

Mathematik für Maschinenbauer, Bauingenieure und Umwelttechniker I

Vorlesungsskriptum WS 2001/02
überarbeitet November 2006

R. Verfürth

Fakultät für Mathematik, Ruhr-Universität Bochum

Inhaltsverzeichnis

Vorbemerkung	9
Kapitel I. Zahlen und Vektoren	11
I.1. Mengen und Abbildungen	11
I.1.1. Mengen	11
I.1.2. Mengenoperationen	11
I.1.3. Abbildungen	12
I.2. Die reellen Zahlen	13
I.2.1. Bezeichnungen	13
I.2.2. Ungleichungen	13
I.2.3. Intervalle	15
I.2.4. Schranken	15
I.2.5. Der Betrag	16
I.2.6. Vollständige Induktion	16
I.2.7. Rekursive Definition	18
I.2.8. Binomialkoeffizienten und Binomische Formel	19
I.3. Die Ebene	20
I.3.1. Kartesische Koordinatensysteme	20
I.3.2. Winkel	21
I.3.3. Sinus, Cosinus	21
I.3.4. Drehungen	22
I.4. Vektoren	24
I.4.1. Vektoren	24
I.4.2. Addition von Vektoren	24
I.4.3. Skalare Vielfache von Vektoren	25
I.4.4. Der Betrag	25
I.4.5. Der Winkel zwischen zwei Vektoren	25
I.4.6. Das Skalarprodukt	26
I.4.7. Das Vektorprodukt	27
I.4.8. Das Spatprodukt	28
I.4.9. Koordinatendarstellungen	29
I.5. Geraden und Ebenen	33
I.5.1. Die Parameterdarstellung einer Geraden	33
I.5.2. Die Koordinatengleichungen einer Geraden	34
I.5.3. Die Momentengleichung einer Geraden	34
I.5.4. Abstand Punkt - Gerade	35
I.5.5. Abstand Gerade - Gerade	36

I.5.6.	Die Parameterdarstellung einer Ebene	37
I.5.7.	Parameterfreie Darstellungen einer Ebene	37
I.5.8.	Die Schnittgerade zweier Ebenen	39
I.5.9.	Die Winkel zwischen zwei Ebenen und zwischen einer Ebene und einer Geraden	40
I.6.	Die komplexen Zahlen	41
I.6.1.	Motivation	41
I.6.2.	Definition	41
I.6.3.	Betrag und Konjugation	41
I.6.4.	Addition und Subtraktion	42
I.6.5.	Multiplikation und Division	43
I.6.6.	Wurzeln komplexer Zahlen	45
Kapitel II.	Lineare Algebra	49
II.1.	Lineare Gleichungssysteme und Matrizen	49
II.1.1.	Matrizen	49
II.1.2.	Rechenregeln	50
II.1.3.	Lineare Gleichungssysteme und Matrizen	52
II.1.4.	Das Gaußsche Eliminationsverfahren	53
II.2.	Die Matrixmultiplikation	59
II.2.1.	Das Matrixprodukt	59
II.2.2.	Die transponierte Matrix	61
II.2.3.	Invertierbare Matrizen	62
II.2.4.	Das Gaußsche Eliminationsverfahren zur Berechnung der inversen Matrix	62
II.2.5.	Die LR-Zerlegung	64
II.3.	Determinanten	68
II.3.1.	Determinanten von 2×2 und 3×3 Matrizen	68
II.3.2.	Die Determinante einer $n \times n$ Matrix	70
II.3.3.	Rechenregeln für Determinanten	71
II.3.4.	Die Cramersche Regel	72
II.3.5.	Kegelschnitte	73
II.4.	Eigenwerte und Eigenvektoren	74
II.4.1.	Definition	74
II.4.2.	Das charakteristische Polynom	75
II.4.3.	Berechnung der Eigenwerte und Eigenvektoren	76
II.4.4.	Numerische Berechnung der Eigenwerte und Eigenvektoren*	78
II.4.5.	Rechenregeln	83
II.4.6.	Ähnliche Matrizen	84
II.4.7.	Orthogonale Matrizen	85
II.4.8.	Symmetrische Matrizen	87
II.4.9.	Die Schursche Normalform	88
II.4.10.	Hauptvektoren	89
II.5.	Quadratische Formen	92

II.5.1.	Definition	92
II.5.2.	Reduktion auf Normalform	94
II.5.3.	Normalformen der ebenen Quadriken	95
II.5.4.	Normalformen der räumlichen Quadriken	96
II.5.5.	Positiv definite Matrizen	97
II.6.	Vektorräume und lineare Abbildungen	98
II.6.1.	Vektorräume	98
II.6.2.	Unterräume	99
II.6.3.	Linearkombination und lineare Hülle	99
II.6.4.	Lineare Abhängigkeit	100
II.6.5.	Basis und Dimension	101
II.6.6.	Skalarprodukte	103
II.6.7.	Normen	104
II.6.8.	Orthogonalität	105
II.6.9.	Lineare Abbildungen	107
II.6.10.	Matrixdarstellung	107
II.6.11.	Komposition linearer Abbildungen	108
II.6.12.	Basiswechsel	108
Kapitel III.	Stetigkeit	111
III.1.	Folgen	111
III.1.1.	Definition	111
III.1.2.	Rechenregeln	111
III.1.3.	Beschränktheit	111
III.1.4.	Monotonie	112
III.1.5.	Teilfolgen	112
III.2.	Grenzwerte von Folgen	112
III.2.1.	Definition	112
III.2.2.	Rechenregeln	113
III.2.3.	Konvergenzkriterien	114
III.2.4.	Die Exponentialfunktion	117
III.2.5.	Uneigentliche Grenzwerte	118
III.3.	Stetigkeit	119
III.3.1.	Definition	119
III.3.2.	Rechenregeln	119
III.3.3.	Eigenschaften stetiger Funktionen	120
III.3.4.	Einseitige Grenzwerte	121
III.3.5.	Polynome	122
III.3.6.	Rationale Funktionen	124
III.3.7.	Trigonometrische Funktionen	125
III.3.8.	Exponential- und Logarithmusfunktion	126
Kapitel IV.	Differentiation	129
IV.1.	Die Ableitung einer differenzierbaren Funktion	129
IV.1.1.	Definition der Ableitung	129
IV.1.2.	Deutungen der Ableitung	130

IV.1.3.	Stetigkeit ist notwendig für Differenzierbarkeit	131
IV.1.4.	Differentiationsregeln	132
IV.1.5.	Differentiation von Polynomen und rationalen Funktionen	132
IV.1.6.	Differentiation der trigonometrischen Funktionen	133
IV.1.7.	Kettenregel	133
IV.1.8.	Höhere Ableitungen	135
IV.2.	Anwendungen der Differentiation	136
IV.2.1.	Maxima und Minima einer Funktion	136
IV.2.2.	Der Mittelwertsatz	138
IV.2.3.	Wendepunkte	140
IV.2.4.	Die Regeln von de l'Hôpital	142
IV.2.5.	Die Fixpunktiteration	144
IV.2.6.	Das Newtonverfahren	146
IV.3.	Umkehrfunktionen	150
IV.3.1.	Grundlagen	150
IV.3.2.	n -te Wurzel, rationale Exponenten	152
IV.3.3.	Die Umkehrfunktionen der trigonometrischen Funktionen	154
IV.4.	Die Exponential- und Logarithmusfunktion	157
IV.4.1.	Die Exponentialfunktion	157
IV.4.2.	Die Logarithmusfunktion	160
IV.4.3.	Exponential- und Logarithmusfunktionen zu beliebigen Basen	161
IV.4.4.	Die Hyperbelfunktionen	163
Kapitel V.	Integration	165
V.1.	Das bestimmte Integral	165
V.1.1.	Definition	165
V.1.2.	Elementare Integrationsregeln	166
V.1.3.	Differentiation und Integration	167
V.2.	Integrationsregeln	170
V.2.1.	Linearität	170
V.2.2.	Partielle Integration	171
V.2.3.	Substitutionsregel	174
V.2.4.	Symmetrien	178
V.3.	Integration rationaler Funktionen	179
V.3.1.	Partialbruchzerlegung	179
V.3.2.	Integration	183
V.3.3.	Verallgemeinerte rationale Funktionen	185
V.4.	Uneigentliche Integrale	189
V.4.1.	Definition	189
V.4.2.	Ein Konvergenzkriterium	191
V.4.3.	Die Eulersche Gammafunktion	192
V.5.	Ebene Kurven; Längen- und Flächenmessung	193
V.5.1.	Parameterdarstellung	193

V.5.2.	Tangente und Normale	195
V.5.3.	Bogenlänge	196
V.5.4.	Krümmung und Krümmungskreis	198
V.5.5.	Polardarstellung einer ebenen Kurve	200
V.5.6.	Flächeninhalte	202
V.5.7.	Volumina von Rotationskörpern	204
V.5.8.	Mantelflächen von Rotationskörpern	205
V.6.	Numerische Integration*	205
V.6.1.	Quadraturformeln*	205
V.6.2.	Zusammengesetzte Quadraturformeln*	207
V.6.3.	Romberg Verfahren*	209
Kapitel VI. Gewöhnliche Differentialgleichungen I		
	Skalare Gleichungen	213
VI.1.	Einführung	213
VI.1.1.	Beispiele	213
VI.1.2.	Grundbegriffe	214
VI.1.3.	Geometrische Deutung	215
VI.1.4.	Eindeutigkeitsfragen	216
VI.2.	Differentialgleichungen 1. Ordnung	217
VI.2.1.	Trennung der Variablen	217
VI.2.2.	Variation der Konstanten	221
VI.2.3.	Homogene Differentialgleichungen	225
VI.2.4.	Bernoulli-Differentialgleichung	227
VI.2.5.	Ricatti-Differentialgleichung	228
VI.3.	Differentialgleichungen 2. Ordnung	229
VI.3.1.	Lineare Differentialgleichungen mit konstanten Koeffizienten	229
VI.3.2.	Die homogene Gleichung	229
VI.3.3.	Die inhomogene Gleichung	232
VI.3.4.	Differentialgleichungen vom Typ $y'' = f(x, y')$	237
VI.3.5.	Differentialgleichungen vom Typ $y'' = f(y, y')$	238
VI.4.	Numerische Lösung gewöhnlicher Differentialgleichungen*	239
VI.4.1.	Motivation*	239
VI.4.2.	Einschrittverfahren*	242
VI.4.3.	Runge-Kutta Verfahren*	244
VI.4.4.	Stabilität*	247
	Zusammenfassung	251
	Index	257

Vorbemerkung

In einigen Abschnitten dieses Skriptums stellen wir numerische Verfahren vor, unter anderem zur Lösung linearer und nicht linearer Gleichungssysteme, zur Berechnung von Eigenwerten und Eigenvektoren, zur numerischen Integration und zur numerischen Lösung gewöhnlicher Differentialgleichungen.

Diese Abschnitte werden in der Vorlesung nicht behandelt und sind durch * gekennzeichnet. Sie sollen lediglich einen ersten Überblick über für Ingenieure wichtige numerische Methoden geben und richten sich besonders an solche Studierende, die die Vorlesung „Numerische Mathematik für Maschinenbauer und Umwelttechniker“ nicht hören.

In jener Vorlesung werden die hier vorgestellten numerischen Methoden und weitere ausführlicher behandelt. Ein Skriptum der Vorlesung steht unter der Adresse

www.ruhr-uni-bochum.de/num1/skripte/index.html

zur Verfügung.

Zu den meisten hier vorgestellten Algorithmen geben wir entsprechende Java-Programme an. Unter der Adresse

www.ruhr-uni-bochum.de/num1/demo/index.html

findet man das Java-Applet `Numerics` mitsamt Benutzeranleitung, das diese und einige weitere Algorithmen demonstriert.

KAPITEL I

Zahlen und Vektoren

I.1. Mengen und Abbildungen

I.1.1. Mengen. Eine grundlegende Fähigkeit des menschlichen Geistes ist es, Objekte zu einem Ganzen zusammenfassen zu können. So fassen wir die Einwohner Bochums zu einem Ganzen zusammen, das wir die Bevölkerung Bochums nennen; die unter deutscher Flagge fahrenden Handelsschiffe fassen wir zur deutschen Handelsflotte zusammen; die Äpfel in einem Korb zu einem „Korb Äpfel“ usw. Ein solches Ganzes nennen wir eine MENGE; die zu einer Menge zusammengefassten Objekte bilden die ELEMENTE dieser Menge.

Mengen werden mit Großbuchstaben bezeichnet. Man schreibt $a \in A$ bzw. $a \notin A$, wenn das Objekt a ein bzw. kein Element der Menge A ist.

$A = \{a_1, a_2, \dots, a_n\}$ kennzeichnet eine Menge mit endlich vielen Elementen. Hierbei kommt es nicht auf die Reihenfolge der Elemente an.

Häufig beschreibt man eine Menge A durch die Angabe einer Eigenschaft E , die genau allen Elementen von A zukommt:

$$A = \{a : a \text{ hat die Eigenschaft } E\}$$

bzw.

$$A = \{a \in X : a \text{ hat die Eigenschaft } E\},$$

wenn besonders hervorgehoben werden soll, dass die Elemente von A in der Grundmenge X enthalten sind.

\emptyset bezeichnet die LEERE MENGE, die keine Elemente enthält.

B heißt TEILMENGE von A , kurz $B \subset A$, wenn jedes Element von B auch in der Menge A enthalten ist. Die Beziehung $B \subset A$ heißt INKLUSION.

Zwei Mengen A und B sind GLEICH, kurz $A = B$, wenn beide Inklusionen $A \subset B$ und $B \subset A$ gelten.

I.1.2. Mengenoperationen. Es gibt folgende Mengenoperationen:

$A \cap B = \{x : x \in A \text{ und } x \in B\}$	DURCHSCHNITT
$A \cup B = \{x : x \in A \text{ oder } x \in B\}$	VEREINIGUNG

$A \setminus B = \{x : x \in A \text{ und } x \notin B\}$	KOMPLEMENT
$A \times B = \{(a, b) : a \in A, b \in B\}$	PRODUKT

BEISPIEL I.1.1. Für

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

$$B = \{3, 4\}$$

erhalten wir

$$A \cap B = \{3\},$$

$$A \cup B = \{1, 2, 3, 4\},$$

$$A \setminus B = \{1, 2\},$$

$$A \times B = \{(1, 3), (1, 4), (2, 3), (2, 4), (3, 3), (3, 4)\}.$$

Zwei Mengen A und B heißen DISJUNKT, wenn $A \cap B = \emptyset$ ist.

I.1.3. Abbildungen. Eine **ABBILDUNG** oder **FUNKTION** von einer Menge A nach einer Menge B , kurz

$$f : A \rightarrow B,$$

ist eine Vorschrift, die jedem Element $a \in A$ genau ein Element $f(a) \in B$ zuordnet. Man nennt A den **DEFINITIONSBEREICH** und die Elemente von A die **ARGUMENTE** der Abbildung f ; $f(a)$ heißt das **BILD** von a unter f . Für eine Teilmenge C von A heißt die Menge

$$f(C) = \{f(c) : c \in C\}$$

das **BILD** von C unter f ; $f(A)$ heißt der **WERTEBEREICH** der Abbildung.

Ist $f(A) = B$, heißt die Abbildung **SURJEKTIV**.

Haben verschiedene Argumente stets verschiedene Bilder, d.h. $f(a) \neq f(a')$ für alle $a \neq a'$, so heißt die Abbildung **INJEKTIV**.

Ist eine Abbildung injektiv und surjektiv, heißt sie **BIJEKTIV**.

Zu einer bijektiven Abbildung $f : A \rightarrow B$ gibt es eine eindeutig bestimmte **UMKEHRABBILDUNG** $f^{-1} : B \rightarrow A$ mit der Eigenschaft

$$f^{-1}(b) = a \iff f(a) = b \text{ für alle } a \in A, b \in B.$$

Sind $f : A \rightarrow B$ und $g : C \rightarrow D$ zwei Abbildungen mit $f(A) \subset C$, so bezeichnet $g \circ f : A \rightarrow D$ die Abbildung, die jedem $a \in A$ das Bild $g(f(a))$ zuordnet; sie heißt **KOMPOSITION** von f und g .

Zwei Abbildungen $f : A \rightarrow B$ und $g : C \rightarrow D$ sind genau dann **GLEICH**, kurz $f = g$, wenn gilt $A = C$, $B = D$ und $f(a) = g(a)$ für jedes $a \in A$.

I.2. Die reellen Zahlen

I.2.1. Bezeichnungen. Wir verwenden die folgenden Bezeichnungen für Zahlenmengen:

$\mathbb{N} = \{0, 1, 2, \dots\}$	NATÜRLICHE ZAHLEN
$\mathbb{N}^* = \{1, 2, \dots\} = \mathbb{N} \setminus \{0\}$	
$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$	GANZE ZAHLEN
$\mathbb{Q} = \left\{ \frac{z}{n} : z \in \mathbb{Z}, n \in \mathbb{N}^* \right\}$	RATIONALE ZAHLEN
\mathbb{R}	REELLE ZAHLEN

Jede reelle Zahl kann als unendlicher Dezimalbruch dargestellt werden. Die rationalen Zahlen sind genau die reellen Zahlen, die eine abbrechende oder periodische Dezimalbruchdarstellung haben, z.B.

$$\begin{aligned}
 1.414 &= \frac{1414}{1000} \\
 &= \frac{707}{500}, \\
 0.\overline{314} &= \frac{314}{999}, \\
 0.17\overline{314} &= \frac{17}{100} + \frac{314}{100 \cdot 999} \\
 &= \frac{17297}{99900}.
 \end{aligned}$$

Reelle Zahlen, die nicht rational sind, heißen **IRRATIONAL**. Beispiele für irrationale Zahlen sind $\sqrt{2} = 1.4142\dots$ und $\pi = 3.14159\dots$ das Verhältnis von Kreisumfang zu Kreisdurchmesser.

Zur Veranschaulichung von \mathbb{R} benutzt man auch die von $-$ nach $+$ orientierte **ZAHLENGRADE**, auf der jeder Punkt einer Zahl entspricht.

I.2.2. Ungleichungen. Für je zwei reelle Zahlen a, b gilt stets genau eine der folgenden Beziehungen:

$a < b$	a ist kleiner als b
$a = b$	a ist gleich b
$a > b$	a ist größer als b

Es gelten die folgenden Rechenregeln:

$a < b$ und $b < c$	$\implies a < c$
$a < b$ und $c \in \mathbb{R}$	$\implies a + c < b + c$

$$\begin{array}{ll}
a < b \text{ und } c > 0 & \implies ac < bc \\
a < b \text{ und } c < d & \implies a + c < b + d \\
a < b \text{ und } c < 0 & \implies bc < ac \\
a \neq 0 & \implies a^2 > 0 \\
a > 0 & \implies \frac{1}{a} > 0 \\
b > a > 0 & \implies 0 < \frac{1}{b} < \frac{1}{a} \\
0 < a < b & \implies 0 < a^n < b^n \text{ für alle } n \in \mathbb{N}^* \\
0 < a < b & \implies 0 < a^{1/n} < b^{1/n} \text{ für alle } n \in \mathbb{N}^* \\
1 < a & \implies 0 < a^m < a^n \text{ für alle } m, n \in \mathbb{Z} \text{ mit } m < n.
\end{array}$$

Neben den Ungleichungszeichen $<$ und $>$ benutzt man noch die Zeichen \leq und \geq :

$$\begin{array}{l}
a \leq b \iff a < b \text{ oder } a = b \\
a \geq b \iff a > b \text{ oder } a = b
\end{array}$$

Es gelten folgende Rechenregeln

$$\begin{array}{ll}
a \leq b \text{ und } b \leq c & \implies a \leq c \\
a \leq b \text{ und } c \in \mathbb{R} & \implies a + c \leq b + c \\
a \leq b \text{ und } c \geq 0 & \implies ac \leq bc \\
a \leq b \text{ und } c \leq 0 & \implies bc \leq ac \\
a \leq b \text{ und } b \leq a & \implies a = b
\end{array}$$

BEISPIEL I.2.1. Für alle reellen Zahlen $a \geq 0$, $b \geq 0$ gilt

$$\sqrt{ab} \leq \frac{1}{2}(a + b).$$

DENN: Es ist $(a - b)^2 \geq 0$. Hieraus folgt

$$\begin{array}{l}
0 \leq (a - b)^2 \\
= a^2 - 2ab + b^2 \quad | +4ab \\
\implies 4ab \leq a^2 + 2ab + b^2 \\
= (a + b)^2 \\
\implies 2\sqrt{ab} \leq a + b
\end{array}$$

$$\implies \sqrt{ab} \leq \frac{1}{2}(a+b).$$

I.2.3. Intervalle. Mit Hilfe der Ungleichungen können wir Intervalle definieren:

$$\begin{aligned} [a, b] &= \{x \in \mathbb{R} : a \leq x \leq b\} \text{ abgeschlossenes Intervall} \\ (a, b) &= \{x \in \mathbb{R} : a < x < b\} \text{ offenes Intervall} \\ [a, b) &= \{x \in \mathbb{R} : a \leq x < b\} \text{ halboffenes Intervall} \\ (a, b] &= \{x \in \mathbb{R} : a < x \leq b\} \text{ halboffenes Intervall} \end{aligned}$$

Zur Vermeidung von Fallunterscheidungen ist es günstig, die Symbole $-\infty$ (MINUS UNENDLICH) und ∞ (PLUS UNENDLICH) einzuführen und $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ festzulegen. Dann ist

$$\begin{aligned} (-\infty, a) &= \{x \in \mathbb{R} : x < a\} \\ (-\infty, a] &= \{x \in \mathbb{R} : x \leq a\} \\ [a, \infty) &= \{x \in \mathbb{R} : x \geq a\} \\ (a, \infty) &= \{x \in \mathbb{R} : x > a\}. \end{aligned}$$

I.2.4. Schranken. Eine Menge S reeller Zahlen heißt NACH OBEN BESCHRÄNKT, wenn es eine reelle Zahl b gibt mit $S \subset (-\infty, b]$. In diesem Fall heißt b eine OBERE SCHRANKE von S . Analog heißt S nach UNTEN BESCHRÄNKT, wenn es eine reelle Zahl a gibt mit $S \subset [a, \infty)$. a heißt dann eine UNTERE SCHRANKE von S . Eine nach unten *und* nach oben beschränkte Menge heißt BESCHRÄNKT.

Eine wesentliche Grundannahme über die reellen Zahlen ist das

VOLLSTÄNDIGKEITSAXIOM: Jede nach oben beschränkte Menge S reeller Zahlen besitzt eine kleinste obere Schranke s .

Diese kleinste obere Schranke s heißt das SUPREMUM von S , kurz $s = \sup S$. $\sup S$ muss nicht zu der Menge S gehören. Zu jeder noch so kleinen Zahl $\varepsilon > 0$ gibt es aber stets ein $x \in S$ mit $\sup S - \varepsilon < x \leq \sup S$.

Aus dem Vollständigkeitsaxiom folgt, dass jede nach unten beschränkte Menge S reeller Zahlen eine größte untere Schranke r besitzt. Diese heißt INFIMUM von S , kurz $r = \inf S$.

Es gilt die Beziehung

$$\inf S = -\sup\{-s : s \in S\}.$$

BEISPIEL I.2.2. $\sup\{x \in \mathbb{R} : x^2 \leq 2\} = \sqrt{2}$, $\inf\{x \in \mathbb{R} : x^2 \geq 2\} = \sqrt{2}$, $\inf\{\frac{1}{n} : n \in \mathbb{N}^*\} = 0$.

I.2.5. Der Betrag. Der BETRAG $|a|$ einer reellen Zahl a ist definiert durch

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

Es gelten folgende Rechenregeln:

$$\begin{aligned} -|a| &\leq a \leq |a| \\ |-a| &= |a| \\ |a| &\geq 0 \\ |a| = 0 &\iff a = 0 \\ |a+b| &\leq |a| + |b| \quad \text{DREIECKSUNGLEICHUNG} \\ |a-b| &\geq ||a| - |b|| \\ |ab| &= |a||b| \\ \left|\frac{a}{b}\right| &= \frac{|a|}{|b|} \quad \text{falls } b \neq 0. \end{aligned}$$

$|a-b|$ ist der Abstand der zu a und b gehörenden Punkte auf der Zahlengeraden. Also gilt für alle $a, x \in \mathbb{R}$ und $\varepsilon > 0$

$$|a-x| \leq \varepsilon \iff a - \varepsilon \leq x \leq a + \varepsilon.$$

Wird für eine Zahl a eine Näherung \bar{a} angegeben, so bedeutet die Schreibweise $a = \bar{a} \pm \varepsilon$ dasselbe wie $|a - \bar{a}| \leq \varepsilon$.

