

Mathematik für Informatiker

Matthias Künzer

Universität Stuttgart

22. Februar 2022

Inhalt

1	Grundlagen	6
1.1	Aussagen	6
1.2	Mengen	8
1.3	Quantoren	10
1.4	Abbildungen	11
1.5	Relationen	13
1.6	Induktion	18
1.7	Rekursive Definition	21
1.8	Binomischer Lehrsatz	23
1.9	Euklidischer Algorithmus	24
1.10	Primzahlen	27
1.11	Zifferndarstellung	28
2	Körper	29
2.1	Restklassen	29
2.2	Monoide und Gruppen	31
2.3	Kommutative Ringe und Körper	32
2.4	Polynome	35
2.5	Polynome über einem Körper	36
2.6	Konstruktion von Körpererweiterungen	39
2.7	Der Körper \mathbb{C} der komplexen Zahlen	40
2.8	Der Körper \mathbb{F}_4 mit 4 Elementen	41
2.9	Der Körper \mathbb{F}_8 mit 8 Elementen	42
2.10	Der Körper \mathbb{F}_9 mit 9 Elementen	43
2.11	Körper: Eine Übersicht	44
2.12	Mehr zu den komplexen Zahlen \mathbb{C}	44
3	Lineare Algebra	50
3.1	Matrizen	50
3.2	Gaußverfahren	53
3.3	Vektorräume	58
3.3.1	Definition	58
3.3.2	Dimension	59
3.3.3	Unterräume	62
3.4	Der Standardvektorraum $\mathbb{R}^{n \times 1}$	65
3.4.1	Geometrische Interpretation	65
3.4.2	Skalarprodukt	66
3.4.3	Kreuzprodukt in $\mathbb{R}^{3 \times 1}$	67
3.5	Lineare Abbildungen	68
3.5.1	Definition	68
3.5.2	Koordinatenvektoren und beschreibende Matrizen	71
3.6	Determinanten	74
3.7	Normalform	78
3.7.1	Eigenwerte, Eigenvektoren, Eigenräume	78
3.7.2	Direkte Summen	81
3.7.3	Haupträume	81

3.7.4	Jordansche Normalform	86
3.8	Matrizen über \mathbb{R} und über \mathbb{C}	93
3.8.1	Gram-Schmidt	93
3.8.1.1	Gram-Schmidt über \mathbb{R}	93
3.8.1.2	Hermiteisches Skalarprodukt und Gram-Schmidt über \mathbb{C}	95
3.8.2	Symmetrische und orthogonale Matrizen über \mathbb{R}	97
3.8.3	Hermiteische und unitäre Matrizen über \mathbb{C}	99
3.8.4	Eine informelle Übersicht	103
4	Funktionen in einer Variablen	104
4.1	Grenzwerte	104
4.2	Stetigkeit	109
4.3	Grenzwerte von Folgen	111
4.4	Limes superior	113
4.5	Reihen	114
4.5.1	Reihenbegriff	114
4.5.2	Konvergenzkriterien	116
4.6	Differentialrechnung	119
4.6.1	Differenzierbarkeit	119
4.6.2	Ableitungsregeln	120
4.6.3	Sinus und Cosinus	122
4.6.4	Die Grenzwertregel nach l'Hôpital	125
4.6.5	Taylor	126
4.6.6	Monotonie	132
4.6.7	Begriffserweiterung von \mathbb{R} nach \mathbb{C}	133
4.6.8	Potenzreihen	135
4.6.9	Exponentialfunktion	137
4.6.10	Logarithmus	140
4.6.11	Ableitung von Umkehrfunktionen	142
4.7	Integralrechnung	144
4.7.1	Integralbegriff	144
4.7.2	Eigenschaften des Integrals	147
4.7.3	Hauptsatz	148
4.7.4	Integrationstechniken	152
4.7.4.1	Partielle Integration	152
4.7.4.2	Komplexwertige Integranden	152
4.7.4.3	Substitution	153
4.7.4.4	Partialbruchzerlegung	156
4.7.5	Uneigentliche Integrale	160
5	Funktionen in mehreren Variablen	164
5.1	Partielle Ableitungen	164
5.2	Definitheit	167
5.3	Lokale Extremstellen	168
5.4	Lokale Extremstellen unter Nebenbedingungen	171
6	Gewöhnliche Differentialgleichungen	178
6.1	Begriff	178
6.2	Differentialgleichungen erster Ordnung	179
6.2.1	Richtungsfeld	179
6.2.2	Separierbare Differentialgleichungen	180
6.2.3	Ähnlichkeits-Differentialgleichungen	183
6.2.4	Lineare Differentialgleichungen	185
6.2.5	Existenz und Eindeutigkeit einer Lösung	188

6.3	Lineare Differentialgleichungssysteme erster Ordnung	190
6.3.1	Der allgemeine Fall	190
6.3.2	Der Fall konstanter Koeffizienten	197
6.3.2.1	Matrixexponentialfunktion	197
6.3.2.2	Lösungsverfahren	204
6.4	Lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten	210
6.5	Weitere zurückführbare Differentialgleichungen	215
6.5.1	Bernoullische Differentialgleichung	215
6.5.2	Riccatische Differentialgleichung	216

Vorwort

Inhalt sind Lineare Algebra und Analysis.

Vorausgesetzt wird Abiturwissen, wenn wir auch Teile davon wiederholen werden.

Als Grundlage diente das Skript von P. LESKY.

Solange das Skript im Entstehen ist, wird wegen laufender Änderungen empfohlen, es nicht auszudrucken, sondern als File regelmäßig zu aktualisieren.

Dank für Verbesserungen geht an Mathias Ritter, Sebastian Nitsche, Monika Truong, Roman Haak, Mišo Gavrilovic, Niklas Schmidt, André Sperrle, Anton Schmid, Stefan Sawatzky, Marvin Stratmann, Fabian Leeske, Alexandru Olaru, David Augustat, Mohamed Ben Salha, Skander El-Habib Ben Hassine, Daniel Schweikert, Florian Bauer, Bastian Neuhold, Elias Schwesig und Studenten im Hörsaal und im Forum.

Für weitere Hinweise auf Fehler und Unklarheiten bin ich dankbar.

Stuttgart, Sommer 2020

Matthias Künzer

Kapitel 1

Grundlagen

1.1 Aussagen

Definition. Eine *Aussage* A ist eine Behauptung, der eindeutig entweder der Wahrheitswert “wahr” oder der Wahrheitswert “falsch” zugewiesen werden kann.

Ist A wahr, so sagen wir, A *gilt*.

Ist A falsch, so sagen wir, A *gilt nicht*.

Beispiel.

- (1) “Jede Primzahl ist ungerade” ist eine Aussage. Diese Aussage hat den Wahrheitswert falsch, sie gilt also nicht.
- (2) “Jede Primzahl ist ≥ 2 ” ist eine Aussage. Diese Aussage gilt.
- (3) “Jede Primzahl ist schön” ist keine Aussage.

Definition. Seien A, B, C Aussagen.

- (1) Die Aussage $\neg A$ gelte genau dann, wenn A nicht gilt.
Wir sprechen $\neg A$ als “nicht A ”.
- (2) Die Aussage $A \vee B$ gelte genau dann, wenn A gilt oder B gilt.
Wir sprechen $A \vee B$ als “ A oder B ”.
- (3) Die Aussage $A \wedge B$ gelte genau dann, wenn A gilt und B gilt.
Wir sprechen $A \wedge B$ als “ A und B ”.
- (4) Sei $(A \Rightarrow B) := (\neg A \vee B)$. Dabei sei “:=” das Symbol für “ist definiert als”.
Wir sprechen $A \Rightarrow B$ als “aus A folgt B ” oder “ A impliziert B ”.

Man sagt dann auch, es ist A *hinreichend* für B .

Man sagt dann auch, es ist B *notwendig* für A .

Man kürzt auch $(A \Rightarrow B) \wedge (B \Rightarrow C)$ ab zu $A \Rightarrow B \Rightarrow C$.

- (5) Sei $(A \Leftrightarrow B) := (A \Rightarrow B) \wedge (B \Rightarrow A)$. Wir sprechen $A \Leftrightarrow B$ als “ A ist äquivalent zu B ”.

Man kürzt auch $(A \Leftrightarrow B) \wedge (B \Leftrightarrow C)$ ab zu $A \Leftrightarrow B \Leftrightarrow C$.

Die zugehörige Wahrheitstafel hat folgende Form. In dieser wird “wahr” als “w” sowie “falsch” durch “f” abgekürzt.

A	B	$\neg A$	$A \vee B$	$A \wedge B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	w	f	f	f
f	w	w	w	f	w	f
f	f	w	f	f	w	w

Beispiel.

- (1) Für $x \in \mathbb{R}$ ist $(x^2 < 9) \Leftrightarrow ((-3 < x) \wedge (x < 3))$. Letzteres faßt man auch zu $-3 < x < 3$ zusammen.
- (2) Für $x \in \mathbb{R}$ ist $(x > 1) \implies (x > 0)$. Damit $x > 1$ ist, ist es also notwendig, daß $x > 0$ ist. Umgekehrt ist es hinreichend, $x > 1$ zu wissen, um auf $x > 0$ schließen zu können.

Beispiel. Seien A und B Aussagen.

- (1) Es ist $(\neg A \vee \neg B) \Leftrightarrow \neg(A \wedge B)$.
- (2) Es ist $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B) \Leftrightarrow (B \vee \neg A) \Leftrightarrow (\neg B \Rightarrow \neg A)$.
Die Variante $(\neg B \Rightarrow \neg A)$ heißt auch *Kontraposition*.
- (3) Es ist $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B) \Leftrightarrow \neg(A \wedge \neg B)$

In anderen Worten, wenn wir unter der Voraussetzung, daß A gilt, annehmen, daß B nicht gilt, so müssen wir zeigen, daß dies nicht sein kann. Dies ist die Grundlage von Widerspruchsbeweisen.

Beispiel: Wir wollen für $x \in \mathbb{R}$ mit $x > 0$ zeigen, daß $(x^2 = 2) \Rightarrow \neg(x \in \mathbb{Q})$ gilt. Wir müssen zeigen, daß $(x^2 = 2) \wedge (x \in \mathbb{Q})$ nicht sein kann.

Annahme, es ist $(x^2 = 2) \wedge (x \in \mathbb{Q})$. Wir schreiben $x = \frac{a}{b}$ als gekürzten Bruch. Dann ist $\frac{a^2}{b^2} = 2$, also $a^2 = 2b^2$. Also ist a gerade. Also ist b ungerade, da $\frac{a}{b}$ gekürzt war. Nun aber ist a^2 durch 4 teilbar, aber $2b^2$ nicht durch 4 teilbar. Wir haben den gewünschten *Widerspruch*.

1.2 Mengen

Definition. Eine *Menge* M ist eine Ansammlung von Elementen.

Wir schreiben $x \in M$, wenn x ein Element von M ist. Sonst schreiben wir $x \notin M$.

Wir schreiben $M = \{x_1, x_2, \dots, x_k\}$, wenn M genau die Elemente x_1, \dots, x_k enthält.

Enthält eine endliche Menge M genau ℓ verschiedene Elemente, so schreiben wir $|M| := \ell$ für die *Kardinalität* von M .

Es gibt auch eine axiomatische Mengenlehre, in welcher der Begriff der Menge formaler gefaßt wird.

Beispiel.

- (1) Die leere Menge $\emptyset = \{ \}$ enthält kein Element.
- (2) Es ist $\{3, 2, 3, 1\} = \{1, 2, 3\}$. Es ist also $|\{3, 2, 3, 1\}| = 3$.
- (3) Es ist $|\{\emptyset, 1, \{1\}, \{1, 2\}, \{2, 1\}\}| = 4$.
- (4) Es ist $\mathbb{N} = \{1, 2, 3, \dots\}$ die Menge der natürlichen Zahlen. Diese Menge ist unendlich.
- (5) Es ist $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ die Menge der ganzen Zahlen.
- (6) Es ist \mathbb{R} die Menge der reellen Zahlen, d.h. der abbrechenden oder nichtabbrechenden Dezimalbrüche.
- (7) Es ist \mathbb{Q} die Menge der rationalen Zahlen, d.h. der abbrechenden oder periodischen Dezimalbrüche.

Definition. Seien M und N Mengen.

Gilt für jedes $x \in M$, daß auch $x \in N$ liegt, dann heißt M eine *Teilmenge* von N , geschrieben $M \subseteq N$.

Wir schreiben auch $(M \subset N) :\Leftrightarrow ((M \subseteq N) \wedge (M \neq N))$, gesprochen: M ist eine *echte Teilmenge* von N . Dabei sei “ $:\Leftrightarrow$ ” das Symbol für “ist definitionsgemäß äquivalent zu”.

Beispiel.

- (1) Es ist $\{1, 3\} \subset \{3, 4, 1\}$.
- (2) Es ist $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

(3) Sei M eine Menge. Sei $A(x)$ eine Aussage, die von $x \in M$ abhängt.

Dann ist $\{x \in M : A(x)\} \subseteq M$.

Z.B. ist $\{x \in \mathbb{R} : |x - 1| < 2\} \subseteq \mathbb{R}$.

(4) Seien $a, b \in \mathbb{R}$. Seien wie folgt *Intervalle* als Teilmengen von \mathbb{R} definiert.

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} \\ [a, b[&:= \{x \in \mathbb{R} : a \leq x < b\} \\]a, b] &:= \{x \in \mathbb{R} : a < x \leq b\} \\]a, b[&:= \{x \in \mathbb{R} : a < x < b\} \\]-\infty, b] &:= \{x \in \mathbb{R} : x \leq b\} \\]-\infty, b[&:= \{x \in \mathbb{R} : x < b\} \\ [a, +\infty[&:= \{x \in \mathbb{R} : a \leq x\} \\]a, +\infty[&:= \{x \in \mathbb{R} : a < x\} \\]-\infty, +\infty[&:= \mathbb{R} \end{aligned}$$

In der Literatur wird auch $(a, b]$ anstelle von $]a, b]$ geschrieben, usf.

(5) Sei $a \in \mathbb{R}$. Seien folgende Teilmengen von \mathbb{R} definiert.

$$\begin{aligned} \mathbb{R}_{\geq a} &:= [a, +\infty[\\ \mathbb{R}_{> a} &:=]a, +\infty[\\ \mathbb{R}_{\leq a} &:=]-\infty, a] \\ \mathbb{R}_{< a} &:=]-\infty, a[\end{aligned}$$

Analog für \mathbb{N} , \mathbb{Z} und \mathbb{Q} . Z.B. ist $\mathbb{Z}_{\geq 1} = \mathbb{N}$.

Definition. Sei M eine Menge. Die *Potenzmenge* von M ist definiert als die Menge aller Teilmengen von M , d.h.

$$\text{Pot}(M) := \{X : X \subseteq M\}.$$

Beispiel.

(1) Es ist $\text{Pot}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

(2) Ist M endlich, dann ist $|\text{Pot}(M)| = 2^{|M|}$.

Definition. Seien M und N Mengen.

(1) Sei $M \cap N := \{x : x \in M \wedge x \in N\}$ ihre *Schnittmenge*. Gesprochen: M geschnitten N .

(2) Sei $M \cup N := \{x : x \in M \vee x \in N\}$ ihre *Vereinigungsmenge*. Gesprochen: M vereinigt N .

(3) Sei $M \setminus N := \{x : x \in M \wedge x \notin N\}$ ihre *Differenzmenge*. Gesprochen: M ohne N .

Beispiel.

Sei $M := \{1, 2, 3\}$ und $N := \{3, 4\}$.

Es ist $M \cap N = \{3\}$. Es ist $M \cup N = \{1, 2, 3, 4\}$. Es ist $M \setminus N = \{1, 2\}$.

Definition. Sei $k \geq 0$. Sei M_i eine Menge für $1 \leq i \leq k$.

Sei $\prod_{i=1}^k M_i := \{(m_1, m_2, \dots, m_k) : m_i \in M_i \text{ für } 1 \leq i \leq k\}$ ihr *kartesisches Produkt*.

Dabei sind zwei Elemente $(m_1, m_2, \dots, m_k), (m'_1, m'_2, \dots, m'_k) \in \prod_{i=1}^k M_i$, *Tupel* genannt, genau dann gleich, wenn $m_i = m'_i$ ist für $1 \leq i \leq k$.

Ist $k \geq 1$, so schreiben wir auch $M_1 \times M_2 \times \dots \times M_k := \prod_{i=1}^k M_i$.

Sei M eine Menge. Wir schreiben auch $M^k := \prod_{i=1}^k M$.

Beispiel.

- (1) Sei M eine Menge. Es ist $M^0 = \{()\}$, also $|M^0| = 1$.
- (2) Es ist $\{1, 3\} \times \{5, 7, 8\} = \{(1, 5), (1, 7), (1, 8), (3, 5), (3, 7), (3, 8)\}$.
- (3) Sind M und N endlich, dann ist $|M \times N| = |M| \cdot |N|$.
- (4) Es kann \mathbb{R}^2 veranschaulicht werden durch die Punkte einer Ebene. Für $(x, y) \in \mathbb{R}^2$ bezeichnet dabei üblicherweise x die horizontale und y die vertikale Koordinate eines Punktes.

1.3 Quantoren

Definition. Sei M eine Menge. Sei $A(x)$ eine von $x \in M$ abhängige Aussage.

- (1) Es bedeutet $\forall x \in M : A(x)$: “Für alle $x \in M$ gilt die Aussage $A(x)$ ”.
Es heißt “ \forall ” auch *Allquantor*.
Häufig schreiben wir in Worten auch kurz “für” anstelle von “für alle”.
- (2) Es bedeutet $\exists x \in M : A(x)$: “Es gibt ein $x \in M$ so, daß die Aussage $A(x)$ gilt”.
Es heißt “ \exists ” auch *Existenzquantor*.

Bemerkung. Sei M eine Menge. Sei $A(x)$ eine von $x \in M$ abhängige Aussage.

- (1) Es ist $\neg(\forall x \in M : A(x)) \Leftrightarrow (\exists x \in M : \neg A(x))$.
- (2) Es ist $\neg(\exists x \in M : A(x)) \Leftrightarrow (\forall x \in M : \neg A(x))$.

Beispiel.

- (1) Die Aussage $\forall x \in \mathbb{R} : x^2 \geq 0$ gilt.
- (2) Die Aussage $\exists x \in \mathbb{Z} : x^2 = 25$ gilt.
- (3) Die Aussage $\exists x \in \mathbb{Z} : x^2 = 26$ gilt nicht.
Mit anderen Worten, die negierte Aussage $\forall x \in \mathbb{Z} : x^2 \neq 26$ gilt.
- (4) Die Aussage $\forall (a, b) \in \mathbb{R}^2 \exists x \in \mathbb{R} : x^3 + ax + b = 0$ gilt.

1.4 Abbildungen

Definition. Seien M und N Mengen.

Eine *Abbildung* $f : M \rightarrow N$ ist eine Vorschrift, die jedem Element x aus M ein Element $f(x)$ aus N zuordnet.

Es heißt eine Abbildung auch eine *Funktion*.

Wir schreiben auch $f : M \rightarrow N : m \mapsto f(m)$.

Wir schreiben auch

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ m & \mapsto & f(m) . \end{array}$$

Es heißt M der *Definitionsbereich* von f . Es heißt N der *Zielbereich* von f .

Für $M' \subseteq M$ heie $f(M') := \{ f(m') : m' \in M' \} \subseteq N$ das *Bild* von M' unter f .

Für $N' \subseteq N$ heie $f^{-1}(N') := \{ m \in M : f(m) \in N' \} \subseteq M$ das *Urbild* von N' unter f .

Beispiel. Sei die Abbildung $q : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto q(x) := x^2$ betrachtet.

Es ist $q^{-1}(\{4\}) = \{-2, 2\}$. Es ist $q^{-1}(\{-1\}) = \emptyset$.

Definition. Sei M eine Menge. Die Abbildung $\text{id}_M : M \rightarrow M : m \mapsto \text{id}_M(m) := m$ heißt *Identitt* auf M . Wir krzen manchmal auch $\text{id} := \text{id}_M$ ab.

Definition. Sei $f : M \rightarrow N$ eine Abbildung.

- (1) Es heißt f *injektiv*, wenn $|f^{-1}(\{n\})| \leq 1$ ist fr $n \in N$.
- (2) Es heißt f *surjektiv*, wenn $|f^{-1}(\{n\})| \geq 1$ ist fr $n \in N$.

(3) Es heißt f *bijektiv*, wenn $|f^{-1}(\{n\})| = 1$ ist für $n \in N$.

Ist f bijektiv, so definieren wir die *Umkehrabbildung* $f^{-1} : N \rightarrow M$ durch

$$\{f^{-1}(n)\} := f^{-1}(\{n\}) .$$

Es ist dann $f(f^{-1}(n)) = n$ für $n \in N$, da $f^{-1}(n)$ im Urbild von $\{n\}$ liegt.

Es ist dann $f^{-1}(f(m)) = m$ für $m \in M$, da $m \in f^{-1}(\{f(m)\})$ liegt, und in dieser Menge nur $f^{-1}(f(m))$ liegt.

Bemerkung.

(1) *Es ist $f : M \rightarrow N$ genau dann injektiv, wenn für $m, m' \in M$ aus $m \neq m'$ auch $f(m) \neq f(m')$ folgt.*

Denn ein $n \in N$ mit $|f^{-1}(\{n\})| > 1$ existiert genau dann, wenn es zwei verschiedene Elemente von M gibt, die unter f auf dasselbe Element abgebildet werden.

(2) *Es ist $f : M \rightarrow N$ genau dann surjektiv, wenn $f(M) = N$ ist.*

Denn für $n \in N$ ist genau dann $|f^{-1}(\{n\})| \geq 1$, wenn $n \in f(M)$ liegt.

Beispiel. Die Abbildung $q : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ ist weder injektiv noch surjektiv.

Definition. Sei $f : M \rightarrow N$ eine Abbildung.

Seien $M' \subseteq M$ und $N' \subseteq N$ so gegeben, daß $f(M') \subseteq N'$ ist.

Dann ist die *Einschränkung* $f|_{M'}^{N'} : M' \rightarrow N' : m' \mapsto f(m')$ definiert.

Es ist also $f|_{M'}^{N'}(m') = f(m')$ für $m' \in M'$, wobei aber $f|_{M'}^{N'}$ den Definitionsbereich M' und den Zielbereich N' hat.

Wir schreiben auch $f|_{M'} := f|_{M'}^N$.

Falls $f(M) \subseteq N'$ ist, so schreiben wir auch $f|^{N'} := f|_M^{N'}$.

Beispiel. Sei wieder die Abbildung $q : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ betrachtet.

Es ist $q|_{\mathbb{R}_{\geq 0}} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ injektiv, aber nicht surjektiv.

Es ist $q|_{\mathbb{R}_{> 0}}^{\mathbb{R}_{> 0}} : \mathbb{R} \rightarrow \mathbb{R}_{> 0}$ surjektiv, aber nicht injektiv.

Es ist $q|_{\mathbb{R}_{> 0}}^{\mathbb{R}_{> 0}} : \mathbb{R}_{> 0} \rightarrow \mathbb{R}_{> 0}$ bijektiv.

Ihre Umkehrfunktion $w := (q|_{\mathbb{R}_{> 0}}^{\mathbb{R}_{> 0}})^{-1}$ ist gegeben durch $w : \mathbb{R}_{> 0} \rightarrow \mathbb{R}_{> 0} : x \mapsto w(x) := \sqrt{x}$.

Definition. Seien $f : M \rightarrow N$ und $g : N \rightarrow P$ Abbildungen.

Das *Kompositum* von f und g ist die Abbildung

$$\begin{array}{ccc} M & \xrightarrow{g \circ f} & P \\ m & \mapsto & (g \circ f)(m) := g(f(m)) . \end{array}$$

Es wird $g \circ f$ gesprochen als “ g nach f ”. Die Operation (\circ) selbst heißt *Komposition*.

$$M \xrightarrow{f} N \xrightarrow{g} P$$

$$\quad \quad \quad \curvearrowright$$

$$\quad \quad \quad g \circ f$$

Beispiel. Sei $q : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$. Sei $s : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sin(x)$.

Es ist $(q \circ s)(x) = \sin(x)^2$ für $x \in \mathbb{R}$.

Es ist $(s \circ q)(x) = \sin(x^2)$ für $x \in \mathbb{R}$.

Bemerkung. Sei $f : M \rightarrow N$ eine bijektive Abbildung.

Es ist $f \circ f^{-1} = \text{id}_N$. Es ist $f^{-1} \circ f = \text{id}_M$.

1.5 Relationen

Definition. Seien M und N Mengen. Eine *Relation* R zwischen M und N ist eine Teilmenge $R \subseteq M \times N$.

Man schreibt $m R n : \Leftrightarrow (m, n) \in R$ für $m \in M$ und $n \in N$.

Eine Relation $R \subseteq M \times M$ zwischen M und M heißt auch kurz Relation auf M .

Definition. Sei $f : M \rightarrow N$ eine Abbildung. Der *Graph*

$$\Gamma(f) := \{ (m, f(m)) : m \in M \}$$

von f ist eine Relation zwischen M und N .

Man kann eine Abbildung von M nach N auch als eine Relationen R zwischen M und N definieren, bei denen für jedes $m \in M$ genau ein $n \in N$ so existiert, daß $(m, n) \in R$ ist. Bei diesem Vorgehen wird also eine Abbildung mit ihrem Graphen identifiziert.

Definition. Sei M eine Menge. Sei $R \subseteq M \times M$ eine Relation auf M .

- (1) Es heißt R *reflexiv*, falls $m R m$ ist für $m \in M$.
- (2) Es heißt R *symmetrisch*, falls $(m R m' \Leftrightarrow m' R m)$ ist für $m, m' \in M$.
- (3) Es heißt R *identitiv*, falls $((m R m' \wedge m' R m) \Rightarrow m = m')$ ist für $m, m' \in M$.
- (4) Es heißt R *transitiv*, falls $((m R m' \wedge m' R m'') \Rightarrow m R m'')$ ist für $m, m', m'' \in M$.

Identitive Relationen werden auch als antisymmetrisch bezeichnet.

Beispiel. Die Relation $R := \{(1, 2), (2, 1), (2, 3)\} \subseteq \{1, 2, 3\} \times \{1, 2, 3\}$ auf $\{1, 2, 3\}$ ist nicht reflexiv, nicht symmetrisch, nicht identitiv und nicht transitiv.

Definition. Sei M eine Menge.

Eine Relation (\leq) auf M heißt *Ordnungsrelation*, falls sie reflexiv, identitiv und transitiv ist.

Wir schreiben hier in der Regel $(m \leq m') :\Leftrightarrow ((m, m') \in (\leq))$ für $m, m' \in M$, d.h. wir lassen in der Kurzschreibung die Klammern um \leq weg.

Eine Ordnungsrelation (\leq) auf M heißt *linear*, falls $m \leq m' \vee m' \leq m$ gilt für $m, m' \in M$.

Beispiel.

- (1) Sei M eine Menge. Auf $\text{Pot}(M)$ ist (\subseteq) eine Ordnungsrelation. Diese ist nicht linear, falls $|M| \geq 2$ ist.
- (2) Auf \mathbb{R} ist (\leq) eine lineare Ordnungsrelation.

Definition. Sei M eine Menge.

Eine Relation (\sim) auf M heißt *Äquivalenzrelation*, falls sie reflexiv, symmetrisch und transitiv ist.

Wir schreiben hier in der Regel $(m \sim m') :\Leftrightarrow ((m, m') \in (\sim))$ für $m, m' \in M$, d.h. wir lassen in der Kurzschreibung die Klammern um \sim weg.

Beispiel. Sei M eine Menge.

- (1) Die Relation der Gleichheit $(=)$ ist eine Äquivalenzrelation auf M . Sie ist in jeder Äquivalenzrelation auf M enthalten.
- (2) Die Relation $(\sim) := M \times M$ ist eine Äquivalenzrelation auf M . Mit ihr ist $m \sim m'$ für alle $m, m' \in M$.

Beispiel. Sei $m \in \mathbb{Z}$.

Sei $m\mathbb{Z} := \{mx : x \in \mathbb{Z}\}$.

Für $a, b \in \mathbb{Z}$ sei

$$(a \equiv_m b) :\Leftrightarrow (a - b \in m\mathbb{Z}),$$

gesprochen “ a kongruent modulo m zu b ”.

Wir zeigen, daß (\equiv_m) eine Äquivalenzrelation auf \mathbb{Z} ist.

Reflexivität: Für $a \in \mathbb{Z}$ ist $a \equiv_m a$ wegen $a - a = 0 = m \cdot 0 \in m\mathbb{Z}$.

Symmetrie: Für $a, b \in \mathbb{Z}$ mit $a \equiv_m b$ ist $a - b \in m\mathbb{Z}$, also $a - b = m \cdot x$ für ein $x \in \mathbb{Z}$, also $b - a = m \cdot (-x)$, also $b \equiv_m a$.

Transitivität: Für $a, b, c \in \mathbb{Z}$ mit $a \equiv_m b$ und $b \equiv_m c$ ist $a - b = m \cdot x$ für ein $x \in \mathbb{Z}$ und $b - c = m \cdot y$ für ein $y \in \mathbb{Z}$. Folglich ist

$$a - c = a - b + b - c = m \cdot x + m \cdot y = m \cdot (x + y).$$

Es folgt $a \equiv_m c$.

Z.B. ist $8 \equiv_3 2$, aber $\neg(8 \equiv_3 0)$, was auch $8 \not\equiv_3 0$ geschrieben wird.

Beispiel.

Sei $P := \{ (a, b) \in \mathbb{Z} \times \mathbb{Z} : b \neq 0 \}$.

Sei $((a, b) \sim (a', b')) : \Leftrightarrow (a \cdot b' = a' \cdot b)$ für $(a, b), (a', b') \in P$.

Wir zeigen, daß (\sim) eine Äquivalenzrelation auf P ist.

Reflexivität: Für $(a, b) \in P$ gilt $(a, b) \sim (a, b)$ wegen $a \cdot b = a \cdot b$.

Symmetrie: Für $(a, b), (a', b') \in P$ mit $(a, b) \sim (a', b')$ ist auch $(a', b') \sim (a, b)$ wegen $a' \cdot b = a \cdot b'$.

Transitivität: Für $(a, b), (a', b'), (a'', b'') \in P$ mit $(a, b) \sim (a', b')$ und $(a', b') \sim (a'', b'')$ ist

$$a \cdot b' \cdot b'' = a' \cdot b \cdot b'' = b \cdot a' \cdot b'' = b \cdot a'' \cdot b',$$

woraus nach Division durch $b' \neq 0$ sich

$$a \cdot b'' = a'' \cdot b$$

ergibt. Somit ist $(a, b) \sim (a'', b'')$.

Definition. Sei M eine Menge. Sei (\sim) eine Äquivalenzrelation auf M .

Für $m \in M$ sei $[m]_{(\sim)} := \{ x \in M : x \sim m \}$ die Äquivalenzklasse von m .

Oft schreiben wir kurz $[m] := [m]_{(\sim)}$.

Sei $M/(\sim) := \{ [m]_{(\sim)} : m \in M \}$ die Menge der Äquivalenzklassen.

Bemerkung. Sei M eine Menge. Sei (\sim) eine Äquivalenzrelation auf M .

Jedes Element von M liegt in genau einer Äquivalenzklasse.

Sei dazu $m \in M$ gegeben.

Zum einen liegt $m \in [m]$, da (\sim) reflexiv ist.

Sei zum anderen $m \in [x]$ und $m \in [y]$ für gewisse Elemente $x, y \in M$. Wir wollen $[x] \stackrel{!}{=} [y]$ zeigen.

Wir zeigen $[x] \stackrel{!}{\subseteq} [y]$. Sei $z \in [x]$. Wir haben $z \stackrel{!}{\in} [y]$ zu zeigen.

Es ist $z \sim x$. Ferner ist $m \sim x$ und $m \sim y$.

Symmetrie gibt $x \sim m$.

Aus $z \sim x \sim m$ folgt mit Transitivität, daß $z \sim m$ ist.

Aus $z \sim m \sim y$ folgt mit Transitivität, daß $z \sim y$ ist.

Also ist $z \in [y]$.

Nach Vertauschung der Rollen von x und y zeigt man mit demselben Argument, daß $[y] \subseteq [x]$ ist.

Insgesamt ist also $[x] = [y]$.

Beispiel. Auf \mathbb{Z} hat (\equiv_3) die Äquivalenzklassen

$$\begin{aligned} [0] &= \{ 3z + 0 : z \in \mathbb{Z} \} \\ [1] &= \{ 3z + 1 : z \in \mathbb{Z} \} \\ [2] &= \{ 3z + 2 : z \in \mathbb{Z} \}. \end{aligned}$$

Es ist z.B. $[7] = [1]$.

Beispiel. Sei P wie im vorvorigen Beispiel. Wir haben die bijektive Abbildung

$$\begin{aligned} \mathbb{Q} &\xrightarrow{f} P/(\sim) \\ \frac{a}{b} &\mapsto [(a, b)]_{(\sim)}. \end{aligned}$$

In der Tat ist für $(a, b), (a', b') \in P$ genau dann $\frac{a}{b} = \frac{a'}{b'}$, wenn $a \cdot b' = a' \cdot b$ ist, d.h. wenn $(a, b) \sim (a', b')$ ist.

Damit ist das Bild $[(a, b)]_{(\sim)}$ eines Bruchs $\frac{a}{b}$ unabhängig davon, wie der Bruch auf der linken Seite geschrieben wird, so daß die angegebene Vorschrift tatsächlich eine Abbildung liefert.

Ferner folgt aus $[(a, b)]_{(\sim)} = [(a', b')]_{(\sim)}$ auch $\frac{a}{b} = \frac{a'}{b'}$. Also ist f injektiv.

Nach Konstruktion ist f surjektiv.

Definition.

Sei M eine Menge. Sei R eine Relation auf M .

Sei die Relation S auf M für $x, y \in M$ definiert wie folgt.

$$xSy \quad :\Leftrightarrow \quad (xRy \vee yRx \vee x = y)$$

Dann ist S reflexiv und symmetrisch.

Sei die Relation (\sim) auf M für $x, y \in M$ definiert wie folgt.

$$x \sim y \quad :\Leftrightarrow \quad \left(\begin{array}{l} \text{es gibt ein } k \in \mathbb{Z}_{\geq 0} \text{ und Elemente } z_i \in M \text{ für } 0 \leq i \leq k \text{ mit} \\ x = z_0 S z_1 S z_2 S \dots S z_k = y \end{array} \right)$$

Dann ist (\sim) eine Äquivalenzrelation, genannt die von R erzeugte Äquivalenzrelation.

Es ist $R \subseteq S \subseteq (\sim) \subseteq M \times M$.

Bemerkung. Sei M eine Menge. Sei R eine Relation auf M . Seien S und (\sim) gebildet wie in voriger Definition. Es ist also insbesondere (\sim) die von R erzeugte Äquivalenzrelation.

Ist $(\hat{\sim})$ eine Äquivalenzrelation auf M mit $R \subseteq (\hat{\sim})$, dann ist $(\sim) \subseteq (\hat{\sim})$.

Zeigen wir zunächst $S \stackrel{!}{\subseteq} (\sim)$. Seien dazu $x, y \in M$ mit xSy gegeben. Also gilt $xRy \vee yRx \vee x = y$. Also gilt $x \sim y \vee y \sim x \vee x = y$. Also gilt $x \sim y$.

Zeigen wir nun $(\sim) \stackrel{!}{\subseteq} (\hat{\sim})$. Seien dazu $x, y \in M$ mit $x \sim y$ gegeben. Dann gibt es eine Kette $x = z_0Sz_1Sz_2S \dots Sz_k = y$. Also ist $x = z_0 \hat{\sim} z_1 \hat{\sim} z_2 \hat{\sim} \dots \hat{\sim} z_k = y$. Somit ist $x \hat{\sim} y$.

Beispiel. Sei auf der Menge $\{1, 2, 3, 4, 5\}$ die Relation $R := \{(1, 3), (2, 4), (3, 5), (5, 5)\}$ gegeben. Als Tafel:

R	1	2	3	4	5
1			×		
2				×	
3					×
4					
5					×

Wir bilden wie oben die Relation

$$S = \{(1, 1), (1, 3), (2, 2), (2, 4), (3, 1), (3, 3), (3, 5), (4, 2), (4, 4), (5, 3), (5, 5)\}.$$

Als Tafel:

S	1	2	3	4	5
1	×		×		
2		×		×	
3	×		×		×
4		×		×	
5			×		×

Für die von R erzeugte Äquivalenzrelation ist nun $S \subseteq (\sim)$. Wegen $1 \sim 3 \sim 5$ ist auch $1 \sim 5$ und $5 \sim 1$.

Nun ist aber

$(\hat{\sim})$	1	2	3	4	5
1	×		×		×
2		×		×	
3	×		×		×
4		×		×	
5	×		×		×

eine Äquivalenzrelation, nämlich die mit den Äquivalenzklassen $\{1, 3, 5\}$ und $\{2, 4\}$.

Es ist $(\hat{\sim})$ in (\sim) enthalten, denn wir haben zu $S \subseteq (\sim)$ ja nur $(1, 5)$ und $(5, 1)$ dazugenommen, welche auch in (\sim) liegen.

Umgekehrt ist aber $(\hat{\sim})$ auch eine Äquivalenzrelation, die R enthält; auch dies nach Konstruktion. Gemäß Bemerkung ist also auch $(\sim) \subseteq (\hat{\sim})$.

Es ist also $(\hat{\sim}) = (\sim)$ die von R erzeugte Äquivalenzrelation.

1.6 Induktion

Sei M eine Menge. Sei (\leq) eine Ordnungsrelation auf M .

Wir schreiben $x < y \iff (x \leq y \wedge x \neq y)$ für $x, y \in M$.

Definition.

(1) Sei $X \subseteq M$. Sei $x \in X$.

Es heißt x ein *minimales* Element von X , falls für alle $y \in X$ nicht $y < x$ gilt.

(2) Es heißt M durch (\leq) *induktiv geordnet*, falls jede nichtleere Teilmenge $X \subseteq M$ wenigstens ein minimales Element besitzt.

Bemerkung. *Es ist M induktiv geordnet genau dann, wenn es in M keine Kette von Elementen der Form $z_0 > z_1 > z_2 > \dots$ mit $z_j \in M$ für $j \geq 0$ gibt.*

Denn gibt es eine solche Kette, so bildet diese eine nichtleere Teilmenge ohne minimales Element.

Ist umgekehrt M nicht induktiv geordnet, dann gibt es eine nichtleere Teilmenge $X \subseteq M$ ohne minimales Element. Sei $z_0 \in X$ gewählt.

Es ist z_0 in X nicht minimal. Also gibt es ein $z_1 \in X$ mit $z_0 > z_1$.

Es ist z_1 in X nicht minimal. Also gibt es ein $z_2 \in X$ mit $z_1 > z_2$.

Usf. Also gibt es in X , mithin in M , eine Kette $z_0 > z_1 > z_2 > \dots$

Beispiel.

(1) Ist M endlich, so ist M induktiv geordnet, wie nach Bemerkung folgt.

(2) Sei $k \in \mathbb{Z}$. Es ist $\mathbb{Z}_{\geq k}$ induktiv geordnet, wie nach Bemerkung folgt.

(3) Es ist \mathbb{R} nicht induktiv geordnet. Z.B. hat die nichtleere Teilmenge $]0, 1[$ kein minimales Element.

(4) Es ist \mathbb{Z} nicht induktiv geordnet. Z.B. hat die nichtleere Teilmenge \mathbb{Z} kein minimales Element.

(5) Ordnen wir \mathbb{N}^2 durch $(a, b) \leq (a', b') \iff (a \leq a') \wedge (b \leq b')$ für $(a, b), (a', b') \in \mathbb{N}^2$, dann ist \mathbb{N}^2 dadurch induktiv geordnet. Auch dies folgt nach Bemerkung.

Lemma (allgemeines Induktionsprinzip).

Sei M durch (\leq) induktiv geordnet.

Sei $A(x)$ eine von $x \in M$ abhängige Aussage.

Sei bekannt, daß für $x \in M$ die Aussage $A(x)$ dann gilt, wenn die Aussage $A(z)$ für alle $z \in M$ mit $z < x$ gilt.

Dann gilt $A(x)$ für alle $x \in M$.

Annahme, die Aussage des Lemmas gilt nicht. Dann ist

$$N := \{ y \in M : \neg A(y) \} \neq \emptyset.$$

Da M induktiv geordnet ist, gibt es ein minimales Element $x \in N$.

Sei $z \in M$ mit $z < x$ gegeben. Wegen der Minimalität von x ist $z \in M \setminus N$, und also gilt $A(z)$.

Nach Voraussetzung gilt also $A(x)$. Dies ist ein *Widerspruch* zu $x \in N$.

Lemma (spezielles Induktionsprinzip).

Sei $k \in \mathbb{Z}$.

Sei $A(z)$ eine von $z \in \mathbb{Z}_{\geq k}$ abhängige Aussage.

Es gelte $A(k)$.

Sei zudem bekannt, daß für $z \in \mathbb{Z}_{\geq k+1}$ die Implikation $A(z-1) \Rightarrow A(z)$ gilt.

Dann gilt $A(z)$ für $z \in \mathbb{Z}_{\geq k}$.

Es heißt $A(k)$ auch *Induktionsanfang*. Es heißt die Implikation $A(z-1) \Rightarrow A(z)$ für $z \in \mathbb{Z}_{\geq k+1}$ auch *Induktionsschritt*. Beim Nachweis des Induktionsschritts bezieht man sich auf die Aussage $A(z-1)$ als *Induktionsvoraussetzung*.

Wir begründen dies. Sei daran erinnert, daß $\mathbb{Z}_{\geq k}$ induktiv geordnet ist.

Nach allgemeinem Induktionsprinzip genügt es also sicherzustellen, daß für $x \in \mathbb{Z}_{\geq k}$ aus der Gültigkeit von $A(z)$ für $k \leq z < x$ folgt, daß $A(x)$ gilt.

Fall $x = k$. Es gilt $A(k)$.

Fall $x \geq k+1$. Es gilt $A(z)$ für $k \leq z < x$. Also gilt $A(x-1)$. Dies impliziert $A(x)$.

Beispiel. Sei $n \in \mathbb{Z}_{\geq 1}$. Sei

$$f(n) := 1 + 2 + \dots + n.$$

Wir wollen zeigen, daß für $n \in \mathbb{Z}_{\geq 1}$

$$f(n) \stackrel{!}{=} \frac{1}{2} n(n+1)$$

gilt. Wir schreiben $g(n) := \frac{1}{2} n(n+1)$. Wir haben $f(n) \stackrel{!}{=} g(n)$ für $n \in \mathbb{Z}_{\geq 1}$ zu zeigen.

Wir führen eine Induktion über $\mathbb{Z}_{\geq 1} = \mathbb{N}$, nach dem speziellen Induktionsprinzip.

Induktionsanfang. Es ist $f(1) = 1$. Es ist $g(1) = \frac{1}{2} \cdot 1 \cdot (1 + 1) = 1$.

Induktionsschritt. Sei $n \geq 2$. Als Induktionsvoraussetzung dürfen wir $f(n-1) = g(n-1)$ verwenden. Daraus müssen wir $f(n) \stackrel{!}{=} g(n)$ folgern.

In der Tat wird

$$\begin{aligned} f(n) &= 1 + 2 + \dots + (n-1) + n \\ &= f(n-1) + n \\ &= g(n-1) + n \\ &= \frac{1}{2}(n-1)((n-1) + 1) + n \\ &= \frac{1}{2}(n^2 - n) + n \\ &= \frac{1}{2}(n^2 + n) = g(n). \end{aligned}$$

Beispiel. Es heie $(m, n) \in \mathbb{N}^2$ *teilerfremd*, wenn es kein $k \in \mathbb{Z}_{\geq 2}$ gibt mit $m \equiv_k 0$ und $n \equiv_k 0$.

Da \mathbb{N}^2 wie im obigen Beispiel induktiv geordnet ist, gilt dies auch fur die Teilmenge

$$T := \{ (m, n) \in \mathbb{N}^2 : (m, n) \text{ ist teilerfremd} \}.$$

Wir behaupten, da es fur $(m, n) \in T$ Zahlen $s, t \in \mathbb{Z}$ gibt mit $s \cdot m + t \cdot n = 1$.

Wir fhren eine Induktion ber T , nach dem allgemeinen Induktionsprinzip.

Sei $(m, n) \in T$ gegeben.

Sei die Behauptung bekannt fur $(m', n') \in T$ mit $(m', n') < (m, n)$, d.h. $m' \leq m$ und $n' \leq n$ und $(m', n') \neq (m, n)$. Verwendet man dies, so spricht man auch hier von der Induktionsvoraussetzung.

Wir haben daraus die Behauptung fur (m, n) zu folgern.

Fall $m = n$. Es ist $m = n = 1$ und $1 \cdot m + 0 \cdot n = 1$.

Fall $m < n$. Es ist auch $(m, n-m) \in T$. *Annahme*, nicht. Dann gibt es ein $k \in \mathbb{Z}_{\geq 2}$ mit $m = k \cdot a$ und $n-m = k \cdot b$ fur gewisse $a, b \in \mathbb{Z}$. Also ist $n = (n-m) + m = k \cdot b + k \cdot a = k \cdot (a+b)$. *Widerspruch* zu $(m, n) \in T$.

Nun ist $(m, n-m) < (m, n)$. Mit Induktionsvoraussetzung gibt es $s', t' \in \mathbb{Z}$ mit $s' \cdot m + t' \cdot (n-m) = 1$. Es folgt

$$1 = s' \cdot m + t' \cdot (n-m) = (s' - t') \cdot m + t' \cdot n.$$

Wir knnen also $s := s' - t'$ und $t := t'$ whlen und haben so die Behauptung fur (m, n) gefolgert.

Fall $m > n$. Wie der vorige Fall, nur mit vertauschten Rollen von m und n .

Z.B. ist $(5, 7) \in T$ und in der Tat $3 \cdot 5 + (-2) \cdot 7 = 1$.

1.7 Rekursive Definition

Sei $k \in \mathbb{Z}$. Sei M eine Menge.

Sei $m \in M$.

Sei $u : M \rightarrow M$ eine Abbildung.

Man kann eine Abbildung $f : \mathbb{Z}_{\geq k} \rightarrow M$ dadurch festlegen, daß man als *Rekursionsanfang*

$$f(k) := m$$

und als *Rekursionsschritt* für $z \in \mathbb{Z}_{\geq k+1}$

$$f(z) := u(f(z-1))$$

setzt. Diesemfalls spricht man von einer *rekursiven* Definition der Abbildung f .

Beispiel. Sei $f : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}$ rekursiv definiert durch

$$f(0) := 1$$

und

$$f(z) := 2f(z-1) + 1$$

für $z \geq 1$.

Wir erhalten $f(0) = 1, f(1) = 3, f(2) = 7, f(3) = 15, \dots$

Um zu zeigen, daß $f(z) = 2^{z+1} - 1 =: g(z)$ ist für $z \geq 0$, führen wir eine Induktion nach $z \geq 0$.

Induktionsanfang: $f(0) = 1, g(0) = 2^{0+1} - 1 = 1$, also $f(0) = g(0)$.

Induktionsschritt: Sei $z \geq 1$. Es wird

$$f(z) = 2f(z-1) + 1 \stackrel{\text{IV}}{=} 2g(z-1) + 1 = 2 \cdot (2^{(z-1)+1} - 1) + 1 = 2^{z+1} - 1 = g(z).$$

Variante. Man kann auch $m_1, m_2 \in M$ und eine Abbildung $u : M \times M \rightarrow M$ wählen und f durch $f(k) := m_1, f(k+1) = m_2$ und $f(z) := u(f(z-1), f(z-2))$ für $z \geq k+2$ rekursiv definieren.

Weitere Varianten sind möglich, sofern der zu definierende Funktionswert zu einem Ausdruck in bereits bekannten Funktionswerten gesetzt wird.

Definition. Sei $k \in \mathbb{Z}$. Sei $f : \mathbb{Z}_{\geq k} \rightarrow \mathbb{R}$ eine Abbildung. Wir setzen rekursiv

$$\sum_{j=k}^{k-1} f(j) := 0$$

und

$$\sum_{j=k}^n f(j) := \left(\sum_{j=k}^{n-1} f(j) \right) + f(n)$$

für $n \geq k$.

Dies läuft auf

$$\sum_{j=k}^n f(j) = f(k) + f(k+1) + \dots + f(n)$$

hinaus.

Definition. Sei $k \in \mathbb{Z}$. Sei $f : \mathbb{Z}_{\geq k} \rightarrow \mathbb{R}$ eine Abbildung. Wir setzen rekursiv

$$\prod_{j=k}^{k-1} f(j) := 1$$

und

$$\prod_{j=k}^n f(j) := \left(\prod_{j=k}^{n-1} f(j) \right) \cdot f(n)$$

für $n \geq k$.

Dies läuft auf

$$\prod_{j=k}^n f(j) = f(k) \cdot f(k+1) \cdot \dots \cdot f(n)$$

hinaus.

Beispiel. Wie in §1.6 überprüft, ist

$$\sum_{j=1}^n j = \frac{1}{2} n(n+1)$$

für $n \in \mathbb{Z}_{\geq 1}$.

Beispiel. Für $x \in \mathbb{R}$ und $n \in \mathbb{Z}_{\geq 0}$ kann man $x^n := \prod_{j=1}^n x$ setzen.

Insbesondere ist $x^0 = 1$. Insbesondere ist $0^0 = 1$.

Definition. Sei $n \in \mathbb{Z}_{\geq 0}$.

Sei

$$n! := \prod_{j=1}^n j.$$

Es wird $n!$ als “ n Fakultät” gesprochen.

Wir erhalten $0! = 1$, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, ...

Man kann auch informell $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ schreiben.

Lemma (Binomischer Lehrsatz).

Seien $x, y \in \mathbb{R}$. Sei $n \in \mathbb{Z}_{\geq 0}$. Dann ist

$$\begin{aligned} (x+y)^n &= \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \\ &= \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y^1 + \binom{n}{2} x^{n-2} y^2 + \dots + \binom{n}{n} x^{n-n} y^n. \end{aligned}$$

Dazu führen wir eine Induktion nach $n \in \mathbb{Z}_{\geq 0}$.

Induktionsanfang. Es ist $(x+y)^0 = 1 = \binom{0}{0} x^0 y^0$.

Induktionsschritt. Wir berechnen für $n \in \mathbb{Z}_{\geq 1}$

$$\begin{aligned} (x+y)^n &= (x+y)^{n-1} \cdot (x+y) \\ &\stackrel{\text{IV}}{=} \left(\sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} y^k \right) \cdot (x+y) \\ &= \left(\sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-k} y^k \right) + \left(\sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} y^{k+1} \right) \\ &\stackrel{\ell = k+1}{=} \left(\sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-k} y^k \right) + \left(\sum_{\ell=1}^n \binom{n-1}{\ell-1} x^{n-\ell} y^\ell \right) \\ &\stackrel{\text{umbenennen}}{=} \left(\sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-k} y^k \right) + \left(\sum_{k=1}^n \binom{n-1}{k-1} x^{n-k} y^k \right) \\ &\stackrel{\text{abspalten}}{=} x^n + \left(\sum_{k=1}^{n-1} \binom{n-1}{k} x^{n-k} y^k \right) + \left(\sum_{k=1}^{n-1} \binom{n-1}{k-1} x^{n-k} y^k \right) + y^n \\ &= x^n + \left(\sum_{k=1}^{n-1} \left(\binom{n-1}{k} + \binom{n-1}{k-1} \right) x^{n-k} y^k \right) + y^n \\ &= x^n + \left(\sum_{k=1}^{n-1} \binom{n}{k} x^{n-k} y^k \right) + y^n \\ &= \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k. \end{aligned}$$

Beispiel. Seien $x, y \in \mathbb{R}$.

(1) Es ist $(x+y)^2 = \binom{2}{0} x^{2-0} y^0 + \binom{2}{1} x^{2-1} y^1 + \binom{2}{2} x^{2-2} y^2 = x^2 + 2xy + y^2$.

(2) Es ist

$$\begin{aligned} (x+y)^5 &= \binom{5}{0} x^{5-0} y^0 + \binom{5}{1} x^{5-1} y^1 + \binom{5}{2} x^{5-2} y^2 + \binom{5}{3} x^{5-3} y^3 + \binom{5}{4} x^{5-4} y^4 + \binom{5}{5} x^{5-5} y^5 \\ &= x^5 + 5x^4 y + 10x^3 y^2 + 10x^2 y^3 + 5x y^4 + y^5. \end{aligned}$$

1.9 Euklidischer Algorithmus

Definition. Seien $d, n \in \mathbb{Z}$. Wir schreiben

$$d|n \quad :\Leftrightarrow \quad n \equiv_d 0 \quad \Leftrightarrow \quad n \in d\mathbb{Z},$$

gesprochen “ d teilt n ”.

Bemerkung. Auf \mathbb{N} ist $(|)$ eine Ordnungsrelation.

Lemma. Seien $m, n \in \mathbb{Z}$. Es gibt ein eindeutiges $g \in \mathbb{Z}_{\geq 0}$ mit folgenden Eigenschaften.

- (1) Es ist $g|m$ und $g|n$.
- (2) Es gibt $s, t \in \mathbb{Z}$ mit $g = s \cdot m + t \cdot n$.

Wir schreiben $\text{ggT}(m, n) := g$, genannt *größter gemeinsamer Teiler* von m und n .

Falls $m = n = 0$ ist, dann ist zwingend $g = 0$.

Falls nicht $m = n = 0$ ist, dann argumentieren wir wie folgt.

Existenz: Sei $g := \min(\mathbb{Z}_{\geq 1} \cap \{u \cdot m + v \cdot n : u, v \in \mathbb{Z}\})$. Dann erfüllt g die Eigenschaft (2).

Wir haben (1) zu zeigen. Division mit Rest gibt $m = g \cdot k + r$ mit $k, r \in \mathbb{Z}$ und mit $0 \leq r < g$. Wir haben $r \stackrel{!}{=} 0$ zu zeigen. *Annahme*, es ist $r \geq 1$. Wegen $r = m - g \cdot k$ ist dann auch $r \in \mathbb{Z}_{\geq 1} \cap \{u \cdot m + v \cdot n : u, v \in \mathbb{Z}\}$, im *Widerspruch* zur Minimalität von g . Es folgt $g|m$. Genauso folgt $g|n$.

Eindeutigkeit: Sei $g \in \mathbb{Z}_{\geq 0}$ mit $g|m, g|n$ und $g = s \cdot m + t \cdot n$ gegeben, wobei $s, t \in \mathbb{Z}$. Sei $\tilde{g} \in \mathbb{Z}_{\geq 0}$ mit $\tilde{g}|m, \tilde{g}|n$ und $\tilde{g} = \tilde{s} \cdot m + \tilde{t} \cdot n$ gegeben, wobei $\tilde{s}, \tilde{t} \in \mathbb{Z}$. Wir haben $g \stackrel{!}{=} \tilde{g}$ zu zeigen.

Es sind $g, \tilde{g} \geq 1$, da 0 außer 0 keine Zahl teilt.

Da $g|m$ und $g|n$, gilt auch $g|(\tilde{s} \cdot m + \tilde{t} \cdot n) = \tilde{g}$. Genauso folgt $\tilde{g}|g$. Die Identivität der Ordnungsrelation $(|)$ auf \mathbb{N} gibt also $g = \tilde{g}$.

Bemerkung. Sind $m, n, d \in \mathbb{Z}$ mit $d|m$ und $d|n$ gegeben, dann ist $d | \text{ggT}(m, n)$.

Denn aus $\text{ggT}(m, n) = sm + tn$ für gewisse $s, t \in \mathbb{Z}$ folgt $d | sm + tn = \text{ggT}(m, n)$.

Bemerkung. Für $m, n, z \in \mathbb{Z}$ ist $\text{ggT}(m, n) = \text{ggT}(m, n + zm)$.

Sei $g := \text{ggT}(m, n)$. Zum einen ist $g|m$ und $g|n + zm$. Zum anderen folgt aus $g = sm + tn$ mit $s, t \in \mathbb{Z}$ auch $g = (s - tz)m + t(n + zm)$.

Algorithmus (Euklid). Seien $m, n \in \mathbb{Z}$ mit $0 \leq m < n$ gegeben.

Wir wollen diesenfalls $\text{ggT}(m, n)$ konstruieren.

Wir schreiben $u_1 := n$ und $u_2 := m$.

Sind für $k \geq 1$ bereits u_k und u_{k+1} konstruiert, so verfahren wir wie folgt.

Falls $u_{k+1} = 0$ ist, dann beenden wir den Algorithmus und setzen $\ell := k + 1$.

Falls $u_{k+1} \geq 1$ ist, dann gibt Division mit Rest

$$u_k = u_{k+1} \cdot h_k + u_{k+2}$$

mit $h_k, u_{k+2} \in \mathbb{Z}$ und mit $0 \leq u_{k+2} \leq u_{k+1} - 1$.

Dies gibt ganze Zahlen

$$u_1 > u_2 > u_3 > \dots > u_\ell = 0.$$

Da nun $\text{ggT}(u_k, u_{k+1}) = \text{ggT}(u_{k+1} \cdot h_k + u_{k+2}, u_{k+1}) = \text{ggT}(u_{k+2}, u_{k+1}) = \text{ggT}(u_{k+1}, u_{k+2})$ ist für $1 \leq k \leq \ell - 2$, ist

$$\text{ggT}(m, n) = \text{ggT}(u_1, u_2) = \text{ggT}(u_2, u_3) = \dots = \text{ggT}(u_{\ell-1}, u_\ell) = \text{ggT}(u_{\ell-1}, 0) = u_{\ell-1}.$$

Nun wollen wir noch $s, t \in \mathbb{Z}$ mit $sm + tn = \text{ggT}(m, n)$ bestimmen.

Wir setzen $w_{\ell-1} := 0$ und $w_{\ell-2} := 1$, sowie rekursiv $w_k := w_{k+2} - h_k \cdot w_{k+1}$ für $1 \leq k \leq \ell - 3$.

Es wird

$$w_{k+1} \cdot u_k + w_k \cdot u_{k+1} = w_{k+1} \cdot (u_{k+1} \cdot h_k + u_{k+2}) + (w_{k+2} - h_k \cdot w_{k+1}) \cdot u_{k+1} = w_{k+2} \cdot u_{k+1} + w_{k+1} \cdot u_{k+2}$$

für $1 \leq k \leq \ell - 3$ und insbesondere

$$w_1 \cdot m + w_2 \cdot n = w_1 \cdot u_2 + w_2 \cdot u_1 = \underbrace{w_{\ell-2}}_{=1} \cdot u_{\ell-1} + \underbrace{w_{\ell-1}}_{=0} \cdot u_{\ell-2} = \text{ggT}(m, n).$$

Wir können also $s := w_1$ und $t := w_2$ wählen.

Man vergleiche auch mit dem letzten Beispiel in §1.6.

Beispiel. Wir wollen $\text{ggT}(27, 60)$ mit dem Euklidischen Algorithmus berechnen.

Sei also $u_1 = n := 60$ und $u_2 = m := 27$.

Division mit Rest gibt folgendes.

$$\begin{aligned} \underbrace{60}_{u_1} &= \underbrace{27}_{u_2} \cdot \underbrace{2}_{h_1} + \underbrace{6}_{u_3} \\ \underbrace{27}_{u_2} &= \underbrace{6}_{u_3} \cdot \underbrace{4}_{h_2} + \underbrace{3}_{u_4} \\ \underbrace{6}_{u_3} &= \underbrace{3}_{u_4} \cdot \underbrace{2}_{h_3} + \underbrace{0}_{u_5} \end{aligned}$$

Wir brechen also bei $\ell := 5$ ab und erhalten

$$\text{ggT}(60, 27) = u_{\ell-1} = u_4 = 3.$$

Wir setzen dann $w_4 := 0$, $w_3 := 1$ und

$$\begin{aligned} w_2 &:= w_4 - h_2 \cdot w_3 = 0 - 4 \cdot 1 = -4 \\ w_1 &:= w_3 - h_1 \cdot w_2 = 1 - 2 \cdot (-4) = 9 \end{aligned}$$

Somit wird

$$\text{ggT}(60, 27) = 3 = w_1 \cdot u_2 + w_2 \cdot u_1 = 9 \cdot 27 + (-4) \cdot 60.$$

1.10 Primzahlen

Wir erinnern daran, daß eine Zahl $p \in \mathbb{Z}_{\geq 2}$ Primzahl heißt, wenn $\{d \in \mathbb{Z}_{\geq 1} : d|p\} = \{1, p\}$ ist.

Bemerkung. Seien $a, b \in \mathbb{Z}$. Sei $p|a \cdot b$. Dann ist $p|a \vee p|b$.

Annahme, es ist $\neg(p|a)$ und $\neg(p|b)$. Da $\text{ggT}(p, a)$ ein Teiler von p ist, ist $\text{ggT}(p, a) \in \{1, p\}$. Da $\neg(p|a)$, folgt $\text{ggT}(p, a) = 1$. Somit gibt es $s, t \in \mathbb{Z}$ mit $sa+tp = 1$. Also ist $sab+tpb = b$. Die linke Seite wird von p geteilt, die rechte nicht. *Widerspruch*.

Bemerkung. Sei $n \in \mathbb{Z}_{\geq 2}$. Dann gibt es auf eindeutige Weise ein $k \in \mathbb{N}$ und Primzahlen $p_1 \leq p_2 \leq \dots \leq p_k$ mit

$$n = p_1 \cdot p_2 \cdot \dots \cdot p_k.$$

Existenz. Induktion über $\mathbb{Z}_{\geq 2}$. Falls n prim ist, dann sind wir fertig. Falls n nicht prim ist, dann können wir $n = a \cdot b$ schreiben mit $a, b \in \mathbb{Z}$ und $2 \leq a \leq n-1$ und $2 \leq b \leq n-1$. Nach Induktion können wir a und b als Produkt von Primzahlen schreiben. Also können wir auch n als Produkt von Primzahlen schreiben, und wir können die Faktoren anordnen.

Eindeutigkeit. Seien $k, \ell \in \mathbb{N}$ und Primzahlen $p_1 \leq p_2 \leq \dots \leq p_k$ und $q_1 \leq q_2 \leq \dots \leq q_\ell$ mit

$$n = p_1 \cdot p_2 \cdot \dots \cdot p_k = q_1 \cdot q_2 \cdot \dots \cdot q_\ell$$

gegeben. Division durch alle Faktoren, die auf beiden Seiten auftauchen, gibt

$$p_{u_1} \cdot \dots \cdot p_{u_s} = q_{v_1} \cdot \dots \cdot q_{v_t}$$

mit $s, t \in \mathbb{Z}_{\geq 0}$ mit $1 \leq u_1 < u_2 < \dots < u_s \leq k$ und $1 \leq v_1 < v_2 < \dots < v_t \leq \ell$ und $p_{u_i} \neq q_{v_j}$ für $1 \leq i \leq s$ und $1 \leq j \leq t$.

Wir haben $s \stackrel{!}{=} 0$ und $t \stackrel{!}{=} 0$ zu zeigen.

Annahme, $s \geq 1$.

Dann ist p_{u_1} ein Teiler von $q_{v_1} \cdot \dots \cdot q_{v_t}$, nicht aber von q_{v_1} .

Mit voriger Bemerkung ist p_{u_1} ein Teiler von $q_{v_2} \cdot \dots \cdot q_{v_t}$, nicht aber von q_{v_2} .

Mit voriger Bemerkung ist p_{u_1} ein Teiler von $q_{v_3} \cdot \dots \cdot q_{v_t}$, nicht aber von q_{v_3} .

Usf.

Am Ende erhalten wir, daß p_{u_1} ein Teiler von 1 ist und somit einen *Widerspruch*.

Genauso zeigt man $t = 0$.

Bemerkung. Seien $a, b \in \mathbb{Z}_{\geq 1}$ gegeben.

Wir schreiben $a = p_1^{\alpha_1} \cdot \dots \cdot p_s^{\alpha_s}$ und $b = p_1^{\beta_1} \cdot \dots \cdot p_s^{\beta_s}$ mit $s \in \mathbb{Z}_{\geq 1}$ und $p_1 < \dots < p_s$ prim und $\alpha_i, \beta_i \in \mathbb{Z}_{\geq 0}$ für $1 \leq i \leq s$.

Dann ist $a|b$ äquivalent zu $\alpha_i \leq \beta_i$ für $1 \leq i \leq s$.

Ferner ist $\text{ggT}(a, b) = p_1^{\min(\alpha_1, \beta_1)} \cdot \dots \cdot p_s^{\min(\alpha_s, \beta_s)}$.

Letzteres wird von der zweiten Bemerkung in §1.9 erzwungen.

Beispiel. Es ist $18 = 2^1 \cdot 3^2 \cdot 5^0$. Es ist $15 = 2^0 \cdot 3^1 \cdot 5^1$. Also ist $\text{ggT}(18, 15) = 2^0 \cdot 3^1 \cdot 5^0 = 3$.

1.11 Zifferndarstellung

Sei $d \in \mathbb{Z}_{\geq 2}$. Wir wollen natürliche Zahlen d -adisch schreiben.

Bemerkung. Sei $n \in \mathbb{N}$. Dann gibt es eindeutig bestimmte $k \in \mathbb{Z}_{\geq 0}$ und $a_i \in \mathbb{Z}$ mit $0 \leq a_i \leq d - 1$ für $0 \leq i \leq k$ mit

$$n = a_k \cdot d^k + a_{k-1} \cdot d^{k-1} + \dots + a_0 \cdot d^0.$$

Man schreibt dann auch

$$n = (a_k a_{k-1} \dots a_0)_d.$$

Für $d = 10$ wird in der Regel der Index weggelassen.

Beispiel. Sei $n := 29$.

Es ist $29 = 29_{10}$, in der *Dezimaldarstellung*.

Es ist $29 = 11101_2$, in der *Binärdarstellung*.

Es ist $29 = 1(13)_{16}$.

Damit dies für $d = 16$ besser aussieht, schreibt man hier gern $A := 10$, $B := 11$, $C := 12$, $D := 13$, $E := 14$ und $F := 15$.

Also ist $29 = 1D_{16}$, in der *Hexadezimaldarstellung*.

Kapitel 2

Körper

2.1 Restklassen

Sei $n \in \mathbb{Z}$.

Wir erinnern an die Äquivalenzrelation (\equiv_n) auf \mathbb{Z} aus §1.5, für welche

$$a \equiv_n b \Leftrightarrow a - b \in n\mathbb{Z} = \{nz : z \in \mathbb{Z}\}$$

ist für $a, b \in \mathbb{Z}$.

Definition. Sei $z \in \mathbb{Z}$. Wir schreiben

$$[z]_n := [z]_{\equiv_n} = \{w \in \mathbb{Z} : w \equiv_n z\}$$

für die Äquivalenzklasse von z , genannt *Restklasse* modulo n .

Definition. Sei $\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/(\equiv_n) = \{[z]_n : z \in \mathbb{Z}\}$ die Menge der Restklassen modulo n , gesprochen “ \mathbb{Z} modulo $n\mathbb{Z}$ ”.

Falls $n \in \mathbb{Z}_{\geq 1}$ ist, ist $\mathbb{Z}/n\mathbb{Z} = \{[0]_n, [1]_n, \dots, [n-1]_n\}$.

Beispiel. Es ist $\mathbb{Z}/3\mathbb{Z} = \{[0]_3, [1]_3, [2]_3\}$. Es ist $[7]_3 = [1]_3$.

Definition. Sei

$$\begin{aligned} (+) : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\rightarrow \mathbb{Z}/n\mathbb{Z} \\ ([a]_n, [b]_n) &\mapsto [a]_n + [b]_n := [a+b]_n \end{aligned}$$

und

$$\begin{aligned} (\cdot) : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\rightarrow \mathbb{Z}/n\mathbb{Z} \\ ([a]_n, [b]_n) &\mapsto [a]_n \cdot [b]_n := [a \cdot b]_n \end{aligned}$$

Man beachte, daß das Ergebnis jeweils nicht vom gewählten Element aus der Restklasse abhängt.

Seien nämlich $a, a', b, b' \in \mathbb{Z}$ mit $[a]_n = [a']_n$ und $[b]_n = [b']_n$ gegeben.

Dann ist $a - a' = nz$ und $b - b' = nw$ für gewisse $z, w \in \mathbb{Z}$.

Also ist auch

$$(a + b) - (a' + b') = n(z + w)$$

und also

$$[a + b]_n = [a' + b']_n.$$

Ferner ist auch

$$a \cdot b - a' \cdot b' = a \cdot b - a \cdot b' + a \cdot b' - a' \cdot b' = a \cdot (b - b') + (a - a') \cdot b' = anw + nzb' = n(aw + zb')$$

und also

$$[a \cdot b]_n = [a' \cdot b']_n.$$

Damit ist $[a]_n + [b]_n = [a']_n + [b']_n$ und $[a]_n \cdot [b]_n = [a']_n \cdot [b']_n$ sichergestellt, so daß wir nun im nachhinein gerechtfertigt haben, daß $(+)$ und (\cdot) überhaupt Abbildungen sind, die dem Paar $([a]_n, [b]_n) = ([a']_n, [b']_n)$ ein jeweils eindeutig erklärtes Bild zuordnen.

Beispiel. Es ist $[1]_3 = [7]_3$. Also ist $[7]_3 \cdot [7]_3 = [1]_3 \cdot [1]_3 = [1 \cdot 1]_3 = [1]_3$. Man hätte auch $[7]_3 \cdot [7]_3 = [7 \cdot 7]_3 = [49]_3 = [1]_3$ rechnen können.

Beispiel. Wir erhalten die folgenden Tafeln für $\mathbb{Z}/3\mathbb{Z}$.

$(+)$	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[1]_3$	$[1]_3$	$[2]_3$	$[0]_3$
$[2]_3$	$[2]_3$	$[0]_3$	$[1]_3$

(\cdot)	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[0]_3$	$[0]_3$
$[1]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[2]_3$	$[0]_3$	$[2]_3$	$[1]_3$

Konvention. Wird angekündigt, daß in $\mathbb{Z}/n\mathbb{Z}$ gerechnet wird, so dürfen die eckigen Klammern bei den Restklassen auch entfallen.

Dies ist ein Mißbrauch von Bezeichnungen, der mit Vorsicht zu handhaben ist.

Beispiel. Wir erhalten die folgenden Tafeln für $\mathbb{Z}/3\mathbb{Z}$.

$(+)$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

(\cdot)	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Eigenschaften, die diese Abbildungen $(+)$ und (\cdot) haben, behandeln wir im nächsten Abschnitt.

2.2 Monoide und Gruppen

Sei M eine Menge.

Sei $(*) : M \times M \rightarrow M : (x, y) \mapsto x * y$ eine Abbildung.

Definition. Es heißt $(M, *)$ ein *Monoid*, falls (1, 2) gelten.

- (1) Es ist $(x * y) * z = x * (y * z)$ für $x, y, z \in M$.
- (2) Es gibt ein $e \in M$ mit $x * e = x = e * x$ für $x \in M$.

Eigenschaft (1) heißt *Assoziativität*. Wir schreiben $x * y * z := (x * y) * z = x * (y * z)$.

Das Element e aus (2) heißt *neutrales Element*.

Oft schreiben wir kurz $M := (M, *)$.

Bemerkung. Sind e und e' neutrale Elemente des Monoids M , dann ist $e = e * e' = e'$.

Definition. Ein Monoid $M = (M, *)$ mit neutralem Element e heißt *Gruppe*, falls (3) gilt.

- (3) Für $x \in M$ gibt es ein $y \in M$ mit $x * y = e = y * x$.

Das Element y aus (3) heißt *inverses Element* zu x .

Bemerkung. Sei M eine Gruppe. Sei $x \in M$. Seien y und y' zu x inverse Elemente. Dann ist $y = y * e = y * x * y' = e * y' = y'$.

Definition. Eine Monoid $M = (M, *)$ heißt *abelsch*, falls (4) gilt.

- (4) Für $x, y \in M$ ist $x * y = y * x$.

Eigenschaft (4) heißt *Kommutativität*.

Beispiel.

- (1) Es ist (\mathbb{N}, \cdot) ein abelsches Monoid, aber keine Gruppe. Das neutrale Element ist 1.
- (2) Es ist $(\mathbb{Z}, +)$ eine abelsche Gruppe. Das neutrale Element ist 0.
- (3) Es ist $(\mathbb{Q}, +)$ eine abelsche Gruppe. Das neutrale Element ist 0.
- (4) Es ist (\mathbb{Q}, \cdot) ein abelsches Monoid, aber keine Gruppe. Das neutrale Element ist 1.
- (5) Es ist $(\mathbb{Q} \setminus \{0\}, \cdot)$ eine abelsche Gruppe. Das neutrale Element ist 1.
- (6) Sei $n \in \mathbb{Z}$. Es ist $(\mathbb{Z}/n\mathbb{Z}, +)$ eine abelsche Gruppe. Das neutrale Element ist $0 = [0]_n$.
- (7) Sei $n \in \mathbb{Z}$. Es ist $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ ein abelsches Monoid. Das neutrale Element ist $1 = [1]_n$.

Definition. Sei $n \in \mathbb{N}$. Wir betrachten die Menge $X_n := \{1, 2, \dots, n\}$.

Sei $S_n := \{X_n \xrightarrow{\sigma} X_n : \sigma \text{ ist eine bijektive Abbildung}\}$.

Wir haben die Komposition $(\circ) : S_n \times S_n \rightarrow S_n : (\sigma, \rho) \mapsto \sigma \circ \rho$.

Es ist $S_n = (S_n, \circ)$ eine Gruppe, genannt *symmetrische Gruppe* von Grad n .

Wir schreiben eine bijektive Abbildung $\sigma : X_n \rightarrow X_n$ in der Form $\sigma =: \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$.

So etwa ist $\text{id} := \text{id}_{X_n} = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$ das neutrale Element von S_n .

Es ist $|S_n| = n!$.

Für $1 \leq i, j \leq n$ schreibt man auch (i, j) für die Abbildung, die i und j vertauscht, genannt *Transposition*.

Beispiel. Sei $n = 3$. Es ist

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}.$$

Sei $\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$. Sei $\rho := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$. Es wird

$$\sigma \circ \rho = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

und

$$\rho \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Insbesondere ist $\sigma \circ \rho \neq \rho \circ \sigma$. Also ist S_3 eine nichtabelsche Gruppe.

2.3 Kommutative Ringe und Körper

Sei K eine Menge.

Sei $(+): K \times K \rightarrow K : (x, y) \mapsto x + y$ eine Abbildung, genannt *Addition*.

Sei $(\cdot): K \times K \rightarrow K : (x, y) \mapsto x \cdot y$ eine Abbildung, genannt *Multiplikation*.

Definition. Es heißt $(K, +, \cdot)$ ein *kommutativer Ring*, falls (1, 2, 3) gelten.

- (1) Es ist $(K, +)$ eine abelsche Gruppe.
- (2) Es ist (K, \cdot) ein abelsches Monoid.
- (3) Es ist $(x + x') \cdot (y + y') = x \cdot y + x \cdot y' + x' \cdot y + x' \cdot y'$ für $x, x', y, y' \in K$.

Eigenschaft (3) heißt *Distributivität*.

Das neutrale Element bezüglich $(+)$ wird 0_K geschrieben, oder kurz $0 := 0_K$.

Das inverse Element bezüglich $(+)$ zu $x \in K$ wird $-x$ geschrieben. Wir schreiben auch $x - y := x + (-y)$ für $x, y \in K$.

Das neutrale Element bezüglich (\cdot) wird 1_K geschrieben, oder kurz $1 := 1_K$.

Oft schreiben wir kurz $xy := x \cdot y$ für $x, y \in K$.

Oft schreiben wir kurz $K := (K, +, \cdot)$.

Etwas allgemeiner kann man auch Ringe einführen. Bei diesen wird anstelle von (2) nur verlangt, daß (K, \cdot) ein Monoid ist.

Bemerkung. Sei K ein kommutativer Ring. Sei $x \in K$.

- (1) Es ist $0 \cdot x = 0 \cdot x + 0 \cdot x - 0 \cdot x = (0 + 0) \cdot x - 0 \cdot x = 0 \cdot x - 0 \cdot x = 0$.
- (2) Es ist $(-1) \cdot x = (-1) \cdot x + x - x = (-1) \cdot x + 1 \cdot x - x = (-1 + 1) \cdot x - x = 0 \cdot x - x \stackrel{(1)}{=} -x$.

Definition. Sei $K = (K, +, \cdot)$ ein kommutativer Ring.

Es heißt K ein *Körper*, falls (4, 5) gelten.

- (4) Es ist $0_K \neq 1_K$.
- (5) Für $x \in K \setminus \{0\}$ gibt es ein $y \in K$ mit $x \cdot y = 1_K$.

Das inverse Element bezüglich (\cdot) zu $x \in K \setminus \{0\}$ wird x^{-1} geschrieben. Wir schreiben auch $\frac{x}{y} := x \cdot y^{-1}$ für $x, y \in K$ mit $y \neq 0$.

Bemerkung. Ein Körper ist also eine Menge K , zusammen mit einer Addition $(+)$ und einer Multiplikation (\cdot) , derart, daß für Elemente $x, y, z, w \in K$ folgende Regeln gelten.

- Es ist $(x + y) + z = x + (y + z)$.
- Es ist $x + y = y + x$.

- Es ist $x + 0 = x$.
- Es ist $x + (-x) = 0$.
- Es ist $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
- Es ist $x \cdot y = y \cdot x$.
- Es ist $x \cdot 1 = x$.
- Es ist $x \cdot x^{-1} = 1$, falls $x \neq 0$.
- Es ist $(x + y) \cdot (z + w) = x \cdot z + x \cdot w + y \cdot z + y \cdot w$.

Ferner sollte noch $0 \neq 1$ sein.

Kurz, ein Körper ist ein Rechenbereich, in welchem alle vier Grundrechenarten "plus", "minus", "mal" und "geteilt" durchführbar sind, wobei man nicht durch 0 teilen darf.

Es hat sich herausgestellt, daß man die Vektorrechnung über einem beliebigen Körper aufbauen kann, ohne daß dabei zusätzliche Schwierigkeiten entstehen.

Beispiel.

- (1) Es ist \mathbb{Q} ein Körper.
- (2) Es ist \mathbb{R} ein Körper.
- (3) Es ist \mathbb{Z} ein kommutativer Ring, aber kein Körper. In der Tat gibt es z.B. zu 2 kein multiplikativ inverses Element.
- (4) Es ist \mathbb{N} kein kommutativer Ring. In der Tat gibt es in \mathbb{N} kein neutrales Element der Addition, und auch z.B. zu 3 kein additiv inverses Element.
- (5) Sei $n \in \mathbb{Z}$. Es ist $\mathbb{Z}/n\mathbb{Z}$ ein kommutativer Ring.

Bemerkung. Sei K ein Körper. Seien $x, y \in K$. Ist $x \cdot y = 0$, dann ist $x = 0$ oder $y = 0$.

Annahme, es ist $x \neq 0$ und $y \neq 0$. Es folgt $1 = x \cdot y \cdot y^{-1} \cdot x^{-1} = 0 \cdot y^{-1} \cdot x^{-1} = 0$.
Widerspruch.

Lemma. Sei $n \in \mathbb{Z}_{>2}$.

Genau dann ist $\mathbb{Z}/n\mathbb{Z}$ ein Körper, wenn n eine Primzahl ist.

Denn ist n nicht prim, so können wir $n = a \cdot b$ schreiben mit $[a]_n \neq [0]_n$ und $[b]_n \neq [0]_n$, aber $[a]_n \cdot [b]_n = [0]_n$. Gemäß vorstehender Bemerkung ist also $\mathbb{Z}/n\mathbb{Z}$ kein Körper.

Sei umgekehrt n prim. Sei $x \in \mathbb{Z}$ mit $[x]_n \neq [0]_n$ gegeben.

