form*

$$U' = x_0 + U = \{x_0 + u | u \in U\},$$

wobei $x_0 \in A^{-1}(b)$ eine spezielle Lösung von (*) ist und $U \subset \mathbb{K}^n$ der Lösungsraum des zugehörigen homogenen Systems $Ax = 0$ ist.

Beweis.

- (i) ist klar. Im Fall (ii) ist klar, dass $x_0 + U \subset U'$. Wir zeigen $U' \subset x_0 + U$. Sei also v eine Lösung von (*). Dann gilt $A(v - x_0) = b - b = 0$ und somit $v = x_0 + (v - x_0) \in x_0 + U$.

Affine Unterräume

Definition

- *Ein affiner Unterraum eines Vektorraums V ist eine Teilmenge der Form $v + U$, wobei $U \subset V$ ein Untervektorraum ist und $v \in V$.*

Bemerkung

- Demnach ist der Lösungsraum eines inhomogenen linearen Gleichungssystems $Ax = b$, $A \in L(\mathbb{K}^n, \mathbb{K}^m)$, ein affiner Unterraum von $\mathbb{K}^n$.
- Der Lösunsraum ist genau dann ein Untervektorraum von $\mathbb{K}^n$, wenn das Gleichungssytem homogen ist, d.h. wenn $b = 0$.

Der Begriff der Gruppe

Definition

- *Eine Gruppe $(G, \cdot)$ ist eine Menge G zusammen mit einer assoziativen Verknüpfung $\cdot : G \times G \to G$, so dass*

(i) *es ein neutrales Element $e \in G$ gibt mit*

$$e \cdot a = a \cdot e = a \quad \text{für alle} \quad a \in G \quad \text{und}$$

(ii) *zu jedem $a \in G$ ein Inverses a^{-1}, so dass $a \cdot a^{-1} = a^{-1} \cdot a = e$.*

Bemerkung

- Statt (i) und (ii), genügt es die folgenden Bedingungen zu fordern, aus denen bereits die Eindeutigkeit von e und a^{-1} folgt (vgl. Fischer):

(i') Es gibt $e \in G$ mit $e \cdot a = a$ für alle $a \in G$ und

(ii') zu jedem $a \in G$ ein a^{-1}, so dass $a^{-1} \cdot a = e$.

Beispiele von Gruppen I

(i) Die Bijektionen $\varphi : X \to X$ einer Menge X in sich bilden eine Gruppe, die mit $\mathrm{Bij}(X)$ bezeichnet wird. Die Vernüpfung ist die Verkettung, das neutrale Element ist Id_X und das Inverse einer Bijektion ist ihre Umkehrabbildung.

(ii) Im Fall einer endlichen Menge X nennt man die Bijektionen $\sigma \in \mathrm{Bij}(X)$ auch Permutationen von X. Die Permutationsgruppe $\mathrm{Bij}(X)$ ist dann eine endliche Gruppe mit $n!$ Elementen, wobei $n = \mathrm{card}(X)$.

(iii) Die Permutationsgruppe $S_n := \mathrm{Bij}(X_n)$ der n-elementigen Menge $X_n = \{1, 2, \ldots, n\}$ heißt die symmetrische Gruppe.

(iv) Sei V ein VR. Die Isomorphismen $V \to V$ heißen auch Automorphismen des Vektorraums V und bilden die sogenannte Automorphismengruppe $\mathrm{Aut}(V)$ von V.

Beispiele von Gruppen II

(v) Im Fall eines endlichdimensionalen Vektorraums V spricht man auch von der allgemeinen linearen Gruppe und schreibt dafür $\mathrm{GL}(V) := \mathrm{Aut}(V)$.

- Man setzt $\mathrm{GL}(n, \mathbb{K}) := \mathrm{GL}(\mathbb{K}^n)$. Die Elemente von $\mathrm{GL}(n, \mathbb{K})$ heißen invertierbare Matrizen.
- Es gilt

$$\begin{aligned} \mathrm{GL}(n, \mathbb{K}) &= \{A \in \mathrm{Mat}(n, \mathbb{K}) | \mathrm{rg}(A) = n\} \\ &= \{A \in \mathrm{Mat}(n, \mathbb{K}) | \ker(A) = 0\}, \end{aligned}$$

  wobei wir der Übersichtlichkeit halber ab jetzt den Nullvektorraum $\{0\}$ einfach mit 0 bezeichnen.

(vi) $\mathrm{GL}(1, \mathbb{K}) = (\mathbb{K}^*, \cdot)$ ist eine kommutative Gruppe.

(vii) Wie bereits bemerkt ist für jeden Vektorraum V durch $(V, +)$ eine kommutative Gruppe gegeben.

Beispiele von Gruppen III

(viii) Das kartesische Produkt $G_1 \times G_2$ von Gruppen G_1, G_2 ist mit der komponentenweisen Verknüpfung wieder eine Gruppe. $G_1 \times G_2$ ist genau dann kommutativ, wenn G_1 und G_2 kommutativ sind.

(ix) Die Kreislinie $S^1 := \{z \in \mathbb{C} | |z| = 1\}$ ist mit der Multiplikation komplexer Zahlen eine kommutative Gruppe.

Gruppenmorphismen

Definition

- *Eine Abbildung $\varphi : G \to H$ zwischen Gruppen G, H heißt ein Gruppenmorphismus, falls*

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$$

  für alle $a, b \in G$.
- *Ein bijektiver Gruppenmorphismus heißt auch Isomorphismus von Gruppen. Ein Isomorphismus einer Gruppe in sich heißt auch Automorphismus.*
- *Zwei Gruppen G und H heißen isomorph, falls es einen Isomorphismus von Gruppen $\varphi : G \to H$ gibt.*
- *Eine Untergruppe einer Gruppe G ist eine Teilmenge $H \subset G$, die mit $a, b \in H$ auch $a \cdot b$ und a^{-1} enthält.*

Gruppenmorphismen

Übungsaufgaben/Beispiele I

- Jeder Gruppenmorphismus $\varphi : G \to H$ bildet das neutrale Element in G auf das neutrale Element in H ab und erfüllt

$$\varphi(a^{-1}) = \varphi(a)^{-1}$$

  für alle $a \in G$.
- Das Inverse eines Isomorphismus $\varphi : G \to H$ ist wieder ein Isomorphismus.
- Jede Untergruppe $H \subset G$ einer Gruppe G ist mit der induzierten Verknüpfung wieder eine Gruppe.
- Das Bild $\varphi(K) \subset H$ einer Untergruppe $K \subset G$ unter einem Gruppenmorphismus $\varphi : G \to H$ ist wieder eine Untergruppe.
- Das Urbild $\varphi^{-1}(K) \subset G$ einer Untergruppe $K \subset H$ unter einem Gruppenmorphismus $\varphi : G \to H$ ist wieder eine Untergruppe.

Gruppenmorphismen

Übungsaufgaben/Beispiele II

- Die Automorphismen einer Gruppe G bilden eine Untergruppe $\mathrm{Aut}(G) \subset \mathrm{Bij}(G)$.
- Die Gruppen $(\mathbb{R}, +)$ und $(\mathbb{R}_+, \cdot)$ sind isomorph. Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}_+$ ist ein Isomorphismus, ebenso wie ihre Umkehrfunktion $\ln : \mathbb{R}_+ \to \mathbb{R}$.
- Die Abbildung $t \mapsto \exp(it)$ definiert einen surjektiven aber nicht injektiven Gruppenmorphismus $\varphi : \mathbb{R} \to S^1$.

Die symmetrische Gruppe

- Jede Permutation $\sigma \in S_n$ wird durch ein Diagramm
$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$
beschrieben.
- Eine Transposition $\tau = \tau_{ij}$ ist eine Permutation, die zwei Zahlen $i < j$ vertauscht und alle anderen Zahlen unverändert läßt.
- Beispielsweise enthält S_3 genau 3 Transpositionen:
$$\begin{aligned} \tau_{12} &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \\ \tau_{13} &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \\ \tau_{23} &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \end{aligned}$$

Die symmetrische Gruppe

Satz

Die symmetrische Gruppe S_n, $n \in \mathbb{N}$, ist nur für $n \leq 2$ kommutativ.

Beweis.

- Es ist klar, dass S_1 und S_2 kommutativ sind. S_3 ist nicht kommutativ, denn z.B. ist $\tau_{12} \circ \tau_{13} \neq \tau_{13} \circ \tau_{12}$. Letzteres folgt bereits aus:

$$\tau_{12} \circ \tau_{13}(1) = \tau_{12}(3) = 3 \neq \tau_{13} \circ \tau_{12}(1) = \tau_{13}(2) = 2$$

- Die Abbildung

$$\sigma \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & 4 & \cdots & n \end{pmatrix}$$

  ist ein injektiver Gruppenmorphismus $S_3 \to S_n$ $(n \geq 3)$. Das Bild ist eine nicht kommutative Untergruppe von S_n. Insbesondere ist S_n für $n \geq 3$ nicht kommutativ. $\square$

Die symmetrische Gruppe

Satz

Für alle $\sigma \in S_n$ $(n \geq 2)$ gibt es Transpositionen $\tau_1, \ldots \tau_k \in S_n$, so dass $\sigma = \tau_1 \circ \cdots \circ \tau_k$.

Beweis.

- Für jede Transposition τ gilt $\tau^{-1} = \tau$ und somit $\mathrm{Id} = \tau \circ \tau$.
- Sei $\mathrm{Id} \neq \sigma \in S_n$. Dann existiert ein $1 \leq i_1 \leq n$, so dass $\sigma(i) = i$ für alle $1 \leq i \leq i_1 - 1$ und $\sigma(i_1) > i_1$. Wir setzen $\tau_1 := \tau_{i_1 \sigma(i_1)}$ und $\sigma_1 := \tau_1 \circ \sigma$. Dann gilt $\sigma_1(i) = i$ für alle $1 \leq i \leq i_1$.
- Falls $\sigma_1 \neq \mathrm{Id}$, so gibt es $i_1 < i_2 \leq n$, so dass $\sigma_1(i) = i$ für alle $1 \leq i \leq i_2 - 1$ und $\sigma_1(i_2) > i_2$ und wir setzen das Iterationsverfahren fort.
- Nach endlich vielen (genauer: nach $k \leq n - 1$) Schritten erhalten wir $\sigma_k = \tau_k \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{Id}$ und somit $\sigma = \tau_1 \circ \cdots \circ \tau_k$. $\square$

Das Vorzeichen einer Permutation

Definition

- *Das Vorzeichen* $\epsilon(\sigma)$ *einer Permutation* $\sigma \in S_n$ *ist definiert als*

$$(*) \quad \epsilon(\sigma) := \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \in \{-1, +1\}.$$

  (Man beachte, dass im Zähler und Nenner bis auf das Vorzeichen das selbe Produkt steht.)