BEISPIEL I.2.3. $\bar{a} = 1.414215$ ist ein Näherungswert für $a = \sqrt{2}$. Es ist $\bar{a}^2 \geq 2$, folglich $\frac{2}{\bar{a}} \leq \sqrt{2} \leq \bar{a}$ und damit $|\bar{a} - \sqrt{2}| \leq |\bar{a} - \frac{2}{\bar{a}}| \leq 0.0000003$ oder $\sqrt{2} = 1.414215 \pm 3 \cdot 10^{-6}$.

I.2.6. Vollständige Induktion. Das INDUKTIONSPRINZIP ist eines der wichtigsten Hilfsmittel der Mathematik. Es lautet:

Sei $A(n)$ eine Aussage über die ganze Zahl n . Gelingt es zu zeigen:

- INDUKTIONSANFANG: Die Aussage $A(n_0)$ ist richtig für eine ganze Zahl n_0 ; und

- **INDUKTIONSSCHRITT:** Aus der Annahme, dass die Aussage $A(n)$ richtig ist für ein beliebiges $n \geq n_0$, folgt die Gültigkeit der Aussage $A(n+1)$.

Dann ist die Aussage $A(n)$ für *jede* ganze Zahl $n \geq n_0$ richtig.

BEISPIEL I.2.4. Für alle $n \in \mathbb{N}^*$ gilt

$$\sum_{k=1}^n k = 1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

INDUKTIONSANFANG: $\sum_{k=1}^1 k = 1 = \frac{1 \cdot (1+1)}{2}.$

INDUKTIONSSCHRITT: Annahme: $\sum_{k=1}^n k = \frac{n(n+1)}{2}.$ Dann folgt

$$\begin{aligned} \sum_{k=1}^{n+1} k &= 1 + 2 + \dots + n + (n+1) \\ &= \sum_{k=1}^n k + (n+1) \\ &= \frac{n(n+1)}{2} + n+1 \quad (\text{wegen Induktionsannahme}) \\ &= \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}. \end{aligned}$$

BEISPIEL I.2.5. Für alle $n \in \mathbb{N}^*$ gilt

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

INDUKTIONSANFANG: $\sum_{k=1}^1 k^2 = 1 = \frac{1 \cdot 2 \cdot 3}{6}.$

INDUKTIONSSCHRITT:

$$\begin{aligned} \sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^n k^2 + (n+1)^2 \\ &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \quad (\text{Induktionsannahme}) \end{aligned}$$

$$\begin{aligned}
&= \frac{n+1}{6} [n(2n+1) + 6(n+1)] \\
&= \frac{n+1}{6} [2n^2 + 7n + 6] \\
&= \frac{n+1}{6} (n+2)(2n+3).
\end{aligned}$$

BEISPIEL I.2.6 (GEOMETRISCHE SUMMENFORMEL). Für jede reelle Zahl $q \neq 1$ und jedes $n \in \mathbb{N}$ gilt

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

INDUKTIONSANFANG: $\sum_{k=0}^0 q^k = q^0 = 1 = \frac{1 - q}{1 - q}.$

INDUKTIONSSCHRITT:

$$\begin{aligned}
\sum_{k=0}^{n+1} q^k &= \sum_{k=0}^n q^k + q^{n+1} \\
&= \frac{1 - q^{n+1}}{1 - q} + q^{n+1} \quad (\text{wegen Induktionsannahme}) \\
&= \frac{1 - q^{n+1} + q^{n+1}(1 - q)}{1 - q} \\
&= \frac{1 - q^{n+2}}{1 - q}.
\end{aligned}$$

BEISPIEL I.2.7 (BERNOULLI-UNGLEICHUNG). Für alle reellen Zahlen $x \geq -1$ und alle $n \in \mathbb{N}^*$ gilt

$$(1 + x)^n \geq 1 + nx.$$

INDUKTIONSANFANG: $(1 + x)^1 = 1 + x \geq 1 + 1 \cdot x.$

INDUKTIONSSCHRITT:

$$\begin{aligned}
(1 + x)^{n+1} &= \underbrace{(1 + x)^n}_{\geq 1 + nx} \underbrace{(1 + x)}_{\geq 0} \\
&\geq (1 + nx)(1 + x) \quad (\text{wegen Induktionsannahme}) \\
&= 1 + (n + 1)x + \underbrace{nx^2}_{\geq 0} \\
&\geq 1 + (n + 1)x.
\end{aligned}$$

I.2.7. Rekursive Definition. Diese Methode beruht auf dem Induktionsprinzip:

Eine Größe A_n wird für $n \in \mathbb{N}$ dadurch definiert, dass man

- A_0 festlegt und
- A_k für $k \leq n$ als bekannt ansieht und A_{n+1} durch diese ausdrückt.

BEISPIEL I.2.8. Die POTENZEN a^n werden für beliebiges $a \in \mathbb{R}$ und $n \in \mathbb{N}$ rekursiv definiert durch $a^0 = 1$ und $a^{n+1} = a \cdot a^n$.

BEISPIEL I.2.9 (FAKULTÄTEN). $n!$ (sprich n Fakultät) wird für $n \in \mathbb{N}$ rekursiv definiert durch $0! = 1$ und $(n+1)! = (n+1) \cdot n!$. Es ist

$$n! = \begin{cases} 1 \cdot 2 \cdot \dots \cdot n & \text{für } n \geq 1, \\ 1 & \text{für } n = 0. \end{cases}$$

I.2.8. Binomialkoeffizienten und Binomische Formel. Für zwei natürliche Zahlen n, k mit $k \leq n$ bezeichnet man die Zahl

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{1 \cdot \dots \cdot k}$$

als BINOMIALKOEFFIZIENT.

Es gilt:

Eine n elementige Menge besitzt genau $\binom{n}{k}$ k elementige Teilmengen. Mit anderen Worten: Es gibt genau $\binom{n}{k}$ Möglichkeiten aus n Elementen k verschiedene Elemente auszuwählen.

BEISPIEL I.2.10. Beim Zahlenlotto „6 aus 49“ gibt es

$$\binom{49}{6} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 13983816$$

Möglichkeiten 6 „Richtige“ zu ziehen.

Für die Binomialkoeffizienten gelten folgende Rechenregeln:

$$\begin{aligned} \binom{n}{0} &= \binom{n}{n} = 1, \\ \binom{n}{k} &= \binom{n}{n-k}, \\ \binom{n+1}{k} &= \binom{n}{k-1} + \binom{n}{k}. \end{aligned}$$

Weiter gilt:

BINOMISCHE FORMEL: Für alle reellen Zahlen x, y und alle natürlichen Zahlen n ist

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

BEISPIEL I.2.11. Aus der Binomischen Formel folgt

$$2^n = (1 + 1)^n = \sum_{k=0}^n \binom{n}{k}.$$

I.3. Die Ebene

I.3.1. Kartesische Koordinatensysteme. In einer Ebene E entsteht ein KARTESISCHES KOORDINATENSYSTEM durch die Vorgabe eines Punktes 0 und zweier aufeinander senkrecht stehender Zahlengeraden, der x - und der y -ACHSE, deren Ursprung jeweils in 0 liegt. Dabei muss die y -Achse durch eine Drehung in positiver Richtung, d.h. gegen den Uhrzeigersinn, aus der x -Achse hervorgehen. Fällt man für einen beliebigen Punkt $P_0 \in E$ die Lote auf die Achsen, so bestimmen die beiden Fußpunkte die x - bzw. y -Koordinate x_0 bzw. y_0 von P_0 und man schreibt $P_0 = (x_0, y_0)$. Der Punkt $0 = (0, 0)$ heißt der NULLPUNKT oder URSPRUNG des Koordinatensystems.

Nach Festlegen eines kartesischen Koordinatensystems gibt es zu jedem Zahlenpaar $(x, y) \in \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ genau einen Punkt $X \in E$ mit $X = (x, y)$ und umgekehrt. Man kann daher Teilmengen von \mathbb{R}^2 als Punktmenge in E veranschaulichen und umgekehrt geometrische Gebilde wie Kurven oder Gebiete in E durch Funktionen, Gleichungen oder Ungleichungen beschreiben.

BEISPIEL I.3.1 (GRAPH EINER FUNKTION). Sei $I \in \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion. Die Punktmenge

$$G_f = \{(x, y) : x \in I, y = f(x)\}$$

in der mit einem kartesischen Koordinatensystem versehenen Ebene heißt Graph der Funktion.

BEISPIEL I.3.2. Nach Wahl eines kartesischen Koordinatensystems betrachten wir Punktmenge $C \subset E$ der Form

$$C = \{(x, y) : F(x, y) = 0\}$$

mit einer gegebenen Funktion $F : \mathbb{R}^2 \rightarrow \mathbb{R}$. Wir sagen in diesem Fall: Die Gleichung $F(x, y) = 0$ beschreibt die Menge C . Man beachte dabei,

dass verschiedene Gleichungen die selbe Menge beschreiben können.

(1) Die Gleichung

$$(b_2 - a_2)(x - a_1) - (b_1 - a_1)(y - a_2) = 0$$

beschreibt eine GERADE durch die Punkte $A = (a_1, a_2)$ und $B = (b_1, b_2)$.

(2) Die Gleichung

$$(x - a_1)^2 + (y - a_2)^2 - r^2 = 0$$

beschreibt einen KREIS mit Mittelpunkt $A = (a_1, a_2)$ und Radius r .

(3) Die Gleichungen

$$\begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 &= 0, \\ \frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 &= 0, \\ y^2 - 2px &= 0 \end{aligned}$$

mit $a > 0$, $b > 0$, $p \neq 0$ beschreiben eine ELLIPSE, eine HYPERBEL und eine PARABEL.

BEISPIEL I.3.3. (1) Die Halbebene unterhalb bzw. oberhalb der Geraden $x - y = 0$ wird durch die Ungleichung $x > y$ bzw. $x < y$ beschrieben. Soll die Gerade zu der Halbebene hinzugehören, muss $>$ bzw. $<$ durch \geq bzw. \leq ersetzt werden.

(2) Die Ungleichung $x^2 + y^2 < 1$ bzw. $x^2 + y^2 > 1$ beschreibt das Innere bzw. das Äußere des Kreises um den Nullpunkt mit Radius 1.

I.3.2. Winkel. Der Winkel α entstehe durch Drehung eines Zeigers um einen Punkt der Ebene. Die Länge des zugehörigen Einheitskreisbogens sei l . Wir nennen l bzw. $-l$ das **BOGENMASS** von α und schreiben $\alpha = l$ bzw. $\alpha = -l$, wenn die Drehung im positiven Sinn, d.h. gegen den Uhrzeiger, bzw. im negativen Sinn, d.h. mit dem Uhrzeiger, erfolgt. Ein Winkel von α° besitzt bei positiver Drehung das Bogenmaß $\frac{\pi}{180}\alpha$.

I.3.3. Sinus, Cosinus. Wird in der mit kartesischen Koordinaten versehenen Ebene der vom Ursprung zum Punkt $(1, 0)$ weisende Zeiger um den Winkel α gedreht, dann bewegt sich die Spitze auf dem Rand des Einheitskreises um den Ursprung von dem Punkt $(1, 0)$ auf einen Punkt P , dessen Koordinaten mit $\cos \alpha$ und $\sin \alpha$ bezeichnet werden: $P = (\cos \alpha, \sin \alpha)$. Die so für alle $\alpha \in \mathbb{R}$ definierten Funktionen $\alpha \mapsto \cos \alpha$ und $\alpha \mapsto \sin \alpha$ heißen **COSINUS** und **SINUS**. Für alle $\alpha \in \mathbb{R}$ gilt

$$(\cos \alpha)^2 + (\sin \alpha)^2 = 1.$$

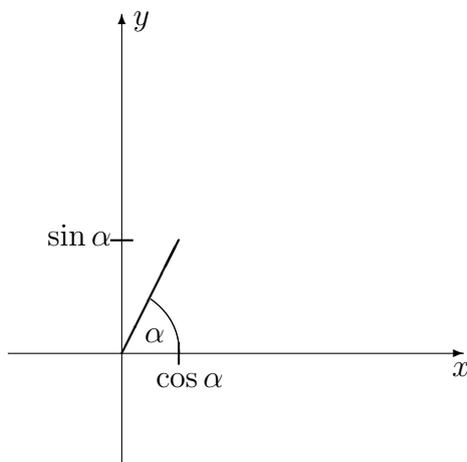


ABBILDUNG I.3.1. Sinus und Cosinus

TABELLE I.3.1. Einige Werte von Sinus und Cosinus

α	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos \alpha$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\sin \alpha$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1

Wegen dieser Eigenschaft und obiger Definition lassen sich leicht die in Tabelle I.3.1 angegebenen Funktionswerte bestimmen.

Weiter gilt für alle α

$$\begin{aligned} \cos(\pi + \alpha) &= -\cos \alpha, & \sin(\pi + \alpha) &= -\sin \alpha, \\ \cos(2\pi + \alpha) &= \cos \alpha, & \sin(2\pi + \alpha) &= \sin \alpha. \end{aligned}$$

I.3.4. Drehungen. Das kartesische Koordinatensystem (x', y') gehe aus dem (x, y) -System durch eine Drehung um den Ursprung mit dem Winkel α hervor, d.h. man dreht die Koordinatenachsen um den Winkel α . Hat ein Punkt P im ursprünglichen System die Koordinaten (x, y) und im neuen System die Koordinaten (x', y') , so gelten die folgenden TRANSFORMATIONSFORMELN (vgl. Abbildung I.3.2):

$$\begin{array}{ll} x = x' \cos \alpha - y' \sin \alpha & x' = x \cos \alpha + y \sin \alpha \\ y = x' \sin \alpha + y' \cos \alpha & y' = -x \sin \alpha + y \cos \alpha \\ P \text{ im ursprünglichen System} & P \text{ im gedrehten System} \end{array}$$

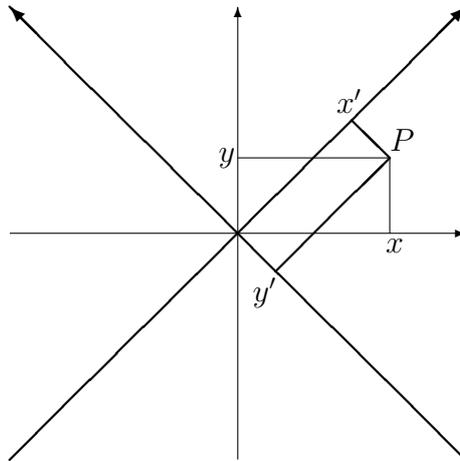


ABBILDUNG I.3.2. Drehung eines kartesischen Koordinatensystems

BEISPIEL I.3.4. Betrachte die Drehung um $\alpha = \frac{\pi}{6}$. Wegen $\cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}$ und $\sin \frac{\pi}{6} = \frac{1}{2}$ gelten die Transformationsformeln

$$\begin{aligned} x &= \frac{\sqrt{3}}{2}x' - \frac{1}{2}y' & x' &= \frac{\sqrt{3}}{2}x + \frac{1}{2}y, \\ y &= \frac{1}{2}x' + \frac{\sqrt{3}}{2}y' & y' &= -\frac{1}{2}x + \frac{\sqrt{3}}{2}y. \end{aligned}$$

Speziell hat der Punkt $M = (2, 3)$ die neuen Koordinaten $M = (\sqrt{3} + \frac{3}{2}, -1 + \frac{3}{2}\sqrt{3})$. Daher besitzt der Kreis $(x - 2)^2 + (y - 3)^2 = 25$ um M mit Radius 5 im neuen System die Darstellung

$$\left(x' - \sqrt{3} - \frac{3}{2}\right)^2 + \left(y' + 1 - \frac{3}{2}\sqrt{3}\right)^2 = 25.$$

Für die Parabel $y = x^2$ ergibt sich im neuen System die Darstellung

$$\frac{1}{2}x' + \frac{\sqrt{3}}{2}y' - \left(\frac{\sqrt{3}}{2}x' - \frac{1}{2}y'\right)^2 = 0$$

bzw. nach „Ausquadrieren“

$$-\frac{3}{4}x'^2 + \frac{\sqrt{3}}{2}x'y' - \frac{1}{4}y'^2 + \frac{1}{2}x' + \frac{\sqrt{3}}{2}y' = 0.$$

Wir halten nun das kartesische Koordinatensystem fest und bilden jeden Punkt $X = (x, y)$ durch eine Drehung um den Ursprung mit dem Winkel α auf den Punkt $X' = (x', y')$ ab. Dann gilt folgende Transformationsformel:

$$\begin{aligned} x' &= x \cos \alpha - y \sin \alpha \\ y' &= x \sin \alpha + y \cos \alpha \end{aligned}$$

BEISPIEL I.3.5. Durch Drehung um $\alpha = \frac{\pi}{6}$ geht der Punkt $M = (2, 3)$ über in den Punkt $M' = (\sqrt{3} - \frac{3}{2}, 1 + \frac{3}{2}\sqrt{3})$. Der Kreis $(x-2)^2 + (y-3)^2 = 25$ geht über in den Kreis $(x - \sqrt{3} + \frac{\sqrt{3}}{2})^2 + (y - 1 - \frac{3}{2}\sqrt{3})^2 = 25$. Die Parabel C mit der Darstellung $y = x^2$ geht über in die Parabel C' mit der Darstellung

$$-\frac{1}{2}x + \frac{\sqrt{3}}{2}y = \left[\frac{\sqrt{3}}{2}x + \frac{1}{2}y\right]^2$$

bzw. nach „Ausquadrieren“

$$-\frac{1}{2}x + \frac{\sqrt{3}}{2}y - \frac{3}{4}x^2 - \frac{\sqrt{3}}{2}xy - \frac{1}{4}y^2 = 0.$$

I.4. Vektoren

I.4.1. Vektoren. Zu je zwei Punkten P und Q gibt es genau eine Parallelverschiebung des Raumes, die den Punkt P in den Punkt Q abbildet. Diese Parallelverschiebung wird mit \overrightarrow{PQ} bezeichnet und heißt VEKTOR von P nach Q . Der Vektor \overrightarrow{PQ} wird dargestellt durch einen Pfeil, der von P nach Q zeigt; seine Länge ist der Abstand der Punkte P und Q . Wird unter \overrightarrow{PQ} ein Punkt R in einen Punkt S verschoben, dann hat offenbar \overrightarrow{RS} die gleiche Wirkung wie \overrightarrow{PQ} , d.h. $\overrightarrow{PQ} = \overrightarrow{RS}$. Zwei gleich lange und gleich gerichtete Pfeile im Raum stellen somit denselben Vektor dar. Statt „der Pfeil \vec{a} stellt einen Vektor dar“ sagt man kurz „ \vec{a} ist ein Vektor“ und berücksichtigt, dass \vec{a} im Raum frei parallel verschoben werden kann und nicht an einen Punkt gebunden ist.

Den zu \vec{a} gleich langen, aber entgegen gesetzten Vektor bezeichnen wir mit $-\vec{a}$; er macht die durch \vec{a} bewirkte Parallelverschiebung rückgängig.

Der NULLVEKTOR $\vec{0}$ bezeichnet die „Verschiebung“ des Raumes, bei der gar nichts bewegt wird.

I.4.2. Addition von Vektoren. Führt man zwei Parallelverschiebungen $\vec{a} = \overrightarrow{PQ}$ und $\vec{b} = \overrightarrow{QR}$ nacheinander aus, ergibt sich wieder eine Parallelverschiebung, nämlich $\vec{c} = \overrightarrow{PR}$. Man nennt \vec{c} die Summe von \vec{a} und \vec{b} und schreibt $\vec{c} = \vec{a} + \vec{b}$. Haben die Pfeile \vec{a} und \vec{b} gleichen Anfangspunkt, so gewinnt man $\vec{a} + \vec{b}$ aus der in Abbildung I.4.1 dargestellten Parallelogrammregel.

Es gelten folgende Rechenregeln:

$\vec{a} + \vec{0} = \vec{a}$ $\vec{a} + (-\vec{a}) = \vec{0}$ $\vec{a} + \vec{b} = \vec{b} + \vec{a} \qquad \text{KOMMUTATIVGESETZ}$

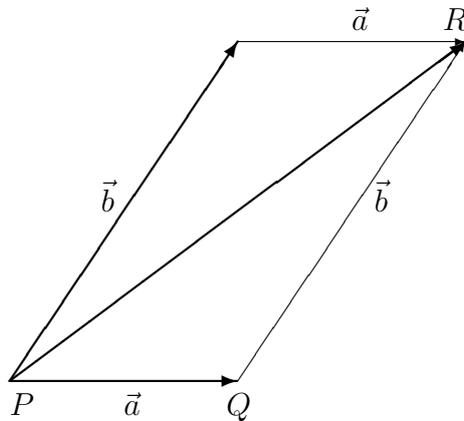


ABBILDUNG I.4.1. Parallelogrammregel

$$\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c} \quad \text{ASSOZIATIVGESETZ.}$$

I.4.3. Skalare Vielfache von Vektoren. Zu einer reellen Zahl $\alpha \geq 0$ und einem Vektor \vec{a} bezeichnet $\alpha\vec{a}$ denjenigen Vektor, der dieselbe Richtung hat wie \vec{a} , aber die α -fache Länge. Im Fall $\alpha < 0$ setzt man $\alpha\vec{a} = -(|\alpha|\vec{a})$. Insbesondere ist $0\vec{a} = \vec{0}$ und $\alpha\vec{0} = \vec{0}$ für jeden Vektor \vec{a} und jede Zahl α . Es gelten folgende Rechenregeln:

$$\begin{aligned} \alpha(\beta\vec{a}) &= (\alpha\beta)\vec{a} \\ \alpha(\vec{a} + \vec{b}) &= (\alpha\vec{a}) + (\alpha\vec{b}) \\ (\alpha + \beta)\vec{a} &= (\alpha\vec{a}) + (\beta\vec{a}). \end{aligned}$$

I.4.4. Der Betrag. Die Länge eines Vektors \vec{a} , das ist für $\vec{a} = \overrightarrow{PQ}$ die Länge der Strecke von P nach Q , nennt man seinen BETRAG und schreibt $|\vec{a}|$. Es ist $|\vec{0}| = 0$. Es gelten folgende Eigenschaften:

$$\begin{aligned} |\alpha\vec{a}| &= |\alpha||\vec{a}| \quad \text{insb. } |-\vec{a}| = |\vec{a}| \\ |\vec{a} + \vec{b}| &\leq |\vec{a}| + |\vec{b}| \quad \text{DREIECKSUNGLEICHUNG.} \end{aligned}$$

I.4.5. Der Winkel zwischen zwei Vektoren. Trägt man zwei von $\vec{0}$ verschiedene Vektoren \vec{a} und \vec{b} von einem Punkt P aus ab, so bezeichnet man den kleineren der beiden positiv gemessenen Winkel, den die Pfeile \vec{a} und \vec{b} im Scheitel P bilden, als den WINKEL zwischen

\vec{a} und \vec{b} , kurz $\angle(\vec{a}, \vec{b})$. Insbesondere ist also $0 \leq \angle(\vec{a}, \vec{b}) \leq \pi$, und es gelten die Regeln:

$$\begin{aligned}\angle(\vec{a}, \vec{b}) &= \angle(\vec{b}, \vec{a}), \\ \angle(\vec{a}, t\vec{a}) &= 0 \quad , \text{ falls } t > 0, \\ \angle(\vec{a}, t\vec{a}) &= \pi \quad , \text{ falls } t < 0, \\ \angle(-\vec{a}, \vec{b}) &= \pi - \angle(\vec{a}, \vec{b}).\end{aligned}$$

Man nennt \vec{a} ORTHOGONAL bzw. SENKRECHT zu \vec{b} , kurz $\vec{a} \perp \vec{b}$, wenn $\angle(\vec{a}, \vec{b}) = \frac{\pi}{2}$ ist. Zur Vermeidung von Fallunterscheidungen legt man fest, dass der Nullvektor orthogonal ist zu jedem beliebigen Vektor.