Wir *behaupten*, daß die Abbildung $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z} : [z]_n \mapsto [x]_n \cdot [z]_n$ injektiv ist. Seien $z, z' \in \mathbb{Z}$ mit $[x]_n \cdot [z]_n = [x]_n \cdot [z']_n$ gegeben. Es folgt $n \mid x \cdot z - x \cdot z' = x \cdot (z - z')$. Da n nicht x teilt und da n prim ist, folgt, daß n die Differenz $z - z'$ teilt. Folglich ist $[z]_n = [z']_n$. Dies zeigt die *Behauptung*.

Da $\mathbb{Z}/n\mathbb{Z}$ eine endliche Menge ist, ist f auch surjektiv. Somit gibt es ein Element $y \in \mathbb{Z}$ mit $f([y]_n) = [1]_n$. Mit anderen Worten, es ist $[x]_n \cdot [y]_n = [1]_n$.

Ferner ist $[0]_n \neq [1]_n$.

Bezeichnung. Ist p eine Primzahl, dann schreiben wir

$$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}.$$

(Körper heißt auf englisch field.)

Beispiel.

(1) Es ist $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z} = \{ [0]_3, [1]_3, [2]_3 \} = \{ 0, 1, 2 \}$ ein Körper.

(2) Es ist $\mathbb{Z}/4\mathbb{Z}$ kein Körper. In der Tat ist $[2]_4 \neq [0]_4$, aber $[2]_4 \cdot [2]_4 = [0]_4$.

Bemerkung. Formeln wie $(x + y)^2 = x^2 + 2xy + y^2$, wie $(x + y) \cdot (x - y) = x^2 - y^2$ und auch wie der Binomische Lehrsatz aus §1.8 gelten für x und y aus jedem kommutativen Ring K . Hierbei ist $2 = 1_K + 1_K$ zu lesen, etc.

2.4 Polynome

Sei K ein kommutativer Ring.

Definition. Ein *Polynom* mit Koeffizienten in K in der formalen Variablen X sei eine Summe der Form

$$f(X) = \sum_{j=0}^n a_j X^j,$$

wobei $n \geq 0$ und $a_j \in K$ für $0 \leq j \leq n$.

Wir schreiben auch

$$f(X) = \sum_{j \geq 0} a_j X^j,$$

wobei es ein $n \in \mathbb{Z}_{>0}$ so gebe, daß $a_j = 0$ ist für $j > n$.

Die Elemente a_j heißen *Koeffizienten* von $f(X)$. Zwei Polynome sind genau dann gleich, wenn ihre Koeffizienten bei X^j übereinstimmen für $j \geq 0$.

Sei die Menge aller Polynome mit Koeffizienten in K bezeichnet als

$$K[X] := \{ f(X) : f(X) \text{ ist ein Polynom mit Koeffizienten in } K \}.$$

Strenggenommen ist ein Polynom ein Tupel $(a_j)_{j \in \mathbb{Z}_{\geq 0}}$ mit $a_j \in K$. Mit anderen Worten, eine Abbildung $\mathbb{Z}_{\geq 0} \rightarrow K : j \mapsto a_j$.

Unsere Schreibung als Polynom in einer formalen Variablen X soll die folgenden Additions- und Multiplikationsregeln suggerieren.

Definition. Seien $f(X) = \sum_{j \geq 0} a_j X^j$ und $g(X) = \sum_{j \geq 0} b_j X^j$ in $K[X]$.

Dann sei

$$\begin{aligned} f(X) + g(X) &:= \sum_{j \geq 0} (a_j + b_j) X^j \\ f(X) \cdot g(X) &:= \sum_{\ell \geq 0} \left(\sum_{j=0}^{\ell} a_j \cdot b_{\ell-j} \right) X^{\ell}. \end{aligned}$$

Es ist $(K[X], +, \cdot)$ ein kommutativer Ring.

Beispiel. In $\mathbb{R}[X]$ ist $(X^2 + 2X + 7) \cdot (5X + 3) = 5X^3 + 13X^2 + 41X + 21$.

Der Koeffizient bei X^2 ergab sich hierbei als $1 \cdot 3 + 2 \cdot 5 + 7 \cdot 0 = 13$.

Bemerkung. Sei $f(X) = \sum_{j=0}^n a_j X^j \in K[X]$. Dies liefert die *polynomiale Funktion*

$$\begin{aligned} K &\rightarrow K \\ x &\mapsto f(x) := \sum_{j=0}^n a_j x^j. \end{aligned}$$

Sei z.B. $f(X) := X^2 - X \in \mathbb{F}_2[X]$. Dann ist $f(0) = 0$ und $f(1) = 0$. Die zugehörige polynomiale Funktion ist also konstant null. Aber $f(X)$ ist nicht das Nullpolynom.

Falls dagegen K ein Körper mit unendlich vielen Elementen ist, ist für Polynome $f(X), g(X) \in K[X]$ genau dann $f(X) = g(X)$, wenn die zugehörige polynomiale Abbildungen übereinstimmen, d.h. wenn $f(x) = g(x)$ ist für $x \in K$. Diesenfalls nennt man eine polynomiale Funktion etwas mißbräuchlich auch wieder kurz Polynom.

Definition. Ein Element $t \in K$ mit $f(t) = 0$ heißt *Nullstelle* von $f(X) \in K[X]$.

2.5 Polynome über einem Körper

Sei K ein Körper.

Definition. Sei $f(X) = \sum_{j \geq 0} a_j X^j \in K[X]$.

Ist $f(X) \neq 0$, dann sei

$$\deg(f(X)) := \max \{ j \in \mathbb{Z}_{\geq 0} : a_j \neq 0 \}$$

der *Grad* von $f(X)$ (englisch **degree**).

Zusätzlich setzen wir $\deg(0) := -\infty$. Dabei vereinbaren wir $-\infty < z$ für $z \in \mathbb{Z}_{\geq 0}$, sowie $(-\infty) + z := -\infty$ für $z \in \mathbb{Z}_{\geq 0}$ und $(-\infty) + (-\infty) := -\infty$.

Somit erhalten wir die Abbildung $\deg : K[X] \rightarrow \mathbb{Z}_{\geq 0} \cup \{-\infty\}$.

Sei $f(X) \neq 0$. Sei $d := \deg(f(X))$. Es heißt a_d der *Leitkoeffizient* von $f(X)$. Es heißt $f(X)$ *normiert*, falls $a_d = 1$ ist.

Beispiel. In $\mathbb{R}[X]$ ist $\deg(X + 2X^3 - 1) = 3$.

Bemerkung. Seien $f(X), g(X) \in K[X]$.

Es ist $\deg(f(X) \cdot g(X)) = \deg(f(X)) + \deg(g(X))$.

Es ist $\deg(f(X) + g(X)) \leq \max\{\deg(f(X)), \deg(g(X))\}$.

Definition. Sei $f(X) \in K[X]$ ein normiertes Polynom mit $\deg(f(X)) \geq 1$.

Es heißt $f(X)$ *irreduzibel*, falls aus

$$f(X) = u(X) \cdot v(X)$$

mit $u(X), v(X) \in K[X]$ folgt, daß $\deg(u(X)) = 0$ oder $\deg(v(X)) = 0$ ist.

Beispiel.

- (1) Es ist $X^2 + 1 \in \mathbb{R}[X]$ irreduzibel. Denn ein Teiler von Grad 1 zöge eine Nullstelle von $f(X)$ nach sich, die es nicht gibt.
- (2) Es ist $X^2 - 2 \in \mathbb{R}[X]$ nicht irreduzibel, da $X^2 - 2 = (X + \sqrt{2}) \cdot (X - \sqrt{2})$.
- (3) Es ist $X^2 + X + 1 \in \mathbb{F}_2[X]$ irreduzibel. Denn ein Teiler von Grad 1 zöge eine Nullstelle von $f(X)$ nach sich, die es nicht gibt.
- (4) Es ist $X^4 + 2X^2 + 1 \in \mathbb{R}[X]$ nicht irreduzibel, da $X^4 + 2X^2 + 1 = (X^2 + 1)^2$. Dieses Polynom hat keine Nullstelle in \mathbb{R} . Die Abwesenheit von Nullstellen garantiert bei Polynomen von Grad ≥ 4 also nicht mehr die Irreduzibilität.

Bemerkung (Polynomdivision).

Seien $f(X) = \sum_{j \geq 0} a_j X^j$ und $g(X) = \sum_{j \geq 0} b_j X^j$ aus $K[X]$. Sei dabei $g(X) \neq 0$.

Es gibt Polynome $h(X), r(X) \in K[X]$ mit $\deg(r(X)) < \deg(g(X))$ und mit

$$f(X) = g(X) \cdot h(X) + r(X).$$

Dazu führen wir eine Induktion über $\deg(f(X)) \in \mathbb{Z}_{\geq 0} \cup \{-\infty\}$. Dies ist möglich, da $\mathbb{Z}_{\geq 0} \cup \{-\infty\}$ induktiv geordnet ist.

Fall $\deg(f(X)) < \deg(g(X))$. Wähle $h(X) = 0$ und $r(X) = f(X)$.

Fall $\deg(f(X)) \geq \deg(g(X))$. Sei $d := \deg(f(X))$. Sei $e := \deg(g(X))$. Sei

$$u(X) := f(X) - X^{d-e} \cdot a_d \cdot b_e^{-1} \cdot g(X).$$

Dann ist $\deg(u(X)) < \deg(f(X))$. Mit Induktion können wir also

$$u(X) = k(X) \cdot g(X) + r(X)$$

schreiben, wobei $k(X), r(X) \in K[X]$ und $\deg(r(X)) < \deg(g(X))$ ist. Es wird

$$f(X) = u(X) + X^{d-e} \cdot a_d \cdot b_e^{-1} \cdot g(X) = (k(X) + X^{d-e} \cdot a_d \cdot b_e^{-1}) \cdot g(X) + r(X),$$

wie verlangt.

Beispiel. In der Praxis sollte man den Induktionsschritt tatsächlich iteriert durchführen.

In $\mathbb{R}[X]$ wird:

$$\begin{array}{r}
 X^6 + X + 1 \\
 -(X^6 - X^5 + 2X^4) \\
 \hline
 X^5 - 2X^4 + X + 1 \\
 -(X^5 - X^4 + 2X^3) \\
 \hline
 -X^4 - 2X^3 + X + 1 \\
 -(-X^4 + X^3 - 2X^2) \\
 \hline
 -3X^3 + 2X^2 + X + 1 \\
 -(-3X^3 + 3X^2 - 6X) \\
 \hline
 -X^2 + 7X + 1 \\
 -(-X^2 + X - 2) \\
 \hline
 6X + 3
 \end{array}
 = (X^2 - X + 2) \cdot (X^4 + X^3 - X^2 - 3X - 1) + (6X + 3)$$

Man kann auch nur jeweils den Teil des Polynoms $f(X)$, das geteilt werden soll, übertragen, bei dem die jeweilige Subtraktion tatsächlich einen Effekt hat. Das spart noch etwas Schreibarbeit:

$$\begin{array}{r}
 X^6 + X + 1 \\
 -(X^6 - X^5 + 2X^4) \\
 \hline
 X^5 - 2X^4 \\
 -(X^5 - X^4 + 2X^3) \\
 \hline
 -X^4 - 2X^3 \\
 -(-X^4 + X^3 - 2X^2) \\
 \hline
 -3X^3 + 2X^2 + X \\
 -(-3X^3 + 3X^2 - 6X) \\
 \hline
 -X^2 + 7X + 1 \\
 -(-X^2 + X - 2) \\
 \hline
 6X + 3
 \end{array}
 = (X^2 - X + 2) \cdot (X^4 + X^3 - X^2 - 3X - 1) + (6X + 3)$$

Bemerkung (Abdividieren von Nullstellen).

Sei $f(X) \in K[X]$.

Ist $a \in K$ eine Nullstelle von $f(X)$, dann gibt es ein $h(X) \in K[X]$ mit

$$f(X) = (X - a) \cdot h(X) .$$

Polynomdivision gibt $f(X) = (X - a) \cdot h(X) + r(X)$ mit $\deg(r(X)) < \deg(X - a) = 1$. Also ist $r(X) = c$ für ein $c \in K$. Es folgt $0 = f(a) = (a - a) \cdot h(a) + c = c$. Somit ist $f(X) = (X - a) \cdot h(X)$.

Bemerkung. Ist $f(X) \in K[X]$ normiert mit $\deg(f(X)) \in \{2, 3\}$, dann ist $f(X)$ genau dann irreduzibel, wenn es keine Nullstelle in K hat.

2.6 Konstruktion von Körpererweiterungen

Sei K ein Körper.

Sei $f(X) = \sum_{j=0}^n a_j X^j \in K[X]$ ein irreduzibles Polynom, wobei $n := \deg(f(X)) \in \mathbb{Z}_{\geq 1}$.

Definition. Für $g(X) \in K[X]$ sagen wir, es *teilt* $f(X)$ das Polynom $g(X)$, wenn es $h(X) \in K[X]$ gibt mit $f(X) \cdot h(X) = g(X)$. Wir schreiben diesenfalls $f(X) \mid g(X)$.

Definition. Sei auf $K[X]$ wie folgt eine Äquivalenzrelation ($\equiv_{f(X)}$) definiert.

Für $u(X), v(X) \in K[X]$ sei

$$u(X) \equiv_{f(X)} v(X) \quad :\Leftrightarrow \quad f(X) \mid (u(X) - v(X)) .$$

Sei $[u(X)]_{f(X)}$ die diesbezügliche Äquivalenzklasse von $u(X)$.

Sei

$$K[X]/f(X)K[X] := \{ [u(X)]_{f(X)} : u(X) \in K[X] \}$$

die Menge der Äquivalenzklassen.

Lemma. *Setzen wir*

$$\begin{aligned} [u(X)]_{f(X)} + [v(X)]_{f(X)} &:= [u(X) + v(X)]_{f(X)} \\ [u(X)]_{f(X)} \cdot [v(X)]_{f(X)} &:= [u(X) \cdot v(X)]_{f(X)} \end{aligned}$$

für $u(X), v(X) \in K[X]$, so ist $K[X]/f(X)K[X] = (K[X]/f(X)K[X], +, \cdot)$ ein Körper.

Dies zeigt man ähnlich wie das Lemma in §2.3.

Bemerkung. *Schreibt man*

$$x := [X]_{f(X)} ,$$

so ist $f(x) = f([X]_{f(X)}) = [f(X)]_{f(X)} = 0$, d.h.

$$x^n = -a_{n-1}x^{n-1} - a_{n-2}x^{n-2} - \dots - a_0x^0.$$

Ferner lässt sich jedes Element $K[X]/f(X)K[X]$ eindeutig in der Form

$$b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \dots + b_0x^0$$

mit $b_j \in K$ für $j \in [0, n-1]$ schreiben.

Dies folgt aus Polynomdivision durch $f(X)$.

Bemerkung. Ist K endlich, dann ist $|K[X]/f(X)K[X]| = |K|^n$.

In den folgenden Abschnitten werden vier Beispiele vorgestellt.

2.7 Der Körper \mathbb{C} der komplexen Zahlen

Sei $K = \mathbb{R}$. Wir verwenden das irreduzible Polynom $f(X) = X^2 + 1 \in \mathbb{R}[X]$.

Sei

$$\mathbb{C} := \mathbb{R}[X]/(X^2 + 1)\mathbb{R}[X]$$

der Körper der *komplexen Zahlen*.

Wir schreiben $i := [X]_{X^2+1}$. Somit ist jedes Element von \mathbb{C} eindeutig in der Form

$$a + bi$$

schreibbar, mit $a, b \in \mathbb{R}$. Es ist also

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}.$$

Wegen $f(X) = X^2 + 1$ ist ferner

$$i^2 = -1.$$

Somit können wir Elemente von \mathbb{C} addieren, subtrahieren und multiplizieren. Z.B. ist

$$\begin{aligned} (2 + 5i) + (4 - 3i) &= 6 + 2i \\ (2 + 5i) \cdot (4 - 3i) &= 2 \cdot 4 + 2 \cdot (-3) \cdot i + 5 \cdot 4 \cdot i + 5 \cdot (-3) \cdot \underbrace{i^2}_{=-1} = 23 + 14i. \end{aligned}$$

Sei $a + bi \neq 0$. Es ist

$$\frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 + b^2}.$$

Z.B. ist

$$\frac{1}{2 + 5i} = \frac{2 - 5i}{(2 + 5i)(2 - 5i)} = \frac{2 - 5i}{2^2 + 5^2} = \frac{2 - 5i}{29}.$$

Bemerkung. Zum Rechnen in \mathbb{C} merke man sich

$$\boxed{i^2 = -1}.$$

2.8 Der Körper \mathbb{F}_4 mit 4 Elementen

Sei $K = \mathbb{F}_2 = \{0, 1\}$. Wir verwenden das irreduzible Polynom $f(X) = X^2 + X + 1 \in \mathbb{F}_2[X]$.

Sei

$$\mathbb{F}_4 := \mathbb{F}_2[X]/(X^2 + X + 1)\mathbb{F}_2[X]$$

der Körper mit 4 Elementen.

Wir schreiben $\alpha := [X]_{X^2+X+1}$. Somit ist jedes Element von \mathbb{F}_4 eindeutig in der Form

$$a + b\alpha$$

schreibbar, mit $a, b \in \mathbb{F}_2$. Es ist also

$$\mathbb{F}_4 = \{a + b\alpha : a, b \in \mathbb{F}_2\} = \{0, 1, \alpha, 1 + \alpha\}.$$

Wegen $f(X) = X^2 + X + 1$ ist ferner

$$\alpha^2 = -\alpha - 1 = \alpha + 1,$$

letzteres, da in \mathbb{F}_2 ja $1 = -1$ ist.

Somit können wir Elemente von \mathbb{F}_4 addieren, subtrahieren und multiplizieren. Additions- und Multiplikationstafel haben folgende Gestalt.

(+)	0	1	α	$1 + \alpha$
0	0	1	α	$1 + \alpha$
1	1	0	$1 + \alpha$	α
α	α	$1 + \alpha$	0	1
$1 + \alpha$	$1 + \alpha$	α	1	0

(·)	0	1	α	$1 + \alpha$
0	0	0	0	0
1	0	1	α	$1 + \alpha$
α	0	α	$\alpha + 1$	1
$1 + \alpha$	0	$1 + \alpha$	1	α

Z.B. ist darin

$$\begin{aligned} (1 + \alpha) + \alpha &= 1 + 2\alpha = 1 \\ (1 + \alpha) \cdot \alpha &= \alpha + \underbrace{\alpha^2}_{=1+\alpha} = 1. \end{aligned}$$

Es ist

$$\frac{1}{1 + \alpha} = \alpha \quad \text{und} \quad \frac{1}{\alpha} = \alpha + 1.$$

Bemerkung. Vorsicht, es ist zwar $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$, aber

$$\mathbb{F}_4 \neq \mathbb{Z}/4\mathbb{Z}.$$

In der Tat ist \mathbb{F}_4 ein Körper, nicht aber $\mathbb{Z}/4\mathbb{Z}$.

Bemerkung. Zum Rechnen in \mathbb{F}_4 merke man sich

$$\boxed{0 = 2} \quad \text{und} \quad \boxed{\alpha^2 = \alpha + 1}.$$

Es findet \mathbb{F}_4 Anwendung z.B. in der Codierungstheorie.

2.9 Der Körper \mathbb{F}_8 mit 8 Elementen

Sei $K = \mathbb{F}_2 = \{0, 1\}$. Wir verwenden das irreduzible Polynom $f(X) = X^3 + X + 1 \in \mathbb{F}_2[X]$.

Sei

$$\mathbb{F}_8 := \mathbb{F}_2[X]/(X^3 + X + 1)\mathbb{F}_2[X]$$

der Körper mit 8 Elementen.

Wir schreiben $\beta := [X]_{X^3+X+1}$. Somit ist jedes Element von \mathbb{F}_8 eindeutig in der Form

$$a + b\beta + c\beta^2$$

schreibbar, mit $a, b, c \in \mathbb{F}_2$. Es ist also

$$\mathbb{F}_8 = \{a + b\beta + c\beta^2 : a, b, c \in \mathbb{F}_2\} = \{0, 1, \beta, 1 + \beta, \beta^2, 1 + \beta^2, \beta + \beta^2, 1 + \beta + \beta^2\}.$$

Wegen $f(X) = X^3 + X + 1$ ist ferner

$$\beta^3 = -\beta - 1 = \beta + 1,$$

letzteres, da in \mathbb{F}_2 ja $1 = -1$ ist.

Somit können wir Elemente von \mathbb{F}_8 addieren, subtrahieren und multiplizieren. Z.B. erhalten wir folgende Potenzen von β .

$$\begin{aligned} \beta^0 &= 1 \\ \beta^1 &= \beta \\ \beta^2 &= \beta^2 \\ \beta^3 &= \beta + 1 \\ \beta^4 &= (\beta + 1) \cdot \beta = \beta^2 + \beta \\ \beta^5 &= (\beta^2 + \beta) \cdot \beta = \beta^3 + \beta^2 = \beta^2 + \beta + 1 \\ \beta^6 &= (\beta^2 + \beta + 1) \cdot \beta = \beta^3 + \beta^2 + \beta = \beta^2 + 1 \\ \beta^7 &= (\beta^2 + 1) \cdot \beta = \beta^3 + \beta = 1 \end{aligned}$$

Insbesondere haben wir folgende Inverse.

$$\begin{aligned} \beta^{-1} &= \beta^6 = \beta^2 + 1 \\ (\beta^2)^{-1} &= \beta^5 = \beta^2 + \beta + 1 \\ (\beta + 1)^{-1} &= (\beta^3)^{-1} = \beta^4 = \beta^2 + \beta \end{aligned}$$

Bemerkung. Vorsicht, es ist zwar $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$, aber

$$\mathbb{F}_8 \neq \mathbb{Z}/8\mathbb{Z}.$$

In der Tat ist \mathbb{F}_8 ein Körper, nicht aber $\mathbb{Z}/8\mathbb{Z}$.

Bemerkung. Zum Rechnen in \mathbb{F}_8 merke man sich

$$\boxed{0 = 2} \quad \text{und} \quad \boxed{\beta^3 = \beta + 1}.$$

2.10 Der Körper \mathbb{F}_9 mit 9 Elementen

Sei $K = \mathbb{F}_3 = \{0, 1, 2\} = \{0, 1, -1\}$. Wir verwenden das irreduzible Polynom $f(X) = X^2 + 1 \in \mathbb{F}_3[X]$.

Sei

$$\mathbb{F}_9 := \mathbb{F}_3[X]/(X^2 + 1)\mathbb{F}_3[X]$$

der Körper mit 9 Elementen.

Wir schreiben $\iota := [X]_{X^2+1}$. Somit ist jedes Element von \mathbb{F}_9 eindeutig in der Form

$$a + b\iota$$

schreibbar, mit $a, b \in \mathbb{F}_3$. Es ist also

$$\mathbb{F}_9 = \{a + b\iota : a, b \in \mathbb{F}_3\} = \{0, 1, -1, \iota, \iota + 1, \iota - 1, -\iota, -\iota + 1, -\iota - 1\}.$$

Wegen $f(X) = X^2 + 1$ ist ferner

$$\iota^2 = -1.$$

Somit können wir Elemente von \mathbb{F}_9 addieren, subtrahieren und multiplizieren. Z.B. ist

$$(1 + \iota)^3 = 1 + 3\iota + 3\iota^2 + \iota^3 = 1 - \iota.$$

Sei $a + b\iota \neq 0$. Es ist

$$\frac{1}{a + b\iota} = \frac{a - b\iota}{(a + b\iota)(a - b\iota)} = \frac{a - b\iota}{a^2 + b^2}.$$

Z.B. ist

$$\frac{1}{1 + \iota} = \frac{1 - \iota}{(1 + \iota)(1 - \iota)} = \frac{1 - \iota}{1^2 + 1^2} = \frac{1 - \iota}{-1} = -1 + \iota.$$

Bemerkung. Vorsicht, es ist zwar $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$, aber

$$\mathbb{F}_9 \neq \mathbb{Z}/9\mathbb{Z}.$$

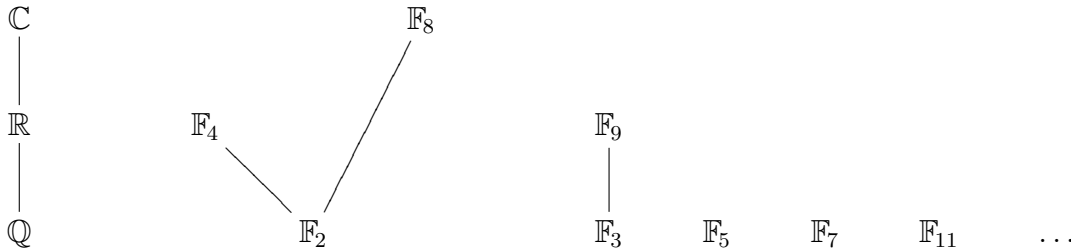
In der Tat ist \mathbb{F}_9 ein Körper, nicht aber $\mathbb{Z}/9\mathbb{Z}$.

Bemerkung. Zum Rechnen in \mathbb{F}_9 merke man sich

$$\boxed{0 = 3} \quad \text{und} \quad \boxed{\iota^2 = -1}.$$

2.11 Körper: Eine Übersicht

Wir haben folgende Körper betrachtet.



Dabei ist $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ für p prim.

Die Striche sollen eine Teilmengenbeziehung andeuten.

So ist $\mathbb{R} = \{a + 0i : a \in \mathbb{R}\} \subseteq \mathbb{C}$. Es entsteht \mathbb{C} aus \mathbb{R} durch Hinzunahme des Elements i mit $i^2 = -1$ und der sich daraus ergebenden Elemente, wie z.B. $5 - 2i$.

Es ist $\mathbb{F}_4 = \{a + 0\alpha : a \in \mathbb{F}_2\} \subseteq \mathbb{F}_8$. Es entsteht \mathbb{F}_4 aus \mathbb{F}_2 durch Hinzunahme des Elements α mit $\alpha^2 = \alpha + 1$ und der sich daraus ergebenden Elemente.

Es ist $\mathbb{F}_8 = \{a + 0\beta + 0\beta^2 : a \in \mathbb{F}_2\} \subseteq \mathbb{F}_8$. Es entsteht \mathbb{F}_8 aus \mathbb{F}_2 durch Hinzunahme des Elements β mit $\beta^3 = \beta + 1$ und der sich daraus ergebenden Elemente.

Es ist $\mathbb{F}_9 = \{a + 0\iota : a \in \mathbb{F}_3\} \subseteq \mathbb{F}_9$. Es entsteht \mathbb{F}_9 aus \mathbb{F}_3 durch Hinzunahme des Elements ι mit $\iota^2 = -1$ und der sich daraus ergebenden Elemente.

2.12 Mehr zu den komplexen Zahlen \mathbb{C}

Definition. Sei $z = a + bi \in \mathbb{C}$, wobei $a, b \in \mathbb{R}$.

(1) Sei $|z| = |a + bi| := \sqrt{a^2 + b^2}$ der *Betrag* von z .

Man beachte, daß auch für $a \in \mathbb{R}$ sich $|a| = |a + 0i| = \sqrt{a^2}$ ergibt.

(2) Sei $\operatorname{Re}(z) = \operatorname{Re}(a + bi) := a$ der *Realteil* von z .

(3) Sei $\operatorname{Im}(z) = \operatorname{Im}(a + bi) := b$ der *Imaginärteil* von z .

(4) Sei $\bar{z} = \overline{a + bi} := a - bi$ die zu z *komplex konjugierte Zahl*.

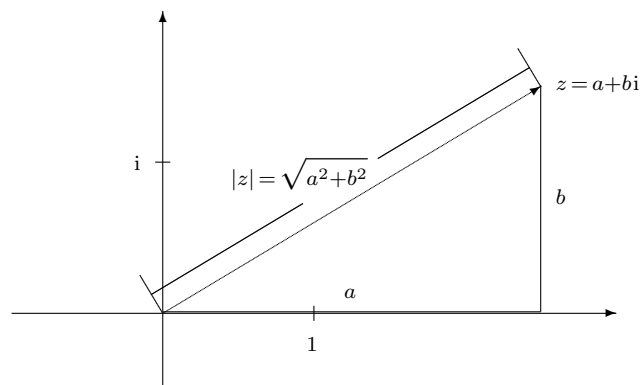
Bemerkung. Es ist $|z|^2 = z \cdot \bar{z}$ für $z \in \mathbb{C}$.

Bemerkung. Es ist $\bar{z} \cdot \bar{w} = \overline{z \cdot w}$ für $z, w \in \mathbb{C}$.

Bemerkung. Es ist $|z \cdot w| = |z| \cdot |w|$ für $z, w \in \mathbb{C}$.

Denn es ist $|z \cdot w| = \sqrt{z \cdot w \cdot \bar{z} \cdot \bar{w}} = \sqrt{(z \cdot \bar{z}) \cdot (w \cdot \bar{w})} = \sqrt{z \cdot \bar{z}} \cdot \sqrt{w \cdot \bar{w}} = |z| \cdot |w|$.

Die komplexe Zahl $z = a + bi$ kann in der *Gaußschen Zahlenebene* als Punkt mit den Koordinaten $(a, b) \in \mathbb{R}^2$ dargestellt werden.

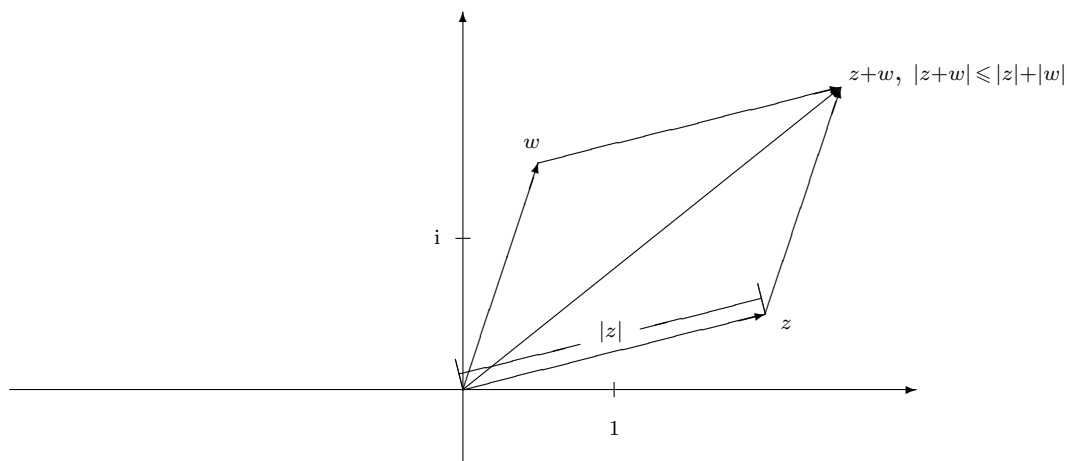


Der Realteil ist dann die horizontale, der Imaginärteil die vertikale Koordinate.

Der Betrag ist der Abstand des Punktes z vom Ursprung 0 .

Addition von komplexen Zahlen $z, w \in \mathbb{C}$ geschieht durch Vektoraddition.

Bemerkung. Für $z, w \in \mathbb{C}$ gilt die *Dreiecksungleichung*: $|z + w| \leq |z| + |w|$.

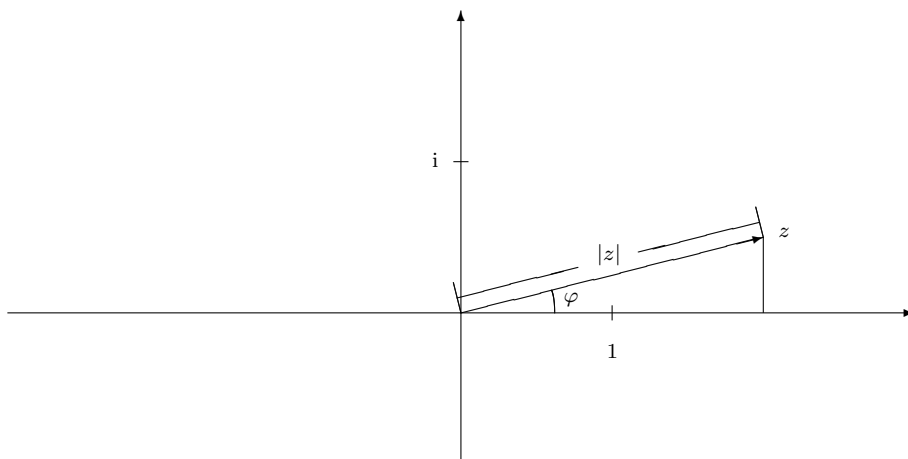


Definition. Sei $z \in \mathbb{C} \setminus \{0\}$. Es gibt genau ein $\varphi \in [0, 2\pi[$ mit

$$z = |z| \cdot (\cos(\varphi) + i \sin(\varphi)) .$$

Es heißt $\arg(z) := \varphi$ das *Argument* von z .

Das Argument von z ist also der Winkel, den die positive reelle Achse mit dem zu z gehörigen Vektor einschließt. Es ist $\operatorname{Re}(z) = |z| \cdot \cos(\varphi)$ und $\operatorname{Im}(z) = |z| \cdot \sin(\varphi)$.



Beispiel. Wir betrachten $z = -1 + i$. Es ist $|-1 + i| = \sqrt{(-1)^2 + 1^2} = \sqrt{2}$.

Es ist

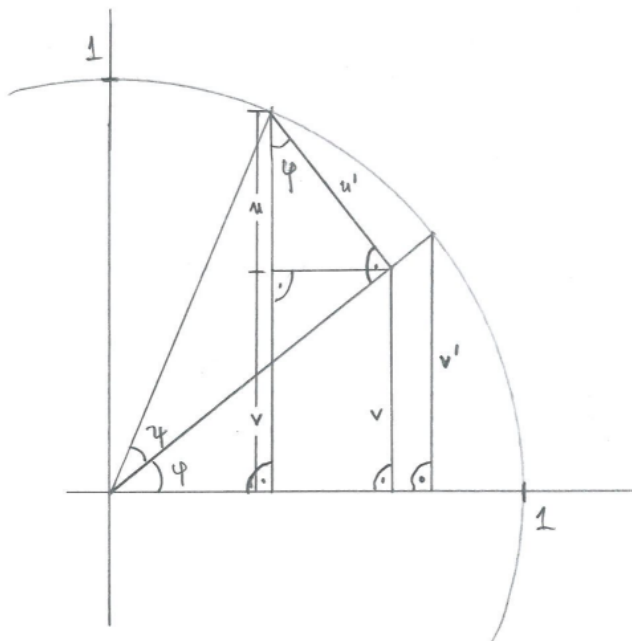
$$-1 + i = \sqrt{2} \left(-\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}} \right) = \sqrt{2} (\cos(3\pi/4) + i \sin(3\pi/4)).$$

Folglich ist $\arg(-1 + i) = 3\pi/4$.

Bemerkung. Seien $\varphi, \psi \in \mathbb{R}$. Es wird

$$\begin{aligned} \sin(\varphi + \psi) &= \sin(\varphi) \cdot \cos(\psi) + \cos(\varphi) \cdot \sin(\psi) \\ \cos(\varphi + \psi) &= \cos(\varphi) \cdot \cos(\psi) - \sin(\varphi) \cdot \sin(\psi). \end{aligned}$$

Wir begründen die erste Gleichheit anhand folgender Skizze.



Es ist $\sin(\varphi + \psi) = v + u$.

Es ist $u' = \sin(\psi)$. Also ist $u = \cos(\varphi) \cdot u' = \cos(\varphi) \cdot \sin(\psi)$.

Es ist $v' = \sin(\varphi)$. Nach Strahlensatz ist also $v = v' \cdot \cos(\psi) = \sin(\varphi) \cdot \cos(\psi)$.

Zusammen ist also $\sin(\varphi + \psi) = v + u = \sin(\varphi) \cdot \cos(\psi) + \cos(\varphi) \cdot \sin(\psi)$.

Die zweite Gleichheit folgt nun dank

$$\begin{aligned} \cos(\varphi + \psi) &= \sin(\pi/2 - \varphi - \psi) \\ &= \sin((\pi/2 - \varphi) + (-\psi)) \\ &= \sin(\pi/2 - \varphi) \cdot \cos(-\psi) + \cos(\pi/2 - \varphi) \cdot \sin(-\psi) \\ &= \cos(\varphi) \cdot \cos(\psi) + \sin(\varphi) \cdot (-\sin(\psi)) . \end{aligned}$$

Später werden wir mit der Formel von de Moivre auch einen rechnerischen Beweis geben können.

Bemerkung. Seien $z, w \in \mathbb{C} \setminus \{0\}$.

Wir schreiben $z = |z| \cdot (\cos(\varphi) + i \sin(\varphi))$ mit $\varphi := \arg(z)$.

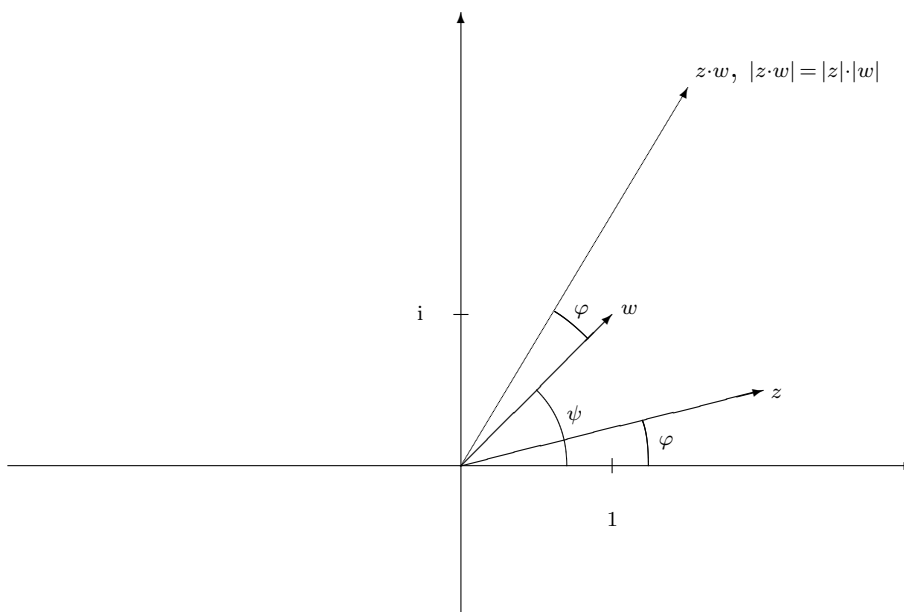
Wir schreiben $w = |w| \cdot (\cos(\psi) + i \sin(\psi))$ mit $\psi := \arg(w)$.

Es wird

$$\begin{aligned} z \cdot w &= |z| \cdot (\cos(\varphi) + i \sin(\varphi)) \cdot |w| \cdot (\cos(\psi) + i \sin(\psi)) \\ &= |z \cdot w| \cdot (\cos(\varphi) \cdot \cos(\psi) - \sin(\varphi) \cdot \sin(\psi) + i \sin(\varphi) \cdot \cos(\psi) + i \cos(\varphi) \cdot \sin(\psi)) \\ &= |z \cdot w| \cdot (\cos(\varphi + \psi) + i \sin(\varphi + \psi)) . \end{aligned}$$

Ist also $\varphi + \psi \in [0, 2\pi[$, dann ist $\arg(z \cdot w) = \arg(z) + \arg(w)$.

Ist hingegen $\varphi + \psi \in [2\pi, 4\pi[$, dann ist $\arg(z \cdot w) = \arg(z) + \arg(w) - 2\pi$.



Insbesondere wird $z^n = (|z| \cdot (\cos(\varphi) + i \sin(\varphi)))^n = |z|^n \cdot (\cos(n\varphi) + i \sin(n\varphi))$ für $n \in \mathbb{Z}$.

Beispiel. Wir betrachten $z = -1 + i$. Gemäß vorigem Beispiel ist

$$-1 + i = \sqrt{2}(\cos(3\pi/4) + i \sin(3\pi/4)).$$

Also wird

$$(-1 + i)^2 = (\sqrt{2})^2(\cos(2 \cdot 3\pi/4) + i \sin(2 \cdot 3\pi/4)) = 2 \cdot (0 + i \cdot (-1)) = -2i.$$

Natürlich hätte man dies auch direkt ausrechnen können:

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

Definition. Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes Polynom $f(X) \in K[X]$ mit $\deg(f(X)) \geq 1$ wenigstens eine Nullstelle in K hat.

Iteriertes Abdividieren von Nullstellen gibt in diesem Fall, daß sich jedes Polynom in $K[X]$ als Produkt von Polynomen von Grad 1 schreiben läßt.

Mit anderen Worten, alle irreduziblen Polynome in $K[X]$ sind dann von der Form $X + a_0$ mit $a_0 \in K$.

Beispiel.

(1) Es ist \mathbb{R} nicht algebraisch abgeschlossen, da z.B. $X^2 + 1$ in \mathbb{R} keine Nullstelle hat.

(2) Es ist \mathbb{C} algebraisch abgeschlossen.

Dies nehmen wir ohne Begründung zur Kenntnis.

(3) Es ist \mathbb{F}_3 nicht algebraisch abgeschlossen, da z.B. $X^3 - X + 1$ in \mathbb{F}_3 keine Nullstelle hat.

Bemerkung. Das Auffinden der Nullstellen eines Polynoms $f(X) \in \mathbb{C}[X]$ ist für kleine Grade noch per Formel möglich, im allgemeinen aber schwierig.

Gewisse Nullstellen kann man aber durch Probieren ermitteln:

Sei $f(X) = \sum_{k=0}^n a_k X^k \in \mathbb{Z}[X]$ von Grad $n \geq 1$ gegeben, mit $a_n \neq 0$ und $a_0 \neq 0$.

Sei $q \in \mathbb{Q}$ eine Nullstelle von $f(X)$. Wir schreiben $q = \frac{u}{v}$ mit teilerfremden Zahlen $u, v \in \mathbb{Z} \setminus \{0\}$.

Dann ist

$$0 = v^n f(q) = v^n \left(\sum_{k=0}^n a_k \left(\frac{u}{v} \right)^k \right) = a_n u^n + \left(\sum_{k=1}^{n-1} a_k u^k v^{n-k} \right) + a_0 v^n.$$

Es folgt, daß v ein Teiler ist von $a_n u^n$ und also von a_n .

Es folgt, daß u ein Teiler ist von a_0v^n und also von a_0 .

Vorsicht, dies schließt auch negative Teiler ein.

Beispiel. Wir wollen $X^3 + \frac{2}{3}X^2 + \frac{11}{3}X - \frac{4}{3}$ in $\mathbb{C}[X]$ in Faktoren von Grad 1 zerlegen.

Zunächst suchen wir die Nullstellen, die in \mathbb{Q} liegen. Diese muß es nicht geben, diese Suche kann also auch erfolglos sein.

Jedenfalls suchen wir dazu die Nullstellen von $3X^3 + 2X^2 + 11X - 4$ der Form $\frac{u}{v}$ mit u und v aus \mathbb{Z} teilerfremd.

Es ist u ein Teiler von -4 , also $-1, +1, -2, +2, -4$ oder $+4$.

Es ist v ein Teiler von 3 , also $-1, +1, -3$ oder $+3$.

Durchprobieren aller 12 Möglichkeiten für $\frac{u}{v}$ gibt die rationale Nullstelle $q = \frac{1}{3}$.

Polynomdivision liefert $X^3 + \frac{2}{3}X^2 + \frac{11}{3}X - \frac{4}{3} = (X - \frac{1}{3})(X^2 + X + 4)$.

Ferner ist $X^2 + X + 4 = (X + \frac{1}{2})^2 + \frac{15}{4} = ((X + \frac{1}{2}) - \frac{i}{2}\sqrt{15}) \cdot ((X + \frac{1}{2}) + \frac{i}{2}\sqrt{15})$.

Zusammen ist also

$$X^3 + \frac{2}{3}X^2 + \frac{11}{3}X - \frac{4}{3} = (X - \frac{1}{3}) \cdot (X + \frac{1}{2} - \frac{i}{2}\sqrt{15}) \cdot (X + \frac{1}{2} + \frac{i}{2}\sqrt{15}).$$

Kapitel 3

Lineare Algebra

Sei K ein Körper.

3.1 Matrizen

Seien $m, n, p, q \geq 0$.

Definition. Eine *Matrix* der Größe $m \times n$ mit Einträgen in K ist eine Tafel

$$A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = (a_{i,j})_{i,j} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n-1} & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & a_{m-1,3} & \cdots & a_{m-1,n-1} & a_{m-1,n} \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n-1} & a_{m,n} \end{pmatrix}$$

mit *Einträgen* $a_{i,j} \in K$. Sie hat m *Zeilen* und n *Spalten*. An *Position* (i, j) von A steht der Eintrag $a_{i,j}$.

Formal ist eine solche Matrix eine Abbildung $\{1, \dots, m\} \times \{1, \dots, n\} \rightarrow K : (i, j) \mapsto a_{i,j}$.

Bei den Indizes merke man sich "Zeilenzähler zuerst, Spaltenzähler später".

Die Menge aller $m \times n$ -Matrizen mit Einträgen in K werde

$$K^{m \times n} := \{ (a_{i,j})_{i,j} : a_{i,j} \in K \text{ für } 1 \leq i \leq m \text{ und } 1 \leq j \leq n \}$$

geschrieben.

Wir identifizieren K mit $K^{1 \times 1}$, schreiben also auch $a_{1,1} = (a_{1,1})$.

Wir identifizieren K^n mit $K^{1 \times n}$, schreiben also auch $(a_{1,1}, \dots, a_{1,n}) = (a_{1,1} \cdots a_{1,n})$.

Beispiel. Es ist $\begin{pmatrix} 7 & -2 & 0 \\ 0 & 1 & -3 \end{pmatrix} \in \mathbb{R}^{2 \times 3}$. Es ist $\begin{pmatrix} 2 & -2+i \\ 0 & 8 \\ i & -3 \end{pmatrix} \in \mathbb{C}^{3 \times 2}$.

Definition. Sei $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$. Sei $A' = (a'_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$.

Sei $B = (b_{j,k})_{1 \leq j \leq n, 1 \leq k \leq p} \in K^{n \times p}$. Sei $s \in K$.

Wir setzen

$$\begin{aligned} A + A' &:= (a_{i,j} + a'_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} && \in K^{m \times n} \\ sA = s \cdot A &:= (s \cdot a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} && \in K^{m \times n} \\ AB = A \cdot B &:= \left(\sum_{j=1}^n a_{i,j} \cdot b_{j,k} \right)_{1 \leq i \leq m, 1 \leq k \leq p} && \in K^{m \times p}. \end{aligned}$$

Sei $0 = 0_{m \times n} := (0)_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in K^{m \times n}$ die *Nullmatrix* der Größe $m \times n$.

Sei $\partial_{i,j} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$ für $1 \leq i \leq n$ und $1 \leq j \leq n$.

Sei

$$E = E_n := (\partial_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

die *Einheitsmatrix* der Größe $n \times n$, wobei leere Einträge als Null gelten sollen.

Für $S \in K^{n \times n}$ setzt man noch $S^0 := E_n$ und $S^k := \underbrace{S \cdot S \cdot \dots \cdot S}_{k \text{ Faktoren}}$ für $k \geq 1$.

Eine Matrix $D = (d_{i,j})_{i,j} \in K^{n \times n}$ heißt *Diagonalmatrix*, wenn $d_{i,j} = 0$ ist für $1 \leq i \leq n$ und $1 \leq j \leq n$ mit $i \neq j$.

Beispiel. Sei $K = \mathbb{R}$.

(1) Es ist $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} + 2 \cdot \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 5 \\ 4 & 8 \end{pmatrix}$.

(2) Es ist $\begin{pmatrix} 3 & 2 & 1 \\ -2 & 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 & 0 & 4 \\ 1 & 1 & 5 & 0 \\ -1 & -4 & -3 & 4 \end{pmatrix} = \begin{pmatrix} -2 & 4 & 7 & 16 \\ 4 & -11 & 16 & 4 \end{pmatrix}$.

(3) Es ist $\begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2$. Es ist $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$.

(4) Es ist $(1 \ 2 \ 3) \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 10$. Es ist $\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} (1 \ 2 \ 3) = \begin{pmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix}$.

(5) Es ist $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$ nicht definiert.

Definition. Ist $A = (a_{i,j})_{i,j} \in K^{m \times n}$, so sei $A^t := (a_{i,j})_{j,i} \in K^{n \times m}$ ihre *Transponierte*.

Beispiel. Für $\begin{pmatrix} 1 & 7 & 3 \\ 2 & 0 & 5 \end{pmatrix} \in \mathbb{R}^{2 \times 3}$ ist $\begin{pmatrix} 1 & 7 & 3 \\ 2 & 0 & 5 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 \\ 7 & 0 \\ 3 & 5 \end{pmatrix} \in \mathbb{R}^{3 \times 2}$.

Bemerkung. Es ist $(K^{m \times n}, +)$ eine abelsche Gruppe.

Seien $A, A' \in K^{m \times n}$. Seien $B, B' \in K^{n \times p}$. Sei $C \in K^{p \times q}$. Seien $\lambda, \lambda', \mu, \mu' \in K$.

(1) *Es ist*

$$\begin{aligned} 1 \cdot A &= A \\ (\lambda + \lambda') \cdot (A + A') &= \lambda \cdot A + \lambda' \cdot A + \lambda \cdot A' + \lambda' \cdot A' \\ (\lambda \cdot A + \lambda' \cdot A') \cdot (\mu \cdot B + \mu' \cdot B') &= \lambda \cdot \mu \cdot A \cdot B + \lambda' \cdot \mu \cdot A' \cdot B + \lambda \cdot \mu' \cdot A \cdot B' + \lambda' \cdot \mu' \cdot A' \cdot B' . \end{aligned}$$

(2) *Es ist*

$$\begin{aligned} A \cdot E_n &= E_m \cdot A &= A \\ (A \cdot B) \cdot C &= A \cdot (B \cdot C) . \end{aligned}$$

Wir schreiben $A \cdot B \cdot C := (A \cdot B) \cdot C = A \cdot (B \cdot C)$.

(3) *Es ist*

$$\begin{aligned} (\lambda \cdot A + \lambda' \cdot A')^t &= \lambda \cdot A^t + \lambda' \cdot A'^t \\ (A \cdot B)^t &= B^t \cdot A^t . \end{aligned}$$

Rechnen wir $A \cdot E_n \stackrel{!}{=} A$ nach. Wir schreiben $A =: (a_{i,j})_{i,j}$. Sei $1 \leq i \leq m$. Sei $1 \leq k \leq n$. An Position (i, k) hat die Matrix $A \cdot E_n$ den Eintrag

$$\sum_{j=1}^n a_{i,j} \cdot \delta_{j,k} = a_{i,k} ,$$

wie dies auch bei der Matrix A der Fall ist.

Rechnen wir $(A \cdot B) \cdot C \stackrel{!}{=} A \cdot (B \cdot C)$ nach. Wir schreiben $A =: (a_{i,j})_{i,j}$, $B =: (b_{j,k})_{j,k}$ und $C =: (c_{k,\ell})_{k,\ell}$.

Sei $1 \leq i \leq m$. Sei $1 \leq \ell \leq q$. An Position (i, ℓ) hat die Matrix $(A \cdot B) \cdot C$ den Eintrag

$$\sum_{k=1}^p \left(\sum_{j=1}^n a_{i,j} \cdot b_{j,k} \right) \cdot c_{k,\ell} = \sum_{k=1}^p \sum_{j=1}^n a_{i,j} \cdot b_{j,k} \cdot c_{k,\ell} .$$

An Position (i, ℓ) hat die Matrix $A \cdot (B \cdot C)$ den Eintrag

$$\sum_{j=1}^n a_{i,j} \cdot \left(\sum_{k=1}^p b_{j,k} \cdot c_{k,\ell} \right) = \sum_{k=1}^p \sum_{j=1}^n a_{i,j} \cdot b_{j,k} \cdot c_{k,\ell} .$$

Das ist dasselbe.

Rechnen wir $(A \cdot B)^t \stackrel{!}{=} B^t \cdot A^t$ nach.

Sei $1 \leq i \leq m$. Sei $1 \leq k \leq p$. An Position (k, i) hat die Matrix $(A \cdot B)^t$ den Eintrag

$$\sum_{j=1}^n a_{i,j} \cdot b_{j,k} .$$

An Position (k, i) hat die Matrix $B^t \cdot A^t$ den Eintrag

$$\sum_{j=1}^n b_{j,k} \cdot a_{i,j} .$$

Das ist dasselbe.

Bemerkung. Sind $A, B \in K^{n \times n}$ von quadratischer Form und ist $A \cdot B = E_n$, dann ist auch $B \cdot A = E_n$.

Diesfalls ist B durch A eindeutig bestimmt, und wir schreiben auch $A^{-1} := B$, genannt die *inverse* Matrix von A .

Zur allgemeinen Berechnung der Inversen siehe letzte Bemerkung in §3.2. Siehe auch das letzte Beispiel in §3.2, Teil (3).

Beispiel. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2}$. Sei $ad - bc \neq 0$.

Wir setzen $B := \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Dann ist

$$A \cdot B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad-bc} \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = E_2.$$

Also ist $B = A^{-1}$. Insbesondere ist auch $B \cdot A = E_2$.

3.2 Gaußverfahren

Seien $m, n \geq 0$.

Definition. Eine Matrix $a \in K^{m \times 1}$ heißt auch ein *Standardvektor* oder kurz ein *Vektor*. Wir schreiben seinen Eintrag an Position $(i, 1)$ oft kurz a_i , für $1 \leq i \leq m$. Also

$$a = (a_i)_{1 \leq i \leq m} = (a_i)_i = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in K^{m \times 1}.$$

Definition. Sei $k \geq 0$. Seien $v_1, \dots, v_k \in K^{m \times 1}$. Wir schreiben kurz

$${}_K \langle v_1, \dots, v_k \rangle := \left\{ \sum_{i=1}^k \lambda_i v_i : \lambda_i \in K \text{ für } 1 \leq i \leq k \right\}$$

für den *Aufspann* oder das *Erzeugnis* von (v_1, \dots, v_k) .

Ist ferner noch $v_0 \in K^{m \times 1}$ gegeben, so schreiben wir kurz

$$v_0 + {}_K \langle v_1, \dots, v_k \rangle := \left\{ v_0 + \sum_{i=1}^k \lambda_i v_i : \lambda_i \in K \text{ für } 1 \leq i \leq k \right\}$$

Problem. Sei $A = (a_{i,j})_{i,j} \in K^{m \times n}$. Sei $b = (b_i)_i \in K^{m \times 1}$. Wir wollen die Lösungsmenge

$$\{x \in K^{n \times 1} : Ax = b\}$$

bestimmen.

Ist $b = 0$, so spricht man von einem *homogenen linearen Gleichungssystem*.

Ist $b \neq 0$, so spricht man von einem *inhomogenen linearen Gleichungssystem*.

Sei allgemein $C = (c_{i,j})_{i,j} \in K^{m \times (n+1)}$ gegeben. Seien $1 \leq k \leq m$ und $1 \leq \ell \leq n$ so, daß $c_{k,\ell} \neq 0$ ist. Wir säubern mit dem Eintrag an Position (k, ℓ) die ℓ -te Spalte von C , indem wir die k -te Zeile mit $c_{k,\ell}^{-1}$ multiplizieren und dann das $c_{i,\ell}$ -fache der entstandenen k -ten Zeile von der i -ten Zeile subtrahieren für alle $1 \leq i \leq m$ mit $i \neq k$. Wir erreichen so, daß die ℓ -te Spalte der umgeformten Matrix bei (k, ℓ) den Eintrag 1 hat, und ansonsten Nullen.

Algorithmus. Es entstehe die Matrix $(A|b) \in K^{m \times (n+1)}$ durch Anfügen von b an A als $(n+1)$ -te Spalte. Wir wollen $(A|b)$ schrittweise umformen. Wir geben hierzu Anweisungen an den Rechner.

Sei $1 \leq \ell_1 \leq n$ minimal so, daß unsere Matrix einen Eintrag ungleich 0 in der ℓ_1 -ten Spalte besitzt. Wähle einen solchen. Vertausche die Zeile mit diesem Eintrag mit der ersten Zeile. Säubere mit dem Eintrag an Position $(1, \ell_1)$ die ℓ_1 -te Spalte.

Sei $1 \leq \ell_2 \leq n$ minimal so, daß unsere Matrix einen Eintrag ungleich 0 in der ℓ_2 -ten Spalte besitzt, der nicht in der ersten Zeile sitzt. Wähle einen solchen. Vertausche die Zeile mit diesem Eintrag mit der zweiten Zeile. Säubere mit dem Eintrag an Position $(2, \ell_2)$ die ℓ_2 -te Spalte.

Sei $1 \leq \ell_3 \leq n$ minimal so, daß unsere Matrix einen Eintrag ungleich 0 in der ℓ_3 -ten Spalte besitzt, der nicht in den Zeilen 1 bis 2 sitzt. Wähle einen solchen. Vertausche die Zeile mit diesem Eintrag mit der dritten Zeile. Säubere mit dem Eintrag an Position $(3, \ell_3)$ die ℓ_3 -te Spalte.

Sei $1 \leq \ell_4 \leq n$ minimal so, daß unsere Matrix einen Eintrag ungleich 0 in der ℓ_4 -ten Spalte besitzt, der nicht in den Zeilen 1 bis 3 sitzt. Wähle einen solchen. Vertausche die Zeile mit diesem Eintrag mit der vierten Zeile. Säubere mit dem Eintrag an Position $(4, \ell_4)$ die ℓ_4 -te Spalte.

Fahre so fort, bis ℓ_{s+1} nicht mehr existiert; was dann der Fall ist, wenn $s+1 > m$ ist oder aber die Zeilen $s+1$ bis m unserer Matrix nur Nulleinträge in den Spalten 1 bis n aufweisen.

Es sind also insgesamt die Spalten ℓ_1, \dots, ℓ_s gesäubert. Numeriere die anderen Spalten aufsteigend durch mit $\ell'_1, \ell'_2, \dots, \ell'_{n-s}$.

Die umgeformte Matrix heiße $(\tilde{A}|\tilde{b})$. Man sagt, \tilde{A} ist in *Zeilenstufenform*. Die *Stufen* von \tilde{A} sind die Einträge 1 in den Spalten ℓ_1, \dots, ℓ_s .

Falls \tilde{b} einen Eintrag ungleich 0 an einer Position $s+1$ bis m hat, dann ist

$$\{x \in K^{n \times 1} : Ax = b\} = \emptyset$$

Falls \tilde{b} nur Nulleinträge an den Positionen $s+1$ bis m hat, dann verfähre man wie folgt.

Man lese das $x_0 \in K^{n \times 1}$ mit Eintrag 0 an Position ℓ'_j für alle $1 \leq j \leq n-s$ ab, das $\tilde{A}x_0 = \tilde{b}$ erfüllt.

Man lese für $1 \leq i \leq n-s$ das $x_i \in K^{n \times 1}$ mit Eintrag 1 an Position ℓ'_i und Eintrag 0 an

Position ℓ'_j für $1 \leq j \leq n - s$ mit $j \neq i$ ab, das $\tilde{A}x_i = 0$ erfüllt.

Dann ist

$$\begin{aligned} \{x \in K^{n \times 1} : Ax = b\} &= \{x_0 + \lambda_1 x_1 + \cdots + \lambda_{n-s} x_{n-s} : \lambda_i \in K \text{ für } 1 \leq i \leq n - s\} \\ &= x_0 + K\langle x_1, \dots, x_{n-s} \rangle. \end{aligned}$$

Ist $b = 0$, so ist auch $x_0 = 0$.

Bemerkung. Eine Matrix $C = (\gamma_{i,j})_{i,j} \in K^{m \times n}$ ist also in Zeilenstufenform, falls es für ein $1 \leq k \leq n$ Stufenspalten mit den Nummern $1 \leq \ell_1 < \ell_2 < \cdots < \ell_k \leq n$ gibt, für welche die Bedingungen (1, 2) gelten.

Sei hierzu noch $\ell_0 := 0$ und $\ell_{k+1} = n + 1$.

(1) Sei $1 \leq i \leq k$, sei also ℓ_i die Nummer einer Stufenspalte.

Es ist $\gamma_{i,\ell_i} = 1$. Es ist $\gamma_{s,\ell_i} = 0$ für $1 \leq s \leq m$ mit $s \neq i$.

(2) Sei $1 \leq j \leq n$ die Nummer einer Nichtstufenspalte, sei also $\ell_i < j < \ell_{i+1}$ für ein $0 \leq i \leq k$. Dann ist $\gamma_{s,j} = 0$ für $i + 1 \leq s \leq m$.

Bemerkung. Sei $A \in K^{m \times n}$. Sei $b \in K^{m \times 1}$.

Es werde $(A|b)$ durch den Algorithmus umgeformt zu $(\tilde{A}|\tilde{b}) \in K^{m \times (n+1)}$. Dann gibt es eine invertierbare Matrix $S \in K^{m \times m}$ mit $S \cdot (A|b) = (\tilde{A}|\tilde{b})$.

Das Produkt zweier invertierbarer Matrizen $S, \tilde{S} \in K^{m \times m}$ ist invertierbar, denn es ist $(S \cdot \tilde{S}) \cdot (\tilde{S}^{-1} \cdot S^{-1}) = E_m$ und $(\tilde{S}^{-1} \cdot S^{-1}) \cdot (S \cdot \tilde{S}) = E_m$.

Also genügt es zu zeigen, daß die Zeilenumformungen in jedem einzelnen Schritt durch Multiplikation mit einer invertierbaren Matrix von links gegeben sind.

Seien $1 \leq k \leq m$ und $1 \leq \ell \leq m$ mit $k \neq \ell$. Wir schreiben $e_{k,\ell} := (\partial_{i,k} \cdot \partial_{j,\ell})_{i,j} \in K^{m \times m}$ für die Matrix mit einem Eintrag 1 an Position (k, ℓ) und ansonsten Nulleinträgen.

Sei $\lambda \in K$. Seien $1 \leq k \leq m$ und $1 \leq \ell \leq m$ mit $k \neq \ell$ gegeben. Die Addition des λ -fachen der ℓ -ten Zeile zur k -ten Zeile ist gegeben durch Multiplikation mit $E_n + \lambda \cdot e_{k,\ell}$ von links. Die Inverse dieser Matrix ist $E_n - \lambda \cdot e_{k,\ell}$.

Sei $\sigma \in S_m$. Eine Vertauschung der Zeilen, bei der die Zeile i in Zeilenposition $\sigma(i)$ vertauscht wird für $1 \leq i \leq m$, ist gegeben durch Multiplikation mit $\sum_{i=1}^m e_{\sigma(i),i}$. Ihr Inverses ist gegeben durch $\sum_{i=1}^m e_{\sigma^{-1}(i),i}$.

Sei $A \in K^{m \times n}$. Die Zeilenstufenform von $A \in K^{m \times n}$ liegt eindeutig fest. Denn seien S, \tilde{S} invertierbare Matrizen mit $S^{-1} \cdot A = B$ und $\tilde{S}^{-1} \cdot A = C$ in Zeilenstufenform. Dann stimmen zunächst die Stufenspaltenpositionen von B und C überein, da diese dadurch gekennzeichnet sind, daß die betreffende Spalte von A nicht im Erzeugnis der vorangegangenen Spalten liegt. Sei r die Anzahl der Stufenspalten in B wie in C . Es ist $S \cdot B = A = \tilde{S} \cdot C$. Also stimmen die ersten r Spalten von S und von \tilde{S} überein. Da die Zeilen $r + 1$ bis m von C Nullzeilen sind, ist auch $S \cdot C = A$. Es folgt $C = S^{-1} \cdot A = B$.

Bemerkung. Sei $A \in K^{n \times n}$.

Dieselbe Umformung wie im Algorithmus, angewandt auf

$$(A | E_n),$$

bei simultanem Mitführen der Spalten im rechten Bereich, gibt

$$(E_n | A^{-1}),$$

falls die Umformung eine solche Matrix im linken Bereich liefert, und ansonsten die Aussage, daß A nicht invertierbar ist.

Denn wird $(E_n | B)$ erreicht mit einer Matrix $B \in K^{n \times n}$, dann gibt es eine invertierbare Matrix $S \in K^{n \times n}$ mit $(S \cdot A | S \cdot E_n) = S \cdot (A | E_n) = (E_n | B)$. Somit ist $S \cdot A = E_n$ und $S = B$, insgesamt also $A^{-1} = S = B$.

Ist umgekehrt E_n links nicht erreichbar, dann gibt es ein $S \in K^{n \times n}$ invertierbar derart, daß $S \cdot A$ nicht invertierbar ist. Dann ist aber auch A nicht invertierbar.

Beispiel.

$$(1) \text{ Sei } A = \begin{pmatrix} 0 & 0 & 2 & 2 \\ 2 & 2 & 4 & 6 \\ 2 & 2 & 3 & 5 \end{pmatrix} \in \mathbb{R}^{3 \times 4}, \text{ sei } b = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \in \mathbb{R}^{3 \times 1}.$$

Wir wollen $\{x \in \mathbb{R}^{4 \times 1} : Ax = b\}$ und zugleich auch noch $\{x \in \mathbb{R}^{4 \times 1} : Ax = 0\}$ bestimmen.

Zunächst bilden wir

$$(A|b) = \left(\begin{array}{cccc|c} 0 & 0 & 2 & 2 & 0 \\ 2 & 2 & 4 & 6 & 2 \\ 2 & 2 & 3 & 5 & 0 \end{array} \right)$$

durch Nebeneinanderstellen.

Es ist $\ell_1 = 1$. Wir säubern mit dem Eintrag an Position $(2, 1)$ die erste Spalte, indem wir die zweite Zeile mit $1/2$ multiplizieren und sodann ihr 2-faches von der dritten Zeile subtrahieren. Schließlich wird noch die zweite Zeile in die erste getauscht. Das gibt:

$$\left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & -1 & -1 & -2 \end{array} \right).$$

Die zweite Spalte muß nun übersprungen werden, da die Einträge an den Positionen $(2, 2)$ und $(3, 2)$ null sind, hier also nicht gesäubert werden kann.

Es wird $\ell_2 = 3$. Wir säubern mit dem Eintrag an Position $(3, 3)$ die dritte Spalte, indem wir die dritte Zeile mit -1 multiplizieren und ihr mit 2-faches von der ersten Zeile und von der zweiten Zeile subtrahieren. Schließlich wird noch die dritte Zeile in die zweite getauscht. Das gibt:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & -4 \end{array} \right) =: (\tilde{A}|\tilde{b}).$$

Die Umformungen sind beendet. Die dabei nicht gesäuberten Spalten von A werden durchnummeriert mit $\ell'_1 := 2$ und $\ell'_2 := 4$.

Wir lesen ab, daß

$$\{x \in \mathbb{R}^{4 \times 1} : Ax = b\} = \{x \in \mathbb{R}^{4 \times 1} : \tilde{A}x = \tilde{b}\} = \emptyset$$

ist, da die dritte Zeile nicht erfüllbar ist.

Ferner lesen wir ab, daß

$$\begin{aligned} \{x \in \mathbb{R}^{4 \times 1} : Ax = 0\} &= \{x \in \mathbb{R}^{4 \times 1} : \tilde{A}x = 0\} \\ &= \left\{ \lambda_1 \begin{pmatrix} -1 \\ \boxed{1} \\ 0 \\ \boxed{0} \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ \boxed{0} \\ -1 \\ \boxed{1} \end{pmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right\} \\ &= \mathbb{R} \left\langle \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\rangle \end{aligned}$$

ist, wobei wir an den Positionen ℓ'_i (also in die Kästchen) jeweils eine 1 und ansonsten Nullen eingesetzt und dann so ergänzt haben, daß jeweils $\tilde{A}x = 0$ ist.

Eine Probe macht man, indem man in die ursprüngliche Gleichung einsetzt.

- (2) Wir suchen alle $q, r, s, t, u, v, w \in \mathbb{R}$, welche das inhomogene lineare Gleichungssystem

$$\begin{array}{rcccccccc} 0q & + & 2r & + & 4s & + & 6t & + & 0u & + & 8v & + & 4w & = & 0 \\ 0q & + & r & + & 2s & + & 3t & + & u & + & 9v & + & 3w & = & 0 \\ 0q & + & r & + & 2s & + & 3t & + & 0u & + & 4v & + & 3w & = & 1 \\ 0q & + & (-1)r & + & (-2)s & + & (-3)t & + & 0u & + & (-4)v & + & (-2)w & = & 0 \end{array}$$

erfüllen. In anderen Worten, wir suchen $\{x := \begin{pmatrix} q \\ r \\ s \\ t \\ u \\ v \\ w \end{pmatrix} \in \mathbb{R}^{7 \times 1} : Ax = b\}$ für

$$(A|b) := \begin{pmatrix} 0 & 2 & 4 & 6 & 0 & 8 & 4 & 0 \\ 0 & 1 & 2 & 3 & 1 & 9 & 3 & 0 \\ 0 & 1 & 2 & 3 & 0 & 4 & 3 & 1 \\ 0 & -1 & -2 & -3 & 0 & -4 & -2 & 0 \end{pmatrix}. \text{ Dazu formen wir wie folgt um.}$$

$$\begin{aligned} (A|b) &\xrightarrow{\text{säubern mit (2,2)}} \begin{pmatrix} 0 & 1 & 2 & 3 & 1 & 9 & 3 & 0 \\ 0 & 0 & 0 & 0 & -2 & -10 & -2 & 0 \\ 0 & 0 & 0 & 0 & -1 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 5 & 1 & 0 \end{pmatrix} \xrightarrow{\text{säubern mit (4,5)}} \begin{pmatrix} 0 & 1 & 2 & 3 & 0 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \\ &\xrightarrow{\text{säubern mit (4,7)}} \begin{pmatrix} 0 & 1 & 2 & 3 & 0 & 4 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 5 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} =: (\tilde{A}|\tilde{b}) \end{aligned}$$

Wir erhalten

$$\begin{aligned} \{x \in \mathbb{R}^{7 \times 1} : Ax = b\} &= \{x \in \mathbb{R}^{7 \times 1} : \tilde{A}x = \tilde{b}\} \\ &= \left\{ \begin{pmatrix} 0 \\ -2 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ -4 \\ 0 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix} : \lambda_i \in \mathbb{R} \text{ für } 1 \leq i \leq 4 \right\} \\ &= \begin{pmatrix} 0 \\ -2 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \mathbb{R} \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -4 \\ 0 \\ 0 \\ -5 \\ 1 \\ 0 \end{pmatrix} \right\rangle. \end{aligned}$$

(3) Für $A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$ wird

$$\left(\begin{array}{cc|cc} 1 & i & 1 & 0 \\ i & -1 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & i & 1 & 0 \\ 0 & 0 & -i & 1 \end{array} \right),$$

und also ist A nicht invertierbar.

Für $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$ wird

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{array} \right),$$

und also $A^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$.

Eine Probe macht man jeweils, indem man $A \cdot A^{-1}$ (oder $A^{-1} \cdot A$) berechnet.

Wer keine Probe macht, ist selbst schuld.

3.3 Vektorräume

Sei K ein Körper.

3.3.1 Definition

Definition. Ein K -Vektorraum besteht aus einer abelschen Gruppe $(V, +)$ und einer Abbildung

$$\begin{aligned} (\cdot) : K \times V &\rightarrow V \\ (\lambda, v) &\mapsto \lambda \cdot v \end{aligned}$$

derart, daß folgende Eigenschaften (V 1–3) gelten.

(V 1) Es ist $1 \cdot v = v$ für $v \in V$.

(V 2) Es ist $(\lambda + \lambda') \cdot (v + v') = \lambda \cdot v + \lambda \cdot v' + \lambda' \cdot v + \lambda' \cdot v'$ für $\lambda, \lambda' \in K$ und $v, v' \in V$.

(V3) Es ist $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$ für $\lambda, \mu \in K$ und $v \in V$.

Wir schreiben oft $\lambda \cdot \mu \cdot v := (\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$ für $\lambda, \mu \in K$ und $v \in V$.

Wir schreiben oft kurz $V := (V, +, \cdot)$. Wir nennen einen K -Vektorraum oft kurz Vektorraum.

Elemente in einem Vektorraum heißen *Vektoren*.

Beispiel.

(1) Seien $m, n \geq 0$. Es ist $K^{m \times n}$, zusammen mit der in §3.1 eingeführten Matrixaddition und der Multiplikation von Elementen aus K mit Matrizen, ein Vektorraum.

Insbesondere ist $K^{m \times 1}$ ein K -Vektorraum, genannt *Standardvektorraum*.

(2) Sei M eine Menge. Sei $K^M := \{f : M \rightarrow K : f \text{ ist eine Abbildung}\}$. Setze

$$\begin{aligned}(f + g)(m) &:= f(m) + g(m) \\ (\lambda \cdot f)(m) &:= \lambda \cdot f(m)\end{aligned}$$

für $f, g \in K^M$, $\lambda \in K$ und $m \in M$. Damit wird K^M zu einem Vektorraum.

Wir fassen so also Abbildungen von M nach K als Vektoren auf.

Bemerkung. Sei V ein Vektorraum, sei $\lambda \in K$ und sei $v \in V$.

Es ist $\lambda \cdot v = 0$ genau dann, wenn $\lambda = 0$ oder $v = 0$.

Es ist $(-1) \cdot v = -v$.

Es ist $0 \cdot v = 0 \cdot v + 0 \cdot v - 0 \cdot v \stackrel{(V2)}{=} (0 + 0) \cdot v - 0 \cdot v = 0 \cdot v - 0 \cdot v = 0$.

Analog zeigt man auch $\lambda \cdot 0 = 0$ für $\lambda \in K$.

Ist umgekehrt $\lambda v = 0$, und ist dabei $\lambda \neq 0$, dann ist $v = \lambda^{-1} \lambda v = 0$.

Es ist $(-1) \cdot v = (-1) \cdot v + 1 \cdot v - 1 \cdot v \stackrel{(V2)}{=} ((-1) + 1) \cdot v - v = 0 \cdot v - v \stackrel{\text{s.o.}}{=} -v$.

3.3.2 Dimension

Sei V ein K -Vektorraum. Sei $n \geq 0$. Sei (v_1, \dots, v_n) ein Tupel aus Vektoren von V .

Definition.

(1) Der *Aufspann* oder das *Erzeugnis* von (v_1, \dots, v_n) ist die Menge

$$\langle v_1, \dots, v_n \rangle = {}_K \langle v_1, \dots, v_n \rangle := \{ \lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_i \in K \text{ für } 1 \leq i \leq n \} \subseteq V$$

aller *Linearkombinationen* dieses Tupels.

- (2) Es heißt das Tupel (v_1, \dots, v_n) *erzeugend* in V , wenn

$$V = {}_K\langle v_1, \dots, v_n \rangle$$

ist.

- (3) Es heißt das Tupel (v_1, \dots, v_n) *linear unabhängig*, wenn mit $\lambda_i \in K$ für $1 \leq i \leq n$ die Gleichung

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0$$

nur für $\lambda_1 = \dots = \lambda_n = 0$ erfüllt ist.

Ansonsten heißt es *linear abhängig*.

- (4) Es heißt das Tupel (v_1, \dots, v_n) eine *K -lineare Basis* oder kurz *Basis* von V , falls es linear unabhängig und erzeugend in V ist.

- (5) Ist (v_1, \dots, v_n) eine Basis von V , so heißt $\dim_K(V) = \dim(V) := n$ die *Dimension* von V .

Hat V eine Basis aus n Vektoren für ein $n \geq 0$, so heißt V *endlichdimensional*.

Hat V für kein $n \geq 0$ eine Basis aus n Vektoren, so heißt V *unendlichdimensional*, und wir schreiben $\dim_K(V) = \dim(V) := \infty$.

Bemerkung.

- (1) *Es ist (v_1, \dots, v_n) genau dann eine Basis von V , wenn für alle $v \in V$ genau ein Tupel $(\lambda_1, \dots, \lambda_n) \in K^n$ existiert mit*

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Zur Eindeutigkeit: Ist $v = \lambda_1 v_1 + \dots + \lambda_n v_n = \lambda'_1 v_1 + \dots + \lambda'_n v_n$ mit $\lambda_i, \lambda'_i \in K$ für $1 \leq i \leq n$, dann ist $\sum_{i=1}^n (\lambda_i - \lambda'_i) v_i = 0$, wegen linearer Unabhängigkeit also $\lambda_i - \lambda'_i = 0$ für $1 \leq i \leq n$.

- (2) Die Dimension von V hängt nicht von der gewählten Basis von V ab. In anderen Worten, sind (v_1, \dots, v_n) und $(v'_1, \dots, v'_{n'})$ Basen von V , dann ist $n = n'$.

Bemerkung. Sei $1 \leq i \leq n$. Wir schreiben $e_i := (\delta_{i,j})_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^{n \times 1}$, wobei die 1 an Position $(i, 1)$ sitzt, für den *i -ten Standardbasisvektor*.

Es ist (e_1, \dots, e_n) eine Basis von $K^{n \times 1}$, die *Standardbasis*. Also ist $\dim_K(K^{n \times 1}) = n$.

Lemma. Sei $m \geq 0$. Sei (v_1, \dots, v_n) ein Tupel von Vektoren in $K^{m \times 1}$.

Habe $A \in K^{m \times n}$ als *i -te Spalte* $v_i \in K^{m \times 1}$ für $1 \leq i \leq n$; d.h. entstehe A durch *Nebeneinanderstellen* der Vektoren v_i .

- (1) Es ist (v_1, \dots, v_n) genau dann ein erzeugendes Tupel in $K^{m \times 1}$, wenn die Zeilenstufenform von A keine Nullzeile enthält.
- (2) Es ist (v_1, \dots, v_n) genau dann linear unabhängig, wenn die Zeilenstufenform von A genau n Nichtnullzeilen enthält, i.e. wenn sie keine Nichtstufenspalten enthält.
- (3) Es ist (v_1, \dots, v_n) genau dann eine Basis von $K^{m \times 1}$, wenn $m = n$ ist und A die Zeilenstufenform E_n hat.

Sei $S \in K^{m \times m}$ invertierbar mit $S \cdot A =: \tilde{A}$ in Zeilenstufenform.

Es ist (v_1, \dots, v_n) erzeugend in $K^{m \times 1}$ genau dann, wenn $(S \cdot v_1, \dots, S \cdot v_n)$ erzeugend in $K^{m \times 1}$ ist. Die Spalten von \tilde{A} bilden genau dann ein erzeugendes Tupel in $K^{m \times 1}$, wenn \tilde{A} keine Nullzeile enthält.

Es ist (v_1, \dots, v_n) linear unabhängig genau dann, wenn $(S \cdot v_1, \dots, S \cdot v_n)$ linear unabhängig ist. Die Spalten von \tilde{A} bilden genau dann ein linear unabhängiges Tupel, wenn \tilde{A} keine Nichtstufenspalte enthält.

Beispiel.

- (1) Sei $K = \mathbb{R}$.

Es ist $\left(\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}\right)$ in $\mathbb{R}^{3 \times 1}$ weder linear unabhängig noch erzeugend.

Denn die zugehörige Matrix wird zur Zeilenstufenform

$$\begin{pmatrix} 2 & 2 & -1 & 2 \\ 1 & -1 & 0 & 2 \\ 1 & 3 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & -1/2 & 1 \\ 0 & -2 & 1/2 & 1 \\ 0 & 2 & -1/2 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1/4 & 3/2 \\ 0 & 1 & -1/4 & -1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

umgeformt, und diese hat 1 (> 0) Nullzeile und 2 (< 4) Nichtnullzeilen.

- (2) Sei $K = \mathbb{R}$.

Es ist $\left(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}\right)$ eine Basis von $\mathbb{R}^{3 \times 1}$, denn die zugehörige Matrix wird zur Zeilenstufenform

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_3$$

umgeformt.