- *Die Permutation heißt gerade, wenn* $\epsilon(\sigma) = +1$ *und ungerade, wenn* $\epsilon(\sigma) = -1$.

Bemerkung

- Die Anzahl der negativen Faktoren im Produkt (*) ist genau die Anzahl der Paare $i < j$ mit $\sigma(i) > \sigma(j)$.
- Die Permutation σ ist also genau dann gerade (bzw. ungerade), wenn diese Anzahl gerade (bzw. ungerade) ist.

Das Vorzeichen einer Permutation

Satz

$\epsilon : S_n \to \{-1, 1\}$ *ist ein Gruppenmorphismus in die Untergruppe* $\{-1, 1\} \subset \mathbb{R}^*$*.*

Beweis.

- Wir berechnen für $\sigma, \tau \in S_n$

$$\begin{aligned}
\epsilon(\sigma \circ \tau) &= \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\
&= \prod_{i<j} \underbrace{\frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)}}_{f_{ij}} \cdot \prod_{i<j} \frac{\tau(j) - \tau(i)}{j - i} = \epsilon(\sigma)\epsilon(\tau),
\end{aligned}$$

- denn wegen (*) $f_{ij} = f_{ji}$ gilt mit $\tau i := \tau(i)$:
$\prod_{i<j} f_{ij} = \prod_{\substack{i<j \\ \tau i<\tau j}} f_{ij} \prod_{\substack{i<j \\ \tau i>\tau j}} f_{ij} \overset{(*)}{=} \prod_{\substack{i<j \\ \tau i<\tau j}} f_{ij} \prod_{\substack{i>j \\ \tau i<\tau j}} f_{ij} =$
$\prod_{\tau i<\tau j} f_{ij} = \epsilon(\sigma)$. □

Das Vorzeichen einer Permutation

Folgerung

- *Sei* $\sigma \in S_n$ *und* $\sigma = \tau_1 \circ \cdots \circ \tau_k$ *eine Darstellung als Produkt von Transpositionen.*
- *Dann gilt*
$$\epsilon(\sigma) = (-1)^k.$$

Beweis.

- Für die Transposition $\tau = \tau_{12}$ gilt $\epsilon(\tau) = -1$, denn $(1, 2)$ ist das einzige Paar (i, j) mit $i < j$ und $\tau(i) > \tau(j)$.
- Sei $\sigma \in S_n$ mit $\sigma(1) = i$ und $\sigma(2) = j$. Dann gilt $\sigma \circ \tau_{12} \circ \sigma^{-1} = \tau_{ij}$ und somit
- $\epsilon(\tau_{ij}) = \epsilon(\sigma)\epsilon(\tau_{12})\epsilon(\sigma^{-1}) = \epsilon(\tau_{12}) = -1.$
- Also $\epsilon(\sigma) = \epsilon(\tau_1 \circ \cdots \circ \tau_k) = \epsilon(\tau_1) \cdots \epsilon(\tau_k) = (-1)^k.$ $\square$

Das Vorzeichen einer Permutation

Beispiel

- Die Permutation

$$\zeta := (23 \cdots n1) := \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$$

- schreibt sich als $\zeta = \tau_{12} \circ \tau_{23} \circ \cdots \circ \tau_{n-1n}$.
- Also ist $\epsilon(\zeta) = (-1)^{n-1}$.
- ζ erzeugt eine n-elementige Untergruppe

$$\langle \zeta \rangle := \{\zeta, \zeta^2, \ldots, \zeta^n = \mathrm{Id}\} \subset S_n.$$

- Die ζ^k, $k = 1, 2, \ldots, n$, heißen zyklische Permutationen.
- Für ungerades n sind also alle zyklischen Permutationen gerade. (ÜA: Für $n = 3$ sind alle ungeraden Permutationen Transpositionen und alle geraden Permutationen zyklisch.)

Determinante

Definition

- *Eine Abbildung* $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *heißt Determinante, wenn sie folgende Eigenschaften hat:*

D1) det *ist linear in jeder Spalte, d.h.*

$$\begin{aligned} & \det(a_1 \cdots a_{i-1}\, \lambda a_i + \mu b_i\, a_{i+1} \cdots a_n) \\ = \quad & \lambda \det(a_1 \cdots\ a_i\ \cdots a_n) + \mu \det(a_1 \cdots a_{i-1}\, b_i\, a_{i+1} \cdots a_n) \end{aligned}$$

für alle Spaltenvektoren $a_1, \ldots, a_n, b_i \in \mathbb{K}^n$ *und* $\lambda, \mu \in \mathbb{K}$.

D2) det *ist alternierend, d.h.* $\det A = 0$, *falls zwei Spalten von* A *übereinstimmen.*

D3) det *ist folgendermaßen normiert:* $\det(\mathbf{1}_n) = 1$, *wobei* $\mathbf{1}_n := (e_1 \cdots e_n)$ *die Einheitsmatrix ist.*

Determinante

- Wir werden zeigen, dass es genau eine Abbildung $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ mit den Eigenschaften D1-D3 gibt.
- Zunächst zeigen wir, dass D1-D3 eine Reihe weiterer Rechenregeln nach sich ziehen.

Determinante

Satz

Sei $A = (a_1 \cdots a_n) \in \mathrm{Mat}(n, \mathbb{K})$*. Aus D1-D2 folgt für jede Permutation* $\sigma \in S_n$*:*

$$(*) \quad \det(a_{\sigma(1)} \cdots a_{\sigma(n)}) = \epsilon(\sigma) \det A.$$

Beweis.

- Aus D1-D2 erhalten wir für $i < j$
  $0 = \det(\cdots a_i + a_j \cdots a_i + a_j \cdots) =$
  $\det(\cdots a_i \cdots a_j \cdots) + \det(\cdots a_j \cdots a_i \cdots)$.
- Hierbei ist die i-te und j-te Spalte angegeben. Die Auslassungspunkte stehen für $a_1 \cdots a_{i-1}$, $a_{i+1} \cdots a_{j-1}$ und $a_{j+1} \cdots a_n$.
- Das beweist (*) für jede Transposition $\sigma = \tau_{ij}$.
- Der allgemeine Fall folgt nun daraus, dass sich jede Perm. als Produkt von Transpositionen schreiben läßt. $\square$

Determinante

Satz

- *Sei* $A = (a_1 \cdots a_n) \in \mathrm{Mat}(n, \mathbb{K})$ *und* $\lambda \in \mathbb{K}$*. Aus D1-D2 folgt, dass Addition des* λ*-fachen der j-ten Spalte von A zur i-ten Spalte von A* ($i \neq j$) *den Wert der Determinante nicht ändert:*

$$\det(a_1 \cdots a_{i-1}\; a_i + \lambda a_j\; a_{i+1} \cdots a_n) = \det(A).$$

Beweis.

Aus der Linearität in der i-ten Spalte folgt:

$$\begin{aligned} & \det(a_1 \cdots a_{i-1}\; a_i + \lambda a_j\; a_{i+1} \cdots a_n) \\ = \quad & \det A + \lambda \det(\cdots a_j \cdots a_j \cdots) \overset{D2}{=} \det A. \end{aligned}$$

□

Determinante

Folgerung

Sei $A = (a_1 \cdots a_n) \in \mathrm{Mat}(n, \mathbb{K})$*. Aus D1-D2 folgt: Falls* $\mathrm{rg}(A) \neq n$*, so ist* $\det(A) = 0$*.*

Beweis.

- $\mathrm{rg}(A) \neq n$ bedeutet, dass sich eine der Spalten a_i als Linearkombination der anderen schreiben läßt:

$$a_i = \sum_{j \neq i} \lambda_j a_j.$$

- Daraus ergibt sich

$$\det A = \det(a_1 \cdots a_{i-1} \underbrace{a_i - \sum_{j \neq i} \lambda_j a_j}_{=0} a_{i+1} \cdots a_n) = 0.$$

Berechnung der Determinante mittels Spaltenumformungen

Satz

(i) *Jede invertierbare Matrix* $A \in \mathrm{GL}(n, \mathbb{K})$ *läßt sich durch wiederholtes Anwenden folgender zwei Spaltenumformungen in eine Diagonalmatrix* $A' = \mathrm{diag}(\lambda_1, \dots, \lambda_n) := (\lambda_1 e_1 \cdots \lambda_n e_n)$ *überführen:*

S1) *Vertauschen von zwei Spalten bei gleichzeitiger Multiplikation einer der beiden mit -1 und*

S2) *Addition eines Vielfachen einer Spalte zu einer anderen Spalte.*

(ii) *Es gilt* $\det A = \det A' = \lambda_1 \cdots \lambda_n \neq 0$.

Beweis.

(ii) Die Spaltenumformungen S1-S2 ändern nicht den Wert der Determinante.

- Somit erhalten wir aus (i):
  $\det A = \det A' \overset{D1}{=} \lambda_1 \cdots \lambda_n \det(\mathbf{1}_n) \overset{D3}{=} \lambda_1 \cdots \lambda_n$.

Berechnung der Determinante mittels Spaltenumformungen

Weiter im Beweis:

(i) Durch Anwendung des Gaußalgorithmus auf die Spalten (statt auf die Zeilen) von A können wir mittels S1-S2 die Matrix A zunächst in untere Dreiecksgestalt:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}$$

bringen.

- Diese Matrix hat Rang n, da die Umformumgen S1-S2 den Rang nicht ändern.
- Also ist das Produkt $\lambda_1 \cdots \lambda_n \neq 0$.
- Durch Umformungen S2 können wir deswegen alle Matrixeinträge unterhalb der Diagonalen eliminieren und erhalten die gewünsche Diagonalgestalt $A' = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. $\square$

Determinante

- Der Gaußalgorithmus liefert auch:

Satz/ÜA

- Für Blockdreiecksmatrizen gilt

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det A \det D.$$

- Hierbei ist A eine $(n \times n)$-Matrix, D eine $(m \times m)$-Matrix, B eine $(n \times m)$-Matrix und C eine $(m \times n)$-Matrix.