I.4.6. Das Skalarprodukt. Das SKALARPRODUKT $\vec{a} \cdot \vec{b}$ zweier Vektoren \vec{a} und \vec{b} ist definiert durch

$$\vec{a} \cdot \vec{b} = \begin{cases} |\vec{a}||\vec{b}| \cos(\angle(\vec{a}, \vec{b})) & , \text{ falls } \vec{a} \neq 0 \text{ und } \vec{b} \neq 0, \\ 0 & , \text{ falls } \vec{a} = 0 \text{ oder } \vec{b} = 0. \end{cases}$$

Es gelten folgende Rechenregeln:

$$\begin{aligned}\vec{a} \cdot \vec{b} &= \vec{b} \cdot \vec{a} && \text{KOMMUTATIVGESETZ} \\ (\alpha\vec{a}) \cdot \vec{b} &= \alpha(\vec{a} \cdot \vec{b}) \\ &= \alpha(\vec{a} \cdot \vec{b}) \\ (\vec{a} + \vec{b}) \cdot \vec{c} &= (\vec{a} \cdot \vec{c}) + (\vec{b} \cdot \vec{c}) && \text{DISTRIBUTIVGESETZ} \\ \vec{a} \cdot \vec{b} = 0 &\iff \vec{a} \perp \vec{b} && \text{ORTHOGONALITÄTSTEST} \\ |\vec{a}| &= \sqrt{\vec{a} \cdot \vec{a}}.\end{aligned}$$

Da der Wertebereich der Cosinus-Funktion das Intervall $[-1, 1]$ ist, folgt aus der Definition des Skalarproduktes:

$$|\vec{a} \cdot \vec{b}| \leq |\vec{a}||\vec{b}| \quad \text{CAUCHY-SCHWARZSCHE UNGLEICHUNG}$$

BEISPIEL I.4.1 (ORTHOGONALE ZERLEGUNGEN). Kräfte werden in der Physik als Vektoren dargestellt. Bewegt sich ein Massenpunkt unter Einwirkung der Kraft \vec{K} vom Punkt P zum Punkt Q , so ist die dabei geleistete Arbeit

$$A = \vec{K} \cdot \vec{S} = |\vec{K}| \cdot |\vec{S}| \cos \angle(\vec{K}, \vec{S}) \quad \text{mit } \vec{S} = \overrightarrow{PQ}.$$

Stellt man \vec{K} als Summe einer Kraftkomponente $\vec{K}_{\vec{S}}$ in Richtung \vec{S} und einer dazu orthogonalen Komponente $\vec{K}_{\vec{S}}^{\perp}$ dar, so ist A allein durch $\vec{K}_{\vec{S}}$ bestimmt. Man definiert daher allgemein:

ORTHOGONALE ZERLEGUNG von \vec{a} längs \vec{b} , falls $\vec{b} \neq 0$:

$$\vec{a} = \vec{a}_{\vec{b}} + \vec{a}_{\vec{b}}^{\perp}$$

mit den Komponenten

$$\vec{a}_{\vec{b}} = \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|^2} \vec{b}$$

in Richtung \vec{b} und

$$\vec{a}_{\vec{b}}^{\perp} = \vec{a} - \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|^2} \vec{b}$$

orthogonal zu \vec{b} .

I.4.7. Das Vektorprodukt. Das VEKTORPRODUKT $\vec{a} \times \vec{b}$ zweier Vektoren \vec{a} und \vec{b} ist der *Vektor* mit den folgenden Eigenschaften:

- (1) Es ist $\vec{a} \times \vec{b} = 0$, falls $\vec{a} = 0$ oder $\vec{b} = 0$ oder \vec{a} parallel zu \vec{b} ist.
- (2) In allen anderen Fällen ist $\vec{a} \times \vec{b}$ der Vektor, der orthogonal zu \vec{a} und \vec{b} ist, mit dem \vec{a} , \vec{b} , $\vec{a} \times \vec{b}$ ein Rechtssystem bilden und dessen Betrag gleich dem Flächeninhalt des von \vec{a} , \vec{b} aufgespannten Parallelogramms ist.

Es gelten folgende Rechenregeln:

$$\begin{aligned} \vec{a} \times \vec{a} &= 0 \\ \vec{a} \times \vec{b} &= -\vec{b} \times \vec{a} \\ \alpha(\vec{a} \times \vec{b}) &= (\alpha\vec{a}) \times \vec{b} \\ &= \vec{a} \times (\alpha\vec{b}) \\ \vec{a} \times (\vec{b} + \vec{c}) &= (\vec{a} \times \vec{b}) + (\vec{a} \times \vec{c}) \\ (\vec{a} + \vec{b}) \times \vec{c} &= (\vec{a} \times \vec{c}) + (\vec{b} \times \vec{c}) \\ \vec{a} \times \vec{b} = 0 &\iff \vec{a} = \vec{0} \text{ oder } \vec{b} = 0 \text{ oder } \vec{a} \text{ parallel zu } \vec{b} \\ |\vec{a} \times \vec{b}|^2 &= |\vec{a}|^2 |\vec{b}|^2 - (\vec{a} \cdot \vec{b})^2. \end{aligned}$$

Zum Nachweis der letzten Beziehung beachte man, dass aus der Definition des Vektorproduktes folgt $|\vec{a} \times \vec{b}| = |\vec{a}||\vec{b}| \sin(\angle(\vec{a}, \vec{b}))$. Wegen $\sin^2 \alpha + \cos^2 \alpha = 1$ für alle α folgt hieraus durch Quadrieren

$$\begin{aligned} |\vec{a} \times \vec{b}|^2 &= |\vec{a}|^2 |\vec{b}|^2 \underbrace{\sin^2(\angle(\vec{a}, \vec{b}))}_{=1-\cos^2(\angle(\vec{a}, \vec{b}))} \\ &= |\vec{a}|^2 |\vec{b}|^2 - \underbrace{|\vec{a}|^2 |\vec{b}|^2 \cos^2(\angle(\vec{a}, \vec{b}))}_{=(\vec{a} \cdot \vec{b})^2} \\ &= |a|^2 |b|^2 - (\vec{a} \cdot \vec{b})^2. \end{aligned}$$

I.4.8. Das Spatprodukt. Das SPATPRODUKT ist eine Kombination aus Vektorprodukt und Skalarprodukt. Für je drei Vektoren \vec{a} , \vec{b} , \vec{c} ist es definiert durch

$$[\vec{a}, \vec{b}, \vec{c}] = \vec{a} \cdot (\vec{b} \times \vec{c}).$$

Es hat folgende geometrische Bedeutung:

Der von den drei Vektoren \vec{a} , \vec{b} , \vec{c} aufgespannte SPAT (auch PARALLELFLÄCHE oder PARALLELEPIPED genannt) hat das Volumen

$$V = |[\vec{a}, \vec{b}, \vec{c}]|.$$

Aus dieser geometrischen Bedeutung ergeben sich unmittelbar folgende Konsequenzen

Das von den Vektoren \vec{a} , \vec{b} , \vec{c} aufgespannte Tetraeder hat das Volumen $\frac{1}{6}|[\vec{a}, \vec{b}, \vec{c}]|$.

Die Vektoren \vec{a} , \vec{b} , \vec{c} sind nicht parallel zu einer Ebene (man sagt „sie sind linear unabhängig“) genau dann, wenn $[\vec{a}, \vec{b}, \vec{c}] \neq 0$ ist.

Die Vektoren $(\vec{a}, \vec{b}, \vec{c})$ bilden ein Rechtssystem genau dann, wenn $[\vec{a}, \vec{b}, \vec{c}] > 0$ ist.

und Rechenregeln

$$\begin{aligned} [\vec{a}, \vec{b}, \vec{c}] &= [\vec{b}, \vec{c}, \vec{a}] = [\vec{c}, \vec{a}, \vec{b}] \\ [\vec{a}, \vec{b}, \vec{c}] &= -[\vec{b}, \vec{a}, \vec{c}] = -[\vec{c}, \vec{b}, \vec{a}] = -[\vec{a}, \vec{c}, \vec{b}]. \end{aligned}$$

I.4.9. Koordinatendarstellungen. Ein KARTESISCHES KOORDINATENSYSTEM im Raum besteht aus drei sich in einem Punkt 0 rechtwinklig schneidenden Zahlengeraden gleicher Längeneinheit und jeweils dem Nullpunkt im Schnittpunkt 0. 0 heißt der URSPRUNG des Koordinatensystems. Die drei Zahlengeraden bezeichnet man als die x -, y - und z -ACHSE derart, dass sie in dieser Reihenfolge ein Rechtssystem bilden. Die drei durch je zwei Koordinatenachsen bestimmten Ebenen nennt man die KOORDINATENEbenen. Die Koordinaten x_0, y_0, z_0 eines Punkte P_0 sind die Lote des Punktes auf die Koordinatenachsen.

Nach Wahl eines kartesischen Koordinatensystems ist jeder Punkt P des Raumes eindeutig bestimmt durch die Angabe seiner Koordinaten x, y, z . In diesem Sinne kann man bei festem Koordinatensystem den Raum mit der Menge \mathbb{R}^3 identifizieren.

Durch ein kartesisches Koordinatensystem werden drei ausgezeichnete Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ definiert, nämlich die Einheitsvektoren in Richtung der positiven x -, y - bzw. z -Achse. Statt $\vec{e}_1, \vec{e}_2, \vec{e}_3$ schreibt man manchmal auch $\vec{e}_x, \vec{e}_y, \vec{e}_z$ oder $\vec{i}, \vec{j}, \vec{k}$. Wir nennen $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ eine KARTESISCHE BASIS und bezeichnen das Koordinatensystem mit $(0; \vec{e}_1, \vec{e}_2, \vec{e}_3)$.

Der Vektor $\vec{a} = \overrightarrow{0A}$ heißt ORTSVEKTOR des Punktes A . Hat A die Koordinaten (a_1, a_2, a_3) , kann sein Ortsvektor \vec{a} in der Form $\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3$ zerlegt werden. Daher schreibt man bei festgelegtem Koordinatensystem abkürzend

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \iff \vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3 = \overrightarrow{0A}$$

mit $A = (a_1, a_2, a_3)$.

BEMERKUNG I.4.2. Man beachte: (a_1, a_2, a_3) bezeichnet die Koordinaten des Punktes A ; $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ bezeichnet den Ortsvektor $\overrightarrow{0A}$.

Sind $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$ zwei Punkte, so ist $\overrightarrow{PQ} = \overrightarrow{0Q} - \overrightarrow{0P}$. Daher folgt

$$\overrightarrow{PQ} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix} \quad \text{falls } P = (p_1, p_2, p_3), Q = (q_1, q_2, q_3).$$

Für Summe, skalares Vielfaches und Betrag von Vektoren gelten in der Koordinatendarstellung folgende Rechenregeln

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} &= \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix} \\ \alpha \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} &= \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \alpha a_3 \end{pmatrix} \\ \left| \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \right| &= \sqrt{a_1^2 + a_2^2 + a_3^2}. \end{aligned}$$

Konstruktionsgemäß sind die Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ paarweise orthogonal und bilden ein Rechtssystem. Zusammen mit den Rechenregeln für Skalarprodukt und Vektorprodukt ergeben sich hieraus die folgenden Koordinatendarstellungen

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} &= a_1 b_1 + a_2 b_2 + a_3 b_3 \\ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} &= \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}. \end{aligned}$$

Aus der Darstellung des Vektorproduktes lassen sich die folgenden nützlichen Identitäten herleiten

$$\begin{aligned} \vec{a} \times (\vec{b} \times \vec{c}) &= (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c} \\ (\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) &= (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{b} \cdot \vec{c})(\vec{a} \cdot \vec{d}). \end{aligned}$$

Vergleichen wir die erste dieser Identitäten mit der Darstellung von $a_{\vec{b}}^\perp$ in Beispiel I.4.1, so erhalten wir die folgende Formel für die orthogonale Komponente von \vec{a} längs \vec{b} :

$$\vec{a}_{\vec{b}}^\perp = \frac{1}{|\vec{b}|^2} \vec{b} \times (\vec{a} \times \vec{b}).$$

Für das Spatprodukt erhält man schließlich die Darstellung

$$\left[\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right] = a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) \\ + a_3(b_1c_2 - b_2c_1)$$

Dies ist der Wert der Determinante (vgl. Abschnitt II.3.1)

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

BEISPIEL I.4.3. Unter der Kraft

$$\vec{K} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \text{ N}$$

werde der Punkt $P = (2, -2, 1)$ nach $Q = (3, 0, 2)$ verschoben. Die Längeneinheit des Koordinatensystems sei 1m. Die geleistete Arbeit ist das Skalarprodukt der Vektoren \vec{K} und \overrightarrow{PQ}

$$\begin{aligned} A &= \vec{K} \cdot \overrightarrow{PQ} \text{ Nm} \\ &= \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3-2 \\ 0-(-2) \\ 2-1 \end{pmatrix} \text{ Nm} \\ &= (3 - 2 + 2) \text{ Nm} \\ &= 3 \text{ Nm.} \end{aligned}$$

BEISPIEL I.4.4. Das DREHMOMENT einer Einzelkraft \vec{K} , die im Punkt P angreift, ist in Bezug auf den Ursprung gegeben durch $\vec{m}_0 = \overrightarrow{0P} \times \vec{K}$. Für \vec{K} und P wie in Beispiel I.4.3 erhalten wir bei den gleichen Maßeinheiten

$$\begin{aligned} \vec{m}_0 &= \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \text{ Nm} \\ &= \begin{pmatrix} -2 \cdot 2 - 1 \cdot (-1) \\ 1 \cdot 3 - 2 \cdot 2 \\ 2 \cdot (-1) - (-2) \cdot 3 \end{pmatrix} \text{ Nm} \\ &= \begin{pmatrix} -3 \\ -1 \\ 4 \end{pmatrix} \text{ Nm.} \end{aligned}$$

BEISPIEL I.4.5. Der FLÄCHENINHALT EINES DREIECKES ABC ist $F = \frac{1}{2} |\overrightarrow{AB} \times \overrightarrow{AC}|$. Für $A = (1, 2, 3)$, $B = (-2, 0, 4)$, $C = (-1, -1, 2)$

ergibt sich

$$\overrightarrow{AB} = \begin{pmatrix} -3 \\ -2 \\ 1 \end{pmatrix}, \quad \overrightarrow{AC} = \begin{pmatrix} -2 \\ -3 \\ -1 \end{pmatrix}$$

und

$$\begin{aligned} F &= \frac{1}{2} \sqrt{|\overrightarrow{AB}|^2 |\overrightarrow{AC}|^2 - (\overrightarrow{AB} \cdot \overrightarrow{AC})^2} \\ &= \frac{1}{2} \sqrt{14 \cdot 14 - 11^2} \\ &= \frac{1}{2} \sqrt{75} \\ &= \frac{5}{2} \sqrt{3}. \end{aligned}$$

BEISPIEL I.4.6 (NICHTORTHOGONALE ZERLEGUNG). Die drei Stäbe eines Tragbockes führen von der Spitze S in Richtung der drei Vektoren

$$\vec{a} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ m}, \quad \vec{b} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \text{ m}, \quad \vec{c} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ m}.$$

In S zieht die Kraft

$$\vec{K} = 60 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \text{ N}$$

und ruft in den Fußpunkten A, B, C die Lagerkräfte $\vec{K}_A = \alpha \vec{a}$, $\vec{K}_B = \beta \vec{b}$, $\vec{K}_C = \gamma \vec{c}$ hervor. Die Zahlen α, β, γ ergeben sich durch die Gleichgewichtsbedingung

$$\vec{K} + \vec{K}_A + \vec{K}_B + \vec{K}_C = 0.$$

Multipliziert man diese Gleichung der Reihe nach mit $\vec{b} \times \vec{c}$, $\vec{c} \times \vec{a}$ und $\vec{a} \times \vec{b}$ skalar, so verschwinden je zwei Summanden und man erhält

$$\begin{aligned} (\vec{b} \times \vec{c}) \cdot \vec{K} + \alpha (\vec{b} \times \vec{c}) \cdot \vec{a} &= 0 \\ \implies \alpha &= -\frac{[\vec{K}, \vec{b}, \vec{c}]}{[\vec{a}, \vec{b}, \vec{c}]} \\ &= -120 \text{ Nm}^{-1} \end{aligned}$$

$$\begin{aligned} (\vec{c} \times \vec{a}) \cdot \vec{K} + \beta (\vec{c} \times \vec{a}) \cdot \vec{b} &= 0 \\ \implies \beta &= -\frac{[\vec{K}, \vec{c}, \vec{a}]}{[\vec{b}, \vec{c}, \vec{a}]} \\ &= 60\sqrt{2} \text{ Nm}^{-1} \end{aligned}$$

$$(\vec{a} \times \vec{b}) \cdot \vec{K} + \gamma (\vec{a} \times \vec{b}) \cdot \vec{c} = 0$$

$$\begin{aligned} \implies \gamma &= -\frac{[\vec{K}, \vec{a}, \vec{b}]}{[\vec{c}, \vec{a}, \vec{b}]} \\ &= 60\sqrt{2} \text{ Nm}^{-1}. \end{aligned}$$

I.5. Geraden und Ebenen

I.5.1. Die Parameterdarstellung einer Geraden. Wir betrachten eine Gerade g durch einen Punkt A in Richtung eines Vektors \vec{c} , $\vec{c} \neq \vec{0}$. Für jeden Punkt X auf g muss der Vektor \overrightarrow{AX} parallel zu \vec{c} sein, d.h. es muss eine Zahl $t \in \mathbb{R}$ geben mit $\overrightarrow{AX} = t\vec{c}$. Man sagt hierzu, g hat die

PUNKT-RICHTUNGS-GLEICHUNG: $\overrightarrow{AX} = t\vec{c}, t \in \mathbb{R}$.

Die Variable t nennt man einen PARAMETER. Jedem Parameter entspricht genau ein Punkt auf g und umgekehrt. Daher nennt man die obige Darstellung auch eine PARAMETERDARSTELLUNG.

Ist ein zweiter Punkt B auf g gegeben, $B \neq A$, kann man $\vec{c} = \overrightarrow{AB}$ wählen.

Führen wir ein kartesisches Koordinatensystem ein, ergeben sich die folgenden Koordinatendarstellungen der Geraden:

PUNKT-RICHTUNGS-GLEICHUNG:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + t \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

ZWEI-PUNKTE-GLEICHUNG:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + t \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix}$$

BEISPIEL I.5.1. Für die Gerade g

$$\text{durch } A = (1, 0, -1) \text{ in Richtung } \vec{c} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

ergibt sich die Punkt-Richtungs-Gleichung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Die Parameterwerte $t = 1$ und $t = 3$ liefern die Punkte $B = (2, 1, 1)$ und $C = (4, 3, 5)$. Daher hat g auch die Zwei-Punkte-Gleichung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 5 \end{pmatrix} + t \begin{pmatrix} -2 \\ -2 \\ -4 \end{pmatrix}.$$

I.5.2. Die Koordinatengleichungen einer Geraden. Löst man jede der drei Gleichungen der Zwei-Punkte-Gleichung einer Geraden nach dem Parameter t auf und setzt die Ergebnisse gleich, erhält man die KOORDINATENGLEICHUNGEN der Geraden durch A, B :

(A)	$\frac{x_1 - a_1}{b_1 - a_1} = \frac{x_2 - a_2}{b_2 - a_2} = \frac{x_3 - a_3}{b_3 - a_3},$	falls $a_i \neq b_i, i = 1, 2, 3$
(B)	$\frac{x_1 - a_1}{b_1 - a_1} = \frac{x_2 - a_2}{b_2 - a_2}, x_3 = a_3,$	falls $a_3 = b_3$
	$\frac{x_1 - a_1}{b_1 - a_1} = \frac{x_3 - a_3}{b_3 - a_3}, x_2 = a_2,$	falls $a_2 = b_2$
	$\frac{x_2 - a_2}{b_2 - a_2} = \frac{x_3 - a_3}{b_3 - a_3}, x_1 = a_1,$	falls $a_1 = b_1$
(C)	$x_1 = a_1, x_2 = a_2,$	falls $a_1 = b_1, a_2 = b_2, a_3 \neq b_3$
	$x_2 = a_2, x_3 = a_3,$	falls $a_2 = b_2, a_3 = b_3, a_1 \neq b_1$
	$x_3 = a_3, x_1 = a_1,$	falls $a_1 = b_1, a_3 = b_3, a_2 \neq b_2$

BEISPIEL I.5.2. Die Gerade aus Beispiel I.5.1 hat die Koordinatengleichungen

$$x_1 - 1 = x_2 = \frac{x_3 + 1}{2}.$$

Aus den Koordinatengleichungen einer Geraden erhält man mit $t = \frac{x_i - a_i}{b_i - a_i}$ für die Indizes i , für die $b_i \neq a_i$ ist, die Zwei-Punkte-Gleichung in Parameterform zurück.

BEISPIEL I.5.3. Die Gerade mit den Koordinatengleichungen

$$\frac{x_1 - 3}{5} = \frac{x_2 + 1}{7}, x_3 = 4$$

besitzt die Parameterdarstellung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} + t \begin{pmatrix} 5 \\ 7 \\ 0 \end{pmatrix}.$$

I.5.3. Die Momentengleichung einer Geraden. Betrachte die Gerade g durch A in Richtung \vec{c} und einen beliebigen Bezugspunkt P . Sei X ein Punkt auf g . Dann ist $\overrightarrow{AX} = \overrightarrow{PX} - \overrightarrow{PA}$ genau dann parallel zu \vec{c} , wenn $\overrightarrow{AX} \times \vec{c} = 0$ ist, d.h. $\overrightarrow{PX} \times \vec{c} = \overrightarrow{PA} \times \vec{c}$. Dies führt auf:

MOMENTENGLEICHUNG bzgl. des Bezugspunktes P :

$$\overrightarrow{PX} \times \vec{c} = \vec{m}_P \quad \text{mit} \quad \vec{m}_P = \overrightarrow{PA} \times \vec{c}.$$

Gemäß Abschnitt I.4.9 ist

$$\overrightarrow{PS} = \frac{1}{|\vec{c}|^2} \vec{c} \times (\overrightarrow{PA} \times \vec{c}) = \frac{1}{|\vec{c}|^2} \vec{c} \times \vec{m}_P$$

der Lotvektor von P auf g . Daher ist

$$\overrightarrow{PX} = \frac{1}{|\vec{c}|^2} \vec{c} \times \vec{m}_P + t\vec{c}$$

eine Parameterdarstellung der durch obige Momentengleichung gegebenen Geraden.