- (3) Ist M eine unendliche Menge, dann ist auch $\dim_K(K^M) = \infty$.
- (4) Es ist $K[X]$ ein K -Vektorraum, für welchen $K \times K[X] : (\lambda, f(X)) \mapsto \lambda \cdot f(X)$ durch Multiplikation innerhalb $K[X]$ gegeben ist, wobei $\lambda \in K$ als Polynom von Grad ≤ 0 verwendet wird.

Es ist $\dim_K(K[X]) = \infty$.

- (5) Es ist \mathbb{C} ein \mathbb{R} -Vektorraum, für welchen $\mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C} : (a, z) \mapsto a \cdot z$ durch Multiplikation innerhalb \mathbb{C} gegeben ist.

Es ist $(1, i)$ eine Basis des \mathbb{R} -Vektorraums \mathbb{C} . Also ist $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

(6) Analog zu (5) haben wir die \mathbb{F}_2 -Vektorräume \mathbb{F}_4 und \mathbb{F}_8 und den \mathbb{F}_3 -Vektorraum \mathbb{F}_9 .

Bemerkung.

(1) Ist (v_1, \dots, v_n) erzeugend in V , dann gibt es ein $k \geq 0$ und $1 \leq i_j \leq n$ für $1 \leq j \leq k$ derart, daß

$$(v_{i_1}, \dots, v_{i_k})$$

eine Basis von V ist.

Ist insbesondere $\dim_K(V) = n$, dann ist das erzeugende Tupel (v_1, \dots, v_n) bereits eine Basis von V .

(2) Ist (v_1, \dots, v_n) linear unabhängig, dann gibt es ein $k \geq 0$ und ein Tupel (w_1, \dots, w_k) von Vektoren in V derart, daß

$$(v_1, \dots, v_n, w_1, \dots, w_k)$$

eine Basis von V ist.

Ist insbesondere $\dim_K(V) = n$, dann ist das linear unabhängige Tupel (v_1, \dots, v_n) bereits eine Basis von V .

3.3.3 Unterräume

Sei V ein K -Vektorraum.

Definition. Ein K -Unterraum oder *Unterraum* von V ist eine Teilmenge $U \subseteq V$ so, daß $0 \in U$ ist und daß für alle $\lambda, \lambda' \in K$ und $u, u' \in U$ auch

$$\lambda u + \lambda' u' \in U$$

ist.

So z.B. sind $0 := \{0\}$ und V Unterräume von V .

Bemerkung. Sei $U \subseteq V$ ein Unterraum. Sei $n \geq 0$ und seien $u_1, \dots, u_n \in U$ und $\lambda_1, \dots, \lambda_n \in K$ gegeben. Dann ist auch

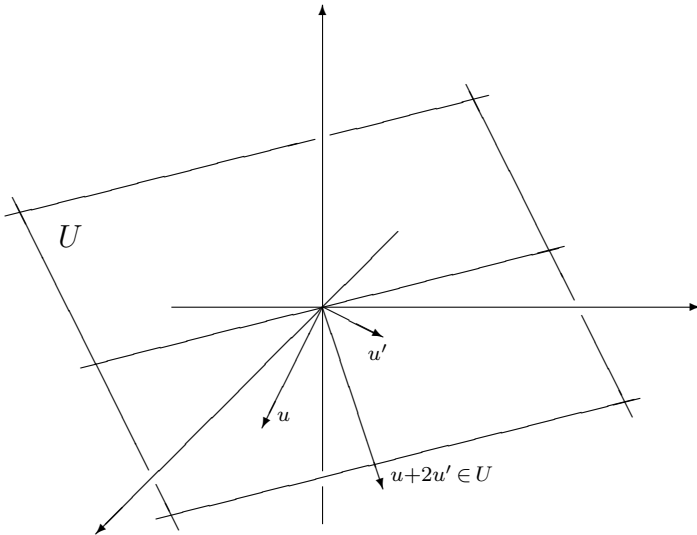
$$\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n \in U.$$

Induktion über $n \geq 0$. Für $n = 0$ ist $0 \in U$.

Sei $n \geq 1$. Es ist $\lambda_1 u_1 + \dots + \lambda_n u_n = 1 \cdot (\lambda_1 u_1 + \dots + \lambda_{n-1} u_{n-1}) + \lambda_n u_n \in U$, da nach Induktionsvoraussetzung der Klammerausdruck in U liegt.

Bemerkung. Ist U ein Unterraum von V , so ist U , zusammen mit den eingeschränkten Abbildungen $(+)|_{U \times U}^U$ und $(\cdot)|_{K \times U}^U$, selbst ein K -Vektorraum.

Skizze eines zweidimensionalen Unterraumes $U \subseteq \mathbb{R}^{3 \times 1}$.



Bemerkung. Sei $n \geq 0$. Sei (v_1, \dots, v_n) ein Tupel von Vektoren in V . Es ist der Aufspann

$$K\langle v_1, \dots, v_n \rangle \subseteq V$$

ein Unterraum.

Beispiel. Seien $m, n \geq 0$. Sei $A \in K^{m \times n}$. Es ist

$$\{x \in K^{n \times 1} : Ax = 0\} \subseteq K^{n \times 1}$$

ein Unterraum.

In der Bezeichnung des Algorithmus von §3.2 ist (x_1, \dots, x_{n-s}) eine Basis des Unterraums $\{x \in K^{n \times 1} : Ax = 0\}$.

Beispiel. Sei $k \in \mathbb{Z}$. Es ist

$$K[X]_{\leq k} := \{f(X) \in K[X] : \deg(f(X)) \leq k\} \subseteq K[X]$$

ein Unterraum.

Bemerkung. Seien $T, U \subseteq V$ Unterräume. Dann sind auch $T \cap U$ und

$$T + U := \{t + u : t \in T, u \in U\}$$

Unterräume von V .

Bemerkung. Sei V endlichdimensional. Seien $T, U \subseteq V$ Unterräume.

(1) Es ist U endlichdimensional mit $\dim(U) \leq \dim(V)$.

Es ist $\dim(U) = 0$ genau dann, wenn $U = 0$ ist.

Es ist $\dim(U) = \dim(V)$ genau dann, wenn $U = V$ ist.

(2) Es ist $\dim(T) + \dim(U) = \dim(T + U) + \dim(T \cap U)$.

Zu (2). Ist (w_1, \dots, w_m) eine Basis von $T \cap U$, so können wir diese ergänzen zu einer Basis $(w_1, \dots, w_m, t_1, \dots, t_k)$ von T und zu einer Basis $(w_1, \dots, w_m, u_1, \dots, u_\ell)$ von U . Man rechnet nach, daß $(w_1, \dots, w_m, t_1, \dots, t_k, u_1, \dots, u_\ell)$ eine Basis von $T + U$ ist.

Die fragliche Gleichung ergibt sich nun aus $(m + k) + (m + \ell) = (m + k + \ell) + m$.

Bemerkung. Seien $m, n \geq 0$. Sei (v_1, \dots, v_n) ein Tupel von Vektoren in $K^{m \times 1}$.

Habe $A \in K^{m \times n}$ als i -te Spalte $v_i \in K^{m \times 1}$ für $1 \leq i \leq n$.

Wir bringen A in die Zeilenstufenform \tilde{A} . Die Stufenspalten seien in den Positionen

$$1 \leq \ell_1 < \ell_2 < \dots < \ell_s \leq n.$$

Sei $U := {}_K\langle v_1, \dots, v_n \rangle$. Eine Basis von U ist gegeben durch

$$(v_{\ell_1}, \dots, v_{\ell_s}).$$

Bemerkung. Seien $m, n, p \geq 0$.

Sei (t_1, \dots, t_n) ein linear unabhängiges Tupel von Vektoren in $K^{m \times 1}$. Sei

$$T := {}_K\langle t_1, \dots, t_n \rangle \subseteq K^{m \times 1}.$$

Sei (u_1, \dots, u_p) ein linear unabhängiges Tupel von Vektoren in $K^{m \times 1}$. Sei

$$U := {}_K\langle u_1, \dots, u_p \rangle \subseteq K^{m \times 1}.$$

Schreibe $v_i := t_i$ für $1 \leq i \leq n$ und $v_i := u_{i-n}$ für $n+1 \leq i \leq n+p$.

Habe $A \in K^{m \times (n+p)}$ die i -te Spalte v_i für $1 \leq i \leq n+p$. D.h. entstehe A durch Nebeneinanderstellen aller Vektoren t_j und u_j .

Sei \tilde{A} die Zeilenstufenform von A . Die Stufenspalten von \tilde{A} seien in den Positionen

$$1 \leq \ell_1 < \ell_2 < \dots < \ell_s \leq n.$$

(1) Eine Basis von $T + U$ ist gegeben durch

$$(v_{\ell_1}, \dots, v_{\ell_s}).$$

(2) Sei, in der Bezeichnung von §3.2, (x_1, \dots, x_{n+p-s}) die dort gefundene Basis von

$\{x \in K^{(n+p) \times 1} : Ax = 0\}$. Schreibe $x_i = \begin{pmatrix} \xi_{1,i} \\ \vdots \\ \xi_{n+p,i} \end{pmatrix}$ für $1 \leq i \leq n+p-s$. Dann ist

$$(\xi_{1,1} t_1 + \dots + \xi_{n,1} t_n, \dots, \xi_{1,n+p-s} t_1 + \dots + \xi_{n,n+p-s} t_n)$$

eine Basis von $T \cap U$.

Denn $A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_{n+p} \end{pmatrix} = 0$ ist äquivalent zu $\xi_1 t_1 + \dots + \xi_n t_n = -(\xi_{n+1} u_1 + \dots + \xi_{n+p} u_p)$, und diese Bedingung charakterisiert Elemente in $T \cap U$.

Beispiel. Sei $K = \mathbb{R}$.

Sei $T := \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \subseteq \mathbb{R}^{5 \times 1}$.

Sei $U := \left\langle \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle \subseteq \mathbb{R}^{5 \times 1}$.

Wir wollen Basen für $T + U$ und $T \cap U$ bestimmen. Sei uns dazu bekannt, daß die erzeugenden Tupel von T und von U linear unabhängig sind.

Wir können also wie folgt rechnen.

$$A := \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 & 1 & -1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Es ist $\ell_1 = 1$, $\ell_2 = 2$, $\ell_3 = 3$ und $\ell_4 = 5$. Also ist eine Basis von $T + U$ gegeben durch

$$\left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

Eine Basis von $\{x \in \mathbb{R}^{6 \times 1} : Ax = 0\}$ ist gegeben durch $\left(\begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right)$. Also ist eine Basis von $T \cap U$ gegeben durch

$$\left(1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right) = \left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \right).$$

3.4 Der Standardvektorraum $\mathbb{R}^{n \times 1}$

Sei $n \geq 0$.

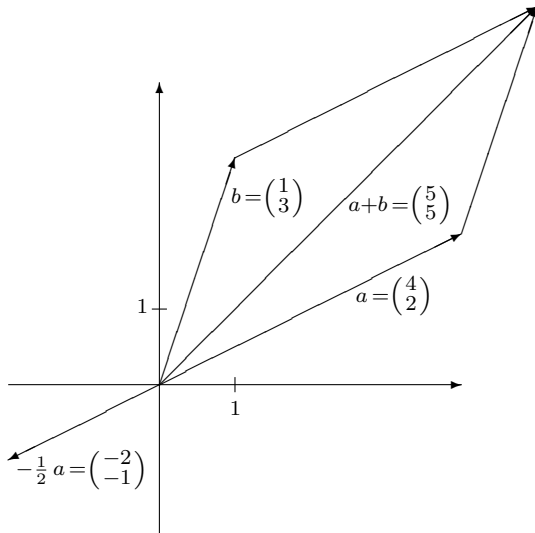
3.4.1 Geometrische Interpretation

Die geometrische Interpretation eines Vektors $a = (a_i)_i \in \mathbb{R}^{n \times 1}$ ist die eines Pfeils von $(0, \dots, 0)$ nach (a_1, \dots, a_n) in \mathbb{R}^n .

Die geometrische Interpretation der Vektoraddition ist das Aneinandersetzen von Pfeilen.

Die geometrische Interpretation des λ -fachen eines Vektors ist das Strecken um den Faktor λ (Richtungsumkehr, falls $\lambda < 0$).

Beispiel. Sei $n = 2$.



3.4.2 Skalarprodukt

Seien $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^{n \times 1}$.

Definition. Das Produkt $a^t b = a^t \cdot b \in \mathbb{R}$ heißt *Skalarprodukt* von a und b .

Es ist $a^t b = \sum_{i=1}^n a_i b_i = a_1 b_1 + \dots + a_n b_n = b^t a$.

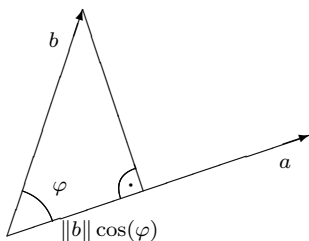
Die *Norm* oder *Länge* von a ist als $\|a\| := (a^t a)^{1/2} = (a_1^2 + \dots + a_n^2)^{1/2}$ definiert.

Lemma (Cauchy-Schwarz). Es ist $|a^t b| \leq \|a\| \cdot \|b\|$.

Bemerkung. Seien $a, b \neq 0$. Ist φ der von a und b eingeschlossene Winkel, so ist

$$a^t b = \|a\| \cdot \|b\| \cdot \cos(\varphi).$$

Insbesondere stehen a und b genau dann orthogonal aufeinander, wenn $a^t b = 0$ ist.



Beispiel. Es ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix}^t \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$. Also stehen $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ orthogonal aufeinander, d.h. senkrecht aufeinander.

Beispiel. Sei $a = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Sei $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Es ist $a^t b = 5$. Also hat der von a und b eingeschlossene Winkel φ den Cosinus

$$\cos(\varphi) = \frac{a^t b}{\|a\| \cdot \|b\|} = \frac{5}{\sqrt{13} \cdot \sqrt{2}}.$$

3.4.3 Kreuzprodukt in $\mathbb{R}^{3 \times 1}$

Seien $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^{3 \times 1}$.

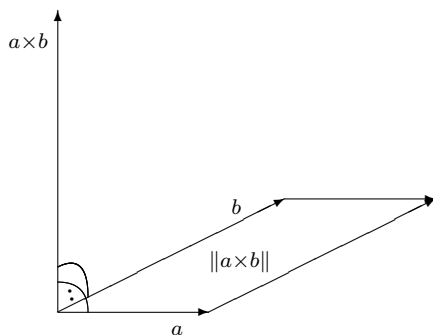
Das *Kreuzprodukt* von a und b ist definiert durch

$$a \times b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Bemerkung.

- (1) Es steht $a \times b$ orthogonal auf a und auf b .
- (2) Es ist $\|a \times b\|$ der Flächeninhalt des von a und b aufgespannten Parallelogramms. Insbesondere ist $a \times a = 0$.

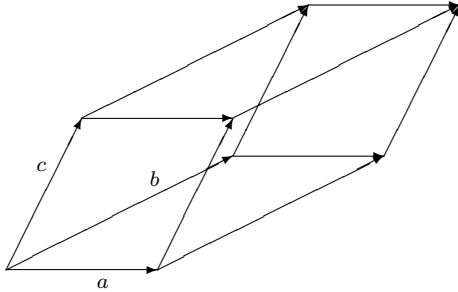
Wenn, wie üblich, die x_1 -Achse bzw. die x_2 -Achse bzw. die x_3 -Achse in Richtung des Daumens bzw. Mittelfingers bzw. Zeigefingers der linken Hand eingezeichnet werden, und a bzw. b in Richtung des Daumens bzw. Mittelfingers der linken Hand zeigt, dann zeigt $a \times b$ in Richtung des Zeigefingers.



Bemerkung. Seien $a, a', b, b', c \in \mathbb{R}^{3 \times 1}$. Seien $\lambda, \lambda', \mu, \mu' \in \mathbb{R}$.

- (1) Es ist $(\lambda a + \lambda' a') \times (\mu b + \mu' b') = \lambda \mu (a \times b) + \lambda' \mu (a' \times b) + \lambda \mu' (a \times b') + \lambda' \mu' (a' \times b')$.
- (2) Es ist $a \times b = -(b \times a)$.

(3) Es ist $|a^t(b \times c)|$ der Rauminhalt des von a , b und c aufgespannten Parallelepipeds.



Denn dieser Rauminhalt ist gleich der Grundfläche $\|b \times c\|$, multipliziert mit der Höhe $|a^t(b \times c)|/\|b \times c\|$ des Parallelepipeds.

Beispiel.

(1) Es ist $\begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} \times \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 2 - 5 \cdot 1 \\ 5 \cdot 3 - 3 \cdot 2 \\ 3 \cdot 1 - 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 9 \\ -3 \end{pmatrix}$.

(2) Seien $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^{2 \times 1}$. Das von $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ und $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ aufgespannte Parallelogramm hat den Flächeninhalt $\left\| \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0 \\ 0 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \right\| = |a_1 b_2 - a_2 b_1|$.

3.5 Lineare Abbildungen

Sei K ein Körper. Seien K -Vektorräume U , V und W gegeben.

3.5.1 Definition

Definition. Eine Abbildung $f : U \rightarrow V$ heißt K -linear oder linear, falls

$$f(\lambda \cdot u + \lambda' \cdot u') = \lambda \cdot f(u) + \lambda' \cdot f(u')$$

ist für $\lambda, \lambda' \in K$ und für $u, u' \in U$.

Bemerkung. Seien $m, n \geq 0$. Sei $A \in K^{m \times n}$. Die Abbildung

$$\begin{aligned} \text{mult}_A : K^{n \times 1} &\rightarrow K^{m \times 1} \\ x &\mapsto \text{mult}_A(x) := A \cdot x \end{aligned}$$

ist K -linear.

Jede K -lineare Abbildung von $K^{n \times 1}$ nach $K^{m \times 1}$ ist von dieser Form.

Beispiel. Die Abbildung $\varphi : K[X] \rightarrow K[X] : f(X) \mapsto X \cdot f(2) + f(X^2)$ ist K -linear.

Beispiel. Die Abbildung $\varphi : \mathbb{F}_4 \rightarrow \mathbb{F}_4 : x \mapsto x^2$ ist \mathbb{F}_2 -linear, aber nicht \mathbb{F}_4 -linear.

Beispiel. Die Abbildung $c : \mathbb{C} \rightarrow \mathbb{C} : z \mapsto \bar{z}$ ist \mathbb{R} -linear, aber nicht \mathbb{C} -linear.

Bemerkung. Seien K -lineare Abbildungen $U \xrightarrow{f} V \xrightarrow{g} W$ gegeben.

Seien K -Unterräume $U' \subseteq U$ und $V' \subseteq V$ gegeben.

- (1) Es ist $f(0) = 0$, denn $f(0) = f(0) + f(0) - f(0) = f(0+0) - f(0) = f(0) - f(0) = 0$.
- (2) Es ist $f^{-1}(V') \subseteq U$ ein K -Unterraum.
- (3) Es ist $f(U') \subseteq V$ ein K -Unterraum.
- (4) Es ist $\text{id}_U : U \rightarrow U$ eine K -lineare Abbildung.
- (5) Es ist $g \circ f : U \rightarrow W$ eine K -lineare Abbildung.

Definition. Wir erinnern an den Unterraum $0 = \{0\} \subseteq V$.

Für eine K -lineare Abbildung $f : U \rightarrow V$ heißt der K -Unterraum

$$\text{Kern}(f) := f^{-1}(0) = \{u \in U : f(u) = 0\} \subseteq U$$

der *Kern* von f .

Bemerkung. *Es ist eine K -lineare Abbildung $f : U \rightarrow V$ genau dann injektiv, wenn $\text{Kern}(f) = 0$ ist.*

Denn ist f injektiv, dann darf außer 0 kein weiteres Element auf 0 abgebildet werden.

Und umgekehrt, ist $\text{Kern}(f) = 0$ und sind $u, u' \in U$ mit $f(u) = f(u')$ gegeben, dann ist $f(u - u') = f(1 \cdot u + (-1) \cdot u') = 1 \cdot f(u) + (-1) \cdot f(u') = 0$ und also $u - u' \in \text{Kern}(f) = 0$, d.h. $u = u'$.

Lemma. *Sei U endlichdimensional. Sei $f : U \rightarrow V$ eine K -lineare Abbildung.*

Es ist

$$\dim(U) = \dim(\text{Kern}(f)) + \dim(f(U)) .$$

Sei hierzu (x_1, \dots, x_k) eine Basis von $\text{Kern}(f)$. Wir ergänzen diese zu einer Basis $(x_1, \dots, x_k, y_1, \dots, y_\ell)$ von U .

Wir behaupten, daß $(f(y_1), \dots, f(y_\ell))$ eine Basis von $f(U)$ ist.

Erzeugend: Sei $u \in U$ gegeben. Schreibe $u = \lambda_1 x_1 + \dots + \lambda_k x_k + \mu_1 y_1 + \dots + \mu_\ell y_\ell$, wobei $\lambda_i \in K$ für $1 \leq i \leq k$ und $\mu_i \in K$ für $1 \leq i \leq \ell$ ist. Es wird

$$\begin{aligned} f(u) &= f(\lambda_1 x_1 + \dots + \lambda_k x_k + \mu_1 y_1 + \dots + \mu_\ell y_\ell) \\ &= \lambda_1 f(x_1) + \dots + \lambda_k f(x_k) + \mu_1 f(y_1) + \dots + \mu_\ell f(y_\ell) \\ &= \mu_1 f(y_1) + \dots + \mu_\ell f(y_\ell) . \end{aligned}$$

Linear unabhängig: Seien $\mu_1, \dots, \mu_\ell \in K$ mit $\mu_1 f(y_1) + \dots + \mu_\ell f(y_\ell) = 0$ gegeben. Dann ist

$$0 = \mu_1 f(y_1) + \dots + \mu_\ell f(y_\ell) = f(\mu_1 y_1 + \dots + \mu_\ell y_\ell) ,$$

also $\mu_1 y_1 + \dots + \mu_\ell y_\ell \in \text{Kern}(f)$ und somit

$$\mu_1 y_1 + \dots + \mu_\ell y_\ell = \lambda_1 x_1 + \dots + \lambda_k x_k$$

für gewisse $\lambda_i \in K$ für $1 \leq i \leq k$. Die lineare Unabhängigkeit von $(x_1, \dots, x_k, y_1, \dots, y_\ell)$ liefert nun $\mu_i = 0$ für $1 \leq i \leq k$ und $\lambda_i = 0$ für $1 \leq i \leq \ell$, insbesondere ersteres.

Dies zeigt die *Behauptung*. Es folgt $\dim(U) = k + \ell = \dim(\text{Kern}(f)) + \dim(f(U))$.

Bemerkung. Sei U endlichdimensional. Sei (x_1, \dots, x_m) eine Basis von U .

Sei (v_1, \dots, v_m) ein Tupel von Vektoren in V .

Es gibt genau eine lineare Abbildung $f : U \rightarrow V$ mit $f(x_i) = v_i$ für $1 \leq i \leq m$.

Diese ist gegeben durch

$$f(\lambda_1 x_1 + \dots + \lambda_m x_m) = \lambda_1 v_1 + \dots + \lambda_m v_m.$$

für $\lambda_i \in K$ für $1 \leq i \leq m$.

Es ist f injektiv genau dann, wenn (v_1, \dots, v_m) linear unabhängig ist.

Es ist f surjektiv genau dann, wenn (v_1, \dots, v_m) erzeugend in V ist.

Es ist f bijektiv genau dann, wenn (v_1, \dots, v_m) eine Basis von V ist.

Bemerkung. Seien $m, n \geq 0$. Sei $A \in K^{m \times n}$.

Wir betrachten die K -lineare Abbildung $\text{mult}_A : K^{n \times 1} \rightarrow K^{m \times 1} : x \mapsto Ax$.

Wir schreiben auch $\text{Kern}(A) := \text{Kern}(\text{mult}_A)$.

Sei (v_1, \dots, v_n) das Tupel der Spaltenvektoren von A . Diese liegen in $K^{m \times 1}$.

Es ist

$$\begin{aligned} \text{Kern}(A) &= \text{Kern}(\text{mult}_A) \\ &= \{x \in K^{n \times 1} : \text{mult}_A(x) = 0\} \\ &= \{x \in K^{n \times 1} : A \cdot x = 0\}. \end{aligned}$$

Es ist

$$\begin{aligned} \text{mult}_A(K^{n \times 1}) &= \{A \cdot y : y \in K^{n \times 1}\} \\ &= \{\lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_i \in K \text{ für } 1 \leq i \leq n\} \\ &= {}_K \langle v_1, \dots, v_n \rangle. \end{aligned}$$

Eine Basis von $\text{mult}_A(K^{n \times 1})$ ist hier gegeben durch $(v_{\ell_1}, \dots, v_{\ell_s})$, wobei $1 \leq \ell_1 < \dots < \ell_s \leq n$ die Positionen der Stufenstellen in der Zeilenstufenform \tilde{A} von A bezeichnen.

Wir schreiben auch

$$\text{Rang}(A) := \dim(\text{mult}_A(K^{n \times 1})) = s$$

für den *Rang* von A . Es wird so

$$n = \dim(K^{n \times 1}) = \dim(\text{Kern}(\text{mult}_A)) + \dim(\text{mult}_A(K^{n \times 1})) = \dim(\text{Kern}(A)) + \text{Rang}(A).$$

In der Tat ist in der Zeilenstufenform \tilde{A} von A auch n die Anzahl der Spalten von \tilde{A} . Ferner ist $\dim(\text{Kern}(A))$ die Anzahl $n - s$ der Nichtstufenspalten und $\text{Rang}(A)$ die Anzahl s der Stufenspalten. Und es ist $n = (n - s) + s$.

Bemerkung. Seien $m, n \geq 0$. Sei $A \in K^{m \times n}$. Es ist $\text{Rang}(A) = \text{Rang}(A^t)$.

Beispiel. Wir betrachten die lineare Abbildung $\varphi : \mathbb{F}_3[X] \rightarrow \mathbb{F}_3^{\mathbb{F}_3}$, die $f(X)$ auf die Abbildung $\varphi(f(X)) : \mathbb{F}_3 \rightarrow \mathbb{F}_3 : x \mapsto f(x)$ schickt.

Wir bestimmen den Kern von φ . Es ist $\varphi(f(X)) = 0$ genau dann, wenn $f(x) = 0$ ist für $x \in \mathbb{F}_3$, i.e. wenn $f(X)$ die Nullstellen 0, 1 und -1 hat.

Dies ist genau dann der Fall, wenn $f(X)$ teilbar ist durch $X \cdot (X - 1) \cdot (X + 1) = X^3 - X$.

Wir erhalten

$$\text{Kern}(\varphi) = \{ f(X) \in \mathbb{F}_3[X] : f(X) \text{ ist teilbar durch } X^3 - X \}.$$

3.5.2 Koordinatenvektoren und beschreibende Matrizen

Seien die K -Vektorräume U, V und W endlichdimensional.

Sei $B := (b_1, \dots, b_m)$ eine Basis von U .

Sei $C := (c_1, \dots, c_n)$ eine Basis von V .

Sei $D := (d_1, \dots, d_p)$ eine Basis von W .

Definition. Sei $u \in U$ gegeben. Es gibt auf eindeutige Weise $\lambda_1, \dots, \lambda_m \in K$ mit

$$u = \sum_{i=1}^m \lambda_i b_i.$$

Wir schreiben

$${}_B u := (\lambda_i)_i = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} \in K^{m \times 1}$$

für den *Koordinatenvektor* von u bezüglich B .

Beispiel. Es ist $B := (1, X, X^2)$ eine Basis des K -Vektorraums $K[X]_{\leq 2}$.

Es ist $X^2 + 2X - 3 = (-3) \cdot 1 + 2 \cdot X + 1 \cdot X^2$. Also ist ${}_B(X^2 + 2X - 3) = \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$.

Definition. Sei $f : U \rightarrow V$ eine lineare Abbildung.

Für $1 \leq i \leq m$ gibt es auf eindeutige Weise $x_{1,i}, \dots, x_{n,i} \in K$ mit

$$f(b_i) = \sum_{j=1}^n x_{j,i} c_j.$$

Wir schreiben

$${}_C f_B := (x_{j,i})_{j,i} \in K^{n \times m}$$

für die *beschreibende Matrix* von f bezüglich C und B .

D.h. für $1 \leq i \leq m$ ist die i -te Spalte von ${}_C f_B$ gegeben durch ${}_C f(b_i) = \begin{pmatrix} x_{1,i} \\ \vdots \\ x_{n,i} \end{pmatrix}$.

Beispiel. Sei $K = \mathbb{R}$. Sei $U := \mathbb{R}^{3 \times 1}$ und $V := \mathbb{R}^{2 \times 1}$.

Sei $A := \begin{pmatrix} 4 & 3 & 1 \\ 2 & 5 & 5 \end{pmatrix} \in \mathbb{R}^{2 \times 3}$. Sei $f := \text{mult}_A : \mathbb{R}^{3 \times 1} \rightarrow \mathbb{R}^{2 \times 1} : x \mapsto f(x) = Ax$.

(1) Sei $B = (b_1, b_2, b_3) := \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$.

Sei $C = (c_1, c_2) := \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$.

Es ist $f(b_1) = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 4 \cdot c_1 + 2 \cdot c_2$. Also ist ${}_C f(b_1) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

Genauso wird ${}_C f(b_2) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ und ${}_C f(b_3) = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$.

Also ist ${}_C f_B = \begin{pmatrix} 4 & 3 & 1 \\ 2 & 5 & 5 \end{pmatrix} = A$.

(2) Sei $B' = (b'_1, b'_2, b'_3) := \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$.

Sei $C' = (c'_1, c'_2) := \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$.

Es ist $f(b'_1) = \begin{pmatrix} 5 \\ 7 \end{pmatrix} = 7 \cdot c'_1 - 2 \cdot c'_2$. Also ist ${}_{C'} f(b'_1) = \begin{pmatrix} 7 \\ -2 \end{pmatrix}$.

Es ist $f(b'_2) = \begin{pmatrix} 4 \\ 10 \end{pmatrix} = 10 \cdot c'_1 - 6 \cdot c'_2$. Also ist ${}_{C'} f(b'_2) = \begin{pmatrix} 10 \\ -6 \end{pmatrix}$.

Es ist $f(b'_3) = \begin{pmatrix} 1 \\ 5 \end{pmatrix} = 5 \cdot c'_1 - 4 \cdot c'_2$. Also ist ${}_{C'} f(b'_3) = \begin{pmatrix} 5 \\ -4 \end{pmatrix}$.

Somit ist ${}_{C'} f_{B'} = \begin{pmatrix} 7 & 10 & 5 \\ -2 & -6 & -4 \end{pmatrix}$.

Bemerkung. Es hat für $A \in K^{n \times m}$ die Abbildung $\text{mult}_A : K^{m \times 1} \rightarrow K^{n \times 1}$ bezüglich der Standardbasen die beschreibende Matrix A .

Beispiel. Sei $\varphi : \mathbb{R}[X]_{\leq 2} \rightarrow \mathbb{R}[X]_{\leq 3} : f(X) \mapsto X \cdot f(X) - f(1)$.

Wir wählen die Basen $B := (1, X, X^2)$ von $\mathbb{R}[X]_{\leq 2}$ und $C := (1, X, X^2, X^3)$ von $\mathbb{R}[X]_{\leq 3}$.

Es wird $\varphi(1) = -1 + X$, $\varphi(X) = -1 + X^2$ und $\varphi(X^2) = -1 + X^3$.

Also ist ${}_C \varphi_B = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Lemma. Seien $U \xrightarrow{f} V \xrightarrow{g} W$ lineare Abbildungen. Sei $u \in U$.

(1) Es ist ${}_C f(u) = {}_C f_B \cdot {}_B u$.

(2) Es ist ${}_D (g \circ f)_B = {}_D g_C \cdot {}_C f_B$.

Dies sehen wir wie folgt ein.

Sei $r : U \rightarrow K^{m \times 1}$ die bijektive lineare Abbildung, die $b_i \mapsto e_i$ schickt für $1 \leq i \leq m$.

Sei $s : V \rightarrow K^{n \times 1}$ die bijektive lineare Abbildung, die $c_i \mapsto e_i$ schickt für $1 \leq i \leq n$.

Sei $t : W \rightarrow K^{p \times 1}$ die bijektive lineare Abbildung, die $d_i \mapsto e_i$ schickt für $1 \leq i \leq p$.

Dann ist ${}_B u = r(u)$. Genauso ist ${}_C v = s(v)$ für $v \in V$.

Ferner ist $s \circ f \circ r^{-1} : K^{m \times 1} \rightarrow K^{n \times 1}$ die Abbildung, die e_i nach ${}_C f(b_i)$ schickt für $1 \leq i \leq m$. Also ist $s \circ f \circ r^{-1} = \text{mult}_{{}_C f_B}$. Also wird, wie in (1) behauptet,

$${}_C f_B \cdot {}_B u = \text{mult}_{{}_C f_B}({}_B u) = (s \circ f \circ r^{-1})(r(u)) = s(f(u)) = {}_C f(u).$$

Genauso ist $t \circ g \circ s^{-1} = \text{mult}_{Dg_C}$ und $t \circ g \circ f \circ r^{-1} = \text{mult}_{D(g \circ f)_B}$. Also wird

$$\text{mult}_{Dg_C \cdot {}_C f_B} = \text{mult}_{Dg_C} \circ \text{mult}_{{}_C f_B} = (t \circ g \circ s^{-1}) \circ (s \circ f \circ r^{-1}) = t \circ g \circ f \circ r^{-1} = \text{mult}_{D(g \circ f)_B},$$

und (2) folgt.

Beispiel. Wir betrachten die \mathbb{F}_2 -lineare Abbildung $f : \mathbb{F}_4 \rightarrow \mathbb{F}_4 : x \mapsto x^2$.

Bezüglich der Basis $B = (1, \alpha)$ des \mathbb{F}_2 -Vektorraums \mathbb{F}_4 wird ${}_B f_B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ wegen $f(1) = 1^2 = 1 \cdot 1 + 0 \cdot \alpha$ und $f(\alpha) = \alpha^2 = 1 \cdot 1 + 1 \cdot \alpha$.

Es ist $f \circ f : \mathbb{F}_4 \rightarrow \mathbb{F}_4 : x \mapsto x^4 = x$. Also ist $f \circ f = \text{id}$. In der Tat ist auch ${}_B (f \circ f)_B = {}_B f_B \cdot {}_B f_B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 = \text{Bid}_B$.

Bemerkung. Seien B und B' Basen von U . Seien C und C' Basen von V .

Sei $f : U \rightarrow V$ eine lineare Abbildung.

Dann wird

$${}_C \text{id}_C \cdot {}_C f_B \cdot {}_B \text{id}_{B'} = {}_C (\text{id} \circ f \circ \text{id})_{B'} = {}_C f_{B'}.$$

Ferner ist ${}_B \text{id}_{B'} \cdot {}_B \text{id}_B = {}_B \text{id}_B = E_m$ und also ${}_B \text{id}_{B'} = ({}_B \text{id}_B)^{-1}$.

Bemerkung. Ist $U = K^{m \times 1}$ und ist B die Standardbasis von $K^{m \times 1}$, dann ist ${}_B u = u$ für $u \in K^{m \times 1}$.

Ist B' eine weitere Basis von $K^{m \times 1}$, dann sind die Spalten von ${}_B \text{id}_{B'}$ also gerade die Vektoren von B' .

Beispiel. Wir greifen das vorvorige Beispiel wieder auf.

Es war $B' = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$. Also ist ${}_B \text{id}_{B'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

Es war $C' = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$. Also ist ${}_C \text{id}_{C'} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Es war ${}_C f_B = A = \begin{pmatrix} 4 & 3 & 1 \\ 2 & 5 & 5 \end{pmatrix}$. Also können wir auch wie folgt rechnen.

$${}_C f_{B'} = {}_C \text{id}_C \cdot {}_C f_B \cdot {}_B \text{id}_{B'} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 4 & 3 & 1 \\ 2 & 5 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 5 & 4 & 1 \\ 7 & 10 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 10 & 5 \\ -2 & -6 & -4 \end{pmatrix}.$$

3.6 Determinanten

Sei K ein kommutativer Ring. Sei $n \geq 0$.

Definition. Sei $\sigma \in S_n$ gegeben.

Sei kurz $N := \{1, \dots, n\}$ geschrieben. Sei

$$\text{fehl}(\sigma) := |\{(i, j) \in N \times N : \text{es ist } i < j, \text{ aber } \sigma(i) > \sigma(j)\}|$$

der *Fehlstand* von σ . Sei damit

$$\text{sgn}(\sigma) := (-1)^{\text{fehl}(\sigma)}$$

das *Signum* von σ .

Beispiel. Wir betrachten die Elemente id , $\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ und $\tau := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ von S_3 .

Es ist $\text{fehl}(\text{id}) = |\emptyset| = 0$. Also ist $\text{sgn}(\text{id}) = (-1)^0 = +1$.

Es ist $\text{fehl}(\sigma) = |\{(1, 2)\}| = 1$. Also ist $\text{sgn}(\sigma) = (-1)^1 = -1$.

Es ist $\text{fehl}(\tau) = |\{(1, 3), (2, 3)\}| = 2$. Also ist $\text{sgn}(\tau) = (-1)^2 = +1$.

Bemerkung. Es ist $\text{sgn}(\text{id}) = +1$.

Es ist $\text{sgn}(\tau \circ \sigma) = \text{sgn}(\tau) \cdot \text{sgn}(\sigma)$ für $\tau, \sigma \in S_n$.

Definition. Sei $A = (a_{i,j})_{i,j} \in K^{n \times n}$.

Sei

$$\det A = \det(A) := \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n a_{i, \sigma(i)} \in K$$

die *Determinante* von A .

Beispiel.

(1) Sei $n = 1$. Sei $A = (a_{1,1})$. Es ist $\det(a_{1,1}) = \sum_{\sigma \in S_1} \text{sgn}(\sigma) \prod_{i=1}^1 a_{i, \sigma(i)} = \text{sgn}(\text{id}) a_{1,1} = a_{1,1}$.

(2) Sei $n = 2$. Wir schreiben $\sigma_{1,2} := \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in S_2$. Es ist $\text{sgn}(\sigma_{1,2}) = -1$.

Es ist $S_2 = \{\text{id}, \sigma_{1,2}\}$. Also ist

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \text{sgn}(\text{id}) \cdot a_{1,1} \cdot a_{2,2} + \text{sgn}(\sigma_{1,2}) \cdot a_{1,2} \cdot a_{2,1} = a_{1,1} \cdot a_{2,2} - a_{1,2} \cdot a_{2,1}.$$

(3) Sei $n = 3$. Wir schreiben $\sigma_{1,2} := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, $\sigma_{2,3} := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $\sigma_{1,3} := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, $\sigma_{1,2,3} := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $\sigma_{1,3,2} := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ in S_3 . Es ist $\text{sgn}(\sigma_{1,2}) = \text{sgn}(\sigma_{2,3}) = \text{sgn}(\sigma_{1,3}) = -1$. Es ist $\text{sgn}(\sigma_{1,2,3}) = \text{sgn}(\sigma_{1,3,2}) = +1$.

Es ist $S_3 = \{\text{id}, \sigma_{1,2}, \sigma_{2,3}, \sigma_{1,3}, \sigma_{1,2,3}, \sigma_{1,3,2}\}$. Also ist

$$\begin{aligned} & \det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \\ &= \text{sgn}(\text{id}) \cdot a_{1,1} \cdot a_{2,2} \cdot a_{3,3} + \text{sgn}(\sigma_{1,2,3}) \cdot a_{1,2} \cdot a_{2,3} \cdot a_{3,1} + \text{sgn}(\sigma_{1,3,2}) \cdot a_{1,3} \cdot a_{2,1} \cdot a_{3,2} \\ & \quad + \text{sgn}(\sigma_{1,3}) \cdot a_{1,3} \cdot a_{2,2} \cdot a_{3,1} + \text{sgn}(\sigma_{2,3}) \cdot a_{1,1} \cdot a_{2,3} \cdot a_{3,2} + \text{sgn}(\sigma_{1,2}) \cdot a_{1,2} \cdot a_{2,1} \cdot a_{3,3} \\ &= a_{1,1} \cdot a_{2,2} \cdot a_{3,3} + a_{1,2} \cdot a_{2,3} \cdot a_{3,1} + a_{1,3} \cdot a_{2,1} \cdot a_{3,2} \\ & \quad - a_{1,3} \cdot a_{2,2} \cdot a_{3,1} - a_{1,1} \cdot a_{2,3} \cdot a_{3,2} - a_{1,2} \cdot a_{2,1} \cdot a_{3,3} . \end{aligned}$$

Die Definition gibt so die *Regel von Sarrus*.

Bemerkung. Seien $A, B \in K^{n \times n}$. Sei $\lambda \in K$.

- (1) Es ist $\det(E_n) = 1$.
- (2) Es ist $\det(A \cdot B) = \det(A) \cdot \det(B)$.
- (3) Es ist $\det(A^t) = \det(A)$.
- (4) Entstehe A' aus A durch Addition des λ -fachen einer Zeile zu einer anderen.
Dann ist $\det(A') = \det(A)$.
- (5) Entstehe A' aus A durch Multiplikation einer Zeile mit λ .
Dann ist $\det(A') = \lambda \cdot \det(A)$.
- (6) Entstehe A' aus A durch Vertauschen zweier Zeilen.
Dann ist $\det(A') = -\det(A)$.

Es folgen (4, 5, 6) aus (2).

Aus (3) folgt, daß in (4, 5, 6) die entsprechenden Regeln für Spaltenumformungen genauso gelten.

Dank (4, 5, 6) kann über einem Körper das Gaußverfahren zur Berechnung der Determinante verwendet werden.

Computer berechnen aus Zeitgründen die Determinante mit dem Gaußverfahren und nicht direkt nach Definition.

Bemerkung. Sei $A \in K^{n \times n}$.

- (1) Es ist $\det(A^k) = \det(A)^k$ für $k \in \mathbb{Z}_{\geq 0}$.
Ist A invertierbar, dann ist auch $\det(A^{-1}) = \det(A)^{-1}$.
- (2) Es ist $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$.

Beispiel.

(1) Sei $A = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Es ist $\det \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix} = 2 \cdot 5 - 4 \cdot 3 = -2$.

Wir erinnern an $A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} 5 & -4 \\ -3 & 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -5 & 4 \\ 3 & -2 \end{pmatrix}$.

(2) Sei $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 1 \\ 2 & 3 & 3 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$.

Sarrus gibt $\det \begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 1 \\ 2 & 3 & 3 \end{pmatrix} = 1 \cdot 3 \cdot 3 + 0 \cdot 1 \cdot 2 + 1 \cdot 2 \cdot 3 - 1 \cdot 3 \cdot 2 - 1 \cdot 1 \cdot 3 - 0 \cdot 2 \cdot 3 = 6$.

Gauß gibt $\det \begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 1 \\ 2 & 3 & 3 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 3 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix} = 3 \cdot 2 \cdot \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix}$
 $= 3 \cdot 2 \cdot \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 6$.

(3) Sei $A = \begin{pmatrix} 1 & 0 & \alpha & 1 \\ \alpha & \alpha & 1 & 1 \\ 1 & \alpha & 0 & 1 \\ 1 & 1 & 0 & \alpha \end{pmatrix} \in \mathbb{F}_4^{4 \times 4}$. Gauß liefert $\det \begin{pmatrix} 1 & 0 & \alpha & 1 \\ \alpha & \alpha & 1 & 1 \\ 1 & \alpha & 0 & 1 \\ 1 & 1 & 0 & \alpha \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & \alpha & 1 \\ 0 & \alpha & \alpha & 1+\alpha \\ 0 & 1 & \alpha & 1+\alpha \\ 0 & 1 & \alpha & 1+\alpha \end{pmatrix} =$
 $\alpha \cdot \det \begin{pmatrix} 1 & 0 & \alpha & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & \alpha & 1+\alpha \end{pmatrix} = -\alpha \cdot \det \begin{pmatrix} 1 & 0 & \alpha & 1 \\ 0 & 1 & 1 & 0 \\ 0 & \alpha & \alpha & 1+\alpha \end{pmatrix} = \alpha \cdot \det \begin{pmatrix} 1 & 0 & \alpha & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1+\alpha \\ 0 & 0 & 1+\alpha & 1+\alpha \end{pmatrix} =$
 $\alpha \cdot (1 + \alpha) \cdot \det \begin{pmatrix} 1 & 0 & \alpha & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1+\alpha \\ 0 & 0 & 1 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 & 1+\alpha \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1+\alpha \end{pmatrix} = (1 + \alpha) \cdot \det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 1 + \alpha$.

Lemma. Sei K ein Körper. Sei $A \in K^{n \times n}$ gegeben.

Es ist genau dann $\det(A) \neq 0$, wenn A invertierbar ist.

Zunächst hat die Zeilenstufenform von A genau dann wenigstens eine Nullzeile, wenn A nicht invertierbar ist.

Dann zeigt die Berechnung der Determinante via Gaußverfahren, daß die Zeilenstufenform von A genau dann wenigstens eine Nullzeile hat, wenn $\det(A) = 0$ ist.

Bemerkung.

(1) Seien $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ und $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ in $\mathbb{R}^{2 \times 1}$ gegeben.

Das von a und b aufgespannte Parallelogramm hat den Flächeninhalt

$$\left\| \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \right\| = |\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}|.$$

(2) Seien $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ und $c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ in $\mathbb{R}^{3 \times 1}$ gegeben.

Das von a , b und c aufgespannte Parallelepiped hat den Rauminhalt

$$|a^t (b \times c)| = |(a_1 \ a_2 \ a_3) \begin{pmatrix} b_2 c_3 - b_3 c_2 \\ b_3 c_1 - b_1 c_3 \\ b_1 c_2 - b_2 c_1 \end{pmatrix}| = |a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1| = |\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}|.$$

Bemerkung.

(1) Seien $s, t \geq 0$ mit $s + t = n$ gegeben.

Sei $A \in K^{s \times s}$. Sei $B \in K^{s \times t}$. Sei $C \in K^{t \times s}$. Sei $D \in K^{t \times t}$.

Wir bilden die Blockmatrix $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in K^{n \times n}$. Es ist $\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A) \cdot \det(D)$.

Wir bilden die Blockmatrix $\begin{pmatrix} A & 0 \\ C & D \end{pmatrix} \in K^{n \times n}$. Es ist $\det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det(A) \cdot \det(D)$.

Dies lässt sich auch iteriert anwenden, um größere Blockmatrizen zu behandeln.

(2) Sei $A = (a_{i,j})_{i,j} \in K^{n \times n}$.

Ist $a_{i,j} = 0$ für $1 \leq j < i \leq n$, so heißt A eine *obere Dreiecksmatrix*.

Ist $a_{i,j} = 0$ für $1 \leq i < j \leq n$, so heißt A eine *untere Dreiecksmatrix*.

Ist A eine obere Dreiecksmatrix oder eine untere Dreiecksmatrix, dann folgt aus (1):

$$\det(A) = \prod_{i=1}^n a_{i,i} = a_{1,1} \cdot a_{2,2} \cdot \dots \cdot a_{n,n} .$$

Lemma (Entwicklung). Sei $n \geq 1$. Sei $A = (a_{i,j})_{i,j} \in K^{n \times n}$.

Für $1 \leq i \leq n$ und $1 \leq j \leq n$ entstehe $A_{[i,j]} \in K^{(n-1) \times (n-1)}$ aus A durch Streichen der i -ten Zeile und der j -ten Spalte.

Sei $1 \leq k \leq n$ gewählt.

Es ist

$$\det(A) = \sum_{i=1}^n (-1)^{i+k} \cdot \det(A_{[i,k]}) \cdot a_{i,k} ,$$

genannt *Entwicklung nach der k -ten Spalte*.

Es gilt auch die transponierte Aussage. Es ist

$$\det(A) = \sum_{j=1}^n (-1)^{k+j} \cdot \det(A_{[k,j]}) \cdot a_{k,j} ,$$

genannt *Entwicklung nach der k -ten Zeile*.

Beispiel.

Sei $A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix} \in \mathbb{F}_5^{4 \times 4}$. Es wird

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix} &= (-1)^{2+1} \cdot 2 \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} + (-1)^{2+4} \cdot 1 \cdot \det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \\ &= -2 \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \\ &= -2 \cdot \det \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix} \\ &= -2 \cdot (-1)^{1+3} \cdot 1 \cdot \det \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \\ &= -2 \cdot 3 \\ &= -1 . \end{aligned}$$

3.7 Normalform

Sei K ein Körper. Sei $n \geq 0$. Sei $A = (a_{i,j})_{i,j} \in K^{n \times n}$.

3.7.1 Eigenwerte, Eigenvektoren, Eigenräume

Definition.

- (1) Es heißt $\lambda \in K$ ein Eigenwert von A , falls es ein $x \in K^{n \times 1} \setminus \{0\}$ gibt mit $Ax = \lambda x$.
- (2) Ist λ ein Eigenwert von A , dann heißt jedes $x \in K^{n \times 1} \setminus \{0\}$ mit $Ax = \lambda x$ ein *Eigenvektor* von A zum Eigenwert λ .
- (3) Ist λ ein Eigenwert von A , dann heißt

$$E_A(\lambda) := \{x \in K^{n \times 1} : Ax = \lambda x\} = \{x \in K^{n \times 1} : (A - \lambda E_n)x = 0\} = \text{Kern}(A - \lambda E_n)$$

der *Eigenraum* von A zum Eigenwert λ .

Dieser besteht also aus allen Eigenvektoren von A zum Eigenwert λ und aus dem Nullvektor.

Bemerkung. Sei $B \in K^{n \times n}$.

Es ist genau dann $\text{Kern}(B) = 0$, wenn die Spalten von B ein linear unabhängiges Tupel bilden, d.h. wenn die Zeilenstufenform von B gleich E_n ist, d.h. wenn B invertierbar ist, d.h. wenn $\det(B) \neq 0$ ist.

Umgekehrt ist also genau dann $\text{Kern}(B) \neq 0$, wenn B nicht invertierbar ist, d.h. wenn $\det(B) = 0$ ist.

Bemerkung. Sei $\lambda \in K$.

Es ist genau dann λ ein Eigenwert von A , wenn es ein $0 \neq x \in \text{Kern}(A - \lambda E_n)$ gibt, d.h. wenn $\text{Kern}(A - \lambda E_n) \neq 0$ ist, d.h. wenn $\det(A - \lambda E_n) = 0$ ist.

Definition. Sei

$$\chi_A(X) := \det(A - XE_n) \in K[X]$$

das *charakteristische Polynom* von A .

Bemerkung. Es ist $\lambda \in K$ genau dann ein Eigenwert von A , wenn λ eine Nullstelle von $\chi_A(X)$ ist.

Beispiel.

- (1) Sei $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Es ist $\chi_A(X) := \det \begin{pmatrix} -X & 1 \\ 1 & -X \end{pmatrix} = X^2 - 1 = (X + 1)(X - 1)$. Also hat A die Eigenwerte 1 und -1 in \mathbb{R} .

Wir lösen $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix} \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \rightsquigarrow \left(\begin{smallmatrix} 1 & -1 \\ 0 & 0 \end{smallmatrix} \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ und erhalten den Eigenraum $E_A(1) = \mathbb{R}\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$. Es ist z.B. $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Eigenvektor von A zum Eigenwert 1.

Wir lösen $\left(\begin{smallmatrix} 0 & -(-1) \\ 1 & 0 \end{smallmatrix} \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \rightsquigarrow \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix} \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ und erhalten den Eigenraum $E_A(-1) = \mathbb{R}\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rangle$. Es ist z.B. $\begin{pmatrix} -2 \\ 2 \end{pmatrix}$ ein Eigenvektor von A zum Eigenwert -1 .

- (2) Sei $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Es ist $\chi_A(X) := \det \begin{pmatrix} 0-X & 1 \\ -1 & 0-X \end{pmatrix} = X^2 + 1$. Also hat A keine Eigenwerte in \mathbb{R} .

Definition. Sei $\text{tr}(A) := \sum_{i=1}^n a_{i,i} = a_{1,1} + a_{2,2} + \dots + a_{n,n}$ die *Spur* von A (engl. *trace*).

Bemerkung.

- (1) Es ist $\deg(\chi_A(X)) = n$.
- (2) Der Koeffizient von $\chi_A(X)$ bei X^n ist $(-1)^n$.
- (3) Sei $n \geq 1$. Der Koeffizient von $\chi_A(X)$ bei X^{n-1} ist $(-1)^{n-1} \cdot \text{tr}(A)$.
- (4) Der Koeffizient von $\chi_A(X)$ bei X^0 ist $\det(A)$.

Die Spur ist eine gute Probe bei der Berechnung von $\chi_A(X)$.

Definition. Sei $f(X) \in K[X] \setminus \{0\}$. Sei $\deg(f(X)) =: m$.

- (1) Wir sagen, es *zerfällt* $f(X) \in K[X]$, falls es $u \in K \setminus \{0\}$ und $t_i \in K$ für $1 \leq i \leq m$ gibt mit $f(X) = u \cdot \prod_{i=1}^m (X - t_i)$.
- (2) Sei $t \in K$ eine Nullstelle von $f(X)$, d.h. sei $f(t) = 0$.

Sei $k \geq 1$. Wir sagen, es ist k die *Vielfachheit* von t als Nullstelle von $f(X)$, wenn $(X - t)^k$ ein Teiler von $f(X)$ ist, nicht aber $(X - t)^{k+1}$.

Bemerkung. Genau dann zerfällt jedes Polynom in $K[X] \setminus \{0\}$, wenn K algebraisch abgeschlossen ist.

Insbesondere zerfällt jedes Polynom in $\mathbb{C}[X]$.

Aber z.B. $X^2 + 2X + 5 = (X + 1)^2 + 4 \in \mathbb{R}[X]$ zerfällt nicht.

Wir werden hauptsächlich Matrizen mit zerfallendem charakteristischem Polynom betrachten. In $\mathbb{C}^{n \times n}$ sind das alle Matrizen.

Definition. Sei $\lambda \in K$ ein Eigenwert von A .

- (1) Die *algebraische Vielfachheit* $\text{aV}_A(\lambda)$ von λ als Eigenwert von A ist die Vielfachheit von λ als Nullstelle von $\chi_A(X)$.
- (2) Die *geometrische Vielfachheit* $\text{gV}_A(\lambda)$ von λ ist die Dimension des Eigenraums von A zum Eigenwert λ . Mit anderen Worten, $\text{gV}_A(\lambda) := \dim_K(\text{E}_A(\lambda))$.

Bemerkung. Sei $\lambda \in K$ ein Eigenwert von $A \in K^{n \times n}$. Es ist

$$1 \leq \text{gV}_A(\lambda) \leq \text{aV}_A(\lambda) \leq n.$$

Beispiel. Sei $A := \begin{pmatrix} 3 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$. Es ist

$$\chi_A(X) = \det \begin{pmatrix} 3-X & 1 & 1 \\ 0 & 3-X & 1 \\ 0 & 0 & 5-X \end{pmatrix} = (3-X)^2(5-X) = -(X-3)^2(X-5).$$

Also hat A die Eigenwerte 3 und 5 in \mathbb{R} . Es ist $\text{aV}_A(3) = 2$ und $\text{aV}_A(5) = 1$.

Was Spur und Determinante angeht, es ist auch

$$\chi_A(X) = -(X^2 - 6X + 9)(X - 5) = -X^3 + 11X^2 - 39X + 45.$$

In der Tat ist $\det(A) = 45$ und $(-1)^{3-1} \text{tr}(A) = 11$.

Da $1 \leq \text{gV}_A(5) \leq \text{aV}_A(5) = 1$, folgt $\text{gV}_A(5) = 1$. Wir wollen eine Basis von $\text{E}_A(5)$ berechnen, um dies zu bestätigen. Dazu lösen wir $\left(\begin{array}{ccc|c} 3-5 & 1 & 1 & 0 \\ 0 & 3-5 & 1 & 0 \\ 0 & 0 & 5-5 & 0 \end{array} \right) = \left(\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1/2 & -1/2 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & -3/4 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$ und erhalten $\left(\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} \right)$ als Basis von $\text{E}_A(5)$. In der Tat finden wir $\text{gV}_A(5) = \dim_{\mathbb{R}} \text{E}_A(5) = 1$ bestätigt.

Es ist $1 \leq \text{gV}_A(3) \leq \text{aV}_A(3) = 2$, d.h. $\text{gV}_A(3) \in \{1, 2\}$. Um $\text{gV}_A(3)$ zu berechnen, bestimmen wir eine Basis von $\text{E}_A(3)$.

Dazu lösen wir $\left(\begin{array}{ccc|c} 3-3 & 1 & 1 & 0 \\ 0 & 3-3 & 1 & 0 \\ 0 & 0 & 1-3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right)$ als Basis von $\text{E}_A(3)$. Somit ist $\text{gV}_A(3) = \dim_{\mathbb{R}} \text{E}_A(3) = 1$.

Bemerkung. Es ist $A \in K^{n \times n}$. Sei $S \in K^{n \times n}$ invertierbar. Sei $A' := S^{-1} \cdot A \cdot S$.

Es ist

$$\begin{aligned} \chi_{A'}(X) &= \det(S^{-1} \cdot A \cdot S - X \cdot \text{E}_n) \\ &= \det(S^{-1} \cdot (A - X \cdot \text{E}_n) \cdot S) \\ &= \det(S)^{-1} \cdot \det(A - X \cdot \text{E}_n) \cdot \det(S) \\ &= \chi_A(X). \end{aligned}$$

Insbesondere ist $\text{tr}(A') = \text{tr}(A)$ und $\det(A') = \det(A)$.

3.7.2 Direkte Summen

Sei V ein K -Vektorraum. Sei $k \geq 0$. Sei $U_i \subseteq V$ ein Unterraum für $1 \leq i \leq k$.

Definition. Sei M eine Menge. Seien $s, t \geq 0$.

Seien $(m_1, m_2, \dots, m_s) \in M^s$ und $(m'_1, m'_2, \dots, m'_t) \in M^t$ gegeben.

Wir schreiben

$$(m_1, m_2, \dots, m_s) \sqcup (m'_1, m'_2, \dots, m'_t) := (m_1, m_2, \dots, m_s, m'_1, m'_2, \dots, m'_t) \in M^{s+t}$$

für die *Konkatenation* dieser beiden Tupel.

Definition. Wir sagen, die Unterräume U_1, U_2, \dots, U_k bilden eine *direkte Summe*, falls der Schnitt je eines dieser Unterräume mit der Summe der anderen gleich 0 ist.

Um die Direktheit zu signalisieren, schreiben wir auch

$$U_1 + U_2 + \dots + U_k = U_1 \oplus U_2 \oplus \dots \oplus U_k = \bigoplus_{i=1}^k U_i .$$

Bemerkung. Ist B eine Basis von V und ist $B = B_1 \sqcup B_2 \sqcup \dots \sqcup B_k$, dann ist $V = \langle B_1 \rangle \oplus \langle B_2 \rangle \oplus \dots \oplus \langle B_k \rangle$.

Beispiel. Es ist $\mathbb{R}^{3 \times 1} = \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \oplus \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle \oplus \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \oplus \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle$.

Bemerkung. Sei V endlichdimensional.

Es mögen U_1, U_2, \dots, U_k eine direkte Summe bilden.

Sei B_i eine Basis von U_i für $1 \leq i \leq k$.

Dann ist $B_1 \sqcup B_2 \sqcup \dots \sqcup B_k$ eine Basis von $U_1 \oplus U_2 \oplus \dots \oplus U_k$.

3.7.3 Haupträume

Zerfalle $\chi_A(X) \in K[X]$. Seien $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von A . Mithin ist

$$\chi_A(X) = (-1)^n \cdot \prod_{i=1}^k (X - \lambda_i)^{\text{aV}_A(\lambda_i)} .$$

Für $1 \leq i \leq k$ schreiben wir

$$A_{(i)} := A - \lambda_i E_n .$$

Also wird der Eigenraum $E_A(\lambda_i) = \text{Kern}(A_{(i)})$.

Bemerkung. Es ist

$$\begin{aligned}\text{tr}(A) &= \sum_{i=1}^k \lambda_i \cdot \text{aV}_A(\lambda_i) \\ \det(A) &= \prod_{i=1}^k \lambda_i^{\text{aV}_A(\lambda_i)}.\end{aligned}$$

Bemerkung. Es ist

$$\text{Kern}(A_{(i)}) \subseteq \text{Kern}(A_{(i)}^2) \subseteq \text{Kern}(A_{(i)}^3) \subseteq \dots \subseteq \text{Kern}(A_{(i)}^{\text{aV}_A(\lambda_i)}) = \text{Kern}(A_{(i)}^{\text{aV}_A(\lambda_i)+1}) = \dots$$

Es gibt also ab der Stelle $\text{aV}_A(\lambda_i)$ nur noch Gleichheiten in dieser Kette.

In der Praxis werden wir dann werden eine Basis von $\text{Kern}(A_{(i)})$ bestimmen, diese zu einer Basis von $\text{Kern}(A_{(i)}^2)$ ergänzen, diese zu einer Basis von $\text{Kern}(A_{(i)}^3)$ ergänzen, usf. Der Prozeß endet, sobald nur noch Gleichheiten vorliegen.

Definition. Für $1 \leq i \leq k$ sei

$$H_A(\lambda_i) := \text{Kern}(A_{(i)}^{\text{aV}_A(\lambda_i)})$$

der *Hauptraum* von A zum Eigenwert λ_i .

Es ist also $E_A(\lambda_i) \subseteq H_A(\lambda_i)$.

Lemma.

- (1) Es ist $\dim H_A(\lambda_i) = \text{aV}_A(\lambda_i)$ für $1 \leq i \leq k$.
- (2) Es ist

$$K^{n \times 1} = \bigoplus_{i=1}^k H_A(\lambda_i)$$

Basen der Haupträume setzen sich also zu einer Basis von $K^{n \times 1}$ zusammen, per Konkatenation.

Bemerkung.

- (1) Sei $1 \leq i \leq k$. Genau dann ist $E_A(\lambda_i) = H_A(\lambda_i)$, wenn $\text{gV}_A(\lambda_i) = \text{aV}_A(\lambda_i)$ ist.
- (2) Genau dann ist $K^{n \times 1} = \bigoplus_{i=1}^k E_A(\lambda_i)$, wenn $\text{gV}_A(\lambda_i) = \text{aV}_A(\lambda_i)$ ist für $1 \leq i \leq k$. Diesenfalls sagen wir, es ist A *diagonalisierbar*.
- (3) Ist $\text{aV}_A(\lambda_i) = 1$ für $1 \leq i \leq k$, dann auch ist auch $\text{gV}_A(\lambda_i) = 1$ für $1 \leq i \leq k$. Also ist diesenfalls A diagonalisierbar. Ferner ist diesenfalls auch $k = n$.

Mit anderen Worten, ist $\chi_A(X) = (-1)^n \cdot (X - \lambda_1)^1 \cdot (X - \lambda_2)^1 \cdot \dots \cdot (X - \lambda_n)^1$ mit n paarweise verschiedenen Eigenwerten $\lambda_1, \lambda_2, \dots, \lambda_n$ von A , dann ist A diagonalisierbar.

Bemerkung. Sei A diagonalisierbar.

Sei $B = (e_1, \dots, e_n)$ die Standardbasis von $K^{n \times 1}$.

Sei C_i eine Basis von $E_A(\lambda_i)$ für $1 \leq i \leq k$. Sei $C := C_1 \sqcup \dots \sqcup C_k$. Es ist C eine Basis von $K^{n \times 1}$, bestehend aus Eigenvektoren von A ; vgl. vorstehende Bemerkung (2).

Sei $D = (d_{i,j})_{i,j} \in K^{n \times n}$ die Diagonalmatrix, für welche $d_{j,j}$ der Eigenwert zum j -ten Tupteleintrag von C ist, für $1 \leq j \leq n$.

Mit anderen Worten, die Diagonale von D besteht aus den Eigenwerten zu den Eigenvektoren in C , in Reihenfolge wie in C .

Da mult_A den j -ten Eintrag von C auf sein $d_{j,j}$ -faches abbildet, ist ${}_C(\text{mult}_A)_C = D$.

Es ist $S := {}_B \text{id}_C \in K^{n \times n}$ die Matrix mit Spaltentupel C .

Es wird

$$D = {}_C(\text{mult}_A)_C = {}_C \text{id}_B \cdot {}_B(\text{mult}_A)_B \cdot {}_B \text{id}_C = S^{-1} \cdot A \cdot S.$$

Zusammengefaßt erhalten wir also mit der Matrix S , welche in den Spalten unsere Basis aus Eigenvektoren stehen hat, und mit der Diagonalmatrix D , welche auf der Diagonalen die Eigenwerte passend zu den Spalten von S stehen hat,

$$S^{-1} \cdot A \cdot S = D.$$

Auf diese Weise haben wir A diagonalisiert.

Beispiel. Sei wieder $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Es war $E_A(1) = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$ und $E_A(-1) = \langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rangle$.

Mit $S := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ und $D := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ wird also

$$S^{-1} \cdot A \cdot S = D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Probe: Wir überprüfen $A \cdot S \stackrel{!}{=} S \cdot D$. Es ist $A \cdot S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$. Es ist $S \cdot D = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$. Das paßt.

Bemerkung.

- (1) Sei $B \in K^{n \times n}$. Sei $S \in K^{n \times n}$ invertierbar. Es haben B und SB dieselbe Zeilenstufenform.
- (2) Seien $B, B' \in K^{n \times n}$ mit derselben Zeilenstufenform gegeben. Sei $C \in K^{n \times n}$. Dann haben BC und $B'C$ dieselbe Zeilenstufenform.

Denn es gibt eine invertierbare Matrix $S \in K^{n \times n}$ mit $B' = SB$. Es folgt $B'C = SBC$. Also kann (1) angewandt werden.

Diese Aussage ist z.B. verwendbar, wenn B' die Zeilenstufenform von B ist.

- (3) Bei Zeilenstufenformen erlauben wir uns, Nullzeilen nicht zu schreiben.

Beispiel. Sei $A := \begin{pmatrix} 4 & 1 & 2 & -2 & -3 \\ 1 & 3 & 0 & -1 & -2 \\ 0 & -2 & 1 & 2 & 2 \\ 1 & 2 & 2 & 0 & -4 \\ 0 & -2 & -2 & 2 & 5 \end{pmatrix} \in \mathbb{Q}^{5 \times 5}$.

Es wird

$$\begin{aligned}
 \chi_A(X) &= \det \begin{pmatrix} 4-X & 1 & 2 & -2 & -3 \\ 1 & 3-X & 0 & -1 & -2 \\ 0 & -2 & 1-X & 2 & 2 \\ 1 & 2 & 2 & 0-X & -4 \\ 0 & -2 & -2 & 2 & 5-X \end{pmatrix} &= \det \begin{pmatrix} 4-X & 1 & 2 & -2 & -3 \\ 1 & 3-X & 0 & -1 & -2 \\ 0 & 0 & 3-X & 0 & X-3 \\ 1 & 0 & 0 & 2-X & 1-X \\ 0 & -2 & -2 & 2 & 5-X \end{pmatrix} \\
 &= (X-3) \cdot \det \begin{pmatrix} 4-X & 1 & 2 & -2 & -3 \\ 0 & 3-X & 0 & -1 & -2 \\ 1 & 0 & 0 & 2-X & 1-X \\ 0 & -2 & -2 & 2 & 5-X \end{pmatrix} &= (X-3) \cdot \det \begin{pmatrix} 4-X & 1 & 2 & -2 & -1 \\ 0 & 3-X & 0 & -1 & 0 \\ 1 & 0 & 0 & 2-X & 1-X \\ 0 & -2 & -2 & 2 & 3-X \end{pmatrix} \\
 &= -(X-3) \cdot \det \begin{pmatrix} 4-X & 1 & -2 & -1 \\ 1 & 3-X & -1 & -2 \\ 0 & 0 & 2-X & 1-X \\ 0 & -2 & 2 & 3-X \end{pmatrix} &= -(X-3) \cdot \det \begin{pmatrix} 4-X & 1 & -2 & -1 \\ 0 & 3-X & X-3 & X-3 \\ 1 & 0 & 2-X & 1-X \\ 0 & -2 & 2 & 3-X \end{pmatrix} \\
 &= (X-3)^2 \cdot \det \begin{pmatrix} 4-X & 1 & -2 & -1 \\ 0 & 1 & -1 & -1 \\ 1 & 0 & 2-X & 1-X \\ 0 & -2 & 2 & 3-X \end{pmatrix} &= (X-3)^2 \cdot \det \begin{pmatrix} 4-X & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 2-X & 1-X \\ 0 & -2 & 0 & 1-X \end{pmatrix} \\
 &= (X-3)^2(X-1) \cdot \det \begin{pmatrix} 4-X & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 2-X & -1 \\ 0 & -2 & 0 & -1 \end{pmatrix} &= (X-3)^2(X-1) \cdot \det \begin{pmatrix} 4-X & -1 & 0 \\ 0 & 2-X & -1 \\ 0 & 0 & -1 \end{pmatrix} \\
 &= -(X-3)^2(X-1) \cdot \det \begin{pmatrix} 4-X & -1 \\ 1 & 2-X \end{pmatrix} &= -(X-3)^2(X-1)((4-X)(2-X) + 1) \\
 &= -(X-3)^2(X-1)(X^2 - 6X + 9) &= -(X-3)^4(X-1).
 \end{aligned}$$

Folglich haben wir den Eigenwert $\lambda_1 = 1$ mit $\text{aV}_A(1) = 1$ und somit auch $\text{gV}_A(1) = 1$. Insbesondere ist $E_A(1) = H_A(1)$.

Ferner haben wir den Eigenwert $\lambda_2 = 3$ mit $\text{aV}_A(3) = 4$.

Probe: Es ist in der Tat die Spur $\text{tr}(A) = 4 + 3 + 1 + 0 + 5 = 13$ auch berechenbar als $\text{tr}(A) = \lambda_1 \cdot \text{aV}_A(\lambda_1) + \lambda_2 \cdot \text{aV}_A(\lambda_2) = 1 \cdot 1 + 3 \cdot 4 = 13$.

Wir wollen nun schrittweise ergänzte Basen der Haupträume bestimmen.

Zu $\lambda_1 = 1$. Es ist $A_{(1)} = A - \lambda_1 E_5 = \begin{pmatrix} 3 & 1 & 2 & -2 & -3 \\ 1 & 2 & 0 & -1 & -2 \\ 0 & -2 & 0 & 2 & 2 \\ 1 & 2 & 2 & -1 & -4 \\ 0 & -2 & -2 & 2 & 4 \end{pmatrix}$. Wir lösen:

$$\begin{pmatrix} 3 & 1 & 2 & -2 & -3 & 0 \\ 1 & 2 & 0 & -1 & -2 & 0 \\ 0 & -2 & 0 & 2 & 2 & 0 \\ 1 & 2 & 2 & -1 & -4 & 0 \\ 0 & -2 & -2 & 2 & 4 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & -1 & -2 & 0 \\ 0 & -5 & 2 & 1 & 3 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 2 & 0 & -2 & 0 \\ 0 & 1 & 1 & -1 & -2 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 2 & -4 & -2 & 0 \\ 0 & 0 & 2 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Somit ist $E_A(1) = \text{Kern}(A_{(1)}) = \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle = H_A(1)$.

Es ist

$$\left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

eine schrittweise ergänzte Basis von $H_A(1)$, was nur eben in bereits einem Schritt erreicht wurde.

Zu $\lambda_2 = 3$. Es ist $A_{(2)} = A - \lambda_2 E_5 = \begin{pmatrix} 1 & 1 & 2 & -2 & -3 \\ 1 & 0 & 0 & -1 & -2 \\ 0 & -2 & -2 & 2 & 2 \\ 1 & 2 & 2 & -3 & -4 \\ 0 & -2 & -2 & 2 & 2 \end{pmatrix}$.

Zur Berechnung einer Basis von $\text{Kern}(A_{(2)})$ lösen wir:

$$\begin{pmatrix} 1 & 1 & 2 & -2 & -3 & 0 \\ 1 & 0 & 0 & -1 & -2 & 0 \\ 0 & -2 & -2 & 2 & 2 & 0 \\ 1 & 2 & 2 & -3 & -4 & 0 \\ 0 & -2 & -2 & 2 & 2 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & -1 & -1 & 0 \\ 0 & -2 & -2 & 2 & 2 & 0 \\ 0 & 2 & 2 & -2 & -2 & 0 \\ 0 & -2 & -2 & 2 & 2 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -1 & -2 & 0 \\ 0 & 1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -1 & -2 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Somit ist $E_A(3) = \text{Kern}(A_{(2)}) = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle \subset H_A(3)$.

Mit $A' := \begin{pmatrix} 1 & 0 & 0 & -1 & -2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$ ist die Zeilenstufenform von $A_{(2)} \cdot A_{(2)}$ dieselbe wie die von

$$A' \cdot A_{(2)} = \begin{pmatrix} 1 & 0 & 0 & -1 & -2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 2 & -2 & -3 \\ 1 & 0 & 0 & -1 & -2 \\ 0 & -2 & -2 & 2 & 2 \\ 1 & 2 & 2 & -3 & -4 \\ 0 & -2 & -2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 4 & -3 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 2 & 2 \end{pmatrix}.$$

Zur Berechnung einer Ergänzung zu einer Basis von $\text{Kern}(A_{(2)}^2)$ lösen wir:

$$\begin{pmatrix} 0 & 3 & 4 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 2 & 2 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Wir ergänzen die im vorigen Schritt erhaltene Basis zu einer Basis von $\text{Kern}(A_{(2)}^2)$ und

bekommen $\text{Kern}(A_{(2)}^2) = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle \subset H_A(3)$.

Hierzu können die Vektoren aus dem Standardverfahren verwandt werden, die zu Nichtstufenspaltenpositionen gehören, die im vorigen Schritt noch Stufen enthielten. In diesem

Fall ist dies nur $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, gehörig zur Nichtstufenspaltenposition 1.

Mit $A'' := \begin{pmatrix} 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$ ist die Zeilenstufenform von $A_{(2)}^2 \cdot A_{(2)}$ dieselbe wie die von

$$A'' \cdot A_{(2)} = \begin{pmatrix} 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 2 & -2 & -3 \\ 1 & 0 & 0 & -1 & -2 \\ 0 & -2 & -2 & 2 & 2 \\ 1 & 2 & 2 & -3 & -4 \\ 0 & -2 & -2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 2 & 2 \end{pmatrix}.$$

Zur Berechnung einer Ergänzung zu einer Basis von $\text{Kern}(A_{(2)}^3)$ lösen wir:

$$(0 \ -2 \ -2 \ 2 \ 2 \ | \ 0) \rightsquigarrow (0 \ 1 \ 1 \ -1 \ -1 \ | \ 0).$$

Wir ergänzen die im vorigen Schritt erhaltene Basis zu einer Basis von $\text{Kern}(A_{(2)}^3)$ und bekommen

$$\text{Kern}(A_{(2)}^3) = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \rangle = H_A(3).$$

Hierzu können die Vektoren aus dem Standardverfahren verwandt werden, die zu Nichtstufenspaltenpositionen gehören, die im vorigen Schritt noch Stufen enthielten. In diesem

Fall ist dies nur $\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$, gehörig zur Nichtstufenspaltenposition 3.

Nun ist in der Tat die linke Seite $\text{Kern}(A_{(2)}^3)$ von Dimension 4, wie die rechte Seite $H_A(3)$. Kurz, die Dimension hat die algebraische Vielfachheit erreicht. Somit ist

$$\left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$$

eine schrittweise ergänzte Basis von $H_A(3)$.

3.7.4 Jordansche Normalform

Zerfalle $\chi_A(X) \in K[X]$. Seien $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von A . Für $1 \leq i \leq k$ schreiben wir $A_{(i)} := A - \lambda_i E_n$.

Bemerkung. Sei $V \subseteq K^{n \times 1}$ ein Unterraum.

Sei (v_1, \dots, v_a) ein linear unabhängiges Tupel in V . Sei $(v_1, \dots, v_a, x_1, \dots, x_c)$ eine Basis von V .

Seien (y_1, \dots, y_b) ein Tupel von Vektoren in V derart, daß $(v_1, \dots, v_a, y_1, \dots, y_b)$ linear unabhängig ist. Insbesondere ist $0 \leq b \leq c$.

Dann können wir zu einer Basis $(v_1, \dots, v_a, y_1, \dots, y_b, y_{b+1}, \dots, y_c)$ ergänzen mit einer Auswahl (y_{b+1}, \dots, y_c) von (x_1, \dots, x_c) .

Um dies zu erreichen, muß nur sichergestellt sein, daß jeder dabei nicht ausgewählte Vektor von (x_1, \dots, x_c) als Linearkombination in $(v_1, \dots, v_a, y_1, \dots, y_b, y_{b+1}, \dots, y_c)$ geschrieben werden kann.

Konstruktion. Wir wollen nun eine geeignete Basis von $H_A(\lambda_i)$ bestimmen.

Schritt 1. Sei $(x_{1,1}, \dots, x_{1,t_1})$ eine Basis von $\text{Kern}(A_{(i)}^1)$.

Ergänze diese zu einer Basis $(x_{1,1}, \dots, x_{1,t_1}, x_{2,1}, \dots, x_{2,t_2})$ von $\text{Kern}(A_{(i)}^2)$.

Ergänze diese zu einer Basis $(x_{1,1}, \dots, x_{1,t_1}, x_{2,1}, \dots, x_{2,t_2}, x_{3,1}, \dots, x_{3,t_3})$ von $\text{Kern}(A_{(i)}^3)$.

So fahren wir fort.

Wir erhalten schließlich eine Basis

$$(x_{1,1}, \dots, x_{1,t_1}, x_{2,1}, \dots, x_{2,t_2}, x_{3,1}, \dots, x_{3,t_3}, \dots, x_{\ell,1}, \dots, x_{\ell,t_\ell})$$

von $H_A(\lambda_i) = \text{Kern}(A_{(i)}^\ell)$, für das aus der Konstruktion resultierende ℓ .

Das ist leider i.a. noch keine geeignete Basis. Wir müssen diese Konstruktion noch von hinten aufrollen.

Schritt 2. Setze $u_\ell := t_\ell$ und $(y_{\ell,1}, \dots, y_{\ell,u_\ell}) := (x_{\ell,1}, \dots, x_{\ell,t_\ell})$.

Ergänze mit einer Auswahl $(y_{\ell-1,1}, \dots, y_{\ell-1,u_{\ell-1}})$ aus $(x_{\ell-1,1}, \dots, x_{\ell-1,t_{\ell-1}})$ zu einer neuen Basis

$$(x_{1,1}, \dots, x_{1,t_1}, \dots, x_{\ell-2,1}, \dots, x_{\ell-2,t_{\ell-2}}, \\ A_{(i)}^1 y_{\ell,1}, \dots, A_{(i)}^1 y_{\ell,u_\ell}, y_{\ell-1,1}, \dots, y_{\ell-1,u_{\ell-1}})$$

von $\text{Kern}(A_{(i)}^{\ell-1})$; vgl. vorstehende Bemerkung.

Ergänze mit einer Auswahl $(y_{\ell-2,1}, \dots, y_{\ell-2,u_{\ell-2}})$ aus $(x_{\ell-2,1}, \dots, x_{\ell-2,t_{\ell-2}})$ zu einer neuen Basis

$$(x_{1,1}, \dots, x_{1,t_1}, \dots, x_{\ell-3,1}, \dots, x_{\ell-3,t_{\ell-3}}, \\ A_{(i)}^2 y_{\ell,1}, \dots, A_{(i)}^2 y_{\ell,u_\ell}, A_{(i)}^1 y_{\ell-1,1}, \dots, A_{(i)}^1 y_{\ell-1,u_{\ell-1}}, y_{\ell-2,1}, \dots, y_{\ell-2,u_{\ell-2}})$$

von $\text{Kern}(A_{(i)}^{\ell-2})$; vgl. vorstehende Bemerkung.