Determinante

- Wir fassen einige der bisherigen Ergebnisse zusammen:

Folgerung

- *Es gibt höchstens eine Abbildung* $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *mit den Eigenschaften D1-D3.*
- *Unter der Voraussetzung, dass eine solche Abbildung existiert, gilt:*
- $A \in \mathrm{Mat}(n, \mathbb{K})$ *ist genau dann invertierbar, wenn* $\det A \neq 0$.

Determinantenmultiplikationssatz

Satz

Aus D1-D3 folgt für alle $A, B \in \mathrm{Mat}(n, \mathbb{K})$*:*

$$\det(AB) = \det(A)\det(B).$$

Beweis.

- Falls $\mathrm{rg}(A) < n$, so ist $\mathrm{rg}(AB) < n$ und somit $\det(AB) = 0 = \det(A)\det(B) = 0 \cdot \det(B)$.
- Falls $\mathrm{rg}(B) < n$, so ist $\ker(B) \neq 0$ und somit $\ker(AB) \neq 0$. Also ebenfalls $\det(AB) = 0 = \det(A)\det(B) = \det(A) \cdot 0$.
- Seien also A, B invertierbar. Wir benutzen folgenden Hilfssatz:

Lemma

- *Jede invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{K})$ schreibt sich als Produkt*
$$A = DE_1 \cdots E_k,$$
- *wobei $D \in \mathrm{GL}(n, \mathbb{K})$ eine Diagonalmatrix ist und die $E_1, \ldots, E_k$ von folgendem Typ sind:*
- $$\begin{array}{rrcl} (1) & E(i,j) & := & (e_{\tau(1)} \cdots e_{\tau(n)}), \\ (2) & E(i,j,\lambda) & := & (e_1 \cdots e_{i-1}\; e_i + \lambda e_j\; e_{i+1} \cdots e_n). \end{array}$$

  Hierbei ist $\lambda \in \mathbb{K}$, $i \neq j$ und $\tau \in S_n$ die Vertauschung von i und j.

Beweis des Lemmas.

- Wir bemerken zunächst, dass $A \mapsto AE(i,j)$ die i-te und j-te Spalte von A vertauscht und
- $A \mapsto AE(i,j,\lambda)$ das λ-fache der j-ten Spalte von A zur i-ten Spalte von A addiert .
- Desweiteren gilt

$$\begin{aligned} E(i,j)^{-1} &= E(i,j), \\ E(i,j,\lambda)^{-1} &= E(i,j,-\lambda). \end{aligned}$$

- Wegen des Gaußalgorithmus gibt es also Matrizen $F_1, \dots, F_k$ vom Typ (1-2), so dass $D = AF_1 \cdots F_k$ eine Diagonalmatrix ist.
- Daraus folgt mit $E_1 := F_k^{-1}, \dots, E_k := F_1^{-1}$:

$$A = DF_k^{-1} \cdots F_1^{-1} = DE_1 \cdots E_k.$$

□

Beweis des Determinantenmultiplikationssatzes

Beweis.

- Wegen des vorherigen Lemmas genügt es $\det(AB) = \det(A)\det(B)$ für den Fall zu beweisen, dass B eine Diagonalmatrix oder vom Typ (1-2) ist.
- Für eine Diagonalmatrix $B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ gilt $\det B = \lambda_1 \cdots \lambda_n$ und somit

$$\det(AB) = \det(\lambda_1 a_1 \cdots \lambda_n a_n) \overset{D1}{=} \lambda_1 \cdots \lambda_n \det A = \det A \det B.$$

- Für $B = E(i,j)$ gilt $\det B = \det(e_{\tau(1)} \cdots e_{\tau(n)}) = \epsilon(\tau)\det(\mathbf{1}_n) = -1$ und somit

$$\begin{aligned} \det(AB) &= \det(a_{\tau(1)} \cdots a_{\tau(n)}) = \epsilon(\tau)\det A = -\det A \\ &= \det A \det B. \end{aligned}$$

- Für $B = E(i,j,\lambda)$ gilt $\det B = 1$ und somit $\det(AB) = \det(a_1 \cdots a_{i-1}\, a_i + \lambda a_j\, a_{i+1} \cdots a_n) = \det A = \det A \det B$. □

Determinantenmultiplikationssatz

Folgerung

(i)

$$\det : \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K}^*$$

ist ein Gruppenmorphismus.

(ii) *Insbesondere gilt für alle $A \in \mathrm{GL}(n, \mathbb{K})$:*

$$\det(A^{-1}) = (\det A)^{-1}$$

(iii)

$$\mathrm{SL}(n, \mathbb{K}) := \{A \in \mathrm{GL}(n, \mathbb{K}) | \det A = 1\} \subset \mathrm{GL}(n, \mathbb{K}).$$

ist eine Untergruppe (die sogenannte spezielle lineare Gruppe).

Beweis.

- (i-ii) sind klar. (iii) folgt aus $\mathrm{SL}(n, \mathbb{K}) = \det^{-1}(1)$. □

Explizite Formel für die Determinante

Satz

- *Es gibt genau eine Abbildung* $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *mit den Eigenschaften D1-D3.*
- *Die Determinante einer* $(n \times n)$*-Matrix* $A = (a^i_j)_{i,j}$ *ist durch folgende Formel gegeben:*

$$\det A = Det\ A := \sum_{\sigma \in S_n} \epsilon(\sigma) a_1^{\sigma(1)} \cdots a_n^{\sigma(n)}.$$

Beweis

- Da wir bereits wissen, dass es *höchstens* eine Abbildung mit den Eigenschaften D1-D3 gibt, genügt es D1-D3 für die Abbildung *Det* nachzuweisen.

Beweis der expliziten Formel für die Determinante:

D1:

$$\begin{aligned} & \sum_{\sigma \in S_n} \epsilon(\sigma) a_1^{\sigma(1)} \cdots (\lambda a_i^{\sigma(i)} + \mu b_i^{\sigma(i)}) \cdots a_n^{\sigma(n)} \\ = \quad & \lambda Det\, A + \mu \sum_{\sigma \in S_n} \epsilon(\sigma) a_1^{\sigma(1)} \cdots b_i^{\sigma(i)} \cdots a_n^{\sigma(n)}. \end{aligned}$$

Beweis der expliziten Formel für die Determinante:

D2: Es genügt zu zeigen, dass

$$Det\,\underbrace{\left(a_{\tau(1)}\cdots a_{\tau(n)}\right)}_{A_\tau :=} = -DetA$$

für jede Transposition τ. Aus $A = A_\tau$ folgt dann $Det\,A = Det\,A_\tau = -Det\,A$ und somit $Det\,A = 0$.

- Mit der Substitution $\sigma' := \sigma\tau = \sigma \circ \tau$ haben wir $\sigma = \sigma'\tau$, $\epsilon(\sigma') = -\epsilon(\sigma)$ und somit
-
$$\begin{aligned} Det(A_\tau) &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\tau(1)}^{\sigma(1)} \cdots a_{\tau(n)}^{\sigma(n)} \\ &= -\sum_{\sigma' \in S_n} \epsilon(\sigma') a_{\tau(1)}^{\sigma'\tau(1)} \cdots a_{\tau(n)}^{\sigma'\tau(n)} \\ &= -\sum_{\sigma' \in S_n} \epsilon(\sigma') a_1^{\sigma'(1)} \cdots a_n^{\sigma'(n)} = -Det\,A \end{aligned}$$

Beweis der expliziten Formel für die Determinante:

D3: Sei nun $A = \mathbf{1}_n$, d.h.

$$a^i_j = \delta^i_j := \begin{cases} 1, & \text{falls} \quad i = j \\ 0 & \text{sonst} \end{cases} \qquad (\delta^i_j \quad \text{heißt Kroneckersymbol}).$$

- Daraus folgt

$$\epsilon(\sigma)\delta_1^{\sigma(1)} \cdots \delta_n^{\sigma(n)} = \begin{cases} 1, & \text{falls} \quad \sigma = \mathrm{Id} \\ 0 & \text{sonst} \end{cases}$$

und somit

$$Det(A) = Det(\mathbf{1}_n) = \sum_{\sigma \in S_n} \epsilon(\sigma)\delta_1^{\sigma(1)} \cdots \delta_n^{\sigma(n)} = 1.$$

$\square$

Determinante eines Endomorphismus

Definition

- *Sei V ein $\mathbb{K}$-Vektorraum der Dimension n, $F \in \mathrm{End}(V)$ ein Endorphismus und $A \in \mathrm{Mat}(n, \mathbb{K})$ seine darstellende Matrix bez. einer Basis B von V.*
- *Wir definieren:*

$$\det F := \det A.$$

- *Behauptung:* Diese Definition hängt nicht von der Wahl der Basis B ab und ist somit sinnvoll.
- Sei B' eine weitere Basis von V dann gilt:

$$A = \phi_B^{-1} \circ F \circ \phi_B, \quad A' := M_{B'}(F) = \phi_{B'}^{-1} \circ F \circ \phi_{B'}$$

- und somit $A' = SAS^{-1}$, $S = \phi_{B'}^{-1} \circ \phi_B$.
- Also in der Tat
  $\det A' = \det(SAS^{-1}) = \det S \det A (\det S)^{-1} = \det A$.