I.5.4. Abstand Punkt - Gerade. Betrachte die Gerade g durch A in Richtung \vec{c} und einen Punkt P . Das Lot \overrightarrow{PS} von P auf g ist gemäß Abschnitt I.4.9 gegeben durch

$$\overrightarrow{PS} = \frac{1}{|\vec{c}|^2} \vec{c} \times (\overrightarrow{PA} \times \vec{c}).$$

Seine Länge $d = |\overrightarrow{PS}|$ ist der Abstand von P zu g . Da gemäß Abschnitt I.4.7

$$\begin{aligned} |\vec{c} \times (\overrightarrow{PA} \times \vec{c})|^2 &= |\vec{c}|^2 |\overrightarrow{PA} \times \vec{c}|^2 - \underbrace{((\overrightarrow{PA} \times \vec{c}) \cdot \vec{c})^2}_{=0} \\ &= |\vec{c}|^2 |\overrightarrow{PA} \times \vec{c}|^2 \end{aligned}$$

ist, ergibt sich für den Abstand von P zu g

$$d = \frac{|\overrightarrow{PA} \times \vec{c}|}{|\vec{c}|}.$$

BEISPIEL I.5.4. Der Abstand des Punktes $P = (0, 1, 0)$ zu der Geraden g aus Beispiel I.5.1 ist

$$\begin{aligned} d &= \frac{\left| \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right|}{\left| \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right|} \\ &= \frac{\sqrt{3 \cdot 6 - (-2)^2}}{\sqrt{6}} \end{aligned}$$

$$\begin{aligned}
&= \sqrt{\frac{14}{6}} \\
&= \frac{\sqrt{21}}{3}.
\end{aligned}$$

I.5.5. Abstand Gerade - Gerade. Wir wollen den Abstand d der beiden Geraden g_1 durch A in Richtung \vec{u} und g_2 durch B in Richtung \vec{v} bestimmen. Falls die Geraden parallel sind, ist dieser Abstand gleich demjenigen des Punktes B zur Geraden g_1 , und wir können das Ergebnis von Abschnitt I.5.4 ausnutzen. Falls die Geraden nicht parallel sind, ist d der Betrag der Komponente von \overrightarrow{AB} in Richtung des Vektors $\vec{n} = \vec{u} \times \vec{v}$, der auf \vec{u} und \vec{v} senkrecht steht. Indem wir in Beispiel I.5.1 $\vec{a} = \overrightarrow{AB}$ und $\vec{b} = \vec{u} \times \vec{v}$ setzen, erhalten wir

$$d = \frac{|\overrightarrow{AB} \cdot (\vec{u} \times \vec{v})|}{|\vec{u} \times \vec{v}|} = \frac{|[\overrightarrow{AB}, \vec{u}, \vec{v}]|}{|\vec{u} \times \vec{v}|}.$$

Insgesamt liefert dies:

Der Abstand d der Geraden durch A in Richtung \vec{u} zu der Geraden durch B in Richtung \vec{v} ist

$$d = \begin{cases} \frac{|\overrightarrow{AB} \times \vec{u}|}{|\vec{u}|} & \text{falls } \vec{u} \text{ und } \vec{v} \text{ parallel sind,} \\ \frac{|[\overrightarrow{AB}, \vec{u}, \vec{v}]|}{|\vec{u} \times \vec{v}|} & \text{falls } \vec{u} \times \vec{v} \neq 0. \end{cases}$$

BEISPIEL I.5.5. Die Geraden

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

und

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} + t \begin{pmatrix} 5 \\ 7 \\ 0 \end{pmatrix}$$

aus Beispiel I.5.1 und I.5.3 haben den Abstand

$$\begin{aligned}
d &= \frac{\left| \left[\begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 5 \\ 7 \\ 0 \end{pmatrix} \right] \right|}{\left| \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 5 \\ 7 \\ 0 \end{pmatrix} \right|} \\
&= \frac{28}{\sqrt{6 \cdot 74 - 12^2}}
\end{aligned}$$

$$\begin{aligned}
 &= \frac{28}{\sqrt{300}} \\
 &= \frac{14}{15}\sqrt{3}.
 \end{aligned}$$

I.5.6. Die Parameterdarstellung einer Ebene. Wir betrachten die Ebene E durch den Punkt A , die von den beiden von Null verschiedenen und nicht parallelen Vektoren \vec{u} und \vec{v} erzeugt wird. Ein Punkt X liegt genau dann in E , wenn sich der Vektor \overrightarrow{AX} in der Form $\overrightarrow{AX} = s\vec{u} + t\vec{v}$ darstellen lässt. Dies führt auf die

PARAMETERDARSTELLUNG von E :

$$\overrightarrow{AX} = s\vec{u} + t\vec{v} \quad \text{mit } s, t \in \mathbb{R}.$$

Führen wir ein kartesisches Koordinatensystem ein, ergibt sich hieraus die Koordinatendarstellung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + s \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Sind zwei weitere von A verschiedene Punkte B und C in E gegeben, sodass A, B und C nicht auf einer Geraden liegen, so kann man $\vec{u} = \overrightarrow{AB}$ und $\vec{v} = \overrightarrow{AC}$ wählen. Dies führt auf die

DREI-PUNKTE-GLEICHUNG von E :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + s \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix} + t \begin{pmatrix} c_1 - a_1 \\ c_2 - a_2 \\ c_3 - a_3 \end{pmatrix}.$$

I.5.7. Parameterfreie Darstellungen einer Ebene. Wir betrachten die Ebene E durch die drei Punkte A, B, C , die paarweise verschieden sind und die nicht auf einer Geraden liegen. Offensichtlich liegt ein Punkt X genau dann in E , wenn der Vektor \overrightarrow{AX} von den Vektoren \overrightarrow{AB} und \overrightarrow{AC} linear abhängig ist. Wegen Abschnitt I.4.8 ist dies äquivalent zu

$$[\overrightarrow{AX}, \overrightarrow{AB}, \overrightarrow{AC}] = 0.$$

Man nennt diese Darstellung die parameterfreie Drei-Punkte-Formel in DETERMINANTENFORM. Obige Gleichung bedeutet $\overrightarrow{AX} \cdot (\overrightarrow{AB} \times \overrightarrow{AC}) =$

0, d.h. der Vektor $0 \neq \vec{n} = \overrightarrow{AB} \times \overrightarrow{AC}$ steht senkrecht auf E . Man nennt \vec{n} einen **NORMALENVEKTOR** von E und

$$\overrightarrow{AX} \cdot \vec{n} = 0 \quad A \text{ Punkt in } E, \vec{n} \text{ Normalenvektor von } E$$

eine **NORMALENGLEICHUNG** von E . In kartesischen Koordinaten lautet mit

$$A = (a_1, a_2, a_3), \vec{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$$

die Normalengleichung

$$n_1x_1 + n_2x_2 + n_3x_3 = c \quad \text{mit } c = n_1a_1 + n_2a_2 + n_3a_3.$$

Ist insbesondere $|\vec{n}| = 1$, nennt man diese Darstellung auch **HESSE-NORMALFORM** der Ebene.

Wird E durch einen Punkt A und einen Normalenvektor \vec{n} gegeben und ist P ein beliebiger Punkt, dann ist der Lotvektor \overrightarrow{PS} von P auf E gleich der Komponente von \overrightarrow{PA} in Richtung \vec{n} . Wegen Beispiel I.4.1 (S. 26) und $\vec{n} \neq 0$ ist daher

$$\overrightarrow{PS} = \frac{\overrightarrow{PA} \cdot \vec{n}}{|\vec{n}|^2} \vec{n}.$$

Insbesondere ist

$$d = \frac{|\overrightarrow{PA} \cdot \vec{n}|}{|\vec{n}|}$$

der **ABSTAND** von P zur Ebene E durch A senkrecht zu \vec{n} .

BEISPIEL I.5.6. Betrachte die Ebene E durch die Punkte $A = (1, 0, 0)$, $B = (2, 1, 1)$ und $C = (3, 0, 2)$. Die Parameterdarstellung von E lautet

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}.$$

Ein Normalenvektor ist

$$\vec{n} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}.$$

Die Normalengleichung lautet

$$2x_1 - 2x_3 = 2.$$

Der Abstand des Punktes $P = (4, 5, -3)$ zu E ist

$$\begin{aligned} d &= \frac{\left| \begin{pmatrix} 3 \\ 5 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \right|}{\left| \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \right|} \\ &= \frac{12}{\sqrt{8}} \\ &= 3\sqrt{2}. \end{aligned}$$

I.5.8. Die Schnittgerade zweier Ebenen. Zwei nicht parallele Ebenen

$$\begin{aligned} E_1 : & \quad n_1x_1 + n_2x_2 + n_3x_3 = c_1 \\ E_2 : & \quad m_1x_1 + m_2x_2 + m_3x_3 = c_2 \end{aligned}$$

schneiden sich in einer Geraden g . Das Vektorprodukt der beiden Normalenvektoren \vec{n} und \vec{m} ist parallel zu beiden Ebenen und zeigt daher in Richtung von g . Kennt man einen Punkt A auf g , erhält man mit

$$\vec{OX} = \vec{OA} + t(\vec{n} \times \vec{m})$$

sofort eine Parameterdarstellung von g .

BEISPIEL I.5.7. Betrachte die Ebene E_1 aus Beispiel I.5.6 und die Ebene E_2 durch die Punkte $F = (0, 0, 1)$, $G = (-1, 2, 0)$ und $H = (3, -1, 3)$. Ein Normalenvektor von E_1 ist gemäß Beispiel I.5.6

$$\vec{n} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}.$$

Ein Normalenvektor von E_2 ist

$$\vec{m} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \times \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ -5 \end{pmatrix}.$$

Also ist die Schnittgerade g parallel zu

$$\vec{n} \times \vec{m} = \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix}.$$

Zur Bestimmung eines Punktes P auf g stellen wir die Normalengleichungen von E_1 und E_2 auf.

$$\begin{aligned} E_1 : & \quad 2x_1 - 2x_3 = 2 \\ E_2 : & \quad 3x_1 - x_2 - 5x_3 = -5. \end{aligned}$$

Wählen wir $x_3 = 1$, erhalten wir aus der Gleichung für E_1 $x_1 = 1$. Einsetzen dieser Werte in die Gleichung für E_2 liefert $x_2 = 6$. Also lautet die Parameterdarstellung von g

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 1 \end{pmatrix} + t \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix}.$$

I.5.9. Die Winkel zwischen zwei Ebenen und zwischen einer Ebene und einer Geraden. Die Winkel zwischen zwei sich schneidenden Ebenen stimmen überein mit den beiden Winkeln zwischen den zugehörigen Normalenvektoren. Sind also \vec{n} und \vec{m} Normalenvektoren der beiden Ebenen, dann sind ϕ_1 mit

$$\cos \phi_1 = \frac{|\vec{n} \cdot \vec{m}|}{|\vec{n}| |\vec{m}|}, \quad 0 \leq \phi_1 \leq \frac{\pi}{2}$$

und $\phi_2 = \pi - \phi_1$ die beiden Winkel zwischen den Ebenen. In der Regel wird nur ϕ_1 berücksichtigt.

Als Winkel zwischen einer Geraden g in Richtung \vec{c} und einer Ebene E mit Normalenvektor \vec{n} bezeichnet man den Winkel zwischen \vec{c} und der zu \vec{n} orthogonalen Komponente $\frac{1}{|\vec{n}|^2} \vec{n} \times (\vec{c} \times \vec{n})$. Für diesen Winkel ϕ gilt $\phi = \frac{\pi}{2} - \angle(\vec{c}, \vec{n})$. Daher ist der Winkel zwischen g und E bestimmt durch

$$\sin \phi = \frac{|\vec{n} \cdot \vec{c}|}{|\vec{n}| |\vec{c}|}.$$

BEISPIEL I.5.8. Für die beiden Ebenen aus Beispiel I.5.7 erhalten wir den Winkel

$$\begin{aligned} \cos \phi &= \frac{\left| \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -1 \\ -5 \end{pmatrix} \right|}{\left| \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \right| \left| \begin{pmatrix} 3 \\ -1 \\ -5 \end{pmatrix} \right|} \\ &= \frac{16}{\sqrt{8 \cdot 35}} \\ &= \frac{4}{35} \sqrt{70} \\ \implies \phi &\approx 17.0239^\circ. \end{aligned}$$

I.6. Die komplexen Zahlen

I.6.1. Motivation. Die Notwendigkeit, Gleichungen der Form

$$x + 5 = 3 \text{ bzw. } 2x = 5 \text{ bzw. } x^2 = 2$$

zu lösen, führt auf die Konstruktion der ganzen bzw. rationalen bzw. reellen Zahlen. Der Wunsch, quadratische Gleichungen

$$x^2 + ax + b = 0$$

einheitlich für alle reellen Werte von a, b – wie z.B. $a = 0, b = -1$ – lösen zu können, führt auf die Konstruktion der komplexen Zahlen. Weitere wichtige Anwendungen der komplexen Zahlen werden wir später kennen lernen.

Im Folgenden definieren wir die komplexen Zahlen und die entsprechenden Grundrechenarten geometrisch. Dabei sollen die gleichen Rechengesetze wie bei den reellen Zahlen gelten.

I.6.2. Definition. Wie in Abschnitt I.3.1 führen wir in der Ebene ein kartesisches Koordinatensystem ein. Jeder Punkt P der Ebene ist dann eindeutig charakterisiert durch seine Koordinaten x und y . Statt (x, y) schreiben wir $x + iy$ und nennen diesen Ausdruck eine KOMPLEXE ZAHL. Die Menge der komplexen Zahlen nennen wir \mathbb{C} :

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}.$$

Für eine komplexe Zahl $z = x + iy$ heißen x bzw. y der REALTEIL bzw. IMAGINÄRTEIL von z , kurz

$$\operatorname{Re}(x + iy) = x, \quad \operatorname{Im}(x + iy) = y.$$

Jedem Paar (x, y) reeller Zahlen entspricht also genau eine komplexe Zahl $z = x + iy$. Umgekehrt entspricht jeder komplexen Zahl z genau ein Paar $(\operatorname{Re} z, \operatorname{Im} z)$ reeller Zahlen. In diesem Sinne können wir die Mengen \mathbb{C} und \mathbb{R}^2 identifizieren.

I.6.3. Betrag und Konjugation. Nach Wahl eines kartesischen Koordinatensystems entspricht einer komplexen Zahl $z = x + iy$ der Punkt $P = (x, y)$ der Ebene. Dieser hat einen eindeutig definierten Abstand zum Ursprung des Koordinatensystems. Dies ist der Betrag des Ortsvektors \vec{OP} und gemäß Abschnitt I.4.9 gleich $\sqrt{x^2 + y^2}$. Wir nennen diese Größe den BETRAG der komplexen Zahl z und schreiben hierfür $|z|$:

$$|x + iy| = \sqrt{x^2 + y^2}$$

bzw.

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

Durch Spiegelung des Punktes $P = (x, y)$ an der x -Achse geht dieser in den Punkt $\overline{P} = (x, -y)$ über. Entsprechend geht die komplexe Zahl $z = x + iy$ über in die komplexe Zahl $x - iy$. Wir nennen diese die zu z KONJUGIERT KOMPLEXE ZAHL und schreiben dafür \bar{z} :

$$\overline{x + iy} = x - iy$$

bzw.

$$\bar{z} = (\operatorname{Re} z) - i(\operatorname{Im} z).$$

Da eine nochmalige Spiegelung den Punkt \overline{P} wieder in den Punkt P überführt, gilt die Regel

$$\overline{(\bar{z})} = z.$$

I.6.4. Addition und Subtraktion. Wir betrachten zwei komplexe Zahlen $z = x + iy$ und $w = u + iv$. Diesen entsprechen die Punkte $P = (x, y)$ und $Q = (u, v)$. Die zugehörigen Ortsvektoren $\overrightarrow{0P}$ und $\overrightarrow{0Q}$ können wir addieren und erhalten somit den Ortsvektor eines neuen Punktes R . Die zugehörige komplexe Zahl nennen wir die Summe von z und w , kurz $z + w$. Aus Abschnitt I.4.9 ergibt sich unmittelbar (vgl. Abbildung I.6.1):

$$(x + iy) + (u + iv) = (x + u) + i(y + v)$$

bzw.

$$z + w = (\operatorname{Re} z + \operatorname{Re} w) + i(\operatorname{Im} z + \operatorname{Im} w).$$

Analog lässt sich die Differenz $z - w$ der Zahlen z und w berechnen:

$$(x + iy) - (u + iv) = (x - u) + i(y - v)$$

bzw.

$$z - w = (\operatorname{Re} z - \operatorname{Re} w) + i(\operatorname{Im} z - \operatorname{Im} w).$$

Für die Konjugation gilt:

$$\overline{z \pm w} = \overline{z} \pm \overline{w}.$$

BEISPIEL I.6.1. Für $z = 2 + 3i$, $w = 1 + 2i$ erhalten wir $z + w = 3 + 5i$ und $z - w = 1 + i$.

I.6.5. Multiplikation und Division. Der Punkt $P = (x, y)$, der zu der komplexen Zahl $z = x + iy$ gehört, ist eindeutig bestimmt durch seinen Abstand r zum Ursprung und durch den Winkel φ zwischen der positiven x -Achse und dem Ortsvektor $\overrightarrow{0P}$, d.h. $z = r \cos \varphi + ir \sin \varphi = r(\cos \varphi + i \sin \varphi)$. Bezeichne mit s und ψ die entsprechenden Größen für die komplexe Zahl $w = u + iv$. Wir definieren das Produkt von z und w , kurz $z \cdot w$, nun durch folgende Vorschrift: Multipliziere die Abstände r und s und addiere die Winkel φ und ψ (vgl. Abbildung I.6.1).

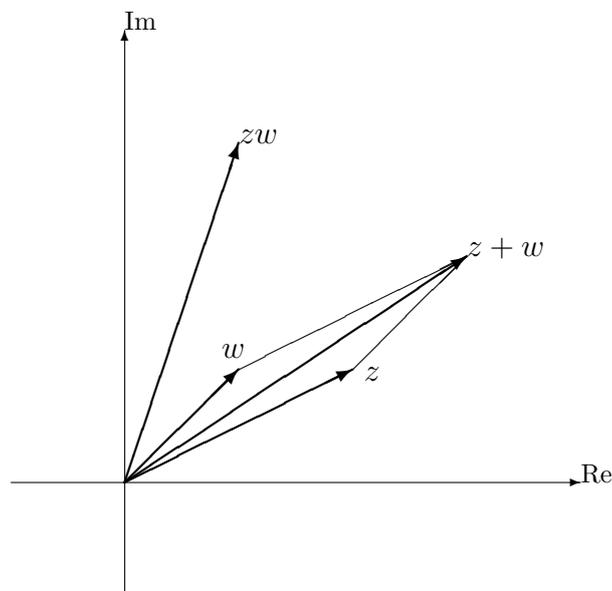


ABBILDUNG I.6.1. Addition und Multiplikation zweier komplexer Zahlen

Speziell erhalten wir für die komplexe Zahl i , d.h. $x = 0$, $y = 1$, den Wert $r = 1$ und $\varphi = \frac{\pi}{2}$. Gemäß unserer Regel hat dann $i^2 = i \cdot i$ den Abstand 1 und den Winkel π . Dies entspricht aber der komplexen Zahl -1 . Es gilt also die Regel

$$i^2 = -1.$$

Man kann zeigen, dass für die Addition und Multiplikation komplexer Zahlen die gleichen Rechenregeln gelten wie bei den reellen Zahlen. Daher ergibt sich folgende Regel für das Produkt zweier komplexer Zahlen:

$$(x + iy) \cdot (u + iv) = (xu - yv) + i(xv + yu)$$

bzw.

$$z \cdot w = ((\operatorname{Re} z)(\operatorname{Re} w) - (\operatorname{Im} z)(\operatorname{Im} w)) \\ + i((\operatorname{Re} z)(\operatorname{Im} w) + (\operatorname{Im} z)(\operatorname{Re} w)).$$

Der Quotient $\frac{z}{w}$, $w \neq 0$, ist die komplexe Zahl, deren Produkt mit w die Zahl z ergibt. Also muss sich dieser Quotient nach folgender Regel berechnen: Dividiere die Abstände r und s und subtrahiere die Winkel φ und ψ .

Andererseits entspricht der Zahl $\frac{\bar{w}}{|w|^2}$ der Abstand $\frac{1}{s}$ und der Winkel $2\pi - \psi$. Daher hat die Multiplikation von z mit $\frac{\bar{w}}{|w|^2}$ den gleichen Effekt wie die Division von z durch w . Daher erhalten wir die folgende Regel für den Quotienten zweier komplexer Zahlen

$$\frac{x + iy}{u + iv} = \frac{(x + iy) \cdot (u - iv)}{u^2 + v^2} \\ = \frac{xu + yv}{u^2 + v^2} + i \frac{-xv + yu}{u^2 + v^2}$$

bzw.

$$\frac{z}{w} = \frac{z \cdot \bar{w}}{|w|^2}.$$

Für die Konjugation und den Betrag ergibt sich

$$\overline{z \cdot w} = \bar{z} \cdot \bar{w} \\ \overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}. \\ |z \cdot w| = |z| |w| \\ \left|\frac{z}{w}\right| = \frac{|z|}{|w|}.$$

BEISPIEL I.6.2. Für $z = 2+3i$, $w = 1+2i$ erhalten wir $z \cdot w = -4+7i$ und $\frac{z}{w} = \frac{8}{5} - \frac{1}{5}i$.

BEISPIEL I.6.3 (ADDITIONSTHEOREME FÜR DEN SINUS UND COSINUS). Betrachte die komplexen Zahlen

$$\begin{aligned}z &= \cos \alpha + i \sin \alpha, \\w &= \cos \beta + i \sin \beta.\end{aligned}$$

Sie haben beide den Betrag 1 und die Winkel α bzw. β . Daher hat $z \cdot w$ den Betrag 1 und den Winkel $\alpha + \beta$. Damit folgt

$$\begin{aligned}\cos(\alpha + \beta) + i \sin(\alpha + \beta) &= z \cdot w \\&= (\cos \alpha + i \sin \alpha) \cdot (\cos \beta + i \sin \beta) \\&= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\&\quad + i(\cos \alpha \sin \beta + \sin \alpha \cos \beta).\end{aligned}$$

Vergleichen wir die Real- und Imaginärteile, erhalten wir die Additionstheoreme

$$\begin{aligned}\cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin(\alpha + \beta) &= \sin \alpha \cos \beta + \cos \alpha \sin \beta.\end{aligned}$$

Da $\frac{z}{w} = z \cdot \bar{w}$ (wegen $|w| = 1$) den Betrag 1 und den Winkel $\alpha - \beta$ hat, erhalten wir analog

$$\begin{aligned}\cos(\alpha - \beta) &= \cos \alpha \cos \beta + \sin \alpha \sin \beta \\ \sin(\alpha - \beta) &= \sin \alpha \cos \beta - \cos \alpha \sin \beta.\end{aligned}$$

I.6.6. Wurzeln komplexer Zahlen. Gegeben seien eine komplexe Zahl z und eine natürliche Zahl $n \geq 2$. Wir betrachten die Gleichung

$$w^n = \underbrace{w \cdot \dots \cdot w}_{n\text{-mal}} = z$$

mit einer unbekanntem komplexen Zahl w . Jede Lösung dieser Gleichung nennen wir eine n -TE WURZEL von z . Ist w eine n -te Wurzel von z schreiben wir auch $w = \sqrt[n]{z}$.

Sind r , s und φ , ψ die Abstände und Winkel zu z bzw. w , muss gemäß unserer Multiplikationsregel gelten: $s^n = r$ und $n \cdot \psi = \varphi$, d.h. $s = \sqrt[n]{r}$, $\psi = \frac{\varphi}{n}$. Betrachte nun eine natürliche Zahl $k \in \{1, \dots, n-1\}$ und die komplexe Zahl \tilde{w} mit Abstand $\tilde{s} = \sqrt[n]{r}$ und Winkel $\tilde{\psi} = \frac{\varphi}{n} + \frac{2\pi k}{n}$. Dann hat \tilde{w}^n den Abstand $\tilde{s}^n = r$ und den Winkel $n\tilde{\psi} = \varphi + 2\pi k$. Sie stellt daher auch die Zahl z dar. D.h., \tilde{w} ist auch eine n -te Wurzel von z . Damit ergibt sich:

Die komplexe Zahl z mit Abstand r und Winkel φ hat genau n verschiedene n -te Wurzeln. Diese haben alle den Abstand $\sqrt[n]{r}$ und die Winkel $\frac{\varphi}{n} + \frac{2\pi k}{n}$, $k = 0, 1, \dots, n - 1$.