Ergänze mit einer Auswahl $(y_{\ell-3,1}, \dots, y_{\ell-3,u_{\ell-3}})$ aus $(x_{\ell-3,1}, \dots, x_{\ell-3,t_{\ell-3}})$ zu einer neuen Basis

$$(x_{1,1}, \dots, x_{1,t_1}, \dots, x_{\ell-4,1}, \dots, x_{\ell-4,t_{\ell-4}}, \\ A_{(i)}^3 y_{\ell,1}, \dots, A_{(i)}^3 y_{\ell,u_\ell}, A_{(i)}^2 y_{\ell-1,1}, \dots, A_{(i)}^2 y_{\ell-1,u_{\ell-1}}, A_{(i)}^1 y_{\ell-2,1}, \dots, A_{(i)}^1 y_{\ell-2,u_{\ell-2}}, y_{\ell-3,1}, \dots, y_{\ell-3,u_{\ell-3}})$$

von $\text{Kern}(A_{(i)}^{\ell-3})$; vgl. vorstehende Bemerkung.

So fahren wir fort, bis eine neue Basis

$$(A_{(i)}^{\ell-1} y_{\ell,1}, \dots, A_{(i)}^{\ell-1} y_{\ell,u_\ell}, A_{(i)}^{\ell-2} y_{\ell-1,1}, \dots, A_{(i)}^{\ell-2} y_{\ell-1,u_{\ell-1}}, \dots, A_{(i)}^0 y_{1,1}, \dots, A_{(i)}^0 y_{1,u_1})$$

von $\text{Kern}(A_{(i)}^1)$ konstruiert ist.

Nun wird noch sortiert.

Definition. Sei $1 \leq m \leq \ell$. Sei $1 \leq j \leq u_m$.

Das aus vorstehender Konstruktion resultierende Tupel

$$(A_{(i)}^{m-1} y_{m,j}, \dots, A_{(i)}^1 y_{m,j}, A_{(i)}^0 y_{m,j})$$

heißt *Hauptvektorkette* der Länge m .

Die aus der Konstruktion resultierenden Hauptvektorketten in $H_A(\lambda_i)$ geben per Konkatination nebeneinandergesetzt eine Basis von $H_A(\lambda_i)$, für $1 \leq i \leq k$.

Da $K^{n \times 1} = \bigoplus_{i=1}^k H_A(\lambda_i)$, können diese Basen per Konkatination nebeneinandergesetzt werden zu einer Basis von $K^{n \times 1}$.

Definition. Sei $\lambda \in K$. Sei $m \geq 1$. Sei

$$J_m(\lambda) := \lambda \cdot E_m + \sum_{i=1}^{m-1} e_{i,i+1} = \begin{pmatrix} \lambda & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda \end{pmatrix} \in K^{m \times m}$$

der *Jordanblock* der Größe $m \times m$ zum Eigenwert λ .

Satz. Sei das Tupel C_i die Konkatenation aller in der Konstruktion aufgetretenen Hauptvektorketten im Hauptraum $H_A(\lambda_i)$ in einer gewählten Reihenfolge, für $1 \leq i \leq k$. Es ist C_i eine Basis von $H_A(\lambda_i)$.

Sei $(m_{i,1}, \dots, m_{i,s_i})$ das Tupel der Längen dieser Hauptvektorketten, in der gewählten Reihenfolge der Konkatenation.

Sei $C := C_1 \sqcup C_2 \sqcup \dots \sqcup C_k$.

Es ist C eine Basis von $K^{n \times 1}$, genannt eine *Jordanbasis* von $K^{n \times 1}$ für A .

Sei $S \in K^{n \times n}$ die Matrix mit Spaltentupel C .

Sei $J \in K^{n \times n}$ die Blockdiagonalmatrix mit den Blöcken

$$J_{m_{1,1}}(\lambda_1), \dots, J_{m_{1,s_1}}(\lambda_1), \quad J_{m_{2,1}}(\lambda_2), \dots, J_{m_{2,s_2}}(\lambda_2), \quad \dots \quad , J_{m_{k,1}}(\lambda_k), \dots, J_{m_{k,s_k}}(\lambda_k) .$$

Also

$$J = \begin{pmatrix} \boxed{J_{m_{1,1}}(\lambda_1)} & & & & & & & & & & \\ & \boxed{J_{m_{1,2}}(\lambda_1)} & & & & & & & & & \\ & & & & & & & & & & \\ & & & \dots & & & & & & & \\ & & & & & \boxed{J_{m_{k,s_{k-1}}}(\lambda_k)} & & & & & \\ & & & & & & & \boxed{J_{m_{k,s_k}}(\lambda_k)} & & & \end{pmatrix} \in K^{n \times n} .$$

Dann ist $S^{-1}AS = J$.

Es heißt J eine *Jordansche Normalform* oder *Jordanform* von A .

Bemerkung. Ist A diagonalisierbar, dann ist das Spaltentupel der im Satz verwandten Matrix S eine Basis aus Eigenvektoren. Ferner ist die Jordansche Normalform von A dann eine Diagonalmatrix mit den Eigenwerten von A auf der Diagonalen, in der Reihenfolge passend zu den Eigenvektoren in S .

Beispiel. Wir setzen das letzte Beispiel von §3.7.3 fort. Es hatte

$$A = \begin{pmatrix} 4 & 1 & 2 & -2 & -3 \\ 1 & 3 & 0 & -1 & -2 \\ 0 & -2 & 1 & 2 & 2 \\ 1 & 2 & 2 & 0 & -4 \\ 0 & -2 & -2 & 2 & 5 \end{pmatrix} \in \mathbb{Q}^{5 \times 5}$$

den Eigenwert $\lambda_1 = 1$ mit $aV_A(1) = 1$ und den Eigenwert $\lambda_2 = 3$ mit $aV_A(3) = 4$.

Wir wollen den Hauptraum beim Eigenwert $\boxed{\lambda_1 = 1}$ mit einer Basis aus Hauptvektorketten ausstatten:

Es war $H_A(1) = E_A(1)$.

Eine schrittweise ergänzte Basis von $H_A(1)$ war gegeben durch

$$\left(\underbrace{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}}_{=: x_{1,1} =: y_{1,1}} \right)$$

Hier waren wir schon nach dem ersten Schritt fertig.

Wir erhalten die aus einer Hauptvektorkette von Länge 1 bestehende Basis von $H_A(1)$:

$$\left(\underbrace{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}} \right)$$

von $H_A(1)$.

Wir wollen den Hauptraum beim Eigenwert $\lambda_2 = 3$ mit einer Basis aus Hauptvektorketten ausstatten:

$$\text{Es ist } A_{(2)} = A - 3E_5 = \begin{pmatrix} 1 & 1 & 2 & -2 & -3 \\ 1 & 0 & 0 & -1 & -2 \\ 0 & -2 & -2 & 2 & 2 \\ 1 & 2 & 2 & -3 & -4 \\ 0 & -2 & -2 & 2 & 2 \end{pmatrix}.$$

Eine schrittweise ergänzte Basis von $H_A(3) = \text{Kern}(A_{(2)}^3)$ war gegeben durch

$$\left(\underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{=: x_{1,1}}, \underbrace{\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{=: x_{1,2}}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{2,1}}, \underbrace{\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{3,1}} \right).$$

$$\text{Wir setzen } y_{3,1} := x_{3,1} = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

$$\text{Es ist } A_{(2)} y_{3,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Wir ergänzen $(A_{(2)} y_{3,1})$ mit der leeren Auswahl von $(x_{2,1})$ zur Basis

$$(x_{1,1}, x_{1,2}, A_{(2)} y_{3,1}) = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

von $\text{Kern}(A_{(2)}^2)$.

$$\text{Es ist } A_{(2)}^2 y_{3,1} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Wir ergänzen $(A_{(2)}^2 y_{3,1})$ mit der Auswahl $(y_{1,1}) := (x_{1,2})$ von $(x_{1,1}, x_{1,2})$ zur Basis

$$(A_{(2)}^2 y_{3,1}, y_{1,1}) = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

von $\text{Kern}(A_{(2)}^1)$.

Somit haben wir die Hauptvektorketten $(A_{(2)}^2 y_{3,1}, A_{(2)} y_{3,1}, y_{3,1})$ und $(y_{1,1})$.

Wir erhalten die aus einer Hauptvektorkette von Länge 3 und aus einer Hauptvektorkette von Länge 1 bestehende Basis von $H_A(3)$:

$$\left(\underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}}_{\text{Länge 3}}, \underbrace{\begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}}_{\text{Länge 1}} \right)$$

Mit

$$S := \begin{pmatrix} 0 & 1 & 1 & 0 & 2 \\ 1 & 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad J := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

ist also $S^{-1}AS = J$ eine Jordansche Normalform von A .

Als Probe überprüfen wir $AS = SJ$.

Beispiel. Sei $K = \mathbb{F}_4$. Wir schreiben in diesem Beispiel α^2 statt $\alpha + 1$. Sei

$$A := \begin{pmatrix} \alpha^2 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ \alpha & \alpha & 0 & 0 & \alpha & 0 & \alpha \\ 0 & 1 & 0 & \alpha & \alpha & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & \alpha & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha \end{pmatrix} \in \mathbb{F}_4^{7 \times 7}.$$

Wir berechnen das charakteristische Polynom.

$$\begin{aligned} \chi_A(X) &= \det \begin{pmatrix} \alpha^2 - X & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 - X & 1 & 0 & 1 & 0 & 1 \\ \alpha & \alpha & 0 - X & 0 & \alpha & 0 & \alpha \\ 0 & 1 & 0 & \alpha - X & \alpha & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & \alpha - X & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha - X & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha - X \end{pmatrix} = \det \begin{pmatrix} \alpha^2 - X & 0 & 1 & 0 \\ 1 & 1 - X & 1 & 0 \\ \alpha & \alpha & 0 - X & 0 \\ 0 & 1 & 0 & \alpha - X \end{pmatrix} \cdot \det \begin{pmatrix} \alpha - X & \alpha^2 & 0 \\ 0 & \alpha - X & 0 \\ 0 & \alpha^2 & \alpha - X \end{pmatrix} \\ &= (\alpha - X) \cdot \det \begin{pmatrix} \alpha^2 - X & 0 & 1 \\ 1 & 1 - X & 1 \\ \alpha & \alpha & 0 - X \end{pmatrix} \cdot (\alpha - X) \cdot \det \begin{pmatrix} \alpha - X & 0 \\ \alpha^2 & \alpha - X \end{pmatrix} \\ &= (\alpha - X)^4 \cdot \det \begin{pmatrix} \alpha^2 - X & 0 & 1 \\ 1 & 1 - X & 1 \\ \alpha & \alpha & 0 - X \end{pmatrix} = (X - \alpha)^4 \cdot \det \begin{pmatrix} 0 & 0 & 1 \\ X + \alpha & X + 1 & 1 \\ X^2 + \alpha^2 X + \alpha & \alpha & X \end{pmatrix} \\ &= (X - \alpha)^4 \cdot \det \begin{pmatrix} X + \alpha & X + 1 \\ X^2 + \alpha^2 X + \alpha & \alpha \end{pmatrix} = (X - \alpha)^4 \cdot \det \begin{pmatrix} 1 + \alpha & X + 1 \\ X^2 + \alpha^2 X & \alpha \end{pmatrix} \\ &= (X - \alpha)^4 \cdot (1 + X^3 + \alpha^2 X^2 + X^2 + \alpha^2 X) = (X - \alpha)^4 \cdot (X^3 + \alpha X^2 + \alpha^2 X + 1) = (X - \alpha)^7. \end{aligned}$$

Wir haben also den einzigen Eigenwert $\lambda_1 = \alpha$ mit der algebraischen Vielfachheit $aV_A(\alpha) = 7$.

Wir wollen den Hauptraum beim Eigenwert $\lambda_1 = \alpha$ mit einer Basis aus Hauptvektorketten ausstatten:

Es ist

$$A_{(1)} := \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 \\ \alpha & \alpha & \alpha & 0 & \alpha & 0 & \alpha \\ 0 & 1 & 0 & 0 & \alpha & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 \end{pmatrix}.$$

Wir dürfen bei der Berechnung von $\text{Kern}(A_{(1)}^i)$ aufhören, sobald $\dim \text{Kern}(A_{(1)}^i) = \dim H_A(\alpha) = \text{aV}_A(\alpha) = 7$ ist.

Zur Berechnung einer Basis von $\text{Kern}(A_{(1)})$ lösen wir:

$$\left(\begin{array}{ccccccc|c} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 & 0 \\ \alpha & \alpha & \alpha & 0 & \alpha & 0 & \alpha & 0 \\ 0 & 1 & 0 & 0 & \alpha & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccccc|c} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \alpha & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccccc|c} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Es wird

$$E_A(\alpha) = \text{Kern}(A_{(1)}) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \subset H_A(\alpha).$$

Mit $A' := \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$ ist $\text{Kern}(A_{(1)} \cdot A_{(1)}) = \text{Kern}(A' \cdot A_{(1)})$.

Es ist $A' \cdot A_{(1)} = \begin{pmatrix} \alpha^2 & \alpha & \alpha^2 & 0 & \alpha^2 & \alpha^2 & \alpha^2 \\ 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Zur Berechnung einer Ergänzung zu einer Basis von $\text{Kern}(A_{(1)}^2)$ lösen wir:

$$\left(\begin{array}{ccccccc|c} \alpha^2 & \alpha & \alpha^2 & 0 & \alpha^2 & \alpha^2 & \alpha^2 & 0 \\ 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccccc|c} 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccccccc|c} 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Das Standardverfahren, angewandt auf die neue Nichtstufenspaltenpositionen 2 und 5, gibt nun zwei ergänzende Vektoren zu

$$\text{Kern}(A_{(1)}^2) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \alpha^2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \subset H_A(\alpha).$$

Mit $A'' := \begin{pmatrix} 1 & \alpha^2 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$ ist $\text{Kern}(A_{(1)}^2 \cdot A_{(1)}) = \text{Kern}(A'' \cdot A_{(1)})$.

Es ist $A'' \cdot A_{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Die Umformung in Zeilenstufenform entfällt.

Das Standardverfahren, angewandt auf die neue Nichtstufenspaltenpositionen 1 und 6, gibt nun zwei ergänzende Vektoren zu

$$\text{Kern}(A_{(1)}^3) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \alpha^2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle = H_A(\alpha).$$

Als schrittweise ergänzte Basis von $H_A(\alpha)$ haben wir also erhalten :

$$\left(\underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{1,1}}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{1,2}}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{=: x_{1,3}}, \underbrace{\begin{pmatrix} \alpha^2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{2,1}}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{2,2}}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{=: x_{3,1}}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{=: x_{3,2}} \right)$$

Wir setzen $y_{3,1} := x_{3,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ und $y_{3,2} := x_{3,2} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

Es ist $A_{(1)}y_{3,1} = \begin{pmatrix} 1 \\ 1 \\ \alpha \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ und $A_{(1)}y_{3,2} = \begin{pmatrix} 0 \\ 0 \\ \alpha^2 \\ \alpha^2 \\ 0 \\ 0 \\ \alpha^2 \end{pmatrix}$.

Wir ergänzen $(x_{1,1}, x_{1,2}, x_{1,3}, A_{(1)}y_{3,1}, A_{(1)}y_{3,2})$ mit der leeren Auswahl von $(x_{2,1}, x_{2,2})$ zur Basis

$$(x_{1,1}, x_{1,2}, x_{1,3}, A_{(1)}y_{3,1}, A_{(1)}y_{3,2}) = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \alpha \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \alpha^2 \\ \alpha^2 \\ 0 \\ 0 \\ \alpha^2 \end{pmatrix} \right)$$

von $\text{Kern}(A_{(1)}^2)$.

Es ist $A_{(1)}^2y_{3,1} = \begin{pmatrix} \alpha^2 \\ 0 \\ \alpha^2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ und $A_{(1)}^2y_{3,2} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Wir ergänzen $(A_{(1)}^2y_{3,1}, A_{(1)}^2y_{3,2})$ mit der Auswahl $(y_{1,1}) := (x_{1,3})$ von $(x_{1,1}, x_{1,2}, x_{1,3})$ zur Basis

$$(A_{(1)}^2y_{3,1}, A_{(1)}^2y_{3,2}, y_{1,1}) = \left(\begin{pmatrix} \alpha^2 \\ 0 \\ \alpha^2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

von $\text{Kern}(A_{(1)}^1)$.

Somit haben wir die Hauptvektorketten $(A_{(1)}^2 y_{3,1}, A_{(1)} y_{3,1}, y_{3,1})$, $(A_{(1)}^2 y_{3,2}, A_{(1)} y_{3,2}, y_{3,2})$ und $(y_{1,1})$.

Wir erhalten die aus zwei Hauptvektorketten von Länge 3 und aus einer Hauptvektorkette von Länge 1 bestehende Basis von $H_A(\alpha)$:

$$\left(\underbrace{\begin{pmatrix} \alpha^2 \\ 0 \\ \alpha^2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \alpha \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\text{Länge 3}}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \alpha^2 \\ 0 \\ 0 \\ \alpha^2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{\text{Länge 3}}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{\text{Länge 1}} \right)$$

Mit

$$S := \begin{pmatrix} \alpha^2 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \alpha^2 & \alpha & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \alpha^2 & 0 & 1 \end{pmatrix} \quad \text{und} \quad J := \begin{pmatrix} \alpha & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha \end{pmatrix}$$

ist also $S^{-1}AS = J$ eine Jordansche Normalform von A .

Als Probe überprüfen wir $AS = SJ$.

3.8 Matrizen über \mathbb{R} und über \mathbb{C}

3.8.1 Gram-Schmidt

3.8.1.1 Gram-Schmidt über \mathbb{R}

Sei $n \geq 0$. Sei $V \subseteq \mathbb{R}^{n \times 1}$ ein Unterraum. Sei $m := \dim_{\mathbb{R}}(V)$.

Definition. Ein Tupel (d_1, \dots, d_m) mit $d_j \in V \subseteq \mathbb{R}^{n \times 1}$ für $1 \leq j \leq m$ heißt *Orthonormalbasis* von V , falls

$$d_j^t \cdot d_k = \delta_{j,k}$$

ist für $1 \leq j \leq m$ und $1 \leq k \leq m$.

Mit anderen Worten, es sollten alle Vektoren Länge 1 haben und paarweise aufeinander senkrecht stehen.

Eine Orthonormalbasis von V ist insbesondere eine Basis von V . Dazu genügt es, lineare Unabhängigkeit zu wissen. In der Tat, ist $\sum_{j=1}^m \mu_j d_j = 0$ mit $\mu_j \in \mathbb{R}$ für $1 \leq j \leq m$, dann folgt $0 = d_k^t \cdot (\sum_{j=1}^m \mu_j d_j) = \sum_{j=1}^m \mu_j \delta_{j,k} = \mu_k$ für $1 \leq k \leq m$.

Beispiel. Sei $V = \mathbb{R}^{3 \times 1}$ der gesamte Raum.

Es sind $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$ und $\left(\frac{1}{\sqrt{3}} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}\right)$ Orthonormalbasen von $\mathbb{R}^{3 \times 1}$.

Konstruktion (Gram-Schmidt). Sei (c_1, \dots, c_m) eine Basis von V .

Wir konstruieren schrittweise eine Orthonormalbasis von V .

Sei $d'_1 := c_1$.

Sei $d_1 := \|d'_1\|^{-1} d'_1$.

Sei $d'_2 := c_2 - (d_1^t \cdot c_2) \cdot d_1$.

Sei $d_2 := \|d'_2\|^{-1} d'_2$.

Sei $d'_3 := c_3 - (d_1^t \cdot c_3) \cdot d_1 - (d_2^t \cdot c_3) \cdot d_2$.

Sei $d_3 := \|d'_3\|^{-1} d'_3$.

Sei $d'_4 := c_4 - (d_1^t \cdot c_4) \cdot d_1 - (d_2^t \cdot c_4) \cdot d_2 - (d_3^t \cdot c_4) \cdot d_3$.

Sei $d_4 := \|d'_4\|^{-1} d'_4$.

So setzen wir fort, bis d_m konstruiert ist.

Das Gram-Schmidt-Verfahren stammt von Laplace und Cauchy.

Lemma. *Es ist (d_1, \dots, d_m) eine Orthonormalbasis von V .*

Denn zum einen ist wegen linearer Unabhängigkeit von (c_1, \dots, c_m) stets $d'_k \neq 0$, sodaß die Konstruktion durchführbar ist.

Zum anderen ist $d_k^t \cdot d_k = 1$ nach Konstruktion.

Wir zeigen noch $d_j^t \cdot d_k \stackrel{!}{=} 0$ für $1 \leq j < k \leq m$ per Induktion nach k . Es genügt, $d_j^t \cdot d'_k \stackrel{!}{=} 0$ zu zeigen. In der Tat wird

$$d_j^t \cdot d'_k = d_j^t \cdot \left(c_k - \sum_{\ell=1}^{k-1} (d_\ell^t \cdot c_k) \cdot d_\ell \right) \stackrel{\text{IV}}{=} d_j^t \cdot c_k - \sum_{\ell=1}^{k-1} (d_\ell^t \cdot c_k) \cdot d_{j,\ell} = d_j^t \cdot c_k - d_j^t \cdot c_k = 0.$$

Bemerkung. Bei der Berechnung beachte man, daß $\|\mu x\|^{-1} \mu x = \|x\|^{-1} x$ ist für $x \in \mathbb{R}^{n \times 1} \setminus \{0\}$ und $\mu \in \mathbb{R}_{>0}$. In diesem Sinne dürfen positive Vorfaktoren, die bei d'_j auftreten, bei der Berechnung von d_j also unterschlagen werden.

Beispiel. Wir suchen eine Orthonormalbasis von $V := \mathbb{R} \left(\underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}}_{=: c_1}, \underbrace{\begin{pmatrix} 2 \\ 1 \\ 3 \\ 0 \end{pmatrix}}_{=: c_2}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}}_{=: c_3} \right) \subseteq \mathbb{R}^{4 \times 1}$.

Es wird $d'_1 = c_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$.

Es wird $d_1 = \|d'_1\|^{-1} d'_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$.

$$\text{Es wird } d'_2 = c_2 - (d_1^t \cdot c_2) \cdot d_1 = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 0 \end{pmatrix} - \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}^t \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \\ 0 \end{pmatrix}\right) \cdot \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 0 \end{pmatrix} - \frac{5}{3} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 3 \\ 4 \\ -5 \end{pmatrix}.$$

Zwischenprobe: $d_1^t d'_2 = 0$.

$$\text{Es wird } d_2 = \|d'_2\|^{-1} d'_2 = \frac{1}{\sqrt{51}} \begin{pmatrix} 1 \\ 3 \\ 4 \\ -5 \end{pmatrix}.$$

Es wird

$$\begin{aligned} d'_3 &= c_3 - (d_1^t \cdot c_3) \cdot d_1 - (d_2^t \cdot c_3) \cdot d_2 \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} - \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}^t \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}\right) \cdot \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \left(\frac{1}{\sqrt{51}} \begin{pmatrix} 1 \\ 3 \\ 4 \\ -5 \end{pmatrix}^t \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}\right) \cdot \frac{1}{\sqrt{51}} \begin{pmatrix} 1 \\ 3 \\ 4 \\ -5 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{17} \begin{pmatrix} 1 \\ 3 \\ 4 \\ -5 \end{pmatrix} = \frac{4}{17} \begin{pmatrix} -4 \\ 5 \\ 1 \\ 3 \end{pmatrix}. \end{aligned}$$

Zwischenprobe: $d_1^t d'_3 = 0$ und $d_2^t d'_3 = 0$.

$$\text{Es wird } d_3 = \|d'_3\|^{-1} d'_3 = \frac{1}{\sqrt{51}} \begin{pmatrix} -4 \\ 5 \\ 1 \\ 3 \end{pmatrix}.$$

Wir erhalten die Orthonormalbasis $(d_1, d_2, d_3) = \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{51}} \begin{pmatrix} 1 \\ 3 \\ 4 \\ -5 \end{pmatrix}, \frac{1}{\sqrt{51}} \begin{pmatrix} -4 \\ 5 \\ 1 \\ 3 \end{pmatrix}\right)$ von V .

Bemerkung. Sind $c_1, c_2 \in \mathbb{R}^{3 \times 1}$ gegeben mit $\|c_1\| = 1$, $\|c_2\| = 1$ und $c_1^t c_2 = 0$, dann ist $(c_1, c_2, c_1 \times c_2)$ eine Orthonormalbasis von $\mathbb{R}^{3 \times 1}$.

3.8.1.2 Hermitesches Skalarprodukt und Gram-Schmidt über \mathbb{C}

Sei $n \geq 0$. Sei $V \subseteq \mathbb{C}^{n \times 1}$ ein Unterraum. Sei $m := \dim_{\mathbb{C}}(V)$.

Definition. Seien $s, t \geq 0$. Sei $A = (a_{j,k})_{j,k} \in \mathbb{C}^{s \times t}$ gegeben. Sei

$$\bar{A} := (\bar{a}_{j,k})_{j,k} \in \mathbb{C}^{s \times t}$$

die zu A komplex konjugierte Matrix.

Definition. Seien $x = (x_j)_j, y = (y_j)_j \in \mathbb{C}^{n \times 1}$.

Das Produkt $\bar{x}^t y = \bar{x}^t \cdot y \in \mathbb{C}$ heißt *hermitesches Skalarprodukt* von x und y .

Es ist $\bar{x}^t y = \sum_{j=1}^n \bar{x}_j y_j = \bar{x}_1 y_1 + \dots + \bar{x}_n y_n = \overline{y^t x}$.

Die *Norm* oder *Länge* von x ist als

$$\|x\| := (\bar{x}^t \cdot x)^{1/2} = (\bar{x}_1 \cdot x_1 + \dots + \bar{x}_n \cdot x_n)^{1/2} = (|x_1|^2 + \dots + |x_n|^2)^{1/2}$$

definiert.

Definition. Ein Tupel (d_1, \dots, d_m) mit $d_j \in V \subseteq \mathbb{C}^{n \times 1}$ für $1 \leq j \leq m$ heißt *Orthonormalbasis* von V , falls

$$\bar{d}_j^\dagger \cdot d_k = \delta_{j,k}$$

ist für $1 \leq j \leq m$ und $1 \leq k \leq m$.

Eine Orthonormalbasis von V ist insbesondere eine Basis von V .

Konstruktion (Gram-Schmidt). Sei (c_1, \dots, c_m) eine Basis von V .

Wir konstruieren schrittweise eine Orthonormalbasis von V .

Sei $d'_1 := c_1$.

Sei $d_1 := \|d'_1\|^{-1} d'_1$.

Sei $d'_2 := c_2 - (\bar{d}_1^\dagger \cdot c_2) \cdot d_1$.

Sei $d_2 := \|d'_2\|^{-1} d'_2$.

Sei $d'_3 := c_3 - (\bar{d}_1^\dagger \cdot c_3) \cdot d_1 - (\bar{d}_2^\dagger \cdot c_3) \cdot d_2$.

Sei $d_3 := \|d'_3\|^{-1} d'_3$.

Sei $d'_4 := c_4 - (\bar{d}_1^\dagger \cdot c_4) \cdot d_1 - (\bar{d}_2^\dagger \cdot c_4) \cdot d_2 - (\bar{d}_3^\dagger \cdot c_4) \cdot d_3$.

Sei $d_4 := \|d'_4\|^{-1} d'_4$.

So setzen wir fort, bis d_m konstruiert ist.

Lemma. Es ist (d_1, \dots, d_m) eine Orthonormalbasis von V .

Bemerkung. Bei der Berechnung beachte man, daß $\|\mu x\|^{-1} \mu x = \|x\|^{-1} x$ ist für $x \in \mathbb{C}^{n \times 1} \setminus \{0\}$ und $\mu \in \mathbb{R}_{>0}$. In diesem Sinne dürfen positive reelle Vorfaktoren, die bei d'_j auftreten, bei der Berechnung von d_j also unterschlagen werden.

Beispiel. Wir konstruieren eine Orthonormalbasis von $V := \mathbb{C} \left\langle \underbrace{\begin{pmatrix} 1+i \\ 2 \\ i \end{pmatrix}}_{=: c_1}, \underbrace{\begin{pmatrix} i \\ 2 \\ 1 \end{pmatrix}}_{=: c_2} \right\rangle \subseteq \mathbb{C}^{3 \times 1}$.

Es wird $d'_1 = c_1$ und $d_1 = \|d'_1\|^{-1} d'_1 = \frac{1}{\sqrt{7}} \begin{pmatrix} 1+i \\ 2 \\ i \end{pmatrix}$.

Es wird

$$d'_2 = c_2 - (\bar{d}_1^\dagger \cdot c_2) \cdot d_1 = \begin{pmatrix} i \\ 2 \\ 1 \end{pmatrix} - \left(\frac{1}{\sqrt{7}} \begin{pmatrix} 1-i \\ 2 \\ -i \end{pmatrix}^\dagger \cdot \begin{pmatrix} i \\ 2 \\ 1 \end{pmatrix} \right) \cdot \frac{1}{\sqrt{7}} \begin{pmatrix} 1+i \\ 2 \\ i \end{pmatrix} = \begin{pmatrix} i \\ 2 \\ 1 \end{pmatrix} - \frac{1+2i}{7} \begin{pmatrix} 1+i \\ 2 \\ i \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 1+4i \\ -2+3i \\ 9-i \end{pmatrix}.$$

Zwischenprobe: $\bar{d}_1^\dagger d'_2 = \frac{1}{7\sqrt{7}} ((1-i)(1+4i) + 2(-2+3i) + (-i)(9-i)) = 0$.

Es wird $d_2 = \frac{1}{4\sqrt{7}} \begin{pmatrix} 1+4i \\ -2+3i \\ 9-i \end{pmatrix}$.

Wir erhalten die Orthonormalbasis $(d_1, d_2) = \left(\frac{1}{\sqrt{7}} \begin{pmatrix} 1+i \\ 2 \\ i \end{pmatrix}, \frac{1}{4\sqrt{7}} \begin{pmatrix} 1+4i \\ -2+3i \\ 9-i \end{pmatrix} \right)$ von V .

3.8.2 Symmetrische und orthogonale Matrizen über \mathbb{R}

Sei $n \geq 0$. Sei $A = (a_{j,k})_{j,k} \in \mathbb{R}^{n \times n}$.

Definition. Es heißt A *symmetrisch*, falls $A = A^t$ ist.

Mit anderen Worten, A ist symmetrisch, falls $a_{j,k} = a_{k,j}$ ist für $1 \leq j \leq n$ und $1 \leq k \leq n$.

Beispiel. Es ist $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 3 & 4 \\ 2 & 4 & 9 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$ symmetrisch.

Bemerkung. Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, dann ist A diagonalisierbar.

Definition. Es heißt A *orthogonal* oder *Orthogonalmatrix*, falls $A^t \cdot A = E_n$ ist.

Dann ist $A^{-1} = A^t$ und also auch $A \cdot A^t = E_n$.

Bemerkung. Sei für $1 \leq k \leq n$ die k -te Spalte von A mit $x_k := (a_{j,k})_j \in \mathbb{R}^{n \times 1}$ bezeichnet.

Es ist A genau dann orthogonal, wenn (x_1, x_2, \dots, x_n) eine Orthonormalbasis von $\mathbb{R}^{n \times 1}$ ist.

In der Tat bedeutet $A^t \cdot A = E_n$ genau, daß $x_j^t \cdot x_k = \delta_{j,k}$ ist für $1 \leq j \leq n$ und $1 \leq k \leq n$.

Beispiel.

(1) Es ist E_n orthogonal.

(2) Es ist $\frac{1}{\sqrt{42}} \begin{pmatrix} \sqrt{14} & \sqrt{3} & -5 \\ \sqrt{14} & 2\sqrt{3} & 4 \\ \sqrt{14} & -3\sqrt{3} & 1 \end{pmatrix}$ orthogonal.

Bemerkung. Ist A orthogonal, dann ist $\det(A) \in \{-1, +1\}$.

Denn dann ist $(\det(A))^2 = \det(A^t) \cdot \det(A) = \det(A^t \cdot A) = \det(E_n) = 1$.

Bemerkung. Sind $S, T \in \mathbb{R}^{n \times n}$ orthogonal, dann sind auch $S \cdot T$ und S^{-1} orthogonal.

Denn dann ist $(S \cdot T)^t \cdot S \cdot T = T^t \cdot S^t \cdot S \cdot T = T^t \cdot T = E_n$.

Ferner ist $(S^{-1})^t \cdot S^{-1} = (S^t)^t \cdot S^{-1} = S \cdot S^{-1} = E_n$.

Bemerkung. Sei A symmetrisch.

Sei x ein Eigenvektor von A mit Eigenwert $\lambda \in \mathbb{R}$. Sei y ein Eigenvektor von A mit Eigenwert μ . Sei $\lambda \neq \mu$. Dann ist $x^t \cdot y = 0$.

Denn dann ist $x^t \cdot A \cdot y = x^t \cdot \mu y = \mu(x^t \cdot y)$ und zugleich $x^t \cdot A \cdot y = x^t \cdot A^t \cdot y = (A \cdot x)^t \cdot y =$

$(\lambda x)^t \cdot y = \lambda(x^t \cdot y)$. Die Differenz ist $0 = (\mu - \lambda)(x^t \cdot y)$. Da $\mu - \lambda \neq 0$, folgt $x^t \cdot y = 0$.

Konstruktion (Orthogonales Diagonalisieren).

Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch.

Seien $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ die paarweise verschiedenen Eigenwerte von A .

Sei B_j eine Orthonormalbasis von $E_A(\lambda_j)$ für $1 \leq j \leq m$.

Da A diagonalisierbar ist, ist $\mathbb{R}^{n \times 1} = \bigoplus_{j=1}^m E_A(\lambda_j)$.

Also ist $B := B_1 \sqcup B_2 \sqcup \dots \sqcup B_m$ eine Basis von $\mathbb{R}^{n \times 1}$.

Sei $S \in \mathbb{R}^{n \times n}$ die Matrix mit Spaltentupel B . Dann ist S orthogonal.

Sei $D \in \mathbb{R}^{n \times n}$ die Diagonalmatrix, deren Eintrag an Position (k, k) der Eigenwert der k -ten Spalte von S ist, für $1 \leq k \leq n$.

Dann ist

$$S^{-1} \cdot A \cdot S = S^t \cdot A \cdot S = D.$$

Da S orthogonal ist, sagen wir auch, es ist A *orthogonal diagonalisierbar*.

Beispiel. Es ist $A := \begin{pmatrix} -1 & -3 & -3 & 3 \\ -3 & -1 & 3 & -3 \\ -3 & 3 & -1 & -3 \\ 3 & -3 & -3 & -1 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$ symmetrisch.

Wir wollen A orthogonal diagonalisieren.

Zunächst wird

$$\begin{aligned} \chi_A(X) &= \det \begin{pmatrix} -1-X & -3 & -3 & 3 \\ -3 & -1-X & 3 & -3 \\ -3 & 3 & -1-X & -3 \\ 3 & -3 & -3 & -1-X \end{pmatrix} = \det \begin{pmatrix} -4-X & 0 & 0 & 4+X \\ 0 & -4-X & 0 & -4-X \\ 0 & 0 & -4-X & -4-X \\ 3 & -3 & -3 & -1-X \end{pmatrix} \\ &= (X+4)^3 \cdot \det \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 3 & -3 & -3 & -1-X \end{pmatrix} = (X+4)^3 \cdot \det \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & -3 & -3 & 8-X \end{pmatrix} = (X-8) \cdot (X+4)^3. \end{aligned}$$

Es wird $\lambda_1 = 8$ und $\lambda_2 = -4$.

Es wird $\text{aV}_A(8) = \text{gV}_A(8) = 1$. Es wird $\text{aV}_A(-4) = \text{gV}_A(-4) = 3$.

Wir berechnen eine Basis von $E_A(8)$.

$$\left(\begin{array}{cccc|c} -9 & -3 & -3 & 3 & 0 \\ -3 & -9 & 3 & -3 & 0 \\ -3 & 3 & -9 & -3 & 0 \\ 3 & -3 & -3 & -9 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 3 & 1 & 1 & -1 & 0 \\ 1 & 3 & -1 & 1 & 0 \\ 1 & -1 & 3 & 1 & 0 \\ 1 & -1 & -1 & -3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & -1 & -3 & 0 \\ 0 & 4 & 4 & 8 & 0 \\ 0 & 4 & 0 & 4 & 0 \\ 0 & 0 & 4 & 4 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 4 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Also ist $\left(\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right)$ eine Basis von $E_A(8)$.

Gram-Schmidt gibt dann die Orthonormalbasis $\left(\frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right)$ von $E_A(8)$.

Wir berechnen eine Basis von $E_A(-4)$.

$$\left(\begin{array}{cccc|c} 3 & -3 & -3 & 3 & 0 \\ -3 & 3 & 3 & -3 & 0 \\ -3 & 3 & 3 & -3 & 0 \\ 3 & -3 & -3 & 3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Also ist $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right)$ eine Basis von $E_A(-4)$.

Wir verwenden das Gram-Schmidt-Verfahren, um zu einer Orthonormalbasis von $E_A(-4)$ zu gelangen.

Es wird $d_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Es wird $d'_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ 0 \end{pmatrix}$. Also wird $d_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$.

Es wird $d'_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} - \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)^t \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left(\frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}\right)^t \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 1/3 \\ 1 \end{pmatrix}$. Also wird $d_3 = \frac{1}{2\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$.

Dies gibt die Orthonormalbasis $(d_1, d_2, d_3) = \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{2\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}\right)$ von $E_A(-4)$.

Zusammen wird mit der Orthogonalmatrix

$$S := \frac{1}{2\sqrt{3}} \begin{pmatrix} \sqrt{3} & \sqrt{6} & \sqrt{2} & -1 \\ -\sqrt{3} & \sqrt{6} & -\sqrt{2} & 1 \\ -\sqrt{3} & 0 & 2\sqrt{2} & 1 \\ \sqrt{3} & 0 & 0 & 3 \end{pmatrix}$$

und der Diagonalmatrix

$$D := \begin{pmatrix} 8 & & & \\ & -4 & & \\ & & -4 & \\ & & & -4 \end{pmatrix}$$

dann $S^t S = E_n$ und $S^{-1} A S = S^t A S = D$.

3.8.3 Hermitesche und unitäre Matrizen über \mathbb{C}

Sei $n \geq 0$. Sei $A = (a_{j,k})_{j,k} \in \mathbb{C}^{n \times n}$.

Definition. Es heißt A *hermitesch*, falls $A = \bar{A}^t$ ist.

Mit anderen Worten, A ist hermitesch, falls $a_{j,k} = \bar{a}_{k,j}$ ist für $1 \leq j \leq n$ und $1 \leq k \leq n$.

Beispiel. Es ist $\begin{pmatrix} 2 & 1-i & 3 \\ 1+i & 4 & i \\ 3 & -i & 5 \end{pmatrix} \in \mathbb{C}^{3 \times 3}$ hermitesch.

Definition. Es heißt A *unitär*, falls $\bar{A}^t \cdot A = E_n$ ist.

Dann ist $A^{-1} = \bar{A}^t$ und also auch $A \cdot \bar{A}^t = E_n$.

Bemerkung. Sei für $1 \leq k \leq n$ die k -te Spalte von A mit $x_k := (a_{j,k})_j \in \mathbb{C}^{n \times 1}$ bezeichnet.

Es ist A genau dann unitär, wenn (x_1, x_2, \dots, x_n) eine Orthonormalbasis von $\mathbb{C}^{n \times 1}$ ist.

In der Tat bedeutet $\bar{A}^t \cdot A = E_n$ genau, daß $\bar{x}_j^t \cdot x_k = \delta_{j,k}$ ist für $1 \leq j \leq n$ und $1 \leq k \leq n$.

Beispiel.

(1) Es ist E_n unitär.

(2) Es ist $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$ unitär.

Bemerkung. Sind $S, T \in \mathbb{C}^{n \times n}$ unitär, dann sind auch $S \cdot T$ und S^{-1} unitär.

Denn dann ist $\overline{(S \cdot T)}^t \cdot S \cdot T = \bar{T}^t \cdot \bar{S}^t \cdot S \cdot T = \bar{T}^t \cdot T = E_n$.

Ferner ist $\overline{(S^{-1})}^t \cdot S^{-1} = \overline{(\bar{S}^t)}^t \cdot S^{-1} = S \cdot S^{-1} = E_n$.

Definition. Es heißt $A \in \mathbb{C}^{n \times n}$ unitär diagonalisierbar, wenn es eine unitäre Matrix $S \in \mathbb{C}^{n \times n}$ und eine Diagonalmatrix $D \in \mathbb{C}^{n \times n}$ gibt mit

$$S^{-1} \cdot A \cdot S = \bar{S}^t \cdot A \cdot S = D.$$

Lemma. Genau dann ist A unitär diagonalisierbar, wenn $A \cdot \bar{A}^t = \bar{A}^t \cdot A$ ist.

Bemerkung.

(1) Ist A hermitesch, dann ist A unitär diagonalisierbar wegen $A \cdot \bar{A}^t = A^2 = \bar{A}^t \cdot A$. Insbesondere ist A dann diagonalisierbar.

(2) Ist A unitär, dann ist A unitär diagonalisierbar wegen $A \cdot \bar{A}^t = E_n = \bar{A}^t \cdot A$. Insbesondere ist A dann diagonalisierbar.

Bemerkung. Sei $A \in \mathbb{C}^{n \times n}$ unitär diagonalisierbar.

(1) Es ist A genau dann hermitesch, wenn alle Eigenwerte von A reell sind.

(2) Es ist A genau dann unitär, wenn alle Eigenwerte von A Betrag 1 haben.

Bemerkung. Sei A hermitesch.

Sei x ein Eigenvektor von A mit Eigenwert $\lambda \in \mathbb{R}$. Sei y ein Eigenvektor von A mit Eigenwert $\mu \in \mathbb{R}$. Sei $\lambda \neq \mu$. Dann ist $\bar{x}^t \cdot y = 0$.

Denn dann ist $\bar{x}^t \cdot A \cdot y = \bar{x}^t \cdot \mu y = \mu(\bar{x}^t \cdot y)$ und zugleich $\bar{x}^t \cdot A \cdot y = \bar{x}^t \cdot \bar{A}^t \cdot y = \overline{(A \cdot x)}^t \cdot y = \overline{(\lambda x)}^t \cdot y = \lambda(\bar{x}^t \cdot y)$. Die Differenz ist $0 = (\mu - \lambda)(\bar{x}^t \cdot y)$. Da $\mu - \lambda \neq 0$, folgt $\bar{x}^t \cdot y = 0$.

Konstruktion (Unitäres Diagonalisieren).

Sei $A \in \mathbb{C}^{n \times n}$ hermitesch.

Seien $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ die paarweise verschiedenen Eigenwerte von A .

Sei B_j eine Orthonormalbasis von $E_A(\lambda_j)$ für $1 \leq j \leq m$.

Da A diagonalisierbar ist, ist $\mathbb{C}^{n \times 1} = \bigoplus_{j=1}^m E_A(\lambda_j)$.

Also ist $B := B_1 \sqcup B_2 \sqcup \dots \sqcup B_m$ eine Basis von $\mathbb{C}^{n \times 1}$.

Sei $S \in \mathbb{C}^{n \times n}$ die Matrix mit Spaltentupel B . Dann ist S unitär.

Sei $D \in \mathbb{R}^{n \times n}$ die Diagonalmatrix, deren Eintrag an Position (k, k) der Eigenwert der k -ten Spalte von S ist, für $1 \leq k \leq n$.

Dann ist

$$S^{-1} \cdot A \cdot S = \bar{S}^t \cdot A \cdot S = D.$$

Beispiel. Es ist $A := \begin{pmatrix} 1 & i & i & 1 \\ -i & 1 & 1 & -i \\ -i & 1 & 3 & i \\ 1 & i & -i & 3 \end{pmatrix} \in \mathbb{C}^{4 \times 4}$ hermitesch.

Wir wollen A unitär diagonalisieren.

Zunächst wird

$$\begin{aligned} \chi_A(X) &= \det \begin{pmatrix} 1-X & i & i & 1 \\ -i & 1-X & 1 & -i \\ -i & 1 & 3-X & i \\ 1 & i & -i & 3-X \end{pmatrix} = \det \begin{pmatrix} 1-X & i & i & 1 \\ -iX & -X & 0 & 0 \\ 0 & 0 & 4-X & 4i-iX \\ 1 & i & -i & 3-X \end{pmatrix} = X(X-4) \cdot \det \begin{pmatrix} 1-X & i & i & 1 \\ -i & -1 & 0 & 0 \\ 0 & 0 & -1 & -i \\ 1 & i & -i & 3-X \end{pmatrix} \\ &= X(X-4) \cdot \det \begin{pmatrix} 2-X & i & i & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 2 & i & -i & 2-X \end{pmatrix} = X(X-4) \cdot \det \begin{pmatrix} 2-X & & & \\ & 2-X & & \\ & & 2-X & \\ & & & 2-X \end{pmatrix} = X(X-4)(X^2 - 4X + 4 - 4) \\ &= X^2(X-4)^2. \end{aligned}$$

Es wird $\lambda_1 = 0$ und $\lambda_2 = 4$.

Es wird $\text{a}V_A(0) = \text{g}V_A(0) = 2$. Es wird $\text{a}V_A(4) = \text{g}V_A(4) = 2$.

Wir berechnen eine Basis von $E_A(0)$.

$$\left(\begin{array}{cccc|c} 1 & i & i & 1 & 0 \\ -i & 1 & 1 & -i & 0 \\ -i & 1 & 3 & i & 0 \\ 1 & i & -i & 3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & i & i & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2i & 0 \\ 0 & 0 & -2i & 2 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & i & 0 & 2 & 0 \\ 0 & 0 & 1 & i & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Damit ist $\left(\begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ -i \\ 1 \end{pmatrix} \right)$ eine Basis von $E_A(0)$.

Gram-Schmidt gibt zunächst $d_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Sodann wird $d'_2 = \begin{pmatrix} -2 \\ 0 \\ -i \\ 1 \end{pmatrix} - \left(\frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right)^t \begin{pmatrix} -2 \\ 0 \\ -i \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -i \\ 1 \end{pmatrix} + i \begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ i \\ -i \\ 1 \end{pmatrix}$.

Also wird $d_2 = \frac{1}{2} \begin{pmatrix} -1 \\ i \\ -i \\ 1 \end{pmatrix}$.

Eine Orthonormalbasis von $E_A(0)$ ist mithin gegeben durch $\left(\frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} -1 \\ i \\ -i \\ 1 \end{pmatrix}\right)$.

Wir berechnen eine Basis von $E_A(4)$.

$$\left(\begin{array}{cccc|c} -3 & i & i & 1 & 0 \\ -i & -3 & 1 & -i & 0 \\ -i & 1 & -1 & i & 0 \\ 1 & i & -i & -1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & i & -i & -1 & 0 \\ 0 & 4i & -2i & -2 & 0 \\ 0 & -4 & 2 & -2i & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & -i/2 & -1/2 & 0 \\ 0 & 1 & -1/2 & i/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Damit ist $\left(\begin{pmatrix} i/2 \\ 1/2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ -i/2 \\ 0 \\ 1 \end{pmatrix}\right)$ eine Basis von $E_A(4)$.

Also ist auch $\left(\begin{pmatrix} i \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -i \\ 0 \\ 2 \end{pmatrix}\right)$ eine Basis von $E_A(4)$.

Gram-Schmidt gibt zunächst $d_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} i \\ 2 \\ 1 \\ 0 \end{pmatrix}$.

Sodann wird $d'_2 = \begin{pmatrix} 1 \\ -i \\ 0 \\ 2 \end{pmatrix} - \left(\frac{1}{\sqrt{6}} \begin{pmatrix} -i \\ 1 \\ 2 \\ 0 \end{pmatrix}\right)^t \begin{pmatrix} 1 \\ -i \\ 0 \\ 2 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} i \\ 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -i \\ 0 \\ 2 \end{pmatrix} + \frac{i}{3} \begin{pmatrix} i \\ 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -2i/3 \\ 2i/3 \\ 2 \end{pmatrix}$.

Es wird $d_2 = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 \\ -i \\ i \\ 3 \end{pmatrix}$.

Eine Orthonormalbasis von $E_A(4)$ ist mithin gegeben durch $\left(\frac{1}{\sqrt{6}} \begin{pmatrix} i \\ 2 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 \\ -i \\ i \\ 3 \end{pmatrix}\right)$.

Zusammen wird mit der unitären Matrix

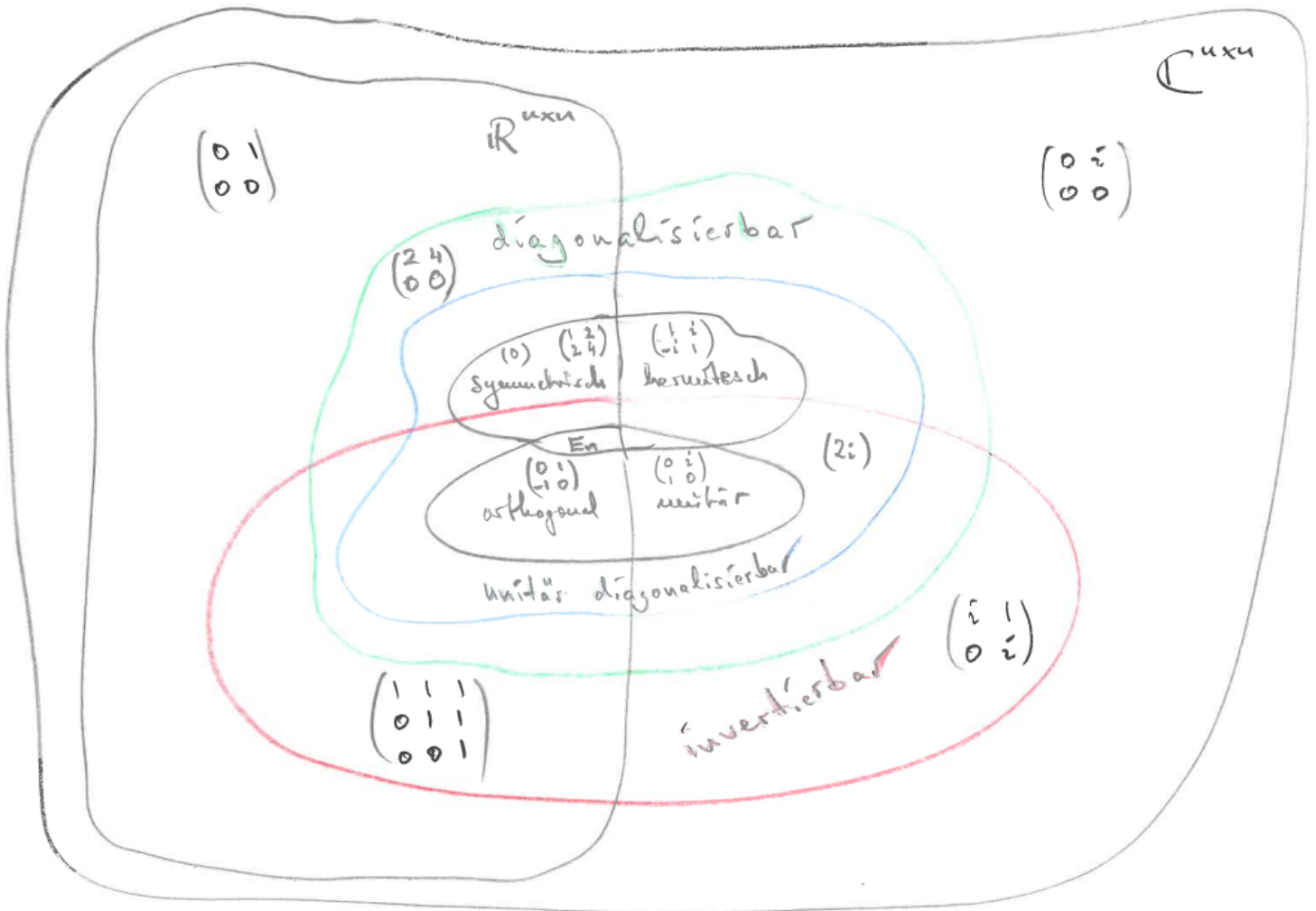
$$S := \frac{1}{2\sqrt{3}} \begin{pmatrix} -i\sqrt{6} & -\sqrt{3} & i\sqrt{2} & 1 \\ \sqrt{6} & i\sqrt{3} & \sqrt{2} & -i \\ 0 & -i\sqrt{3} & 2\sqrt{2} & i \\ 0 & \sqrt{3} & 0 & 3 \end{pmatrix}$$

und der Diagonalmatrix

$$D := \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 4 & \\ & & & 4 \end{pmatrix}$$

dann $\bar{S}^t S = E_4$ und $S^{-1} A S = \bar{S}^t A S = D$.

3.8.4 Eine informelle Übersicht



Kapitel 4

Funktionen in einer Variablen

4.1 Grenzwerte

Sei $D \subseteq \mathbb{R}$. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion.

Sei $g : D \rightarrow \mathbb{R}$ eine Funktion. Sei $h : D \rightarrow \mathbb{R}$ eine Funktion.

Definition.

(1) Sei $+\infty$ ein Element, für welches $x < +\infty$ gilt für $x \in \mathbb{R}$.

Gesprochen wird $+\infty$ als *plus unendlich*.

Oft schreibt man auch $\infty := +\infty$.

(2) Sei $-\infty$ ein Element, für welches $-\infty < x$ gilt für $x \in \mathbb{R}$.

Gesprochen wird $-\infty$ als *minus unendlich*.

Bei Intervallen wie $[0, +\infty[= \mathbb{R}_{\geq 0}$ oder $] -\infty, 5[$ haben wir diese beiden Elemente auch schon verwendet.

Definition.

(1) Sei $x \in \mathbb{R}$.

Für $\varepsilon \in \mathbb{R}_{>0}$ heißt $U_\varepsilon(x) :=]x - \varepsilon, x + \varepsilon[$ die ε -Umgebung von x .

(2) Für $\varepsilon \in \mathbb{R}_{>0}$ heißt $U_\varepsilon(+\infty) :=]\varepsilon, +\infty[$ die ε -Umgebung von $+\infty$.

(3) Für $\varepsilon \in \mathbb{R}_{>0}$ heißt $U_\varepsilon(-\infty) :=]-\infty, -\varepsilon[$ die ε -Umgebung von $-\infty$.

Definition. Sei $x \in \mathbb{R} \cup \{-\infty, +\infty\}$.

Wir sagen, x hängt an D , falls für jedes $\varepsilon \in \mathbb{R}_{>0}$ sich $D \cap U_\varepsilon(x) \neq \emptyset$ ergibt.

Sei

$$\overline{D} := \{x \in \mathbb{R} \cup \{-\infty, +\infty\} : x \text{ hängt an } D\}$$

der *Abschluß* von D in $\mathbb{R} \cup \{-\infty, +\infty\}$.

Es ist $D \subseteq \overline{D}$.

Beispiel.

(1) Es ist $\overline{[0, 1[} = [0, 1]$.

Zum einen hängt 1 an $[0, 1[$, da für jedes $\varepsilon \in]0, 1]$ sich

$$[0, 1[\cap U_\varepsilon(1) = [0, 1[\cap]1 - \varepsilon, 1 + \varepsilon[=]1 - \varepsilon, 1[\neq \emptyset$$

ergibt. Für $\varepsilon \in]1, +\infty[$ ist diese Schnittmenge dann erst recht nicht leer.

Zum anderen ist $[0, 1[\subseteq \overline{[0, 1[}$.

Schließlich hängt auch kein Element von $\mathbb{R} \setminus [0, 1]$ an $[0, 1[$.

(2) Es ist $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Dies können wir auch als Abkürzung verwenden.

Definition. Sei $x_0 \in \overline{D}$. Sei $y_0 \in \overline{\mathbb{R}}$.

Es heißt y_0 ein *Grenzwert* von f bei x_0 , falls für jedes $\varepsilon \in \mathbb{R}_{>0}$ ein $\delta \in \mathbb{R}_{>0}$ existiert mit

$$f(U_\delta(x_0) \cap D) \subseteq U_\varepsilon(y_0).$$

Diese Bedingung läßt sich äquivalent auch schreiben als

$$U_\delta(x_0) \cap D \subseteq f^{-1}(U_\varepsilon(y_0)).$$

Ein Grenzwert wird auch als ein *Limes* bezeichnet.

Bemerkung. Sei $x_0 \in \overline{D}$. Es gibt höchstens einen Grenzwert von f bei x_0 .

Annahme, es sind $y_0, \tilde{y}_0 \in \overline{\mathbb{R}}$ zwei Grenzwerte von f bei x_0 mit $y_0 \neq \tilde{y}_0$.

Dann gibt es $\varepsilon, \tilde{\varepsilon} \in \mathbb{R}_{>0}$ mit $U_\varepsilon(y_0) \cap U_{\tilde{\varepsilon}}(\tilde{y}_0) = \emptyset$.

Wähle $\delta \in \mathbb{R}_{>0}$ mit $f(U_\delta(x_0) \cap D) \subseteq U_\varepsilon(y_0)$.

Wähle $\tilde{\delta} \in \mathbb{R}_{>0}$ mit $f(U_{\tilde{\delta}}(x_0) \cap D) \subseteq U_{\tilde{\varepsilon}}(\tilde{y}_0)$.

Da x_0 in \overline{D} liegt, also an D hängt, gibt es ein $x_1 \in U_\delta(x_0) \cap U_{\tilde{\delta}}(x_0) \cap D$. Also ist

$$f(x_1) \in f(U_\delta(x_0) \cap D) \cap f(U_{\tilde{\delta}}(x_0) \cap D) \subseteq U_\varepsilon(y_0) \cap U_{\tilde{\varepsilon}}(\tilde{y}_0) = \emptyset.$$

Wir haben einen *Widerspruch*.

Definition. Sei $x_0 \in \overline{D}$.

- (1) Es heißt f *konvergent* in x_0 , wenn f bei x_0 einen Grenzwert hat, der in \mathbb{R} liegt.
- (2) Es heißt f *transvergent* (oder *bestimmt divergent*) in x_0 , wenn f bei x_0 einen Grenzwert hat, der in $\{-\infty, +\infty\}$ liegt.
- (3) Es heißt f *divergent* in x_0 , wenn f bei x_0 keinen Grenzwert hat.

Hat f in $x_0 \in \overline{D}$ einen Grenzwert $y_0 \in \overline{\mathbb{R}}$, so wird dieser

$$y_0 =: \lim_{x \rightarrow x_0} f(x) =: \lim_{x \in D, x \rightarrow x_0} f(x)$$

geschrieben.

Vorsicht: Bei Transvergenz existiert also ein Grenzwert, aber eben nur einer, der gleich $-\infty$ oder gleich $+\infty$ ist. Bei Divergenz existiert hingegen kein Grenzwert.

Beispiel. Sei $D := \mathbb{R} \setminus \{0\}$.

Wir betrachten die Funktion $f : D \rightarrow \mathbb{R} : x \mapsto f(x) := \frac{1}{x^2}$.

- (1) Sei $x_0 := +\infty \in \overline{D}$. Es ist

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow +\infty} \frac{1}{x^2} = 0 =: y_0.$$

Insbesondere ist f also konvergent in $x_0 = +\infty$.

Um dies nach Definition zu bestätigen, berechnen wir

$$\begin{aligned} f^{-1}(U_\varepsilon(y_0)) &= \{x \in D : f(x) \in U_\varepsilon(y_0)\} \\ &= \{x \in D : \frac{1}{x^2} \in]0 - \varepsilon, 0 + \varepsilon[\} \\ &= \{x \in D : -\varepsilon < \frac{1}{x^2} < +\varepsilon \} \\ &= \{x \in D : \frac{1}{x^2} < \varepsilon \} \\ &= \{x \in D : \frac{1}{|x|} < \sqrt{\varepsilon} \} \\ &= \{x \in D : |x| > \varepsilon^{-1/2} \} \\ &=]-\infty, -\varepsilon^{-1/2}[\cup]\varepsilon^{-1/2}, +\infty[. \end{aligned}$$

Mit $\delta := \varepsilon^{-1/2}$ ist also $U_\delta(x_0) \cap D = U_{\varepsilon^{-1/2}}(+\infty) \cap D =]\varepsilon^{-1/2}, +\infty[\subseteq f^{-1}(U_\varepsilon(y_0))$, wie gewünscht.

- (2) Sei $x_0 := 0 \in \overline{D}$. Es ist

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow 0} \frac{1}{x^2} = +\infty =: y_0.$$

Insbesondere ist f also transvergent in $x_0 = 0$.

Um dies nach Definition zu bestätigen, berechnen wir

$$\begin{aligned}
 f^{-1}(U_\varepsilon(y_0)) &= \{x \in D : f(x) \in U_\varepsilon(y_0)\} \\
 &= \{x \in D : \frac{1}{x^2} \in]\varepsilon, +\infty[\} \\
 &= \{x \in D : \varepsilon < \frac{1}{x^2}\} \\
 &= \{x \in D : |x| < \varepsilon^{-1/2}\} \\
 &=]-\varepsilon^{-1/2}, +\varepsilon^{-1/2}[\setminus \{0\} \\
 &= U_{\varepsilon^{-1/2}}(0) \cap D .
 \end{aligned}$$

Mit $\delta := \varepsilon^{-1/2}$ ist also $U_\delta(x_0) \cap D = U_{\varepsilon^{-1/2}}(0) \cap D \subseteq f^{-1}(U_\varepsilon(y_0))$, wie gewünscht, wobei hier sogar Gleichheit vorliegt.

Bemerkung. Sei $x_0 \in \overline{D}$.

Seien f und g konvergent in x_0 .

Seien also $\lim_{x \rightarrow x_0} f(x)$ und $\lim_{x \rightarrow x_0} g(x)$ existent und in \mathbb{R} gelegen.

- (1) Es ist $\lim_{x \rightarrow x_0} (f(x) + g(x)) = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x)$.
- (2) Es ist $\lim_{x \rightarrow x_0} (f(x) \cdot g(x)) = \lim_{x \rightarrow x_0} f(x) \cdot \lim_{x \rightarrow x_0} g(x)$.
- (3) Sei $\lim_{x \rightarrow x_0} g(x) \neq 0$.

$$\text{Es ist } \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}.$$

- (4) Sei $c \in \mathbb{R}$. Sei $f(x) > 0$ für $x \in D$.

$$\text{Es ist } \lim_{x \rightarrow x_0} f(x)^c = (\lim_{x \rightarrow x_0} f(x))^c.$$

Aussage (4) gründet dabei auf der sogenannten Stetigkeit der Abbildung $\mathbb{R}_{>0} \rightarrow \mathbb{R} : t \mapsto t^c$, die wir unten noch betrachten wollen.

In eingeschränkter Weise kann man mit $+\infty$ auch rechnen, wie folgende Bemerkung zeigen soll.

Bemerkung. Sei $x_0 \in \overline{D}$.

- (1) Es ist $\lim_{x \rightarrow x_0} f(x) = +\infty$ genau dann, wenn $\lim_{x \rightarrow x_0} -f(x) = -\infty$ ist.
- (2) Ist $\lim_{x \rightarrow x_0} f(x) \in \mathbb{R} \cup \{+\infty\}$ und $\lim_{x \rightarrow x_0} g(x) = +\infty$, dann ist

$$\lim_{x \rightarrow x_0} (f(x) + g(x)) = +\infty .$$

(3) Ist $\lim_{x \rightarrow x_0} f(x) \in \mathbb{R}_{>0} \cup \{+\infty\}$ und $\lim_{x \rightarrow x_0} g(x) = +\infty$, dann ist

$$\lim_{x \rightarrow x_0} (f(x) \cdot g(x)) = +\infty.$$

(4) Ist $\lim_{x \rightarrow x_0} f(x) \in \mathbb{R}$ und $\lim_{x \rightarrow x_0} g(x) = +\infty$, dann ist $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0$.

(5) Ist $\lim_{x \rightarrow x_0} f(x) \in \mathbb{R}_{>0}$ und $g(x) > 0$ für $x \in D$ und $\lim_{x \rightarrow x_0} g(x) = 0$, dann ist $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = +\infty$.

Entsprechendes gilt dank (1) bei Vorliegen negativer Vorzeichen.

Vorsicht ist aber bei Ausdrücken etwa der Form $(+\infty) - (+\infty)$ und $\frac{+\infty}{+\infty}$ angebracht, wo ein Ergebnis nicht durch eine simple Rechenoperation zu bekommen ist.

Beispiel.

(1) Es ist $\lim_{x \rightarrow +\infty} x = +\infty$.

Also ist $\lim_{x \rightarrow +\infty} x^k = +\infty$ für $k \in \mathbb{Z}_{\geq 1}$.

Also ist $\lim_{x \rightarrow +\infty} x^{-k} = 0$ für $k \in \mathbb{Z}_{\geq 1}$.

(2) Es ist $\lim_{x \rightarrow +\infty} \frac{x^3 - x}{2x^3 + x^2 + 1} = \lim_{x \rightarrow +\infty} \frac{1 - x^{-2}}{2 + x^{-1} + x^{-3}} = \frac{\lim_{x \rightarrow +\infty} 1 - x^{-2}}{\lim_{x \rightarrow +\infty} 2 + x^{-1} + x^{-3}} = \frac{1}{2}$.

(3) Es ist $\lim_{x \rightarrow +\infty} \frac{x^4 - x}{2x^3 + x^2 + 1} = \lim_{x \rightarrow +\infty} \frac{1 - x^{-3}}{2x^{-1} + x^{-2} + x^{-4}} = +\infty$.

(4) Es ist $\lim_{x \rightarrow +\infty} \frac{x^3 - x}{2x^4 + x^2 + 1} = \lim_{x \rightarrow +\infty} \frac{x^{-1} - x^{-3}}{2 + x^{-2} + x^{-4}} = 0$.

(5) Es ist $\lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = \lim_{x \rightarrow 1} \frac{x + 1}{1} = 2$.

Bemerkung. Sei $x_0 \in \overline{D}$.

Sei $\varepsilon \in \mathbb{R}_{>0}$. Sei $E := D \cap U_\varepsilon(x_0)$. Dann ist auch $x_0 \in \overline{E}$.

Es existiert $\lim_{x \in D, x \rightarrow x_0} f(x)$ genau dann, wenn $\lim_{x \in E, x \rightarrow x_0} f(x)$ existiert, und diesenfalls ist

$$\lim_{x \in D, x \rightarrow x_0} f(x) = \lim_{x \in E, x \rightarrow x_0} f(x).$$

Auf der rechten Seite ist hierbei genau genommen die eingeschränkte Funktion $f|_E$ gemeint.

Die Einschränkung des Definitionsbereichs auf $E = D \cap U_\varepsilon(x_0)$ hat also weder auf die Existenz noch auf den Wert des Limes einen Einfluß.

Lemma (Sandwich). Sei $x_0 \in \overline{D}$. Sei

$$g(x) \leq f(x) \leq h(x)$$

für $x \in D$. Sei zudem $\lim_{x \rightarrow x_0} g(x) = \lim_{x \rightarrow x_0} h(x)$. Dann ist

$$\lim_{x \rightarrow x_0} g(x) = \lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} h(x).$$

Zusatz: Bereits aus $g(x) \leq f(x)$ für $x \in D$ und $\lim_{x \rightarrow x_0} g(x) = +\infty$ folgt $\lim_{x \rightarrow x_0} f(x) = +\infty$.

Mit der vorstehenden Bemerkung kann man für das Sandwich-Lemma manchmal zu einem kleineren Definitionsbereich übergehen, falls man die erforderliche Ungleichung nicht auf ganz D kennt.

Beispiel. Wir wollen $\lim_{x \rightarrow +\infty} \frac{x}{x + \cos(x)}$ bestimmen, wobei die vorliegende Funktion auf $D :=]2, +\infty[$ definiert sei.

Da $-1 \leq \cos(x) \leq +1$, ist

$$\frac{x}{x+1} \leq \frac{x}{x+\cos(x)} \leq \frac{x}{x-1}$$

für $x \in D$.

Zum einen ist $\lim_{x \rightarrow +\infty} \frac{x}{x+1} = \lim_{x \rightarrow +\infty} \frac{1}{1+x^{-1}} = \frac{1}{1+\lim_{x \rightarrow +\infty} x^{-1}} = \frac{1}{1+0} = 1$.

Zum anderen ist $\lim_{x \rightarrow +\infty} \frac{x}{x-1} = \lim_{x \rightarrow +\infty} \frac{1}{1-x^{-1}} = \frac{1}{1-\lim_{x \rightarrow +\infty} x^{-1}} = \frac{1}{1-0} = 1$.

Diese beiden Grenzwerte stimmen überein. Also ist auch $\lim_{x \rightarrow +\infty} \frac{x}{x+\cos(x)} = 1$.

Später werden wir mittels der Regel von l'Hôpital noch interessantere Grenzwerte bestimmen können.

Bemerkung. Seien $D, E \subseteq \mathbb{R}$. Sei $f : D \rightarrow \mathbb{R}$ und $g : E \rightarrow \mathbb{R}$ Funktionen mit $f(D) \subseteq E$. Sei $x_0 \in D$ gegeben. Seien $\lim_{x \rightarrow x_0} f(x) =: y_0$ und $\lim_{y \rightarrow y_0} g(y)$ in $\overline{\mathbb{R}}$ existent.

Dann ist $\lim_{x \rightarrow x_0} g(f(x)) = \lim_{y \rightarrow y_0} g(y)$.

4.2 Stetigkeit

Sei $D \subseteq \mathbb{R}$. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion.

Definition.

- (1) Sei $x_0 \in D$. Es heißt f stetig in x_0 , wenn f in x_0 konvergiert.
- (2) Es heißt f stetig, falls für alle $x_0 \in D$ gilt, daß f stetig in x_0 ist.

Bemerkung. Sei $x_0 \in D$.

Ist f stetig in x_0 , dann ist

$$f(x_0) = \lim_{x \rightarrow x_0} f(x).$$

Annahme, nicht. Schreibe $y_0 := \lim_{x \rightarrow x_0} f(x)$. Sei $\varepsilon := |y_0 - f(x_0)|/2 \in \mathbb{R}_{>0}$. Dann ist $f(x_0) \notin U_\varepsilon(y_0)$.

Wähle $\delta \in \mathbb{R}_{>0}$ derart, daß $f(U_\delta(x_0) \cap D) \subseteq U_\varepsilon(y_0)$ ist. Insbesondere liegt dann auch $f(x_0) \in f(U_\delta(x_0) \cap D) \subseteq U_\varepsilon(y_0)$. Wir haben einen *Widerspruch*.

Beispiel. Sei $D = \mathbb{R}$ und

$$h : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \begin{cases} 0 & \text{falls } x < 0 \\ 2 & \text{falls } x \geq 0 \end{cases}$$

Es hat also h bei 0 einen Sprung mit ‘‘Sprunghöhe’’ 2.

Wir wollen überprüfen, daß h nicht stetig ist. Genauer gesagt wollen wir überprüfen, daß h in $x_0 := 0$ nicht stetig ist.

Annahme, h ist stetig in 0. Dann ist $2 = h(0) = \lim_{x \rightarrow 0} h(x)$.

Sei $\varepsilon := 1$, die halbe ‘‘Sprunghöhe’’. Wähle $\delta \in \mathbb{R}_{>0}$ mit $h(U_\delta(0)) \subseteq U_\varepsilon(2) =]1, 3[$.

Da $-\delta/2 \in U_\delta(0)$, ist daher auch $0 = h(-\delta/2) \in]1, 3[$. Wir haben einen *Widerspruch*.

Bemerkung. Sei $n \in \mathbb{Z}_{\geq 0}$. Seien $a_i \in \mathbb{R}$ für $0 \leq i \leq n$.

Dann ist die polynomiale Funktion $p : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto p(x) := \sum_{i=0}^n a_i x^i$ stetig.

Bemerkung. Sei $c \in \mathbb{R}$. Dann ist $p : \mathbb{R}_{>0} \rightarrow \mathbb{R} : x \mapsto x^c$ stetig.

Bemerkung. Die Funktionen $\sin : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sin(x)$, $\cos : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \cos(x)$ und $b : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$ sind stetig.

Für die Definition von x^c , $\sin(x)$, $\cos(x)$ greifen wir hier auf Schulwissen zurück. Wir werden daran nochmals im Zusammenhang mit der Exponentialfunktion und mit Potenzreihen erinnern.

Bemerkung. Seien $f, g : D \rightarrow \mathbb{R}$ stetige Funktionen.

Dann sind auch deren Summe $f + g : D \rightarrow \mathbb{R} : x \mapsto f(x) + g(x)$ und deren Produkt $f \cdot g : D \rightarrow \mathbb{R} : x \mapsto f(x) \cdot g(x)$ stetige Funktionen.

Ist $g(x) \neq 0$ für $x \in D$, dann ist auch deren Quotient $\frac{f}{g} : D \rightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}$ stetig.

Bemerkung.

Sei $D \subseteq \mathbb{R}$ und sei $f : D \rightarrow \mathbb{R}$ eine stetige Funktion.

Sei $E \subseteq \mathbb{R}$ und sei $g : E \rightarrow \mathbb{R}$ eine stetige Funktion.

Sei $f(D) \subseteq E$. Dann ist auch $g \circ f|_D : D \rightarrow \mathbb{R} : x \mapsto g(f(x))$ eine stetige Funktion.

Folgende Limesvertauschungsregel zeigt den Sinn des Begriffs der stetigen Funktion.

Lemma. Sei $D \subseteq \mathbb{R}$. Sei $f : D \rightarrow \mathbb{R}$. Sei $x_0 \in \overline{D}$. Sei $y_0 := \lim_{x \rightarrow x_0} f(x)$ in \mathbb{R} existent.

Sei $f(D) \cup \{y_0\} \subseteq E \subseteq \mathbb{R}$. Sei $g : E \rightarrow \mathbb{R}$ stetig.

Dann ist

$$\lim_{x \rightarrow x_0} g(f(x)) = g(\lim_{x \rightarrow x_0} f(x)) = g(y_0).$$

Kurz, stetige Funktionen vertauschen mit Limiten.

Beispiel. Vom vorletzten Beispiel in §4.1, Teil (2), wissen wir: $\lim_{x \rightarrow +\infty} \frac{x^3 - x}{2x^3 + x^2 + 1} = \frac{1}{2}$.

Nun ist $g : \mathbb{R}_{>0} \rightarrow \mathbb{R} : x \mapsto g(x) := x^{1/2}$ stetig.

Also ist $\lim_{x \rightarrow +\infty} \sqrt{\frac{x^3 - x}{2x^3 + x^2 + 1}} = \lim_{x \rightarrow +\infty} \left(\frac{x^3 - x}{2x^3 + x^2 + 1}\right)^{1/2} = \left(\lim_{x \rightarrow +\infty} \frac{x^3 - x}{2x^3 + x^2 + 1}\right)^{1/2} = \frac{1}{\sqrt{2}}$.

Bemerkung. Seien $a, b \in \mathbb{R}$ mit $a < b$. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig.

Es existieren Maximum $\max\{f(x) : x \in [a, b]\}$ und Minimum $\min\{f(x) : x \in [a, b]\}$ von f auf $[a, b]$.

Mit anderen Worten, es gibt $s, t \in [a, b]$ mit $f(s) \leq f(x) \leq f(t)$ für $x \in [a, b]$.

4.3 Grenzwerte von Folgen

Sei $k \in \mathbb{Z}$.

Definition. Eine Funktion $f : \mathbb{Z}_{\geq k} \rightarrow \mathbb{R}$ heißt *reelle Folge* oder *Folge*.

Wir schreiben auch $f_n := f(n)$ für den Funktionswert bei $n \in \mathbb{Z}_{\geq k}$.

Wir schreiben auch $f =: (f_n)_{n \geq k}$ für die Folge als ganze.

Bemerkung. Es ist $+\infty \in \overline{\mathbb{Z}_{\geq k}}$.

Für eine Folge $(f_n)_{n \geq k}$ verfügen wir also über den Begriff des Grenzwerts

$$\lim_{n \rightarrow \infty} f_n := \lim_{n \in \mathbb{Z}_{\geq k}, n \rightarrow +\infty} f_n$$

Auch eine Folge kann konvergieren, transvergieren oder divergieren.

Man schreibt auch kurz $f_n \rightarrow a$ statt $\lim_{n \rightarrow \infty} f_n = a$, gesprochen “ f_n geht gegen a ”.

Bemerkung. Sei $D \subseteq \mathbb{R}$. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion.

Sei $x_0 \in \overline{D}$ so, daß f bei x_0 einen Grenzwert in $\overline{\mathbb{R}}$ hat.

Sei $(a_n)_{n \geq k}$ eine Folge mit $a_n \in D$ für $n \in \mathbb{Z}_{\geq k}$ und mit $\lim_{n \rightarrow \infty} a_n = x_0$.

Dann ist $\lim_{n \rightarrow \infty} f(a_n) = \lim_{x \rightarrow x_0} f(x)$, wie aus der letzten Bemerkung in §4.1 folgt.

Bemerkung. Vorstehende Bemerkung wird oft im Spezialfall $(a_n)_{n \geq k} = (n)_{n \geq k}$ und $x_0 = +\infty$ verwendet. Diesfalls gilt dann

$$\lim_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} f(n) = \lim_{x \rightarrow +\infty} f(x).$$

Kurz, falls der Funktionsgrenzwert existiert, stimmt er mit dem Folgentgrenzwert überein.

Folgende Beispiele können wir hier nur ohne Begründung zur Kenntnis nehmen. Mit später eingeführten Techniken wie l'Hôpital können wir teilweise Begründungen nachreichen.

Beispiel. Sei $q \in \mathbb{R}$.

Ist $q \in]-\infty, -1]$, dann existiert $\lim_{n \rightarrow \infty} q^n$ nicht.

Ist $q \in]-1, +1[$, dann ist $\lim_{n \rightarrow \infty} q^n = 0$.

Ist $q = 1$, dann ist $\lim_{n \rightarrow \infty} q^n = 1$.

Ist $q \in]+1, +\infty[$, dann ist $\lim_{n \rightarrow \infty} q^n = +\infty$.

Beispiel. Sei $c \in \mathbb{R}_{>0}$. Es ist $\lim_{n \rightarrow \infty} c^{1/n} = 1$.

Beispiel. Es ist $\lim_{n \rightarrow \infty} n^{1/n} = 1$.

Beispiel. Es ist $\lim_{n \rightarrow \infty} (n!)^{1/n} = +\infty$.

Beispiel. Es ist $\lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n =: e \approx 2,718281828459$.

Beispiel. Wir wollen $\lim_{n \rightarrow \infty} (n^2 + 1)^{1/n} = 1$ nachweisen, unter Verwendung der vorangegangenen Beispiele.

Für $n \geq 1$ ist $(n^2)^{1/n} \leq (n^2 + 1)^{1/n} \leq (2n^2)^{1/n}$.