Orientierung

Definition

- *Sei V ein endlichdimensionaler reeller VR.*
- *Zwei Basen $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_n)$ heißen gleich orientiert, wenn der Basiswechsel $F = \phi_{B'} \circ \phi_B^{-1} : V \to V$ positive Determinante hat.*

Bemerkungen

- Beachte, dass $Fb_i = b'_i$.
- Die darstellende Matrix von F bez. der Basis B ist $M_B(F) = \phi_B^{-1} \circ \phi_{B'} : \mathbb{R}^n \to \mathbb{R}^n$. Daher gilt

$$\det(\phi_{B'} \circ \phi_B^{-1}) = \det F = \det M_B(F) = \det(\phi_B^{-1} \circ \phi_{B'}).$$

Äquivalenzrelation

Definition

- *Eine Relation auf einer Menge X ist eine Teilmenge $R \subset X \times X$. Wir schreiben $x \sim y$ für $(x, y) \in R$.*
- *Eine Relation $R \subset X \times X$ heißt Äquivalenzrelation, wenn sie folgende Eigenschaften hat:*
  (i) $x \sim x$ *für alle* $x \in X$,
  (ii) $x \sim y \Longrightarrow y \sim x$ *und*
  (iii) $x \sim y \sim z \Longrightarrow x \sim z$.
- *Jedes Element $x \in X$ definiert eine Äquivalenzklasse*

$$[x] := \{y \in X | y \sim x\}.$$

Orientierung

Satz

- *Sei V ein reeller VR der Dimension $n \in \mathbb{N}$. Wir schreiben $B \sim B'$, falls B und B' gleich orientierte Basen von V sind.*

a) *Die Relation "$\sim$" ist eine Äquivalenzrelation auf der Menge der Basen von V.*

b) *Jede Basis B definiert eine Äquivalenzklasse*

$$[B] := \{B' \quad \text{Basis von} \quad V | B' \sim B\}$$

und es gibt genau zwei Äquivalenzklassen.

Beweis.

a) ist eine einfache Übungsaufgabe.

b) Sei $B = (b_1, \ldots, b_n)$ eine Basis. Jede Basis ist entweder gleich orientiert zu B oder zu $(-b_1, b_2, \ldots, b_n)$. Daher gibt es genau zwei Äquivalenzklassen. $\square$

Orientierung

Definition

- *Sei V ein reeller Vektorraum der Dimension $n \in \mathbb{N}$.*
- *Eine Orientierung von V ist eine Äquivalenzklasse gleich orientierter Basen von V.*

Bemerkungen

- Jede Basis $B = (b_1, \ldots, b_n)$ definiert eine Orientierung $[B]$.
- Wie wir gesehen haben gibt es genau zwei Orientierungen $[B]$ und $[B]^{op} = [(-b_1, b_2, \ldots, b_n)]$. Letztere heißt die zu $[B]$ entgegengesetzte (oder umgekehrte) Orientierung.
- Automorphismen $F \in \mathrm{GL}(V)$ mit $\det F > 0$ heißen orientierungserhaltend, denn $FB = (Fb_1, \ldots, Fb_n) \in [B]$ für jede Basis $B = (b_1, \ldots, b_n)$.
- $F \in \mathrm{GL}(V)$ mit $\det F < 0$ heißt orientierungsumkehrend, denn $FB = (Fb_1, \ldots, Fb_n) \in [B]^{op}$ für jede Basis B.

Orientierung

Beispiele I

- Die kanonische Orientierung des $\mathbb{R}^n$ ist die durch die kanonische Basis definierte Orientierung $[(e_1, \ldots, e_n)]$.
- Die Basen (e_2, e_3, e_1) und $(-e_2, e_1, e_3)$ definieren z.B. die kanonische Orientierung des $\mathbb{R}^3$.
- Die Basis (e_2, e_1, e_3) definiert die entgegengesetzte Orientierung $[(-e_1, e_2, e_3)]$.
- Die Drehung (um die z-Achse)

$$D = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ist wegen $\det D = 1 > 0$ orientierungserhaltend.

Orientierung

Beispiele II

- Die Spiegelung (an der (x, y)-Ebene)

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

ist wegen $\det S = -1 < 0$ orientierungsumkehrend.

Transponierte Matrix

Definition

Sei $A = (a^i_j)_{i,j}$ eine $(m \times n)$-Matrix. Die $(n \times m)$-Matrix A^t mit den Einträgen

$$(a^t)^i_j := a^j_i$$

heißt die zu A transponierte Matrix.

Satz

Für alle $A \in \mathrm{Mat}(m, n, \mathbb{K})$ und $B \in \mathrm{Mat}(n, p, \mathbb{K})$ gilt

$$(AB)^t = B^t A^t.$$

Beweis.

Für $C = AB$ und $D = B^t A^t$ gilt

$$(c^t)^k_i = c^i_k = \sum_{j=1}^{n} a^i_j b^j_k = \sum_{j=1}^{n} (b^t)^k_j (a^t)^j_i = (d)^k_i. \qquad \square$$

Transponierte Matrix

Folgerung

(i) *$\alpha : A \mapsto (A^{-1})^t$, $\mathrm{GL}(n, \mathbb{K}) \to \mathrm{GL}(n, \mathbb{K})$, ist ein Gruppenautomorphismus.*

(ii) *α ist eine Involution, d.h. $\alpha \circ \alpha = \mathrm{Id}$.*

(iii) *Für alle $A \in \mathrm{GL}(n, \mathbb{K})$ gilt $(A^{-1})^t = (A^t)^{-1}$.*

Beweis.

(ii-iii) Transponieren von $AA^{-1} = \mathbf{1}_n$ liefert $\alpha(A)A^t = \mathbf{1}_n$. Also sind $\alpha(A)$ und A^t invertierbar und $\alpha(A) = (A^t)^{-1}$. Letztere Gleichung ist genau (iii) und und impliziert (ii), denn $A \mapsto (A^t)^{-1}$ ist die Umkehrabbildung zu $\alpha : A \mapsto (A^{-1})^t$.

(i) Für $A, B \in \mathrm{GL}(n, \mathbb{K})$ gilt $(AB)^{-1} = B^{-1}A^{-1}$ und somit

$$\alpha(AB) = (B^{-1}A^{-1})^t = (A^{-1})^t(B^{-1})^t = \alpha(A)\alpha(B).$$

Also ist $\alpha = \alpha^{-1}$ ein Gruppenisomorphismus. $\square$

Transponierte Matrix

Satz

Für jede Matrix $A \in \mathrm{Mat}(n, \mathbb{K})$ gilt

$$\det A = \det A^t.$$

Beweis.

Die Substitution $\sigma' = \sigma^{-1}$ liefert wegen $\epsilon(\sigma') = \epsilon(\sigma)$.

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_1^{\sigma(1)} \cdots a_n^{\sigma(n)} \\
&= \sum_{\sigma' \in S_n} \epsilon(\sigma') a^1_{\sigma'(1)} \cdots a^n_{\sigma'(n)} \\
&= \det A^t
\end{aligned}$$

$\square$

Transponierte Matrix

Folgerung

(i) *Es gibt genau eine Abbildung* $\det : \mathrm{Mat}(n, \mathbb{K}) \to \mathbb{K}$ *mit den folgenden Eigenschaften:*

- D1') $\det$ *ist linear in jeder Zeile,*
- D2') *es gilt* $\det A = 0$*, falls zwei Zeilen von* A *übereinstimmen und*
- D3') $\det(\mathbf{1}_n) = 1$.

(ii) *Addition des* λ*-fachen der* j*-ten Zeile der Matrix* A *zur* i*-ten Zeile* $(i \neq j)$ *ändert nicht den Wert der Determinante.*

(iii) *Vertauschen der* i*-ten und* j*-ten Zeile* $(i \neq j)$ *ändert die Determinante nur um ein Vorzeichen.*

(iv) *Insbesondere läßt sich die Determinante durch Anwendung des Gaußalgorithmus auf die Zeilen der Matrix berechnen.*

Berechnung von Determinanten

- Die Berechnung der Determinante einer Matrix $A \in \mathrm{Mat}(n, \mathbb{K})$ mittels der Formel

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) a^1_{\sigma(1)} \cdots a^n_{\sigma(n)}$$

  ist wegen des schnellen Wachstums von $\mathrm{card}(S_n) = n!$ für große n sehr aufwändig.
- Beachte, dass $n!$ schneller wächst als jede Exponentialfunktion a^n, $a > 1$.

Berechnung von Determinanten

- Für $n \leq 3$ erhalten wir:

$n = 1$

$$a = a_1^1 \in \mathrm{Mat}(1, \mathbb{K}) = \mathbb{K}, \quad \det a = a,$$

$n = 2$

$$\det \begin{pmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{pmatrix} = a_1^1 a_2^2 - a_2^1 a_1^2,$$

$n = 3$

$$\det \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{pmatrix}$$
$$= \quad a_1^1 a_2^2 a_3^3 + a_2^1 a_3^2 a_1^3 + a_3^1 a_1^2 a_2^3 - a_3^1 a_2^2 a_1^3 - a_2^1 a_1^2 a_3^3 - a_1^1 a_3^2 a_2^3,$$

dabei entsprechen die ersten drei Terme den zyklischen Permutationen (123), (231) und (312), die letzten drei den Transpositionen τ_{13}, τ_{12} und τ_{23}.

- Das erste Tripel beginnt mit dem Produkt der Hauptdiagonalelemente, das zweite mit dem Produkt der Nebendiagonalelemente.

Matrixinversion mittels Gaußalgorithmus

- Sei $A \in \mathrm{GL}(n, \mathbb{K})$ und $y \in \mathbb{K}^n$. Der Gaußalgorithmus liefert ein Verfahren um die eindeutige Lösung $x = A^{-1}y$ der Vektorgleichung $Ax = y$ zu bestimmen.
- Dafür wird die Matrix $(A|y)$ durch Zeilenumformungen in die Form $(\mathbf{1}_n|y')$ gebracht. $y' = A^{-1}y = x$ ist dann die gesuchte Lösung.
- Insbesondere berechnet der Gaußalgorithmus die Abbildung $y \mapsto A^{-1}y$, sprich die Inverse Matrix A^{-1}.
- Wir werden nun die Inverse Matrix durch Determinanten ausdrücken.

Streichungsmatrix

Definition

- *Sei* $A = (a^i_j)_{i,j} \in \mathrm{Mat}(n, \mathbb{K})$*. Die Matrix* $A_{ij} \in \mathrm{Mat}(n-1, \mathbb{K})$*, die aus* A *durch Streichen der* i*-ten Zeile und* j*-ten Spalte entsteht heißt* (i,j)*-te Streichungsmatrix von* A*.*

Lemma

(i) $(A^t)_{ij} = (A_{ji})^t$ *und*

(ii) $(-1)^{i+j} \det A_{ij} = \det(a_1 \cdots a_{j-1} e_i a_{j+1} \cdots a_n)$.

Beweis

(i) ist klar.

Beweis.