Dies liefert folgende Regel:

$$\begin{aligned} z &= |z| \cos \varphi + i|z| \sin \varphi \\ \Rightarrow \sqrt[n]{z} &= \sqrt[n]{|z|} \cos\left(\frac{\varphi + 2\pi k}{n}\right) + i \sqrt[n]{|z|} \sin\left(\frac{\varphi + 2\pi k}{n}\right) \\ k &= 0, 1, \dots, n - 1. \end{aligned}$$

BEISPIEL I.6.4. Für $z = i$ und $n = 2$ erhalten wir

$$\begin{aligned} |z| &= 1 \\ \varphi &= \frac{\pi}{2} \end{aligned}$$

und damit die zwei Wurzeln

$$\begin{aligned} \sqrt{z} &= \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \\ &= \frac{\sqrt{2}}{2} + i \frac{\sqrt{2}}{2} \end{aligned}$$

und

$$\begin{aligned} \sqrt{z} &= \cos \frac{5\pi}{4} + i \sin \frac{5\pi}{4} \\ &= -\frac{\sqrt{2}}{2} - i \frac{\sqrt{2}}{2}. \end{aligned}$$

Für $z = -4\sqrt{2} + i4\sqrt{2}$ und $n = 3$ erhalten wir

$$\begin{aligned} |z| &= 8 \\ \varphi &= \frac{3\pi}{4} \end{aligned}$$

und damit die drei Wurzeln

$$\begin{aligned} \sqrt[3]{z} &= 2 \cos \frac{\pi}{4} + i2 \sin \frac{\pi}{4} \\ &= \sqrt{2} + i\sqrt{2} \\ \sqrt[3]{z} &= 2 \cos\left(\frac{\pi}{4} + \frac{2\pi}{3}\right) + i2 \sin\left(\frac{\pi}{4} + \frac{2\pi}{3}\right) \\ &= -\frac{\sqrt{2}(1 + \sqrt{3})}{2} + i \frac{\sqrt{2}(\sqrt{3} - 1)}{2} \end{aligned}$$

und

$$\begin{aligned}\sqrt[3]{z} &= 2 \cos\left(\frac{\pi}{4} + \frac{4\pi}{3}\right) + i2 \sin\left(\frac{\pi}{4} + \frac{4\pi}{3}\right) \\ &= \frac{\sqrt{2}(\sqrt{3}-1)}{2} - i \frac{\sqrt{2}(\sqrt{3}+1)}{2}.\end{aligned}$$

BEISPIEL I.6.5 (QUADRATISCHE GLEICHUNGEN). Betrachte die quadratische Gleichung

$$x^2 + 4x + 13 = 0.$$

Sie hat die Diskriminante

$$4^2 - 4 \cdot 13 = -36$$

und besitzt daher keine reellen Lösungen. Aus den Formeln von Vieta ergeben sich formal die Lösungen

$$x_{1,2} = -2 \pm \sqrt{4 - 13} = -2 \pm \sqrt{-9}.$$

Wegen $i^2 = -1$, d.h. $i = \sqrt{-1}$, erhalten wir die zwei komplexen Lösungen

$$x_{1,2} = -2 \pm 3i.$$

KAPITEL II

Lineare Algebra

II.1. Lineare Gleichungssysteme und Matrizen

II.1.1. Matrizen. Eine $m \times n$ MATRIX ist ein rechteckiges Zahlenschema der Form

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

Die reellen oder komplexen Zahlen a_{ij} heißen ELEMENTE der Matrix. Wir schreiben auch abkürzend $A = (a_{ij})_{m \times n}$ oder nur $A = (a_{ij})$, wenn die Zahlen m und n feststehen. Die Matrix heißt REELL, wenn *alle* Elemente reell sind, sonst KOMPLEX. Sofern nicht explizit anders erwähnt, betrachten wir im Folgenden stets reelle Matrizen.

Die $m \times 1$ -Matrizen bzw. $1 \times n$ -Matrizen heißen SPALTENVEKTOREN bzw. ZEILENVEKTOREN. Sie haben die Form

$$\mathbf{S} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad \text{bzw.} \quad \mathbf{Z} = (a_1, \dots, a_n).$$

Man beachte, dass man bei nur einer Zeile oder Spalte keinen zusätzlichen Spalten- bzw. Zeilenindex benötigt.

Die Matrix $A = (a_{ij})_{m \times n}$ besteht aus m Zeilenvektoren mit je n Elementen

$$\mathbf{Z}_i = (a_{i1}, a_{i2}, \dots, a_{in}) \quad i = 1, 2, \dots, m$$

bzw. aus n Spaltenvektoren mit je m Elementen

$$\mathbf{S}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad j = 1, 2, \dots, n.$$

Je nachdem, ob wir die Zeilen oder die Spalten von A hervorheben wollen, schreiben wir

$$A = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \\ \vdots \\ \mathbf{Z}_m \end{pmatrix} \quad \text{oder} \quad A = (\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_n).$$

Das (i, j) -te Element a_{ij} von A gehört zum i -ten Zeilenvektor \mathbf{Z}_i und zum j -ten Spaltenvektor \mathbf{S}_j .

Zwei Matrizen A und B sind genau dann GLEICH, wenn sie die gleiche Anzahl von Zeilen und Spalten haben und wenn $a_{ij} = b_{ij}$ gilt für alle i, j .

BEISPIEL II.1.1. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & -1 \end{pmatrix}$$

hat 2 Zeilen und 3 Spalten, ist also eine 2×3 -Matrix. Sie hat die Zeilenvektoren

$$\mathbf{Z}_1 = (1, 2, 3) \quad \text{und} \quad \mathbf{Z}_2 = (4, 0, -1)$$

und die Spaltenvektoren

$$\mathbf{S}_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad \mathbf{S}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \text{und} \quad \mathbf{S}_3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Die Menge aller reellen bzw. komplexen $m \times n$ Matrizen bezeichnen wir mit $\mathbb{R}^{m \times n}$ bzw. $\mathbb{C}^{m \times n}$. Insbesondere ist $\mathbb{R}^{1 \times n}$ die Menge aller reellen Zeilenvektoren mit n Elementen und $\mathbb{R}^{m \times 1}$ die Menge aller reellen Spaltenvektoren mit m Elementen. Zur Abkürzung schreiben wir auch \mathbb{R}^n statt $\mathbb{R}^{n \times 1}$ und fassen die Elemente von \mathbb{R}^n stets als Spaltenvektoren auf.

II.1.2. Rechenregeln. Matrizen mit gleicher Anzahl von Zeilen und Spalten können addiert und subtrahiert werden. Dies geschieht elementweise:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \pm \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} \pm b_{11} & \cdots & a_{1n} \pm b_{1n} \\ \vdots & & \vdots \\ a_{m1} \pm b_{m1} & \cdots & a_{mn} \pm b_{mn} \end{pmatrix}.$$

Matrizen können zudem mit einem Zahlenfaktor multipliziert werden. Auch dies geschieht elementweise:

$$\lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

BEISPIEL II.1.2. Es ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & -1 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 5 \\ 1 & 6 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 8 \\ 5 & 6 & 2 \end{pmatrix}$$

und

$$3 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 0 & -3 \end{pmatrix}.$$

Die Matrix

$$0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{m \times n},$$

deren sämtliche Elemente verschwinden, heißt NULLMATRIX. Die Nullmatrizen in $\mathbb{R}^{1 \times n}$ und $\mathbb{R}^{m \times 1}$ nennt man auch NULLVEKTOR.

Es gelten die zu den Zahlen analogen Rechenregeln:

$$\begin{aligned} A + B &= B + A \\ (A + B) + C &= A + (B + C) \\ A + 0 &= A \\ A - A &= 0 \\ (\lambda\mu)A &= \lambda(\mu A) \\ 1A &= A \\ (\lambda + \mu)A &= \lambda A + \mu A \\ \lambda(A + B) &= \lambda A + \lambda B. \end{aligned}$$

Aus den Rechenregeln folgt, dass sich jeder Spaltenvektor

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^{n \times 1}$$

eindeutig schreiben lässt als $\mathbf{a} = a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n$ mit den Vektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Das System dieser Vektoren heißt NATÜRLICHE BASIS des \mathbb{R}^n . Ganz analog heißt das System der Vektoren

$$\mathbf{e}'_1 = (1, 0, \dots, 0), \mathbf{e}'_2 = (0, 1, 0, \dots, 0), \dots, \mathbf{e}'_n = (0, \dots, 0, 1)$$

die NATÜRLICHE BASIS des $\mathbb{R}^{1 \times n}$.

II.1.3. Lineare Gleichungssysteme und Matrizen. Ein LINEARES GLEICHUNGSSYSTEM MIT m GLEICHUNGEN UND n UNBEKANNTEN x_1, \dots, x_n , kurz LGS, hat die Form

$$(II.1.1) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

mit den Koeffizienten a_{ij} und den Absolutgliedern b_i . (Kommt eine Unbekannte in einer Gleichung nicht vor, hat sie dort den Koeffizienten 0.)

Für ein solches LGS schreibt man

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

oder kurz

$$A\mathbf{x} = \mathbf{b}$$

mit der Koeffizientenmatrix $A = (a_{ij})_{m \times n}$, dem Spaltenvektor $\mathbf{x} \in \mathbb{R}^n$ mit den unbekanntenen Komponenten x_i und dem Spaltenvektor $\mathbf{b} \in \mathbb{R}^m$ der RECHTEN SEITE.

Ein LGS heißt HOMOGEN, wenn $\mathbf{b} = \mathbf{0}$ ist. Andernfalls heißt es INHOMOGEN.

Ein Spaltenvektor \mathbf{c} mit den Komponenten c_1, \dots, c_n heißt eine Lösung des LGS (II.1.1), wenn für $x_i = c_i$ ($i = 1, \dots, n$) alle m Gleichungen erfüllt sind, d.h. $A\mathbf{c} = \mathbf{b}$ gilt.

Ein homogenes LGS $A\mathbf{x} = \mathbf{0}$ besitzt stets mindestens eine Lösung, nämlich den Nullvektor. Diese Lösung nennt man auch die TRIVIALE LÖSUNG.

Nicht jedes LGS besitzt eine Lösung. Es tritt stets einer der folgenden drei Fälle auf:

- Das Gleichungssystem besitzt KEINE Lösung.
BEISPIEL: (zwei parallele Geraden)

$$\begin{aligned} 2x_1 + x_2 &= 1 \\ 2x_1 + x_2 &= 2. \end{aligned}$$

- Das Gleichungssystem besitzt GENAU EINE Lösung.
BEISPIEL: (zwei nicht parallele Geraden)

$$\begin{aligned} 2x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2. \end{aligned}$$

Hier ist $x_1 = -1$, $x_2 = 3$ die eindeutige Lösung.

- Das Gleichungssystem besitzt UNENDLICH VIELE Lösungen.
BEISPIEL: (eine Geraden)

$$2x_1 + x_2 = 1.$$

Hier ist $x_1 = a, x_2 = 1 - 2a$ für jedes $a \in \mathbb{R}$ eine Lösung.

II.1.4. Das Gaußsche Eliminationsverfahren. Das bekannteste und wichtigste Verfahren zur Lösung linearer Gleichungssysteme $A\mathbf{x} = \mathbf{b}$ ist das Gaußsche Eliminationsverfahren. Zu seiner Beschreibung ist es zweckmäßig, die Koeffizientenmatrix A und die rechte Seite \mathbf{b} zur erweiterten Koeffizientenmatrix (A, \mathbf{b}) zusammenzufassen, indem man \mathbf{b} als $(n + 1)$ -ten Spaltenvektor hinzufügt.

Das Gaußsche Eliminationsverfahren beruht auf folgenden Grundoperationen:

- Vertauschen zweier Zeilen von (A, \mathbf{b}) oder zweier Spalten von A . Wurden zwei Spalten vertauscht, müssen die Komponenten des Vektors \mathbf{x} der Unbekannten entsprechend umnummeriert werden.
- Multiplikation einer Zeile von (A, \mathbf{b}) mit einer von Null verschiedenen Zahl.
- Addition bzw. Subtraktion des α -fachen einer Zeile von (A, \mathbf{b}) von einer anderen Zeile.

Diese Operationen ändern die Lösungsmenge des LGS nicht, d.h. geht (B, \mathbf{c}) durch solche Operationen aus (A, \mathbf{b}) hervor, so besitzt das LGS $B\mathbf{x} = \mathbf{c}$ genau die gleichen Lösungen wie das LGS $A\mathbf{x} = \mathbf{b}$.

Das Gaußsche Eliminationsverfahren besteht aus drei Etappen:

- dem Eliminationsteil,
- dem Lösbarkeitstest,
- dem Rücklösungsteil.

Am Ende des Eliminationsteils hat die erweiterte Koeffizientenmatrix folgende Struktur

$$\begin{matrix} r \\ \\ m-r \end{matrix} \left\{ \begin{matrix} \left(\begin{array}{cccccccccc} \bullet & * & * & \dots & \dots & \dots & \dots & * & * \\ 0 & \bullet & * & \dots & \dots & \dots & \dots & * & * \\ 0 & 0 & \bullet & \dots & \dots & \dots & \dots & * & * \\ \vdots & \vdots & & \ddots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & & \bullet & * & \dots & * & * \\ 0 & & & & & 0 & \dots & 0 & * \\ \vdots & & & & & \vdots & \dots & \vdots & \vdots \\ 0 & & & & & 0 & \dots & 0 & * \end{array} \right) \\ \underbrace{\hspace{10em}}_{n+1} \end{matrix} \right.$$

Dabei steht \bullet für eine von Null verschiedene Zahl und $*$ für eine beliebige Zahl.

Die Zahl r heißt der RANG von A , kurz

$$r = \text{Rang } A.$$

Der Rang von A ist eindeutig bestimmt. Es gilt

$$\text{Rang } A \leq \min\{m, n\} \quad \text{für } A \in \mathbb{R}^{m \times n}.$$

Beim Lösungstest wird entschieden, ob das LGS keine, genau eine oder unendlich viele Lösungen besitzt. Falls es keine Lösung besitzt, wird das Verfahren beendet. Im Rücklösungsteil werden sämtliche Lösungen des LGS ermittelt.

ALGORITHMUS II.1.3. GAUSSSCHE ELIMINATIONSVERFAHREN zur Lösung des LGS $A\mathbf{x} = \mathbf{b}$.

(1) ELIMINATIONSTEIL:

(a) Setze $i = 1$.

(b) PIVOTSUCHE: Bestimme zwei Indizes i_0, j_0 mit $i \leq i_0 \leq m$ und $i \leq j_0 \leq n$ und $a_{i_0 j_0} \neq 0$. Falls keine derartigen Indizes existieren, setze $r = i - 1$ und gehe zum Lösbarkeitstest (2), andernfalls fahre mit Schritt (c) fort.

(c) ZEILEN- / SPALTENTAUSSCH: Falls $i_0 \neq i$ ist, vertausche die Zeilen i und i_0 von (A, \mathbf{b}) . Falls $j_0 \neq i$ ist, vertausche die Spalten i und j_0 von (A, \mathbf{b}) und merke die entsprechende Umnummerierung der Komponenten des Lösungsvektors. Falls $i = m$ ist, setze $r = m$ und gehe zum Lösbarkeitstest (2), andernfalls fahre mit Schritt (d) fort.

(d) ELIMINATION: Für $k = i + 1, \dots, m$ subtrahiere das $\frac{a_{ki}}{a_{ii}}$ -fache der i -ten Zeile von (A, \mathbf{b}) von der k -ten Zeile von (A, \mathbf{b}) .

(e) Falls $i \leq m - 1$ ist, erhöhe i um 1 und gehe zu Schritt (b) zurück.

(2) LÖSBARKEITSTEST: Falls $r < m$ und mindestens eine der Zahlen b_{r+1}, \dots, b_m von Null verschieden ist, besitzt das LGS keine Lösung. Beende in diesem Fall das Verfahren. Andernfalls gehe zum Rücklösungsteil.

(3) RÜCKLÖSUNGSTEIL:

(a) Setze

$$x_r = \frac{b_r}{a_{rr}}$$

und berechne für $i = r - 1, r - 2, \dots, 1$ sukzessive

$$x_i = \frac{1}{a_{ii}} \{ b_i - a_{ir}x_r - a_{ir-1}x_{r-1} - \dots - a_{ii+1}x_{i+1} \}.$$

(b) Falls $r = n$ ist, ist

$$\mathbf{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}$$

mit x_1, \dots, x_r aus (a) die eindeutige Lösung des LGS. Beende in diesem Fall das Verfahren. Andernfalls setze

$$\mathbf{u}_0 = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{n \times 1}$$

und fahre mit Schritt (c) fort.

(c) Führe für $j = r + 1, \dots, n$ folgende Schritte aus:

(i) Führe Schritt (a) durch mit dem Vektor

$$\begin{pmatrix} -a_{1j} \\ \vdots \\ -a_{rj} \end{pmatrix}$$

an Stelle des Vektors

$$\begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}.$$

(ii) Setze

$$\mathbf{u}_{j-r} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-te Komponente}$$

mit x_1, \dots, x_r aus Schritt (a).

(iii) Die Lösungsmenge des LGS ist

$$\{\mathbf{u}_0 + \alpha_1 \mathbf{u}_1 + \dots + \alpha_{n-r} \mathbf{u}_{n-r} : \alpha_1, \dots, \alpha_{n-r} \in \mathbb{R}\}.$$

Beende das Verfahren.

BEMERKUNG II.1.4. (1) Wenn man nur daran interessiert ist zu entscheiden, ob das LGS eine eindeutige Lösung besitzt und wie diese Lösung lautet, kann man sich bei der Pivotsuche in Schritt (1b) auf die i -te Spalte beschränken, d.h. man setzt $j_0 = i$. Dann entfällt die Notwendigkeit eines Spaltentausches mit entsprechender Umnummerierung der Unbekannten.

(2) Bei der Implementierung des Gaußschen Eliminationsverfahrens auf einem Computer sollte man bei der Pivotsuche die Indizes i_0, j_0 so bestimmen, dass

$$|a_{i_0 j_0}| = \max\{|a_{k\ell}| : i \leq k \leq m, j \leq \ell \leq n\}$$

ist. Die Abfrage „ $a_{i_0 j_0} = 0$?“ ist zu ersetzen durch „ $|a_{i_0 j_0}| \leq \varepsilon$ “, wobei ε eine kleine positive Zahl ist. Die Wahl von ε hängt von der Rechengenauigkeit ab. Typische Werte auf gebräuchlichen Computern sind $\varepsilon = 10^{-6}$ oder $\varepsilon = 10^{-8}$.

Das folgende Java-Programm realisiert Algorithmus II.1.3 für quadratische Matrizen, d.h. $n = m$:

```
// Gaussian elimination
public void gaussElimination() throws LinearAlgebraException {
    elimination();
    backSolve();
} // end of gaussElimination
// elimination part of Gaussian elimination
public void elimination() throws LinearAlgebraException {
    for ( int i = 0; i < dim-1; i++) {
        int jp = pivot(i); // find pivot, j is pivotelement
        if ( Math.abs( a[jp][i] ) < EPS ) // pivot = 0 ??
            throw new LinearAlgebraException(
                " WARNING: zero pivot in "+i+"-th elimination step");
        if ( jp > i ) { // swap equations i and j
            swap(b, i, jp);
            swap(a, i, jp);
        }
        for ( int k = i+1; k < dim; k++ ) // elimination
            substractEquations(k, i, a[k][i]/a[i][i]);
    }
    if ( Math.abs( a[dim-1][dim-1] ) < EPS )
        throw new LinearAlgebraException(
            " WARNING: zero pivot in "+(dim-1)+"-th elimination step");
} // end of elimination
// find pivot
public int pivot(int i) {
    int pivotElement = i;
    double pivotValue = Math.abs( a[i][i] );
    for ( int j = i+1; j < dim; j++ ) {
        if ( Math.abs( a[j][i] ) > pivotValue ) {
            pivotElement = j;
            pivotValue = Math.abs( a[j][i] );
        }
    }
    return pivotElement;
} // end of pivot
```

```

// subtract factor times equation i from equation k
public void subtractEquations(int k, int i, double factor) {
    b[k] -= b[i]*factor;
    for( int l = i+1; l < dim; l++ )
        a[k][l] -= a[i][l]*factor;
} // end of subtractEquations
// solution part of Gaussian elimination
public void backSolve() {
    x[dim-1] = b[dim-1]/a[dim-1][dim-1];
    for ( int i = dim-2; i >= 0; i-- )
        x[i] = (b[i] - innerProduct(a[i], x, i+1, dim))/a[i][i];
} // end of backSolve
// swap components i and j of vector u
public void swap(double[] u, int i, int j) {
    double v = u[i];
    u[i] = u[j];
    u[j] = v;
} // end of swap
// inner product of sections first, ..., last-1 of two vectors
public double innerProduct(double[] u, double[] v, int first,
    int last) {
    double prod = 0;
    for( int i = first; i < last; i++ )
        prod += u[i]*v[i];
    return prod;
} // end of inner product

```

BEISPIEL II.1.5. Betrachte das LGS

$$\begin{aligned}x_1 + 2x_2 + 3x_3 &= 4 \\4x_1 + 5x_2 + 6x_3 &= 0 \\7x_1 + 8x_2 + 10x_3 &= 4.\end{aligned}$$

Jede dieser Gleichungen beschreibt eine Ebene im Raum. Daher besteht die Lösungsmenge aus keinem Punkt, genau einem Punkt, einer Geraden oder aus einer Ebene. Es ist

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}.$$

Das Gaußsche Eliminationsverfahren liefert

$$\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 10 & 4 \end{array} \rightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -16 \\ 0 & -6 & -11 & -24 \end{array} \rightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -16 \\ 0 & 0 & 1 & 8 \end{array}.$$

Also ist $\text{Rang } A = 3$ und

$$\begin{aligned}x_3 &= 8 \\x_2 &= -\frac{1}{3}\{-16 + 6 \cdot 8\} \\ &= -\frac{32}{3}\end{aligned}$$

$$\begin{aligned} x_1 &= 4 - 2 \cdot \left(-\frac{32}{3}\right) - 3 \cdot 8 \\ &= \frac{4}{3}. \end{aligned}$$

Die eindeutige Lösung des LGS ist

$$\begin{pmatrix} \frac{4}{3} \\ -\frac{32}{3} \\ 8 \end{pmatrix}.$$

BEISPIEL II.1.6. Wir ersetzen in Beispiel II.1.5 den Koeffizienten 10 von x_3 in der dritten Gleichung durch 9:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}.$$

Das Gaußsche Eliminationsverfahren liefert jetzt

$$\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 4 \end{array} \rightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -16 \\ 0 & -6 & -12 & -24 \end{array} \rightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -16 \\ 0 & 0 & 0 & 8 \end{array}.$$

Also ist $\text{Rang } A = 2$. Wegen $b_3 = 8$ hat das LGS keine Lösung.

BEISPIEL II.1.7. Wir ersetzen in Beispiel II.1.6 die letzte Komponente von \mathbf{b} durch -4 :

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ -4 \end{pmatrix}.$$

Das Gaußsche Eliminationsverfahren liefert jetzt

$$\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & -4 \end{array} \rightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -16 \\ 0 & -6 & -12 & -32 \end{array} \rightarrow \begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -16 \\ 0 & 0 & 0 & 0 \end{array}.$$

Also ist $\text{Rang } A = 2$. Wegen $b_3 = 0$ besitzt das LGS unendlich viele Lösungen. Schritt (3a) liefert

$$\begin{aligned} x_2 &= \frac{16}{3} \\ x_1 &= 4 - 2 \cdot \frac{16}{3} \\ &= -\frac{20}{3} \end{aligned}$$

und somit

$$\mathbf{u}_0 = \begin{pmatrix} -\frac{20}{3} \\ \frac{16}{3} \\ 0 \end{pmatrix}.$$

Schritt (3c) liefert

$$\begin{aligned} x_2 &= \frac{6}{-3} \\ &= -2 \\ x_1 &= -3 - 2 \cdot (-2) \\ &= 1 \end{aligned}$$

und somit

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Daher ist die Lösungsmenge des LGS die Gerade

$$\left\{ \left(\begin{pmatrix} -\frac{20}{3} \\ \frac{16}{3} \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

II.2. Die Matrixmultiplikation

II.2.1. Das Matrixprodukt. Wir definieren das Produkt eines Zeilenvektors \mathbf{a} mit n Elementen und eines Spaltenvektors \mathbf{b} mit ebenfalls n Elementen durch

$$\mathbf{a}\mathbf{b} = (a_1, \dots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \dots + a_nb_n.$$

BEISPIEL II.2.1. Es ist

$$(2, 3, 0, 4) \begin{pmatrix} 1 \\ -1 \\ 5 \\ 2 \end{pmatrix} = 2 - 3 + 0 + 8 = 7.$$

Wir definieren das Produkt einer $k \times \ell$ -Matrix A mit den Zeilenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_k$ mit einer $\ell \times m$ -Matrix B mit den Spaltenvektoren $\mathbf{b}_1, \dots, \mathbf{b}_m$ durch

$$\begin{aligned}
 AB &= \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} (\mathbf{b}_1, \dots, \mathbf{b}_m) \\
 &= \begin{pmatrix} \mathbf{a}_1 \mathbf{b}_1 & \mathbf{a}_1 \mathbf{b}_2 & \dots & \mathbf{a}_1 \mathbf{b}_m \\ \mathbf{a}_2 \mathbf{b}_1 & \mathbf{a}_2 \mathbf{b}_2 & \dots & \mathbf{a}_2 \mathbf{b}_m \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_k \mathbf{b}_1 & \mathbf{a}_k \mathbf{b}_2 & \dots & \mathbf{a}_k \mathbf{b}_m \end{pmatrix} \\
 &= \left(\sum_{\mu=1}^{\ell} a_{i\mu} b_{\mu j} \right)_{\substack{1 \leq i \leq k \\ 1 \leq j \leq m}}.
 \end{aligned}$$

Beachte:

Das Produkt AB ist nur definiert, wenn die Spaltenzahl von A gleich der Zeilenzahl von B ist. Die Zeilen- bzw. Spaltenzahl des Produktes AB ist gleich der Zeilenzahl von A bzw. der Spaltenzahl von B .