Es ist

$$\lim_{n \rightarrow \infty} (n^2)^{1/n} = \lim_{n \rightarrow \infty} n^{1/n} \cdot n^{1/n} = \left(\lim_{n \rightarrow \infty} n^{1/n} \right) \cdot \left(\lim_{n \rightarrow \infty} n^{1/n} \right) = 1 \cdot 1 = 1.$$

Es ist

$$\lim_{n \rightarrow \infty} (2n^2)^{1/n} = \lim_{n \rightarrow \infty} 2^{1/n} \cdot (n^2)^{1/n} = \left(\lim_{n \rightarrow \infty} 2^{1/n} \right) \cdot \left(\lim_{n \rightarrow \infty} (n^2)^{1/n} \right) = 1 \cdot 1 = 1.$$

Mit Sandwich-Lemma folgt also

$$\lim_{n \rightarrow \infty} (n^2 + 1)^{1/n} = 1.$$

Beispiel. Es ist

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \sqrt{n}(\sqrt{n} - \sqrt{n+1}) &= \lim_{n \rightarrow \infty} \sqrt{n} \cdot \frac{(\sqrt{n} - \sqrt{n+1})(\sqrt{n} + \sqrt{n+1})}{\sqrt{n} + \sqrt{n+1}} \\
 &= \lim_{n \rightarrow \infty} \sqrt{n} \cdot \frac{(\sqrt{n})^2 - (\sqrt{n+1})^2}{\sqrt{n} + \sqrt{n+1}} = - \lim_{n \rightarrow \infty} \frac{\sqrt{n}}{\sqrt{n} + \sqrt{n+1}} \\
 &= - \lim_{n \rightarrow \infty} \frac{1}{1 + \sqrt{1 + \frac{1}{n}}} = - \frac{1}{1 + \lim_{n \rightarrow \infty} \sqrt{1 + \frac{1}{n}}} \\
 &= - \frac{1}{1 + \sqrt{\lim_{n \rightarrow \infty} (1 + \frac{1}{n})}} = -\frac{1}{2}.
 \end{aligned}$$

Wem das unglaublich vorkommt, der setze einmal in den vorliegenden Ausdruck mit dem Taschenrechner große Werte für n ein.

4.4 Limes superior

Sei $\ell \in \mathbb{Z}$. Sei $(a_k)_{k \geq \ell}$ eine reelle Folge.

Definition. Sei $m \in \mathbb{Z}$.

Sei $\mathbb{Z}_{\geq m} \rightarrow \mathbb{Z}_{\geq \ell} : j \mapsto k_j$ eine Abbildung, für welche $k_j < k_{j+1}$ ist für $j \in \mathbb{Z}_{\geq m}$.

Dann heißt $(a_{k_j})_{j \geq m}$ eine *Teilfolge* von $(a_k)_{k \geq \ell}$.

Beispiel. Sei $(a_k)_{k \geq 1} = ((-1)^k + \frac{1}{k})_{k \geq 1}$.

(1) Sei $k_j := 2j$ für $j \in \mathbb{Z}_{\geq 1}$.

Wir erhalten die Teilfolge $(a_{k_j})_{j \geq 1} = ((-1)^{2j} + \frac{1}{2j})_{j \geq 1} = (1 + \frac{1}{2j})_{j \geq 1}$.

(2) Sei $k_j := 2j + 1$ für $j \in \mathbb{Z}_{\geq 1}$.

Wir erhalten die Teilfolge $(a_{k_j})_{j \geq 1} = ((-1)^{2j+1} + \frac{1}{2j+1})_{j \geq 1} = (-1 + \frac{1}{2j+1})_{j \geq 1}$.

Bemerkung. Hat die Folge $(a_k)_{k \geq \ell}$ den Grenzwert $x \in \overline{\mathbb{R}}$, dann hat auch jede ihrer Teilfolgen diesen Grenzwert x .

Nun führen wir einen Begriff ein für Grenzwerte von Teilfolgen von $(a_k)_{k \geq \ell}$. Nach vorstehender Bemerkung liefert dies nichts Neues, wenn $(a_k)_{k \geq \ell}$ einen Grenzwert hat. Wirklich nötig wird die Begrifflichkeit nur, falls $(a_k)_{k \geq \ell}$ keinen Grenzwert hat.

Definition. Ein Grenzwert einer Teilfolge von $(a_k)_{k \geq \ell}$ heißt *Häufungspunkt* dieser Folge.

Hierzu können natürlich nur Teilfolgen herangezogen werden, die einen Grenzwert haben, d.h. die konvergent oder transvergent sind.

Beispiel. Wir setzen das vorige Beispiel fort. Sei $(a_k)_{k \geq 1} = ((-1)^k + \frac{1}{k})_{k \geq 1}$.

- (1) Die Teilfolge $(1 + \frac{1}{2j})_{j \geq 1}$ hat den Grenzwert 1.
- (2) Die Teilfolge $(-1 + \frac{1}{2j+1})_{j \geq 1}$ hat den Grenzwert -1 .

Also sind -1 und 1 Häufungspunkte dieser Folge.

Bemerkung. Findet man konvergente oder transvergente Teilfolgen von $(a_k)_{k \geq \ell}$ derart, daß höchstens endlich viele Zahlen in $\mathbb{Z}_{\geq \ell}$ nicht als Index in einer dieser Teilfolgen auftauchen, dann liefern die Grenzwerte dieser Teilfolgen bereits alle Häufungspunkte.

Beispiel. Im vorigen Beispiel ist jedes $k \in \mathbb{Z}_{\geq 2}$ gleich $2j$ für ein $j \geq 1$ oder gleich $2j + 1$ für ein $j \geq 1$.

Der Index $k = 1$ wurde nicht abgedeckt. Endlich viele Ausnahmen stören aber nicht.

Also ist die Menge aller Häufungspunkte von $(a_k)_{k \geq 1} = ((-1)^k + \frac{1}{k})_{k \geq 1}$ gleich $\{-1, +1\}$.

Bemerkung. Die Menge der Häufungspunkte von $(a_k)_{k \geq \ell}$ ist nicht leer und hat ein maximales Element.

Definition. Sei der *Limes superior* $\limsup_{k \rightarrow \infty} a_k$ definiert als der maximale Häufungspunkt von $(a_k)_{k \geq \ell}$.

Bemerkung. Existiert $\lim_{k \rightarrow \infty} a_k$, dann ist $\limsup_{k \rightarrow \infty} a_k = \lim_{k \rightarrow \infty} a_k$. In diesem Sinne ist der Limes superior ein Ersatz für den Limes.

Hat umgekehrt die Folge $(a_k)_{k \geq \ell}$ mehr als einen Häufungspunkt, dann hat sie keinen Grenzwert.

Die Folge $(a_k)_{k \geq \ell}$ hat aber immer einen Limes superior.

Beispiel. Wir setzen voriges Beispiel fort.

Es ist $\limsup_{k \rightarrow \infty} (-1)^k + \frac{1}{k} = \max\{-1, 1\} = 1$.

4.5 Reihen

4.5.1 Reihenbegriff

Definition. Sei $\ell \in \mathbb{Z}$. Sei eine Folge $(a_k)_{k \geq \ell}$ gegeben.

Eine *Reihe* ist die Folge der *Teilsummen* $(\sum_{k=\ell}^n a_k)_{n \geq \ell}$.

Sei

$$\sum_{k=\ell}^{\infty} a_k := \lim_{n \rightarrow \infty} \sum_{k=\ell}^n a_k$$

der *Grenzwert* oder kurz *Wert* dieser Reihe. Häufig wird auch dieser Wert als Reihe angesprochen.

Auch eine Reihe kann konvergieren, transvergieren oder divergieren.

Bemerkung. Sei $q \in \mathbb{R} \setminus \{1\}$ gegeben.

$$(1) \text{ Es ist } \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q} \text{ für } n \in \mathbb{Z}_{\geq 0}.$$

$$(2) \text{ Es ist } \sum_{k=0}^{\infty} q^k = \frac{1}{1 - q} \text{ falls } |q| < 1.$$

Falls $|q| \geq 1$ ist, konvergiert diese Reihe nicht.

Die Summe in (1) heißt *geometrische Summe*. Die Reihe in (2) heißt *geometrische Reihe*.

Zu (1). Schreibe $x := \sum_{k=0}^n q^k$. Dann ist

$$x \cdot q = \sum_{k=0}^n q^k \cdot q = \sum_{k=0}^n q^{k+1} = \sum_{\ell=1}^{n+1} q^{\ell} = x - q^0 + q^{n+1}.$$

Also ist $x \cdot (1 - q) = 1 - q^{n+1}$. Es folgt $x = \frac{1 - q^{n+1}}{1 - q}$.

Zu (2). Es ist

$$\sum_{k=0}^{\infty} q^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n q^k \stackrel{(1)}{=} \lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q},$$

da $\lim_{n \rightarrow \infty} q^{n+1} = 0$ wegen $|q| < 1$.

Beispiel. Es ist $\sum_{k=1}^{\infty} \frac{1}{3^k} = -1 + \sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k = -1 + \frac{1}{1 - \frac{1}{3}} = -1 + \frac{3}{2} = \frac{1}{2}$.

Bemerkung. Sei $\ell \in \mathbb{Z}$. Sei $(a_k)_{k \geq \ell}$ eine konvergente Folge.

Dann ist $\sum_{k=\ell}^{\infty} a_{k+1} - a_k = (\lim_{k \rightarrow \infty} a_k) - a_{\ell}$.

Eine solche Reihe heißt auch eine *Teleskopreihe*.

In der Tat ist für $n \geq \ell$

$$\sum_{k=\ell}^n a_{k+1} - a_k = \left(\sum_{k=\ell}^n a_{k+1}\right) - \left(\sum_{k=\ell}^n a_k\right) = \left(\sum_{k=\ell+1}^{n+1} a_k\right) - \left(\sum_{k=\ell}^n a_k\right) = a_{n+1} - a_{\ell}.$$

Somit wird

$$\sum_{k=\ell}^{\infty} a_{k+1} - a_k = \lim_{n \rightarrow \infty} \sum_{k=\ell}^n a_{k+1} - a_k = \left(\lim_{n \rightarrow \infty} a_{n+1} \right) - a_{\ell} = \left(\lim_{k \rightarrow \infty} a_k \right) - a_{\ell}.$$

Beispiel. Wir wollen $\sum_{n=3}^{\infty} \frac{-1}{n^2 - n}$ berechnen.

Wir setzen hierzu an: $\frac{-1}{n^2 - n} = \frac{-1}{n(n-1)} = \frac{u}{n} + \frac{v}{n-1}$, wobei $u, v \in \mathbb{R}$ noch gesucht sind. Hierzu rechnen wir die rechte Seite aus: $\frac{u}{n} + \frac{v}{n-1} = \frac{u(n-1) + vn}{n(n-1)}$. Wir wählen also $u = 1$ und $v = -1$. Somit erkennen wir, daß eine Teleskopreihe vorliegt und rechnen

$$\sum_{n=3}^{\infty} \frac{-1}{n^2 - n} = \sum_{n=3}^{\infty} \frac{1}{n} - \frac{1}{n-1} = \sum_{n=3}^{\infty} \frac{1}{(n+1)-1} - \frac{1}{n-1} = \lim_{n \rightarrow +\infty} \frac{1}{n-1} - \frac{1}{3-1} = -\frac{1}{2}.$$

Wem das unglaublich vorkommt, der summiere einmal von 3 bis n für ein großes n mit einem geeigneten Taschenrechner.

Beispiel. Es ist $\sum_{k=1}^{\infty} \frac{1}{k} = +\infty$. Diese Reihe heißt auch *harmonische* Reihe. Die harmonische Reihe transvergiert also gegen $+\infty$.

Das sieht man wie folgt.

$$\text{Es ist } \sum_{k=2^0}^{2^1-1} \frac{1}{k} = 1 \geq \frac{1}{2}.$$

$$\text{Es ist } \sum_{k=2^1}^{2^2-1} \frac{1}{k} \geq \sum_{k=2^1}^{2^2-1} 2^{-2} = \frac{1}{2}.$$

$$\text{Es ist } \sum_{k=2^2}^{2^3-1} \frac{1}{k} \geq \sum_{k=2^2}^{2^3-1} 2^{-3} = \frac{1}{2}.$$

$$\text{Es ist } \sum_{k=2^3}^{2^4-1} \frac{1}{k} \geq \sum_{k=2^3}^{2^4-1} 2^{-4} = \frac{1}{2}.$$

Usf.

$$\text{Somit ist } \sum_{k=1}^{2^n-1} \frac{1}{k} \geq \frac{n}{2}.$$

Also geht die Folge der Teilsummen dieser Reihe gegen $+\infty$.

$$\text{Nochmal zum Vergleich: } \sum_{k=1}^{\infty} \frac{1}{k} = +\infty, \text{ aber } \lim_{k \rightarrow \infty} \frac{1}{k} = 0.$$

4.5.2 Konvergenzkriterien

Das Bestimmen eines Reihenwerts ist oft schwierig bis unmöglich. Nur festzustellen, daß eine Reihe konvergiert, ist oft viel einfacher.

Sei $\ell \in \mathbb{Z}$. Sei eine Folge $(a_k)_{k \geq \ell}$ gegeben. Wir betrachten die Reihe $\sum_{k=\ell}^{\infty} a_k$.

Bemerkung. Ist die Reihe $\sum_{k=\ell}^{\infty} a_k$ konvergent, dann ist $\lim_{k \rightarrow \infty} a_k = 0$.

Das letzte Beispiel in §4.5.1 zeigt, daß hier die Umkehrung nicht gilt.

Beispiel. Die Reihe $\sum_{k=1}^{\infty} (-1)^k$ konvergiert nicht, da $\lim_{k \rightarrow \infty} (-1)^k$ nicht gleich 0 ist, sondern vielmehr gar nicht existiert.

Definition. Die Reihe $\sum_{k=\ell}^{\infty} a_k$ heißt *absolut konvergent*, falls die Reihe $\sum_{k=\ell}^{\infty} |a_k|$ konvergiert.

Bemerkung. Falls die Reihe $\sum_{k=\ell}^{\infty} a_k$ absolut konvergiert, dann konvergiert sie auch.

Lemma (Leibniz-Kriterium).

Sei $a_k \cdot a_{k+1} < 0$ für $k \in \mathbb{Z}_{\geq \ell}$, d.h. seien die Elemente a_k von abwechselndem Vorzeichen.

Sei $|a_k| \geq |a_{k+1}|$ für $k \in \mathbb{Z}_{\geq \ell}$. Sei $\lim_{k \rightarrow \infty} |a_k| = 0$.

Dann ist die Reihe $\sum_{k=\ell}^{\infty} a_k$ konvergent.

Beispiel. Wir wollen die Konvergenz der Reihe $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ untersuchen.

Es ist also $(a_k)_{k \geq 1} = \left(\frac{(-1)^k}{k}\right)_{k \geq 1}$.

Es sind die Elemente a_k von abwechselndem Vorzeichen.

Es ist $\left|\frac{(-1)^k}{k}\right| = \frac{1}{k}$. Also ist zum einen $|a_k| = \frac{1}{k} \geq \frac{1}{k+1} = |a_{k+1}|$ für $k \in \mathbb{Z}_{\geq 1}$. Ferner ist $\lim_{k \rightarrow \infty} |a_k| = \lim_{k \rightarrow \infty} \frac{1}{k} = 0$.

Mit dem Leibnizkriterium folgt, daß die Reihe $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ konvergiert.

Diese Reihe ist also konvergent, nicht aber absolut konvergent, da die harmonische Reihe transvergiert.

Es ist $\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -\ln(2) \approx -0,693147$. Das nachzurechnen wäre aber schwieriger als die bloße Konvergenzuntersuchung.

Lemma (Wurzelkriterium).

Ist $\limsup_{k \rightarrow \infty} |a_k|^{1/k} < 1$, dann ist die Reihe $\sum_{k=\ell}^{\infty} a_k$ absolut konvergent.

Bei der Bildung dieses Limes superior dürfen nur Folgenglieder mit $k \geq 1$ und $k \geq \ell$ berücksichtigt werden.

Beispiel. Wir wollen die Konvergenz der Reihe $\sum_{k=1}^{\infty} \frac{1}{k!}$ untersuchen.

Es ist $|a_k|^{1/k} = \frac{1}{k!^{1/k}}$, was gemäß Beispiel in §4.3 gegen 0 geht.

Also ist hier $\limsup_{k \rightarrow \infty} |a_k|^{1/k} = \limsup_{k \rightarrow \infty} \frac{1}{k!^{1/k}} = \lim_{k \rightarrow \infty} \frac{1}{k!^{1/k}} = 0 < 1$.

Wir schließen, daß $\sum_{k=1}^{\infty} \frac{1}{k!}$ konvergiert, dies natürlich auch absolut.

Es ist $\sum_{k=1}^{\infty} \frac{1}{k!} = e - 1$. Das wird sich später klären, im Kontext der Exponentialfunktion.

Lemma (Quotientenkriterium). Sei $a_k \neq 0$ für $k \in \mathbb{Z}_{\geq \ell}$.

Ist $\limsup_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| < 1$, dann ist die Reihe $\sum_{k=\ell}^{\infty} a_k$ absolut konvergent.

Beispiel. Wir wollen nochmal die Konvergenz der Reihe $\sum_{k=1}^{\infty} \frac{1}{k!}$ untersuchen.

Es ist $\left| \frac{a_{k+1}}{a_k} \right| = \frac{1/(k+1)!}{1/k!} = \frac{k!}{(k+1)!} = \frac{1}{k+1} \rightarrow 0 < 1$.

Wir schließen abermals, daß $\sum_{k=1}^{\infty} \frac{1}{k!}$ konvergiert, dies natürlich auch absolut.

Lemma (Majorantenkriterium).

Wir wollen die Konvergenz der Reihe $\sum_{k=\ell}^{\infty} a_k$ untersuchen.

Sei ein $m \in \mathbb{Z}_{\geq \ell}$ und eine reelle Folge $(b_k)_{k \geq m}$ gefunden mit $|a_k| \leq b_k$ für $k \in \mathbb{Z}_{\geq m}$ und mit $\sum_{k=m}^{\infty} b_k$ konvergent.

Dann ist $\sum_{k=\ell}^{\infty} a_k$ absolut konvergent.

Die Reihe $\sum_{k=m}^{\infty} b_k$ heißt dann *konvergente Majorante* der Reihe $\sum_{k=\ell}^{\infty} a_k$.

Beispiel. Wir wollen die Konvergenz der Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ untersuchen.

Es ist $\left| \frac{1}{k^2} \right| = \frac{1}{k^2} \leq \frac{1}{k^2 - k}$ für $k \in \mathbb{Z}_{\geq 3}$. Es ist $\sum_{k=3}^{\infty} \frac{1}{k^2 - k} = \frac{1}{2}$, insbesondere ist diese Reihe konvergent; siehe vorletztes Beispiel in §4.5.1.

Gemäß Majorantenkriterium ist dann auch $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergent, dies natürlich auch absolut.

Es ist $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \approx 1,644934$. Dies ist mit unseren Mitteln nicht nachzurechnen – unter Verwendung von Fourierreihen könnte man es bestätigen.

Bemerkung. Sei eine Reihe $\sum_{k=\ell}^{\infty} a_k$ zu untersuchen.

Sei dazu eine weitere Reihe $\sum_{k=\ell}^{\infty} b_k$ gefunden worden und ein $m \in \mathbb{Z}_{\geq \ell}$ mit $a_k \geq b_k \geq 0$ für $k \in \mathbb{Z}_{\geq m}$ und mit $\sum_{k=\ell}^{\infty} b_k = +\infty$.

Wäre dann $\sum_{k=\ell}^{\infty} a_k$ konvergent, dann wäre nach Majorantenkriterium auch $\sum_{k=\ell}^{\infty} b_k$ konvergent, was *nicht* so ist.

Also ist auch $\sum_{k=\ell}^{\infty} a_k = +\infty$.

4.6 Differentialrechnung

4.6.1 Differenzierbarkeit

Sei $D \subseteq \mathbb{R}$. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion.

Definition. Es heißt $D \subseteq \mathbb{R}$ eine *offene* Teilmenge, wenn es für jedes $x \in D$ ein $\delta \in \mathbb{R}_{>0}$ gibt mit $U_\delta(x) \subseteq D$.

Sei in diesem §4.6.1 die Teilmenge $D \subseteq \mathbb{R}$ als offen vorausgesetzt.

Beispiel. Es ist $\mathbb{R} \subseteq \mathbb{R}$ offen. Es ist $\mathbb{R}_{>3} \subseteq \mathbb{R}$ offen. Es ist $\mathbb{R} \setminus \{-1\} \subseteq \mathbb{R}$ offen.

Es ist $\mathbb{R}_{\geq 3} \subseteq \mathbb{R}$ nicht offen, da es kein $\delta \in \mathbb{R}_{>0}$ gibt mit $U_\delta(3) \subseteq \mathbb{R}_{\geq 3}$.

Definition. Sei $x_0 \in D$.

Es heißt f *differenzierbar in x_0* , falls die auf $D \setminus \{x_0\}$ definierte Funktion $\frac{f(x)-f(x_0)}{x-x_0}$ in $x_0 \in D \subseteq \overline{D \setminus \{x_0\}}$ konvergiert. Diesfalls schreiben wir

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Bemerkung. Es ist $\frac{f(x)-f(x_0)}{x-x_0}$ die Steigung der Geraden durch die beiden Punkte $(x_0, f(x_0))$ und $(x, f(x))$ auf dem Graphen von f . Läßt man nun x gegen x_0 gehen, um als Grenzwert $f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x)-f(x_0)}{x-x_0}$ zu erhalten, so erkennt man, daß $f'(x_0)$ die Steigung der Tangente am Graph von f durch den Punkt $(x_0, f(x_0))$ ist.

Bemerkung. Es ist alternativ auch $\lim_{x \rightarrow x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h \rightarrow 0} \frac{f(x_0+h)-f(x_0)}{h}$.

Definition. Es heißt $f : D \rightarrow \mathbb{R}$ *differenzierbar*, wenn f in x differenzierbar ist für $x \in D$.

Wir erhalten diesfalls die Funktion $f' : D \rightarrow \mathbb{R} : x \mapsto f'(x)$, genannt *Ableitung* von f .

Man schreibt auch $\frac{d}{dx} f(x) := f'(x)$.

Manchmal schreibt man auch $(f(x))' := f'(x)$.

Man schreibt auch $f''(x) = (f')'(x)$ für die *zweite Ableitung* von f an der Stelle x , falls existent. Man schreibt auch $f'''(x) = ((f')')'(x)$ für die *dritte Ableitung* von f an der Stelle x , falls existent.

Beispiel. Wir wollen einmal direkt nach Definition untersuchen, ob die Funktion $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x) := x^3$ differenzierbar ist.

Bei $x_0 \in \mathbb{R}$ ist

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{(x_0 + h)^3 - x_0^3}{h} = \frac{x_0^3 + 3x_0^2h + 3x_0h^2 + h^3 - x_0^3}{h} = 3x_0^2 + 3x_0h + h^2$$

nach dem Binomischen Lehrsatz aus §1.8. Folglich wird

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \rightarrow 0} 3x_0^2 + 3x_0h + h^2 = 3x_0^2.$$

Insbesondere ist f differenzierbar.

Man schreibt auch $\frac{d}{dx} x^3 = 3x^2$.

Manchmal schreibt man auch $(x^3)' = 3x^2$.

Bemerkung. Für $q \in \mathbb{Q} \setminus \{0\}$ ist $\frac{d}{dx} x^q = qx^{q-1}$.

Hat q in gekürzter Schreibweise einen geraden Nenner, muß dabei $x > 0$ vorausgesetzt werden.

Bemerkung. Ist $f : D \rightarrow \mathbb{R}$ in x differenzierbar, dann ist f auch stetig.

Definition. Die Funktion $f : D \rightarrow \mathbb{R}$ heißt *stetig differenzierbar*, wenn f differenzierbar und f' stetig ist.

Eine Funktion, die differenzierbar, nicht aber stetig differenzierbar ist, ist $x^2 \cdot \sin(1/x)$, bei 0 mit 0 fortgesetzt. Dieses Beispiel werden wir unten in §4.6.3 nachholen, sobald wir die nötigen Regeln kennen.

4.6.2 Ableitungsregeln

Seien $D, E \subseteq \mathbb{R}$ offen.

Lemma. Seien $f, g : D \rightarrow \mathbb{R}$ differenzierbar. Seien $a, b \in \mathbb{R}$ konstant.

(1) Es ist $(a \cdot f(x) + b \cdot g(x))' = a \cdot f'(x) + b \cdot g'(x)$ für $x \in D$.

(2) Es ist $(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$ für $x \in D$.

(3) Sei $g(x) \neq 0$ für $x \in D$.

$$\text{Es ist } \left(\frac{f(x)}{g(x)} \right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2} \text{ für } x \in D.$$

Es heißt (2) auch *Produktregel*. Es heißt (3) auch *Quotientenregel*.

Beweis. Wir wollen nur (2) nachweisen. Es ist

$$\begin{aligned}
 & (f(x) \cdot g(x))' \\
 = & \lim_{h \rightarrow 0} (f(x+h) \cdot g(x+h) - f(x) \cdot g(x)) \cdot h^{-1} \\
 = & \lim_{h \rightarrow 0} (f(x+h) \cdot g(x+h) - f(x+h) \cdot g(x) + f(x+h) \cdot g(x) - f(x) \cdot g(x)) \cdot h^{-1} \\
 = & \lim_{h \rightarrow 0} f(x+h) \cdot \lim_{h \rightarrow 0} (g(x+h) - g(x)) \cdot h^{-1} + \lim_{h \rightarrow 0} (f(x+h) - f(x)) \cdot h^{-1} \cdot g(x) \\
 = & f(x) \cdot g'(x) + f'(x) \cdot g(x).
 \end{aligned}$$

□

Beispiel. Es ist $\frac{d}{dx} x = 1$.

Also ist $\frac{d}{dx} x^2 = \frac{d}{dx} (x \cdot x) = 1 \cdot x + x \cdot 1 = 2x$. Die Produktregel liefert also das bekannte Resultat erneut.

Beispiel. Mit der Quotientenregel wird

$$\frac{d}{dx} \frac{x}{x^2 + 1} = \frac{1 \cdot (x^2 + 1) - x \cdot 2x}{(x^2 + 1)^2} = \frac{1 - x^2}{(x^2 + 1)^2}.$$

Lemma (Kettenregel). Sei $f : D \rightarrow \mathbb{R}$ differenzierbar. Sei $g : E \rightarrow \mathbb{R}$ differenzierbar. Sei $f(D) \subseteq E$.

Für $x \in D$ ist

$$\frac{d}{dx} g(f(x)) = g'(f(x)) \cdot f'(x).$$

Der Faktor $g'(f(x))$ heißt hierbei auch *äußere Ableitung*. Der Faktor $f'(x)$ heißt hierbei auch *innere Ableitung*.

Beweis. Sei $h(x) := g(f(x))$ für $x \in D$. Unter Beachtung von $\lim_{\tilde{x} \rightarrow x} f(\tilde{x}) = f(x) =: y$ wird

$$\begin{aligned}
 h'(x) &= \lim_{\tilde{x} \rightarrow x} (h(\tilde{x}) - h(x)) \cdot (\tilde{x} - x)^{-1} \\
 &= \lim_{\tilde{x} \rightarrow x} (g(f(\tilde{x})) - g(f(x))) \cdot (\tilde{x} - x)^{-1} \\
 &= \lim_{\tilde{x} \rightarrow x} (g(f(\tilde{x})) - g(f(x))) \cdot (f(\tilde{x}) - f(x))^{-1} \cdot (f(\tilde{x}) - f(x)) \cdot (\tilde{x} - x)^{-1} \\
 &= \lim_{\tilde{x} \rightarrow x} (g(f(\tilde{x})) - g(f(x))) \cdot (f(\tilde{x}) - f(x))^{-1} \cdot \lim_{\tilde{x} \rightarrow x} (f(\tilde{x}) - f(x)) \cdot (\tilde{x} - x)^{-1} \\
 &= \lim_{\tilde{y} \rightarrow y} (g(\tilde{y}) - g(y)) \cdot (\tilde{y} - y)^{-1} \cdot \lim_{\tilde{x} \rightarrow x} (f(\tilde{x}) - f(x)) \cdot (\tilde{x} - x)^{-1} \\
 &= g'(y) \cdot f'(x) \\
 &= g'(f(x)) \cdot f'(x).
 \end{aligned}$$

□

Beispiel. Sei $f(x) := x^2$. Sei $g(x) := x^3$.

Es wird $g(f(x)) = x^6$. Also ist $\frac{d}{dx} x^6 = 6x^5$.

Alternativ wird mit der Kettenregel

$$\frac{d}{dx} x^6 = \frac{d}{dx} g(f(x)) = g'(f(x)) \cdot f'(x) = 3f(x)^2 \cdot f'(x) = 3(x^2)^2 \cdot 2x = 6x^5.$$

Beispiel. Es wird

$$\frac{d}{dx} (x^2 + 1)^3 = 3(x^2 + 1)^2 \cdot 2x.$$

Beispiel. Es wird

$$\frac{d}{dx} (1 + \sqrt{x})^{\frac{1}{3}} = \frac{1}{3}(1 + \sqrt{x})^{-\frac{2}{3}} \cdot \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{6(1 + \sqrt{x})^{2/3}\sqrt{x}}.$$

Bemerkung. Wir können nun die Quotientenregel aus der Kettenregel und der Produktregel herleiten.

Denn in der beschriebenen Situation wird, wenn wir $u(t) := t^{-1}$ setzen und $u'(t) = -t^{-2}$ verwenden,

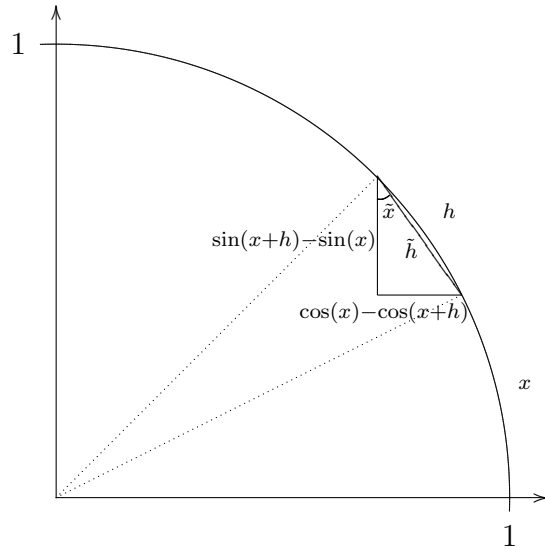
$$\begin{aligned} \frac{d}{dx} \frac{f(x)}{g(x)} &= \frac{d}{dx} (f(x) \cdot g(x)^{-1}) \\ &= f'(x) \cdot g(x)^{-1} + f(x) \cdot \frac{d}{dx} (g(x)^{-1}) \\ &= f'(x) \cdot g(x)^{-1} + f(x) \cdot \frac{d}{dx} (u(g(x))) \\ &= f'(x) \cdot g(x)^{-1} + f(x) \cdot u'(g(x)) \cdot g'(x) \\ &= f'(x) \cdot g(x)^{-1} + f(x) \cdot (-g(x)^{-2}) \cdot g'(x) \\ &= \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}. \end{aligned}$$

4.6.3 Sinus und Cosinus

Bemerkung. Es ist

$$\begin{aligned} \frac{d}{dx} \sin(x) &= \cos(x) \\ \frac{d}{dx} \cos(x) &= -\sin(x). \end{aligned}$$

Dies wollen wir anhand einer Skizze begründen. Dabei sind x und h Winkel im Bogenmaß.



Es ist

$$\frac{\sin(x+h) - \sin(x)}{h} \approx \frac{\sin(x+h) - \sin(x)}{\tilde{h}} = \cos(\tilde{x}) \approx \cos(x)$$

und

$$\frac{\cos(x+h) - \cos(x)}{h} \approx -\frac{\cos(x) - \cos(x+h)}{\tilde{h}} = -\sin(\tilde{x}) \approx -\sin(x),$$

wobei die Näherungen umso besser sind, je kleiner $|h|$ ist.

Beispiel. Es ist $\frac{d}{dx} \sin(5x) = \cos(5x) \cdot (5x)' = \cos(5x) \cdot 5 = 5 \cos(5x)$.

Beispiel. Es ist $\tan(x) = \frac{\sin(x)}{\cos(x)}$ und also

$$\frac{d}{dx} \tan(x) = \frac{d}{dx} \frac{\sin(x)}{\cos(x)} = \frac{(\sin(x))' \cdot \cos(x) - \sin(x) \cdot (\cos(x))'}{\cos(x)^2} = \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2}.$$

Man kann die Rechnung aber auch anders zu Ende führen.

$$\frac{d}{dx} \tan(x) \stackrel{\text{siehe oben}}{=} \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = 1 + \tan(x)^2.$$

Beides ist richtig.

Beispiel. Sei

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x) := \begin{cases} x^2 \sin(1/x) & \text{falls } x \in \mathbb{R} \setminus \{0\} \\ 0 & \text{falls } x = 0 \end{cases}$$

Für $x \in \mathbb{R} \setminus \{0\}$ können wir die Ableitung mit den Regeln berechnen. Es wird

$$f'(x) = 2x \sin(1/x) + x^2 \cos(1/x) \cdot (-x^{-2}) = 2x \sin(1/x) - \cos(1/x).$$

Für $x = 0$ können wir die Ableitung direkt nach Definition berechnen. Es wird

$$f'(0) = \lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} h \sin(1/h) = 0,$$

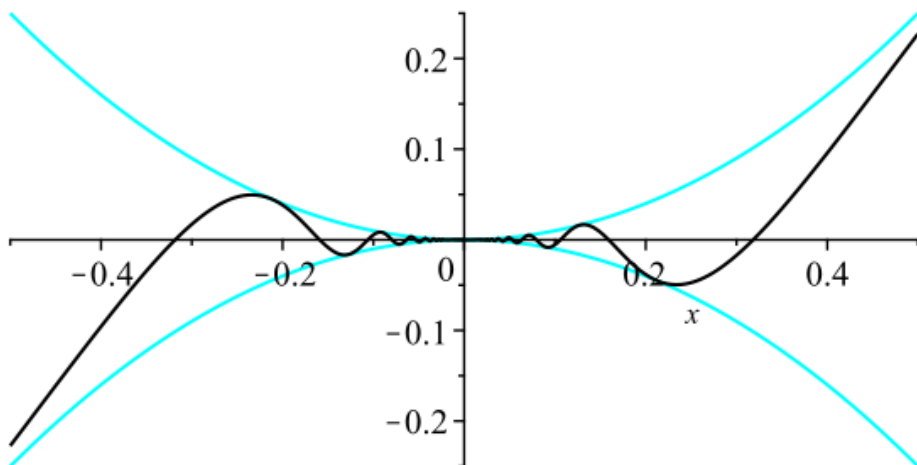
dank des Sandwiches $-|h| \leq h \sin(1/h) \leq |h|$.

Wir behaupten, daß f' nicht stetig ist.

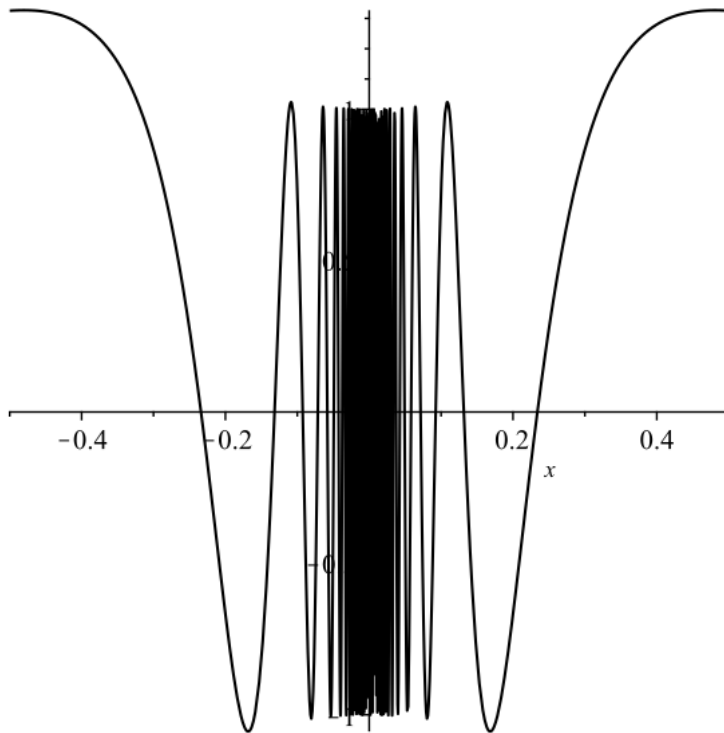
Annahme, f' ist stetig. Dann ist $\lim_{n \rightarrow \infty} f'(\frac{1}{2n\pi}) = f'(\lim_{n \rightarrow \infty} \frac{1}{2n\pi}) = f'(0) = 0$. Aber $\lim_{n \rightarrow \infty} f'(\frac{1}{2n\pi}) = \lim_{n \rightarrow \infty} \frac{2}{2n\pi} \sin(2n\pi) - \cos(2n\pi) = -1$. *Widerspruch*.

Im Ergebnis ist f also differenzierbar, aber nicht stetig differenzierbar.

Graph von f in schwarz, zwischen den Graphen von x^2 und $-x^2$ in blau :



Graph von f' :



4.6.4 Die Grenzwertregel nach l'Hôpital

Sei $D \subseteq \mathbb{R}$ offen. Sei $x_0 \in \overline{D}$.

Lemma (l'Hôpital). Seien $f : D \rightarrow \mathbb{R}$ und $g : D \rightarrow \mathbb{R}$ differenzierbare Funktionen.

- (1) Sei $\lim_{x \rightarrow x_0} f(x) = 0$ und $\lim_{x \rightarrow x_0} g(x) = 0$.

Dann ist

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

falls letzterer Grenzwert in $\overline{\mathbb{R}}$ existiert.

- (2) Sei $\lim_{x \rightarrow x_0} f(x) \in \{-\infty, +\infty\}$ und $\lim_{x \rightarrow x_0} g(x) \in \{-\infty, +\infty\}$.

Dann ist

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

falls letzterer Grenzwert in $\overline{\mathbb{R}}$ existiert.

Die Regel von l'Hôpital ist nicht mit der Quotientenregel zu verwechseln. Diese beiden Regeln gehören gedanklich in verschiedene Schubladen.

Bemerkung. Falls $x_0 \in D$ liegt und f und g stetig sind, so ist die Voraussetzung in Teil (1) des vorstehenden Lemmas erfüllt, falls

$$f(x_0) = 0 \quad \text{und} \quad g(x_0) = 0.$$

Beispiel.

$$(1) \text{ Es ist } \lim_{x \rightarrow 0} \frac{\sin(x)}{x} \stackrel{\text{l'H.}}{=} \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = \cos(0) = 1.$$

$$(2) \text{ Es ist } \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{x^2} \stackrel{\text{l'H.}}{=} \lim_{x \rightarrow 0} \frac{-\sin(x)}{2x} \stackrel{\text{l'H.}}{=} \lim_{x \rightarrow 0} \frac{-\cos(x)}{2} = \frac{-\cos(0)}{2} = -\frac{1}{2}.$$

(3) Es ist

$$\begin{aligned} \lim_{x \rightarrow 0} \sin(x)^{-2} - x \sin(x)^{-3} &= \lim_{x \rightarrow 0} \frac{\sin(x) - x}{\sin(x)^3} \\ &\stackrel{\text{l'H.}}{=} \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{3 \sin(x)^2 \cos(x)} \\ &\stackrel{\text{l'H.}}{=} \lim_{x \rightarrow 0} \frac{-\sin(x)}{3(2 \sin(x) \cos(x)^2 - \sin(x)^3)} \\ &= \lim_{x \rightarrow 0} \frac{-1}{3(2 \cos(x)^2 - \sin(x)^2)} \\ &= -\frac{1}{6}. \end{aligned}$$

$$(4) \text{ Es ist } \lim_{x \rightarrow +\infty} \frac{x}{2x + \sin(x)} = \lim_{x \rightarrow +\infty} \frac{1}{2 + \frac{\sin(x)}{x}} = \frac{1}{2}.$$

Es versagt aber l'Hôpital, da $\lim_{x \rightarrow +\infty} \frac{(x)'}{(2x + \sin(x))'} = \lim_{x \rightarrow +\infty} \frac{1}{2 + \cos(x)}$ nicht existiert.

4.6.5 Taylor

Sei $D \subseteq \mathbb{R}$ offen. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion.

Definition. Wir definieren rekursiv folgende Funktionen, sofern existent.

Sei $f^{(0)}(x) := f(x)$ für $x \in D$.

Sei $f^{(n)}(x) := \frac{d}{dx} f^{(n-1)}(x)$ für $x \in D$, sofern existent.

Es heißt $f^{(n)} : D \rightarrow \mathbb{R}$ dann die n -te Ableitung von f , gesprochen “ f oben n ”.

Natürlich sind auch weiterhin die Bezeichnungen

$$f(x) = f^{(0)}(x), \quad f'(x) = f^{(1)}(x), \quad f''(x) = f^{(2)}(x), \quad f'''(x) = f^{(3)}(x)$$

in Gebrauch.

Beispiel. Sei $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} : x \mapsto x^{-1}$.

Es ist $f^{(n)}(x) = (-1)^n \cdot n! \cdot x^{-(n+1)}$ für $x \in \mathbb{R} \setminus \{0\}$.

Dies wollen wir mit Induktion überprüfen.

Induktionsanfang: Für $n = 0$ ist in der Tat $x^{-1} = f^{(0)}(x) = (-1)^0 \cdot 0! \cdot x^{-(0+1)}$ für $x \in \mathbb{R} \setminus \{0\}$.

Induktionsschritt. Sei $n \geq 1$. Sei die Formel für $n - 1$ bekannt. Dann wird

$$\begin{aligned} f^{(n)}(x) &= \frac{d}{dx} f^{(n-1)}(x) \\ &\stackrel{\text{IV}}{=} \frac{d}{dx} (-1)^{n-1} \cdot (n-1)! \cdot x^{-((n-1)+1)} \\ &= (-1)^{n-1} \cdot (n-1)! \cdot \frac{d}{dx} x^{-n} \\ &= (-1)^{n-1} \cdot (n-1)! \cdot (-n) x^{-(n+1)} \\ &= (-1)^n \cdot n! \cdot x^{-(n+1)}. \end{aligned}$$

Definition. Sei $n \geq 0$. Falls $f^{(k)} : D \rightarrow \mathbb{R}$ existiert für $0 \leq k \leq n$ und $f^{(n)}$ stetig ist, dann heißt f eine n -fach stetig differenzierbare Funktion.

Oft sind Funktionen beliebig oft differenzierbar, und damit n -fach stetig differenzierbar für jedes $n \geq 0$.

Satz (Taylor). Sei $x_0 \in D$.

Sei $n \geq 0$. Sei $f : D \rightarrow \mathbb{R}$ eine $(n+1)$ -fach stetig differenzierbare Funktion.

Das *Taylorpolynom* der Stufe n von f zum Entwicklungspunkt x_0 sei

$$T_n(f, x, x_0) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

für $x \in \mathbb{R}$.

Das *Restglied* nach Lagrange der Stufe n zum Entwicklungspunkt x_0 sei

$$R_n(f, x, x_0, \vartheta) := \frac{f^{(n+1)}(x_0 + \vartheta(x - x_0))}{(n+1)!} (x - x_0)^{n+1}.$$

für $x \in \mathbb{R}$ und $\vartheta \in [0, 1]$.

Für $x \in D$ gibt es ein $\vartheta \in [0, 1]$ mit

$$f(x) = T_n(f, x, x_0) + R_n(f, x, x_0, \vartheta),$$

vorausgesetzt, $[x_0, x] \subseteq D$ falls $x_0 \leq x$, $[x, x_0] \subseteq D$ falls $x \leq x_0$.

Das Taylorpolynom der Stufe n von f zum Entwicklungspunkt x_0 ist das Polynom von Grad $\leq n$, welches die Funktion f in der Nähe von x_0 bestmöglich annähert.

Das ϑ ist die eingebaute Ungenauigkeit – man weiß in der Praxis, es liegt $\vartheta \in [0, 1]$, aber man weiß es nicht genauer. Für Abschätzungen reicht das aber oft.

Auf eine Formel zu hoffen, die kein solches ϑ enthält, wäre zu optimistisch, schließlich ist nicht jede Funktion ein Polynom, und auch nicht ein Polynom plus ein simpler Ausdruck.

Bemerkung. Das Taylorpolynom von Stufe 0 zum Entwicklungspunkt x_0 ist

$$T_0(f, x, x_0) = \sum_{k=0}^0 \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \frac{f^{(0)}(x_0)}{0!} (x - x_0)^0 = f(x_0).$$

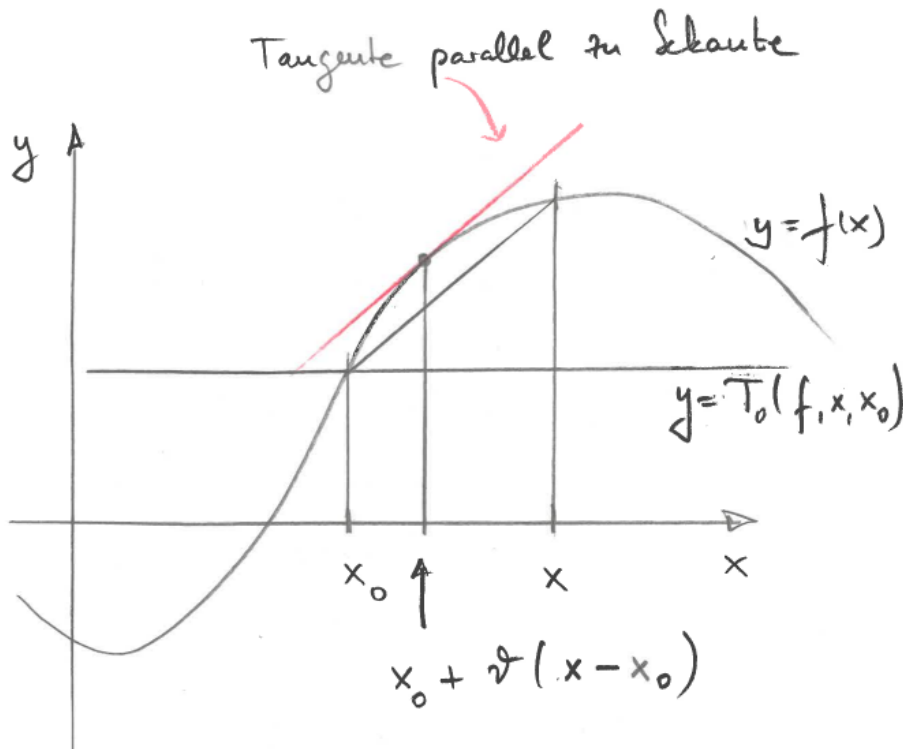
Der Satz von Taylor gibt

$$\begin{aligned} f(x) &= T_0(f, x, x_0) + R_0(f, x, x_0, \vartheta) \\ &= f(x_0) + f^{(1)}(x_0 + \vartheta(x - x_0)) \cdot (x - x_0) \\ &= f(x_0) + f'(x_0 + \vartheta(x - x_0)) \cdot (x - x_0). \end{aligned}$$

für ein $\vartheta \in [0, 1]$. Also ist

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(x_0 + \vartheta(x - x_0)).$$

Mit anderen Worten, die Steigung der Sekante an den Stellen x und x_0 ist gleich der Steigung der Tangente an der Zwischenstelle $x_0 + \vartheta(x - x_0)$.



Es hieße $\vartheta = 0$, daß die Zwischenstelle $x_0 + \vartheta(x - x_0)$ auf x_0 läge.

Es hieße $\vartheta = 1$, daß die Zwischenstelle $x_0 + \vartheta(x - x_0)$ auf x läge.

Tatsächlich haben wir in der Skizze $\vartheta \approx 0,38 \in [0, 1]$.

Beispiel. Das Taylorpolynom der Stufe 1 zum Entwicklungspunkt x_0 ist

$$\begin{aligned} T_1(f, x, x_0) &= \sum_{k=0}^1 \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \frac{f^{(0)}(x_0)}{0!} (x - x_0)^0 + \frac{f^{(1)}(x_0)}{1!} (x - x_0)^1 \\ &= f(x_0) + f'(x_0)(x - x_0). \end{aligned}$$

Der Graph von $T_1(f, x, x_0)$ ist die Tangente an den Graphen von f an der Stelle x_0 .

Beispiel. Sei $D := \mathbb{R}_{>0}$. Sei $f : D \rightarrow \mathbb{R} : x \mapsto f(x) = \sqrt{x}$.

Sei nach $T_1(f, x, 1)$, $T_2(f, x, 1)$ und $T_3(f, x, 1)$ gefragt.

Sei dann nach einem $C \in \mathbb{R}_{>0}$ gefragt mit $|f(x) - T_3(f, x, 1)| \leq C \cdot |x - 1|^4 = C \cdot (x - 1)^4$ für $x \in [\frac{1}{2}, \frac{3}{2}]$.

Der Faktor $(x - 1)^4$ hat hierbei den Sinn, die Abschätzung in der Nähe von 0 sehr genau zu machen. Ist etwa $x = 1,1$, dann vergegenwärtige man sich, daß $(x - 1)^4 = 0,0001$ ist.

Wir rechnen.

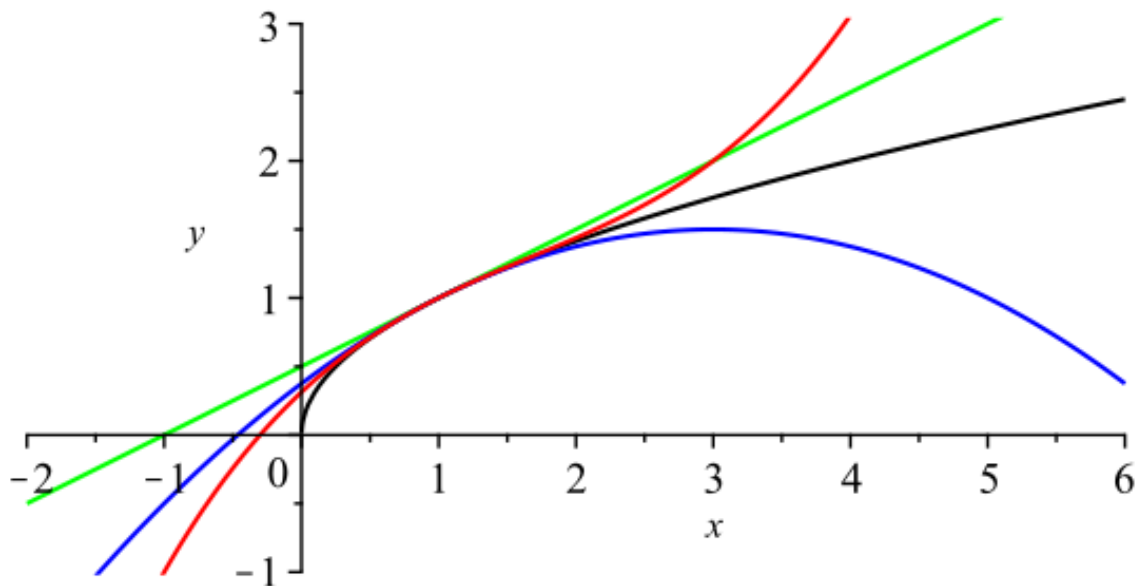
$$\begin{array}{ll} f^{(0)}(x) = x^{1/2} & f^{(0)}(1) = 1 \\ f^{(1)}(x) = \frac{1}{2} \cdot x^{-1/2} & f^{(1)}(1) = \frac{1}{2} \\ f^{(2)}(x) = \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot x^{-3/2} & f^{(2)}(1) = -\frac{1}{4} \\ f^{(3)}(x) = \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot x^{-5/2} & f^{(3)}(1) = \frac{3}{8} \\ f^{(4)}(x) = \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot \left(-\frac{5}{2}\right) \cdot x^{-7/2} & f^{(4)}(1) = -\frac{15}{16} x^{-7/2} \end{array}$$

Wir erhalten folgende Taylorpolynome.

$$\begin{aligned} T_1(f, x, 1) &= f^{(0)}(1) + f^{(1)}(1)(x - 1) \\ &= 1 + \frac{1}{2}(x - 1) \\ T_2(f, x, 1) &= f^{(0)}(1) + f^{(1)}(1)(x - 1) + \frac{1}{2!} f^{(2)}(1)(x - 1)^2 \\ &= 1 + \frac{1}{2}(x - 1) - \frac{1}{8}(x - 1)^2 \\ T_3(f, x, 1) &= f^{(0)}(1) + f^{(1)}(1)(x - 1) + \frac{1}{2!} f^{(2)}(1)(x - 1)^2 + \frac{1}{3!} f^{(3)}(1)(x - 1)^3 \\ &= 1 + \frac{1}{2}(x - 1) - \frac{1}{8}(x - 1)^2 + \frac{1}{16}(x - 1)^3 \end{aligned}$$

Hier sollte man das so stehenlassen und nicht z.B. $(x - 1)^2$ noch ausmultiplizieren. Daß Potenzen von $(x - 1)$ dastehen, ist im Sinne des Näherungsvorhabens zweckmäßig.

In folgender Skizze ist der Graph von $f(x) = \sqrt{x}$ schwarz, der von $T_1(f, x, 1)$, also von der Tangente, grün, der von $T_2(f, x, 1)$ blau und der von $T_3(f, x, 1)$ rot.



Für ein $\vartheta \in [0, 1]$ gilt nun folgendes.

$$\begin{aligned}
 |f(x) - T_3(f, x, 1)| &= |R_3(f, x, 1, \vartheta)| \\
 &= \left| \frac{f^{(4)}(1 + \vartheta(x-1))}{4!} (x-1)^4 \right| \\
 &= \frac{\frac{15}{16} (1 + \vartheta(x-1))^{-7/2}}{24} (x-1)^4 \\
 &= \frac{5}{128(1 + \vartheta(x-1))^{7/2}} (x-1)^4
 \end{aligned}$$

Wir brauchen eine Abschätzung für $x \in [\frac{1}{2}, \frac{3}{2}]$, d.h. $|x-1| \leq \frac{1}{2}$. Was auch immer $\vartheta \in [0, 1]$ für einen Wert hat, jedenfalls ist $1 + \vartheta(x-1) \in [\frac{1}{2}, \frac{3}{2}]$. Somit können wir wie folgt fortsetzen.

$$\frac{5}{128(1 + \vartheta(x-1))^{7/2}} (x-1)^4 \leq \frac{5}{128(1/2)^{7/2}} (x-1)^4 = \frac{5}{16} \sqrt{2} \cdot (x-1)^4$$

Wir können also $C := \frac{5}{16} \sqrt{2}$ setzen und erhalten insgesamt

$$|f(x) - T_3(f, x, 1)| \leq \frac{5}{16} \sqrt{2} \cdot (x-1)^4$$

für $x \in [\frac{1}{2}, \frac{3}{2}]$.

Korollar (Taylor). Sei $x_0 \in D$.

Sei $f : D \rightarrow \mathbb{R}$ eine beliebig oft differenzierbare Funktion.

Die *Taylorreihe* von f zum Entwicklungspunkt x_0 sei

$$T_\infty(f, x, x_0) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

für $x \in \mathbb{R}$, wobei die Konvergenzuntersuchung dieser Reihe noch aussteht.

Sei $x \in D$ gegeben. Falls für jede Folge $(\vartheta_n)_{n \geq 0}$ mit Werten $\vartheta_n \in [0, 1]$ die Folge

$$(\mathbf{R}_n(f, x, x_0, \vartheta_n))_{n \geq 0} = \left(\frac{f^{(n+1)}(x_0 + \vartheta_n(x - x_0))}{(n+1)!} (x - x_0)^{n+1} \right)_{n \geq 0}$$

den Grenzwert 0 hat, dann ist

$$f(x) = T_\infty(f, x, x_0).$$

Die Taylorreihe hat also gegenüber dem Taylorpolynom der Stufe n dann den Vorzug, den exakten Funktionswert zu liefern, nicht nur eine Näherung.

Beispiel. Sei $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x) := \cos(x)$.

Für $m \geq 0$ ist

$$\begin{array}{ll} f^{(4m+0)}(x) = \cos(x) & f^{(4m+0)}(0) = 1 \\ f^{(4m+1)}(x) = -\sin(x) & f^{(4m+1)}(0) = 0 \\ f^{(4m+2)}(x) = -\cos(x) & f^{(4m+2)}(0) = -1 \\ f^{(4m+3)}(x) = \sin(x) & f^{(4m+3)}(0) = 0. \end{array}$$

Und jedenfalls ist $|f^{(k)}(x)| \leq 1$ für $k \in \mathbb{Z}_{\geq 0}$ und $x \in \mathbb{R}$.

Wir erhalten zum einen

$$\begin{aligned} T_\infty(f, x, 0) &= \frac{f^{(0)}(0)}{0!} x^0 + \frac{f^{(1)}(0)}{1!} x^1 + \frac{f^{(2)}(0)}{2!} x^2 + \frac{f^{(3)}(0)}{3!} x^3 + \frac{f^{(4)}(0)}{4!} x^4 + \frac{f^{(5)}(0)}{5!} x^5 + \dots \\ &= \frac{x^0}{0!} - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} \pm \dots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}. \end{aligned}$$

Zum anderen ist für jedes $x \in \mathbb{R}$ und jede Folge $(\vartheta_n)_{n \geq 0}$ mit Werten $\vartheta_n \in [0, 1]$

$$|\mathbf{R}_n(f, x, 0, \vartheta_n)| = \left| \frac{f^{(n+1)}(\vartheta_n x)}{(n+1)!} \cdot x^{n+1} \right| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Dabei ist

$$0 \leq \frac{|x|^{n+1}}{(n+1)!} = \frac{|x|^{m+1}}{(m+1)!} \cdot \frac{|x|}{m+2} \cdot \frac{|x|}{m+3} \cdot \dots \cdot \frac{|x|}{n+1} \leq \frac{|x|^{m+1}}{(m+1)!} \cdot \left(\frac{1}{2}\right)^{n-m}$$

für $2|x| \leq m \leq n$, was dank Sandwich $\lim_{n \rightarrow \infty} \frac{|x|^{n+1}}{(n+1)!} = 0$ nach sich zieht. Was wiederum dank Sandwich $\lim_{n \rightarrow \infty} R_n(f, x, 0, \vartheta_n) = 0$ nach sich zieht.

Somit konvergiert nach Korollar für jedes $x \in \mathbb{R}$ die Taylorreihe gegen die gegebene Funktion. Es ist also für $x \in \mathbb{R}$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} .$$

Beispiel. Sei $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x) := \sin(x)$. Für $m \geq 0$ ist

$$\begin{array}{ll} f^{(4m+0)}(x) = \sin(x) & f^{(4m+0)}(0) = 0 \\ f^{(4m+1)}(x) = \cos(x) & f^{(4m+1)}(0) = 1 \\ f^{(4m+2)}(x) = -\sin(x) & f^{(4m+2)}(0) = 0 \\ f^{(4m+3)}(x) = -\cos(x) & f^{(4m+3)}(0) = -1 . \end{array}$$

Und jedenfalls ist $|f^{(k)}(x)| \leq 1$ für $k \in \mathbb{Z}_{\geq 0}$ und $x \in \mathbb{R}$.

Wir erhalten zum einen

$$\begin{aligned} T_{\infty}(f, x, 0) &= \frac{f^{(0)}(0)}{0!} x^0 + \frac{f^{(1)}(0)}{1!} x^1 + \frac{f^{(2)}(0)}{2!} x^2 + \frac{f^{(3)}(0)}{3!} x^3 + \frac{f^{(4)}(0)}{4!} x^4 + \frac{f^{(5)}(0)}{5!} x^5 + \dots \\ &= \frac{x^1}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} \pm \dots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} . \end{aligned}$$

Zum anderen ist für jedes $x \in \mathbb{R}$ und jede Folge $(\vartheta_n)_{n \geq 0}$ mit Werten $\vartheta_n \in [0, 1]$

$$|R_n(f, x, 0, \vartheta_n)| = \left| \frac{f^{(n+1)}(\vartheta_n x)}{(n+1)!} \cdot x^{n+1} \right| \leq \frac{|x|^{n+1}}{(n+1)!} \rightarrow 0 .$$

Somit konvergiert nach Korollar für jedes $x \in \mathbb{R}$ die Taylorreihe gegen die gegebene Funktion. Es ist also für $x \in \mathbb{R}$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} .$$

4.6.6 Monotonie

Sei $D \subseteq \mathbb{R}$ ein offenes Intervall. Sei $f : D \rightarrow \mathbb{R}$ eine Funktion.

Definition.

- (1) Es heißt f *monoton wachsend*, falls für $x_1, x_2 \in D$ mit $x_1 \leq x_2$ dann $f(x_1) \leq f(x_2)$ ist.
- (2) Es heißt f *monoton fallend*, falls für $x_1, x_2 \in D$ mit $x_1 \leq x_2$ dann $f(x_1) \geq f(x_2)$ ist.

Bemerkung. Sei $f : D \rightarrow \mathbb{R}$ stetig differenzierbar.

- (1) Es ist genau dann $f'(x) \geq 0$ für $x \in D$, wenn f monoton wachsend ist.
- (2) Es ist genau dann $f'(x) \leq 0$ für $x \in D$, wenn f monoton fallend ist.

Zu (1). Sei $f'(x) \geq 0$ für $x \in D$. Seien $x_1, x_2 \in D$ mit $x_1 \leq x_2$ gegeben. Zu zeigen ist $f(x_1) \leq f(x_2)$. Der Satz von Taylor gibt dafür ein $\vartheta \in [0, 1]$ mit

$$f(x_2) = T_0(f, x_2, x_1) + R_0(f, x_2, x_1, \vartheta) = f(x_1) + \underbrace{f'(x_1 + \vartheta(x_2 - x_1))}_{\geq 0} \cdot \underbrace{(x_2 - x_1)}_{\geq 0}.$$

Sei umgekehrt f monoton wachsend. Dann ist $\frac{f(x) - f(x_0)}{x - x_0} \geq 0$ für $x_0, x \in D$ mit $x \neq x_0$ und also auch $f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$ für $x_0 \in D$.

Zu (2) dann analog.

4.6.7 Begriffserweiterung von \mathbb{R} nach \mathbb{C}

Bemerkung. Konvergenz von Folgen und Reihen komplexer Zahlen ist definiert wie im Reellen, wobei nun für $\varepsilon \in \mathbb{R}_{\geq 0}$ und $z_0 \in \mathbb{C}$

$$U_\varepsilon^{\mathbb{C}}(z_0) = \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}$$

bedeute. Es ist also $U_\varepsilon^{\mathbb{C}}(z_0)$ in der Zahlenebene \mathbb{C} eine Kreisscheibe ohne Rand mit Mittelpunkt z_0 und Radius ε .

Wir schreiben ferner $\partial U_\varepsilon^{\mathbb{C}}(z_0) = \{z \in \mathbb{C} : |z - z_0| = \varepsilon\}$ für den Rand der Kreisscheibe $U_\varepsilon^{\mathbb{C}}(z_0)$.

Wir brauchen noch die Festlegung $U_{+\infty}^{\mathbb{C}}(z_0) := \mathbb{C}$ und $\partial U_{+\infty}^{\mathbb{C}}(z_0) := \emptyset$.

Wenn man nun den Betrag überall als Betrag komplexer Zahlen liest, gelten geometrische Reihe, Wurzel-, Quotienten- und Majorantenkriterium weiter. Auch hier ist jede absolut konvergente Reihe auch konvergent.

Bemerkung. Sei $(a_k)_{k \geq \ell}$ mit $\ell \in \mathbb{Z}$ und $a_k \in \mathbb{C}$ gegeben.

Man kann für Konvergenz Real- und Imaginärteil getrennt untersuchen. Beträge vertauschen mit Grenzwerten.

- (1) Sei $a \in \mathbb{C}$. Es ist $\lim_{k \rightarrow \infty} a_k = a$ genau dann, wenn $\lim_{k \rightarrow \infty} \operatorname{Re}(a_k) = \operatorname{Re}(a)$ und $\lim_{k \rightarrow \infty} \operatorname{Im}(a_k) = \operatorname{Im}(a)$ ist.
- (2) Sei $s \in \mathbb{C}$. Es ist $\sum_{k=\ell}^{\infty} a_k = s$ genau dann, wenn $\sum_{k=\ell}^{\infty} \operatorname{Re}(a_k) = \operatorname{Re}(s)$ und $\sum_{k=\ell}^{\infty} \operatorname{Im}(a_k) = \operatorname{Im}(s)$ ist.

(3) Sei $\lim_{k \rightarrow \infty} a_k = a \in \mathbb{C}$. Dann ist $\lim_{k \rightarrow \infty} |a_k| = |a|$.

Beispiel. Es wird

$$\sum_{k=0}^{\infty} \frac{1}{(i+1)^k} = \sum_{k=0}^{\infty} \left(\frac{1}{i+1} \right)^k = \frac{1}{1 - \frac{1}{i+1}} = \frac{i+1}{i+1-1} = 1-i,$$

was zulässig ist, da $\left| \frac{1}{i+1} \right| = \frac{1}{\sqrt{2}} < 1$ ist.

Insbesondere ist diese Reihe konvergent.

Das Wurzelkriterium liefert ebenfalls, daß diese Reihe konvergent ist. Denn

$$\left| \frac{1}{(i+1)^k} \right|^{1/k} = \left(\frac{1}{|i+1|^k} \right)^{1/k} = \frac{1}{|i+1|} = \frac{1}{\sqrt{2}} \rightarrow \frac{1}{\sqrt{2}} < 1.$$

Das Quotientenkriterium liefert ebenfalls, daß diese Reihe konvergent ist. Denn

$$\left| \frac{\frac{1}{(i+1)^{k+1}}}{\frac{1}{(i+1)^k}} \right| = \left| \frac{1}{i+1} \right| = \frac{1}{\sqrt{2}} \rightarrow \frac{1}{\sqrt{2}} < 1.$$

Natürlich ist das ein besonders einfacher Fall, bei beiden Kriterien letzten Endes auf eine konstante Folge gestoßen zu sein. Und natürlich interessieren einen beide Kriterien normalerweise nicht mehr, sobald man sogar ausrechnen konnte, gegen welchen Grenzwert die Reihe konvergiert.

Definition. Sei $D \subseteq \mathbb{R}$ offen. Sei $f : D \rightarrow \mathbb{C}$ eine Funktion.

Für $x \in D$ sei $f'(x) := \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$, sofern existent.

Bemerkung.

- (1) Die Ableitungsregeln gelten auch für Funktionen auf $D \subseteq \mathbb{R}$ mit Werten in \mathbb{C} .
- (2) Schreiben wir $f(x) = a(x) + ib(x)$ mit $a(x), b(x) \in \mathbb{R}$ für $x \in D$, dann ist

$$f'(x) = a'(x) + ib'(x).$$

Man kann also die Ableitung berechnen, indem man Real- und Imaginärteil ableitet.

- (3) Es heißt f stetig, falls $D \rightarrow \mathbb{R} : x \mapsto \operatorname{Re}(f(x))$ und $D \rightarrow \mathbb{R} : x \mapsto \operatorname{Im}(f(x))$ stetig sind.

Beispiel. Für $x \in \mathbb{R}$ setzen wir $f(x) := \frac{1}{ix+1} \in \mathbb{C}$. Dann ist $f'(x) = -\frac{1}{(ix+1)^2} \cdot i$.

Bemerkung. Für Taylorpolynome und das Restglied nach Lagrange braucht man reellwertige Funktionen. Will man dennoch für eine komplexwertige Funktion eine Restgliedabschätzung erreichen, so muß man dies für Real- und Imaginärteil separat durchführen.

4.6.8 Potenzreihen

Definition. Sei $z_0 \in \mathbb{C}$. Sei $(a_n)_{n \geq 0}$ eine Folge komplexer Zahlen.

Dann heißt

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

die *Potenzreihe* mit Koeffizienten a_k für $k \geq 0$ um den *Entwicklungspunkt* z_0 .

Definition. Sei eine Potenzreihe $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ gegeben, mit Koeffizienten $a_k \in \mathbb{C}$ für $k \geq 0$ und Entwicklungspunkt $z_0 \in \mathbb{C}$.

Wir bilden die *Wurzelfolge* $(|a_k|^{1/k})_{k \geq 1}$ dieser Potenzreihe. Es heißt

$$\rho := \left(\limsup_{k \rightarrow \infty} |a_k|^{1/k} \right)^{-1}$$

der *Konvergenzradius* und $U_\rho^{\mathbb{C}}(z_0)$ die *Konvergenzkreisscheibe* dieser Potenzreihe.

Hierbei ist gegebenenfalls $0^{-1} = +\infty$ und $(+\infty)^{-1} = 0$ zu lesen.

Lemma. Sei $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ gegeben, mit $a_k \in \mathbb{C}$ für $k \geq 0$ und $z_0 \in \mathbb{C}$.

Sei ρ ihr Konvergenzradius.

Es konvergiert $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ absolut für $z \in U_\rho^{\mathbb{C}}(z_0)$

Es konvergiert $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ nicht für $z \in \mathbb{C} \setminus (U_\rho^{\mathbb{C}}(z_0) \cup \partial U_\rho^{\mathbb{C}}(z_0))$.

Mit anderen Worten, die Reihe konvergiert absolut für ein z auf der Konvergenzkreisscheibe; die Reihe konvergiert nicht für ein z , das weder auf der Konvergenzkreisscheibe noch auf ihrem Rand liegt.

Das Lemma macht keine Aussage, was für $z \in \partial U_\rho^{\mathbb{C}}(z_0)$ passiert, also falls z auf dem Rand der Konvergenzkreisscheibe liegt.

Bemerkung. Sei $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ eine Potenzreihe. Gebe es ein m mit $a_k \neq 0$ für $k \geq m$.

Wir bilden die *Quotientenfolge* $(\frac{|a_{k+1}|}{|a_k|})_{k \geq m}$ dieser Potenzreihe. Ist $\lim_{k \rightarrow \infty} \frac{|a_{k+1}|}{|a_k|}$ existent, dann ist der Konvergenzradius unserer Potenzreihe auch gegeben durch

$$\rho = \left(\lim_{k \rightarrow \infty} \frac{|a_{k+1}|}{|a_k|} \right)^{-1}.$$

Vorsicht, falls der Limes der Quotientenfolge nicht existiert, darf nicht ersatzweise der Limes superior verwandt werden, um den Konvergenzradius zu berechnen.

Beispiel. Wir betrachten die Potenzreihe $\sum_{k=1}^{\infty} \frac{1}{k} (z - i)^k$.

Der Entwicklungspunkt ist $z_0 = i$. Die Koeffizienten sind $a_k = \frac{1}{k}$ für $k \geq 1$ und $a_0 = 0$.

Für die Wurzelfolge gilt

$$|a_k|^{1/k} = \left| \frac{1}{k} \right|^{1/k} = \frac{1}{k^{1/k}} \rightarrow \frac{1}{1} = 1.$$

Der Konvergenzradius ist somit

$$\rho = (\limsup_{k \rightarrow \infty} |a_k|^{1/k})^{-1} = (\lim_{k \rightarrow \infty} |a_k|^{1/k})^{-1} = 1^{-1} = 1.$$

Wir betrachten alternativ hierzu die Quotientenfolge. Es ist

$$\frac{|a_{k+1}|}{|a_k|} = \frac{1/(k+1)}{1/k} = \frac{k}{k+1} \rightarrow 1.$$

Dies bestätigt den Konvergenzradius $\rho = (\lim_{k \rightarrow \infty} \frac{|a_{k+1}|}{|a_k|})^{-1} = 1^{-1} = 1$.

Folglich ist $\sum_{k=1}^{\infty} \frac{1}{k} (z-i)^k$ absolut konvergent für $z \in U_1^{\mathbb{C}}(i)$, d.h. für $z \in \mathbb{C}$ mit $|z-i| < 1$. Die Reihe ist nicht konvergent für $z \in \mathbb{C} \setminus (U_1^{\mathbb{C}}(i) \cup \partial U_1^{\mathbb{C}}(i))$, d.h. für $z \in \mathbb{C}$ mit $|z-i| > 1$.

Ableiten wollen wir hingegen eine Potenzreihe nur, falls ihr Entwicklungspunkt x_0 in \mathbb{R} liegt und ihr Argument x aus \mathbb{R} stammt:

Bemerkung. Sei $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ gegeben, mit $a_k \in \mathbb{C}$ für $k \geq 0$ und $x_0 \in \mathbb{R}$.

Sei ρ ihr Konvergenzradius.

Dann können wir definieren:

$$f : U_{\rho}(x_0) =]x_0 - \rho, x_0 + \rho[\rightarrow \mathbb{C} : x \mapsto f(x) := \sum_{k=0}^{\infty} a_k (x - x_0)^k.$$

Für $x \in U_{\rho}(x_0)$ wird

$$f'(x) = \sum_{k=1}^{\infty} a_k k (x - x_0)^{k-1} = \sum_{k=0}^{\infty} a_{k+1} (k+1) (x - x_0)^k.$$

D.h. wir dürfen summandenweise ableiten.

Insbesondere ist f beliebig oft differenzierbar und also auch stetig.

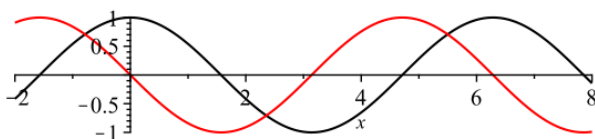
Beispiel. Sei für $x \in \mathbb{R}$

$$f(x) := \cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}.$$

Es wird

$$f'(x) = \sum_{k=1}^{\infty} \frac{(-1)^k}{(2k)!} 2k x^{2k-1} = - \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{(2k-1)!} x^{2k-1} = - \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = -\sin(x).$$

Das bestätigt, was wir schon wußten.



4.6.9 Exponentialfunktion

Definition. Für $z \in \mathbb{C}$ sei

$$e^z := \exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!} = \frac{z^0}{0!} + \frac{z^1}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots$$

Diese Potenzreihe konvergiert für $z \in \mathbb{C}$, da ihr Konvergenzradius

$$\rho = \left(\limsup_{k \rightarrow \infty} \left| \frac{1}{k!} \right|^{1/k} \right)^{-1} = \left(\lim_{k \rightarrow \infty} \frac{1}{(k!)^{1/k}} \right)^{-1} = 0^{-1} = +\infty$$

ist. Wir erhalten also eine Funktion

$$\exp : \mathbb{C} \rightarrow \mathbb{C} : z \mapsto \exp(z) .$$

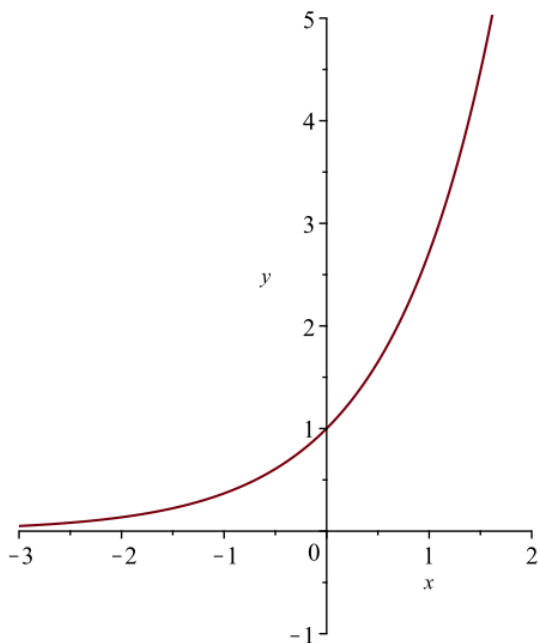
Bemerkung.

Es ist $\exp(0) = \sum_{k=0}^{\infty} \frac{0^k}{k!} = 1 + 0 + 0 + \dots = 1$.

Es ist $\exp(1) = e \approx 2,718281828459$. Auf den Zusammenhang zur Formel aus §4.3 gehen wir unten noch ein.

Es ist $\exp(x) \geq 1 + x$ für $x \in \mathbb{R}_{\geq 0}$. Insbesondere ist $\lim_{x \rightarrow \infty} \exp(x) = +\infty$.

Der reelle Graph der Exponentialfunktion hat folgende Gestalt.



Lemma. Es ist $\exp(z+w) = \exp(z) \cdot \exp(w)$ für $z, w \in \mathbb{C}$.

Denn

$$\begin{aligned}
 & \exp(z) \cdot \exp(w) \\
 = & \left(\sum_{k=0}^{\infty} \frac{z^k}{k!} \right) \cdot \left(\sum_{\ell=0}^{\infty} \frac{w^\ell}{\ell!} \right) \\
 = & \left(\frac{z^0}{0!} + \frac{z^1}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \right) \cdot \left(\frac{w^0}{0!} + \frac{w^1}{1!} + \frac{w^2}{2!} + \frac{w^3}{3!} + \dots \right) \\
 = & \left(\frac{z^0}{0!} \frac{w^0}{0!} \right) + \left(\frac{z^1}{1!} \frac{w^0}{0!} + \frac{z^0}{0!} \frac{w^1}{1!} \right) + \left(\frac{z^2}{2!} \frac{w^0}{0!} + \frac{z^1}{1!} \frac{w^1}{1!} + \frac{z^0}{0!} \frac{w^2}{2!} \right) + \left(\frac{z^3}{3!} \frac{w^0}{0!} + \frac{z^2}{2!} \frac{w^1}{1!} + \frac{z^1}{1!} \frac{w^2}{2!} + \frac{z^0}{0!} \frac{w^3}{3!} \right) + \dots \\
 = & \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{z^{n-k}}{(n-k)!} \frac{w^k}{k!} \right) \\
 = & \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{1}{n!} \binom{n}{k} z^{n-k} w^k \right) \\
 = & \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n \\
 = & \exp(z+w).
 \end{aligned}$$

Im vorletzten Schritt wird der Binomische Lehrsatz aus §1.8 verwendet.

Die Anwendung des Distributivitätsgesetzes auf unendliche Summen in der vorstehenden Rechnung sollte man noch rechtfertigen, aber das ist eine mathematische Feinheit; Stichwort Cauchy-Produktformel.

Bemerkung. Für $n \in \mathbb{Z}_{\geq 0}$ ist mit vorstehendem Lemma, iteriert angewandt, nun $\exp(n) = \exp(1)^n = e^n$. Dies verwendet man zur Rechtfertigung der Ausdehnung dieser Schreibweise zu $\exp(z) = e^z$ für $z \in \mathbb{C}$.

Lemma. Für $x \in \mathbb{R}$ ist

$$\frac{d}{dx} \exp(x) = \exp(x).$$

Denn summandenweises Ableiten gibt

$$\frac{d}{dx} \exp(x) = \frac{d}{dx} \sum_{k=0}^{\infty} \frac{x^k}{k!} = \sum_{k=1}^{\infty} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp(x).$$

Beispiel. Es ist $(e^{x^2-1})' = e^{x^2-1} \cdot 2x$ nach Kettenregel.

Weiter wird $(e^{x^2-1})'' = (e^{x^2-1} \cdot 2x)' = e^{x^2-1} \cdot 2x \cdot 2x + e^{x^2-1} \cdot 2 = e^{x^2-1} \cdot (4x^2 + 2)$ nach Produkt- und Kettenregel.

Beispiel. Es ist $\lim_{x \rightarrow +\infty} \frac{x}{e^x} \stackrel{\text{r.H.}}{=} \lim_{x \rightarrow +\infty} \frac{1}{e^x} = 0$.

Lemma (de Moivre). Für $z \in \mathbb{C}$ ist

$$\exp(iz) = \cos(z) + i \sin(z).$$

Hierbei haben wir die Potenzreihen für \sin und \cos auch im Komplexen gelten lassen.

Denn

$$\begin{aligned}
 & \exp(iz) \\
 = & \frac{(iz)^0}{0!} + \frac{(iz)^1}{1!} + \frac{(iz)^2}{2!} + \frac{(iz)^3}{3!} + \frac{(iz)^4}{4!} + \frac{(iz)^5}{5!} + \frac{(iz)^6}{6!} + \frac{(iz)^7}{7!} + \frac{(iz)^8}{8!} + \frac{(iz)^9}{9!} + \frac{(iz)^{10}}{10!} + \frac{(iz)^{11}}{11!} + \dots \\
 = & \frac{z^0}{0!} + i \frac{z^1}{1!} - \frac{z^2}{2!} - i \frac{z^3}{3!} + \frac{z^4}{4!} + i \frac{z^5}{5!} - \frac{z^6}{6!} - i \frac{z^7}{7!} + \frac{z^8}{8!} + i \frac{z^9}{9!} - \frac{z^{10}}{10!} - i \frac{z^{11}}{11!} + \dots \\
 = & \left(\frac{z^0}{0!} - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \frac{z^8}{8!} - \frac{z^{10}}{10!} \pm \dots \right) + i \left(\frac{z^1}{1!} - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \frac{z^9}{9!} - \frac{z^{11}}{11!} \pm \dots \right) \\
 = & \cos(z) + i \sin(z) .
 \end{aligned}$$

Beispiel. Es ist $e^{i\pi} = \exp(i\pi) = \cos(\pi) + i \sin(\pi) = -1$.