(ii) Die $(n \times n)$-Matrix $(a_1 \cdots a_{j-1} e_i a_{j+1} \cdots a_n)$ läßt sich durch Addition von Vielfachen der j-ten Spalte zu den anderen Spalten in folgende Form bringen:

$$B_{ij} = \begin{pmatrix} a_1^1 & \cdots & a_{j-1}^1 & 0 & a_{j+1}^1 & \cdots & a_n^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \cdots & a_{j-1}^{i-1} & 0 & a_{j+1}^{i-1} & \cdots & a_n^{i-1} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_1^{i+1} & \cdots & a_{j-1}^{i+1} & 0 & a_{j+1}^{i+1} & \cdots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^n & \cdots & a_{j-1}^n & 0 & a_{j+1}^n & \cdots & a_n^n \end{pmatrix}$$

- Die Matrix B_{ij} läßt sich durch $i-1$ Zeilen- und $j-1$ Spaltenvertauschungen in die Blockdiagonalgestalt $\begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$ bringen. Also $\det(a_1 \cdots a_{j-1} e_i a_{j+1} \cdots a_n) = \det B_{ij} = (-1)^{(i-1)+(j-1)} \det A_{ij} = (-1)^{i+j} \det A_{ij}$. $\square$

Berechnung der Inversen Matrix mittels Determinanten

Satz

- *Sei* $A = (a^i_j)_{i,j} \in \mathrm{Mat}(n, \mathbb{K})$ *und* $\widetilde{A} = (\widetilde{a}^i_j)_{i,j}$ *definiert durch*
$$\widetilde{a}^i_j = (-1)^{i+j} \det A_{ji}.$$
- *Dann gilt* $\widetilde{A}A = A\widetilde{A} = (\det A)\mathbf{1}_n$.

Beweis.

1) Wir zeigen zuerst $\widetilde{A}A = (\det A)\mathbf{1}_n$:

$$\begin{aligned}
\sum_j \widetilde{a}^i_j a^j_k &\overset{Lemma}{=} \sum_j a^j_k \det(a_1 \cdots a_{i-1} e_j \, a_{i+1} \cdots a_n) \\
&= \det(a_1 \cdots a_{i-1} \underbrace{\sum_j a^j_k e_j}_{=a_k} a_{i+1} \cdots a_n) = \delta^i_k \det A.
\end{aligned}$$

Weiter im Beweis:

2) Aus Lemma Teil (i) folgt $\det(A^t)_{ji} = \det(A_{ij})^t = \det A_{ij}$ und somit

$$\widetilde{A^t} = \left(\widetilde{A}\right)^t.$$

- Aus 1) folgt dann

$$\left(\widetilde{A}\right)^t A^t = \widetilde{A^t} A^t = (\det A^t)\mathbf{1}_n = (\det A)\mathbf{1}_n.$$

- Transponieren der letzten Gleichung liefert: $A\widetilde{A} = (\det A)\mathbf{1}_n$.
  $\square$

Entwicklungssatz von Laplace

Folgerung

- *Für jede Matrix* $A \in \mathrm{Mat}(n, \mathbb{K})$ *gilt:*

(Z)

$$\det A = \sum_j (-1)^{i+j} a^i_j \det A_{ij}$$

(Entwicklung nach der i-ten Zeile) und

(S)

$$\det A = \sum_i (-1)^{i+j} a^i_j \det A_{ij}$$

(Entwicklung nach der j-ten Spalte).

Beweis.
Auf der rechten Seite von (Z) bzw (S) steht genau der i-te bzw. j-te Diagonaleintrag der Matrix $A\widetilde{A} = (\det A)\mathbf{1}_n$ bzw. der Matrix $\widetilde{A}A = (\det A)\mathbf{1}_n$. $\square$

Cramersche Regel

Folgerung

- *Sei* $A \in \mathrm{GL}(n, \mathbb{K})$ *und* $b \in \mathbb{K}^n$. *Dann gilt:*

(i) $A^{-1} = (\det A)^{-1}\widetilde{A}$.

(ii) *Die eindeutig bestimmte Lösung* $x = (x^i)_i$ *der Gleichung* $Ax = b$ *berechnet sich nach der folgenden* *Cramerschen Regel*

$$x^i = (\det A)^{-1} \det(a_1 \cdots a_{i-1}\, b\, a_{i+1} \cdots a_n).$$

Beweis.

(i) folgt wg. $\det A \neq 0$ aus $\tilde{A}A = (\det A)\mathbf{1}_n$.

(ii) Nach dem Lemma Teil (ii) ist $\widetilde{a}^i_j = \det(a_1 \cdots a_{i-1} e_j a_{i+1} \cdots a_n)$ und somit

- $x^i = (\det A)^{-1} \sum_j \widetilde{a}^i_j b^j =$
  $(\det A)^{-1} \det(a_1 \cdots a_{i-1}\, b\, a_{i+1} \cdots a_n)$. □

Beispiel

- Sei
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, \mathbb{K}),$$
- d.h. $\det A = ad - bc \neq 0$.
- Dann ergibt sich aus dem vorherigen Satz:
$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Eigenwerte

Definition

- *Sei F ein Endomorphismus eines $\mathbb{K}$-Vektorraums V.*
- *$\lambda \in \mathbb{K}$ heißt Eigenwert von F, wenn es einen Vektor $0 \neq v \in V$ gibt, so dass*

$$F(v) = \lambda v.$$

- *Jeder solche Vektor heißt Eigenvektor von F zum Eigenwert λ.*
- *Der Unterraum*

$$V_\lambda := \{v \in V | F(v) = \lambda v\}$$

heißt Eigenraum von F zum Eigenwert λ.

Eigenwerte

Beispiele

1) Sei $V = \mathbb{R}^2$ und

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{End}(V).$$

- $e_1 + e_2$ ist ein Eigenvektor zum Eigenwert 1,
- $e_1 - e_2$ ein Eigenvektor zum Eigenwert -1,
- $V_1 = \mathbb{R}(e_1 + e_2)$, $V_{-1} = \mathbb{R}(e_1 - e_2)$,
- Da $V = V_1 \oplus V_{-1}$, hat F keine weiteren Eigenwerte.

2) Die Drehung

$$D_\varphi := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

hat keine reellen Eigenwerte, falls φ kein ganzzahliges Vielfaches von π ist.

Charakteristisches Polynom

- Wir nehmen der Einfachheit halber an, dass $\mathbb{K}$ ein unendlicher Körper ist. Dann brauchen wir zwischen Polynomen $P(t) \in \mathbb{K}[t]$ und polynomialen Funktionen $P : \mathbb{K} \to \mathbb{K}$ nicht zu unterscheiden.

Definition

- *Sei V ein $\mathbb{K}$-VR der Dimension $n < \infty$ und $F \in \mathrm{End}(V)$.*
- *Das charakteristische Polynom $P_F : \mathbb{K} \to \mathbb{K}$, $t \mapsto P_F(t)$, von F ist definiert durch*

$$\begin{aligned} P_F(t) &:= \det(F - t\mathrm{Id}) = \det(A - t\mathbf{1}_n) \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma)(a^1_{\sigma(1)} - t\delta^1_{\sigma(1)}) \cdots (a^n_{\sigma(n)} - t\delta^n_{\sigma(n)}), \end{aligned}$$

wobei $A = (a^i_j)_{i,j}$ die darstellende Matrix von F bez. (irgend)einer Basis von V ist.

Charakteristisches Polynom

Satz

- *Das char. Polynom von* $F \in \mathrm{End}(V)$ *hat Grad* $n = \dim V$.
- *Schreibt man*

$$P_F(t) = \sum_{k=0}^{n} \alpha_{n-k}(F)(-t)^k$$

dann ist α_j *homogen vom Grad* j, *d.h.*

- $\alpha_j(\lambda F) = \lambda^j \alpha_j(F)$ *für alle* $\lambda \in \mathbb{K}$ *und* $F \in \mathrm{End}(V)$.
- *Desweiteren gilt:*

$$\begin{aligned}
\alpha_0(F) &= 1 \\
\alpha_1(F) &= \mathrm{spur}(F) := \mathrm{spur}(A) := a_1^1 + \cdots + a_n^n \\
&\vdots \\
\alpha_n(F) &= \det F = \det A.
\end{aligned}$$

Charakteristisches Polynom

Beweis.

- Das Polynom $(a^1_{\sigma(1)} - t\delta^1_{\sigma(1)})\cdots(a^n_{\sigma(n)} - t\delta^n_{\sigma(n)})$ hat für jede Permutation $\sigma \neq \mathrm{Id}$ den Grad $\leq n-2$.
- Daher gilt

$$P_F(t) = (a^1_1 - t)\cdots(a^n_n - t) + Q(t)$$

  mit einem Polynom $Q(t)$ vom Grad $\leq n-2$.
- Ausmultiplizieren liefert nun:

$$P_F(t) = (-t)^n + (-t)^{n-1}(a^1_1 + \cdots + a^n_n) + \cdots + P_F(0)$$

  mit $P_F(0) = \det(F - 0\cdot\mathrm{Id}) = \det F$.
- Die Homogenität von α_j folgt für $\lambda \neq 0$ aus $\lambda^n P_F(t) = P_{\lambda F}(\lambda t) = \sum_{k=0}^n \alpha_{n-k}(\lambda F)(-\lambda t)^k = \sum_{k=0}^n \lambda^k \alpha_{n-k}(\lambda F)(-t)^k$ durch Koeffizientenvergleich. $\square$

Charakteristisches Polynom

Satz

- *Sei V ein endlichdimensionaler VR und $F \in \mathrm{End}(V)$. Dann gilt:*
- *Die Eigenwerte von F sind genau die Nullstellen des charakteristischen Polynoms von F.*

Beweis.

- $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert von F, wenn es einen Vektor $0 \neq v \in V$ gibt mit $F(v) = \lambda v$, d.h. wenn $\ker(F - \lambda \mathrm{Id}) \neq 0$.
- Letzteres ist gleichbedeutend mit $P_F(\lambda) = \det(F - \lambda \mathrm{Id}) = 0$.

$\square$

Diagonalisierbarkeit

Definition

- *Sei V ein $\mathbb{K}$-VR der Dimension n.*
- *$F \in \mathrm{End}(V)$ heißt* **diagonalisierbar**, *wenn es eine Basis $B = (b_1, \ldots, b_n)$ von V gibt, so dass die darstellende Matrix von F Diagonalgestalt hat:*

$$M_B(F) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \quad \lambda_1, \ldots, \lambda_n \in \mathbb{K}.$$

Bemerkung

- Die Basisvektoren b_i sind dann Eigenvektoren von F und die Diagonaleinträge λ_i sind die zugehörigen Eigenwerte λ_i.
- F ist also genau dann diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.