Das Matrixprodukt ist im allgemeinen nicht kommutativ, d.h. es gibt Matrizen A und B mit

$$AB \neq BA.$$

BEISPIEL II.2.2. Für

$$\begin{aligned}
 A &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\
 B &= \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}
 \end{aligned}$$

erhalten wir

$$AB = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}$$

und

$$BA = \begin{pmatrix} 3 & 2 \\ 5 & 4 \end{pmatrix}.$$

BEISPIEL II.2.3. Es ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 & 0 & 4 & -5 \\ 2 & 1 & 3 & 4 \\ 0 & -2 & -6 & 2 \end{pmatrix} = \begin{pmatrix} 7 & -4 & -8 & 9 \\ 22 & -7 & -5 & 12 \end{pmatrix}.$$

Die Matrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

heißt $n \times n$ EINHEITSMATRIX. Wenn die Größe von n aus dem Zusammenhang klar ist, schreiben wir auch einfach I statt I_n .

Für beliebige Matrizen $A, A_1, A_2 \in \mathbb{R}^{k \times \ell}$, $B, B_1, B_2 \in \mathbb{R}^{\ell \times m}$, $C \in \mathbb{R}^{m \times n}$ und $\alpha \in \mathbb{R}$ gelten folgende Rechenregeln:

$$\begin{aligned} (A_1 + A_2)B &= (A_1B) + (A_2B), \\ A(B_1 + B_2) &= (AB_1) + (AB_2), \\ \alpha(AB) &= (\alpha A)B \\ &= A(\alpha B), \\ A(BC) &= (AB)C, \\ I_k A &= A I_\ell. \end{aligned}$$

II.2.2. Die transponierte Matrix. Jeder $m \times n$ -Matrix $A \in \mathbb{R}^{m \times n}$ ordnen wir ihre TRANSPONIERTER MATRIX $A^T \in \mathbb{R}^{n \times m}$ zu, deren i -te Zeile aus den Elementen der i -ten Spalte von A besteht.

BEISPIEL II.2.4. Es ist

$$\begin{aligned} (1 \ 2 \ 3)^T &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \\ \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}^T &= (4, 5, 6), \\ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T &= \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \\ \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^T &= \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}. \end{aligned}$$

Es gelten folgende Rechenregeln:

$$\begin{aligned} (A + B)^T &= A^T + B^T, \\ (\alpha A)^T &= \alpha(A^T), \\ (A^T)^T &= A, \end{aligned}$$

$$(AB)^T = B^T A^T.$$

Eine $n \times n$ Matrix A heißt SYMMETRISCH, wenn $A^T = A$ ist. Sie heißt SCHIEFSYMMETRISCH, wenn $A^T = -A$ ist.

BEISPIEL II.2.5. Die Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

ist symmetrisch; die Matrix

$$\begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{pmatrix}$$

ist schiefsymmetrisch.

II.2.3. Invertierbare Matrizen. Eine $n \times n$ Matrix A heißt INVERTIERBAR oder REGULÄR, wenn es eine $n \times n$ Matrix B gibt mit $AB = BA = I_n$. Falls A invertierbar ist, ist diese Matrix B eindeutig bestimmt. Sie heißt INVERSE MATRIX oder INVERSE von A und wird mit A^{-1} bezeichnet. Ist A nicht invertierbar, nennen wir A auch SINGULÄR.

Ist A invertierbar und $\mathbf{b} \in \mathbb{R}^{n \times 1}$, so ist das LGS $A\mathbf{x} = \mathbf{b}$ eindeutig lösbar. Die Lösung ist $\mathbf{x} = A^{-1}\mathbf{b}$.

Es gelten folgende Rechenregeln:

$$\begin{aligned} A \text{ invertierbar und } \alpha \neq 0 &\implies \alpha A \text{ invertierbar und } (\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}. \\ A \text{ und } B \text{ invertierbar} &\implies AB \text{ invertierbar und } (AB)^{-1} = B^{-1} A^{-1}. \\ A \text{ invertierbar} &\implies A^T \text{ invertierbar und } (A^T)^{-1} = (A^{-1})^T. \end{aligned}$$

II.2.4. Das Gaußsche Eliminationsverfahren zur Berechnung der inversen Matrix. Mit einer Variante des Gaußschen Eliminationsverfahrens aus Abschnitt II.1.4 kann entschieden werden, ob eine $n \times n$ -Matrix A invertierbar ist, und im positiven Fall A^{-1} berechnet werden. Die Idee ist, simultan die LGS $A\mathbf{x} = \mathbf{e}_1, \dots, A\mathbf{x} = \mathbf{e}_n$ zu lösen. Dazu wird die Matrix A durch die $n \times n$ Einheitsmatrix I_n zu (A, I_n) ergänzt. Auf die erweiterte Matrix werden die Grundoperationen aus Abschnitt II.1.4 so lange angewandt, bis entweder entschieden ist, dass A nicht invertierbar ist, oder bis die erweiterte Matrix die Form (I_n, B) hat. Dann ist $B = A^{-1}$.

ALGORITHMUS II.2.6. GAUSSSCHES ELIMINATIONSVERFAHREN zur Berechnung der inversen Matrix.

- (1) ELIMINATIONSTEIL:
- Setze $i = 1$.
 - PIVOTSUCHE: Bestimme einen Index i_0 mit $i \leq i_0 \leq n$ und $a_{i_0 i} \neq 0$. Falls kein derartiger Index existiert, ist die Matrix nicht invertierbar; beende das Verfahren. Andernfalls fahre mit Schritt (c) fort.
 - ZEILENTAUSCH: Vertausche die Zeilen i und i_0 der erweiterten Matrix.
 - ELIMINATION: Für $k = i + 1, \dots, n$ subtrahiere das $\frac{a_{ki}}{a_{ii}}$ -fache der i -ten Zeile der erweiterten Matrix von deren k -ter Zeile.
 - Falls $i < n - 1$ ist, erhöhe i um 1 und gehe zu Schritt (b) zurück. Falls $i = n - 1$ und $a_{nn} = 0$ ist, ist die Matrix nicht invertierbar; beende das Verfahren. Andernfalls gehe zum Rücklösungsteil über.
- (2) RÜCKLÖSUNGSTEIL:
- Für $i = 1, \dots, n$ dividiere die i -te Zeile der erweiterten Matrix durch a_{ii} .
 - Für $i = n, n-1, \dots, 1$ und $k = i-1, i-2, \dots, 1$ subtrahiere das a_{ki} -fache der i -ten Zeile der erweiterten Matrix von deren k -ter Zeile.
 - Die letzten n Spalten der erweiterten Matrix sind die Inverse von A . Beende das Verfahren.

BEMERKUNG II.2.7. Wie in Abschnitt II.1.4 sollte man bei der Implementierung des Verfahrens auf einem Computer den Pivotindex i_0 so bestimmen, dass

$$|a_{i_0 i}| = \max\{|a_{ki}| : i \leq k \leq n\}$$

ist, und die Abfrage „ $|a_{ii}| = 0$?“ durch „ $|a_{ii}| \leq \varepsilon$?“ ersetzen.

BEISPIEL II.2.8. Für die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$$

aus Beispiel II.1.5 (S. 57) erhalten wir folgendes Schema:

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 & 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 & \rightarrow & 0 & -3 & -6 & -4 & 1 & 0 \\ 7 & 8 & 10 & 0 & 0 & 1 & & 0 & -6 & -11 & -7 & 0 & 1 \end{array}$$

$$\rightarrow \begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 & 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 & \rightarrow & 0 & 1 & 2 & \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 & & 0 & 0 & 1 & 1 & -2 & 1 \end{array}$$

$$\rightarrow \begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 6 & -3 \\ 0 & 1 & 0 & -\frac{2}{3} & \frac{11}{3} & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array} \quad \rightarrow \begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{3} & -\frac{4}{3} & 1 \\ 0 & 1 & 0 & -\frac{2}{3} & \frac{11}{3} & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}$$

Es ist

$$A^{-1} = \begin{pmatrix} -\frac{2}{3} & -\frac{4}{3} & 1 \\ -\frac{2}{3} & \frac{11}{3} & -2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Aus der Definition der Invertierbarkeit, Algorithmus II.2.6 und Abschnitt II.1.4 ergibt sich:

Für eine $n \times n$ Matrix sind folgende Aussagen äquivalent:

- A ist invertierbar.
- Es gibt eine $n \times n$ Matrix B mit $BA = I$.
- Es gibt eine $n \times n$ Matrix C mit $AC = I$.
- Das homogene LGS $A\mathbf{x} = \mathbf{0}$ besitzt nur die triviale Lösung $\mathbf{x} = \mathbf{0}$.
- Jedes LGS $A\mathbf{x} = \mathbf{b}$ besitzt eine eindeutige Lösung.
- $\text{Rang } A = n$.

II.2.5. Die LR-Zerlegung. In der Praxis tritt häufig das Problem auf, mehrere LGS $A\mathbf{x} = \mathbf{b}$ mit gleicher Matrix A und verschiedenen rechten Seiten \mathbf{b} lösen zu müssen. Da der Eliminationsteil des Gaußschen Verfahrens zur Lösung linearer Gleichungssysteme, Algorithmus II.1.3 (S. 54), $O(n^3)$ -Operationen, der Rücklösungsteil aber nur $O(n^2)$ -Operationen erfordert, ist es nicht sinnvoll, diesen Algorithmus auf jedes LGS separat anzuwenden. Da andererseits der Rücklösungsteil des Gaußschen Verfahrens zur Bestimmung von A^{-1} , Algorithmus II.2.6, genauso aufwendig ist wie der Eliminationsteil und da die Multiplikation $A^{-1}\mathbf{b}$ für jedes \mathbf{b} $O(n^2)$ Operationen erfordert, lohnt sich die Berechnung von A^{-1} nur dann, wenn die Zahl der rechten Seiten etwa n ist.

Die LR-Zerlegung ist auf diese Problematik zugeschnitten. Die Idee besteht darin, im Algorithmus II.1.3 (S. 54) die Behandlung der Matrix A und der rechten Seite \mathbf{b} zu entkoppeln. Hierzu merkt man sich die Elementaroperationen, die im Eliminationsteil ausgeführt werden, um sie später auf die rechte Seite anzuwenden.

Der Algorithmus zerfällt in den Zerlegungs- und den Lösungsteil. Im Zerlegungsteil werden drei Matrizen L , R und P mit folgenden Eigenschaften berechnet:

- P ist eine Permutationsmatrix, d.h. P entsteht durch Zeilenvertauschungen aus I_n .

- R ist eine OBERE DREIECKSMATRIX mit von Null verschiedenen Diagonalelementen, d.h. R hat die Form

$$\begin{pmatrix} \bullet & * & * & \dots & * \\ 0 & \bullet & * & \dots & * \\ 0 & 0 & \bullet & \dots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \bullet \end{pmatrix}.$$

- L ist eine UNTERE DREIECKSMATRIX mit Diagonalelementen 1, d.h. L hat die Form

$$\begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ * & 1 & 0 & \dots & \dots & 0 \\ * & * & 1 & \dots & \dots & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & 0 \\ * & * & \dots & \dots & * & 1 \end{pmatrix}.$$

- Es ist

$$LR = PA.$$

Im Lösungsteil wird bei gegebener rechter Seite \mathbf{b} das LGS $A\mathbf{x} = \mathbf{b}$ in folgenden Schritten gelöst:

- Berechne $\mathbf{z} = P\mathbf{b}$.
- Löse das LGS $L\mathbf{y} = \mathbf{z}$.
- Löse das LGS $R\mathbf{x} = \mathbf{y}$.

Dabei können der zweite und dritte Schritt wegen der Dreiecksstruktur von L und R analog zu Schritt (3a) von Algorithmus II.1.3 (S. 54) ausgeführt werden.

ALGORITHMUS II.2.9. *Bestimmung der LR-ZERLEGUNG.*

- (0) Setze $\mathbf{p} = (1, 2, \dots, n)$ und $i = 1$.
- (1) Bestimme einen Index i_0 mit $i \leq i_0 \leq n$ und $a_{i_0 i} \neq 0$. Falls kein solcher Index existiert, ist die Matrix A nicht zerlegbar und damit auch nicht invertierbar; beende das Verfahren mit einer entsprechenden Fehlermeldung.
- (2) Vertausche die i -te und die i_0 -te Zeile von A und die i -te und die i_0 -te Komponente von \mathbf{p} .
- (3) Für $k = i + 1, i + 2, \dots, n$ führe folgende Schritte aus:
 - (a) Ersetze a_{ki} durch $\frac{a_{ki}}{a_{ii}}$.
 - (b) Für $j = i + 1, i + 2, \dots, n$ ersetze a_{kj} durch $a_{kj} - a_{ij}a_{ki}$.
- (4) Falls $i < n - 1$ ist, erhöhe i um 1 und gehe zu Schritt 1 zurück. Falls $i = n - 1$ und $a_{nn} = 0$ ist, ist die Matrix A nicht zerlegbar und damit auch nicht invertierbar; beende das Verfahren mit einer entsprechenden Fehlermeldung. Ansonsten ist das Verfahren erfolgreich abgeschlossen.

Nach erfolgreicher Durchführung von Algorithmus II.2.9 enthält die obere Hälfte der Matrix A einschließlich ihrer Diagonalen die obere Hälfte der Matrix R . Die untere Hälfte von A ohne Diagonale enthält die untere Hälfte von L ohne die Diagonale, die durch 1 ergänzt werden muss. Die Matrix P erhält man aus dem Vektor \mathbf{p} wie folgt: In der i -ten Zeile ($1 \leq i \leq n$) von P steht in der p_i -ten Spalte eine 1, alle anderen Elemente der Zeile sind 0.

ALGORITHMUS II.2.10. *Lösen des LGS $A\mathbf{x} = \mathbf{b}$ bei bekannter LR-Zerlegung.*

(1) Für $i = 1, \dots, n$ setze

$$z_i = b_{p_i}.$$

(2) Setze

$$y_1 = z_1$$

und berechne für $i = 2, \dots, n$ sukzessive

$$y_i = z_i - a_{i1}y_1 - \dots - a_{i-1}y_{i-1}.$$

(3) Setze

$$x_n = \frac{y_n}{a_{nn}}$$

und berechne für $i = n-1, n-2, \dots, 1$ sukzessive

$$x_i = \frac{1}{a_{ii}} \{y_i - a_{ii+1}x_{i+1} - \dots - a_{in}x_n\}.$$

Das folgende Java-Programm realisiert die Algorithmen II.2.9 und II.2.10. Die Methoden `pivot`, `swap`, `innerProduct` und `backsolve` sind identisch mit den in Abschnitt II.1.4 angegebenen Methoden gleichen Namens.

```
// LR decomposition
public void lrDecomposition() throws LinearAlgebraException {
    lrElimination();
    permutation();
    forSolve();
    backSolve();
} // end of lrDecomposition
// elimination part of LR-decomposition
public void lrElimination() throws LinearAlgebraException {
    perm = new int[dim];
    for( int i = 0; i < dim; i++ )
        perm[i] = i;
    for ( int i = 0; i < dim-1; i++ ) {
        int jp = pivot(i); // find pivot, j is pivotelement
        if ( Math.abs( a[jp][i] ) < EPS ) // pivot = 0 ??
            throw new LinearAlgebraException(
                " WARNING: zero pivot in "+i+"-th elimination step");
        if ( jp > i ) { // swap equations i and j
            swap(perm, i, jp);
            swap(a, i, jp);
        }
        for ( int k = i+1; k < dim; k++ ) { // elimination
```

```

        a[k][i] /= a[i][i];
        subtractRows(k, i, a[k][i]);
    }
}
if ( Math.abs( a[dim-1][dim-1] ) < EPS )
    throw new LinearAlgebraException(
        " WARNING: zero pivot in "+(dim-1)+"-th elimination step");
} // end of lrElimination
// subtract factor times row i from row k of a
public void subtractRows(int k, int i, double factor) {
    for( int l = i+1; l < dim; l++ )
        a[k][l] -= a[i][l]*factor;
} // end of subtractRows
// forward solution part of LR-algorithm (rhs is x, result is b)
public void forSolve() {
    b[0] = x[0];
    for( int i = 1; i < dim; i++ )
        b[i] = x[i] - innerProduct(a[i], b, 0, i);
} // end of forSolve
// permute right-hand side according to vector perm and store on x
public void permutation() {
    for( int i = 0; i < dim; i++ )
        x[i] = b[ perm[i] ];
} // end of permutation

```

Für den Aufwand der Algorithmen II.2.9 und II.2.10 gilt:

Algorithmus II.2.9 benötigt $O(n^3)$ arithmetische Operationen. Algorithmus II.2.10 benötigt pro rechter Seite $O(n^2)$ Operationen.

Das folgende Beispiel verdeutlicht die Vorgehensweise der Algorithmen II.2.9 und II.2.10.

BEISPIEL II.2.11. Für die Matrix aus Beispiel II.1.5 (S. 57) liefert Algorithmus II.2.9 die LR -Zerlegung:

$$\mathbf{p} = (1, 2, 3)$$

und

$$\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{array} \rightarrow \begin{array}{ccc} 1 & 2 & 3 \\ 4 & -3 & -6 \\ 7 & -6 & -11 \end{array} \rightarrow \begin{array}{ccc} 1 & 2 & 3 \\ 4 & -3 & -6 \\ 7 & 2 & 1 \end{array}$$

Wir erhalten $A = LR$ mit

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}.$$

Für die rechte Seite

$$\mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}$$

ergibt sich $\mathbf{z} = \mathbf{b}$ und

$$\mathbf{y} = \begin{pmatrix} 4 \\ -16 \\ 8 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \frac{4}{3} \\ -\frac{32}{3} \\ 8 \end{pmatrix}.$$

II.3. Determinanten

II.3.1. Determinanten von 2×2 und 3×3 Matrizen. Das von den Vektoren

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{und} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

in der Ebene erzeugte Parallelogramm hat gemäß Abschnitt I.4.7 und I.4.9 die Fläche

$$F = \left| \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right| = |a_1 b_2 - a_2 b_1|.$$

Wir nennen die Zahl

$$\det A = a_1 b_2 - a_2 b_1$$

die DETERMINANTE der 2×2 Matrix

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

Wegen Abschnitt I.4.7 und I.4.9 ist $\det A = 0$ genau dann, wenn die Vektoren \mathbf{a} und \mathbf{b} parallel sind, d.h. $\mathbf{a} = \lambda \mathbf{b}$ für ein $\lambda \in \mathbb{R}$. Wegen Abschnitt II.1.4 ist letzteres genau dann der Fall, wenn $\text{Rang } A < 2$ ist. Also gilt:

$$\det A \neq 0 \iff A \text{ ist invertierbar.}$$

BEISPIEL II.3.1. Es ist

$$\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2.$$

Betrachte nun drei Vektoren

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}.$$

Das von ihnen in einem kartesischen Koordinatensystem erzeugte Spat hat gemäß Abschnitt I.4.8 und I.4.9 das Volumen

$$\begin{aligned}
 V &= |[\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]| \\
 &= |a_{11}(a_{22}a_{33} - a_{32}a_{23}) \\
 &\quad - a_{21}(a_{12}a_{33} - a_{32}a_{13}) \\
 &\quad + a_{31}(a_{12}a_{23} - a_{22}a_{13})|.
 \end{aligned}
 \tag{II.3.1}$$

Wir nennen die Zahl

$$\det A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$$

die DETERMINANTE der 3×3 Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

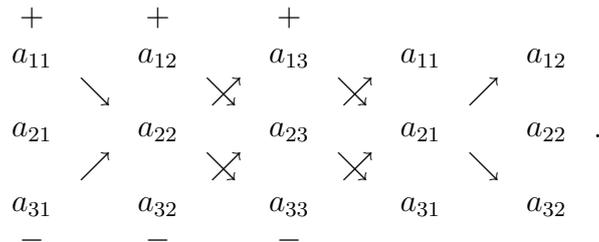
Wie für 2×2 Matrizen gilt:

$$\det A \neq 0 \iff A \text{ ist invertierbar.}$$

Ein Vergleich der Formel (II.3.1) mit der Definition der Determinante einer 2×2 Matrix liefert die Rekursionsformel

$$\begin{aligned}
 \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \\
 &\quad - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \\
 &\quad + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}.
 \end{aligned}$$

Die Determinante einer 3×3 Matrix kann nach der FORMEL VON SARRUS berechnet werden:



Dabei müssen die drei Zahlen auf jedem Pfeil miteinander multipliziert, die Ergebnisse der \searrow -Pfeile addiert und die Ergebnisse der \nearrow -Pfeile subtrahiert werden.

BEISPIEL II.3.2. Es ist

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} &= 1 \cdot 5 \cdot 10 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 \\ &\quad - 7 \cdot 5 \cdot 3 - 8 \cdot 6 \cdot 1 - 10 \cdot 4 \cdot 2 \\ &= 50 + 84 + 96 - 105 - 48 - 80 \\ &= -3. \end{aligned}$$

II.3.2. Die Determinante einer $n \times n$ Matrix. In Verallgemeinerung der Fälle $n = 2$ und $n = 3$ wird die DETERMINANTE einer $n \times n$ Matrix wie folgt rekursiv definiert:

Für $n = 1$, d.h. $A = (a_{11})$, ist $\det A = a_{11}$.

Für $n \geq 2$ ist

$$\begin{aligned} \det A &= a_{11} \det(A_{11}) - a_{21} \det A_{21} \\ &\quad + \dots + (-1)^{n+1} a_{n1} \det A_{n1}. \end{aligned}$$

Dabei ist A_{ij} die $(n-1) \times (n-1)$ Matrix, die durch Streichen der i -ten Zeile und j -ten Spalte aus A entsteht.