Bemerkung. Für $z \in \mathbb{C}$ folgt

$$\begin{aligned}
 \cos(z) &= \frac{1}{2}((\cos(z) + i \sin(z)) + (\cos(-z) + i \sin(-z))) = \frac{1}{2}(e^{iz} + e^{-iz}) \\
 \sin(z) &= \frac{1}{2i}((\cos(z) + i \sin(z)) - (\cos(-z) + i \sin(-z))) = \frac{1}{2i}(e^{iz} - e^{-iz}) .
 \end{aligned}$$

Beispiel. Wir leiten den Satz des Pythagoras her. Für $z \in \mathbb{C}$ wird

$$\begin{aligned}
 \sin(z)^2 + \cos(z)^2 &= \left(\frac{1}{2i}(e^{iz} - e^{-iz}) \right)^2 + \left(\frac{1}{2}(e^{iz} + e^{-iz}) \right)^2 \\
 &= \frac{1}{4}(-e^{2iz} - 2e^{iz}e^{-iz} + e^{-2iz}) + (e^{2iz} + 2e^{iz}e^{-iz} + e^{-2iz}) \\
 &= \frac{1}{4} \cdot 4 = 1 .
 \end{aligned}$$

Definition. Für $z \in \mathbb{C}$ sei

$$\begin{aligned}
 \cosh(z) &:= \frac{1}{2}(e^z + e^{-z}) \\
 \sinh(z) &:= \frac{1}{2}(e^z - e^{-z})
 \end{aligned}$$

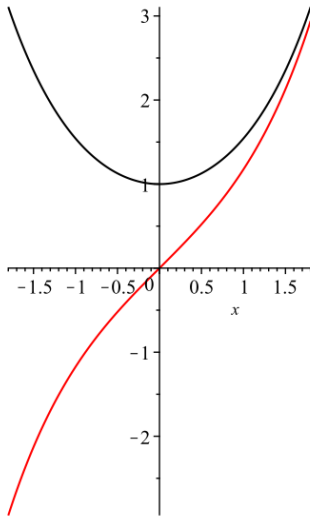
genannt *Cosinus hyperbolicus* und *Sinus hyperbolicus*.

Es ist $\cosh(-z) = \cosh(z)$ und $\sinh(-z) = -\sinh(z)$ für $z \in \mathbb{C}$.

Es ist $\cosh(z)^2 - \sinh(z)^2 = \frac{1}{4}((e^{2z} + 2 + e^{-2z}) - (e^{2z} - 2 + e^{-2z})) = 1$ für $z \in \mathbb{C}$.

Die Definition gibt $\frac{d}{dx} \cosh(x) = \sinh(x)$ und $\frac{d}{dx} \sinh(x) = \cosh(x)$ für $x \in \mathbb{R}$.

Die reellen Graphen dieser Funktionen haben folgende Gestalt. Der Sinus hyperbolicus in rot, der Cosinus hyperbolicus in schwarz.



Den Graphen von \cosh kennt man von der Form einer hängenden Kette.

Bemerkung. Für $z \in \mathbb{C}$ wird nach Binomischem Lehrsatz

$$\left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^n \binom{n}{k} 1^{n-k} \cdot \left(\frac{z}{n}\right)^k = \sum_{k=0}^n \binom{n}{k} n^{-k} z^k .$$

Für festes k und laufendes $n \in \mathbb{Z}_{\geq k}$ wird aus dem Koeffizienten bei z^k beim Grenzübergang $n \rightarrow \infty$ somit

$$\binom{n}{k} n^{-k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!} n^{-k} = \frac{1}{k!} \cdot \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-k+1}{n} \rightarrow \frac{1}{k!} .$$

Also ist auch

$$\lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{\infty} \frac{z^k}{k!} = e^z .$$

Dies erklärt auch, warum $e = e^1 = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$ ist; vgl. drittletztes Beispiel in §4.3.

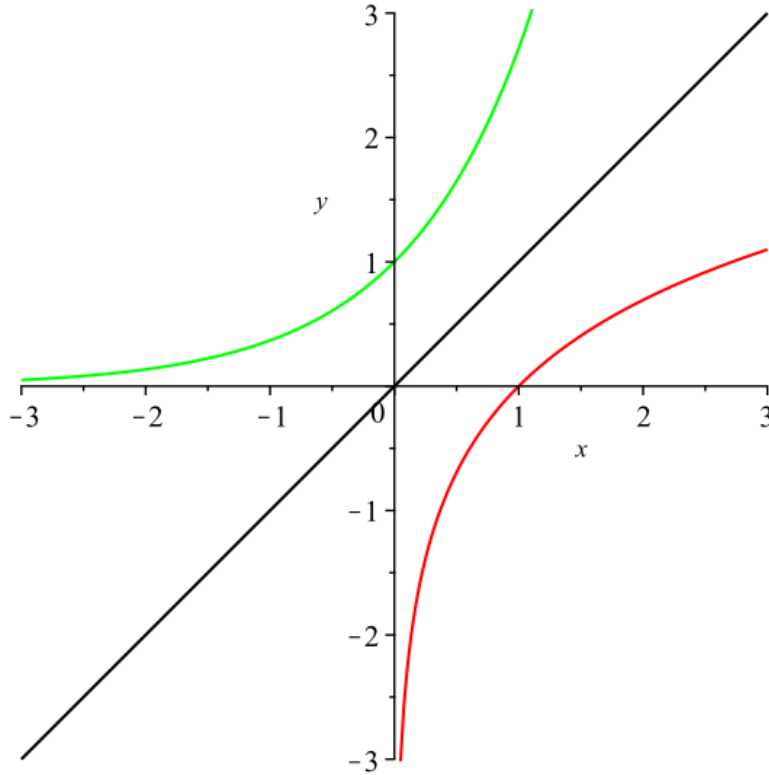
4.6.10 Logarithmus

Die Abbildung $\exp : \mathbb{R} \rightarrow \mathbb{R}_{>0}$, die wir noch im Zielbereich auf $\mathbb{R}_{>0}$ eingeschränkt haben, ist bijektiv, wie man am Graph der Funktion erkennen kann.

Definition. Sei $\ln : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ die Umkehrabbildung von \exp , genannt *Logarithmus* (lat. logarithmus naturalis).

Es ist $\ln(\exp(x)) = x$ für $x \in \mathbb{R}$. Es ist $\exp(\ln(x)) = x$ für $x \in \mathbb{R}_{>0}$.

Der Graph $y = \ln(x)$, in rot, ergibt sich aus dem Graphen $y = \exp(x) = e^x$, in grün, durch Spiegelung an der Geraden $y = x$.



Beispiel. Es ist $\ln(1) = \ln(\exp(0)) = 0$. Es ist $\ln(e) = \ln(\exp(1)) = 1$.

Bemerkung. Für $a, b \in \mathbb{R}_{>0}$ ist $\ln(a \cdot b) = \ln(a) + \ln(b)$ und $\ln(\frac{1}{a}) = -\ln(a)$.

Denn $\ln(a \cdot b) = \ln(\exp(\ln(a)) \cdot \exp(\ln(b))) = \ln(\exp(\ln(a) + \ln(b))) = \ln(a) + \ln(b)$.

Und also $0 = \ln(1) = \ln(a \cdot \frac{1}{a}) = \ln(a) + \ln(\frac{1}{a})$.

Bemerkung. Für $b \in \mathbb{R}_{>1}$ setzt man $\log_b(x) := \frac{\ln(x)}{\ln(b)}$ für $x \in \mathbb{R}_{>0}$.

Definition. Sei $a \in \mathbb{R}_{>0}$. Sei $z \in \mathbb{C}$. Sei

$$a^z := e^{\ln(a) \cdot z}.$$

Für $z \in \mathbb{Q}$ gibt das die übliche Potenz von a mit Exponent z .

Bemerkung. Seien $a, b \in \mathbb{R}_{>0}$. Seien $z, w \in \mathbb{C}$.

(1) Es ist $a^z \cdot b^z = (a \cdot b)^z$.

(2) Es ist $a^{z+w} = a^z \cdot a^w$.

(3) Es ist $(a^b)^z = a^{b \cdot z}$.

(4) Es ist $\frac{d}{dx} a^x = \ln(a) \cdot a^x$ für $x \in \mathbb{R}$.

(5) Sei $b > 1$. Dann ist $b^{\log_b(x)} = x$ für $x \in \mathbb{R}_{>0}$ und $\log_b(b^y) = y$ für $y \in \mathbb{R}$.

Zu (1). Es ist $a^z \cdot b^z = e^{\ln(a) \cdot z} \cdot e^{\ln(b) \cdot z} = e^{\ln(a) \cdot z + \ln(b) \cdot z} = e^{(\ln(a) + \ln(b)) \cdot z} = e^{\ln(a \cdot b) \cdot z} = (a \cdot b)^z$.

Zu (2). Es ist $a^{z+w} = e^{\ln(a) \cdot (z+w)} = e^{\ln(a) \cdot z + \ln(a) \cdot w} = e^{\ln(a) \cdot z} \cdot e^{\ln(a) \cdot w} = a^z \cdot a^w$.

Zu (3). Es ist $(a^b)^z = e^{\ln(a^b) \cdot z} = e^{\ln(e^{\ln(a) \cdot b}) \cdot z} = e^{\ln(a) \cdot b \cdot z} = a^{b \cdot z}$.

Zu (4). Es ist $\frac{d}{dx} a^x = \frac{d}{dx} e^{\ln(a) \cdot x} = e^{\ln(a) \cdot x} \cdot \ln(a) = \ln(a) \cdot a^x$.

Zu (5). Es ist $b^{\log_b(x)} = e^{\ln(b) \cdot \log_b(x)} = e^{\ln(x)} = x$.

Es ist $\log_b(b^y) = \frac{\ln(b^y)}{\ln(b)} = \frac{\ln(e^{\ln(b) \cdot y})}{\ln(b)} = \frac{\ln(b) \cdot y}{\ln(b)} = y$.

Beispiel. Für $a, b \in \mathbb{R}_{>0}$ und $z \in \mathbb{C}$ ist $a^{b^z} := a^{(b^z)} = e^{\ln(a) \cdot b^z}$ und $(a^b)^z = e^{\ln(a) \cdot b \cdot z}$. Und im allgemeinen ist $b^z \neq b \cdot z$. Somit ist im allgemeinen auch $a^{b^z} \neq (a^b)^z$.

Beispiel. Sei $c \in \mathbb{R}_{>0}$. Es ist

$$\lim_{x \rightarrow +\infty} c^{1/x} = \lim_{x \rightarrow +\infty} \exp(\ln(c) \cdot \frac{1}{x}) = \exp(\ln(c)) \cdot \lim_{x \rightarrow +\infty} \frac{1}{x} = \exp(0) = 1.$$

Vgl. zweites Beispiel in §4.3.

4.6.11 Ableitung von Umkehrfunktionen

Sei $D \subseteq \mathbb{R}$ ein offenes Intervall. Sei $E \subseteq \mathbb{R}$ ein offenes Intervall.

Sei $f : D \rightarrow E$ differenzierbar und bijektiv.

Sei $f'(x) > 0$ für $x \in D$ oder aber $f'(x) < 0$ für $x \in D$.

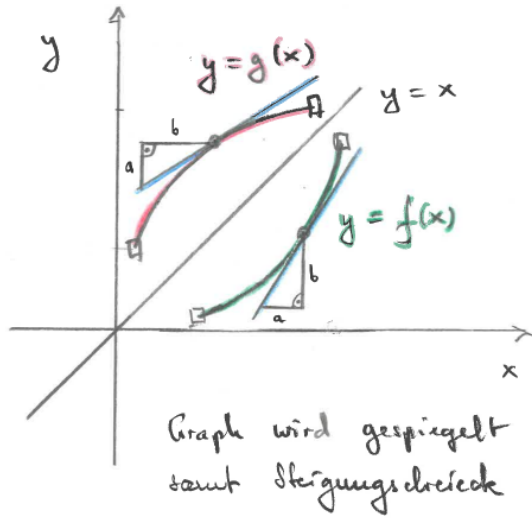
Sei $g : E \rightarrow D$ die Umkehrfunktion von f .

Es ist also $g(f(x)) = x$ für $x \in D$ und $f(g(x)) = x$ für $x \in E$.

Lemma. Für $x \in E$ ist

$$g'(x) = \frac{1}{f'(g(x))}.$$

Denn es ist $1 = \frac{d}{dx} x = \frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x)$.



Beispiel. Wir betrachten die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}_{>0}: x \mapsto f(x) := \exp(x)$.

Ihre Umkehrfunktion ist $g: \mathbb{R}_{>0} \rightarrow \mathbb{R}: x \mapsto g(x) = \ln(x)$.

Wir erhalten

$$\frac{d}{dx} \ln(x) = \frac{1}{x}.$$

Denn es ist $f'(x) = \exp(x)$ und also

$$\frac{d}{dx} \ln(x) = g'(x) = \frac{1}{f'(g(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$

Beispiel. Sei $a \in \mathbb{R} \setminus \{0\}$. Sei $x \in \mathbb{R}_{>0}$. Es wird

$$\frac{d}{dx} x^a = \frac{d}{dx} \exp(\ln(x) \cdot a) = \exp(\ln(x) \cdot a) \cdot \frac{1}{x} \cdot a = x^a \cdot \frac{1}{x} \cdot a = ax^{a-1}.$$

Beispiel. Es ist

$$\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x} \stackrel{\text{r.H.}}{=} \lim_{x \rightarrow +\infty} \frac{1/x}{1} = 0.$$

Also wird

$$\lim_{x \rightarrow +\infty} x^{1/x} = \lim_{x \rightarrow +\infty} \exp(\ln(x) \cdot 1/x) = \exp\left(\lim_{x \rightarrow +\infty} \frac{\ln(x)}{x}\right) \stackrel{\text{s.o.}}{=} \exp(0) = 1.$$

Vgl. drittes Beispiel in §4.3.

Beispiel. Es ist $\lim_{x \rightarrow 0} x \ln(x) = \lim_{x \rightarrow 0} \frac{\ln(x)}{x^{-1}} \stackrel{\text{r.H.}}{=} \lim_{x \rightarrow 0} \frac{x^{-1}}{-x^{-2}} = \lim_{x \rightarrow 0} -x = 0$.

Definition. Sei $f : (-\frac{\pi}{2}, +\frac{\pi}{2}) \rightarrow \mathbb{R} : x \mapsto f(x) := \tan(x) = \frac{\sin(x)}{\cos(x)}$.

Ihre Umkehrfunktion sei $g : \mathbb{R} \rightarrow (-\frac{\pi}{2}, +\frac{\pi}{2}) : x \mapsto g(x) := \arctan(x)$, genannt *Arcustangens*.

Es ist also $\tan(\arctan(x)) = x$ für $x \in \mathbb{R}$ und $\arctan(\tan(x)) = x$ für $x \in (-\frac{\pi}{2}, +\frac{\pi}{2})$.

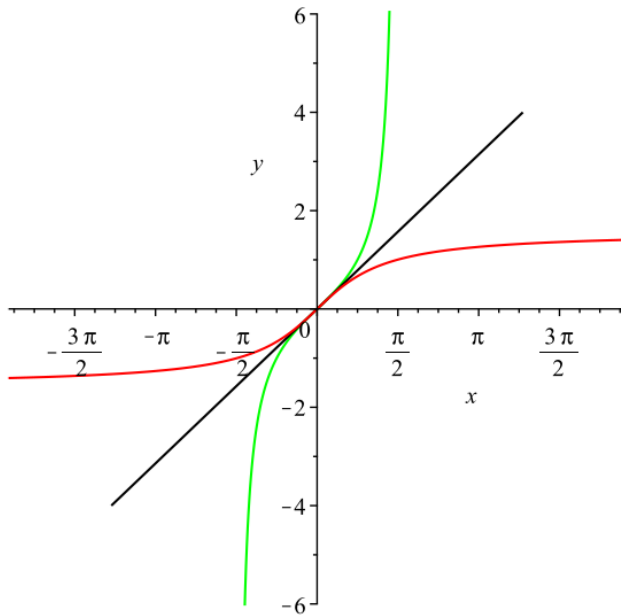
Wir erhalten

$$\frac{d}{dx} \arctan(x) = \frac{1}{1+x^2}.$$

Denn es ist $f'(x) = 1 + \tan(x)^2$ und also

$$\frac{d}{dx} \arctan(x) = g'(x) = \frac{1}{f'(g(x))} = \frac{1}{1 + \tan(\arctan(x))^2} = \frac{1}{1+x^2}.$$

Der Graph $y = \tan(x)$, eingeschränkt auf $(-\frac{\pi}{2}, +\frac{\pi}{2})$, in grün. Der Graph $y = \arctan(x)$ in rot. Diese beiden Graphen gehen durch Spiegelung an $y = x$ auseinander hervor.



Daher ist $\lim_{x \rightarrow +\infty} \arctan(x) = \frac{\pi}{2}$ und $\arctan(-x) = -\arctan(x)$ für $x \in \mathbb{R}$.

4.7 Integralrechnung

4.7.1 Integralbegriff

Seien $a, b \in \mathbb{R}$ mit $a \leq b$ gegeben.

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion. Gebe es $s, t \in \mathbb{R}$ mit $s \leq f(x) \leq t$ für $x \in [a, b]$.

Definition. Sei $M \subseteq \mathbb{R}$ eine nichtleere Teilmenge.

(1) Es heißt $c \in \mathbb{R}$ eine *obere Schranke* an M , falls für $x \in M$ sich $x \leq c$ ergibt.

Falls die Menge der oberen Schranken von M nicht leer ist, so hat sie ein Minimum, genannt *Supremum* von M , geschrieben $\sup(M)$.

Falls das Maximum von M existiert, dann ist $\sup(M) = \max(M)$.

(2) Es heißt $c \in \mathbb{R}$ eine *untere Schranke* an M , falls für $x \in M$ sich $x \geq c$ ergibt.

Falls die Menge der unteren Schranken von M nicht leer ist, so hat sie ein Maximum, genannt *Infimum* von M , geschrieben $\inf(M)$.

Falls das Minimum von M existiert, dann ist $\inf(M) = \min(M)$.

Das Supremum ist ein verallgemeinertes Maximum.

Das Infimum ist ein verallgemeinertes Minimum.

Beispiel. Sei $M := [0, 1[$.

Die Menge der oberen Schranken von M ist $[1, +\infty[$. Ihr Minimum ist $\sup([0, 1[) = 1$.

Man beachte, daß $[0, 1[$ kein Maximum hat, aber ein Supremum.

Die Menge der unteren Schranken von M ist $]-\infty, 0]$. Ihr Maximum ist $\inf([0, 1[) = 0$.

Man beachte, daß $[0, 1[$ ein Minimum hat und infolgedessen $\inf([0, 1[) = \min([0, 1[)$ ist.

Definition. Eine *Unterteilung* von $[a, b]$ ist ein Tupel

$$\underline{x} := (x_0, x_1, \dots, x_k)$$

mit $k \in \mathbb{Z}_{\geq 1}$, mit $x_j \in \mathbb{R}$ für $1 \leq i \leq k$ und mit

$$a = x_0 < x_1 < x_2 < \dots < x_k = b.$$

Sei $\mathcal{U}([a, b])$ die Menge der Unterteilungen von $[a, b]$.

Definition. Sei $\underline{x} = (x_0, \dots, x_k)$ eine Unterteilung von $[a, b]$. Es ist also $\underline{x} \in \mathcal{U}([a, b])$.

(1) Die *Untersumme* von f bezüglich der Unterteilung \underline{x} ist

$$\text{Unter}(f, \underline{x}) := \sum_{j=1}^k (x_j - x_{j-1}) \cdot \inf\{f(x) : x \in [x_{j-1}, x_j]\}.$$

(2) Die *Obersumme* von f bezüglich der Unterteilung \underline{x} ist

$$\text{Ober}(f, \underline{x}) := \sum_{j=1}^k (x_j - x_{j-1}) \cdot \sup\{f(x) : x \in [x_{j-1}, x_j]\}.$$

Definition. Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt *integrierbar* im Sinne von Riemann, wenn in

$$\sup (\{ \text{Unter}(f, \underline{x}) : \underline{x} \in \mathcal{U}([a, b]) \}) \leq \inf (\{ \text{Ober}(f, \underline{x}) : \underline{x} \in \mathcal{U}([a, b]) \})$$

eine Gleichheit vorliegt. Diesfalls heißt

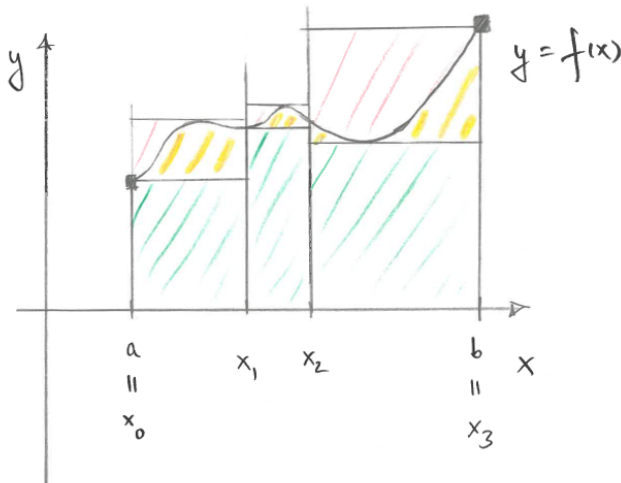
$$\int_a^b f(x) dx := \sup (\{ \text{Unter}(f, \underline{x}) : \underline{x} \in \mathcal{U}([a, b]) \}) = \inf (\{ \text{Ober}(f, \underline{x}) : \underline{x} \in \mathcal{U}([a, b]) \})$$

das *Integral* von f über $[a, b]$.

Insbesondere ist $\text{Unter}(f, \underline{x}) \leq \int_a^b f(x) dx \leq \text{Ober}(f, \underline{x})$ für $\underline{x} \in \mathcal{U}([a, b])$.

Es heißen $f(x)$ der *Integrand* und a und b die *Integrationsgrenzen* dieses Integrals.

Das Integral stammt in dieser Form wohl von Darboux. Riemann hatte eine Variante davon unter Verwendung von Zwischenpunkten entwickelt.



Der Inhalt der grünen Fläche ist die Untersumme. Der Inhalt der grünen und der gelben Fläche ist das Integral. Der Inhalt der grünen, der gelben und der roten Fläche ist die Obersumme.

Beispiel. Sei $n \in \mathbb{Z}_{\geq 1}$. Sei $f : [1, n+1] \rightarrow \mathbb{R} : x \mapsto f(x) := \ln(x)$.

Es ist $f'(x) = \frac{1}{x} > 0$ für $x \in [1, n+1]$.

Sei $a := 1$. Sei $b := n+1$.

Sei $\underline{x} = (x_0, x_1, \dots, x_n) := (1, 2, \dots, n+1)$. Sei also $x_j := j+1$ für $0 \leq j \leq n$.

Es ist $\inf \{ f(x) : x \in [j, j+1] \} = \min \{ \ln(x) : x \in [j, j+1] \} = \ln(j)$ für $0 \leq j \leq n$.

Es ist $\sup\{f(x) : x \in [j, j+1]\} = \max\{\ln(x) : x \in [j, j+1]\} = \ln(j+1)$ für $0 \leq j \leq n$.

Somit ist

$$\begin{aligned}\text{Unter}(f, \underline{x}) &= \sum_{j=1}^n ((j+1) - j) \cdot \ln(j) = \sum_{j=1}^n \ln(j) = \ln(n!) \\ \text{Ober}(f, \underline{x}) &= \sum_{j=1}^n ((j+1) - j) \cdot \ln(j+1) = \sum_{j=1}^n \ln(j+1) = \sum_{j=2}^{n+1} \ln(j) = \ln((n+1)!).\end{aligned}$$

Es folgt

$$\ln(n!) = \text{Unter}(f, \underline{x}) \leq \int_1^{n+1} \ln(x) \, dx \leq \text{Ober}(f, \underline{x}) = \ln((n+1)!).$$

Wir werden dieses Integral auch noch ausrechnen, nicht nur abschätzen.

Beispiel. Sei $f : [0, 1] \rightarrow \mathbb{R}$ dadurch definiert, daß $f(x) := 1$ ist, falls $x \in \mathbb{Q}$ liegt, und $f(x) := 0$ ist, falls $x \in \mathbb{R} \setminus \mathbb{Q}$ liegt.

Da hier bei jeder Unterteilung \underline{x} in jedem Intervall $[x_{j-1}, x_j]$ sowohl ein Element von \mathbb{Q} als auch ein Element von $\mathbb{R} \setminus \mathbb{Q}$ liegt, ist immer $\text{Unter}(f, \underline{x}) = 0$ und $\text{Ober}(f, \underline{x}) = 1$.

Das Supremum aller Untersummen, also 0, ist hier echt kleiner als das Infimum aller Obersummen, also 1.

Somit ist f nicht integrierbar.

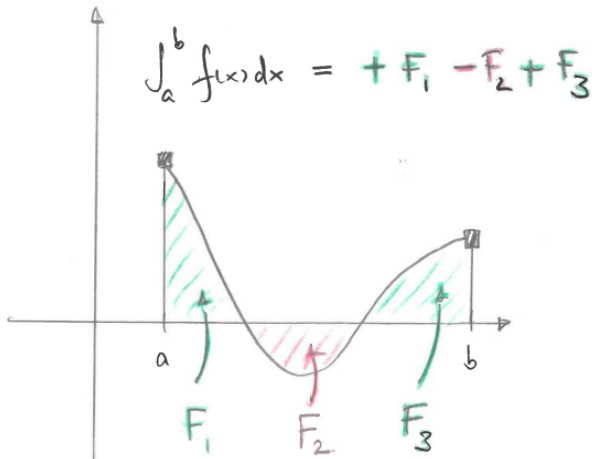
4.7.2 Eigenschaften des Integrals

Seien $a, b \in \mathbb{R}$ mit $a \leq b$ gegeben. Sei $f : [a, b] \rightarrow \mathbb{R}$.

Bemerkung. Ist f stetig, so ist f integrierbar.

Bemerkung. Das Integral $\int_a^b f(x) \, dx$ mißt den Flächeninhalt zwischen dem Graphen von $f : [a, b] \rightarrow \mathbb{R}$ und der x -Achse, wobei die unter der x -Achse liegenden Flächenstücke negativ zu nehmen sind.

Genauer gesagt kann man den Flächeninhalt so definieren.



Bemerkung

(1) Seien $u, v : [a, b] \rightarrow \mathbb{R}$ integrierbar. Seien $\lambda, \mu \in \mathbb{R}$.

$$\text{Es ist } \int_a^b \lambda u(x) + \mu v(x) dx = \lambda \int_a^b u(x) dx + \mu \int_a^b v(x) dx.$$

(2) Seien $a, b, c \in \mathbb{R}$ mit $a \leq b \leq c$. Sei $w : [a, c] \rightarrow \mathbb{R}$ integrierbar.

$$\text{Es ist } \int_a^c w(x) dx = \int_a^b w(x) dx + \int_b^c w(x) dx.$$

(3) Sei f integrierbar. Es ist $\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$.

Dies ist eine Variante der Dreiecksungleichung.

(4) Sei f integrierbar mit $f(x) \geq 0$ für $x \in [a, b]$. Dann ist auch $\int_a^b f(x) dx \geq 0$.

Definition. Sei $\int_b^a f(x) dx := - \int_a^b f(x) dx$.

4.7.3 Hauptsatz

Seien $a, b \in \mathbb{R}$ mit $a \leq b$ gegeben. Sei $D \subseteq \mathbb{R}$ offen mit $[a, b] \subseteq D$.

Sei $f : D \rightarrow \mathbb{R}$ stetig. Sei $g : D \rightarrow \mathbb{R}$ stetig differenzierbar.

Definition. Ist $g'(x) = f(x)$ für $x \in]a, b[$, so heißt g eine *Stammfunktion* von f auf $[a, b]$.

Oft nennt man eine Stammfunktion auch *Aufleitung*.

Gilt dies für jedes solche Intervall in D , so heißt g auch Stammfunktion von f auf D .

Bemerkung. Seien $g, \tilde{g} : D \rightarrow \mathbb{R}$ Stammfunktionen von f auf $[a, b]$.

Dann gibt es ein $c \in \mathbb{R}$ mit $g(x) + c = \tilde{g}(x)$ für $x \in [a, b]$.

Unterschied:

Die Aufleitung einer Funktion ist nur bis auf eine Konstante eindeutig bestimmt.

Die Ableitung einer Funktion ist eindeutig bestimmt.

Satz (Hauptsatz).

(1) Sei g eine Stammfunktion von f auf $[a, b]$. Dann ist $\int_a^b f(x) \, dx = g(b) - g(a)$.

Wir schreiben auch $[g(x)]_{x=a}^{x=b} = [g(x)]_a^b := g(b) - g(a)$, um einen Zwischenschritt einbauen zu können:

$$\int_a^b f(x) \, dx = [g(x)]_a^b = g(b) - g(a) .$$

Dies gilt auch noch, falls $a > b$ ist.

(2) Es ist $h : D \rightarrow \mathbb{R} : x \mapsto \int_a^x f(t) \, dt$ eine Stammfunktion von f auf D , falls D ein Intervall ist.

Auf diese theoretische Weise wie in (2) können wir also für jede stetige Funktion eine Stammfunktion angeben.

Das Auffinden einer Stammfunktion in der Praxis ist dennoch ein Problem.

Definition. Die Aussage, es sei g eine Stammfunktion von f , kann auch wie folgt geschrieben werden.

$$\int f(x) \, dx = [g(x)]$$

Man nennt dabei $\int f(x) \, dx$ auch das *unbestimmte Integral* von f . Die eckigen Klammern sollen daran erinnern, daß eine Stammfunktion nur bis auf Konstante eindeutig festliegt: $[g(x)] = [g(x) + c]$ für jede Konstante $c \in \mathbb{R}$.

Wir können mit solchen Ausdrücken auch rechnen: Seien $\lambda, \tilde{\lambda} \in \mathbb{R}$.

Für stetig differenzierbare Funktionen g und \tilde{g} sei

$$\lambda[g(x)] + \tilde{\lambda}[\tilde{g}(x)] := [\lambda g(x) + \tilde{\lambda} \tilde{g}(x)] .$$

Dann ist für stetige Funktionen f und \tilde{f} auch

$$\lambda \int f(x) \, dx + \tilde{\lambda} \int \tilde{f}(x) \, dx = \int \lambda f(x) + \tilde{\lambda} \tilde{f}(x) \, dx .$$

Dies braucht man, wenn man eine Funktion, die man integrieren will, zu diesem Zweck zerlegt und die einzelnen Teile integriert.

Beispiel.

Es ist $\int x^2 \, dx = [\frac{1}{3}x^3]$, da $x^2 = \frac{d}{dx} \frac{1}{3}x^3$ ist.

So ist $\int_0^1 x^2 \, dx = [\frac{1}{3}x^3]_0^1 = \frac{1}{3}1^3 - \frac{1}{3}0^3 = \frac{1}{3}$.

Beispiel.

(1) Sei $a \in \mathbb{R} \setminus \{-1\}$. Es ist

$$\int x^a \, dx = \left[\frac{x^{a+1}}{a+1} \right],$$

da $x^a = \frac{d}{dx} \frac{x^{a+1}}{a+1}$ ist.

Dies gilt auf \mathbb{R} , falls $a \in \mathbb{Z} \setminus \{-1\}$ liegt, oder falls $a \in \mathbb{Q} \setminus \{-1\}$ liegt und gekürzt ungeraden Nenner hat. Ansonsten gilt es auf $\mathbb{R}_{>0}$.

(2) Auf $\mathbb{R}_{>0}$ ist

$$\int x^{-1} \, dx = [\ln(x)],$$

da $x^{-1} = \frac{d}{dx} \ln(x)$ ist.

Benötigt man die Stammfunktion auf $\mathbb{R} \setminus \{0\}$, so ist $\int x^{-1} \, dx = [\ln(|x|)]$, da für $x < 0$ sich zudem $\frac{d}{dx} \ln(|x|) = \frac{d}{dx} \ln(-x) = \frac{1}{-x} \cdot (-1) = \frac{1}{x}$ ergibt.

(3) Es ist

$$\int e^x \, dx = [e^x],$$

da $e^x = \frac{d}{dx} e^x$ ist.

(4) Es ist

$$\begin{aligned} \int \sin(x) \, dx &= [-\cos(x)] \\ \int \cos(x) \, dx &= [\sin(x)] \\ \int \sinh(x) \, dx &= [\cosh(x)] \\ \int \cosh(x) \, dx &= [\sinh(x)], \end{aligned}$$

weil die jeweilige Ableitung der Funktion auf der rechten Seite den Integranden auf der linken Seite liefert.

(5) Es ist

$$\int \frac{1}{1+x^2} \, dx = [\arctan(x)],$$

da $\frac{1}{1+x^2} = \frac{d}{dx} \arctan(x)$ ist. Somit wird z.B.

$$\int_{-1}^{+1} \frac{1}{1+x^2} \, dx = [\arctan(x)]_{-1}^{+1} = \arctan(+1) - \arctan(-1) = \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{2}.$$

Denn $\tan(\pi/4) = +1$ und $\tan(-\pi/4) = -1$.

Beispiel. Wir setzen das vorletzte Beispiel aus §4.7.1 fort.

Dort hatten wir für $n \in \mathbb{Z}_{\geq 1}$ folgende Ungleichungen festgestellt.

$$\ln(n!) \leq \int_1^{n+1} \ln(x) \, dx \leq \ln((n+1)!).$$

Wir versuchen nun, durch Probieren eine Stammfunktion von $\ln(x)$ zu finden.

Erster Versuch: Es ist $(x \ln(x))' = 1 \cdot \ln(x) + x \cdot x^{-1} = \ln(x) + 1$.

Den störenden konstanten Summanden 1 können wir korrigieren.

Zweiter Versuch: Es ist $(x \ln(x) - x)' = \ln(x) + 1 - 1 = \ln(x)$.

Also ist $\int \ln(x) \, dx = [x \ln(x) - x]$.

Es folgt

$$\begin{aligned} \int_1^{n+1} \ln(x) \, dx &= [x \ln(x) - x]_1^{n+1} \\ &= ((n+1) \ln(n+1) - (n+1)) - (1 \cdot \ln(1) - 1) \\ &= (n+1) \ln(n+1) - n. \end{aligned}$$

Somit wird

$$\ln(n!) \leq (n+1) \ln(n+1) - n \leq \ln((n+1)!).$$

Es ist $(\exp(x))' = \exp(x) > 0$ stets. Anwenden von \exp liefert also die Ungleichungen

$$n! = \exp(\ln(n!)) \leq \exp((n+1) \ln(n+1) - n) \leq \exp(\ln((n+1)!)) = (n+1)!.$$

Die linke Ungleichung lautet $n! \leq \exp((n+1) \ln(n+1) - n) = (n+1)^{(n+1)} \cdot e^{-n}$.

In der rechten Ungleichung ersetzen wir noch $n+1$ durch n und erhalten auf diese Weise $n! \geq \exp(n \ln(n) - (n-1)) = n^n \cdot e^{-(n-1)}$.

Zusammen haben wir folgende Ungleichungen für $n \geq 2$ erhalten.

$$\frac{n^n}{e^{n-1}} \leq n! \leq \frac{(n+1)^{(n+1)}}{e^n}.$$

Diese gelten auch noch für $n = 1$, wie man durch Einsetzen bestätigt, aber nicht mehr für $n = 0$.

Das kann man noch verfeinern, Stichwort Stirlingsche Formel.

Beispiel. Gemäß vorstehendem Beispiel wird $n!^{1/n} \geq \left(\frac{n^n}{e^{n-1}}\right)^{1/n} = \frac{n}{e^{1-\frac{1}{n}}}$ für $n \geq 2$.

Es ist $\lim_{n \rightarrow \infty} \frac{n}{e^{1-\frac{1}{n}}} = \frac{+\infty}{e} = +\infty$. Also ist auch

$$\lim_{n \rightarrow \infty} n!^{1/n} = +\infty.$$

Vgl. viertes Beispiel in §4.3.

4.7.4 Integrationstechniken

4.7.4.1 Partielle Integration

Sei $D \subseteq \mathbb{R}$ offen.

Seien $f, g : D \rightarrow \mathbb{R}$ differenzierbare Funktionen.

Lemma (Partielle Integration). *Es ist*

$$\int f'(x) \cdot g(x) \, dx = [f(x) \cdot g(x)] - \int f(x) \cdot g'(x) \, dx .$$

Denn dies ist gleichbedeutend mit $\int f'(x) \cdot g(x) + f(x) \cdot g'(x) \, dx \stackrel{!}{=} [f(x) \cdot g(x)]$, was aber wegen $f'(x) \cdot g(x) + f(x) \cdot g'(x) = \frac{d}{dx}(f(x) \cdot g(x))$ dank Produktregel zutrifft.

Beispiel. Es ist $\int \sin(x) \cdot x \, dx = [(-\cos(x)) \cdot x] - \int (-\cos(x)) \cdot 1 \, dx = [-\cos(x) \cdot x + \sin(x)]$.

Beispiel. Es ist $\int e^x \cdot x^2 \, dx = [e^x \cdot x^2] - \int e^x \cdot 2x \, dx = [e^x \cdot x^2] - ([e^x \cdot 2x] - \int e^x \cdot 2 \, dx) = [e^x \cdot (x^2 - 2x + 2)]$.

Beispiel. Es ist $\int \ln(x) \, dx = \int 1 \cdot \ln(x) \, dx = [x \cdot \ln(x)] - \int x \cdot x^{-1} \, dx = [x \cdot \ln(x)] - \int 1 \, dx = [x \cdot \ln(x) - x]$.

Beispiel. Es ist $\int e^x \cos(x) \, dx = [e^x \cos(x)] - \int e^x (-\sin(x)) \, dx = [e^x \cos(x)] + \int e^x \sin(x) \, dx = [e^x \cos(x)] + [e^x \sin(x)] - \int e^x \cos(x) \, dx$.

Nun scheint es, als hätten wir im Kreis gerechnet. Nicht ganz, wir können nämlich das entstandene Integral noch nach links bringen: $2 \int e^x \cos(x) \, dx = [e^x \cos(x)] + [e^x \sin(x)]$. Also ergibt sich $\int e^x \cos(x) \, dx = [\frac{1}{2} e^x (\cos(x) + \sin(x))]$.

4.7.4.2 Komplexwertige Integranden

Für die nächsten beiden Techniken in §4.7.4.3 und §4.7.4.4 hilft es, auch Funktionen $f : D \rightarrow \mathbb{C}$ als Integranden zuzulassen, wobei aber nach wie vor $D \subseteq \mathbb{R}$ sei.

Für $a, b \in \mathbb{R}$ mit $a \leq b$ und $[a, b] \subseteq D$ sei dabei

$$\int_a^b f(x) \, dx := \int_a^b \operatorname{Re}(f(x)) \, dx + i \int_a^b \operatorname{Im}(f(x)) \, dx .$$

Die Eigenschaften aus §4.7.2 und der Hauptsatz aus §4.7.3 gelten dann entsprechend auch in diesem Fall. Auch die partielle Integration aus §4.7.4.1 ist anwendbar.

Beispiel. Es ist $\int 2 \cos(x) \, dx = \int e^{ix} + e^{-ix} \, dx = \int e^{ix} \, dx + \int e^{-ix} \, dx$.

Nun ist $\frac{d}{dx} \frac{1}{i} e^{ix} = e^{ix}$ und $\frac{d}{dx} \frac{1}{-i} e^{-ix} = e^{-ix}$ nach Kettenregel.

Dank Hauptsatz ist also $\int e^{ix} dx + \int e^{-ix} dx = [\frac{1}{i} e^{ix}] + [\frac{1}{-i} e^{-ix}] = [-i(e^{ix} - e^{-ix})] = [-i \cdot 2i \sin(x)] = [2 \sin(x)]$.

Insgesamt ist $\int 2 \cos(x) dx = [2 \sin(x)]$, wie wir ja auch schon wußten.

4.7.4.3 Substitution

Sei $D \subseteq \mathbb{R}$ offen. Sei $E \subseteq \mathbb{R}$ offen.

Seien $u : D \rightarrow E \subseteq \mathbb{R}$ differenzierbar. Sei $f : E \rightarrow \mathbb{C}$ differenzierbar.

Lemma (Substitution.) *Es ist*

$$\int f'(u(x)) \cdot u'(x) dx = [f(u(x))].$$

Denn nach Kettenregel ist $\frac{d}{dx} f(u(x)) = f'(u(x)) \cdot u'(x)$.

Bemerkung. Merkt man sich, daß u als $u(x)$ zu lesen ist, dann wird

$$\int f'(u(x)) \cdot u'(x) dx = [f(u(x))] = [f(u)] = \int f'(u) du.$$

Als Merkregel kann man noch $u'(x) = \frac{du}{dx}$ schreiben und erhält so

$$\int f'(u(x)) \cdot \frac{du}{dx} dx = \int f'(u) du.$$

Das ist dann eine Art Kürzungsregel.

Bemerkung. *Seien $a, b \in D$ mit $[a, b] \subseteq D$ gegeben. Es wird*

$$\int_a^b f'(u(x)) \cdot u'(x) dx = \int_{u(a)}^{u(b)} f'(u) du,$$

wenn man auf der rechten Seite u als Variable liest.

Die Integrationsgrenzen sollten also ebenfalls substituiert werden.

Denn

$$\int_a^b f'(u(x)) \cdot u'(x) dx = [f(u(x))]_a^b = f(u(b)) - f(u(a)) = [f(u)]_{u(a)}^{u(b)} = \int_{u(a)}^{u(b)} f'(u) du,$$

wobei in den letzten beiden Ausdrücken u als Variable gelesen wurde.

Beispiel. Wir wollen $\int x \cos(x^2) dx$ berechnen. Wir verwenden $u(x) = x^2$, $u'(x) = 2x$, $f(t) = \sin(t)$ und $f'(t) = \cos(t)$. Dann ist $f'(u(x)) \cdot u'(x) = \cos(x^2) \cdot 2x$. Damit wird

$$\int x \cos(x^2) dx = \frac{1}{2} \int \cos(x^2) \cdot 2x dx = \frac{1}{2} [\sin(x^2)].$$

Wenn man die Integrationsgrenzen mitsubstituiert, dann wird

$$\begin{aligned} \int_0^{\sqrt{\pi/2}} x \cos(x^2) dx &= \frac{1}{2} \int_0^{\sqrt{\pi/2}} \cos(x^2) \cdot 2x dx \\ &= \frac{1}{2} \int_{u(0)}^{u(\sqrt{\pi/2})} \cos(u) du \\ &= \frac{1}{2} \int_0^{\pi/2} \cos(u) du \\ &= \frac{1}{2} [\sin(u)]_0^{\pi/2} \\ &= \frac{1}{2} (\sin(\pi/2) - \sin(0)) = \frac{1}{2}. \end{aligned}$$

Alternativ, wenn man über die Stammfunktion argumentiert, dann wird genauso

$$\begin{aligned} \int_0^{\sqrt{\pi/2}} x \cos(x^2) dx &= \frac{1}{2} \int_0^{\sqrt{\pi/2}} \cos(x^2) \cdot 2x dx \\ &= \frac{1}{2} [\sin(x^2)]_0^{\sqrt{\pi/2}} \\ &= \frac{1}{2} (\sin(\pi/2) - \sin(0)) = \frac{1}{2}. \end{aligned}$$

Beispiel. Sei $a \in \mathbb{C} \setminus \{0\}$. Sei $b \in \mathbb{C}$. Es wird

$$\int f'(ax+b) dx = \frac{1}{a} \int f'(ax+b)(ax+b)' dx = \frac{1}{a} [f(ax+b)] = \left[\frac{1}{a} f(ax+b) \right].$$

Dies bezeichnet man auch als *lineare Substitution*.

So z.B. wird

$$\int \cos(7x) dx = \left[\frac{1}{7} \sin(7x) \right].$$

Beispiel. Wir wollen nochmals $\int e^x \cos(x) dx$ berechnen. Diesesmal ohne Trick, dafür aber mit de Moivre und Substitution.

Es wird $e^x \cos(x) = e^x \cdot \frac{1}{2}(e^{ix} + e^{-ix}) = \frac{1}{2}(e^{x(1+i)} + e^{x(1-i)})$. Lineare Substitution gibt

$$\begin{aligned} \int e^x \cos(x) dx &= \frac{1}{2} (\int e^{x(1+i)} dx + \int e^{x(1-i)} dx) \\ &= \frac{1}{2} \left(\left[\frac{1}{1+i} e^{x(1+i)} \right] + \left[\frac{1}{1-i} e^{x(1-i)} \right] \right) \\ &= \frac{1}{4} \left([(1-i) e^{x(1+i)}] + [(1+i) e^{x(1-i)}] \right) \\ &= \frac{1}{4} [e^x (e^{ix} + e^{-ix}) - i e^x (e^{ix} - e^{-ix})] \\ &= \frac{1}{4} [e^x \cdot 2 \cos(x) - i e^x \cdot 2i \sin(x)] \\ &= \left[\frac{1}{2} e^x (\cos(x) + \sin(x)) \right]. \end{aligned}$$

Vgl. das letzte Beispiel in §4.7.4.1.

Beispiel. Wir wollen $\int \cos(x)^4 dx$ berechnen.

Mit de Moivre und dem Binomischen Lehrsatz ist

$$\begin{aligned}
 \cos(x)^4 &= \left(\frac{1}{2}(e^{ix} + e^{-ix})\right)^4 \\
 &= \frac{1}{16}((e^{ix})^4 + 4(e^{ix})^3 \cdot e^{-ix} + 6(e^{ix})^2 \cdot (e^{-ix})^2 + 4e^{ix} \cdot (e^{-ix})^3 + (e^{-ix})^4) \\
 &= \frac{1}{16}(e^{4ix} + 4e^{2ix} + 6 + 4e^{-2ix} + e^{-4ix}) \\
 &= \frac{1}{16}((e^{4ix} + e^{-4ix}) + 4(e^{2ix} + e^{-2ix}) + 6) \\
 &= \frac{1}{16}(2\cos(4x) + 4 \cdot 2\cos(2x) + 6) \\
 &= \frac{1}{8}\cos(4x) + \frac{1}{2}\cos(2x) + \frac{3}{8}.
 \end{aligned}$$

Also wird mit linearer Substitution

$$\int \cos(x)^4 dx = \int \frac{1}{8}\cos(4x) + \frac{1}{2}\cos(2x) + \frac{3}{8} dx = \left[\frac{1}{32}\sin(4x) + \frac{1}{4}\sin(2x) + \frac{3}{8}x \right].$$

Beispiel. Partielle Integration und Substitution geben

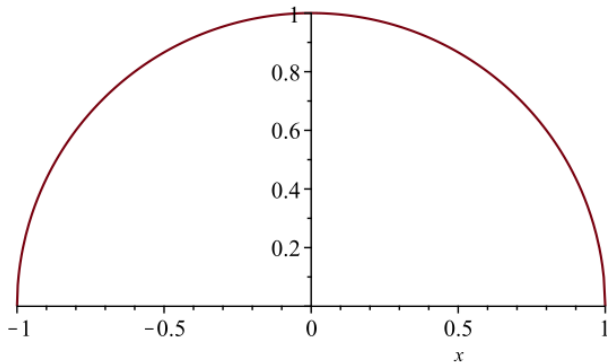
$$\begin{aligned}
 \int \arctan(x) dx &= \int 1 \cdot \arctan(x) dx \\
 &= [x \cdot \arctan(x)] - \int x \cdot \frac{1}{1+x^2} dx \\
 &= [x \cdot \arctan(x)] - \frac{1}{2} \int \frac{1}{1+x^2} \cdot 2x dx \\
 &= [x \cdot \arctan(x)] - \frac{1}{2}[\ln(1+x^2)] \\
 &= [x \cdot \arctan(x) - \frac{1}{2}\ln(1+x^2)].
 \end{aligned}$$

Probe: Es ist $\frac{d}{dx}(x \cdot \arctan(x) - \frac{1}{2}\ln(1+x^2)) = 1 \cdot \arctan(x) + x \cdot \frac{1}{1+x^2} - \frac{1}{2} \frac{1}{1+x^2} \cdot 2x = \arctan(x)$.

Manchmal hilft Substitution auch rückwärts:

Beispiel. Wir wollen den Flächeninhalt A eines Halbkreises mit Radius 1 bestimmen.

Hierzu betrachten wir den Graphen der Funktion $h: [-1, +1] \rightarrow \mathbb{R}: x \mapsto \sqrt{1-x^2}$.



Dies ist ein Halbkreis. Denn der Abstand des Ursprungs zum Punkt $(x, h(x))$ auf dem Graphen ist $\sqrt{x^2 + h(x)^2} = \sqrt{x^2 + (\sqrt{1-x^2})^2} = 1$.

Wir haben also $A = \int_{-1}^{+1} \sqrt{1-u^2} \, du$ zu berechnen.

Nun verwenden wir Substitution rückwärts.

Sei $f(u)$ eine Stammfunktion von $h(u)$, existent nach Hauptsatz. Somit wissen wir $f'(u) = h(u) = \sqrt{1-u^2}$, aber mehr auch nicht.

Wir setzen $u(x) := \sin(x)$. Läuft $x \in [-\pi/2, +\pi/2]$, dann läuft $u(x) = \sin(x) \in [-1, +1]$.

Laut dritter Bemerkung, rückwärts gelesen, wird

$$A = \int_{-1}^{+1} \sqrt{1-u^2} \, du = \int_{u(-\pi/2)}^{u(+\pi/2)} f'(u) \, du = \int_{-\pi/2}^{+\pi/2} f'(u(x)) \cdot u'(x) \, dx.$$

Und das können wir ausrechnen.

Es ist $f'(u(x)) \cdot u'(x) = \sqrt{1-u(x)^2} \cdot u'(x) = \sqrt{1-\sin(x)^2} \cdot \cos(x) = \sqrt{\cos(x)^2} \cdot \cos(x) = \cos(x) \cdot \cos(x) = \cos(x)^2$ für $x \in [-\pi/2, +\pi/2]$, wobei für die vorletzte Gleichheit zu beachten ist, daß dort $\cos(x) \geq 0$ gilt.

Ferner ist $\cos(x)^2 = (\frac{1}{2}(e^{ix} + e^{-ix}))^2 = \frac{1}{4}(e^{2ix} + 2 + e^{-2ix}) = \frac{1}{2} + \frac{1}{2} \cos(2x)$.

Also können wir fortsetzen

$$A = \int_{-\pi/2}^{+\pi/2} f'(u(x)) \cdot u'(x) \, dx = \int_{-\pi/2}^{+\pi/2} \left(\frac{1}{2} + \frac{1}{2} \cos(2x) \right) dx = \left[\frac{x}{2} + \frac{1}{4} \sin(2x) \right]_{-\pi/2}^{+\pi/2} = \frac{\pi}{2}.$$

In der Praxis formuliert dies oft kurz wie folgt. Mit der Substitution $u(x) = \sin(x)$ wird

$$A = \int_{u=-1}^{u=+1} \sqrt{1-u^2} \, du = \int_{x=-\pi/2}^{x=+\pi/2} \sqrt{1-u(x)^2} \cdot \frac{du}{dx} \, dx = \int_{-\pi/2}^{+\pi/2} \cos(x) \cdot \cos(x) \, dx = \dots = \frac{\pi}{2}.$$

4.7.4.4 Partialbruchzerlegung

Sei $m \geq 1$. Seien $s_1, \dots, s_m \in \mathbb{C}$ paarweise verschieden. Seien $e_1, \dots, e_m \in \mathbb{Z}_{\geq 1}$.

Wir bilden das Polynom

$$g(x) = (x - s_1)^{e_1} \cdot (x - s_2)^{e_2} \cdot \dots \cdot (x - s_m)^{e_m}.$$

Sei $f(x)$ ein Polynom mit Koeffizienten in \mathbb{C} und mit $\deg(f(x)) < \deg(g(x))$. Wir wollen

$$\int \frac{f(x)}{g(x)} \, dx$$

berechnen.

Man beachte noch, daß für jeden Bruch $\frac{h(x)}{g(x)}$ von Polynomen der Zähler mittels Polynomdivision $h(x) = q(x) \cdot g(x) + f(x)$ geschrieben werden kann, mit Polynomen $q(x)$ und $f(x)$ mit $\deg(f(x)) < \deg(g(x))$, was $\frac{h(x)}{g(x)} = q(x) + \frac{f(x)}{g(x)}$ liefert. Wenn wir Brüche von Polynomen integrieren wollen, dürfen wir dank Polynomdivision also annehmen, daß der Zähler kleineren Grad als der Nenner hat.

Wir wollen nun $\frac{f(x)}{g(x)}$ in eine Summe von Brüchen zerlegen, die wir integrieren können.

Wir machen folgenden Ansatz. Gesucht seien Konstanten $a_{\ell,j} \in \mathbb{C}$ mit

$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{f(x)}{(x-s_1)^{e_1} \cdot (x-s_2)^{e_2} \cdot \dots \cdot (x-s_m)^{e_m}} \\ &\stackrel{!}{=} \sum_{\ell=1}^m \sum_{j=1}^{e_\ell} \frac{a_{\ell,j}}{(x-s_\ell)^j} \\ &= \frac{a_{1,1}}{(x-s_1)^1} + \frac{a_{1,2}}{(x-s_1)^2} + \dots + \frac{a_{1,e_1}}{(x-s_1)^{e_1}} \\ &\quad + \frac{a_{2,1}}{(x-s_2)^1} + \frac{a_{2,2}}{(x-s_2)^2} + \dots + \frac{a_{2,e_2}}{(x-s_2)^{e_2}} \\ &\quad + \dots \\ &\quad + \frac{a_{m,1}}{(x-s_m)^1} + \frac{a_{m,2}}{(x-s_m)^2} + \dots + \frac{a_{m,e_m}}{(x-s_m)^{e_m}} \end{aligned}$$

für $x \in \mathbb{C}$ mit $g(x) \neq 0$. Eine solche Zerlegung heißt dann *Partialbruchzerlegung*.

Dieser Ansatz ist so konstruiert, daß es für die Konstanten $a_{\ell,j}$ genau eine Lösung geben muß.

Um die Konstanten $a_{\ell,j}$ zu bestimmen, multipliziere man diesen Ansatz auf beiden Seiten mit $g(x)$.

Wir schreiben $d := \deg(g(x))$. Die Anzahl der zu bestimmenden unbekannt Konstanten $a_{\ell,j}$ ist gerade $\sum_{\ell=1}^m e_\ell = d$.

Man erhält auf der linken Seite das Polynom $f(x)$.

Man erhält auf der rechten Seite ein Polynom in der Variablen x von Grad $d-1$, in welchem noch die unbekannt Konstanten $a_{\ell,j}$ auftreten.

Koeffizientenvergleich bei x^0, x^1, \dots, x^{d-1} liefert nun d Gleichungen in den d unbekannt Konstanten $a_{\ell,j}$.

Man löse dieses inhomogene lineare Gleichungssystem.

Dann bleiben Terme der Form $\frac{a_{\ell,j}}{(x-s_\ell)^j}$ zu integrieren.

Bemerkung. Falls $j \geq 2$ ist, dann ist $\int \frac{1}{(x-s_\ell)^j} dx = \left[\frac{1}{(1-j)(x-s_\ell)^{j-1}} \right]$.

Bemerkung. Falls $j = 1$ ist, dann ist $\int \frac{1}{x-s_\ell} dx = [\ln(|x-s_\ell|)]$, falls s_ℓ reell ist.

Bemerkung. Seien $f(x)$ und $g(x)$ Polynome mit reellen Koeffizienten.

Tritt ein s_{ℓ_1} auf, das nicht reell ist, dann tritt auch ein s_{ℓ_2} auf mit $s_{\ell_2} = \bar{s}_{\ell_1}$, wobei $\ell_1, \ell_2 \in \{1, \dots, m\}$.

Kurz, die echt komplexen Nullstellen des Nenners $g(x)$ treten in komplex konjugierten Paaren auf.

Dann kann man wie folgt zusammenfassen.

Sei $s = a + bi$, mit $a, b \in \mathbb{R}$ und $b \neq 0$.

$$(1) \text{ Es ist } \int \frac{1}{x-s} + \frac{1}{x-\bar{s}} dx = \int \frac{1}{x-(a+bi)} + \frac{1}{x-(a-bi)} dx = [\ln((x-a)^2 + b^2)].$$

$$(2) \text{ Es ist } \int \frac{i}{x-s} - \frac{i}{x-\bar{s}} dx = \int \frac{i}{x-(a+bi)} - \frac{i}{x-(a-bi)} dx = [-2 \arctan(\frac{x-a}{b})].$$

Denn in (1) ist der Integrand $\frac{1}{x-(a+bi)} + \frac{1}{x-(a-bi)} = \frac{x-(a-bi)+x-(a+bi)}{(x-(a+bi))(x-(a-bi))} = \frac{2(x-a)}{(x-a)^2+b^2}$, was auch die Ableitung der rechten Seite ist.

Und in (2) ist der Integrand $\frac{i}{x-(a+bi)} - \frac{i}{x-(a-bi)} = \frac{i(x-(a-bi)-x+(a+bi))}{(x-(a+bi))(x-(a-bi))} = \frac{-2b}{(x-a)^2+b^2}$. Die Ableitung der rechten Seite ist $\frac{d}{dx}(-2 \arctan(\frac{x-a}{b})) = \frac{-2}{1+(\frac{x-a}{b})^2} \cdot \frac{1}{b} = \frac{-2b}{b^2+(x-a)^2}$, was dasselbe ist.

Bemerkung. In der Praxis schreibt man im Ansatz oft statt $a_{1,1}, a_{1,2}, a_{2,1}, \dots$ der Einfachheit halber A, B, C, \dots , um die Doppelindizes nicht schreiben zu müssen.

In der oben ausgeführten Theorie geht diese vereinfachte Schreibung noch nicht, da man die Summen in der dortigen Allgemeinheit zum Ausdruck bringen muß.

Beispiel. Wir wollen $\int \frac{1}{x^2-1} dx$ berechnen.

Wir faktorisieren $x^2 - 1 = (x-1)^1(x+1)^1$.

Wir setzen

$$\frac{1}{(x-1)^1(x+1)^1} \stackrel{!}{=} \frac{A}{(x-1)^1} + \frac{B}{(x+1)^1}$$

an. Wir haben A und B zu bestimmen. Multiplikation mit $(x-1)(x+1)$ gibt

$$1 \stackrel{!}{=} A(x+1) + B(x-1) = (A+B)x + (A-B).$$

Koeffizientenvergleich führt auf folgendes lineare Gleichungssystem.

$$\begin{aligned} A+B &= 0 && \text{(Koeffizient bei } x^1) \\ A-B &= 1 && \text{(Koeffizient bei } x^0) \end{aligned}$$

Dies lösen wir wie folgt.

$$\left(\begin{array}{c|c} 1 & -1 \\ \hline 1 & 1 \end{array} \middle| \begin{array}{c} 0 \\ 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{c|c} 1 & -1 \\ \hline 0 & -2 \end{array} \middle| \begin{array}{c} 0 \\ 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array} \middle| \begin{array}{c} 1/2 \\ -1/2 \end{array} \right).$$

Also $A = 1/2$ und $B = -1/2$. Wir erhalten

$$\int \frac{1}{x^2-1} dx = \int \frac{1/2}{x-1} + \frac{-1/2}{x+1} dx = \left[\frac{1}{2} \ln(|x-1|) - \frac{1}{2} \ln(|x+1|) \right].$$

Beispiel. Wir wollen $\int \frac{2x^2+6x-5}{(x-1)^2(x+2)} dx$ berechnen.

Wir setzen

$$\frac{2x^2 + 6x - 5}{(x-1)^2(x+2)^1} \stackrel{!}{=} \frac{A}{(x-1)^1} + \frac{B}{(x-1)^2} + \frac{C}{(x+2)^1}$$

an. Wir haben A , B und C zu bestimmen. Multiplikation mit $(x-1)^2(x+2)$ gibt

$$\begin{aligned} 2x^2 + 6x - 5 &\stackrel{!}{=} A(x-1)(x+2) + B(x+2) + C(x-1)^2 \\ &= (A+C)x^2 + (A+B-2C)x + (-2A+2B+C). \end{aligned}$$

Koeffizientenvergleich führt auf folgendes lineare Gleichungssystem.

$$\begin{aligned} A + C &= 2 && \text{(Koeffizient bei } x^2) \\ A + B - 2C &= 6 && \text{(Koeffizient bei } x^1) \\ -2A + 2B + C &= -5 && \text{(Koeffizient bei } x^0) \end{aligned}$$

Dies lösen wir wie folgt.

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 1 & 1 & -2 & 6 \\ -2 & 2 & 1 & -5 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -3 & 4 \\ 0 & 2 & 3 & -1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -3 & 4 \\ 0 & 0 & 9 & -9 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array} \right).$$

Also $A = 3$, $B = 1$ und $C = -1$. Wir erhalten

$$\int \frac{2x^2+6x-5}{(x-1)^2(x+2)} dx = \int \frac{3}{(x-1)^1} + \frac{1}{(x-1)^2} + \frac{-1}{(x+2)^1} dx = [3 \ln(|x-1|) - \frac{1}{x-1} - \ln(|x+2|)].$$

Beispiel. Wir wollen $\int \frac{x^3+x^2+2x+3}{(x^2+2x+2)^2} dx$ berechnen.

Es ist $x^2 + 2x + 2 = (x+1-i)(x+1+i)$. Also ist $(x^2 + 2x + 2)^2 = (x+1-i)^2(x+1+i)^2$.

Wir setzen

$$\frac{x^3 + x^2 + 2x + 3}{(x+1-i)^2(x+1+i)^2} \stackrel{!}{=} \frac{A}{(x+1-i)^1} + \frac{B}{(x+1-i)^2} + \frac{C}{(x+1+i)^1} + \frac{D}{(x+1+i)^2}$$

an. Wir haben A , B , C und D zu bestimmen. Multiplikation mit $(x+1-i)^2(x+1+i)^2$ gibt

$$\begin{aligned} x^3 + x^2 + 2x + 3 &\stackrel{!}{=} A(x+1-i)(x+1+i)^2 + B(x+1+i)^2 \\ &\quad + C(x+1-i)^2(x+1+i) + D(x+1-i)^2 \\ &= A(x^3 + (3+i)x^2 + (4+2i)x + (2+2i)) + B(x^2 + (2+2i)x + 2i) \\ &\quad + C(x^3 + (3-i)x^2 + (4-2i)x + (2-2i)) + D(x^2 + (2-2i)x - 2i) \\ &= (A+C)x^3 \\ &\quad + ((3+i)A + B + (3-i)C + D)x^2 \\ &\quad + ((4+2i)A + (2+2i)B + (4-2i)C + (2-2i)D)x \\ &\quad + ((2+2i)A + 2iB + (2-2i)C - 2iD). \end{aligned}$$

Koeffizientenvergleich führt auf folgendes lineare Gleichungssystem.

$$\begin{aligned} A + C &= 1 && \text{(Koeffizient bei } x^3) \\ (3 + i)A + B + (3 - i)C + D &= 1 && \text{(Koeffizient bei } x^2) \\ (4 + 2i)A + (2 + 2i)B + (4 - 2i)C + (2 - 2i)D &= 2 && \text{(Koeffizient bei } x^1) \\ (2 + 2i)A + 2iB + (2 - 2i)C - 2iD &= 3 && \text{(Koeffizient bei } x^0) \end{aligned}$$

Dies lösen wir wie folgt.

$$\begin{aligned} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 3+i & 1 & 3-i & 1 & 1 \\ 4+2i & 2+2i & 4-2i & 2-2i & 2 \\ 2+2i & 2i & 2-2i & -2i & 3 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ i & 1 & -i & 1 & -2 \\ 2i & 2+2i & -2i & 2-2i & -2 \\ 2i & 2i & -2i & -2i & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2i & 1 & -2-i \\ 0 & 2+2i & -4i & 2-2i & -2-2i \\ 0 & 2i & -4i & -2i & 1-2i \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2i & 1 & -2-i \\ 0 & 2i & 0 & -2i & 2 \\ 0 & 2i & -4i & -2i & 1-2i \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & -2i & 2 & -2 \\ 0 & 1 & 0 & -1 & -i \\ 0 & 0 & -4i & 0 & -1-2i \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1/2+i/4 \\ 0 & 1 & 0 & -1 & -i \\ 0 & 0 & 1 & 0 & 1/2-i/4 \\ 0 & 0 & 0 & 2 & -3/2+i \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1/2+i/4 \\ 0 & 1 & 0 & 0 & -3/4-i/2 \\ 0 & 0 & 1 & 0 & 1/2-i/4 \\ 0 & 0 & 0 & 1 & -3/4+i/2 \end{array} \right) \end{aligned}$$

Also $A = \frac{1}{2} + \frac{i}{4}$, $B = -\frac{3}{4} - \frac{i}{2}$, $C = \frac{1}{2} - \frac{i}{4}$ und $D = -\frac{3}{4} + \frac{i}{2}$. Mit der dritten Bemerkung erhalten wir folgendes.

$$\begin{aligned} &\int \frac{x^3+x^2+2x+3}{(x^2+2x+2)^2} dx \\ &= \frac{1}{4} \int \frac{2+i}{x+1-i} + \frac{-3-2i}{(x+1-i)^2} + \frac{2-i}{x+1+i} + \frac{-3+2i}{(x+1+i)^2} dx \\ &= \frac{1}{4} \int \left(\frac{i}{x-(-1+i)} - \frac{i}{x-(-1-i)} \right) + 2 \left(\frac{1}{x-(-1+i)} + \frac{1}{x-(-1-i)} \right) + \frac{-3-2i}{(x+1-i)^2} + \frac{-3+2i}{(x+1+i)^2} dx \\ &= \frac{1}{4} \left[(-2) \arctan\left(\frac{x-(-1)}{1}\right) + 2 \ln((x-(-1))^2 + 1^2) - \frac{-3-2i}{x+1-i} - \frac{-3+2i}{x+1+i} \right] \\ &= \left[-\frac{1}{2} \arctan(x+1) + \frac{1}{2} \ln(x^2+2x+2) - \frac{1}{4} \frac{(-3-2i)(x+1+i) + (-3+2i)(x+1-i)}{x^2+2x+2} \right] \\ &= \left[-\frac{1}{2} \arctan(x+1) + \frac{1}{2} \ln(x^2+2x+2) + \frac{1}{2} \frac{3x+1}{x^2+2x+2} \right] \end{aligned}$$

4.7.5 Uneigentliche Integrale

Definition.

- (1) Seien $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R}$ mit $a < b$ gegeben. Sei $f :]a, b] \rightarrow \mathbb{R}$ stetig. Wir definieren das *uneigentliche Integral*

$$\int_a^b f(x) dx := \lim_{u \rightarrow a} \int_u^b f(x) dx .$$

Hierbei ist der Grenzwert der Funktion $]a, b] \rightarrow \mathbb{R} : u \mapsto \int_u^b f(x) dx$ für $u \rightarrow a$ zu bilden. Dabei ist insbesondere $a < u$.

Dieses uneigentliche Integral kann konvergieren, transvergieren oder divergieren, je nach Verhalten des Grenzwerts aus der Definition.

Es heißt a seine *uneigentliche Integrationsgrenze*.

- (2) Seien $a \in \mathbb{R}$ und $b \in \mathbb{R} \cup \{+\infty\}$ mit $a < b$ gegeben. Sei $f : [a, b[\rightarrow \mathbb{R}$ stetig. Wir definieren das *uneigentliche Integral*

$$\int_a^b f(x) \, dx := \lim_{v \rightarrow b} \int_a^v f(x) \, dx .$$

Hierbei ist der Grenzwert der Funktion $[a, b[\rightarrow \mathbb{R} : v \mapsto \int_a^v f(x) \, dx$ für $v \rightarrow b$ zu bilden. Dabei ist insbesondere $v < b$.

Dieses uneigentliche Integral kann konvergieren, transvergieren oder divergieren, je nach Verhalten des Grenzwerts aus der Definition.

Es heißt b seine *uneigentliche Integrationsgrenze*.

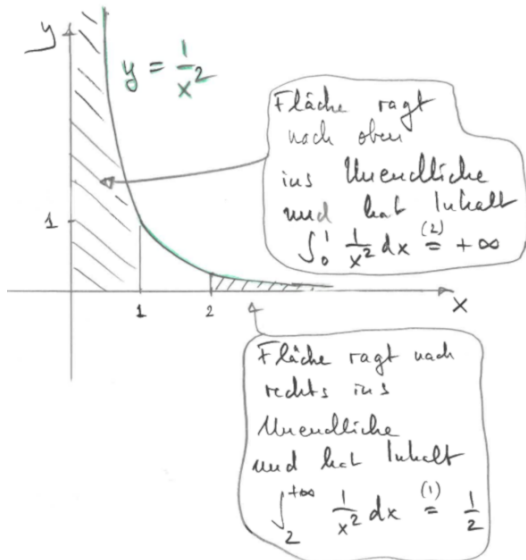
Ein Integral ist also uneigentlich, wenn eine seiner Integrationsgrenzen nicht zum Definitionsbereich des Integranden gehört. Diese Integrationsgrenze heißt dann auch uneigentlich.

Beispiel.

(1) Es wird $\int_2^{+\infty} \frac{1}{x^2} \, dx = \lim_{v \rightarrow +\infty} \int_2^v \frac{1}{x^2} \, dx = \lim_{v \rightarrow +\infty} \left[-\frac{1}{x}\right]_2^v = \lim_{v \rightarrow +\infty} \left(-\frac{1}{v}\right) - \left(-\frac{1}{2}\right) = \frac{1}{2}$.

(2) Es wird $\int_0^1 \frac{1}{x^2} \, dx = \lim_{u \rightarrow 0} \int_u^1 \frac{1}{x^2} \, dx = \lim_{u \rightarrow 0} \left[-\frac{1}{x}\right]_u^1 = \lim_{u \rightarrow 0} \left(-\frac{1}{1}\right) - \left(-\frac{1}{u}\right) = +\infty$.

Hier ist der Integrand $\frac{1}{x^2}$ als auf $]0, 1]$ definiert zu lesen, sodaß sich u nur von der positiven Seite aus der 0 annähern kann.



Lemma (Majorantenkriterium).

- (1) Seien $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R}$ mit $a < b$ gegeben. Sei $f :]a, b] \rightarrow \mathbb{R}$ stetig.

Sei $g :]a, b] \rightarrow \mathbb{R}$ stetig mit $|f(x)| \leq g(x)$ für $x \in]a, b]$. Sei $\int_a^b g(x) \, dx$ konvergent.

Dann ist auch $\int_a^b f(x) \, dx$ konvergent.

(2) Seien $a \in \mathbb{R}$ und $b \in \mathbb{R} \cup \{+\infty\}$ mit $a < b$ gegeben. Sei $f : [a, b[\rightarrow \mathbb{R}$ stetig.

Sei $g : [a, b[\rightarrow \mathbb{R}$ stetig mit $|f(x)| \leq g(x)$ für $x \in [a, b[$. Sei $\int_a^b g(x) dx$ konvergent.

Dann ist auch $\int_a^b f(x) dx$ konvergent.

In beiden Fällen nennt man g eine *konvergente Majorante* von f .

In beiden Fällen ist auch $\int_a^b |f(x)| dx \leq \int_a^b g(x) dx$.

Beispiel. Wir wollen $\int_1^{+\infty} e^{-x^2} dx$ auf Konvergenz untersuchen.

Es ist $x \leq x^2$ für $x \in [1, +\infty[$. Also ist dort auch $|e^{-x^2}| = e^{-x^2} \leq e^{-x}$.

Es ist $\int_1^{+\infty} e^{-x} dx = \lim_{v \rightarrow +\infty} \int_1^v e^{-x} dx = \lim_{v \rightarrow +\infty} [-e^{-x}]_1^v = \lim_{v \rightarrow +\infty} (-e^{-v}) - (-e^{-1}) = e^{-1}$.

Also hat dieses Integral einen Grenzwert in \mathbb{R} , d.h. es konvergiert.

Somit ist e^{-x} auf $[1, +\infty[$ eine konvergente Majorante von e^{-x^2} .

Damit konvergiert $\int_1^{+\infty} e^{-x^2} dx$ nach Majorantenkriterium.

Es ist $\int_1^{+\infty} e^{-x^2} dx \approx 0,1394$. Einen genauen Wert können wir nicht berechnen.

Lemma (Integralkriterium).

Sei $n \in \mathbb{Z}$. Sei $f : [n, +\infty[\rightarrow \mathbb{R}$ eine stetige Funktion mit

$$f(x_1) \geq f(x_2) \geq 0$$

für $n \leq x_1 \leq x_2$.

Dann ist die Reihe $\sum_{j=n}^{\infty} f(j)$ genau dann konvergent, wenn das uneigentliche Integral $\int_n^{+\infty} f(x) dx$ konvergent ist. Diesenfalls gelten folgende Ungleichungen.

$$\int_n^{+\infty} f(x) dx \leq \sum_{j=n}^{\infty} f(j) \leq f(n) + \int_n^{+\infty} f(x) dx$$

Ferner ist genau dann $\sum_{j=n}^{\infty} f(j) = +\infty$, wenn $\int_n^{+\infty} f(x) dx = +\infty$ ist.

Beispiel. Sei $a > 1$. Sei $f : [1, +\infty[\rightarrow \mathbb{R} : x \mapsto f(x) := \frac{1}{x^a}$.

Dann ist $f(x_1) \geq f(x_2) \geq 0$ für $1 \leq x_1 \leq x_2$.

Es ist $\int_1^{+\infty} x^{-a} dx = \lim_{v \rightarrow +\infty} \int_1^v x^{-a} dx = \lim_{v \rightarrow +\infty} [-\frac{x^{1-a}}{a-1}]_1^v = \lim_{v \rightarrow +\infty} (-\frac{v^{1-a}}{a-1}) - (-\frac{1^{1-a}}{a-1}) = \frac{1}{a-1}$.

Also hat dieses Integral einen Grenzwert in \mathbb{R} , d.h. es konvergiert.

Somit konvergiert auch $\sum_{j=1}^{\infty} \frac{1}{j^a}$.

Zudem ist

$$\frac{1}{a-1} = \int_1^{+\infty} x^{-a} dx \leq \sum_{j=1}^{\infty} \frac{1}{j^a} \leq 1^{-a} + \int_1^{+\infty} x^{-a} dx = 1 + \frac{1}{a-1}$$

Beispiel. Sei $f : [2, +\infty[\rightarrow \mathbb{R} : x \mapsto f(x) := \frac{1}{x \ln(x)}$.

Dann ist $f(x_1) \geq f(x_2) \geq 0$ für $1 \leq x_1 \leq x_2$.

Es ist

$$\begin{aligned}
 \int_2^{+\infty} \frac{1}{x \ln(x)} \, dx &= \lim_{v \rightarrow +\infty} \int_2^v \frac{1}{x \ln(x)} \, dx \\
 &= \lim_{v \rightarrow +\infty} \int_2^v \frac{1}{\ln(x)} \cdot \frac{1}{x} \, dx \\
 &= \lim_{v \rightarrow +\infty} \int_{\ln(2)}^{\ln(v)} \frac{1}{u} \, du \\
 &= \lim_{v \rightarrow +\infty} [\ln(u)]_{\ln(2)}^{\ln(v)} \\
 &= \lim_{v \rightarrow +\infty} \ln(\ln(v)) - \ln(\ln(2)) \\
 &= +\infty .
 \end{aligned}$$

Also ist auch $\sum_{j=2}^{\infty} \frac{1}{j \ln(j)} = +\infty$.

Der Extra-Faktor $\ln(j)$ hat es also nicht fertiggebracht, aus der transvergenten Reihe $\sum_{j=2}^{\infty} \frac{1}{j}$ eine konvergente Reihe zu machen.

Kapitel 5

Funktionen in mehreren Variablen

5.1 Partielle Ableitungen

Sei $n \geq 1$.

Definition. Sei $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Es war $\|x\| = (x_1^2 + \dots + x_n^2)^{1/2}$ die Norm von x .

Für $x, y \in \mathbb{R}^n$ ist $\|x - y\|$ der *Abstand* der Punkte x und y .

Definition. Sei $x \in \mathbb{R}^n$. Sei $\delta \in \mathbb{R}_{>0}$. Sei $U_\delta(x) := \{y \in \mathbb{R}^n : \|x - y\| < \delta\}$.

Definition. Es heißt $D \subseteq \mathbb{R}^n$ eine *offene* Teilmenge, wenn es für jedes $x \in D$ ein $\delta \in \mathbb{R}_{>0}$ gibt mit $U_\delta(x) \subseteq D$.

Offene Teilmengen zeichnen sich dadurch aus, daß ihr Rand nicht zu ihnen gehört.

Sei im weiteren Verlauf dieses Abschnitts $D \subseteq \mathbb{R}^n$ offen und

$$f : D \rightarrow \mathbb{R}$$

eine Funktion.

Definition. Sei $x \in D$. Es heißt f *stetig in x* , falls es für jedes $\varepsilon \in \mathbb{R}_{>0}$ ein $\delta \in \mathbb{R}_{>0}$ gibt mit $f(U_\delta(x)) \subseteq U_\varepsilon(f(x))$.

Es heißt f *stetig*, falls sie stetig ist in jedem Punkt $x \in D$.

Bemerkung. Die Funktion $f(x) = x_i$ ist stetig auf \mathbb{R}^n . Konstante Funktionen sind stetig. Summen und Produkte stetiger Funktionen sind stetig. Das Kompositum stetiger Funktionen ist stetig.

Die meisten Funktionen sind also stetig. Unstetige Funktionen können künstlich produziert werden, z.B. über Fallunterscheidungen.

Definition. Sei $x = (x_1, \dots, x_n) \in D$.

Sei $1 \leq i \leq n$. Sei $D_i := \{t \in \mathbb{R} : (x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n) \in D\}$.

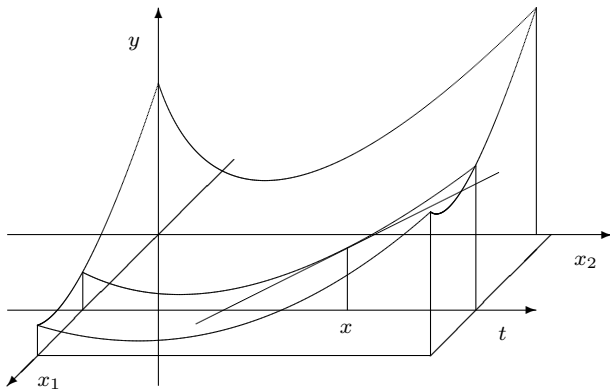
Betrachte die Funktion $D_i \rightarrow \mathbb{R}, t \mapsto f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$. Falls die Ableitung dieser Funktion an der Stelle $t = x_i$ existiert, so heißt f *partiell nach x_i differenzierbar an der Stelle x* , und den Wert dieser Ableitung bezeichnen wir mit

$$f_{x_i}(x_1, \dots, x_n),$$

genannt die *partielle Ableitung von f nach x_i* an der Stelle x .

Die gekrümmte Fläche in folgender Darstellung ist der Graph einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Genauer gesagt, ein Ausschnitt dieses Graphen, der ins Bild gepaßt hat.

Auf dieser Fläche liegen also Punkte (x_1, x_2, y) , die $y = f(x_1, x_2)$ erfüllen.



Die Steigung der eingezeichneten Tangente ist $f_{x_2}(x)$.