Diagonalisierbarkeit

Satz
Das charakteristische Polynom eines diagonalisierbaren Endomorphismus F zerfällt in Linearfaktoren.

Beweis.
Aus $M_B(F) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ folgt

$$\begin{aligned} P_F(t) &= \det(\mathrm{diag}(\lambda_1 - t, \ldots, \lambda_n - t)) \\ &= (\lambda_1 - t) \cdots (\lambda_n - t) \\ &= (-1)^n (t - \lambda_1) \cdots (t - \lambda_n). \end{aligned}$$

□

Diagonalisierbarkeit

Beispiele I

- Sei
$$F = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{End}(\mathbb{K}^2).$$
- Dann gilt $P_F(t) = t^2 + 1$.

(i) Für $\mathbb{K} = \mathbb{R}$ hat $P_F(t)$ keine Nullstellen und zerfällt also nicht in Linearfaktoren.

- Insbesondere ist F nicht diagonalisierbar.
- (*Bemerkung:* F ist die 90°-Drehung und hat auch deshalb keine Eigenvektoren.)

(ii) Für $\mathbb{K} = \mathbb{C}$ zerfällt $P_F(t) = t^2 + 1 = (t + i)(t - i)$.

- Darüber hinaus ist F diagonalisierbar:

$$\begin{aligned} F(e_1 - ie_2) &= e_2 + ie_1 = i(e_1 - ie_2) \\ F(e_1 + ie_2) &= e_2 - ie_1 = -i(e_1 + ie_2). \end{aligned}$$

Diagonalisierbarkeit

Beispiele II

- Sei
$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{End}(\mathbb{K}^2).$$
- $P_F(t) = (t-1)^2$ zerfällt in Linearfaktoren, aber F ist nicht diagonalisierbar:
- 1 ist der einzige Eigenwert und $V_1 = \mathbb{K}e_1$.
- Es gibt also keine Basis aus Eigenvektoren.

Definition

- *Sei F ein Endomorphismus eines endlichdimensionalen $\mathbb{K}$-Vektorraums V und λ eine Nullstelle von P_F.*
- *Wir bezeichnen die Multiplizität von λ mit m_λ:*

$$P_F(t) = (t-\lambda)^{m_\lambda} Q(t), \quad Q \in \mathbb{K}[t], \quad Q(\lambda) \neq 0.$$

Diagonalisierbarkeit

Satz

$n_\lambda := \dim V_\lambda \leq m_\lambda$.

Beweis.

- Sei $B_\lambda = (b_1, \ldots, b_{n_\lambda})$ eine Basis des Eigenraums V_λ.
- Nach dem Basisergänzungssatz können wir B_λ zu einer Basis B von V ergänzen.
- Dann gilt
$$M_B(F) = \begin{pmatrix} \lambda \mathbf{1}_{n_\lambda} & * \\ 0 & D \end{pmatrix},$$
wobei D eine quadratische Matrix ist.
- Daraus ergibt sich $P_F(t) = (-1)^{n_\lambda}(t-\lambda)^{n_\lambda} P_D(t)$ und somit $n_\lambda \leq m_\lambda$.

□

Einschub: Direkte Summe von mehr als zwei Unterräumen

Definition

- *Sei V ein VR und $V_1, V_2, \dots, V_k \subset V$ Unterräume.*
- *Wir schreiben*

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k, \quad \textit{falls}$$

(i)

$$\begin{aligned} V &= V_1 + V_2 + \cdots + V_k \\ &:= \{v_1 + v_2 + \cdots + v_k | v_i \in V_i, i = 1, 2, \dots, k\} \end{aligned}$$

und

(ii) *jedes k-Tupel $(v_1, v_2, \dots, v_k)$ von Vektoren $v_i \in V_i \setminus 0$ linear unabhängig ist.*

Einschub: Direkte Summe von mehr als zwei Unterräumen

Satz/ÜA

Die folgenden Bedingungen sind äquivalent:

(i) $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$,

(ii) Jeder Vektor $v \in V$ hat eine eindeutige Darstellung $v = v_1 + v_2 + \cdots + v_k$ mit $v_i \in V_i$ und

(iii) $V = (\cdots((V_1 \oplus V_2) \oplus V_3) \oplus \cdots \oplus V_{k-1}) \oplus V_k$.

Bemerkung

- Für unendlichdimensionale Vektorräume V sind die folgenden Verallgemeinerungen nützlich.
- Sei $(V_i)_{i \in I}$ eine beliebige (nicht notwendiger Weise endliche) Familie von Unterräumen $V_i \subset V$.
- Wir definieren $\sum_{i \in I} V_i := \mathrm{span}\{\cup_{i \in I} V_i\}$ und schreiben
- $V = \bigoplus_{i \in I} V_i$, falls (i) $V = \sum_{i \in I} V_i$ und (ii) jede Familie $(v_i)_{i \in I}$ von Vektoren $v_i \in V_i \setminus 0$ linear unabhängig ist.

Diagonalisierbarkeit

Satz

- *Sei V ein VR, $F \in \mathrm{End}(V)$ und v_i, $i = 1, \ldots, k$, seien Eigenvektoren zu paarweise verschiedenen Eigenwerten λ_i.*
- *Dann sind $(v_1, \ldots, v_k)$ linear unabhängig.*

Beweis.

- Wir führen den Beweis durch Induktion nach k.
- Für $k = 1$ ist nichts zu zeigen.
- Wir nehmen an, dass $(v_1, \ldots, v_{k-1})$ linear unabhängig sind (wobei $k \geq 2$).

Diagonalisierbarkeit

Weiter im Beweis:

- Aus $0 = \sum_{i=1}^{k} c_i v_i$ folgt einerseits

$$0 = \sum_{i=1}^{k} \lambda_k c_i v_i$$

- und andererseits durch Anwendung von F

$$0 = \sum_{i=1}^{k} \lambda_i c_i v_i.$$

- Subtraktion der beiden Gleichungen liefert

$$0 = \sum_{i=1}^{k-1} (\lambda_k - \lambda_i) c_i v_i$$

  und somit $c_i = 0$ für $i = 1, \ldots, k-1$.
- Es folgt $0 = \sum_{i=1}^{k} c_i v_i = c_k v_k$ und somit auch $c_k = 0$. $\square$

Diagonalisierbarkeit

Satz

- *Sei V ein endlichdimensionaler VR und $F \in \mathrm{End}(V)$.*
- *F ist genau dann diagonalisierbar, wenn*
- *das charakteristische Polynom P_F in Linearfaktoren zerfällt und $m_\lambda = n_\lambda$ für alle Eigenwerte λ.*

Beweis.

- $\lambda_1, \ldots, \lambda_k$ seien die Eigenwerte von F und $U = \bigoplus_{i=1}^{k} V_{\lambda_i} \subset V$ die Summe der Eigenräume.
- Sei B_U eine Basis von U. Nach dem Austauschlemma können wir B_U zu einer Basis B von V ergänzen.

Diagonalisierbarkeit

Weiter im Beweis:

- Die darstellende Matrix von F bez. B hat dann die Gestalt

$$M_B(F) = \begin{pmatrix} A & * \\ 0 & D \end{pmatrix},$$

- wobei A die darstellende Matrix der Einschränkung $F|_U$ bez. B_U ist und D eine $(m \times m)$-Matrix ist.
- Hierbei ist $m = \dim V - \dim U$.
- Daraus folgt $P_F(t) = P_A(t)Q(t)$, wobei $Q(t) = P_D(t)$, falls $m > 0$ und $Q(t) = 1$, falls $m = 0$.
- F ist genau dann diagonalisierbar, wenn $U = V$, d.h. wenn $m = 0$.
- $m = 0$ ist gleichbedeutend mit $P_F(t) = P_A(t) = (\lambda_1 - t)^{n_{\lambda_1}} \cdots (\lambda_k - t)^{n_{\lambda_k}}$, dh. mit (i-ii).

$\square$

Multilineare Abbildungen

Definition

- *V und W seien $\mathbb{K}$-Vektorräume.*
- *Eine Abbildung*

$$\mu : V^k = V \times \cdots \times V \quad \rightarrow \quad W$$
$$(v_1, \ldots, v_k) \quad \mapsto \quad \mu(v_1, \ldots, v_k)$$

heißt multilinear (genauer k-linear), wenn

- *μ in jedem der k Argumente linear ist, d.h. wenn*
-

$$\begin{aligned} & \mu(v_1, \ldots, v_{i-1}, \lambda v + \lambda' v', v_{i+1}, \ldots, v_k) \\ = \quad & \lambda \mu(v_1, \ldots, v, \ldots, v_k) + \lambda' \mu(v_1, \ldots, v', \ldots, v_k) \end{aligned}$$

für alle $i = 1, \ldots, k$, $v, v' \in V$, $\lambda, \lambda' \in \mathbb{K}$.

Multilineare Abbildungen

Beispiele

$k = 1$: 1-lineare Abbildungen $V \to W$ sind genau lineare Abbildungen.

$k = 2$: 2-lineare Abbildungen $V \times V \to W$ nennt man auch bilineare Abbildungen.

$W = \mathbb{K}$: Multilineare Abbildungen $V \times \cdots \times V \to \mathbb{K}$ mit Werten in $\mathbb{K}$ nennt man auch multilineare Formen.

- In den Fällen $k = 1, 2$ spricht man von Linearformen bzw. von Bilinearformen auf V.
- Der Vektorraum aller Linearformen auf V heißt Dualraum von V und wird mit V^* bezeichnet.
- Bespielsweise ist $f \mapsto \int_a^b f(x)dx$ eine Linearform auf dem VR der integrierbaren Funktionen auf dem Intervall $[a, b]$.
- Die Determinante $\det : \mathrm{Mat}(n, \mathbb{K}) = (\mathbb{K}^n)^n \to \mathbb{K}$ ist n-linear.

Dualraum

Satz

Für jeden endlichdimensionalen VR gilt $\dim V^* = \dim V$.

Beweis.