BEISPIEL II.3.3. Für

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & -2 & 1 & 0 \\ 0 & 6 & 3 & -2 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

erhalten wir

$$A_{11} = \begin{pmatrix} -2 & 1 & 0 \\ 6 & 3 & -2 \\ 4 & 3 & 1 \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} 2 & 0 & 1 \\ 6 & 3 & -2 \\ 4 & 3 & 1 \end{pmatrix},$$

$$A_{31} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix},$$

$$A_{41} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 6 & 3 & -2 \end{pmatrix}$$

und

$$\det A = 1 \cdot \det A_{11} - 3 \cdot \det A_{21} + 0 \cdot \det A_{31} - 2 \cdot \det A_{41}$$

$$\begin{aligned}
&= -32 - 3 \cdot 24 - 2 \cdot (-16) \\
&= -72.
\end{aligned}$$

II.3.3. Rechenregeln für Determinanten. Es gelten folgende Rechenregeln für Determinanten:

- (1) Eine Determinante kann durch Entwicklung nach einer beliebigen Zeile oder Spalte berechnet werden:

$$\begin{aligned}
\det A &= (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+2} a_{i2} \det A_{i2} \\
&\quad + \dots + (-1)^{i+n} a_{in} \det A_{in} \\
&\quad \text{(Entwicklung nach der } i\text{-ten Zeile)} \\
&= (-1)^{j+1} a_{1j} \det A_{1j} + (-1)^{j+2} a_{2j} \det A_{2j} \\
&\quad + \dots + (-1)^{j+n} a_{jn} \det A_{jn} \\
&\quad \text{(Entwicklung nach der } j\text{-ten Spalte)}.
\end{aligned}$$

- (2) Hat A obere bzw. untere Dreiecksgestalt, d.h. $a_{ij} = 0$ für $i > j$ bzw. $a_{ij} = 0$ für $i < j$, so ist $\det A$ das Produkt der Diagonalelemente

$$\det A = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}.$$

- (3) Die Determinante ist LINEAR IN JEDER ZEILE bzw. SPALTE, d.h.

$$\begin{aligned}
\det \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \alpha \mathbf{z}_i \\ \vdots \\ \mathbf{z}_n \end{pmatrix} &= \alpha \det \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_i \\ \vdots \\ \mathbf{z}_n \end{pmatrix} \\
\det \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{a} + \mathbf{b} \\ \vdots \\ \mathbf{z}_n \end{pmatrix} &= \det \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{a} \\ \vdots \\ \mathbf{z}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{z}_n \end{pmatrix}
\end{aligned}$$

bzw.

$$\begin{aligned}
\det(\mathbf{S}_1, \dots, \alpha \mathbf{S}_i, \dots, \mathbf{S}_n) &= \alpha \det(\mathbf{S}_1, \dots, \mathbf{S}_i, \dots, \mathbf{S}_n) \\
\det(\mathbf{S}_1, \dots, \mathbf{a} + \mathbf{b}, \dots, \mathbf{S}_n) &= \det(\mathbf{S}_1, \dots, \mathbf{a}, \dots, \mathbf{S}_n) \\
&\quad + \det(\mathbf{S}_1, \dots, \mathbf{b}, \dots, \mathbf{S}_n).
\end{aligned}$$

- (4) Die Determinante ist ALTERNIEREND, d.h. geht \tilde{A} aus A durch Vertauschen zweier Zeilen oder Spalten hervor, dann gilt, $\det \tilde{A} = -\det A$.

Insbesondere ist $\det A = 0$, falls zwei Zeilen oder Spalten von A gleich sind.

- (5) Die Determinante ändert sich nicht, wenn das Vielfache einer Zeile (bzw. Spalte) zu einer anderen Zeile (bzw. Spalte) hinzuaddiert wird.
- (6) $\det(A^T) = \det A$.
- (7) $\det(AB) = (\det A)(\det B)$.
- (8) A ist invertierbar genau dann, wenn $\det A \neq 0$ ist. In diesem Fall ist $\det(A^{-1}) = (\det A)^{-1}$.
- (9) Ist B eine $k \times k$ Matrix, C eine $k \times \ell$ Matrix, D eine $\ell \times k$ Matrix und E eine $\ell \times \ell$ Matrix, so gilt

$$\det \begin{pmatrix} B & C \\ 0 & E \end{pmatrix} = \det B \cdot \det E,$$

$$\det \begin{pmatrix} B & 0 \\ D & E \end{pmatrix} = \det B \cdot \det E.$$

BEMERKUNG II.3.4. Für $n \geq 4$ empfiehlt es sich für die Berechnung von $\det A$ häufig, A mit dem Eliminationsteil des Gaußschen Algorithmus II.1.3 (S. 54) zunächst auf obere Dreiecksgestalt zu transformieren. Dies ändert die Determinante nicht.

BEISPIEL II.3.5. Es ist

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 3 & 4 \end{pmatrix} = 0 \quad \text{da } \mathbf{z}_3 = \mathbf{z}_1 + \mathbf{z}_2$$

und

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 1 & 2 \\ 0 & 1 & 3 & -1 \\ -1 & 2 & 0 & 1 \end{pmatrix} &= \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 4 & 1 & 2 \\ 0 & 1 & 3 & -1 \\ 0 & 4 & 3 & 5 \end{pmatrix} \\ &= \det \begin{pmatrix} 4 & 1 & 2 \\ 1 & 3 & -1 \\ 4 & 3 & 5 \end{pmatrix} \\ &= \det \begin{pmatrix} 0 & -11 & 6 \\ 1 & 3 & -1 \\ 0 & -9 & 9 \end{pmatrix} \\ &= -\det \begin{pmatrix} -11 & 6 \\ -9 & 9 \end{pmatrix} \\ &= 45. \end{aligned}$$

II.3.4. Die Cramersche Regel. Die Matrix $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ sei invertierbar. Dann besitzt das LGS $A\mathbf{x} = \mathbf{b}$ eine eindeutige Lösung. Ersetzen wir die i -te Spalte \mathbf{a}_i von A durch $\mathbf{b} = \mathbf{a}_1x_1 + \dots + \mathbf{a}_nx_n$, so erhalten wir mit den Rechenregeln des vorigen Abschnitts

$$\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)$$

$$\begin{aligned}
 &= \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) \\
 &= x_1 \underbrace{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_1, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}_{=0} + \dots \\
 &\quad + x_i \underbrace{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_i, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}_{=\det A} + \dots \\
 &\quad + x_n \underbrace{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_n, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}_{=0} \\
 &= x_i \det A.
 \end{aligned}$$

Dies beweist die

CRAMERSCHE REGEL: Die Lösung des LGS $A\mathbf{x} = \mathbf{b}$ ist bei invertierbarer Matrix A gegeben durch

$$x_i = \frac{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}{\det A} \quad i = 1, \dots, n.$$

BEISPIEL II.3.6. Die Cramersche Regel ist für die praktische Rechnung bei großen Werten von n , d.h. $n \succeq 4$, UNGEEIGNET. Denn ihr Rechenaufwand ist $n!$ im Vergleich zu n^3 beim Gaußschen Eliminationsverfahren. So würde ein Gigaflop Rechner mit der Cramerschen Regel zur Lösung eines LGS mit 20 Gleichungen und Unbekannten mehr als 77 Jahre benötigen; bei dem Gaußschen Eliminationsverfahren ist die Rechenzeit kleiner als $8\mu\text{sec}$.

II.3.5. Kegelschnitte. Ein Kegelschnitt in der (x, y) -Ebene ist gegeben durch eine Gleichung der Form

$$a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 = 0,$$

in der nicht alle Koeffizienten a_i von Null verschieden sind. Da ein Koeffizient zu 1 normiert werden kann, ist im allgemeinen ein Kegelschnitt durch die Angabe von 5 Punkten $A_i = (x_i, y_i)$, $1 \leq i \leq 5$, auf der Kurve bestimmt. Diese Punkte A_i und der „allgemeine“ Punkt $x = (x, y)$ liefern ein homogenes Gleichungssystem

$$\begin{aligned}
 a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 &= 0 \\
 a_1x_1^2 + a_2y_1^2 + a_3x_1y_1 + a_4x_1 + a_5y_1 + a_6 &= 0 \\
 &\vdots \quad \vdots \\
 a_1x_5^2 + a_2y_5^2 + a_3x_5y_5 + a_4x_5 + a_5y_5 + a_6 &= 0
 \end{aligned}$$

für die Koeffizienten a_1, \dots, a_6 . Da dieses eine nicht triviale Lösung besitzt, muss die Determinante der Koeffizientenmatrix verschwinden. Das ergibt die sog. 5-PUNKTE-GLEICHUNG FÜR DEN ALLGEMEINEN

KEGELSCHNITT:

$$\det \begin{pmatrix} x^2 & y^2 & xy & x & y & 1 \\ x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_5^2 & y_5^2 & x_5 y_5 & x_5 & y_5 & 1 \end{pmatrix} = 0.$$

Spezielle Kegelschnitte sind durch weniger als 5 Punkte festgelegt. So hat z.B. ein Kreis die Gleichung $a(x^2 + y^2) + bx + cy + d = 0$ und ist durch die Angabe von drei Punkten festgelegt. Wie oben ergibt sich die Gleichung

$$\det \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{pmatrix} = 0$$

für den Kreis durch die Punkte (x_i, y_i) , $1 \leq i \leq 3$.

BEISPIEL II.3.7. Der Kreis durch die Punkte $(0, 0)$, $(1, 3)$ und $(2, -1)$ hat die Gleichung

$$\begin{aligned} 0 &= \det \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ 0 & 0 & 0 & 1 \\ 10 & 1 & 3 & 1 \\ 5 & 2 & -1 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} x^2 + y^2 & x & y \\ 10 & 1 & 3 \\ 5 & 2 & -1 \end{pmatrix} \\ &= -7(x^2 + y^2) + 25x + 15y \end{aligned}$$

also $7(x^2 + y^2) - 25x - 15y = 0$.

II.4. Eigenwerte und Eigenvektoren

II.4.1. Definition. Eine Zahl $\lambda \in \mathbb{C}$ heißt EIGENWERT, kurz EW, der komplexen $n \times n$ Matrix A , wenn es mindestens einen von Null verschiedenen komplexen Vektor \mathbf{b} gibt mit

$$A\mathbf{b} = \lambda\mathbf{b}.$$

Jeder derartige Vektor heißt EIGENVEKTOR, kurz EV, von A zum EW λ .

Ist \mathbf{b} ein EV von A zum EW λ und $\alpha \neq 0$ eine komplexe Zahl, so ist $\alpha\mathbf{b}$ ebenfalls ein EV von A zum EW λ .

BEISPIEL II.4.1. Für

$$A = \begin{pmatrix} 0 & -1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

erhalten wir

$$A\mathbf{b} = \begin{pmatrix} -2 \\ -4 \\ 2 \end{pmatrix} = -2\mathbf{b}.$$

Also ist -2 ein EW von A und \mathbf{b} ein zugehöriger EV.

II.4.2. Das charakteristische Polynom. Definitionsgemäß ist $\lambda \in \mathbb{C}$ ein EW der $n \times n$ Matrix A genau dann, wenn das homogene LGS $(A - \lambda I)\mathbf{x} = \mathbf{0}$ eine nicht triviale Lösung besitzt. Gemäß Abschnitt II.3.3 ist dies genau dann der Fall, wenn $\det(A - \lambda I) = 0$ ist.

Die Vorschrift $\lambda \mapsto \det(A - \lambda I)$ beschreibt eine Funktion von \mathbb{C} in \mathbb{C} . Aus den Rechenregeln für Determinanten folgt, dass

$$\det(A - \lambda I) = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0$$

ist mit komplexen Zahlen a_0, \dots, a_{n-1} . D.h., $\det(A - \lambda I)$ ist ein POLYNOM vom Grad n . Wir nennen $\det(A - \lambda I)$ das CHARAKTERISTISCHE POLYNOM von A und bezeichnen es mit $\chi_A(\lambda)$. Es gilt:

Die EW von A sind genau die Nullstellen des charakteristischen Polynoms χ_A .

BEMERKUNG II.4.2. Verschiedentlich wird die Funktion $\lambda \mapsto \det(\lambda I - A)$ auch als charakteristisches Polynom von A bezeichnet. Wegen der Rechenregeln für Determinanten ist

$$\det(\lambda I - A) = (-1)^n \det(A - \lambda I),$$

d.h. die Funktionen $\det(\lambda I - A)$ und $\det(A - \lambda I)$ unterscheiden sich nur durch das Vorzeichen und haben insbesondere die selben Nullstellen.

Man kann zeigen, dass jedes komplexe Polynom in komplexe Linearfaktoren zerlegt werden kann. Daher gibt es ein $r \in \{1, \dots, n\}$, paarweise verschiedene komplexe Zahlen $\lambda_1, \dots, \lambda_r$ und von Null verschiedene natürliche Zahlen k_1, \dots, k_r mit $k_1 + \dots + k_r = n$ und

$$\chi_A(\lambda) = (\lambda_1 - \lambda)^{k_1} \cdot \dots \cdot (\lambda_r - \lambda)^{k_r}.$$

Die Zahl k_i heißt die ALGEBRAISCHE VIELFACHHEIT des EW λ_i . Definitionsgemäß ist die Matrix $A - \lambda_i I$ singular und erfüllt daher $\text{Rang}(A - \lambda_i I) < n$. Die Zahl $n - \text{Rang}(A - \lambda_i I)$ heißt GEOMETRISCHE VIELFACHHEIT des EW λ_i .

Konstruktionsgemäß gilt für jeden Eigenwert:

$$1 \leq \text{geometrische Vielfachheit} \leq \text{algebraische Vielfachheit}.$$

Im allgemeinen stimmen aber algebraische und geometrische Vielfachheit eines EW *nicht* überein.

BEISPIEL II.4.3. Für

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

lautet das charakteristische Polynom

$$\chi_A(\lambda) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2.$$

Also hat A nur den EW 1; die algebraische Vielfachheit ist 2. Aus Abschnitt II.1.4 ergibt sich

$$\text{Rang}(A - I) = \text{Rang} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{Rang} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1.$$

Also ist die geometrische Vielfachheit des EW 1.

Aus den Rechenregeln für Determinanten ergibt sich, dass

$$\chi_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} \text{Spur } A + \dots + \det A$$

ist mit

$$\text{Spur } A = a_{11} + a_{22} + \dots + a_{nn}.$$

Ein Koeffizientenvergleich mit der Zerlegung in Linearfaktoren liefert die nützlichen Formeln

$$\begin{aligned} \det A &= \lambda_1^{k_1} \cdot \dots \cdot \lambda_r^{k_r} \\ \text{Spur } A &= k_1 \lambda_1 + \dots + k_r \lambda_r. \end{aligned}$$

II.4.3. Berechnung der Eigenwerte und Eigenvektoren. Aus dem vorigen Abschnitt ergibt sich folgende Vorgehensweise zur Berechnung der Eigenwerte und Eigenvektoren:

- Stelle das charakteristische Polynom χ_A auf.
- Bestimme alle Nullstellen $\lambda_1, \dots, \lambda_r$ von χ_A .
- Für jeden EW λ_i bestimme mit dem Gaußschen Eliminationsverfahren II.1.3 (S. 54) alle Lösungen des homogenen LGS $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$.

BEISPIEL II.4.4. Betrachte

$$A = \begin{pmatrix} -2 & -8 & -12 \\ 1 & 4 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Das charakteristische Polynom ist

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{pmatrix} -2-\lambda & -8 & -12 \\ 1 & 4-\lambda & 4 \\ 0 & 0 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda) \det \begin{pmatrix} -2-\lambda & -8 \\ 1 & 4-\lambda \end{pmatrix} \\ &= (1-\lambda)[(-2-\lambda)(4-\lambda) + 8] \\ &= (1-\lambda)[-8 - 2\lambda + \lambda^2 + 8] \\ &= (1-\lambda)(\lambda-2)\lambda. \end{aligned}$$

Also sind die EW $\lambda_1 = 0$, $\lambda_2 = 1$ und $\lambda_3 = 2$. Mit dem Gaußschen Eliminationsverfahren ergibt sich:

für den EW 0:

$$\begin{array}{ccc|ccc|ccc} -2 & -8 & -12 & & 1 & 4 & 4 & & 1 & 4 & 4 \\ 1 & 4 & 4 & \rightarrow & 0 & 0 & -4 & \rightarrow & 0 & 0 & 1 \\ 0 & 0 & 1 & & 0 & 0 & 1 & & 0 & 0 & 0 \end{array}$$

$$\Rightarrow x_3 = 0, x_2 = 1, x_1 = -4$$

$$\Rightarrow \begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} \text{ ist ein EV zum EW } 0$$

für den EW 1:

$$\begin{array}{ccc|ccc} -3 & -8 & -12 & & 1 & 3 & 4 \\ 1 & 3 & 4 & \rightarrow & 0 & 1 & 0 \\ 0 & 0 & 0 & & 0 & 0 & 0 \end{array}$$

$$\Rightarrow x_3 = 1, x_2 = 0, x_1 = -4$$

$$\Rightarrow \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} \text{ ist ein EV zum EW } 1$$

für den EW 2:

$$\begin{array}{ccc|ccc|ccc} -4 & -8 & -12 & & 1 & 2 & 4 & & 1 & 2 & 4 \\ 1 & 2 & 4 & \rightarrow & 0 & 0 & 4 & \rightarrow & 0 & 0 & -1 \\ 0 & 0 & -1 & & 0 & 0 & -1 & & 0 & 0 & 0 \end{array}$$

$$\Rightarrow x_3 = 0, x_2 = 1, x_1 = -2$$

$$\implies \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \text{ ist ein EV zum EW } 2.$$

II.4.4. Numerische Berechnung der Eigenwerte und Eigenvektoren*. Die beschriebene Vorgehensweise zur Berechnung der EW und EV ist nur zur „Rechnung per Hand“ bei kleinen Werten von n geeignet. Für die numerische Berechnung mit Computern ist sie zu aufwändig und zu instabil, d.h. mit großen Ungenauigkeiten behaftet. Für diese Aufgabe sind spezielle, besser geeignete Verfahren entwickelt worden.

Die wichtigsten sind:

- (1) Die Potenzmethode und das Verfahren der Rayleigh-Quotienten für die Berechnung des betragsmäßig größten Eigenwertes.
- (2) Das inverse Verfahren von Wielandt und das Verfahren der inversen Rayleigh-Quotienten für die Berechnung des betragsmäßig kleinsten Eigenwertes einer invertierbaren Matrix.
- (3) Das QR-Verfahren zur Berechnung aller Eigenwerte.

Wir werden hier nur die Verfahren aus (1) und (2) besprechen.

ALGORITHMUS II.4.5. POTENZMETHODE zur Berechnung des betragsmäßig größten Eigenwertes.

(0) Gegeben:

A (Matrix)

\mathbf{x}_0 (Startnäherung für einen Eigenvektor mit $\|\mathbf{x}_0\| = 1$)

(1) Für $m = 0, 1, 2, \dots$ führe folgende Schritte aus:

$$\begin{aligned} \mathbf{u}_{m+1} &= A\mathbf{x}_m, \\ \sigma_{m+1} &= \operatorname{sgn}(\langle \mathbf{u}_{m+1}, \mathbf{x}_m \rangle), \\ \lambda_{m+1} &= \sigma_{m+1} \|\mathbf{u}_{m+1}\|, \\ \mathbf{x}_{m+1} &= \sigma_{m+1} \frac{1}{\|\mathbf{u}_{m+1}\|} \mathbf{u}_{m+1}. \end{aligned}$$

Hat A die Eigenwerte μ_1, \dots, μ_n (mehrfache Eigenwerte werden mehrfach aufgeführt!) und gilt $|\mu_1| = \dots = |\mu_p| > |\mu_{p+1}| \geq \dots \geq |\mu_n|$ sowie $\mu_1 = \dots = \mu_p$, kann man zeigen, dass die Zahlen λ_m gegen den Eigenwert μ_1 und die Vektoren \mathbf{x}_m gegen einen Eigenvektor zum Eigenwert μ_1 konvergieren. Dabei wird der Fehler jeweils pro Iteration um den Faktor $\frac{|\mu_{p+1}|}{|\mu_1|}$ reduziert. Die Konvergenzgeschwindigkeit hängt also ab vom Abstand zwischen dem betragsmäßig größten und zweitgrößten Eigenwert.

Falls $|\mu_1| = \dots = |\mu_p| > |\mu_{p+1}| \geq \dots \geq |\mu_n|$ aber $\mu_1 \neq \dots \neq \mu_p$ ist, konvergieren die Zahlen λ_m nach wie vor gegen den Eigenwert μ_1 , aber die

Vektoren \mathbf{x}_m müssen nicht mehr gegen einen zugehörigen Eigenvektor konvergieren.

Das folgende Java-Programm realisiert Algorithmus II.4.5:

```
// power method
public void power() throws LinearAlgebraException {
    double ew = 0.0, diff = 1.0, norm = 1.0;
    int s = 1, iter = 0;
    b = new double[dim];
    norm = Math.sqrt( innerProduct(x, x)/dim );
    if ( norm > EPS )
        multiply(x, x, 1.0/norm);
    else
        throw new LinearAlgebraException(
            " ERROR: zero vector after "+iter+" iterations");
    while ( iter < maxit && diff > EPS ) {
        b = multiply(a, x);
        s = ( innerProduct(b, x) >= 0 ) ? 1 : -1;
        diff = ew;
        norm = Math.sqrt( innerProduct(b, b)/dim );
        if ( norm < EPS )
            throw new LinearAlgebraException(
                " ERROR: zero vector after "+iter+" iterations");
        ew = s*norm;
        residuals[iter] = ew;
        multiply(x, b, 1.0/ew);
        diff = Math.abs( ew - diff );
        iter++;
    }
    iterations = Math.max(iter, 1);
} // end of power
// inner product of two vectors
public double innerProduct(double[] u, double[] v) {
    double prod = 0;
    for( int i = 0; i < dim; i++ )
        prod += u[i]*v[i];
    return prod;
} // end of inner product
// multiply matrix c with vector y
public double[] multiply(double[][] c, double[] y) {
    double[] result = new double[dim];
    for( int i = 0; i < dim; i++ )
        result[i] = innerProduct(c[i], y);
    return result;
} // end of multiply
// multiply vector v with double c and store on u
public void multiply(double[] u, double[] v, double c) {
    for( int i = 0; i < dim; i++ )
        u[i] = v[i]*c;
} // end of multiply
```

ALGORITHMUS II.4.6. RAYLEIGH-QUOTIENTEN-ITERATION zur Bestimmung des betragsmäßig größten Eigenwertes einer symmetrischen Matrix.

- (0) Gegeben: \mathbf{x}_0 (Startwert mit $\|\mathbf{x}_0\| = 1$)
 (1) Für $m = 0, 1, \dots$ führe folgende Schritte aus:

$$\mathbf{u}_{m+1} = A\mathbf{x}_m,$$

$$\mathbf{x}_{m+1} = \frac{1}{\|\mathbf{u}_{m+1}\|} \mathbf{u}_{m+1},$$

$$\lambda_{m+1} = (\mathbf{x}_m, \mathbf{u}_{m+1}).$$

Unter den gleichen Voraussetzungen wie bei Algorithmus II.4.5, kann man zeigen, dass die Zahlen λ_m gegen den betragsmäßig größten Eigenwert von A konvergieren. Dabei wird der Fehler pro Iteration um den Faktor $(\frac{|\mu_{p+1}|}{|\mu_1|})^2$ reduziert. Algorithmus II.4.6 konvergiert also doppelt so schnell wie Algorithmus II.4.5.

Das folgende Java-Programm realisiert Algorithmus II.4.6. Die Methoden `innerProduct` und `multiply` sind die selben wie bei der Potenzmethode.

```
// Rayleigh quotient iteration
public void rayleigh() throws LinearAlgebraException {
    double ew = 0.0, diff = 1.0, norm = 1.0;
    int iter = 0;
    b = new double[dim];
    norm = Math.sqrt( innerProduct(x, x)/dim );
    if ( norm > EPS )
        multiply(x, x, 1.0/norm);
    else
        throw new LinearAlgebraException(
            " ERROR: zero vector after "+iter+" iterations");
    while ( iter < maxit && diff > EPS ) {
        b = multiply(a, x);
        diff = ew;
        ew = innerProduct(b, x)/dim;
        norm = Math.sqrt( innerProduct(b, b)/dim );
        if ( norm < EPS )
            throw new LinearAlgebraException(
                " ERROR: zero vector after "+iter+" iterations");
        residuals[iter] = ew;
        multiply(x, b, 1.0/norm);
        diff = Math.abs( ew - diff );
        iter++;
    }
    iterations = Math.max(iter, 1);
} // end of rayleigh
}
```

Wendet man die Algorithmen II.4.5 und II.4.6 auf die Matrix A^{-1} an und bildet die Kehrwerte der entsprechenden Größen, erhält man einen Algorithmus zur Berechnung des betragsmäßig kleinsten Eigenwertes. Aus Effizienzgründen wird dabei die Matrix A^{-1} *nicht* berechnet, sondern $\mathbf{z} = A^{-1}\mathbf{x}$ durch Lösen des Gleichungssystems $A\mathbf{z} = \mathbf{x}$ bestimmt.