Ist f partiell nach x_i differenzierbar an jeder Stelle $x \in D$, so heißt f *partiell differenzierbar nach x_i* . Diesemfalls resultiert die partielle Ableitung $f_{x_i} : D \rightarrow \mathbb{R}$.

Man schreibt auch $\frac{\partial f}{\partial x_i} := f_{x_i}$.

In anderen Worten, die partielle Ableitung nach x_i erhält man, indem man alle anderen Variablen $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ als konstant betrachtet, nur x_i als variabel, und dann die Funktion in dieser verbliebenen Variablen x_i ableitet.

Beispiel. Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R} : x \mapsto f(x) = f(x_1, x_2, x_3) := x_1^2 \cdot x_2^3 \cdot x_3^7$. Es ist

$$\begin{aligned} f_{x_1}(x_1, x_2, x_3) &= 2x_1 \cdot x_2^3 \cdot x_3^7 \\ f_{x_2}(x_1, x_2, x_3) &= 3x_1^2 \cdot x_2^2 \cdot x_3^7 \\ f_{x_3}(x_1, x_2, x_3) &= 7x_1^2 \cdot x_2^3 \cdot x_3^6. \end{aligned}$$

Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R} : x \mapsto f(x) = f(x_1, x_2) := \exp(x_1 - 2x_2^3) + x_1$. Es ist

$$\begin{aligned} f_{x_1}(x_1, x_2) &= \exp(x_1 - 2x_2^3) + 1 \\ f_{x_2}(x_1, x_2) &= \exp(x_1 - 2x_2^3) \cdot (-6)x_2^2 \end{aligned}$$

Definition. Sei f_{x_i} existent für $i \in \{1, \dots, n\}$. Sei $x \in D$.

Der *Gradient* von f an der Stelle x ist der Vektor

$$\nabla_f(x) := \begin{pmatrix} f_{x_1}(x) \\ f_{x_2}(x) \\ \vdots \\ f_{x_n}(x) \end{pmatrix} \in \mathbb{R}^{n \times 1}.$$

Der Buchstabe ∇ spricht sich "nabla"; vgl. [en.wikipedia.org/wiki/Nabla_\(instrument\)](http://en.wikipedia.org/wiki/Nabla_(instrument)).

Definition. Wir können das partielle Ableiten auch iterieren, sofern die Funktion es zulässt. Wir schreiben dann kurz

$$\begin{aligned} f_{x_i x_j}(x) &:= (f_{x_i})_{x_j}(x) \\ f_{x_i x_j x_k}(x) &:= ((f_{x_i})_{x_j})_{x_k}(x), \end{aligned}$$

wobei $i, j, k \in \{1, \dots, n\}$. Usf.

Definition. Sei $f_{x_i x_j}$ existent für $i, j \in \{1, \dots, n\}$. Sei $x \in D$.

Die *Hessematrix* von f an der Stelle x ist die Matrix

$$H_f(x) := \begin{pmatrix} f_{x_1 x_1}(x) & f_{x_1 x_2}(x) & \dots & f_{x_1 x_n}(x) \\ f_{x_2 x_1}(x) & f_{x_2 x_2}(x) & \dots & f_{x_2 x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1}(x) & f_{x_n x_2}(x) & \dots & f_{x_n x_n}(x) \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Beispiel. Sei wieder $f : \mathbb{R}^3 \rightarrow \mathbb{R} : x \mapsto f(x) = f(x_1, x_2, x_3) := x_1^2 \cdot x_2^3 \cdot x_3^7$.

Gradient und Hessematrix sind wie folgt.

$$\nabla_f(x) = \begin{pmatrix} 2x_1 \cdot x_2^3 \cdot x_3^7 \\ 3x_1^2 \cdot x_2^2 \cdot x_3^7 \\ 7x_1^2 \cdot x_2^3 \cdot x_3^6 \end{pmatrix} \in \mathbb{R}^{3 \times 1}, \quad H_f(x) = \begin{pmatrix} 2x_2^3 \cdot x_3^7 & 6x_1 \cdot x_2^2 \cdot x_3^7 & 14x_1 \cdot x_2^3 \cdot x_3^6 \\ 6x_1 \cdot x_2^2 \cdot x_3^7 & 6x_1^2 \cdot x_2 \cdot x_3^7 & 21x_1^2 \cdot x_2^2 \cdot x_3^6 \\ 14x_1 \cdot x_2^3 \cdot x_3^6 & 21x_1^2 \cdot x_2^2 \cdot x_3^6 & 42x_1^2 \cdot x_2^3 \cdot x_3^5 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$$

Beispiel. Sei wieder $f : \mathbb{R}^2 \rightarrow \mathbb{R} : x \mapsto f(x) = f(x_1, x_2) := \exp(x_1 - 2x_2^3) + x_1$.

Gradient und Hessematrix sind wie folgt.

$$\nabla_f(x) = \begin{pmatrix} \exp(x_1 - 2x_2^3) + 1 \\ -6 \exp(x_1 - 2x_2^3) x_2^2 \end{pmatrix}, \quad H_f(x) = \exp(x_1 - 2x_2^3) \cdot \begin{pmatrix} 1 & -6x_2^2 \\ -6x_2^2 & 36x_2^4 - 12x_2 \end{pmatrix}$$

Lemma (Schwarz). Sei $f_{x_i x_j}$ existent und stetig für $i, j \in \{1, \dots, n\}$. Sei $x \in D$.

Dann ist die Hessematrix $H_f(x) \in \mathbb{R}^{n \times n}$ symmetrisch.

Mit anderen Worten, es ist $f_{x_i x_j}(x) = f_{x_j x_i}(x)$ für $i, j \in \{1, \dots, n\}$.

Bemerkung. Manchmal verwendet man statt der Variablenamen x_1, x_2, x_3, \dots auch Variablenamen x, y, z, \dots oder u, v, w, \dots usf. Dann müssen nicht so viele Indizes notiert werden.

Ist z.B. $f(x, y) = x \sin(2x + y)$, dann ist $f_x(x, y) = \sin(2x + y) + 2x \cos(2x + y)$ und $f_{xy}(x, y) = \cos(2x + y) - 2x \sin(2x + y)$.

5.2 Definitheit

Sei $n \geq 1$. Sei $A = (a_{i,j})_{i,j} \in \mathbb{R}^{n \times n}$ symmetrisch. Wir erinnern daran, daß A orthogonal diagonalisierbar ist, mit einer resultierenden Diagonalmatrix in $\mathbb{R}^{n \times n}$; cf. §3.8.2.

Definition.

- (1) Es heißt A *positiv definit*, falls $x^t Ax > 0$ ist für $x \in \mathbb{R}^{n \times 1} \setminus \{0\}$.
- (2) Es heißt A *negativ definit*, falls $x^t Ax < 0$ ist für $x \in \mathbb{R}^{n \times 1} \setminus \{0\}$.
- (3) Es heißt A *indefinit*, falls es $x, y \in \mathbb{R}^{n \times 1}$ gibt mit $x^t Ax > 0$ und $y^t Ay < 0$.

Beispiel.

- (1) Es ist $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ positiv definit, da $(x_1 \ x_2) \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 3x_2^2 > 0$ ist für $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
- (2) Es ist $\begin{pmatrix} 9 & 5 \\ 5 & -3 \end{pmatrix}$ indefinit, da $(1 \ 0) \begin{pmatrix} 9 & 5 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 9 > 0$ und $(0 \ 1) \begin{pmatrix} 9 & 5 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -3 < 0$ ist.

Bemerkung.

- (1) Es ist A positiv definit genau dann, wenn alle seine Eigenwerte in $\mathbb{R}_{>0}$ liegen.
- (2) Es ist A negativ definit genau dann, wenn alle seine Eigenwerte in $\mathbb{R}_{<0}$ liegen.
- (3) Es ist A indefinit genau dann, wenn von den Eigenwerten von A wenigstens einer in $\mathbb{R}_{>0}$ und wenigstens einer in $\mathbb{R}_{<0}$ liegt.

Bemerkung. Es gibt symmetrische Matrizen in $\mathbb{R}^{n \times n}$, die weder positiv definit noch negativ definit noch indefinit sind.

Zum Beispiel trifft dies für $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ zu, da diese Matrix die Eigenwerte 0 und 1 hat.

Bemerkung. Ist $\det(A) \neq 0$ und ist A weder positiv noch negativ definit, dann ist A indefinit.

Bemerkung. Es ist A negativ definit genau dann, wenn $-A$ positiv definit ist.

Definition. Sei $1 \leq k \leq n$. Der k -te *Hauptminor* von A ist definiert als

$$M_k(A) := \det \begin{pmatrix} a_{1,1} & \dots & a_{1,k} \\ \vdots & & \vdots \\ a_{k,1} & \dots & a_{k,k} \end{pmatrix}.$$

Es ist also $M_k(A)$ die Determinante der Matrix im oberen linken $k \times k$ -Block von A .

Beispiel. Die Hauptminoren von $A := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}$ sind

$$\begin{aligned} M_1(A) &= \det(1) &= 1 \\ M_2(A) &= \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} &= 1 \\ M_3(A) &= \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix} &= 2 \end{aligned}$$

Lemma (Hauptminorenkriterium).

- (1) Es ist A positiv definit genau dann, wenn $M_j(A) > 0$ ist für $1 \leq j \leq n$.
- (2) Es ist A negativ definit genau dann, wenn $(-1)^j \cdot M_j(A) > 0$ ist für $1 \leq j \leq n$.

Für positive Definitheit muß also gelten: $M_1(A) > 0, M_2(A) > 0, M_3(A) > 0, M_4(A) > 0, \dots$

Für negative Definitheit muß also gelten: $M_1(A) < 0, M_2(A) > 0, M_3(A) < 0, M_4(A) > 0, \dots$

Beispiel. Es ist $A := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}$ positiv definit, da $M_1(A) = 1 > 0, M_2(A) = 1 > 0$ und $M_3(A) = 2 > 0$ ist. Vgl. voriges Beispiel.

Es ist $-A$ negativ definit, da $M_1(-A) = -1 < 0, M_2(-A) = 1 > 0$ und $M_3(-A) = -2 < 0$ ist.

5.3 Lokale Extremstellen

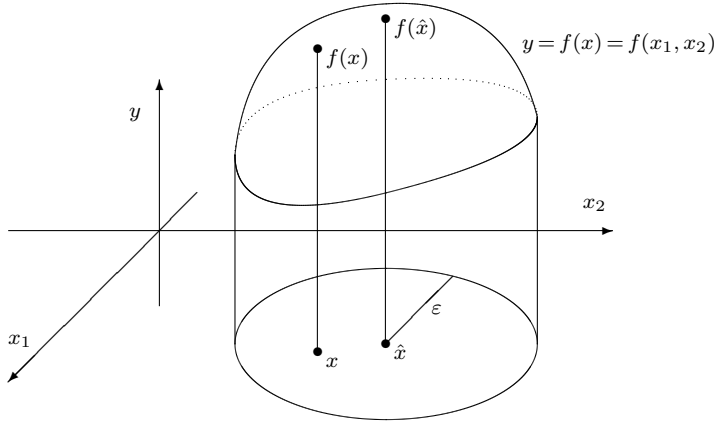
Sei $n \geq 1$. Sei $D \subseteq \mathbb{R}^n$ offen.

Sei $f : D \rightarrow \mathbb{R}$ gegeben mit $f_{x_i x_j} : D \rightarrow \mathbb{R}$ existent und stetig für $i, j \in \{1, \dots, n\}$.

Definition. Sei $\hat{x} \in D$ gegeben.

- (1) Es heißt \hat{x} eine *lokale Maximalstelle* von f , falls es ein $\varepsilon \in \mathbb{R}_{>0}$ gibt mit $f(\hat{x}) \geq f(x)$ für $x \in U_\varepsilon(\hat{x}) \cap D$.
- (2) Es heißt \hat{x} eine *lokale Minimalstelle* von f , falls es ein $\varepsilon \in \mathbb{R}_{>0}$ gibt mit $f(\hat{x}) \leq f(x)$ für $x \in U_\varepsilon(\hat{x}) \cap D$.
- (3) Es heißt \hat{x} eine *lokale Extremstelle* von f , falls \hat{x} eine lokale Maximal- oder eine lokale Minimalstelle von f ist.

Skizze einer lokalen Maximalstelle \hat{x} von f im Fall $n = 2$.



Definition. Sei $\hat{x} \in D$. Es heißt \hat{x} eine *Flachstelle* von f , falls $\nabla_f(\hat{x}) = 0$ ist.

Bemerkung. Ist $n = 2$, dann sind die Flachstellen von f die Stellen, an denen der Graph von f eine horizontale Tangentialebene hat.

Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R} : x = (x_1, x_2) \mapsto f(x) = f(x_1, x_2) = 3x_1^2 + x_2^2 + 2x_1^3$.

Es ist $\nabla_f(x) = \nabla_f(x_1, x_2) = \begin{pmatrix} 6x_1 + 6x_1^2 \\ 2x_2 \end{pmatrix}$.

Es ist $\nabla_f(x) = 0$ genau dann, wenn $6x_1 + 6x_1^2 = 0$ und $2x_2 = 0$ ist, also genau dann, wenn $(x_1, x_2) \in \{(0, 0), (-1, 0)\}$ liegt.

Mit anderen Worten, es hat f die Flachstellen $(0, 0)$ und $(-1, 0)$.

Bemerkung. Ist $\hat{x} \in D$ eine lokale Extremstelle von f , dann ist \hat{x} eine Flachstelle von f .

Hat man also die Flachstellen von f aufgelistet, so befinden sich alle lokalen Extremstellen unter diesen Flachstellen.

Definition. Eine Flachstelle von f , die keine lokale Extremstelle ist, heißt *Sattelstelle* von f .

Lemma. Sei \hat{x} eine Flachstelle von f .

- (1) Ist die Hessematrix $H_f(\hat{x})$ negativ definit, dann ist \hat{x} eine lokale Maximalstelle von f .
- (2) Ist die Hessematrix $H_f(\hat{x})$ positiv definit, dann ist \hat{x} eine lokale Minimalstelle von f .
- (3) Ist die Hessematrix $H_f(\hat{x})$ indefinit, dann ist \hat{x} eine Sattelstelle von f .

Ist die Hessematrix $H_f(\hat{x})$ weder negativ definit noch positiv definit noch indefinit, dann können wir mit diesem Lemma keine Aussage über die Art der Flachstelle \hat{x} machen.

Falls die Hessematrix in keinen der drei Fälle paßt, dann ist für die Flachstelle in der Tat alles möglich. So eine Stelle wird in unseren Aufgaben nicht auf ihre Art hin zu untersuchen sein.

Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R} : x = (x_1, x_2) \mapsto f(x) = f(x_1, x_2) = 3x_1^2 + x_2^2 + 2x_1^3$, mit den Flachstellen $(-1, 0)$ und $(0, 0)$, wie im vorigen Beispiel gesehen.

Es war $\nabla_f(x) = \begin{pmatrix} 6x_1+6x_1^2 \\ 2x_2 \end{pmatrix}$. Also wird $H_f(x) = H_f(x_1, x_2) = \begin{pmatrix} 6+12x_1 & 0 \\ 0 & 2 \end{pmatrix}$.

Flachstelle $(-1, 0)$. Es ist $H_f(-1, 0) = \begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix}$ indefinit. Also ist $(-1, 0)$ eine Sattelstelle.

Flachstelle $(0, 0)$. Es ist $H_f(0, 0) = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}$ positiv definit. Also ist $(0, 0)$ eine lokale Minimalstelle.

Wir haben die Bezeichnung \hat{x} verwendet, um die fixierte Stelle \hat{x} besser vom variablen Punkt x unterscheiden zu können. In der Praxis muß man diese Bezeichnung nicht unbedingt verwenden.

In der Praxis heißen die Variablen x_1, x_2, x_3, \dots auch oft wieder x, y, z, \dots

Unsere Methode verallgemeinert die bekannte Methode im Fall $n = 1$:

Beispiel. Sei $n = 1$. Sei $D = \mathbb{R}$. Sei $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f(x)$.

Es wird $f_x = f'$ und $f_{xx} = f''$.

Also wird $\nabla_f(x) = (f'(x))$ und $H_f(x) = (f''(x))$.

Es ist $\hat{x} \in \mathbb{R}$ eine Flachstelle von f genau dann, wenn $f'(\hat{x}) = 0$ ist.

Sei nun \hat{x} eine Flachstelle von f .

Es ist $\hat{x} \in \mathbb{R}$ eine lokale Maximalstelle von f , falls $(f''(\hat{x}))$ negativ definit ist, d.h. falls $f''(\hat{x}) < 0$ ist.

Es ist $\hat{x} \in \mathbb{R}$ eine lokale Minimalstelle von f , falls $(f''(\hat{x}))$ positiv definit ist, d.h. falls $f''(\hat{x}) > 0$ ist.

Ist hingegen $f''(\hat{x}) = 0$, dann ist $H_f(\hat{x}) = (0)$ weder positiv definit noch negativ definit noch indefinit, und wir machen mit unserer Methode keine Aussage über die Art der Flachstelle \hat{x} .

Beispiel. Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R} : (x, y, z) \mapsto e^{x+y+z}(xy + 2z)$.

Der Gradient von f ist

$$\nabla_f(x, y, z) = \begin{pmatrix} e^{x+y+z}(xy+2z+y) \\ e^{x+y+z}(xy+2z+x) \\ e^{x+y+z}(xy+2z+2) \end{pmatrix} = e^{x+y+z} \cdot \begin{pmatrix} xy+2z+y \\ xy+2z+x \\ xy+2z+2 \end{pmatrix}.$$

Um die Flachstellen zu bestimmen, lösen wir die vektorielle Gleichung $\nabla_f(x, y, z) = 0$. Wir lösen also folgendes Gleichungssystem.

$$\begin{aligned} xy + 2z + y &= 0 \\ xy + 2z + x &= 0 \\ xy + 2z + 2 &= 0 \end{aligned}$$

Aus den ersten beiden Gleichungen folgt $x = y$. Aus den letzten beiden Gleichungen folgt $x = 2$. Einsetzen in die letzte Gleichung gibt $z = -3$.

Wir haben also die Flachstelle $(2, 2, -3)$.

Die Hessematrix von f ist

$$\begin{aligned} H_f(x, y, z) &= \begin{pmatrix} e^{x+y+z}(xy+2z+2y) & e^{x+y+z}(xy+2z+y+x+1) & e^{x+y+z}(xy+2z+y+2) \\ e^{x+y+z}(xy+2z+x+y+1) & e^{x+y+z}(xy+2z+2x) & e^{x+y+z}(xy+2z+x+2) \\ e^{x+y+z}(xy+2z+2+y) & e^{x+y+z}(xy+2z+2+x) & e^{x+y+z}(xy+2z+4) \end{pmatrix} \\ &= e^{x+y+z} \cdot \begin{pmatrix} xy+2z+2y & xy+2z+y+x+1 & xy+2z+y+2 \\ xy+2z+x+y+1 & xy+2z+2x & xy+2z+x+2 \\ xy+2z+2+y & xy+2z+2+x & xy+2z+4 \end{pmatrix}. \end{aligned}$$

Flachstelle $(2, 2, -3)$. Es ist

$$H := H_f(2, 2, -3) = e \cdot \begin{pmatrix} 2 & 3 & 2 \\ 3 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

Die Hauptminoren von H sind $M_1(H) = \det(e \cdot (2)) = 2e$, $M_2(H) = \det(e \cdot \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}) = -5e^2$ und $M_3(H) = \det(H) = -2e^3$. Somit ist H weder positiv noch negativ definit. Da zudem $\det(H) \neq 0$ ist, folgt, daß H indefinit ist.

Somit ist $(2, 2, -3)$ eine Sattelstelle von f .

5.4 Lokale Extremstellen unter Nebenbedingungen

Sei $n \in \mathbb{Z}_{\geq 1}$. Sei $D \subseteq \mathbb{R}^n$ offen.

Sei $f : D \rightarrow \mathbb{R}$ gegeben mit $f_{x_i x_j} : D \rightarrow \mathbb{R}$ existent und stetig für $i, j \in \{1, \dots, n\}$.

Sei $\ell \in \mathbb{Z}_{\geq 1}$. Sei $g_k : D \rightarrow \mathbb{R}$ für $1 \leq k \leq \ell$ gegeben mit $(g_k)_{x_i x_j} : D \rightarrow \mathbb{R}$ existent und stetig für $i, j \in \{1, \dots, n\}$. Wir fassen diese Funktionen zusammen zur Abbildung

$$\begin{aligned} g : D &\rightarrow \mathbb{R}^\ell \\ x &\mapsto g(x) := (g_1(x), \dots, g_\ell(x)). \end{aligned}$$

Sei

$$D_g := \{x \in D : g_k(x) = 0 \text{ für } 1 \leq k \leq \ell\}.$$

Es ist also D_g die Menge der Punkte aus D , an denen alle Funktionen g_k den Wert 0 haben, für $1 \leq k \leq \ell$.

Das Ziel ist die Ermittlung der lokalen Extremstellen von f auf D_g .

Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R} : x = (x_1, x_2) \mapsto f(x_1, x_2) = x_1 \cdot x_2$.

Sei $\ell = 1$. Sei $g_1 : \mathbb{R}^2 \rightarrow \mathbb{R} : x = (x_1, x_2) \mapsto g_1(x_1, x_2) = x_1^2 + x_2^2 - 2$. Es ist $g(x) = (g_1(x))$.

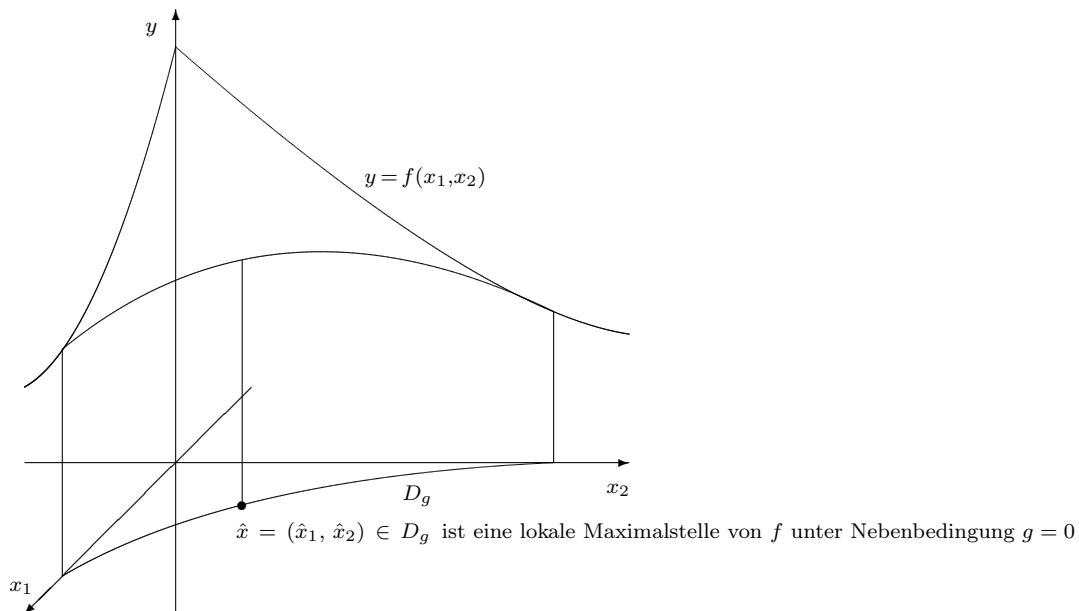
Es ist $D_g = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 - 2 = 0\}$ ein Kreis mit Radius $\sqrt{2}$ um den Mittelpunkt $(0, 0)$.

Es wird uns im folgenden interessieren, wo auf diesem Kreis D_g unsere Funktion f eine lokale Maximalstelle oder eine lokale Minimalstelle hat.

Definition. Sei $\hat{x} \in D_g$.

- (1) Es heißt \hat{x} eine *lokale Maximalstelle* von f unter Nebenbedingung $g = 0$, falls es ein $\varepsilon \in \mathbb{R}_{>0}$ gibt mit $f(\hat{x}) \geq f(x)$ für $x \in U_\varepsilon(\hat{x}) \cap D_g$.
- (2) Es heißt \hat{x} eine *lokale Minimalstelle* von f unter Nebenbedingung $g = 0$, falls es ein $\varepsilon \in \mathbb{R}_{>0}$ gibt mit $f(\hat{x}) \leq f(x)$ für $x \in U_\varepsilon(\hat{x}) \cap D_g$.
- (3) Es heißt \hat{x} eine *lokale Extremstelle* von f unter Nebenbedingung $g = 0$, falls \hat{x} eine lokale Maximal- oder Minimalstelle von f unter Nebenbedingung $g = 0$ ist.

Es ist “unter Nebenbedingung $g = 0$ ” eine feste Wendung, mit der nicht zum Ausdruck kommen soll, daß g als Funktion null ist, sondern, daß die Nullstellenmenge D_g von g der Betrachtung zugrundegelegt wird.



Definition. Sei $\hat{x} \in D_g$.

Sei $N(\hat{x}) \in \mathbb{R}^{n \times \ell}$ die Matrix mit dem Spaltentupel $(\nabla_{g_1}(\hat{x}), \nabla_{g_2}(\hat{x}), \dots, \nabla_{g_\ell}(\hat{x}))$.

Wenn das Spaltentupel von $N(\hat{x})$ linear unabhängig ist und wenn

$$\nabla_f(\hat{x}) = N(\hat{x}) \cdot \lambda$$

eine Lösung $\lambda \in \mathbb{R}^{\ell \times 1}$ besitzt, dann heißt \hat{x} eine *Flachstelle* von f unter Nebenbedingung $g = 0$.

Die Aussage, daß das Spaltentupel von $N(\hat{x})$ linear unabhängig ist, ist hierbei äquivalent zur Aussage, daß der Rang von $N(\hat{x})$ gleich ℓ ist.

Der dann eindeutig festliegende Vektor $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_\ell \end{pmatrix}$ heißt *Lagrange-Multiplikator* zur Flachstelle \hat{x} unter Nebenbedingung $g = 0$.

Eine Flachstelle \hat{x} unter Nebenbedingung $g = 0$ muß also Lösung des folgenden *Lagrange-Gleichungssystems* sein.

$$\begin{aligned} f_{x_1}(\hat{x}) &= \lambda_1 \cdot (g_1)_{x_1}(\hat{x}) + \dots + \lambda_\ell \cdot (g_\ell)_{x_1}(\hat{x}) \\ f_{x_2}(\hat{x}) &= \lambda_1 \cdot (g_1)_{x_2}(\hat{x}) + \dots + \lambda_\ell \cdot (g_\ell)_{x_2}(\hat{x}) \\ &\vdots \\ f_{x_n}(\hat{x}) &= \lambda_1 \cdot (g_1)_{x_n}(\hat{x}) + \dots + \lambda_\ell \cdot (g_\ell)_{x_n}(\hat{x}) \\ g_1(\hat{x}) &= 0 \\ &\vdots \\ g_\ell(\hat{x}) &= 0 \end{aligned}$$

Die letzten ℓ Gleichungen hierin stellen nur sicher, daß tatsächlich $\hat{x} \in D_g$ liegt.

Wir haben also $n + \ell$ Gleichungen für die $n + \ell$ Unbekannten $\hat{x}_1, \dots, \hat{x}_n, \lambda_1, \dots, \lambda_\ell$.

Auch hier benötigt man die Bezeichnung $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)$ nur in der Theorie, um dort zwischen fester Stelle \hat{x} und variabler Stelle x unterscheiden zu können.

Beispiel. Wir setzen voriges Beispiel fort.

Es ist $\nabla_f(x_1, x_2) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$. Es ist $\nabla_{g_1}(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$.

Die Unbekannten x_1, x_2, λ_1 haben das Lagrange-Gleichungssystem

$$\begin{aligned} x_2 &= \lambda_1 \cdot 2x_1 \\ x_1 &= \lambda_1 \cdot 2x_2 \\ x_1^2 + x_2^2 - 2 &= 0 \end{aligned}$$

zu erfüllen.

Setzen wir x_1 aus der zweiten Gleichung in die erste ein, so erhalten wir

$$x_2 = \lambda_1 \cdot 2(\lambda_1 \cdot 2x_2) = 4\lambda_1^2 x_2 .$$

Fall $x_2 = 0$. Dann ist auch $x_1 = 0$ und die dritte Gleichung ist unerfüllbar.

Fall $x_2 \neq 0$. Dann folgt aus $x_2 = 4\lambda_1^2 x_2$, daß $1 = 4\lambda_1^2$ sein muß.

Subfall $\lambda_1 = \frac{1}{2}$. Dank zweiter Gleichung ist $x_1 = x_2$. Die dritte Gleichung gibt $x_1 = 1$ oder $x_1 = -1$. Also erhalten wir in diesem Subfall die folgenden Flachstellen unter Nebenbedingung $g = 0$.

$$\begin{aligned} (1, 1) &\quad \text{mit } \lambda_1 = \frac{1}{2} \\ (-1, -1) &\quad \text{mit } \lambda_1 = \frac{1}{2} \end{aligned}$$

Subfall $\lambda_1 = -\frac{1}{2}$. Dank zweiter Gleichung ist $x_1 = -x_2$. Die dritte Gleichung gibt $x_1 = 1$ oder $x_1 = -1$. Also erhalten wir in diesem Subfall die folgenden Flachstellen unter Nebenbedingung $g = 0$.

$$\begin{aligned} (1, -1) & \text{ mit } \lambda_1 = -\frac{1}{2} \\ (-1, 1) & \text{ mit } \lambda_1 = -\frac{1}{2} \end{aligned}$$

Zur Probe vergewissere man sich, daß all diese Stellen mit ihrem zugehörigem Lagrange-Multiplikator das Lagrange-Gleichungssystem erfüllen und daß die Matrix $N(x) \in \mathbb{R}^{2 \times 1}$ an diesen Stellen Rang 1 hat. Im vorliegenden Fall einer einzigen Spalte $\nabla_{g_1}(x_1, x_2)$ darf diese dazu nicht die Nullspalte sein.

Lemma. Sei $\hat{x} \in D$ eine Flachstelle von f unter Nebenbedingung $g = 0$.

Sei $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_\ell \end{pmatrix}$ der zugehörige Lagrange-Multiplikator.

Sei

$$H := H_f(\hat{x}) - (\lambda_1 H_{g_1}(\hat{x}) + \dots + \lambda_\ell H_{g_\ell}(\hat{x})) \in \mathbb{R}^{n \times n}.$$

Wir lösen das lineare Gleichungssystem $N(\hat{x})^t u = 0$ in $u \in \mathbb{R}^{n \times 1}$.

Sei $(u_1, \dots, u_{n-\ell})$ eine Basis des Lösungsraums $\{u \in \mathbb{R}^{n \times 1} : N(\hat{x})^t u = 0\}$.

Sei $U \in \mathbb{R}^{n \times (n-\ell)}$ die Matrix mit Spaltentupel $(u_1, \dots, u_{n-\ell})$.

Mit anderen Worten, habe U ein linear unabhängiges Spaltentupel aus $n - \ell$ Vektoren, die auf allen Spalten von $N(\hat{x})$ senkrecht stehen.

Wir haben nun die symmetrische Matrix $U^t \cdot H \cdot U \in \mathbb{R}^{(n-\ell) \times (n-\ell)}$ zu betrachten.

- (1) Ist $U^t \cdot H \cdot U$ negativ definit, dann ist \hat{x} eine lokale Maximalstelle von f unter Nebenbedingung $g = 0$.
- (2) Ist $U^t \cdot H \cdot U$ positiv definit, dann ist \hat{x} eine lokale Minimalstelle von f unter Nebenbedingung $g = 0$.
- (3) Ist $U^t \cdot H \cdot U$ indefinit, dann ist \hat{x} eine Sattelstelle von f unter Nebenbedingung $g = 0$.

Ist $U^t \cdot H \cdot U$ weder negativ definit noch positiv definit noch indefinit, dann können wir mit diesem Lemma keine Aussage treffen über die Art der Flachstelle \hat{x} unter Nebenbedingung $g = 0$.

Falls $U^t \cdot H \cdot U$ in keinen der drei Fälle paßt, dann ist für die Flachstelle \hat{x} unter Nebenbedingung $g = 0$ in der Tat alles möglich. So eine Stelle wird in unseren Aufgaben nicht auf ihre Art hin zu untersuchen sein.

Beispiel. Wir setzen voriges Beispiel fort.

Aus $\nabla_f(x_1, x_2) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$ erhalten wir $H_f(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Aus $\nabla_{g_1}(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$ erhalten wir $H_{g_1}(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

Wir haben die folgenden Flachstellen von f unter Nebenbedingung $g = 0$ zu betrachten.

Stelle $(x_1, x_2) = (1, 1)$ mit $\lambda_1 = \frac{1}{2}$. Es ist $N(1, 1) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$. Also können wir $U = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ wählen.

Es ist $H = H_f(1, 1) - \lambda_1 H_{g_1}(1, 1) = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$. Es ist $U^t \cdot H \cdot U = \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-4)$ negativ definit. Wir haben erhalten:

Es ist $(1, 1)$ eine lokale Maximalstelle von f unter Nebenbedingung $g = 0$.

Stelle $(x_1, x_2) = (-1, -1)$ mit $\lambda_1 = \frac{1}{2}$. Es ist $N(-1, -1) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$. Also können wir $U = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ wählen.

Es ist $H = H_f(-1, -1) - \lambda_1 H_{g_1}(-1, -1) = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$. Es ist $U^t \cdot H \cdot U = \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-4)$ negativ definit. Wir haben erhalten:

Es ist $(-1, -1)$ eine lokale Maximalstelle von f unter Nebenbedingung $g = 0$.

Stelle $(x_1, x_2) = (1, -1)$ mit $\lambda_1 = -\frac{1}{2}$. Es ist $N(1, -1) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$. Also können wir $U = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ wählen.

Es ist $H = H_f(1, -1) - \lambda_1 H_{g_1}(1, -1) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Es ist $U^t \cdot H \cdot U = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (4)$ positiv definit. Wir haben erhalten:

Es ist $(1, -1)$ eine lokale Minimalstelle von f unter Nebenbedingung $g = 0$.

Stelle $(x_1, x_2) = (-1, 1)$ mit $\lambda_1 = -\frac{1}{2}$. Es ist $N(-1, 1) = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$. Also können wir $U = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ wählen.

Es ist $H = H_f(-1, 1) - \lambda_1 H_{g_1}(-1, 1) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Es ist $U^t \cdot H \cdot U = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (4)$ positiv definit. Wir haben erhalten:

Es ist $(-1, 1)$ eine lokale Minimalstelle von f unter Nebenbedingung $g = 0$.

Auch hier werden in der Praxis oft statt x_1, x_2, x_3, \dots Variablenamen wie x, y, z, \dots verwendet:

Beispiel. Sei $f : \mathbb{R}^4 \rightarrow \mathbb{R} : (x, y, z, w) \mapsto f(x, y, z, w) := (x + y) \cdot (z + w)$.

Sei $g_1 : \mathbb{R}^4 \rightarrow \mathbb{R} : (x, y, z, w) \mapsto g_1(x, y, z, w) := x^2 + y^2 - 2$.

Sei $g_2 : \mathbb{R}^4 \rightarrow \mathbb{R} : (x, y, z, w) \mapsto g_2(x, y, z, w) := z^2 + w^2 - 2$.

Dann ist $g := (g_1, g_2) : \mathbb{R}^4 \rightarrow \mathbb{R}^2 : (x, y, z, w) \mapsto (g_1(x, y, z, w), g_2(x, y, z, w))$ zu setzen.

Die Aufgabe sei es, die lokalen Extremstellen von f unter Nebenbedingung $g = 0$ zu finden, die in $\mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ liegen.

Es ist $\nabla_f(x, y, z, w) = \begin{pmatrix} z+w \\ z+w \\ x+y \\ x+y \end{pmatrix}$.

Es ist $\nabla_{g_1}(x, y, z, w) = \begin{pmatrix} 2x \\ 2y \\ 0 \\ 0 \end{pmatrix}$ und $\nabla_{g_2}(x, y, z, w) = \begin{pmatrix} 0 \\ 2z \\ 2w \\ 0 \end{pmatrix}$, also $N(x, y, z, w) = \begin{pmatrix} 2x & 0 \\ 2y & 0 \\ 0 & 2z \\ 0 & 2w \end{pmatrix}$.

Das Lagrange-Gleichungssystem hat folgende Gestalt.

$$z + w = \lambda_1 \cdot 2x + \lambda_2 \cdot 0 \quad (1)$$

$$z + w = \lambda_1 \cdot 2y + \lambda_2 \cdot 0 \quad (2)$$

$$x + y = \lambda_1 \cdot 0 + \lambda_2 \cdot 2z \quad (3)$$

$$x + y = \lambda_1 \cdot 0 + \lambda_2 \cdot 2w \quad (4)$$

$$x^2 + y^2 - 2 = 0 \quad (5)$$

$$z^2 + w^2 - 2 = 0 \quad (6)$$

Die Unbekannten sind $x, y, z, w, \lambda_1, \lambda_2$.

Fall $\lambda_1 \neq 0$. Aus (1) und (2) folgt $x = y$. Aus (5) folgt $x = 1$ oder $x = -1$.

Subfall $\lambda_2 \neq 0$. Aus (3) und (4) folgt $z = w$. Aus (6) folgt $z = 1$ oder $z = -1$.

Wir haben die folgenden Flachstellen von f unter Nebenbedingung $g = 0$ erhalten.

$$\begin{array}{ll} (1, 1, 1, 1) & \text{mit } \lambda_1 = 1 \text{ und } \lambda_2 = 1 \\ (1, 1, -1, -1) & \text{mit } \lambda_1 = -1 \text{ und } \lambda_2 = -1 \\ (-1, -1, 1, 1) & \text{mit } \lambda_1 = -1 \text{ und } \lambda_2 = -1 \\ (-1, -1, -1, -1) & \text{mit } \lambda_1 = 1 \text{ und } \lambda_2 = 1 \end{array}$$

Subfall $\lambda_2 = 0$. Aus (3) folgt $x = -y$. Aus (5) folgt $x = 1$ oder $x = -1$. Aus (1) und (2) folgt $z = -w$. Aus (1) folgt $\lambda_1 = 0$, was nach Fallvoraussetzung nicht sein darf.

Fall $\lambda_1 = 0$. Aus (1) und (2) folgt $z = -w$. Aus (6) folgt $z = 1$ oder $z = -1$. Aus (3) und (4) folgt $x = -y$. Aus (5) folgt $x = 1$ oder $x = -1$. Aus (3) folgt $\lambda_2 = 0$.

Wir haben die folgenden Flachstellen von f unter Nebenbedingung $g = 0$ erhalten.

$$\begin{array}{ll} (1, -1, 1, -1) & \text{mit } \lambda_1 = 0 \text{ und } \lambda_2 = 0 \\ (1, -1, -1, 1) & \text{mit } \lambda_1 = 0 \text{ und } \lambda_2 = 0 \\ (-1, 1, 1, -1) & \text{mit } \lambda_1 = 0 \text{ und } \lambda_2 = 0 \\ (-1, 1, -1, 1) & \text{mit } \lambda_1 = 0 \text{ und } \lambda_2 = 0 \end{array}$$

Von diesen 8 Flachstellen von f unter Nebenbedingung $g = 0$ liegt als einzige $(1, 1, 1, 1)$ in $\mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$. Sie hat $\lambda_1 = 1$ und $\lambda_2 = 1$.

Für diese machen wir eine Probe. Einsetzen von $x = 1, y = 1, z = 1, w = 1, \lambda_1 = 1$ und $\lambda_2 = 1$ liefert im Lagrange-Gleichungssystem tatsächlich sechs erfüllte Gleichungen.

Ferner hat die Matrix $N(1, 1, 1, 1) = \begin{pmatrix} 2 & 0 \\ 2 & 0 \\ 0 & 2 \\ 0 & 2 \end{pmatrix}$ den Rang 2, wie gewünscht.

Untersuchen wir, von welcher Art diese Flachstelle von f unter Nebenbedingung $g = 0$ ist.

Es ist $H_f(x, y, z, w) = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$, $H_{g_1}(x, y, z, w) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ und $H_{g_2}(x, y, z, w) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Also ist zu setzen :

$$H := H_f(1, 1, 1, 1) - (\lambda_1 \cdot H_{g_1}(1, 1, 1, 1) + \lambda_2 \cdot H_{g_2}(1, 1, 1, 1)) = \begin{pmatrix} -2 & 0 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & -2 \end{pmatrix}.$$

Es ist $N(1, 1, 1, 1) = \begin{pmatrix} 2 & 0 \\ 2 & 0 \\ 0 & 2 \\ 0 & 2 \end{pmatrix}$. Wir transponieren und lösen das zugehörige homogene lineare

Gleichungssystem: $\begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$. Eine Basis des Lösungsraums ist gegeben durch $\left(\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right)$. Wir können also $U := \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$ wählen. Es ist

$$U^t \cdot H \cdot U = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix}$$

negativ definit.

Somit ist $(1, 1, 1, 1)$ eine lokale Maximalstelle von f unter Nebenbedingung $g = 0$.

Kapitel 6

Gewöhnliche Differentialgleichungen

Schreibweise. Sei $D \subseteq \mathbb{R}$. Sei $u : D \rightarrow \mathbb{R} : x \mapsto u(x)$ eine Funktion.

In diesem Kapitel ist die Schreibweise dafür oft schlicht: $u(x)$, definiert auf D .

Oder auch: u , definiert auf D .

6.1 Begriff

Definition. Sei $n \geq 1$.

Eine *Differentialgleichung* n -ter Ordnung ist eine Gleichung der Form

$$y^{(n)}(x) = f(x, y^{(n-1)}(x), y^{(n-2)}(x), \dots, y^{(0)}(x))$$

für eine Funktion $f : M \rightarrow \mathbb{R}$, wobei $M \subseteq \mathbb{R}^{n+1}$.

Sei $D \subseteq \mathbb{R}$ offen.

Eine *Lösung* dieser Differentialgleichung auf D ist eine n -fach differenzierbare Funktion $y : D \rightarrow \mathbb{R}$, für welche eben $y^{(n)}(x) = f(x, y^{(n-1)}(x), y^{(n-2)}(x), \dots, y^{(0)}(x))$ gilt für $x \in D$.

Man schreibt die Differentialgleichung auch kurz

$$y^{(n)} = f(x, y^{(n-1)}, y^{(n-2)}, \dots, y^{(0)})$$

Fordert man zudem für ein $\hat{x} \in D$ noch

$$y^{(0)}(\hat{x}) = y_0, \quad y^{(1)}(\hat{x}) = y_1, \quad \dots, \quad y^{(n-1)}(\hat{x}) = y_{n-1}$$

für gewisse $y_0, \dots, y_{n-1} \in \mathbb{R}$, so spricht man von *Anfangsbedingungen* an y .

Zunächst ein paar Beispiele, die nur die Problemstellung illustrieren sollen. Es werden Lösungen nur angegeben, noch nicht hergeleitet.

Beispiel. Es ist

$$y'(x) = y(x)$$

eine Differentialgleichung erster Ordnung.

Diese wird auch kurz

$$y' = y$$

geschrieben.

Für jedes $c \in \mathbb{R}$ ist $y(x) = c e^x$ eine Lösung dieser Differentialgleichung auf \mathbb{R} .

Beispiel. Es ist

$$y' = -y/x$$

eine Differentialgleichung erster Ordnung.

Auf $\mathbb{R}_{>0}$ ist $y(x) = \frac{1}{x}$ eine Lösung dieser Differentialgleichung.

Beispiel. Es ist

$$y'' + y = 0$$

eine Differentialgleichung zweiter Ordnung, bei welcher wir noch y nach links gebracht haben.

Es ist $y(x) = \sin(x)$ eine Lösung dieser Differentialgleichung auf \mathbb{R} .

Es ist $y(x) = \cos(x)$ eine Lösung dieser Differentialgleichung auf \mathbb{R} .

Wollen wir zudem die Anfangsbedingungen $y(0) = 1$ und $y'(0) = 3$ erfüllt haben, so finden wir dafür die Lösung $y(x) = 3 \sin(x) + \cos(x)$ auf \mathbb{R} .

Beispiel. Es ist

$$y' = x^2$$

eine Differentialgleichung erster Ordnung. Jede Stammfunktion $\frac{1}{3}x^3 + c$ von x^2 ist eine Lösung, wobei $c \in \mathbb{R}$.

Das Lösen von Differentialgleichungen verallgemeinert also das Integrieren.

6.2 Differentialgleichungen erster Ordnung

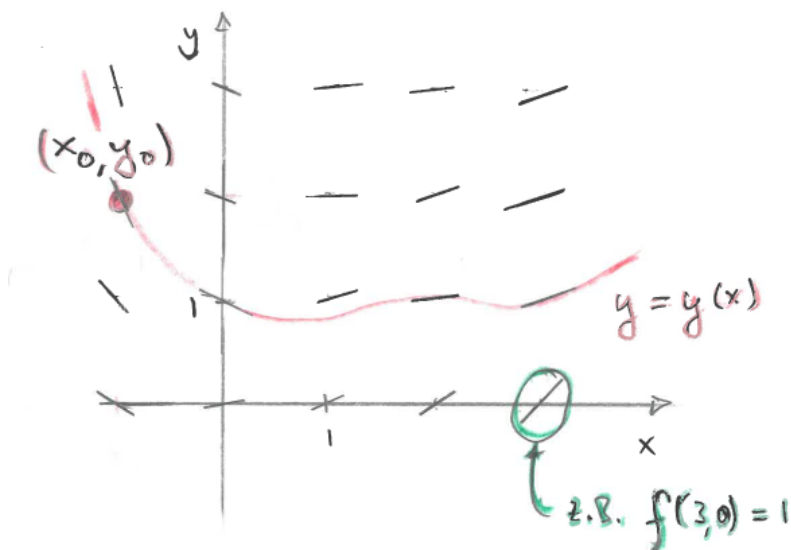
6.2.1 Richtungsfeld

Bevor wir an die Lösungsmethoden gehen, hier eine Veranschaulichung.

Zu lösen sei die Differentialgleichung $y' = f(x, y)$.

Mit anderen Worten, der Graph von $y(x)$ hat an der Stelle $(x, y) \in \mathbb{R}^2$ eine vorgegebene Tangentensteigung $f(x, y)$.

Diese machen wir an einigen Stellen durch eine kurze Strecke kenntlich.



Der Graph $y = y(x)$ einer Lösung mit vorgegebenem Anfangswert $y(x_0) = y_0$ wurde durch dieses Richtungsfeld skizziert.

6.2.2 Separierbare Differentialgleichungen

Seien $f(x)$ und $g(x)$ differenzierbare reellwertige Funktionen auf einem geeigneten Definitionsbereich für die Variable x .

Zu lösen sei die Differentialgleichung

$$y' = \frac{f'(x)}{g'(y)}.$$

Dann muß notwendig

$$\int g'(y(x))y'(x) dx = \int f'(x) dx$$

sein. Substitution auf der linken Seite gibt

$$[g(y)] = \int g'(y) dy = \int f'(x) dx = [f(x)]$$

Also gibt es ein $c \in \mathbb{R}$ mit

$$g(y(x)) = f(x) + c$$

Kann man dies nun nach $y(x)$ auflösen, gegebenenfalls nach Wahl eines hinreichend kleinen Definitionsbereiches für x , so liefert dies eine Lösung der Differentialgleichung.

Eine Probe ist hier verpflichtend. Man setze dazu die gefundene Funktion $y(x)$ in die Differentialgleichung ein und überprüfe die Gleichheit.

Beispiel. Zu lösen sei die Differentialgleichung

$$y' = xy .$$

Wir suchen eine Lösung, die die Anfangsbedingung $y(1) = -3$ erfüllt.

In obiger Bezeichnung sollte $g'(y) = \frac{1}{y}$ und $f'(x) = x$ sein.

Dazu sei $g(y) := \ln(|y|)$ und $f(x) := \frac{1}{2}x^2$.

Zu erfüllen ist

$$[\ln(|y|)] = \int \frac{1}{y} dy = \int \frac{y'}{y} dx \stackrel{!}{=} \int x dx = [\frac{1}{2}x^2] .$$

Es sollte somit

$$\ln(|y|) = \frac{1}{2}x^2 + c$$

sein für ein $c \in \mathbb{R}$. Also $|y| = \exp(\frac{1}{2}x^2 + c)$.

Also $y = \exp(\frac{1}{2}x^2) \cdot \exp(c)$ oder $y = \exp(\frac{1}{2}x^2) \cdot (-\exp(c))$.

Nun können $\exp(c)$ oder $-\exp(c)$ jeden Wert $d \in \mathbb{R} \setminus \{0\}$ annehmen. Somit wird

$$y = y(x) = d \cdot \exp(\frac{1}{2}x^2)$$

die gefundene Lösung unserer Differentialgleichung, bevor wir die Anfangsbedingung beachten.

Wir stellen durch direkte Betrachtung fest, daß dies auch noch für $d = 0$ eine Lösung darstellt, ohne daß dies aus unserer Methode resultiert wäre.

Nun muß noch die Anfangsbedingung erfüllt werden durch geeignete Wahl von d . Es sollte

$$-3 \stackrel{!}{=} y(1) = d \cdot \exp(\frac{1}{2}1^2) = d e^{1/2}$$

sein. Es folgt $d = -3e^{-1/2}$. Also wird

$$y = y(x) = -3e^{-1/2} \exp(\frac{1}{2}x^2) = -3 \exp(\frac{1}{2}(x^2 - 1)) ,$$

definiert auf \mathbb{R} .

Wir machen noch die Probe. Zum einen ist $y'(x) = -3 \exp(\frac{1}{2}(x^2 - 1)) \cdot x = x \cdot y(x)$, also ist die Differentialgleichung erfüllt. Zum anderen ist $y(1) = -3 \exp(\frac{1}{2}(1^2 - 1)) = -3$, also ist die Anfangsbedingung erfüllt.

Ohne Berücksichtigung des Betrags in der Stammfunktion $\ln(|y|)$ hätten wir übrigens keine Lösung für unsere Anfangsbedingung $y(1) = -3$ gefunden.

Beispiel. Zu lösen sei die Differentialgleichung

$$y' = \frac{x^2}{y}.$$

Wir suchen eine Lösung, die die Anfangsbedingung $y(0) = -1$ erfüllt.

In obiger Bezeichnung sollte $g'(y) = y$ und $f'(x) = x^2$ sein.

Dazu sei $g(y) := \frac{1}{2}y^2$ und $f(x) := \frac{1}{3}x^3$.

Zu erfüllen ist

$$[\frac{1}{2}y^2] = \int y \, dy = \int yy' \, dx \stackrel{!}{=} \int x^2 \, dx = [\frac{1}{3}x^3].$$

Es sollte somit

$$\frac{1}{2}y^2 = \frac{1}{3}x^3 + c$$

sein für ein $c \in \mathbb{R}$. Also

$$y = y(x) = \sqrt{\frac{2}{3}x^3 + 2c}.$$

oder

$$y = y(x) = -\sqrt{\frac{2}{3}x^3 + 2c}.$$

Je nach gefundenem c sollte der Definitionsbereich D für x so gewählt werden, daß wir $\frac{2}{3}x^3 + 2c \geq 0$ haben für $x \in D$. Dies ist die gefundene Lösung unserer Differentialgleichung, bevor wir die Anfangsbedingung beachten.

Nun muß noch die Anfangsbedingung erfüllt werden durch geeignete Wahl von c und geeignete Wahl des Vorzeichens.

Es sollte

$$-1 \stackrel{!}{=} y(0) = -\sqrt{\frac{2}{3}0^3 + 2c} = -\sqrt{2c}$$

sein, wobei wir auf der rechten Seite das negative Vorzeichen wählen mußten.

Es folgt $c = 1/2$. Also wird

$$y = y(x) = -\sqrt{\frac{2}{3}x^3 + 1},$$

definiert auf der offenen Teilmenge $D :=]-\sqrt[3]{\frac{3}{2}}, +\infty[\subseteq \mathbb{R}$.

Wir machen noch die Probe. Zum einen ist $y'(x) = -\frac{1}{2}(\frac{2}{3}x^3 + 1)^{-1/2} \cdot 2x^2 = \frac{x^2}{-\sqrt{\frac{2}{3}x^3 + 1}} = \frac{x^2}{y(x)}$,

also ist die Differentialgleichung erfüllt. Zum anderen ist $y(0) = -\sqrt{\frac{2}{3}0^3 + 1} = -1$, also ist die Anfangsbedingung erfüllt.

6.2.3 Ähnlichkeits-Differentialgleichungen

Sei $f(x)$ eine reellwertige Funktion auf einem geeigneten Definitionsbereich für die Variable x .

Zu lösen sei die *Ähnlichkeits-Differentialgleichung*

$$y' = f\left(\frac{y}{x}\right).$$

Wir substituieren $u(x) := \frac{y(x)}{x}$, also $y(x) = u(x) \cdot x$.

Dann wird $y'(x) = u'(x) \cdot x + u(x)$.

Unsere Differentialgleichung wird zu

$$u'(x) \cdot x + u(x) = f(u(x)).$$

Umgeschrieben und abgekürzt wird dies zu

$$u' = (f(u) - u) \cdot x^{-1}.$$

Das ist separierbar. Also

$$\frac{u'}{f(u) - u} = x^{-1}.$$

Zu erfüllen wird

$$[g(u)] := \int \frac{1}{f(u) - u} du = \int \frac{u'}{f(u) - u} dx = \int x^{-1} dx = [\ln(|x|)].$$

Also

$$g(u) = \ln(|x|) + c$$

für ein $c \in \mathbb{R}$. Dies muß nun nach u aufgelöst werden. Sodann führe man die Rücksubstitution $y = u \cdot x$ durch.

Beispiel. Wir suchen eine Lösung der Differentialgleichung

$$y' = 1 + \frac{y^2}{x^2}$$

zur Anfangsbedingung $y(1) = \frac{1+\sqrt{3}}{2}$.

Wir substituieren $y = u \cdot x$ und also $y' = u' \cdot x + u$ und erhalten

$$u' \cdot x + u = 1 + u^2,$$

umgeschrieben zu

$$u' = (u^2 - u + 1) \cdot x^{-1},$$

umgeschrieben zu

$$\frac{u'}{u^2 - u + 1} = x^{-1}.$$

Zu erfüllen wird

$$\int \frac{1}{u^2 - u + 1} du = \int \frac{u'}{u^2 - u + 1} dx = \int x^{-1} dx = [\ln(|x|)].$$

Nun könnte man Partialbruchzerlegung auf das Integral der linken Seite loslassen.

Eine passende Substitution hilft aber auch. Wir schreiben dazu

$$u^2 - u + 1 = (u - \frac{1}{2})^2 + \frac{3}{4} = \frac{3}{4} \left(\left(\frac{2}{\sqrt{3}}(u - \frac{1}{2}) \right)^2 + 1 \right).$$

Wir substituieren linear mit $v := \frac{2}{\sqrt{3}}(u - \frac{1}{2})$. Damit wird

$$\begin{aligned} \int \frac{1}{u^2 - u + 1} du &= \frac{\sqrt{3}}{2} \int \frac{1}{\frac{3}{4}(v^2 + 1)} dv \\ &= \frac{2}{\sqrt{3}} \int \frac{1}{v^2 + 1} dv \\ &= \frac{2}{\sqrt{3}} [\arctan(v)] \\ &= \left[\frac{2}{\sqrt{3}} \arctan\left(\frac{2}{\sqrt{3}}(u - \frac{1}{2})\right) \right]. \end{aligned}$$

Somit können wir fortsetzen. Für ein $c \in \mathbb{R}$ sollte

$$\frac{2}{\sqrt{3}} \arctan\left(\frac{2}{\sqrt{3}}(u - \frac{1}{2})\right) = \ln(|x|) + c$$

sein. Wir lösen schrittweise nach u auf. Zuerst

$$\arctan\left(\frac{2}{\sqrt{3}}(u - \frac{1}{2})\right) = \frac{\sqrt{3}}{2} (\ln(|x|) + c),$$

dann

$$\frac{2}{\sqrt{3}}(u - \frac{1}{2}) = \tan\left(\frac{\sqrt{3}}{2} (\ln(|x|) + c)\right),$$

dann

$$u = \frac{\sqrt{3}}{2} \tan\left(\frac{\sqrt{3}}{2} (\ln(|x|) + c)\right) + \frac{1}{2}.$$

Rücksubstitution liefert

$$y = u \cdot x = \left(\frac{\sqrt{3}}{2} \tan\left(\frac{\sqrt{3}}{2} (\ln(|x|) + c)\right) + \frac{1}{2} \right) \cdot x.$$

Die Anfangsbedingung sollte noch erfüllt sein. Es führt

$$\frac{\sqrt{3}+1}{2} \stackrel{!}{=} y(1) = \left(\frac{\sqrt{3}}{2} \tan\left(\frac{\sqrt{3}}{2} (\ln(|1|) + c)\right) + \frac{1}{2} \right) \cdot 1 = \frac{\sqrt{3}}{2} \tan\left(\frac{\sqrt{3}}{2} c\right) + \frac{1}{2}$$

auf $\tan\left(\frac{\sqrt{3}}{2} c\right) = 1$, also auf $\frac{\sqrt{3}}{2} c = \frac{\pi}{4}$ und also auf $c = \frac{\pi}{2\sqrt{3}}$.

Wir erhalten die Lösung

$$y = y(x) = \left(\frac{\sqrt{3}}{2} \tan\left(\frac{\sqrt{3}}{2} (\ln(|x|) + \frac{\pi}{2\sqrt{3}})\right) + \frac{1}{2} \right) \cdot x = \left(\frac{\sqrt{3}}{2} \tan\left(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}\right) + \frac{1}{2} \right) \cdot x$$

Wir wollen noch einen sinnvollen Definitionsbereich finden. Zunächst sollte $x \neq 0$ sein. Ferner sollte nicht $\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4} = \frac{\pi}{2} + k\pi$ sein für ein $k \in \mathbb{Z}$. Also nicht $\ln(|x|) = \frac{2}{\sqrt{3}}(\frac{\pi}{4} + k\pi)$. Also nicht $x = \pm \exp(\frac{2}{\sqrt{3}}(\frac{\pi}{4} + k\pi))$. Somit können wir als Definitionsbereich nehmen:

$$D := \mathbb{R} \setminus (\{0\} \cup \{s \cdot \exp(\frac{2}{\sqrt{3}}(\frac{\pi}{4} + k\pi)) : k \in \mathbb{Z}, s \in \{-1, +1\}\})$$

Soviel Mengenlehre muß offenbar sein.

Probe: Zum einen wird

$$\begin{aligned} y' &= \frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}) + \frac{1}{2} + (\frac{\sqrt{3}}{2}(1 + \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}))^2) \cdot \frac{\sqrt{3}}{2} \cdot \frac{1}{x} \cdot x \\ &= \frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}) + \frac{1}{2} + \frac{3}{4}(1 + \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}))^2 \\ &= \frac{3}{4} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4})^2 + \frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}) + \frac{5}{4}. \end{aligned}$$

Zum anderen wird

$$\begin{aligned} 1 + \frac{y^2}{x^2} &= 1 + x^{-2} \cdot ((\frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}) + \frac{1}{2}) \cdot x)^2 \\ &= 1 + (\frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}) + \frac{1}{2})^2 \\ &= \frac{3}{4} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4})^2 + \frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|x|) + \frac{\pi}{4}) + \frac{5}{4}. \end{aligned}$$

Das ist dasselbe. Also ist die Differentialgleichung erfüllt. Ferner wird

$$y(1) = (\frac{\sqrt{3}}{2} \tan(\frac{\sqrt{3}}{2} \ln(|1|) + \frac{\pi}{4}) + \frac{1}{2}) \cdot 1 = \frac{\sqrt{3}}{2} \tan(\frac{\pi}{4}) + \frac{1}{2} = \frac{\sqrt{3}+1}{2}.$$

Also ist die Anfangsbedingung erfüllt.

6.2.4 Lineare Differentialgleichungen

Seien $a(x)$ und $b(x)$ reellwertige Funktionen auf einem geeigneten Definitionsbereich für die Variable x .

Zu lösen sei die *lineare* Differentialgleichung

$$y' = a(x) \cdot y + b(x).$$

Ist $b(x)$ die Nullfunktion, so ist dies eine *homogene* lineare Differentialgleichung.

Ist $b(x)$ nicht die Nullfunktion, so ist dies eine *inhomogene* lineare Differentialgleichung. Hiervon wollen wir ausgehen.

Als Vorbereitung lösen wir zunächst die zugehörige homogene lineare Differentialgleichung

$$y' = a(x) \cdot y.$$

Diese ist separierbar. Es wird

$$\frac{y'}{y} = a(x).$$

Also

$$[\ln(|y|)] = \int \frac{1}{y} dy = \int \frac{y'}{y} dx = \int a(x) dx =: [f(x)].$$

Also

$$\ln(|y|) = f(x) + c$$

für ein $c \in \mathbb{R}$. Also

$$y = \exp(f(x)) \cdot e^c \quad \text{oder} \quad y = \exp(f(x)) \cdot (-e^c).$$

Da wir auch die konstante Lösung $y = 0$ haben, kann nun e^c oder $-e^c$ oder 0 jeden Wert $d \in \mathbb{R}$ annehmen. Wir erhalten

$$y = \exp(f(x)) \cdot d$$

als Lösung der homogenen linearen Differentialgleichung $y' = a(x) \cdot y$.

Für unsere inhomogene lineare Differentialgleichung $y' = a(x) \cdot y + b(x)$ machen wir nun den Ansatz, daß wir die Konstante d durch eine zu ermittelnde Funktion $d(x)$ ersetzen und diese Funktion $d(x)$ so zu bestimmen versuchen, daß die Differentialgleichung erfüllt ist. Sei also

$$y = \exp(f(x)) \cdot d(x)$$

angesetzt. Man nennt diesen Ansatz auch *Variation der Konstanten*.

Es sollte also sein

$$\begin{aligned} a(x) \cdot y + b(x) &\stackrel{!}{=} y' \\ &= \exp(f(x)) \cdot f'(x) \cdot d(x) + \exp(f(x)) \cdot d'(x) \\ &= a(x) \cdot \exp(f(x)) \cdot d(x) + \exp(f(x)) \cdot d'(x) \\ &= a(x) \cdot y + \exp(f(x)) \cdot d'(x). \end{aligned}$$

Dies führt uns auf die Bedingung

$$b(x) \stackrel{!}{=} \exp(f(x)) \cdot d'(x)$$

und also auf

$$d'(x) \stackrel{!}{=} b(x) \cdot \exp(-f(x)).$$

Es wird

$$[d(x)] = \int b(x) \cdot \exp(-f(x)) dx =: [g(x)].$$

Somit gibt es ein $h \in \mathbb{R}$ mit

$$d(x) = g(x) + h.$$

Einsetzen in unseren Ansatz liefert nun

$$y = \exp(f(x)) \cdot (g(x) + h) .$$

Beispiel. Wir wollen auf $\mathbb{R}_{>0}$ die inhomogene lineare Differentialgleichung

$$y' = yx^{-1} + x$$

lösen mit der Anfangsbedingung $y(1) = t$ für einen Parameter $t \in \mathbb{R}$.

Dazu lösen wir zunächst die zugehörige homogene lineare Differentialgleichung

$$y' = yx^{-1}$$

Diese ist separierbar. Also

$$\frac{y'}{y} = x^{-1} ,$$

also

$$[\ln(|y|)] = \int \frac{1}{y} dy = \int \frac{y'}{y} dx = \int x^{-1} dx = [\ln(x)] ,$$

also

$$\ln(|y|) = \ln(x) + c$$

für ein $c \in \mathbb{R}$. Es wird

$$y = \exp(\ln(x)) \cdot e^c \quad \text{oder} \quad y = \exp(\ln(x)) \cdot (-e^c) \quad \text{oder} \quad y = 0 .$$

Also

$$y = x \cdot d$$

für ein $d \in \mathbb{R}$.

Nun zur gegebenen inhomogenen linearen Differentialgleichung. Wir machen den Ansatz der Variation der Konstanten.

$$y = x \cdot d(x) .$$

Es soll also sein :

$$yx^{-1} + x \stackrel{!}{=} y' = d(x) + x \cdot d'(x) = yx^{-1} + x \cdot d'(x) .$$

Also

$$d'(x) = 1 .$$

Also

$$d(x) = x + h$$

für ein $h \in \mathbb{R}$. Also

$$y = x(x + h) .$$

Die Anfangsbedingung verlangt

$$t = y(1) = 1 + h .$$

Somit wird $h = t - 1$. Wir erhalten die Lösung

$$y = y(x) = x(x + t - 1) = x^2 + (t - 1)x .$$

Probe: Es ist $y' = 2x + t - 1$. Es ist $yx^{-1} + x = x + t - 1 + x = 2x + t - 1$. Das ist dasselbe. Also ist die Differentialgleichung erfüllt.

Ferner wird $y(1) = 1 \cdot (1 + t - 1) = t$. Also ist die Anfangsbedingung erfüllt.

6.2.5 Existenz und Eindeutigkeit einer Lösung

Lemma. Sei $(\hat{x}, \hat{y}) \in \mathbb{R}^2$. Seien $a, b > 0$. Wir betrachten das Rechteck

$$V := [\hat{x} - a, \hat{x} + a] \times [\hat{y} - b, \hat{y} + b] = \{ (x, y) : |x - \hat{x}| \leq a \text{ und } |y - \hat{y}| \leq b \} \subseteq \mathbb{R}^2 ,$$

welches Mittelpunkt (\hat{x}, \hat{y}) und Kantenlängen $2a$ und $2b$ hat.

Sei $f : V \rightarrow \mathbb{R} : (x, y) \mapsto f(x, y)$ stetig.

Die *Lipschitzbedingung* sei erfüllt: Es gebe ein $L \in \mathbb{R}_{\geq 0}$ mit

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

für alle $(x, y_1), (x, y_2) \in V$.

Sei $m := \max\{|f(x, y)| : (x, y) \in V\}$. Sei $h := \min\{a, bm^{-1}\}$. Sei $D :=]\hat{x} - h, \hat{x} + h[$.

Dann gibt es genau eine differenzierbare Funktion $y : D \rightarrow \mathbb{R}$, welche die Differentialgleichung

$$y'(x) = f(x, y(x))$$

für $x \in D$ und die Anfangsbedingung $y(\hat{x}) = \hat{y}$ erfüllt.

Die technischen Bedingungen stellen also sicher, daß man durch das Richtungsfeld dieser Differentialgleichung ausgehend vom Punkt (\hat{x}, \hat{y}) genau eine Lösung zeichnen kann, wenn man sich dabei nicht weiter als um h links und rechts von dieser Stelle entfernt.

Bemerkung. Ist $f_y(x, y)$ im Innern von V existent und ist dort $|f_y(x, y)| \leq L$, dann ist die Lipschitzbedingung mit der Konstanten L erfüllt.

Gegenbeispiel (bei nicht ganz erfüllten Voraussetzungen).

Sei die Differentialgleichung $y' = 2|y|^{1/2}$ zu lösen mit der Anfangsbedingung $y(0) = 0$.

Für jedes $c \in \mathbb{R}_{\geq 0}$ ist

$$y(x) := \begin{cases} x^2 & \text{für } x \in]0, +\infty[\\ 0 & \text{für } x \in [-c, 0] \\ -(x + c)^2 & \text{für } x \in]-\infty, -c[\end{cases}$$

eine Lösung, mit $y(0) = 0$.

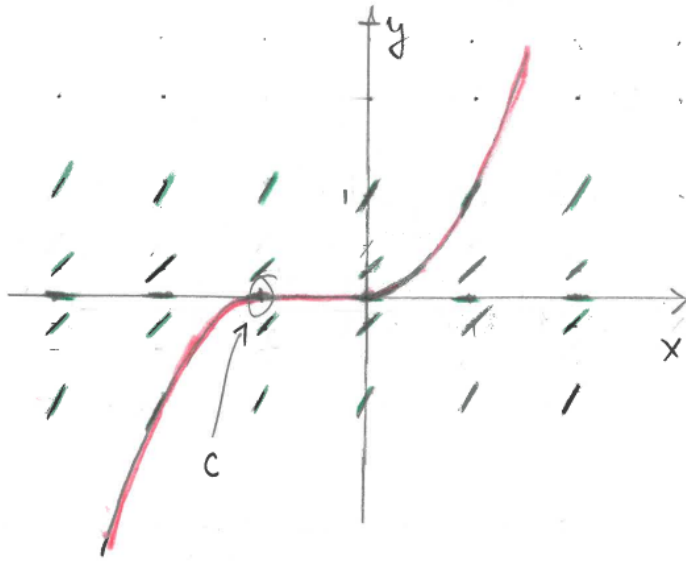
Denn für $x > 0$ ist $y'(x) = 2x = 2|x^2|^{1/2} = 2|y(x)|^{1/2}$.

Für $x \in [-c, 0]$ ist $y'(x) = 0 = 2|0|^{1/2} = 2|y(x)|^{1/2}$.

Für $x < -c$ ist $y'(x) = -2(x+c) = 2|x+c| = 2|-(x+c)^2|^{1/2} = 2|y(x)|^{1/2}$.

Da $c \geq 0$ beliebig wählbar ist, ist hier die Eindeutigkeit der Lösung der Differentialgleichung unter der gestellten Anfangsbedingung nicht erfüllt.

Aber hier ist auch die Lipschitzbedingung nicht erfüllt: $\frac{|f(x,y_1)-f(x,0)|}{|y_1-0|} = \frac{|2|y_1|^{1/2}-0|}{|y_1-0|} = 2|y_1|^{-1/2}$ geht für $y_1 \rightarrow 0$ gegen $+\infty$, bleibt also nicht $\leq L$, egal welches L man wählt.



Hier bleibt es dem Zeichner der Lösung im Richtungsfeld überlassen, wie lange er auf der Geraden $y = 0$ bleiben will.

Lemma. Sei $\hat{x} \in D \subseteq \mathbb{R}$, wobei D ein offenes Intervall sei.

Sei $f : D \times \mathbb{R} \rightarrow \mathbb{R} : (x, y) \mapsto f(x, y)$ stetig.

Die Lipschitzbedingung sei erfüllt: Es gebe ein $L \in \mathbb{R}_{\geq 0}$ mit

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

für alle $(x, y_1), (x, y_2) \in D \times \mathbb{R}$.

Erfülle $y : D \rightarrow \mathbb{R} : x \mapsto y(x)$ die Differentialgleichung $y'(x) = f(x, y(x))$ für $x \in D$.

Erfülle $\tilde{y} : D \rightarrow \mathbb{R} : x \mapsto \tilde{y}(x)$ die Differentialgleichung $\tilde{y}'(x) = f(x, \tilde{y}(x))$ für $x \in D$.

Dann ist

$$|y(x) - \tilde{y}(x)| \leq |y(\hat{x}) - \tilde{y}(\hat{x})| e^{L|x-\hat{x}|}$$

für $x \in D$.

Wieder gilt: Ist f_y existent und $|f_y(x, y)| \leq L$ für $(x, y) \in D \times \mathbb{R}$, dann ist die Lipschitzbedingung mit der Konstanten L erfüllt.

Hat man zwei Lösungen y und \tilde{y} der Differentialgleichung, so kann also der Abstand $|y(\hat{x}) - \tilde{y}(\hat{x})|$ der Anfangswerte zusammen mit L für eine Abschätzung des Abstands $|y(x) - \tilde{y}(x)|$ der Funktionswerte verwendet werden.

Beispiel. Sei $\hat{x} = 0$. Sei $D =]-1, +1[$. Sei $f(x, y) = 2xy$.

Sei $y' = 2xy$ auf D zu lösen.

Dazu rechnen wir: $\int \frac{1}{y} dy = \int \frac{y'}{y} dx = \int 2x dx$, also $[\ln(|y|)] = [x^2]$, also

$$y = e^{x^2} \cdot d$$

für ein $d \in \mathbb{R}$.

Es ist $f_y(x, y) = 2x$. Es ist also $|f_y(x, y)| = 2|x| \leq 2 =: L$ für $(x, y) \in D \times \mathbb{R}$.

Für zwei Lösungen y und \tilde{y} gilt nach vorstehendem Lemma also

$$|y(x) - \tilde{y}(x)| \leq |y(0) - \tilde{y}(0)| e^{2|x-0|}$$

für $x \in D$.

Überprüfen wir dies. Sei $y(x) = e^{x^2}$ und $\tilde{y}(x) = 2e^{x^2}$ für $x \in]-1, +1[$.

Dann ist

$$|y(x) - \tilde{y}(x)| = e^{x^2}$$

und

$$|y(0) - \tilde{y}(0)| e^{2|x-0|} = e^{2|x|}.$$

Und tatsächlich ist $x^2 \leq 2|x|$ und somit auch $e^{x^2} \leq e^{2|x|}$ für $x \in]-1, +1[$.

6.3 Lineare Differentialgleichungssysteme erster Ordnung

6.3.1 Der allgemeine Fall

Wir erweitern nun die Fragestellung und wollen nach vektorwertigen Funktionen suchen, die eine Differentialgleichung erfüllen.

Weiter unten werden wir die Suche nach einer skalarwertigen Funktion, die eine Differentialgleichung n -ter Ordnung erfüllen soll, zurückführen auf die Suche nach einer vektorwertigen Funktion, die eine Differentialgleichung erster Ordnung erfüllen soll.

Sei $n \geq 1$.

Definition. Sei $D \subseteq \mathbb{R}$ offen.

Sei $f : D \rightarrow \mathbb{C}^{n \times 1} : x \mapsto f(x) = (f_j(x))_j = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$ eine vektorwertige Funktion.

Es ist $f'(x) := \begin{pmatrix} f'_1(x) \\ \vdots \\ f'_n(x) \end{pmatrix}$, sofern existent. Diesemfalls heißt f *differenzierbar*.

Beispiel. Ist $f(x) = \begin{pmatrix} x^2 \\ \sin(x) \end{pmatrix}$, dann ist $f'(x) = \begin{pmatrix} 2x \\ \cos(x) \end{pmatrix}$.

Definition. Sei $D \subseteq \mathbb{R}$ offen.

Sei $A : D \rightarrow \mathbb{R}^{n \times n} : x \mapsto A(x) = (a_{j,k}(x))_{j,k}$.

Sei $g : D \rightarrow \mathbb{R}^{n \times 1} : x \mapsto g(x) = (g_j(x))_j$.

Ein *lineares Differentialgleichungssystem* erster Ordnung ist eine Gleichung der Form

$$y'(x) = A(x) \cdot y(x) + g(x)$$

Wir schreiben dies auch kurz

$$y' = A(x) \cdot y + g(x).$$

Ist $g(x) = 0$ für $x \in D$, so heißt dieses *homogen*, ansonsten *inhomogen*.

Eine *Lösung* dieses Differentialgleichungssystems auf D ist eine differenzierbare Funktion

$$y : D \rightarrow \mathbb{R}^{n \times 1} : x \mapsto y(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix},$$

für welche eben $y'(x) = A(x) \cdot y(x) + g(x)$ gilt für $x \in D$.

Fordert man zudem für ein $\hat{x} \in D$ noch

$$y(\hat{x}) = \hat{y}$$

für einen gewissen Vektor $\hat{y} \in \mathbb{R}^{n \times 1}$, so spricht man von einer *Anfangsbedingung* an y .

Bemerkung. Es ist die Gleichung von Vektoren

$$y'(x) = A(x) \cdot y(x) + g(x)$$

äquivalent zum Gleichungssystem

$$\begin{aligned} y'_1(x) &= a_{1,1}(x)y_1(x) + \dots + a_{1,n}(x)y_n(x) + g_1(x) \\ y'_2(x) &= a_{2,1}(x)y_1(x) + \dots + a_{2,n}(x)y_n(x) + g_2(x) \\ &\vdots \\ y'_n(x) &= a_{n,1}(x)y_1(x) + \dots + a_{n,n}(x)y_n(x) + g_n(x). \end{aligned}$$

Kurz,

$$\begin{aligned} y_1' &= a_{1,1}(x)y_1 + \dots + a_{1,n}(x)y_n + g_1(x) \\ y_2' &= a_{2,1}(x)y_1 + \dots + a_{2,n}(x)y_n + g_2(x) \\ &\vdots \\ y_n' &= a_{n,1}(x)y_1 + \dots + a_{n,n}(x)y_n + g_n(x). \end{aligned}$$

Beispiel. Sei $D = \mathbb{R}_{>-1}$.

Auf D suchen wir eine Lösung des homogenen linearen Differentialgleichungssystems

$$y' = \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix} y$$

zur Anfangsbedingung $y(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Zu lösen ist also

$$\begin{aligned} y_1' &= y_1 + xy_2 \\ y_2' &= (x+1)^{-1}y_2 \end{aligned}$$

zur Anfangsbedingung $\begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Die zweite Gleichung gibt $\int \frac{1}{y_2} dy_2 = \int \frac{y_2'}{y_2} dx = \int (x+1)^{-1} dx$, also $[\ln(|y_2|)] = [\ln(|x+1|)]$. Hier ist $|x+1| = x+1$, da $x > -1$. Es wird

$$y_2 = y_2(x) = c \cdot e^{\ln(x+1)} = c(x+1),$$

wobei $c \in \mathbb{R}$.

Die erste Gleichung gibt nun $y_1' = y_1 + x \cdot c(x+1)$.

Die allgemeine Lösung zu $y_1' = y_1$ ist $y_1 = y_1(x) = d \cdot e^x$, wobei $d \in \mathbb{R}$. Wir setzen an mit einer Variation der Konstanten, also $y_1(x) = d(x) \cdot e^x$. Es ist $y_1'(x) = d'(x) \cdot e^x + d(x) \cdot e^x$.

Es muß also $d'(x) \cdot e^x + d(x) \cdot e^x \stackrel{!}{=} d(x) \cdot e^x + cx(x+1)$ sein. Es muß also

$$d'(x) \stackrel{!}{=} cx(x+1)e^{-x} = c(x^2+x)e^{-x}$$

sein. Wir integrieren:

$$\begin{aligned} \int (x^2+x)e^{-x} dx &= [(x^2+x)(-e^{-x})] - \int (2x+1)(-e^{-x}) dx \\ &= [-(x^2+x)e^{-x}] + \int (2x+1)e^{-x} dx \\ &= [-(x^2+x)e^{-x}] + [(2x+1)(-e^{-x})] - \int 2(-e^{-x}) dx \\ &= [-(x^2+x)e^{-x} - (2x+1)e^{-x} - 2e^{-x}] \\ &= [-(x^2+3x+3)e^{-x}] \end{aligned}$$

Also

$$d(x) = -c(x^2+3x+3)e^{-x} + s,$$

mithin

$$y_1 = -c(x^2 + 3x + 3) + s e^x$$

wobei $s \in \mathbb{R}$. Somit löst

$$y = y(x) = \begin{pmatrix} -c(x^2+3x+3)+s e^x \\ c(x+1) \end{pmatrix} = c \cdot \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s \cdot \begin{pmatrix} e^x \\ 0 \end{pmatrix}$$

unser Differentialgleichungssystem. Die Anfangsbedingung verlangt

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} \stackrel{!}{=} y(0) = \begin{pmatrix} -3c+s \\ c \end{pmatrix}.$$

Es wird $c = 1$ und $s = 2$. Somit erhalten wir die Lösung

$$y = y(x) = \begin{pmatrix} -(x^2+3x+3)+2e^x \\ x+1 \end{pmatrix}$$

unseres Differentialgleichungssystems zur Anfangsbedingung $y(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Probe: Es ist $y'(x) = \begin{pmatrix} -(2x+3)+2e^x \\ 1 \end{pmatrix}$. Es ist

$$A(x) \cdot y(x) = \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix} \cdot \begin{pmatrix} -(x^2+3x+3)+2e^x \\ x+1 \end{pmatrix} = \begin{pmatrix} -(x^2+3x+3)+2e^x + x \cdot (x+1) \\ 1 \end{pmatrix} = \begin{pmatrix} -(2x+3)+2e^x \\ 1 \end{pmatrix}.$$

Das ist dasselbe. Also ist das Differentialgleichungssystem erfüllt.