- Sei $B = (b_1, \dots, b_n)$ eine Basis von V.
- Sei $b^i = b_i^* \in V^*$ die Linearform mit
$$b^i(b_j) = \delta_j^i, \quad j = 1, \dots n.$$
- *Behauptung:* Dann ist $(b^1, \dots, b^n)$ eine Basis von V^* (die sogenannte duale Basis) und somit $\dim V^* = n = \dim V$.
- Die b^i sind linear unabhängig, denn aus $\sum_{i=1}^n \lambda_i b^i = 0$ folgt $0 = \sum_{i=1}^n \lambda_i b^i(b_j) = \lambda_j$ für alle j.
- Sie bilden auch ein Erzeugendensystem, denn für jede Linearform $\beta \in V^*$ gilt $\beta = \sum_{i=1}^n \beta(b_i) b^i$.

$\square$

Duale Abbildung

Definition

- *Zu jeder linearen Abbildung $F : V \to W$ gehört eine lineare Abbildung*
$$F^* : W^* \to V^*,$$
- *die definiert ist durch*
$$F^*(w^*)(v) := w^*(F(v)), \quad \textit{für alle} \quad v \in V, \quad w^* \in W^*.$$
- *Sie heißt die zu F duale Abbildung.*

Darstellende Matrix der dualen Abbildung

Satz

- *Sei $F : V \to W$ eine lineare Abbildung zwischen endlichdimensionalen $\mathbb{K}$-Vektorräumen.*
- *Sei $A = M_B^{B'}(F) \in \mathrm{Mat}(m, n, \mathbb{K})$ ihre darst. Matrix bez. Basen $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_m)$ von V bzw. W.*
- *Dann gilt für die darst. Matrix der dualen Abbildung $F^* : W^* \to V^*$ bez. der dualen Basen $B'^* = ({b'_1}^*, \ldots, {b'_m}^*)$ und $B^* = (b_1^*, \ldots, b_n^*)$ von W^* bzw. V^*:*
$$M_{B'^*}^{B^*}(F^*) = A^t \in \mathrm{Mat}(n, m, \mathbb{K}).$$

Beweis.

- Aus $F^*({b'_i}^*)(b_j) = {b'_i}^*(Fb_j) = {b'_i}^*(\sum_{k=1}^m a_j^k b'_k) = \sum_{k=1}^m a_j^k {b'_i}^*(b'_k) = \sum_{k=1}^m a_j^k \delta_k^i = a_j^i$ folgt
- $F^*({b'_i}^*) = \sum_{j=1}^n a_j^i b_j^*$, d.h. $M_{B'^*}^{B^*}(F^*) = A^t$. $\square$

Symmetrische und schiefymmetrische Bilinearformen

Definition

- *Eine Bilinearform*
$$\beta : V \times V \to \mathbb{K}$$
heißt symmetrisch, wenn
-
$$\beta(v, w) = \beta(w, v)$$
für alle $v, w \in V$.
- *Sie heißt schiefsymmetrisch, wenn*
$$\beta(v, w) = -\beta(w, v) \quad \text{für alle} \quad v, w \in V.$$
- *Die Bilinearform β heißt nicht entartet, wenn die lineare Abbildung $v \mapsto \beta(v, \cdot)$, $V \to V^*$, injektiv ist.*
- *(M.a.W. wenn $v = 0$ der einzige Vektor ist, für den $\beta(v, w) = 0$ für alle $w \in V$.)*

Darstellende Matrix einer Bilinearform

Definition

- *Sei $\alpha : V \times V \to \mathbb{K}$ eine Bilinearform und $B = (b_1, \ldots, b_n)$ eine Basis von V.*
- *Die $(n \times n)$-Matrix A mit den Einträgen*

$$\alpha_{ij} := \alpha(b_i, b_j)$$

heißt darstellende Matrix von α.

Bemerkungen/ÜA

- α ist genau dann symmetrisch bzw. schiefsymmetrisch, wenn A symmetrisch bzw. schiefsymmetrisch ist, d.h. wenn $A = A^t$ bzw. $A = -A^t$.
- α ist genau dann nicht entartet, wenn A invertierbar ist.
- (Hinweis: Die darstellende Matrix der linearen Abbildung $v \mapsto \alpha(v, \cdot)$, $V \to V^*$, ist A^t.)

Euklidische Vektorräume

Definition

- *Sei V ein reeller VR.*
- *Eine symmetrische Bilinearform*

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}, \quad (v, w) \mapsto \langle v, w \rangle$$

- *heißt positiv definit, wenn für alle $v \in V \setminus 0$*

$$\langle v, v \rangle > 0.$$

- *Ein (Euklidisches) Skalarprodukt auf V ist eine positiv definite symmetrische Bilinearform $V \times V \to \mathbb{R}$.*
- *Ein Euklidischer Vektorraum ist ein reeller VR zusammen mit einem Skalarprodukt darauf.*

Bemerkung

Skalarprodukte sind nicht entartet, denn aus $\langle v, v \rangle = 0$ folgt $v = 0$.

Euklidische Vektorräume

Beispiel

- Das kanonische Skalarprodukt von $\mathbb{R}^n$ ist definiert durch

$$\langle x, y \rangle := \sum_{i=1}^{n} x^i y^i,$$

wobei $x = \sum_{i=1}^{n} x^i e_i$ und $y = \sum_{i=1}^{n} y^i e_i$.

- Die darstellende Matrix des kanonischen Skalarprodukts von $\mathbb{R}^n$ bez. der kanonischen Basis ist die Einheitsmatrix $\mathbf{1}_n$, denn
-

$$\langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1, & \text{falls} \quad i = j \\ 0 & \text{sonst.} \end{cases}$$

Länge eines Vektors

Definition

- *Sei V ein Euklidischer VR.*
- *Die Länge eines Vektors $v \in V$ ist die Zahl*

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

- *Der Abstand $d(x, y)$ zweier Punkte $x, y \in V$ ist die Länge des Vektors $y - x$:*

$$d(x, y) := \|y - x\|.$$

Die Cauchy-Schwarzsche Ungleichung

Satz

- *Sei V ein Euklidischer VR.*
- *Dann gilt für alle $v, w \in V$:*

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|,$$

- *mit Gleichheit genau dann, wenn v und w linear abhängig sind.*

Beweis.

- Wir können annehmen, dass $v \neq 0$ und $w \neq 0$, denn sonst ist nichts zu zeigen.

Die Cauchy-Schwarzsche Ungleichung

Weiter im Beweis:

- Für alle $\lambda, \mu \in \mathbb{R}$ gilt

$$0 \leq \langle \lambda v + \mu w, \lambda v + \mu w \rangle = \lambda^2 \|v\|^2 + \mu^2 \|w\|^2 + 2\lambda\mu \langle v, w \rangle.$$

- Einsetzen von $\lambda = \sqrt{\frac{\|w\|}{\|v\|}}$ und $\mu = \pm\sqrt{\frac{\|v\|}{\|w\|}}$ liefert die CSU.
- Gleichheit gilt genau dann, wenn für eine dieser beiden Wahlen $\lambda v + \mu w = 0$, d.h. wenn v und w linear abhängig sind. $\square$

Winkel

Definition

- *Sei V ein Euklidischer VR und $v, w \in V \setminus 0$.*
- *Der Winkel $\angle(v, w)$ zwischen v und w ist definiert als*

$$\angle(v, w) := \arccos \underbrace{\frac{\langle v, w \rangle}{\|v\| \|w\|}}_{\in[-1,1]} \in [0, \pi].$$

- *Man sagt, dass $v, w \in V$ senkrecht aufeinander stehen, wenn*

$$\langle v, w \rangle = 0$$

und schreibt dafür $v \perp w$.

Normierte Vektorräume

Definition

- *Sei V eine $\mathbb{K}$-VR, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.*

(i) *Eine Norm auf V ist eine Abbildung*

$$\|\cdot\| : V \to [0, \infty), \quad v \mapsto \|v\|,$$

- *mit folgenden Eigenschaften:*

N1) *Für $v \in V$ gilt genau dann $\|v\| = 0$, wenn $v = 0$.*

N2) $\|\lambda v\| = |\lambda| \cdot \|v\|$ *für alle $v \in V$, $\lambda \in \mathbb{K}$.*

N3) $\|v + w\| \leq \|v\| + \|w\|$ *für alle $v, w \in V$ (Dreiecksungleichung).*

(ii) *Ein normierter Vektorraum ist ein reeller oder komplexer VR zusammen mit einer Norm darauf.*

Normierte Vektorräume

Satz

- *Sei V ein reeller VR. Jedes Skalarprodukt $\langle\cdot,\cdot\rangle$ auf V definiert eine Norm $\|\cdot\|$ auf V:*
- $$\|v\| := \sqrt{\langle v, v\rangle}, \quad v \in V.$$

Beweis.

- N1) folgt daraus, dass $\langle\cdot,\cdot\rangle$ positiv definit ist.
- N2) folgt aus der Bilinearität von $\langle\cdot,\cdot\rangle$:
  $\|\lambda v\|^2 = \langle \lambda v, \lambda v\rangle = \lambda^2\langle v, v\rangle = \lambda^2\|v\|^2$.
- N3) folgt aus der CSU:

$$\begin{aligned} \|v+w\|^2 &= \|v\|^2 + \|w\|^2 + 2\langle v, w\rangle \\ &\leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2. \quad \square \end{aligned}$$

Metrische Räume

Definition

- *Eine Metrik (ein Abstandsbegriff) auf einer Menge X ist eine Abbildung $d : X \times X \to [0, \infty)$ mit folgenden Eigenschaften:*

M1) *Für $x, y \in X$ gilt genau dann $d(x, y) = 0$, wenn $x = y$.*

M2) $d(x, y) = d(y, x)$ *für alle* $x, y \in X$.

M3) $d(x, z) \leq d(x, y) + d(y, z)$ *für alle* $x, y, z \in X$ *(Dreiecksungleichung).*

Metrische Räume

Satz

Jede Norm $\|\cdot\|$ *auf einem reellen oder komplexen VR* V *definiert eine Metrik* $d : V \times V \to [0, \infty)$ *durch*

$$d(x,y) := \|x-y\|, \quad x, y \in V.$$

Beweis.

M1, M2 sind trivial. M3 folgt aus N3:
$\|x-z\| = \|x-y+y-z\| \leq \|x-y\| + \|y-z\|$. $\square$

Beispiel

- Die durch die Euklidische Norm von $\mathbb{R}^n$ definierte Metrik d heißt die Euklidische Metrik des $\mathbb{R}^n$.

Metrische Räume

Definition

Sei (X, d) ein metrischer Raum.

(i) *Eine Folge $x_n \in X$, $n \in \mathbb{N}$, heißt konvergent mit dem Grenzwert $x \in X$,*

wenn die Folge $d(x_n, x) \in \mathbb{R}$ eine Nullfolge ist.

(ii) *Eine Funktion $f : X \to \mathbb{C}$ heißt im Punkt $x \in X$ stetig, wenn $\lim_{n\to\infty} f(x_n) = f(x)$ für jede gegen $x \in X$ konvergente Folge $x_n \in X$.*

- *f heißt stetig, wenn f in allen Punkten $x \in X$ stetig ist.*

Beispiel

Sei $(V, \|\cdot\|)$ ein normierter VR mit zugehöriger Metrik d.

Dann ist $\|\cdot\| : V \to \mathbb{R}$ eine stetige Funktion (bezüglich der Metrik d), denn aus N3 folgt $|\|x\| - \|y\|| \leq \|x - y\|$.

Die Parallelogrammgleichung

Satz

- *Eine Norm $\|\cdot\|$ auf einem reellen VR V wird genau dann durch ein SKP auf V induziert, wenn*
- *für alle $v, w \in V$ die Parallelogrammgleichung gilt:*

$$(*) \quad \|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

Beweis.

"$\Longrightarrow$" Aus $\|u\|^2 = \langle u, u \rangle$ für alle $u \in V$ ergibt sich (*) durch Berechnung der linken Seite mittels der Bilinearität des SKP.

Die Parallelogrammgleichung

Weiter im Beweis:

"$\Longleftarrow$" Wir definieren

$$\langle v, w\rangle := \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right).$$

- Dann gilt offenbar $\langle v, w\rangle = \langle w, v\rangle$ und $\langle v, v\rangle = \|v\|^2$.
- Es genügt daher die Linearität von $\langle\cdot,\cdot\rangle$ im ersten Argument nachzuweisen.
- Für alle $u, u' \in V$ gilt:

$$\begin{aligned}
& \langle u+u', w\rangle + \langle u-u', w\rangle \\
= \quad & \frac{1}{4}\left(\|u+u'+w\|^2 - \|u+u'-w\|^2\right. \\
& \left. + \|u-u'+w\|^2 - \|u-u'-w\|^2\right) \\
\overset{(*)}{=} \quad & \frac{1}{2}(\|u+w\|^2 + \|u'\|^2) - \frac{1}{2}(\|u-w\|^2 + \|u'\|^2) \\
= \quad & \frac{1}{2}(\|u+w\|^2 - \|u-w\|^2) = 2\langle u, w\rangle.
\end{aligned}$$

Die Parallelogrammgleichung

Weiter im Beweis:

- Wir haben gezeigt, dass

$$(1) \quad \langle u + u', w \rangle + \langle u - u', w \rangle = 2\langle u, w \rangle.$$

- Für $u = u'$ erhält man $\langle 2u, w \rangle = 2\langle u, w \rangle$.
- Mit $v = u + u'$ und $v' = u - u' \in V$ erhalten wir daher aus (1):

$$\langle v, w \rangle + \langle v', w \rangle = \langle 2u, w \rangle = \langle v + v', w \rangle.$$

- Daraus folgt leicht $\langle nv, w \rangle = n\langle v, w \rangle$ für alle $n \in \mathbb{Z}$.
- Einsetzen von $v = \frac{1}{n} u$ liefert $\langle u, w \rangle = n\langle \frac{1}{n} u, w \rangle$ für alle $u \in V, n \in \mathbb{Z}^*$ und somit

$$(2) \quad \langle \lambda u, w \rangle = \lambda \langle u, w \rangle$$

  für alle $\lambda \in \mathbb{Q}$.
- Die Stetigkeit der Norm liefert nun die Gleichung (2) für alle $\lambda \in \mathbb{R}$, da es zu jeder reellen Zahl λ eine Folge rationaler Zahlen gibt, die gegen λ konvergiert. $\square$

Die Parallelogrammgleichung

Beispiele

(i) Durch

$$\|x\|_1 := \sum_{i=1}^{n} |x^i|, \quad x = \sum x^i e_i \in \mathbb{R}^n,$$

wird eine Norm auf $\mathbb{R}^n$ definiert.

Diese erfüllt nicht die Parallelogrammgleichung, denn z.B. ist

$$8 = \|e_1 + e_2\|^2 + \|e_1 - e_2\|^2 \neq 2\|e_1\|^2 + 2\|e_2\|^2 = 4.$$

(ii) Die Euklidische Norm auf $\mathbb{R}^n$ ist die durch das kanonische SKP definierte Norm: $\|x\| := \|x\|_2 := \sqrt{\sum_{i=1}^{n}(x^i)^2}$.

(iii) Die Maximumsnorm

$$\|x\|_{max} := \max\{|x^1|, |x^2|, \dots, |x^n|\}$$

auf $\mathbb{R}^n$ erfüllt nicht die Parallelogrammgleichung.

Orthonormale Basen

Satz

- *Sei V ein Euklidischer VR und $(v_i)_{i\in I}$ eine orthogonale Familie von Vektoren $v_i \in V \setminus 0$, d.h.*

$$v_i \perp v_j \quad \textit{für alle} \quad i \neq j.$$

- *Dann ist $(v_i)_{i\in I}$ linear unabhängig.*

Beweis.

- Sei $0 = \sum_{j\in J} \lambda_j v_j$, wobei $J \subset I$ eine endliche Teilmenge ist und $\lambda_j \in \mathbb{R}$.
- Daraus folgt für alle $i \in J$

$$0 = \langle v_i, \sum_{j\in J} \lambda_j v_j \rangle = \sum_{j\in J} \lambda_j \langle v_i, v_j \rangle = \lambda_i \|v_i\|^2$$

und somit $\lambda_i = 0$. $\square$

Orthonormale Basen

Definition

- *Sei V ein n-dimensionaler Euklidischer VR.*

(i) *Eine Vektor $v \in V$ heißt Einheitsvektor, wenn $\|v\| = 1$.*

(ii) *Eine orthogonale Familie $(b_1, \ldots, b_n)$ bestehend aus n Einheitsvektoren $b_i \in V$ heißt Orthonormalbasis von V.*

(iii) *Für jeden Unterraum $U \subset V$ heißt der UR*

$$U^\perp := \{v \in V | v \perp U\}$$

das orthogonale Komplement von U in V.

Beispiel

Die kanonische Basis von $\mathbb{R}^n$ ist orthonormal (bez. des kan. SKP).

Orthonormale Basen

Lemma

- *Sei β eine nichtentartete Bilinearform auf einem endlichdimensionalen VR V und $U \subset V$ ein UR.*

(i) *Dann hat der UR*

$$U' := \quad \{v \in V | \beta(u, v) = 0 \quad \textit{für alle} \quad u \in U\}$$

die Dimension $\dim V - \dim U$*.*

(ii) *Die Unterräume U und U' sind genau dann komplementär, wenn $U \cap U' = 0$.*

Beweis.

Sei $(u_1, \ldots, u_r)$ eine Basis von U. Der Unterraum U' ist genau der Lösungsraum des linearen Gleichungssystems

$$\beta(u_i, v) = 0, \quad \text{für alle} \quad i = 1, \ldots, r.$$

Orthonormale Basen

Weiter im Beweis:

- Somit gilt $\dim U' = \dim V - r$, denn unter der injektiven linearen Abbildung $u \mapsto \beta(u, \cdot)$ gehen die Basisvektoren $u_i \in U$ in linear unabhängige Linearformen $\beta(u_i, \cdot) \in V^*$ über.
- (ii) folgt aus (i) und der Dimensionsformel:

$$\begin{aligned} \dim(U + U') &= \dim U + \dim U' - \dim(U \cap U') \\ &\overset{(i)}{=} \dim V - \dim(U \cap U'). \end{aligned}$$

$\square$

Orthonormale Basen

Satz

- *Sei $U \subset V$ ein UR eines endlichdimensionalen Euklidischen Vektorraums.*
- *Dann ist $U^{\perp}$ ein Komplement zu U in V.*

Beweis.

- Nach dem vorigen Lemma genügt es $U \cap U^{\perp} = 0$ zu zeigen.
- Aus $v \in U \cap U^{\perp}$ folgt $\langle v, v \rangle = 0$ und somit $v = 0$.

$\square$

Das Orthonormalisierungsverfahren von Gram und Schmidt

Satz (Gram-Schmidt)

Jeder endlichdimensionale Euklidische VR V besitzt eine ONB.

Beweis.

- Wir beweisen den Satz durch Induktion nach $n = \dim V \in \mathbb{N}$.
- Falls $\dim V = 1$, so existiert $v \in V \setminus 0$ und $b_1 := \frac{v}{\|v\|}$ ist eine ONB.
- Wir nehmen an, dass jeder Euklid. VR der Dimension $< n$ eine ONB hat und zeigen, dass dann auch jeder n-dimensionale Euklid. VR eine ONB hat.
- Sei $v \in V \setminus 0$. Nach Induktion besitzt der $(n-1)$-dimensionale UR $v^\perp := (\mathbb{R}v)^\perp$ eine ONB $(b_2, \dots, b_n)$.
- Durch $b_1 := \frac{v}{\|v\|}$ wird diese zu einer ONB von V ergänzt.

□

Die orthogonale Gruppe

Definition

- *Sei* $(V, \langle\cdot,\cdot\rangle)$ *ein Euklidischer VR.*
- *Eine surjektive lineare Abbildung* $F : V \to V$ *heißt orthogonal, wenn*

$$(*) \quad \langle F(v), F(w)\rangle = \langle v, w\rangle \quad \textit{für alle} \quad v, w \in V.$$

- *Bemerkung: Aus (*) folgt, dass* F *injektiv ist. Falls* $\dim V < \infty$, *so folgt daraus die Surjektivität. Diese muss dann also nicht gefordert werden.*
- *ÜA: Die orthogonalen Abbildungen bilden eine Untergruppe* $\mathrm{O}(V) \subset \mathrm{Aut}(V)$, *die sogenannte orthogonale Gruppe.*

Beispiel

$A \in \mathrm{O}(n) := \mathrm{O}(\mathbb{R}^n) \iff A^t A = \mathbf{1}_n.$