Ferner ist $y(0) = \begin{pmatrix} -(0^2+3 \cdot 0+3)+2e^0 \\ 0+1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Also ist die Anfangsbedingung erfüllt.

Definition. Sei $D \subseteq \mathbb{R}$.

Für Funktionen $y, \tilde{y} : D \rightarrow \mathbb{R}^{n \times 1}$ und $\lambda \in \mathbb{R}$ sei

$$\begin{aligned} y + \tilde{y} &: D \rightarrow \mathbb{R}^{n \times 1} : x \mapsto (y + \tilde{y})(x) := y(x) + \tilde{y}(x) \\ \lambda \cdot y &: D \rightarrow \mathbb{R}^{n \times 1} : x \mapsto (\lambda \cdot y)(x) := \lambda \cdot y(x) \end{aligned}$$

Damit wird die Menge $(\mathbb{R}^{n \times 1})^D$ der Funktionen von D nach $\mathbb{R}^{n \times 1}$ zu einem \mathbb{R} -Vektorraum.

Lemma. Sei $D \subseteq \mathbb{R}$ ein offenes Intervall.

Sei $A : D \rightarrow \mathbb{R}^{n \times n} : x \mapsto A(x) = (a_{j,k}(x))_{j,k}$ stetig.

Sei $g : D \rightarrow \mathbb{R}^{n \times 1} : x \mapsto g(x) = (g_j(x))_j$ stetig.

Sei $\hat{x} \in D$. Sei $\hat{y} \in \mathbb{R}^{n \times 1}$.

(1) Es gibt es genau eine differenzierbare Funktion $y : D \rightarrow \mathbb{R}^{n \times 1}$ mit

$$y'(x) = A(x) \cdot y(x) + g(x) \quad \text{für } x \in D,$$

für welche $y(\hat{x}) = \hat{y}$ ist.

- (2) Wir betrachten die Menge $L_{A,0}$ der differenzierbaren Funktionen $y : D \rightarrow \mathbb{R}^{n \times 1}$ mit

$$y'(x) = A(x) \cdot y(x) \quad \text{für } x \in D .$$

Es ist $L_{A,0} \subseteq (\mathbb{R}^{n \times 1})^D$ ein Unterraum, auch *Lösungsraum* von $y' = A(x) \cdot y$ genannt. Die Auswertungsabbildung $L_{A,0} \rightarrow \mathbb{R}^{n \times 1} : y \mapsto y(\hat{x})$ ist \mathbb{R} -linear. Dank (1) ist sie bijektiv.

Insbesondere ist $\dim_{\mathbb{R}}(L_{A,0}) = n$.

Eine \mathbb{R} -lineare Basis von $L_{A,0}$ heißt auch *Fundamentalsystem* für $y' = A(x) \cdot y$.

- (3) Wir betrachten die Menge $L_{A,g}$ der differenzierbaren Funktionen $y : D \rightarrow \mathbb{R}^{n \times 1}$ mit

$$y'(x) = A(x) \cdot y(x) + g(x) \quad \text{für } x \in D ,$$

auch *Lösungsmenge* von $y' = A(x) \cdot y + g(x)$ genannt.

Sei $y_{[0]} : D \rightarrow \mathbb{R}^{n \times 1}$ ein Element von $L_{A,g}$. Wir nennen $y_{[0]}$ auch eine *partikuläre Lösung* von $y' = A(x) \cdot y + g(x)$.

Dann ist

$$L_{A,g} = \{ y_{[0]} + y : y \in L_{A,0} \} = y_{[0]} + L_{A,0} .$$

- (4) Sei $(y_{[1]}, \dots, y_{[n]})$ ein Tupel von Lösungen von $y' = A(x) \cdot y$ auf D .

Für $x \in D$ sei $F(x) \in \mathbb{R}^{n \times n}$ die Matrix mit Spaltentupel $(y_{[1]}(x), \dots, y_{[n]}(x))$.

Dann heißt $w(x) := \det(F(x))$ die *Wronski-Determinante* dieses Tupels.

Die folgenden Aussagen (a, b, c) sind äquivalent.

- (a) Es ist $w(x) \neq 0$ für ein $x \in D$.
- (b) Es ist $w(x) \neq 0$ für alle $x \in D$.
- (c) Es ist $(y_{[1]}, \dots, y_{[n]})$ ein Fundamentalsystem für $y' = A(x) \cdot y$.

Insbesondere ist $w(\hat{x}) \neq 0$ eine notwendige und hinreichende Bedingung dafür, daß ein Fundamentalsystem vorliegt.

- (5) Sei $(y_{[1]}, \dots, y_{[n]})$ ein Fundamentalsystem für $y' = A(x) \cdot y$ auf D .

Jede Lösung von $y' = A(x) \cdot y$ ist also von der Form

$$y(x) = c_1 y_{[1]}(x) + \dots + c_n y_{[n]}(x) ,$$

mit $c_1, \dots, c_n \in \mathbb{R}$.

Wir machen den Ansatz der *Variation der Konstanten*, um eine partikuläre Lösung für $y' = A(x) \cdot y + g(x)$ zu erhalten.

Sei dazu

$$y(x) = c_1(x) y_{[1]}(x) + \dots + c_n(x) y_{[n]}(x)$$

angesetzt, mit gesuchten differenzierbaren Funktionen $c_1(x), \dots, c_n(x)$ auf D mit Werten in \mathbb{R} .

Dann wird

$$\begin{aligned} y'(x) &= (c_1(x)y'_{[1]}(x) + \dots + c_n(x)y'_{[n]}(x)) + (c'_1(x)y_{[1]}(x) + \dots + c'_n(x)y_{[n]}(x)) \\ \stackrel{!}{=} A(x) \cdot y(x) + g(x) &= A(x) \cdot (c_1(x)y_{[1]}(x) + \dots + c_n(x)y_{[n]}(x)) + g(x), \end{aligned}$$

also

$$c'_1(x)y_{[1]}(x) + \dots + c'_n(x)y_{[n]}(x) \stackrel{!}{=} g(x)$$

zu lösen. Eine einzelne Lösung genügt, um damit die partikuläre Lösung

$$y_{[0]}(x) = c_1(x)y_{[1]}(x) + \dots + c_n(x)y_{[n]}(x)$$

zu bilden.

Beispiel. Wir greifen das vorige Beispiel wieder auf. Es ist $D := \mathbb{R}_{>-1}$.

Wir haben gesehen, daß das homogene lineare Differentialgleichungssystem

$$y' = \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix} y$$

gelöst wird von

$$y = y(x) = \begin{pmatrix} -c(x^2+3x+3)+se^x \\ c(x+1) \end{pmatrix},$$

wobei $c, s \in \mathbb{R}$. Mit anderen Worten, mit $A(x) := \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix}$, definiert auf D , wird

$$L_{A,0} = \left\{ \begin{pmatrix} -c(x^2+3x+3)+se^x \\ c(x+1) \end{pmatrix} : s, c \in \mathbb{R} \right\} = \left\{ c \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s \begin{pmatrix} e^x \\ 0 \end{pmatrix} : c, s \in \mathbb{R} \right\},$$

mit den Funktionen darin definiert auf D .

Für jedes $\hat{y} \in \mathbb{R}^{2 \times 1}$ gibt es dank Theorie genau ein $y \in L_{A,0}$ mit $y(0) = \hat{y}$, da die Auswertungsabbildung $L_{A,0} \rightarrow \mathbb{R}^{2 \times 1} : y \mapsto y(0)$ bijektiv ist.

Wir betrachten das Tupel von Lösungen unseres Differentialgleichungssystems

$$\left(\underbrace{\begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix}}_{y_{[1]}(x)}, \underbrace{\begin{pmatrix} e^x \\ 0 \end{pmatrix}}_{y_{[2]}(x)} \right),$$

mit den Funktionen darin definiert auf D . Seine Wronski-Determinante ist

$$w(x) = \det \begin{pmatrix} -(x^2+3x+3) & e^x \\ x+1 & 0 \end{pmatrix} = -e^x(x+1).$$

Es ist $w(0) = -1 \neq 0$. Also ist

$$(y_{[1]}(x), y_{[2]}(x)) = \left(\begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix}, \begin{pmatrix} e^x \\ 0 \end{pmatrix} \right)$$

ein Fundamentalsystem für unser Differentialgleichungssystem, d.h. eine \mathbb{R} -lineare Basis von $L_{A,0}$.

Dies zeigt nun auch durch direkte Betrachtung, daß die Auswertungsabbildung

$$\begin{aligned} L_{A,0} &\rightarrow \mathbb{R}^{2 \times 1} \\ y &\mapsto y(0) \\ c \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s \begin{pmatrix} e^x \\ 0 \end{pmatrix} &\mapsto c \begin{pmatrix} -3 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

bijektiv ist.

Beispiel. Wir greifen das vorige Beispiel wieder auf. Es ist $D := \mathbb{R}_{>-1}$.

Wir wollen das inhomogene lineare Differentialgleichungssystem

$$y' = \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix} y + \begin{pmatrix} -x^2 \\ x+1 \end{pmatrix}$$

lösen, mit $A(x) = \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix}$ und $g(x) := \begin{pmatrix} -x^2 \\ x+1 \end{pmatrix}$.

Aus dem vorigen Beispiel ist uns $L_{A,0}$ bereits bekannt. Wir müssen also noch eine partikuläre Lösung finden.

Wir machen den Ansatz der Variation der Konstanten: Wir wollen für

$$y(x) = \underbrace{c(x) \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix}}_{y_{[1]}(x)} + \underbrace{s(x) \begin{pmatrix} e^x \\ 0 \end{pmatrix}}_{y_{[2]}(x)}$$

geeignete Funktionen $c(x)$ und $s(x)$ auf D bestimmen. Gelten soll

$$\begin{aligned} y'(x) &= (c(x) \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix})' + s(x) \begin{pmatrix} e^x \\ 0 \end{pmatrix}' + (c'(x) \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s'(x) \begin{pmatrix} e^x \\ 0 \end{pmatrix}) \\ \stackrel{!}{=} A(x) \cdot y(x) + g(x) &= (c(x)A(x) \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s(x)A(x) \begin{pmatrix} e^x \\ 0 \end{pmatrix}) + \begin{pmatrix} -x^2 \\ x+1 \end{pmatrix}. \end{aligned}$$

Unter Beachtung dessen, daß wir zwei Lösungen des zugehörigen homogenen Differentialgleichungssystems verwendet haben, vereinfacht sich dies zu

$$c'(x) \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s'(x) \begin{pmatrix} e^x \\ 0 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} -x^2 \\ x+1 \end{pmatrix}.$$

Es sollte also $c'(x) = 1$ sein. Wir können $c(x) = x$ wählen.

Damit sollte dann $-(x^2 + 3x + 3) + s'(x)e^x = -x^2$ sein, also $s'(x) = (3x + 3)e^{-x}$.

Wir integrieren: $\int (3x + 3)e^{-x} dx = [(3x + 3)(-e^{-x})] - \int 3(-e^{-x}) dx = [(-3x - 6)e^{-x}]$.

Wir können also $s(x) = (-3x - 6)e^{-x}$ wählen.

Eine partikuläre Lösung ist mithin

$$\begin{aligned} y_{[0]}(x) &= c(x) \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s(x) \begin{pmatrix} e^x \\ 0 \end{pmatrix} \\ &= x \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + (-3x-6)e^{-x} \begin{pmatrix} e^x \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -x^3-3x^2-6x-6 \\ x(x+1) \end{pmatrix}. \end{aligned}$$

Somit erhalten wir die Lösungsmenge

$$\begin{aligned} L_{A,g} &= \begin{pmatrix} -x^3-3x^2-6x-6 \\ x(x+1) \end{pmatrix} + L_{A,0} \\ &= \left\{ \begin{pmatrix} -x^3-3x^2-6x-6 \\ x(x+1) \end{pmatrix} + c \begin{pmatrix} -(x^2+3x+3) \\ x+1 \end{pmatrix} + s \begin{pmatrix} e^x \\ 0 \end{pmatrix} : c, s \in \mathbb{R} \right\}. \end{aligned}$$

Probe: Wir müssen für unsere partikuläre Lösung noch die Probe machen.

Zum einen wird

$$y'_{[0]}(x) = \begin{pmatrix} -x^3-3x^2-6x-6 \\ x(x+1) \end{pmatrix}' = \begin{pmatrix} -3x^2-6x-6 \\ 2x+1 \end{pmatrix}.$$

Zum anderen wird

$$\begin{aligned} A(x) \cdot y_{[0]}(x) + g(x) &= \begin{pmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{pmatrix} \begin{pmatrix} -x^3-3x^2-6x-6 \\ x(x+1) \end{pmatrix} + \begin{pmatrix} -x^2 \\ x+1 \end{pmatrix} \\ &= \begin{pmatrix} -x^3-3x^2-6x-6+x^2(x+1) \\ x \end{pmatrix} + \begin{pmatrix} -x^2 \\ x+1 \end{pmatrix} \\ &= \begin{pmatrix} -3x^2-6x-6 \\ 2x+1 \end{pmatrix}. \end{aligned}$$

Das ist dasselbe.

6.3.2 Der Fall konstanter Koeffizienten

Sei $n \geq 1$.

Wir wollen ein Differentialgleichungssystem der Form $y' = Ay + g(x)$ betrachten, mit einer konstanten Matrix $A \in \mathbb{R}^{n \times n}$, die also unabhängig ist von x . Hierzu brauchen wir eine Erweiterung des Definitionsbereichs der Exponentialfunktion auf Matrizen. Es ist nützlich, hierbei in $\mathbb{C}^{n \times n}$ zu arbeiten, um auch den Fall $A \in \mathbb{R}^{n \times n} \subseteq \mathbb{C}^{n \times n}$ mit Eigenwerten in \mathbb{C} , die nicht alle in \mathbb{R} liegen, gut behandeln zu können.

6.3.2.1 Matrixexponentialfunktion

Sei $A \in \mathbb{C}^{n \times n}$.

Definition. Sei $(B_\ell)_{\ell \geq 0}$ eine Folge von Matrizen $B_\ell \in \mathbb{C}^{n \times n}$.

Wir schreiben $B_\ell = (b_{\ell;j,k})_{j,k}$.

Wir definieren den Wert der Reihe

$$\sum_{\ell=0}^{\infty} B_\ell := \left(\sum_{\ell=0}^{\infty} b_{\ell;j,k} \right)_{j,k} \in \mathbb{C}^{n \times n},$$

sofern existent. Kurz, es wird eintragsweise summiert.

Definition. Sei die Matrixexponentialfunktion definiert durch

$$e^A = \exp(A) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} A^\ell \in \mathbb{C}^{n \times n}.$$

Dieser Reihenwert existiert in $\mathbb{C}^{n \times n}$ für jedes $A \in \mathbb{C}^{n \times n}$.

Beispiel. Es ist

$$\exp\left(\begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix}\right) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix}^\ell = \begin{pmatrix} \sum_{\ell=0}^{\infty} \frac{1}{\ell!} 2^\ell & 0 \\ 0 & \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (\pi i)^\ell \end{pmatrix} = \begin{pmatrix} \exp(2) & 0 \\ 0 & \exp(\pi i) \end{pmatrix} = \begin{pmatrix} e^2 & 0 \\ 0 & -1 \end{pmatrix}.$$

Bemerkung. Sei $u \in \mathbb{C}^{n \times 1}$ ein konstanter Vektor.

Sei

$$y(x) := e^{Ax} \cdot u,$$

definiert auf \mathbb{R} . Dann ist

$$y'(x) = A \cdot e^{Ax} \cdot u.$$

Beispiel. Wir setzen das vorige Beispiel fort. Es ist

$$e^{\begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix} x} = \exp\left(\begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix} x\right) = \begin{pmatrix} \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (2x)^\ell & 0 \\ 0 & \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (\pi i x)^\ell \end{pmatrix} = \begin{pmatrix} e^{2x} & 0 \\ 0 & e^{\pi i x} \end{pmatrix}.$$

Sei

$$y(x) := \exp\left(\begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix} x\right) \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3e^{2x} \\ e^{\pi i x} \end{pmatrix}.$$

Es ist auf der einen Seite

$$y'(x) = \begin{pmatrix} 3e^{2x} \\ e^{\pi i x} \end{pmatrix}' = \begin{pmatrix} 6e^{2x} \\ \pi i e^{\pi i x} \end{pmatrix}.$$

Es ist auf der anderen Seite

$$\begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix} \cdot \exp\left(\begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix} x\right) \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \pi i \end{pmatrix} \cdot \begin{pmatrix} e^{2x} & 0 \\ 0 & e^{\pi i x} \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 6e^{2x} \\ \pi i e^{\pi i x} \end{pmatrix}.$$

Das ist in der Tat dasselbe.

Bemerkung. Seien $A, B \in \mathbb{C}^{n \times n}$. Sei $S \in \mathbb{C}^{n \times n}$ invertierbar.

- (1) Ist $AB = BA$, dann ist $\exp(A + B) = \exp(A) \cdot \exp(B)$.
 Insbesondere ist $\exp(A) \cdot \exp(-A) = E_n$, also $\exp(-A) = \exp(A)^{-1}$.
- (2) Es ist $\exp(S^{-1}AS) = S^{-1} \exp(A)S$.
- (3) Es ist $\det(\exp(A)) = \exp(\operatorname{tr}(A))$.

Erinnerung. Wir können für die Matrix $A \in \mathbb{C}^{n \times n}$ eine invertierbare Matrix S so finden, daß $S^{-1}AS$ in Jordanscher Normalform ist; vgl. §3.7.4.

Wir gehen dazu wie folgt vor.

- (1) Man bestimme das charakteristische Polynom

$$\chi_A(X) = \det(A - XE_n) = (-1)^n \cdot (X - \lambda_1)^{a_{V_A}(\lambda_1)} \cdot \dots \cdot (X - \lambda_k)^{a_{V_A}(\lambda_k)}.$$

Es sind $\lambda_1, \dots, \lambda_k$ die Eigenwerte von A . Es ist $a_{V_A}(\lambda_j)$ die algebraische Vielfachheit von λ_j .

- (2) Für jedes $1 \leq j \leq k$ bestimmen wir wie folgt eine geeignete Basis des Hauptraums $H_A(\lambda_j)$.

Wir kürzen $A_{(j)} := A - \lambda_j E_n$ ab.

Schritt 1.

Man bestimme eine Basis des Eigenraums $E_A(\lambda_j) = \operatorname{Kern}(A_{(j)})$.

Man ergänze diese zu einer Basis von $\operatorname{Kern}(A_{(j)}^2)$.

Man ergänze diese zu einer Basis von $\operatorname{Kern}(A_{(j)}^3)$.

Usf.

Dies beende man, sobald man $a_{V_A}(\lambda_j) = \dim(H_A(\lambda_j))$ Basisvektoren vorliegen hat. Wir haben so eine schrittweise ergänzte Basis von $H_A(\lambda_j)$ erhalten.

Noch zwei Anmerkungen zur praktischen Durchführung.

Bei der obigen Berechnung müssen zur Bestimmung der Basen Zeilenstufenformen erstellt werden. Die Zeilenstufenform für einen Ergänzungsschritt wird von rechts mit $A_{(j)}$ multipliziert, das Produkt wird in Zeilenstufenform gebracht. Dies ist dann die Zeilenstufenform für den nächsten Ergänzungsschritt.

Nullzeilen darf man in diesen Zeilenstufenformen weglassen.

Man muß in jedem Ergänzungsschritt die Basisvektoren ergänzen, die in der Zeilenstufenform zu Nichtstufenspalten gehören, an deren Position im vorigen Schritt noch Stufenspalten waren.

Schritt 2. In Schritt 1 habe sich $H_A(\lambda_j) = \operatorname{Kern}(A_{(j)}^\ell)$ ergeben.

Man multipliziere die Vektoren aus der Ergänzung zu $\text{Kern}(A_{(j)}^\ell)$ mit $A_{(j)}$. Mit diesen Produkten ersetze man Vektoren aus der Ergänzung zu $\text{Kern}(A_{(j)}^{\ell-1})$ so, daß wieder eine Basis dieses Kerns entsteht.

Man multipliziere die nun verwendeten Vektoren aus der Ergänzung zu $\text{Kern}(A_{(j)}^{\ell-1})$ mit $A_{(j)}$. Mit diesen Produkten ersetze man Vektoren aus der Ergänzung zu $\text{Kern}(A_{(j)}^{\ell-2})$ so, daß wieder eine Basis dieses Kerns entsteht.

Man multipliziere die nun verwendeten Vektoren aus der Ergänzung zu $\text{Kern}(A_{(j)}^{\ell-2})$ mit $A_{(j)}$. Mit diesen Produkten ersetze man Vektoren aus der Ergänzung zu $\text{Kern}(A_{(j)}^{\ell-3})$ so, daß wieder eine Basis dieses Kerns entsteht.

Usf., bis man zur Basis von $\text{Kern}(A_{(j)}^1)$ gelangt ist.

Nun sortiere man noch die am Ende verwendeten Vektoren zu Hauptvektorketten. Ist y ein Vektor, der bei der Ergänzung zu $\text{Kern}(A_{(j)}^m)$ nicht ersetzt wurde, so gibt er Anlaß zur Hauptvektorkette $(A_{(j)}^{m-1}y, \dots, A_{(j)}^0y)$.

Diese Hauptvektorketten ergeben nebeneinandergestellt unsere geeignete Basis von $H_A(\lambda_j)$.

- (3) Die in (2) für jedes $1 \leq j \leq k$ gefundenen geeigneten Basen der Haupträume ergeben nebeneinandergestellt eine Jordanbasis von $\mathbb{C}^{n \times 1}$.

Diese Jordanbasis wird in die Spalten der Matrix $S \in \mathbb{C}^{n \times n}$ geschrieben. Dann ist $S^{-1}AS =: J$ in Jordanscher Normalform.

Jede Hauptvektorkette gibt Anlaß zu einem Jordanblock auf der Blockdiagonalen von J .

Die Kantenlänge eines Jordanblocks ist gleich der Länge der zugehörigen Hauptvektorkette. Die Diagonaleinträge eines Jordanblocks sind gleich dem Eigenwert, bezüglich dem die zugehörige Hauptvektorkette gebildet wurde.

Erinnerung an häufig auftretende Spezialfälle.

- (1) Ist in Schritt (2) in vorstehender Erinnerung bereits $E_A(\lambda_j) = H_A(\lambda_j)$, dann reduziert sich dieser Schritt zu einer Bestimmung einer Basis von $E_A(\lambda_j)$. Jeder Basisvektor ist dann eine Hauptvektorkette von Länge 1, und die zugehörigen Jordanblöcke sind 1×1 -Matrizen mit Eintrag λ_j .
- (2) Trifft $E_A(\lambda_j) = H_A(\lambda_j)$ für alle Eigenwerte λ_j zu, dann hat S in den Spalten eine Basis aus Eigenvektoren stehen, sortiert nach zugehörigen Eigenwerten. Es ist dann $S^{-1}AS = J$ eine Diagonalmatrix mit den zu den Eigenvektoren in den Spalten von S gehörigen Eigenwerten auf der Diagonalen. Diesemfalls ist A diagonalisierbar.
- (3) Es ist $gV_A(\lambda_j) = \dim(E_A(\lambda_j))$ die geometrische Vielfachheit und $aV_A(\lambda_j) = \dim(H_A(\lambda_j))$ die algebraische Vielfachheit von λ_j .

Es ist $1 \leq gV_A(\lambda_j) \leq aV_A(\lambda_j) \leq n$.

Genau dann ist $E_A(\lambda_j) = H_A(\lambda_j)$, wenn $gV_A(\lambda_j) = aV_A(\lambda_j)$ ist. Trifft dies zu für alle Eigenwerte λ_j , dann ist A diagonalisierbar.

- (4) Falls $aV_A(\lambda_j) = 1$ ist, dann ist auch $gV_A(\lambda_j) = 1$. Trifft dies zu für alle Eigenwerte λ_j , dann ist A diagonalisierbar.

Bemerkung. Sei A eine Blockdiagonalmatrix mit den Blöcken A_1, \dots, A_m .

Dann ist $\exp(A)$ die Blockdiagonalmatrix mit den Blöcken $\exp(A_1), \dots, \exp(A_m)$.

Beispiel. Sei $\lambda \in \mathbb{C}$. Sei $x \in \mathbb{R}$.

- (1) Es ist $J_1(\lambda) = (\lambda)$. Es ist $\exp(J_1(\lambda)x) = \exp((\lambda)x) = (\exp(\lambda x)) = (e^{\lambda x}) = e^{\lambda x} \cdot (1)$.

- (2) Es ist $J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Es ist $J_2(\lambda)^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Für $k \geq 1$ ist $J_2(\lambda)^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}$, wie Induktion über k ergeben wird.

Für $k = 1$ trifft dies zu. Ist $k \geq 2$ und trifft es für $k - 1$ zu, so wird

$$J_2(\lambda)^k = J_2(\lambda) \cdot J_2(\lambda)^{k-1} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \cdot \begin{pmatrix} \lambda^{k-1} & (k-1)\lambda^{k-2} \\ 0 & \lambda^{k-1} \end{pmatrix} = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}.$$

Für unser $x \in \mathbb{R}$ wird

$$\exp(J_2(\lambda)x) = \exp\left(\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} x\right) = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^k x^k = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k & \sum_{k=1}^{\infty} \frac{1}{k!} k \lambda^{k-1} x^k \\ 0 & \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k \end{pmatrix}.$$

Dabei wird $\sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k = \exp(\lambda x)$. Ferner wird

$$\sum_{k=1}^{\infty} \frac{1}{k!} k \lambda^{k-1} x^k = x \cdot \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \lambda^{k-1} x^{k-1} = x \cdot \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k = x \exp(\lambda x).$$

Also wird

$$\exp(J_2(\lambda)x) = \exp\left(\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} x\right) = \begin{pmatrix} e^{\lambda x} & x e^{\lambda x} \\ 0 & e^{\lambda x} \end{pmatrix} = e^{\lambda x} \cdot \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

- (3) Es ist $J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$.

Es ist $J_3(\lambda)^0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Es ist $J_3(\lambda)^1 = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$.

Für $k \geq 2$ ist $J_3(\lambda)^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{pmatrix}$, wie Induktion über k ergeben wird.

Für $k = 2$ trifft dies zu. Ist $k \geq 3$ und trifft es für $k - 1$ zu, so wird

$$J_3(\lambda)^k = J_3(\lambda) \cdot J_3(\lambda)^{k-1} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \cdot \begin{pmatrix} \lambda^{k-1} & (k-1)\lambda^{k-2} & \binom{k-1}{2}\lambda^{k-3} \\ 0 & \lambda^{k-1} & (k-1)\lambda^{k-2} \\ 0 & 0 & \lambda^{k-1} \end{pmatrix} = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{pmatrix}.$$

Wir beachten hierbei $\binom{k-1}{1} + \binom{k-1}{2} = \binom{k}{2}$.

Für unser $x \in \mathbb{R}$ wird

$$\begin{aligned} \exp(J_3(\lambda)x) &= \exp\left(\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} x\right) = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^k x^k \\ &= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k & \sum_{k=1}^{\infty} \frac{1}{k!} k \lambda^{k-1} x^k & \sum_{k=2}^{\infty} \frac{1}{k!} \binom{k}{2} \lambda^{k-2} x^k \\ 0 & \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k & \sum_{k=1}^{\infty} \frac{1}{k!} k \lambda^{k-1} x^k \\ 0 & 0 & \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k \end{pmatrix}. \end{aligned}$$

Wie in (2) wird dabei $\sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k = \exp(\lambda x)$ und $\sum_{k=1}^{\infty} \frac{1}{k!} k \lambda^{k-1} x^k = x \exp(\lambda x)$.
Ferner wird

$$\sum_{k=2}^{\infty} \frac{1}{k!} \binom{k}{2} \lambda^{k-2} x^k = \frac{x^2}{2} \cdot \sum_{k=2}^{\infty} \frac{1}{(k-2)!} \lambda^{k-2} x^{k-2} = \frac{x^2}{2} \cdot \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k x^k = \frac{x^2}{2} \exp(\lambda x).$$

Also wird

$$\exp(J_3(\lambda)x) = \exp\left(\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} x\right) = \begin{pmatrix} e^{\lambda x} & x e^{\lambda x} & x^2 e^{\lambda x}/2 \\ 0 & e^{\lambda x} & x e^{\lambda x} \\ 0 & 0 & e^{\lambda x} \end{pmatrix} = e^{\lambda x} \begin{pmatrix} 1 & x & x^2/2 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}.$$

Konstruktion. Wir wollen $\exp(Ax)$ berechnen für $x \in \mathbb{R}$.

Man finde $S \in \mathbb{C}^{n \times n}$ invertierbar mit $S^{-1}AS =: J$ in Jordanscher Normalform.

Es wird $\exp(Ax) = \exp(S(S^{-1}AS)xS^{-1}) = S \exp(Jx)S^{-1}$.

Beispiel. Sei $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$. Wir wollen $\exp(Ax)$ berechnen für $x \in \mathbb{R}$.

Es ist $\chi_A(X) = \det \begin{pmatrix} -X & 1 \\ -1 & -X \end{pmatrix} = X^2 + 1 = (X - i)(X - (-i))$.

Es ist also $\lambda_1 = i$ und $\lambda_2 = -i$, jeweils mit algebraischer Vielfachheit 1. Somit ist A diagonalisierbar.

Zu $\lambda_1 = i$. Wir formen um: $A - iE_2 = A_{(1)} = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}$.

Also hat $E_A(i) = \text{Kern}(A - iE_2)$ die Basis $\left(\begin{pmatrix} -i \\ 1 \end{pmatrix}\right)$.

Zu $\lambda_2 = -i$. Wir formen um: $A + iE_2 = A_{(2)} = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}$.

Also hat $E_A(-i) = \text{Kern}(A + iE_2)$ die Basis $\left(\begin{pmatrix} i \\ 1 \end{pmatrix}\right)$.

Es wird $S = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$ und $S^{-1}AS = J = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Es wird $\exp(Jx) = \begin{pmatrix} \exp(ix) & 0 \\ 0 & \exp(-ix) \end{pmatrix} = \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix}$.

Wir rechnen.

$$\begin{aligned} \exp(Ax) &= S \exp(Jx) S^{-1} \\ &= \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix} \cdot \frac{1}{-2i} \begin{pmatrix} 1 & -i \\ -1 & -i \end{pmatrix} \\ &= \frac{1}{2i} \begin{pmatrix} -i e^{ix} & i e^{-ix} \\ e^{ix} & e^{-ix} \end{pmatrix} \cdot \begin{pmatrix} -1 & i \\ 1 & i \end{pmatrix} \\ &= \frac{1}{2i} \begin{pmatrix} i(e^{ix} + e^{-ix}) & e^{ix} - e^{-ix} \\ -(e^{ix} - e^{-ix}) & i(e^{ix} + e^{-ix}) \end{pmatrix} \\ &= \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix}. \end{aligned}$$

Zur Probe rechnen wir $\det(\exp(Ax)) = \det \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} = 1$ aus. Dies stimmt tatsächlich mit $\exp(\operatorname{tr}(Ax)) = \exp(0) = 1$ überein.

Beispiel. Sei $A = \begin{pmatrix} 3 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{C}^{4 \times 4}$. Wir wollen $\exp(Ax)$ berechnen für $x \in \mathbb{R}$.

$$\begin{aligned} \text{Wir berechnen zuerst das charakteristische Polynom } \chi_A(X) &= \det \begin{pmatrix} 3-X & 1 & 1 & -1 \\ 0 & 1-X & -1 & 1 \\ 0 & 0 & 2-X & 0 \\ 1 & 0 & 0 & 2-X \end{pmatrix} \\ &= (2-X) \det \begin{pmatrix} 3-X & 1 & -1 \\ 0 & 1-X & 1 \\ 1 & 0 & 2-X \end{pmatrix} = (2-X) \det \begin{pmatrix} 3-X & 0 & -1 \\ 0 & 2-X & 1 \\ 1 & 2-X & 2-X \end{pmatrix} = (2-X)^2 \det \begin{pmatrix} 3-X & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 2-X \end{pmatrix} \\ &= (2-X)^2 \det \begin{pmatrix} 3-X & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1-X \end{pmatrix} = (2-X)^2 \det \begin{pmatrix} 3-X & -1 \\ 1 & 1-X \end{pmatrix} = (2-X)^2 (X^2 - 4X + 4) \\ &= (X-2)^4. \end{aligned}$$

Wir haben den Eigenwert $\lambda_1 = 2$ mit algebraischer Vielfachheit $\operatorname{aV}_A(2) = 4$.

Zu $\lambda_1 = 2$. Schritt 1. Wir formen um und lassen dabei Nullzeilen wegfallen:

$$A_{(1)} = A - 2E_4 = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix}.$$

Wir erhalten die Basis $\left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right)$ von $\operatorname{Kern}(A_{(1)})$.

Wir formen um:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix} \cdot A_{(1)} = \begin{pmatrix} 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix}.$$

Wir ergänzen mit $\left(\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$ zur Basis $\left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$ von $\operatorname{Kern}(A_{(1)}^2)$.

Wir formen um:

$$\begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix} \cdot A_{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}.$$

Wir ergänzen mit $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$ zur Basis $\left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$ von $\operatorname{Kern}(A_{(1)}^3)$.

Nun beenden wir die Ergänzung, da $\operatorname{aV}_A(2) = 4$ Basisvektoren vorliegen und somit eine Basis von $H_A(2)$ gefunden ist.

Schritt 2. Wir müssen noch den Ersetzungsprozeß durchlaufen.

Es wird $A_{(1)} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. Mit diesem ersetzen wir $\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ und erhalten so die Basis $\left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$ von $\operatorname{Kern}(A_{(1)}^2)$.

Es wird $A_{(1)} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$. Mit diesem ersetzen wir $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ und erhalten so die Basis $\left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right)$ von $\operatorname{Kern}(A_{(1)})$.

Nach Vorschrift. Natürlich haben wir festgestellt, daß dies zufällig auf keine Änderung hinausläuft. Erwähnen wir diese überflüssige Ersetzung, dann haben wir aber die Orientierung über die Kettenbildung.

Es resultieren die Hauptvektorketten $\left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right)$ und $\left(\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}\right)$.

Somit ist $S = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$ und $S^{-1}AS = J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Wir berechnen noch S^{-1} .

$$\begin{aligned} \left(\begin{array}{cccc|cccc} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 1 \end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & -1 & 0 & 1 \end{array}\right) \\ &\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right) \end{aligned}$$

Wir erhalten $S^{-1} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

Es ist $\exp(Jx) = \exp\left(\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} x\right) = \begin{pmatrix} e^{2x} & x e^{2x} & x^2 e^{2x}/2 & 0 \\ 0 & e^{2x} & x e^{2x} & 0 \\ 0 & 0 & e^{2x} & 0 \\ 0 & 0 & 0 & e^{2x} \end{pmatrix}$, vgl. obiges Beispiel, (3) und (1).

Es wird

$$\begin{aligned} \exp(Ax) &= S \exp(Jx) S^{-1} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} e^{2x} & x e^{2x} & x^2 e^{2x}/2 & 0 \\ 0 & e^{2x} & x e^{2x} & 0 \\ 0 & 0 & e^{2x} & 0 \\ 0 & 0 & 0 & e^{2x} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \frac{1}{2} e^{2x} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2x & x^2 & 0 \\ 0 & 2 & 2x & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \frac{1}{2} e^{2x} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x^2 & 2-2x+x^2 & 2-2x+x^2 & 2x-x^2 \\ 2x & -2+2x & -2+2x & 2-2x \\ 2 & 2 & 2 & -2 \\ 0 & 0 & 2 & 0 \end{pmatrix} \\ &= \frac{1}{2} e^{2x} \begin{pmatrix} 2x+2 & 2x & 2x & -2x \\ x^2 & 2-2x+x^2 & -2x+x^2 & 2x-x^2 \\ 0 & 0 & 2 & 0 \\ x^2+2x & x^2 & x^2 & 2-x^2 \end{pmatrix}. \end{aligned}$$

6.3.2.2 Lösungsverfahren

Konstruktion. Sei $A \in \mathbb{R}^{n \times n}$.

Wir wollen das homogene lineare Differentialgleichungssystem $y' = Ay$ lösen.

Dazu ziehen wir die Matrix

$$\exp(Ax)$$

heran; vgl. §6.3.2.1. Sei das Tupel der Spalten dieser Matrix gleich

$$(y_{[1]}(x), \dots, y_{[n]}(x)),$$

mit anderen Worten, sei $y_{[k]}(x) := \exp(Ax) e_k$ für $1 \leq k \leq n$.

Dann besteht dieses Tupel aus Lösungen von $y' = Ay$ auf \mathbb{R} , da $y'_{[k]}(x) = A \exp(Ax) e_k = Ay_{[k]}(x)$ ist für $x \in \mathbb{R}$ und $1 \leq k \leq n$; vgl. erste Bemerkung in §6.3.2.1.

Die Wronski-Determinante dieses Tupels ist

$$w(x) = \det(\exp(Ax)) = \exp(\operatorname{tr}(Ax)) \neq 0,$$

für $x \in \mathbb{R}$, da die Exponentialfunktion keine Nullstelle hat. Man kann auch argumentieren, es ist $\exp(Ax)$ invertierbar mit $\exp(Ax)^{-1} = \exp(-Ax)$ und deswegen $w(x) \neq 0$ für $x \in \mathbb{R}$.

Also ist

$$(y_{[1]}(x), \dots, y_{[n]}(x)) \text{ ein Fundamentalsystem für } y' = Ay,$$

d.h. eine \mathbb{R} -lineare Basis des Lösungsraums dieses Differentialgleichungssystems.

Mit anderen Worten, jede Lösung von $y' = Ay$ auf \mathbb{R} ist von der Form

$$y(x) = c_1 y_{[1]}(x) + \dots + c_n y_{[n]}(x) = \exp(Ax)c,$$

wobei $c_1, \dots, c_n \in \mathbb{R}$, zusammengefaßt zum Koeffizientenvektor $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^{n \times 1}$. Verschiedene Koeffizientenvektoren geben verschiedene Lösungen.

Beispiel. Wir wollen alle Lösungen des Differentialgleichungssystems

$$\begin{aligned} y_1' &= -y_2 \\ y_2' &= y_1 \end{aligned}$$

auf \mathbb{R} finden. Dann wollen wir insbesondere die Lösung zum Anfangswert $y(\frac{\pi}{2}) = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ bestimmen.

Zunächst schreiben wir dieses Differentialgleichungssystem um zu

$$y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y.$$

Es ist also $A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Es ist $\exp(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x) = \exp(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (-x)) = \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix}$; vgl. vorletztes Beispiel in §6.3.2.1.

Also ist $(\begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}, \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix})$ ein Fundamentalsystem für $y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y$, definiert auf \mathbb{R} .

Für jede Lösung aus dem Fundamentalsystem machen wir eine Probe.

Für den ersten Vektor $y(x) = \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}$ ist $y_1(x) = \cos(x)$ und $y_2(x) = \sin(x)$. Es wird

$$\begin{aligned} y_1'(x) &= -\sin(x) = -y_2(x) \\ y_2'(x) &= \cos(x) = y_1(x) \end{aligned}$$

Für den zweiten Vektor $y(x) = \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix}$ ist $y_1(x) = -\sin(x)$ und $y_2(x) = \cos(x)$. Es wird

$$\begin{aligned} y_1'(x) &= -\cos(x) = -y_2(x) \\ y_2'(x) &= -\sin(x) = y_1(x) \end{aligned}$$

Ende der Probe.

Jede Lösung von $y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y$ ist also von der Form

$$y(x) = c_1 \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix} + c_2 \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix},$$

wobei $c_1, c_2 \in \mathbb{R}$.

Nun wollen wir noch die Lösung $y(x)$ mit $y(\frac{\pi}{2}) = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ ermitteln. Einsetzen von $x = \frac{\pi}{2}$ ergibt die Bedingung

$$y(\frac{\pi}{2}) = c_1 \begin{pmatrix} \cos(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) \end{pmatrix} + c_2 \begin{pmatrix} -\sin(\frac{\pi}{2}) \\ \cos(\frac{\pi}{2}) \end{pmatrix} = c_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 4 \\ 5 \end{pmatrix},$$

was $c_1 = 5$ und $c_2 = -4$ liefert. Wir erhalten somit die Lösung

$$y(x) = 5 \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix} - 4 \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix} = \begin{pmatrix} 5 \cos(x) + 4 \sin(x) \\ 5 \sin(x) - 4 \cos(x) \end{pmatrix}$$

auf \mathbb{R} von $y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y$ zum Anfangswert $y(\frac{\pi}{2}) = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$.

Alternativ können wir für den Anfangswert auch aus $y(x) = \exp(Ax)c$ mit $c \in \mathbb{R}^{2 \times 1}$ schließen, daß $y(\frac{\pi}{2}) = \exp(A\frac{\pi}{2})c = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ sein sollte, also $c = \exp(A(-\frac{\pi}{2})) \begin{pmatrix} 4 \\ 5 \end{pmatrix}$, also

$$\begin{aligned} y(x) &= \exp(Ax) \exp(A(-\frac{\pi}{2})) \begin{pmatrix} 4 \\ 5 \end{pmatrix} \\ &= \exp(A(x - \frac{\pi}{2})) \begin{pmatrix} 4 \\ 5 \end{pmatrix} \\ &= \begin{pmatrix} \cos(x - \frac{\pi}{2}) & -\sin(x - \frac{\pi}{2}) \\ \sin(x - \frac{\pi}{2}) & \cos(x - \frac{\pi}{2}) \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} \\ &= \begin{pmatrix} \sin(x) & \cos(x) \\ -\cos(x) & \sin(x) \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \sin(x) + 5 \cos(x) \\ -4 \cos(x) + 5 \sin(x) \end{pmatrix}. \end{aligned}$$

Beispiel. Sei $A = \begin{pmatrix} 3 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{4 \times 4}$. Wir wollen die Lösung des Differentialgleichungssystems $y' = Ay$ zur Anfangsbedingung $y(-1) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ bestimmen.

Es ist

$$\exp(Ax) = \frac{1}{2} e^{2x} \begin{pmatrix} 2x+2 & 2x & 2x & -2x \\ x^2 & 2-2x+x^2 & -2x+x^2 & 2x-x^2 \\ 0 & 0 & 2 & 0 \\ x^2+2x & x^2 & x^2 & 2-x^2 \end{pmatrix};$$

vgl. letztes Beispiel in §6.3.2.1.

Wir erhalten also das Fundamentalsystem

$$\left(\frac{1}{2} e^{2x} \begin{pmatrix} 2x+2 \\ x^2 \\ 0 \\ x^2+2x \end{pmatrix}, \frac{1}{2} e^{2x} \begin{pmatrix} 2x \\ 2-2x+x^2 \\ 0 \\ x^2 \end{pmatrix}, \frac{1}{2} e^{2x} \begin{pmatrix} 2x \\ -2x+x^2 \\ 2 \\ x^2 \end{pmatrix}, \frac{1}{2} e^{2x} \begin{pmatrix} -2x \\ 2x-x^2 \\ 0 \\ 2-x^2 \end{pmatrix} \right)$$

für $y' = Ay$. Mit anderen Worten, es ist jede Lösung von $y' = Ay$ von der Form

$$y(x) = \frac{1}{2} e^{2x} \left(c_1 \begin{pmatrix} 2x+2 \\ x^2 \\ 0 \\ x^2+2x \end{pmatrix} + c_2 \begin{pmatrix} 2x \\ 2-2x+x^2 \\ 0 \\ x^2 \end{pmatrix} + c_3 \begin{pmatrix} 2x \\ -2x+x^2 \\ 2 \\ x^2 \end{pmatrix} + c_4 \begin{pmatrix} -2x \\ 2x-x^2 \\ 0 \\ 2-x^2 \end{pmatrix} \right),$$

wobei $c_1, c_2, c_3, c_4 \in \mathbb{R}$.

Damit wollen wir das Anfangswertproblem lösen. Sein soll

$$y(-1) = \frac{1}{2} e^{-2} \left(c_1 \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} -2 \\ 5 \\ 0 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} -2 \\ 3 \\ 2 \\ 1 \end{pmatrix} + c_4 \begin{pmatrix} -2 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right) \stackrel{!}{=} \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir lösen dazu ein lineares Gleichungssystem wie folgt.

$$\begin{aligned} \left(\begin{array}{cccc|c} 0 & -2 & -2 & 2 & 4e^2 \\ 1 & 5 & 3 & -3 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ -1 & 1 & 1 & 1 & 0 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & -1 & -1 & 0 \\ 0 & -2 & -2 & 2 & 4e^2 \\ 0 & 6 & 4 & -2 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 & -2e^2 \\ 0 & 3 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & -2 & -2e^2 \\ 0 & 1 & 1 & -1 & -2e^2 \\ 0 & 0 & -1 & 2 & 6e^2 \\ 0 & 0 & 1 & 0 & 0 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & -2 & -2e^2 \\ 0 & 1 & 0 & -1 & -2e^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 6e^2 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 4e^2 \\ 0 & 1 & 0 & 0 & e^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3e^2 \end{array} \right) \end{aligned}$$

Das gibt $c_1 = 4e^2$, $c_2 = e^2$, $c_3 = 0$ und $c_4 = 3e^2$. Wir erhalten somit die Lösung

$$\begin{aligned} y(x) &= \frac{1}{2} e^{2x} \left(4e^2 \begin{pmatrix} 2x+2 \\ x^2 \\ 0 \\ x^2+2x \end{pmatrix} + e^2 \begin{pmatrix} 2x \\ 2-2x+x^2 \\ 0 \\ x^2 \end{pmatrix} + 3e^2 \begin{pmatrix} -2x \\ 2x-x^2 \\ 0 \\ 2-x^2 \end{pmatrix} \right) \\ &= \frac{1}{2} e^{2x+2} \begin{pmatrix} 4x+8 \\ 2x^2+4x+2 \\ 0 \\ 2x^2+8x+6 \end{pmatrix} = e^{2x+2} \begin{pmatrix} 2x+4 \\ x^2+2x+1 \\ 0 \\ x^2+4x+3 \end{pmatrix} \end{aligned}$$

von $y' = Ay$ zum Anfangswert $y(-1) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Alternativ können wir für den Anfangswert auch aus $y(x) = \exp(Ax)c$ mit $c \in \mathbb{R}^{4 \times 1}$ schließen, daß $y(-1) = \exp(A(-1))c = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ sein sollte, also $c = \exp(A) \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, also

$$\begin{aligned} y(x) &= \exp(Ax) \exp(A) \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &= \exp(A(x+1)) \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &= \frac{1}{2} e^{2(x+1)} \begin{pmatrix} (2(x+1)+2) \cdot 2 \\ (x+1)^2 \cdot 2 \\ 0 \cdot 2 \\ ((x+1)^2+2(x+1)) \cdot 2 \end{pmatrix} = e^{2x+2} \begin{pmatrix} 2x+4 \\ x^2+2x+1 \\ 0 \\ x^2+4x+3 \end{pmatrix}. \end{aligned}$$

Bemerkung. Die Matrixexponentialfunktion ist nur im Falle einer konstanten Koeffizientenmatrix zum Lösen eines homogenen linearen Differentialgleichungssystems zu gebrauchen.

Um dies einzusehen, betrachten wir einmal die nichtkonstante Koeffizientenmatrix

$$A(x) := (x)$$

der Größe 1×1 . Wir wollen $y' = A(x)y$ lösen. Mit anderen Worten, wir wollen $y'_1 = xy_1$ lösen.

Diese Differentialgleichung ist separierbar. Zu lösen ist $\frac{y_1'}{y_1} = x$, also $[\ln(|y_1|)] = \int \frac{1}{y_1} dy_1 = \int \frac{y_1'}{y_1} dx = [\frac{1}{2}x^2]$, also $y_1(x) = d e^{x^2/2}$ für ein $d \in \mathbb{R}$. Das ist die richtige allgemeine Lösung.

Nun ist $\exp(A(x) \cdot x) = \exp((x^2)) = (e^{x^2})$.

Und $y_1(x) = e^{x^2}$ ist keine Lösung zu $y_1' = xy_1$. Denn es ist $y_1'(x) = 2x e^{x^2}$, was auf \mathbb{R} nicht dieselbe Funktion ist wie $xy_1(x) = x e^{x^2}$. Es unterscheiden sich die Werte etwa bei $x = 1$, da $2e \neq e$.

Konstruktion. Sei $A \in \mathbb{R}^{n \times n}$.

Sei $g : \mathbb{R} \rightarrow \mathbb{R}^{n \times 1} : x \mapsto g(x) = (g_j(x))_j$ nicht die Nullfunktion.

Wir wollen das inhomogene lineare Differentialgleichungssystem $y' = Ay + g(x)$ lösen.

Wir bestimmen eine partikuläre Lösung $y_{[0]}(x)$ von $y' = Ay + g(x)$.

Wir bestimmen ein Fundamentalsystem $(y_{[1]}(x), \dots, y_{[n]}(x))$ von $y' = Ay$, des zugehörigen homogenen linearen Differentialgleichungssystems.

Jede Lösung von $y' = Ay + g(x)$ auf \mathbb{R} ist von der Form

$$y(x) = y_{[0]}(x) + c_1 y_{[1]}(x) + \dots + c_n y_{[n]}(x) = y_{[0]}(x) + \exp(Ax)c,$$

wobei $c_1, \dots, c_n \in \mathbb{R}$, zusammengefaßt zum Koeffizientenvektor $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^{n \times 1}$.

Zur Bestimmung einer partikulären Lösung $y_{[0]}(x)$ von $y' = Ay + g(x)$ machen wir den Ansatz der *Variation der Konstanten*

$$y_{[0]}(x) = \exp(Ax)c(x)$$

für eine zu bestimmende Funktion $c : \mathbb{R} \rightarrow \mathbb{R}^{n \times 1} : x \mapsto c(x)$.

Es wird

$$y_{[0]}'(x) = A \exp(Ax)c(x) + \exp(Ax)c'(x)$$

und

$$Ay_{[0]}(x) + g(x) = A \exp(Ax)c(x) + g(x).$$

Gleichsetzen liefert die Bedingung

$$\exp(Ax)c'(x) \stackrel{!}{=} g(x),$$

d.h.

$$c'(x) \stackrel{!}{=} \exp(Ax)^{-1}g(x) = \exp(-Ax)g(x) = \exp(A \cdot (-x))g(x).$$

Hierbei erhält man $\exp(A \cdot (-x))$ aus $\exp(Ax)$ durch Ersetzen von x durch $-x$.

Dann löst man dies durch Integration in jedem Vektoreintrag und erhält so $c(x)$. Damit bilde man $y_{[0]}(x) = \exp(Ax)c(x)$, was eine partikuläre Lösung $y_{[0]}(x)$ auf \mathbb{R} ist, wie gesucht.

Beispiel. Wir wollen alle Lösungen des inhomogenen linearen Differentialgleichungssystems

$$\begin{aligned} y_1' &= -y_2 \\ y_2' &= y_1 + x \end{aligned}$$

auf \mathbb{R} finden.

Zunächst schreiben wir dieses Differentialgleichungssystem um zu

$$y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y + \begin{pmatrix} 0 \\ x \end{pmatrix} .$$

Es ist also $A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ und $g(x) = \begin{pmatrix} 0 \\ x \end{pmatrix}$.

Es ist $\exp(Ax) = \exp\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x\right) = \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix}$; vgl. erstes Beispiel in §6.3.2.2.

Ein Fundamentalsystem des zugehörigen homogenen linearen Differentialgleichungssystems $y' = Ay$ war dort also gegeben durch $(y_{[1]}(x), y_{[2]}(x)) = \left(\begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}, \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix}\right)$.

Variation der Konstanten: Wir machen also den Ansatz $y_{[0]}(x) = \exp(Ax)c(x)$, mit einem zu bestimmenden $c(x) = \begin{pmatrix} c_1(x) \\ c_2(x) \end{pmatrix}$.

Es wird $y'_{[0]}(x) = A \exp(Ax)c(x) + \exp(Ax)c'(x)$ und $Ay_{[0]}(x) + g(x) = A \exp(Ax)c(x) + \begin{pmatrix} 0 \\ x \end{pmatrix}$. Zu lösen ist also

$$\exp(Ax)c'(x) \stackrel{!}{=} \begin{pmatrix} 0 \\ x \end{pmatrix} ,$$

d.h.

$$c'(x) = \begin{pmatrix} c'_1(x) \\ c'_2(x) \end{pmatrix} \stackrel{!}{=} \exp(A \cdot (-x)) \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} \cos(-x) & -\sin(-x) \\ \sin(-x) & \cos(-x) \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} x \sin(x) \\ x \cos(x) \end{pmatrix} .$$

Es ist $\int x \sin(x) dx = [x(-\cos(x))] + \int 1 \cdot \cos(x) dx = [-x \cos(x) + \sin(x)]$.

Es ist $\int x \cos(x) dx = [x \sin(x)] - \int 1 \cdot \sin(x) dx = [x \sin(x) + \cos(x)]$.

Wir brauchen nur eine partikuläre Lösung. Wir können also $c_1(x) = -x \cos(x) + \sin(x)$ und $c_2(x) = x \sin(x) + \cos(x)$ wählen und erhalten

$$\begin{aligned} y_{[0]}(x) &= \exp(Ax)c(x) \\ &= \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix} \begin{pmatrix} -x \cos(x) + \sin(x) \\ x \sin(x) + \cos(x) \end{pmatrix} \\ &= \begin{pmatrix} -x \cos(x)^2 + \cos(x) \sin(x) - x \sin(x)^2 - \sin(x) \cos(x) \\ -x \sin(x) \cos(x) + \sin(x)^2 + x \cos(x) \sin(x) + \cos(x)^2 \end{pmatrix} \\ &= \begin{pmatrix} -x \\ 1 \end{pmatrix} . \end{aligned}$$

Somit ist jede Lösung von $y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y + \begin{pmatrix} 0 \\ x \end{pmatrix}$ von der Form

$$y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} = y_{[0]}(x) + c_1 y_{[1]}(x) + c_2 y_{[2]}(x) = \begin{pmatrix} -x \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix} + c_2 \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix} ,$$

wobei $c_1, c_2 \in \mathbb{R}$.

Einzeln geschrieben also

$$\begin{aligned} y_1(x) &= -x + c_1 \cos(x) - c_2 \sin(x) \\ y_2(x) &= 1 + c_1 \sin(x) + c_2 \cos(x) \end{aligned}$$

Bemerkung. Sei $A \in \mathbb{R}^{n \times n}$. Sei ein Fundamentalsystem für $y' = Ay$ gesucht.

Ein solches steht in den Spalten von $\exp(Ax)$.

Sei $S \in \mathbb{C}^{n \times n}$ invertierbar mit $J := S^{-1}AS$ in Jordanform. Sei einmal der günstige Fall angenommen, daß $S \in \mathbb{R}^{n \times n}$ gefunden wurde. Dies ist möglich, falls die Eigenwerte von $A \in \mathbb{C}^{n \times n}$ alle in \mathbb{R} liegen.

Dann kann man die volle Berechnung von $\exp(Ax)$ auch noch vermeiden.

Es enthält $\exp(Ax) = S \exp(Jx)S^{-1}$ in den Spalten ein Fundamentalsystem. Also enthält auch $\exp(Ax)S = S \exp(Jx)$ in den Spalten ein Fundamentalsystem.

Kurz: Ist S reell, dann enthält auch

$$S \exp(Jx)$$

in den Spalten ein Fundamentalsystem für $y' = Ay$.

6.4 Lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten

Sei $n \geq 1$. Sei $D \subseteq \mathbb{R}$ offen.

Wir betrachten auf D die Differentialgleichung n -ter Ordnung

$$(*) \quad y^{(n)} = a_{n-1}y^{(n-1)} + a_{n-2}y^{(n-2)} + \dots + a_0y^{(0)} + b(x),$$

wobei $a_0, \dots, a_{n-1} \in \mathbb{R}$ und $b : D \rightarrow \mathbb{R} : x \mapsto b(x)$.

Konstruktion. Wir setzen

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} := \begin{pmatrix} y^{(0)} \\ y^{(1)} \\ \vdots \\ y^{(n-1)} \end{pmatrix}.$$

Unsere Differentialgleichung n -ter Ordnung für y wird zu folgendem linearen Differentialgleichungssystem erster Ordnung mit konstanten Koeffizienten.

$$\begin{aligned} z_1' &= z_2 \\ z_2' &= z_3 \\ &\vdots \\ z_{n-1}' &= z_n \\ z_n' &= a_0z_1 + a_1z_2 + \dots + a_{n-1}z_n + b(x). \end{aligned}$$

Die ersten $n - 1$ Gleichungen darin geben nur den Zusammenhang zwischen den aufeinanderfolgenden Ableitungen von y wieder.

Die letzte Gleichung darin gibt die Differentialgleichung (*) wieder.

Unter Verwendung einer Matrix können wir dieses System wie folgt schreiben.

$$z' = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ a_0 & a_1 & \dots & \dots & a_{n-1} \end{pmatrix} z + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(x) \end{pmatrix}$$

Leergelassene Einträge seien dabei gleich 0.

Dies lösen wir wie in der ersten und der zweiten Konstruktion in §6.3.2.2.

Beispiel. Wir betrachten die Differentialgleichung $y'' + y = x$.

Wir suchen die Lösung zu den Anfangsbedingungen $y(0) = 1$ und $y'(0) = 0$.

Wir setzen $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} y \\ y' \end{pmatrix}$ und erhalten das folgende Differentialgleichungssystem.

$$\begin{aligned} z_1' &= z_2 \\ z_2' &= -z_1 + x \end{aligned}$$

Mit anderen Worten, es liegt folgendes Differentialgleichungssystem vor.

$$z' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} z + \begin{pmatrix} 0 \\ x \end{pmatrix} .$$

Es ist $\exp\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x\right) = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix}$; vgl. vorletztes Beispiel in §6.3.2.1.

Der Ansatz der Variation der Konstanten $z(x) = \exp\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x\right) \cdot c(x)$ gibt $\begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} c'(x) = \begin{pmatrix} 0 \\ x \end{pmatrix}$. Also $c'(x) = \begin{pmatrix} \cos(-x) & \sin(-x) \\ -\sin(-x) & \cos(-x) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} -x \sin(x) \\ x \cos(x) \end{pmatrix}$. Partielle Integration gibt

$$\begin{aligned} [c_1(x)] &= \int -x \sin(x) \, dx = [x \cos(x) - \sin(x)] \\ [c_2(x)] &= \int x \cos(x) \, dx = [x \sin(x) + \cos(x)] . \end{aligned}$$

Das liefert die partikuläre Lösung

$$z_{[0]}(x) = \exp\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x\right) \cdot c(x) = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} \cdot \begin{pmatrix} x \cos(x) - \sin(x) \\ x \sin(x) + \cos(x) \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix} .$$

Also wird

$$z(x) = \begin{pmatrix} x \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix} + c_2 \begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix} ,$$

wobei $c_1, c_2 \in \mathbb{R}$.

Als Anfangsbedingung haben wir

$$z(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} \cos(0) \\ -\sin(0) \end{pmatrix} + c_2 \begin{pmatrix} \sin(0) \\ \cos(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

zu erfüllen. Es muß $c_1 = 1$ und $c_2 = -1$ sein. Wir erhalten so die Lösung

$$z(x) = \begin{pmatrix} x \\ 1 \end{pmatrix} + \begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix} - \begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix} .$$

Dies müssen wir noch rückübersetzen. Es war $z_1 = y$. Also erhalten wir für die Differentialgleichung $y'' + y = x$ die Lösung

$$y(x) = x + \cos(x) - \sin(x)$$

zu den Anfangsbedingungen $y(0) = 1$ und $y'(0) = 0$.

Probe. Es wird $y'(x) = 1 - \sin(x) - \cos(x)$ und $y''(x) = -\cos(x) + \sin(x)$. Somit wird

$$y''(x) + y(x) = (-\cos(x) + \sin(x)) + (x + \cos(x) - \sin(x)) = x.$$

Ferner wird $y(0) = 0 + \cos(0) - \sin(0) = 1$ und $y'(0) = 1 - \sin(0) - \cos(0) = 0$.

Beispiel. Sei $t \in \mathbb{R}_{\geq 0}$ ein Parameter.

Wir betrachten die Differentialgleichung $y'' - 2ty' + y = 0$.

Wir suchen die Lösung zu den Anfangsbedingungen $y(0) = 0$ und $y'(0) = 1$, in Abhängigkeit vom Parameter t .

Wir setzen $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} y \\ y' \end{pmatrix}$ und erhalten das folgende Differentialgleichungssystem.

$$\begin{aligned} z_1' &= z_2 \\ z_2' &= -z_1 + 2tz_2 \end{aligned}$$

Mit anderen Worten, es liegt folgendes Differentialgleichungssystem vor.

$$z' = \begin{pmatrix} 0 & 1 \\ -1 & 2t \end{pmatrix} z.$$

Die Anfangsbedingungen für y und y' werden zur Anfangsbedingung $z(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ für z .

Sei $A_t := \begin{pmatrix} 0 & 1 \\ -1 & 2t \end{pmatrix} \in \mathbb{C}^{2 \times 2}$. Es ist $\chi_{A_t}(X) = \det \begin{pmatrix} -X & 1 \\ -1 & 2t-X \end{pmatrix} = X^2 - 2tX + 1$.

Das charakteristische Polynom hat die Nullstellen $t + \sqrt{t^2 - 1}$ und $t - \sqrt{t^2 - 1}$.

Fall $t \in \mathbb{R}_{\geq 0} \setminus \{1\}$. Wir haben die Eigenwerte $\lambda_1 := t + \sqrt{t^2 - 1}$ und $\lambda_2 := t - \sqrt{t^2 - 1}$, beide mit algebraischer Vielfachheit 1. Die Matrix A_t ist also diagonalisierbar.

Falls $t < 1$ ist, lesen wir hierbei $\sqrt{t^2 - 1} = i\sqrt{1 - t^2}$.

Wir kürzen ab zu $w := \sqrt{t^2 - 1}$.

Nebenrechnung: Es ist $(t - w)(t + w) = t^2 - w^2 = t^2 - (t^2 - 1) = 1$.

Zu $\lambda_1 = t + w$. Es wird $\begin{pmatrix} -t-w & 1 \\ -1 & t-w \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -t+w \\ 1 & -t-w \end{pmatrix}$. Also hat $E_A(\lambda_1)$ die Basis $\left(\begin{pmatrix} t-w \\ 1 \end{pmatrix}\right)$.

Zu $\lambda_2 = t - w$. Es wird $\begin{pmatrix} -t+w & 1 \\ -1 & t+w \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -t-w \\ 1 & -t-w \end{pmatrix}$. Also hat $E_A(\lambda_2)$ die Basis $\left(\begin{pmatrix} t+w \\ 1 \end{pmatrix}\right)$.

Mit $S_t := \begin{pmatrix} t-w & t+w \\ 1 & 1 \end{pmatrix}$ wird also $S_t^{-1} A_t S_t = \begin{pmatrix} t+w & 0 \\ 0 & t-w \end{pmatrix} =: J_t$.

Wir können S_t nach Formel invertieren zu

$$S_t^{-1} = \frac{1}{(t-w)-(t+w)} \begin{pmatrix} 1 & -t-w \\ -1 & t-w \end{pmatrix} = \frac{1}{2w} \begin{pmatrix} -1 & t+w \\ 1 & -t+w \end{pmatrix}$$

vgl. letztes Beispiel in §3.1.

Für $x \in \mathbb{R}$ ist

$$\exp(J_t x) = \begin{pmatrix} e^{(t+w)x} & 0 \\ 0 & e^{(t-w)x} \end{pmatrix} = e^{tx} \begin{pmatrix} e^{wx} & 0 \\ 0 & e^{-wx} \end{pmatrix}.$$

Also wird

$$\begin{aligned} \exp(A_t x) &= S_t \cdot \exp(J_t x) \cdot S_t^{-1} \\ &= \begin{pmatrix} t-w & t+w \\ 1 & 1 \end{pmatrix} \cdot e^{tx} \begin{pmatrix} e^{wx} & 0 \\ 0 & e^{-wx} \end{pmatrix} \cdot \frac{1}{2w} \begin{pmatrix} -1 & t+w \\ 1 & -t+w \end{pmatrix} \\ &= \frac{e^{tx}}{2w} \begin{pmatrix} t-w & t+w \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -e^{wx} & e^{wx}(t+w) \\ e^{-wx} & -e^{-wx}(t-w) \end{pmatrix} \\ &= \frac{e^{tx}}{2w} \begin{pmatrix} -(t-w)e^{wx} + (t+w)e^{-wx} & (t-w)e^{wx}(t+w) - (t+w)e^{-wx}(t-w) \\ -e^{wx} + e^{-wx} & e^{wx}(t+w) - e^{-wx}(t-w) \end{pmatrix} \\ &= \frac{e^{tx}}{2w} \begin{pmatrix} -(t-w)e^{wx} + (t+w)e^{-wx} & e^{wx} - e^{-wx} \\ -e^{wx} + e^{-wx} & e^{wx}(t+w) - e^{-wx}(t-w) \end{pmatrix}. \end{aligned}$$

Also wird

$$z(x) = \frac{e^{tx}}{2w} \begin{pmatrix} -(t-w)e^{wx} + (t+w)e^{-wx} & e^{wx} - e^{-wx} \\ -e^{wx} + e^{-wx} & e^{wx}(t+w) - e^{-wx}(t-w) \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

wobei $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in \mathbb{R}^{2 \times 1}$.

Speziell wird, unter Verwendung von $\exp(A_t \cdot 0) = E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,

$$z(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Somit wird

$$z(x) = \frac{e^{tx}}{2w} \begin{pmatrix} e^{wx} - e^{-wx} \\ e^{wx}(t+w) - e^{-wx}(t-w) \end{pmatrix}$$

die Lösung zur Anfangsbedingung $z(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Es ist $y = z_1$. Also wird

$$y(x) = \frac{e^{tx}}{2w} (e^{wx} - e^{-wx})$$

die Lösung zu den Anfangsbedingungen $y(0) = 0$ und $y'(0) = 1$.

Subfall $t \in [0, 1[$. Unsere Lösung wird zu

$$y(x) = \frac{e^{tx}}{2i\sqrt{1-t^2}} (e^{ix\sqrt{1-t^2}} - e^{-ix\sqrt{1-t^2}}) = \frac{e^{tx}}{\sqrt{1-t^2}} \sin(x\sqrt{1-t^2}).$$

Subfall $t \in]1, +\infty[$. Unsere Lösung wird zu

$$y(x) = \frac{e^{tx}}{2\sqrt{t^2-1}} (e^{x\sqrt{t^2-1}} - e^{-x\sqrt{t^2-1}}) = \frac{e^{tx}}{\sqrt{t^2-1}} \sinh(x\sqrt{t^2-1}).$$

Fall $t = 1$. Wir haben den Eigenwert $\lambda_1 = 1$ von $A_1 = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$. Es ist $A_1 - E_2 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$.

Es wird $\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \rightsquigarrow (1 \ -1)$. Also hat $E_{A_1}(1) = \text{Kern}((A_1 - E_2)^1)$ die Basis $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$.

Es wird $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Also können wir mit $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ ergänzen zur Basis $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ von $\text{Kern}((A_1 - E_2)^2) = H_{A_1}(1) = \mathbb{C}^{2 \times 1}$.

Schließlich wird $\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, mit welchem wir den Vektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ersetzen.

Wir erhalten so die einzige Hauptvektorkette $\left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$.

Sei demgemäß $S_1 := \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$. Es wird $S_1^{-1} A_1 S_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} =: J_1$.

Für $x \in \mathbb{R}$ ist $\exp(J_1 x) = \begin{pmatrix} e^x & x e^x \\ 0 & e^x \end{pmatrix} = e^x \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$; vgl. drittes Beispiel in §6.3.2.1. Also wird

$$\exp(A_1 x) = S_1 \cdot \exp(J_1 x) \cdot S_1^{-1} = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \cdot e^x \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} = e^x \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x & -x-1 \\ 1 & -1 \end{pmatrix} = e^x \begin{pmatrix} -x+1 & x \\ -x & x+1 \end{pmatrix}.$$

Also wird

$$z(x) = e^x \begin{pmatrix} -x+1 & x \\ -x & x+1 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

wobei $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in \mathbb{R}^{2 \times 1}$.

Speziell wird, unter Verwendung von $\exp(A_1 \cdot 0) = E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,

$$z(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Somit wird

$$z(x) = e^x \begin{pmatrix} x \\ x+1 \end{pmatrix}$$

die Lösung zur Anfangsbedingung $z(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Es ist $y = z_1$. Also wird

$$y(x) = x e^x$$

die Lösung zu den Anfangsbedingungen $y(0) = 0$ und $y'(0) = 1$.

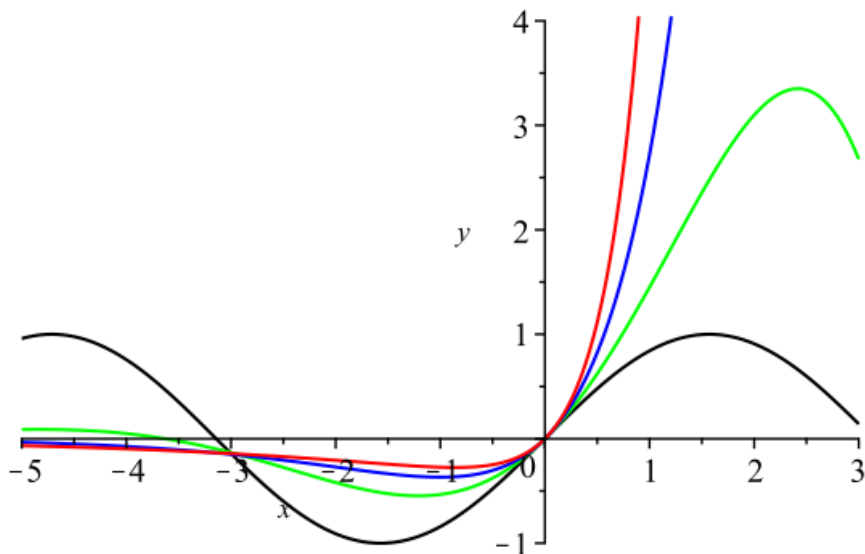
Fassen wir die Resultate für die Lösung von $y'' - 2ty' + y = 0$ auf \mathbb{R} zu den Anfangsbedingungen $y(0) = 0$ und $y'(0) = 1$ zusammen.

Ist $t \in [0, 1[$, so erhalten wir $y(x) = \frac{e^{tx}}{\sqrt{1-t^2}} \sin(x\sqrt{1-t^2})$.

Ist $t = 1$, so erhalten wird $y(x) = x e^x$.

Ist $t \in]1, +\infty[$, so erhalten wir $y(x) = \frac{e^{tx}}{\sqrt{t^2-1}} \sinh(x\sqrt{t^2-1})$.

Folgendes Schaubild enthält den Graphen der Lösung für $t = 0$ (schwarz), für $t = \frac{1}{2}$ (grün), für $t = 1$ (blau) und für $t = \frac{3}{2}$ (rot).



6.5 Weitere zurückführbare Differentialgleichungen

6.5.1 Bernoullische Differentialgleichung

Seien $a(x)$ und $b(x)$ reellwertige Funktionen auf einem geeigneten Definitionsbereich für die Variable x .

Sei $t \in \mathbb{R} \setminus \{0, 1\}$.

Zu lösen sei die *Bernoullische* Differentialgleichung

$$y' = a(x)y + b(x)y^t.$$

Dazu substituieren wir wie folgt zu einer gesuchten Funktion $u = u(x)$.

$$\begin{aligned} u &= y^{1-t} \\ u' &= (1-t)y^{-t} \cdot y' \end{aligned}$$

Wir erhalten die substituierte Differentialgleichung

$$u' = (1-t)y^{-t}(a(x)y + b(x)y^t) = (1-t)a(x)u + (1-t)b(x).$$

Diese ist linear und kann wie in §6.2.4 gelöst werden.

Die gefundene Lösung $u = u(x)$ substituieren wir zurück zu einer Lösung

$$y = y(x) = u(x)^{\frac{1}{1-t}}$$

der Bernoullischen Differentialgleichung.

Beispiel. Wir suchen alle Lösungen von

$$y' = y + y^2 .$$

Substitution $u = y^{1-2} = y^{-1}$ führt auf $u' = -y^{-2}y'$ und also auf

$$u' = -y^{-2}(y + y^2) = -y^{-1} - 1 = -u - 1 .$$

Es hat $u' = -u$ die Lösungen $u(x) = c e^{-x}$, wobei $c \in \mathbb{R}$.

Variation der Konstanten gibt den Ansatz $u = c(x)e^{-x}$, mit einer zu bestimmenden Funktion $c = c(x)$. Damit wird $u' = -u - 1$ zu $c'(x)e^{-x} - c(x)e^{-x} = -c(x)e^{-x} - 1$, also zu $c'(x) = -e^x$, also zu $c(x) = -e^x + d$, wobei $d \in \mathbb{R}$. Also wird $u = u(x) = c(x)e^{-x} = -1 + d e^{-x}$.

Schließlich wird

$$y = \frac{1}{u(x)} = \frac{1}{d e^{-x} - 1}$$

zur einer Lösung von $y' = y + y^2$, wobei $d \in \mathbb{R}$.

Diese ist definiert an allen Stellen, bei denen $d e^{-x} - 1 \neq 0$ ist.

Falls $d \leq 0$ ist, ist diese Lösung auf \mathbb{R} definiert.

Falls $d > 0$ ist, dann ist $d e^{-x} - 1 = 0$ genau dann, wenn $e^{-x} = d^{-1}$ ist, also $e^x = d$, also $x = \ln(d)$. Somit ist diesenfalls der Definitionsbereich $\mathbb{R} \setminus \{\ln(d)\}$ möglich.

Probe. Aus $y(x) = (d e^{-x} - 1)^{-1}$ folgt $y'(x) = -(d e^{-x} - 1)^{-2}(-d e^{-x}) = \frac{d e^{-x}}{(d e^{-x} - 1)^2}$.

Auf der anderen Seite ist $y(x) + y(x)^2 = (d e^{-x} - 1)^{-1} + (d e^{-x} - 1)^{-2} = \frac{d e^{-x} - 1 + 1}{(d e^{-x} - 1)^2} = \frac{d e^{-x}}{(d e^{-x} - 1)^2}$. Das ist dasselbe.

6.5.2 Riccatische Differentialgleichung

Seien $a(x)$, $b(x)$ und $c(x)$ reellwertige Funktionen auf einem geeigneten Definitionsbereich D für die Variable x .

Zu lösen sei die *Riccatische* Differentialgleichung

$$y' = a(x)y^2 + b(x)y + c(x) .$$

Zur Verfügung stehe uns eine partikuläre Lösung $\eta = \eta(x)$.

Es sei also $\eta'(x) = a(x)\eta(x)^2 + b(x)\eta(x) + c(x)$ für $x \in D$ bekannt.

Dazu setzen wir mit einer gesuchten Funktion $v = v(x)$ wie folgt an.

$$\begin{aligned} y &= \eta(x) + \frac{1}{v} \\ y' &= \eta'(x) - \frac{v'}{v^2} \end{aligned}$$

Wir erhalten die Differentialgleichung

$$\begin{aligned}\eta'(x) - \frac{v'}{v^2} &= a(x)\left(\eta(x) + \frac{1}{v}\right)^2 + b(x)\left(\eta(x) + \frac{1}{v}\right) + c(x) \\ &= a(x)\eta(x)^2 + 2a(x)\frac{\eta(x)}{v} + \frac{a(x)}{v^2} + b(x)\eta(x) + \frac{b(x)}{v} + c(x) .\end{aligned}$$

Da $\eta(x)$ eine Lösung ist, wird dies zu

$$-\frac{v'}{v^2} = 2a(x)\frac{\eta(x)}{v} + \frac{a(x)}{v^2} + \frac{b(x)}{v} ,$$

also zur Differentialgleichung

$$v' = -(2a(x)\eta(x) + b(x))v - a(x) .$$

Diese ist linear und kann wie in §6.2.4 gelöst werden.

Die gefundene Lösung $v = v(x)$ substituieren wir zurück zu einer Lösung

$$y = y(x) = \eta(x) + \frac{1}{v(x)}$$

der Riccatischen Differentialgleichung.

Beispiel. Wir suchen alle Lösungen von

$$y' = y + y^2 e^x + e^{-x} .$$

Die partikuläre Lösung $\eta(x) = -e^{-x}$ sei uns bekannt, sei es durch Probieren, sei es durch entsprechende Aufgabenstellung.

Wir machen den Ansatz $y = -e^{-x} + \frac{1}{v}$ und erhalten

$$\begin{aligned}y' &= e^{-x} - v^{-2}v' \\ y + y^2 e^x + e^{-x} &= -e^{-x} + v^{-1} + (-e^{-x} + v^{-1})^2 e^x + e^{-x} \\ &= v^{-1} + (e^{-2x} - 2e^{-x}v^{-1} + v^{-2})e^x \\ &= -v^{-1} + e^{-x} + e^x v^{-2} .\end{aligned}$$

Gleichsetzen liefert

$$e^{-x} - v^{-2}v' = -v^{-1} + e^{-x} + e^x v^{-2} ,$$

also

$$v' = v - e^x .$$

Wir lösen $v' = v$ mit $v = v(x) = e^x t$, wobei $t \in \mathbb{R}$. Variation der Konstanten $t = t(x)$ gibt $v = e^x t$ und $v' = e^x t + e^x t'$. Wir brauchen also

$$e^x t + e^x t' = e^x t - e^x$$

und somit

$$t' = -1 .$$

Wir erhalten

$$t = t(x) = d - x$$

mit $d \in \mathbb{R}$. In den Ansatz eingesetzt ergibt sich

$$v = e^x t = e^x(d - x).$$

In den Ansatz eingesetzt ergibt sich

$$y = y(x) = -e^{-x} + \frac{1}{v} = -e^{-x} + \frac{e^{-x}}{d - x},$$

wobei $d \in \mathbb{R}$. Diese Lösung ist definiert auf $\mathbb{R} \setminus \{d\}$.

Probe. Zum einen wird

$$y'(x) = e^{-x} + \frac{-e^{-x}(d-x) - e^{-x}(-1)}{(d-x)^2} = e^{-x} + \frac{e^{-x}(x-d+1)}{(d-x)^2}$$

Zum anderen wird

$$\begin{aligned} y(x) + y(x)^2 e^x + e^{-x} &= -e^{-x} + \frac{e^{-x}}{d-x} + \left(-e^{-x} + \frac{e^{-x}}{d-x}\right)^2 e^x + e^{-x} \\ &= \frac{e^{-x}}{d-x} + e^{-x} \left(1 - \frac{2}{d-x} + \frac{1}{(d-x)^2}\right) \\ &= e^{-x} + \frac{e^{-x}(d-x-2(d-x)+1)}{(d-x)^2} \\ &= e^{-x} + \frac{e^{-x}(x-d+1)}{(d-x)^2}. \end{aligned}$$

Das ist dasselbe.

Literatur

- [1] KÜNZER, M., *Lineare Algebra für Informatiker*, Skript, pnp.mathematik.uni-stuttgart.de/lexmath/kuenzer/skript.pdf, Ulm, 2004.
- [2] KÜNZER, M., *Mathematik für Wirtschaftswissenschaftler*, Skript, pnp.mathematik.uni-stuttgart.de/lexmath/kuenzer/wsk.pdf, Stuttgart, 2004.
- [3] LESKY, P., *Vorlesung Mathematik I+II für Studierende der Fächer inf, swt, msv*, info.mathematik.uni-stuttgart.de/Mathe2InfSS19/skr/Skript-1und2.pdf, Stuttgart, 2019.