ALGORITHMUS II.4.7. INVERSE ITERATION VON WIELANDT *zur Berechnung des betragsmäßig kleinsten Eigenwertes.*

- (0) Gegeben: \mathbf{x}_0 (Startwert mit $\|\mathbf{x}_0\| = 1$)
- (1) Für $m = 0, 1, \dots$ führe folgende Schritte aus:

(a) Löse das LGS

$$A\mathbf{u}_{m+1} = \mathbf{x}_m.$$

(b) Berechne

$$\begin{aligned}\sigma_{m+1} &= \operatorname{sgn}((\mathbf{u}_{m+1}, \mathbf{x}_m)), \\ \rho_{m+1} &= \frac{\sigma_{m+1}}{\|\mathbf{u}_{m+1}\|}, \\ \mathbf{x}_{m+1} &= \sigma_{m+1} \frac{1}{\|\mathbf{u}_{m+1}\|} \mathbf{u}_{m+1}.\end{aligned}$$

Hat die Matrix A die Eigenwerte μ_1, \dots, μ_n und gilt $|\mu_1| \geq \dots \geq |\mu_{n-r-1}| > |\mu_{n-r}| = \dots = |\mu_n| > 0$ sowie $\mu_{n-r} = \dots = \mu_n$, kann man zeigen, dass die Zahlen ρ_m gegen den Eigenwert μ_n und die Vektoren \mathbf{x}_m gegen einen Eigenvektor von A zum Eigenwert μ_n konvergieren.

Das folgende Java-Programm realisiert Algorithmus II.4.7. Die Methoden `innerProduct` und `multiply` sind die selben wie bei der Potenzmethode. Die Methoden `lrElimination`, `permutation`, `forSolve` und `backSolve` lösen ein lineares Gleichungssystem mit der LR-Zerlegung und sind in Abschnitt II.2.5 wiedergegeben.

```
// inverse power method
public void inversePower() throws LinearAlgebraException {
    double ew = 0.0, diff = 1.0, norm = 1.0;
    int s = 1, iter = 0;
    b = new double[dim];
    d = new double[dim];
    norm = Math.sqrt( innerProduct(x, x)/dim );
    if ( norm > EPS )
        multiply(x, x, 1.0/norm);
    else
        throw new LinearAlgebraException(
            " ERROR: zero vector after "+iter+" iterations");
    lrElimination();
    while ( iter < maxit && diff > EPS ) {
        copy(b, x); // methods of permutation, forSolve, backSolve
        copy(d, x); // take rhs b and give solution x
        permutation();
        forSolve();
        backSolve();
        copy(b, x); // store solution on b
        copy(x, d); // retrieve old x
        s = ( innerProduct(b, x) >= 0 ) ? 1 : -1;
        diff = ew;
        norm = Math.sqrt( innerProduct(b, b)/dim );
        if ( norm < EPS )
            throw new LinearAlgebraException(
                " ERROR: zero vector after "+iter+" iterations");
        ew = s/norm;
        residuals[iter] = ew;
        multiply(x, b, ew);
        diff = Math.abs( ew - diff );
        iter++;
    }
}
```

```

        iterations = Math.max(iter, 1);
    }    // end of inversePower
}
// copy vector v to vector u
public void copy(double[] u, double[] v) {
    for( int i = 0; i < dim; i++ )
        u[i] = v[i];
}    // end of copy

```

ALGORITHMUS II.4.8. INVERSE RAYLEIGH-QUOTIENTEN-ITERATION zur Bestimmung des betragsmäßig kleinsten Eigenwertes einer symmetrischen Matrix.

- (0) Gegeben: \mathbf{x}_0 (Startwert mit $\|\mathbf{x}_0\| = 1$)
 (1) Für $m = 0, 1, \dots$ führe folgende Schritte aus:
 (a) Löse das LGS

$$A\mathbf{u}_{m+1} = \mathbf{x}_m.$$

- (b) Berechne

$$\mathbf{x}_{m+1} = \frac{1}{\|\mathbf{u}_{m+1}\|} \mathbf{u}_{m+1},$$

$$\rho_{m+1} = (\mathbf{x}_m, \mathbf{u}_{m+1})^{-1}.$$

Unter den gleichen Voraussetzungen wie bei Algorithmus II.4.7, kann man zeigen, dass die Zahlen ρ_m gegen den Eigenwert μ_n von A konvergieren. Allerdings konvergiert Algorithmus II.4.8 doppelt so schnell wie Algorithmus II.4.7.

Das folgende Java-Programm realisiert Algorithmus II.4.8. Die Methoden `innerProduct` und `multiply` sind die selben wie bei der Potenzmethode. Die Methoden `lrElimination`, `permutation`, `forSolve` und `backSolve` lösen ein lineares Gleichungssystem mit der LR-Zerlegung und sind in Abschnitt II.2.5 wiedergegeben.

```

// inverse Rayleigh quotient iteration
public void inverseRayleigh() throws LinearAlgebraException {
    double ew = 0.0, diff = 1.0, norm = 1.0;
    int iter = 0;
    b = new double[dim];
    d = new double[dim];
    norm = Math.sqrt( innerProduct(x, x)/dim );
    if ( norm > EPS )
        multiply(x, x, 1.0/norm);
    else
        throw new LinearAlgebraException(
            " ERROR: zero vector after "+iter+" iterations");
    lrElimination();
    while ( iter < maxit && diff > EPS ) {
        copy(b, x);    // methods of permutation, forSolve, backSolve
        copy(d, x);    // take rhs b and give solution x
        permutation();
        forSolve();
        backSolve();
        copy(b, x);    // store solution on b
    }
}

```

```

copy(x, d);          // retrieve old x
diff = ew;
ew = dim/innerProduct(b, x);
norm = Math.sqrt( innerProduct(b, b)/dim );
if ( norm < EPS )
    throw new LinearAlgebraException(
        " ERROR: zero vector after "+iter+" iterations");
residuals[iter] = ew;
multiply(x, b, 1.0/norm);
diff = Math.abs( ew - diff );
iter++;
}
iterations = Math.max(iter, 1);
} // end of inverseRayleigh
}

```

Hat man eine Näherung $\tilde{\mu}$ für einen EW berechnet, kann man einen zugehörigen EV berechnen, indem man das homogene LGS $(A - \tilde{\mu}I)\mathbf{x} = \mathbf{0}$ mit dem Gaußschen Eliminationsverfahren II.1.3 (S. 54) löst. Diese naive Vorgehensweise ist aber nicht empfehlenswert, da sie numerisch instabil, d.h. anfällig gegen Rundungsfehler und Fehler bei der Berechnung von $\tilde{\mu}$ ist. Wesentlich effizienter ist die Anwendung von Algorithmus II.4.7 auf die Matrix $A - \tilde{\mu}I$.

Zur Beschreibung der Vorgehensweise nehmen wir an, dass wir eine Näherung $\tilde{\mu}$ für einen Eigenwert μ_i der Matrix A bestimmt haben, und wollen einen zugehörigen Eigenvektor $\tilde{\mathbf{y}}$ bestimmen. Wir setzen voraus, dass $\tilde{\mu}$ dichter an μ_i liegt als an jedem anderen Eigenwert von A , d.h.

$$|\tilde{\mu} - \mu_i| < |\tilde{\mu} - \mu_j| \text{ für jeden Eigenwert } \mu_j \text{ von } A \text{ mit } \mu_j \neq \mu_i.$$

Dann ist $\mu_i - \tilde{\mu}$ der betragsmäßig kleinste Eigenwert der Matrix $A - \tilde{\mu}I$ und jeder andere Eigenwert dieser Matrix hat einen echt größeren Betrag. Außerdem ist jeder Eigenvektor von $A - \tilde{\mu}I$ zum Eigenwert $\mu_i - \tilde{\mu}$ ein Eigenvektor von A zum Eigenwert μ_i . Daher können wir Algorithmus II.4.7 auf die Matrix $A - \tilde{\mu}I$ anwenden und erhalten eine Folge von Vektoren \mathbf{x}_m , die gegen einen Eigenvektor von A zum Eigenwert μ_i konvergiert.

Im allgemeinen liefern schon wenige Iterationen eine hinreichend genaue Approximation für einen solchen Eigenvektor.

Man beachte, dass in jeder Iteration des Verfahrens ein lineares Gleichungssystem der Form

$$(A - \tilde{\mu}I)\mathbf{u}_{m+1} = \mathbf{x}_m$$

gelöst werden muss.

II.4.5. Rechenregeln. Es gelten u.a. folgende nützliche Regeln zu Eigenwerten und Eigenvektoren:

- A habe den EV \mathbf{b} zum EW λ . Dann haben

$$\alpha A, \quad A + \beta I, \quad \alpha_m A^m + \dots + \alpha_1 A + \alpha_0 I$$

ebenfalls den EV \mathbf{b} , jedoch jeweils zum EW

$$\alpha\lambda, \quad \lambda + \beta, \quad \alpha_m\lambda^m + \dots + \alpha_1\lambda + \alpha_0.$$

- A und A^T haben die gleichen EW, aber zu jeweils verschiedenen EV.
- A ist genau dann invertierbar, wenn alle EW von Null verschieden sind. In diesem Fall ist λ genau dann ein EW von A , wenn λ^{-1} ein EW von A^{-1} ist; die zugehörigen EV stimmen überein.

II.4.6. Ähnliche Matrizen. Zwei $n \times n$ Matrizen A und B heißen **ÄHNLICH**, wenn es eine invertierbare $n \times n$ Matrix T gibt mit $TAT^{-1} = B$. Wegen

$$\begin{aligned} \det(TAT^{-1} - \lambda I) &= \det(T(A - \lambda I)T^{-1}) \\ &= \det T \cdot \det(A - \lambda I) \cdot (\det T)^{-1} \\ &= \det(A - \lambda I) \end{aligned}$$

haben ähnliche Matrizen dasselbe charakteristische Polynom. Daher gilt:

Ähnliche Matrizen haben die gleichen Eigenwerte.

Ist \mathbf{u} ein EV von A zum EW λ und $B = TAT^{-1}$, so ist $T\mathbf{u}$ ein EV von B zum EW λ , denn

$$B(T\mathbf{u}) = TAT^{-1}(T\mathbf{u}) = T\mathbf{A}\mathbf{u} = T(\lambda\mathbf{u}) = \lambda T\mathbf{u}.$$

Eine Matrix A heißt **DIAGONALISIERBAR**, wenn sie zu einer Diagonalmatrix D ähnlich ist. Die Diagonalelemente von D sind dann die EW von A . Ist $A = TDT^{-1}$, so ist der i -te Spaltenvektor von T ein EV von A zum EW D_{ii} . Ist A diagonalisierbar, stimmen die algebraischen und geometrischen Vielfachheiten der EW überein. Beispiel II.4.3 zeigt daher, dass nicht jede Matrix diagonalisierbar ist. Andererseits kann man zeigen, dass z.B. jede reelle symmetrische Matrix diagonalisierbar ist (vgl. Abschnitt II.4.8).

BEISPIEL II.4.9. Betrachte die Matrix A aus Beispiel II.4.4 und bezeichne mit T die Matrix, deren Spalten die drei EV aus Beispiel II.4.4 sind:

$$T = \begin{pmatrix} -4 & -4 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Mit dem Gaußschen Eliminationsverfahren II.2.6 ergibt sich

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

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 4 & 4 & 0 & 0 & 1 & 0.5 & 2 & 2 \end{array} \rightarrow \begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0.5 & 2 & 2 & 0 & 0 & 1 & 0.5 & 2 & 2 \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & -0.5 & -1 & -2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0.5 & 2 & 2 \end{array}.$$

Also ist

$$T^{-1} = \begin{pmatrix} -0.5 & -1 & -2 \\ 0 & 0 & 1 \\ 0.5 & 2 & 2 \end{pmatrix}$$

und

$$T^{-1}AT = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Mit

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

ergibt sich dann beispielsweise

$$\begin{aligned} A^{10} &= (TDT^{-1})^{10} \\ &= TD^{10}T^{-1} \\ &= \begin{pmatrix} -4 & -4 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1024 \end{pmatrix} \begin{pmatrix} -0.5 & -1 & -2 \\ 0 & 0 & 1 \\ 0.5 & 2 & 2 \end{pmatrix} \\ &= \begin{pmatrix} -4 & -4 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 512 & 2048 & 2048 \end{pmatrix} \\ &= \begin{pmatrix} -1024 & -4096 & -4100 \\ 512 & 2048 & 2048 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

II.4.7. Orthogonale Matrizen. Zwei Spaltenvektoren $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^{n \times 1}$ heißen ORTHOGONAL, wenn $\mathbf{s}_1^T \mathbf{s}_2 = 0$ ist. Ein Spaltenvektor $\mathbf{s} \in \mathbb{R}^{n \times 1}$ heißt NORMIERT, wenn $\mathbf{s}^T \mathbf{s} = 1$ ist. Zwei Zeilenvektoren $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^{1 \times n}$ heißen orthogonal, wenn die Spaltenvektoren $\mathbf{z}_1^T, \mathbf{z}_2^T$ orthogonal sind, d.h. $\mathbf{z}_1 \mathbf{z}_2^T = 0$. Analog heißt ein Zeilenvektor $\mathbf{z} \in \mathbb{R}^{1 \times n}$ normiert, wenn der Spaltenvektor \mathbf{z}^T normiert ist, d.h. $\mathbf{z} \mathbf{z}^T = 1$.

Eine $n \times n$ Matrix A heißt ORTHOGONAL, wenn ihre Spaltenvektoren bzw. – was dasselbe ist – ihre Zeilenvektoren normiert und paarweise

orthogonal sind. Aus den Rechenregeln für die Matrixmultiplikation folgt:

$$A \text{ orthogonal} \iff AA^T = A^T A = I.$$

Also gilt:

$$A \text{ orthogonal} \iff A^T = A^{-1}.$$

Wegen

$$1 = \det I = \det(AA^T) = \det A \cdot \det A^T = (\det A)^2$$

gilt außerdem:

$$A \text{ orthogonal} \implies |\det A| = 1.$$

BEMERKUNG II.4.10. Sei $A = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ eine orthogonale 3×3 Matrix mit $\det A = 1$. Dann ist $(0; \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ ein kartesisches Koordinatensystem (vgl. Abschnitt I.4.9). Die Abbildung $\mathbf{x} \mapsto A\mathbf{x}$ beschreibt den Übergang von dem Koordinatensystem $(0; \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ in das Koordinatensystem $(0; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Geometrisch wird sie durch die Hintereinanderschaltung von Drehungen und Spiegelungen beschrieben. Insbesondere bleiben unter dieser Abbildung die Längen von Vektoren erhalten.

BEISPIEL II.4.11. Die Drehmatrix

$$\begin{pmatrix} \cos \omega & \sin \omega \\ -\sin \omega & \cos \omega \end{pmatrix}$$

ist orthogonal.

BEISPIEL II.4.12. Sei $\mathbf{u} \in \mathbb{R}^{n \times 1}$ ein normierter Spaltenvektor. Die Matrix

$$H(\mathbf{u}) = (\mathbf{e}_1 - 2u_1\mathbf{u}, \mathbf{e}_2 - 2u_2\mathbf{u}, \dots, \mathbf{e}_n - 2u_n\mathbf{u})$$

heißt die zu \mathbf{u} gehörige HOUSEHOLDER-MATRIX. Wegen

$$\begin{aligned} (\mathbf{e}_i - 2u_i\mathbf{u})^T(\mathbf{e}_i - 2u_i\mathbf{u}) &= \underbrace{\mathbf{e}_i^T \mathbf{e}_i}_{=1} - 4u_i \underbrace{\mathbf{e}_i^T \mathbf{u}}_{=u_i} + 4u_i^2 \underbrace{\mathbf{u}^T \mathbf{u}}_{=1} \\ &= 1 \end{aligned}$$

und für $i \neq j$

$$\begin{aligned} (\mathbf{e}_i - 2u_i\mathbf{u})^T(\mathbf{e}_j - 2u_j\mathbf{u}) &= \underbrace{\mathbf{e}_i^T \mathbf{e}_j}_{=0} - 2u_j \underbrace{\mathbf{e}_i^T \mathbf{u}}_{=u_i} - 2u_i \underbrace{\mathbf{u}^T \mathbf{e}_j}_{=u_j} + 4u_i u_j \underbrace{\mathbf{u}^T \mathbf{u}}_{=1} \\ &= 0 \end{aligned}$$

ist sie orthogonal. Ist speziell $n = 3$, so beschreibt sie geometrisch eine Spiegelung an der zu \mathbf{u} senkrechten Ebene durch den Nullpunkt.

II.4.8. Symmetrische Matrizen. Wir erinnern daran, dass nicht jede Matrix diagonalisierbar ist und dass i.a. die algebraische und die geometrische Vielfachheit eines EW nicht übereinstimmen. Es gilt jedoch:

Sei A eine symmetrische $n \times n$ Matrix, d.h. $A^T = A$. Dann gilt:

- Alle EW von A sind reell.
- Für jeden EW von A stimmen algebraische und geometrische Vielfachheit überein.
- A ist diagonalisierbar; die Matrix T in $A = TDT^{-1}$ kann als orthogonale Matrix gewählt werden.
- Es gibt eine Basis des \mathbb{R}^n , die aus paarweise orthogonalen EV von A besteht.

BEISPIEL II.4.13. Für

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

erhalten wir das charakteristische Polynom

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{pmatrix} 1 - \lambda & 3 & 0 \\ 3 & -2 - \lambda & -1 \\ 0 & -1 & 1 - \lambda \end{pmatrix} \\ &= (1 - \lambda) \det \begin{pmatrix} -2 - \lambda & -1 \\ -1 & 1 - \lambda \end{pmatrix} - 3 \det \begin{pmatrix} 3 & -1 \\ 0 & 1 - \lambda \end{pmatrix} \\ &= (1 - \lambda)[(-2 - \lambda)(1 - \lambda) - 1] - 9(1 - \lambda) \\ &= (1 - \lambda)(-2 + \lambda + \lambda^2 - 1 - 9) \\ &= (1 - \lambda)(\lambda^2 + \lambda - 12) \\ &= (1 - \lambda)(\lambda - 3)(\lambda + 4). \end{aligned}$$

Also sind die EW $\lambda_1 = 1$, $\lambda_2 = 3$ und $\lambda_3 = -4$. Das Gaußsche Eliminationsverfahren II.1.3 (S. 54) liefert:

für $\lambda_1 = 1$:

$$\begin{array}{ccc|ccc|ccc} 0 & 3 & 0 & & 3 & -3 & -1 & & 3 & -3 & -1 \\ 3 & -3 & -1 & \rightarrow & 0 & -1 & 0 & \rightarrow & 0 & -1 & 0 \\ 0 & -1 & 0 & & 0 & 3 & 0 & & 0 & 0 & 0 \end{array}$$

$$\implies x_3 = 1, x_2 = 0, x_1 = \frac{1}{3}$$

$$\implies \begin{pmatrix} \frac{\sqrt{10}}{10} \\ 0 \\ \frac{3\sqrt{10}}{10} \end{pmatrix} \text{ ist normierter EV zu } \lambda_1 = 1$$

für $\lambda_2 = 3$:

$$\begin{array}{ccc} -2 & 3 & 0 \\ 3 & -5 & -1 \\ 0 & -1 & -2 \end{array} \rightarrow \begin{array}{ccc} -2 & 3 & 0 \\ 0 & -0.5 & -1 \\ 0 & -1 & -2 \end{array} \rightarrow \begin{array}{ccc} -2 & 3 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{array}$$

$$\implies x_3 = 1, x_2 = -2, x_1 = -3$$

$$\implies \begin{pmatrix} -\frac{3\sqrt{14}}{14} \\ -\frac{\sqrt{14}}{7} \\ \frac{\sqrt{14}}{14} \end{pmatrix} \text{ ist normierter EV zu } \lambda_2 = 3$$

für $\lambda_3 = -4$:

$$\begin{array}{ccc} 5 & 3 & 0 \\ 3 & 2 & -1 \\ 0 & -1 & 5 \end{array} \rightarrow \begin{array}{ccc} 5 & 3 & 0 \\ 0 & 0.2 & -1 \\ 0 & -1 & 5 \end{array} \rightarrow \begin{array}{ccc} 5 & 3 & 0 \\ 0 & -1 & 5 \\ 0 & 0 & 0 \end{array}$$

$$\implies x_3 = 1, x_2 = 5, x_1 = -3$$

$$\implies \begin{pmatrix} -\frac{3\sqrt{35}}{35} \\ \frac{\sqrt{35}}{7} \\ \frac{\sqrt{35}}{35} \end{pmatrix} \text{ ist normierter EV zu } \lambda_3 = -4.$$

Also ist $A = TDT^{-1}$ mit

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{pmatrix},$$

$$T = \begin{pmatrix} \frac{\sqrt{10}}{10} & -\frac{3\sqrt{14}}{14} & -\frac{3\sqrt{35}}{35} \\ 0 & -\frac{\sqrt{14}}{7} & \frac{\sqrt{35}}{7} \\ \frac{3\sqrt{10}}{10} & \frac{\sqrt{14}}{14} & \frac{\sqrt{35}}{35} \end{pmatrix}.$$

II.4.9. Die Schursche Normalform. Matrizen wie

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

mit mehrfachen EW, deren algebraische Vielfachheit größer ist als ihre geometrische Vielfachheit, sind nicht zu Diagonalmatrizen ähnlich. Für diese Matrizen gibt es keine Basis aus EV. Dementsprechend muss man das Konzept der Diagonalisierbarkeit abschwächen. Wir wollen daher zeigen, dass *jede* Matrix A zu einer oberen Dreiecksmatrix S ähnlich

ist, d.h. dass es eine invertierbare Matrix B gibt mit $B^{-1}AB = S$. Die Diagonalelemente von S sind dann natürlich die EW von A . Eine solche Matrix S nennt man SCHURSCHE NORMALFORM von A .

Zum Beweis unserer Behauptung betrachten wir einen EW λ von A . Zu λ gibt es mindestens einen EV \mathbf{u} . Diesen können wir zu einer Basis des \mathbb{C}^n ergänzen. Die so gefundenen Vektoren fassen wir als Spaltenvektoren zu einer Matrix B_1 zusammen, d.h. $B_1 = (\mathbf{u}, *, \dots, *)$. Dann gilt

$$\begin{aligned} B_1^{-1}AB_1 &= B_1^{-1}(\lambda\mathbf{u}, *, \dots, *) \\ &= \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} \end{aligned}$$

mit einer $(n-1) \times (n-1)$ -Matrix A_1 . Mit A_1 können wir obigen Prozess wiederholen und erhalten eine $(n-1) \times (n-1)$ -Matrix \tilde{B}_2 mit

$$\tilde{B}_2^{-1}A_1\tilde{B}_2 = \begin{pmatrix} \mu & * & \dots & * \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix}$$

mit einer $(n-2) \times (n-2)$ -Matrix A_2 . Für

$$B_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \tilde{B}_2 & \\ 0 & & & \end{pmatrix}$$

gilt dann

$$B_2^{-1}B_1^{-1}AB_1B_2 = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \mu & 0 & \dots & 0 \\ 0 & 0 & & & \\ \vdots & \vdots & & A_2 & \\ 0 & 0 & & & \end{pmatrix}.$$

Nach $n-1$ Schritten erhalten wir also Matrizen B_1, B_2, \dots, B_{n-1} , sodass

$$B_{n-1}^{-1} \cdot \dots \cdot B_1^{-1}AB_1 \cdot \dots \cdot B_{n-1} = S$$

obere Dreiecksgestalt hat. $B = B_1 \cdot \dots \cdot B_{n-1}$ und S sind die gesuchten Matrizen.

II.4.10. Hauptvektoren. Wir betrachten eine beliebige Matrix A und ihre Schursche Normalform S . Dann ist $S = B^{-1}AB$ mit einer invertierbaren Matrix B . Sei λ ein m -facher EW von A , $m \geq 2$. Indem

wir ggf. B durch BP mit einer geeigneten Permutationsmatrix P ersetzen, können wir annehmen, dass die ersten m Diagonalelemente von S gleich λ sind. Dann ist $S - \lambda I$ von der Form

$$S - \lambda I = \begin{pmatrix} E & F \\ 0 & G \end{pmatrix}$$

mit

$$E = \begin{pmatrix} 0 & * & \dots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & * \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{C}^{m \times m},$$

$$F \in \mathbb{C}^{m \times (n-m)},$$

$$G \in \mathbb{C}^{(n-m) \times (n-m)}.$$

Wie man leicht nachrechnet, ist $E^m = 0$. Daher sind die ersten m Spalten von $(S - \lambda I)^m$ gleich Null, d.h.

$$(S - \lambda I)^m \mathbf{e}_i = 0 \quad \text{für } i = 1, \dots, m.$$

Wegen

$$\begin{aligned} (S - \lambda I)^m &= (B^{-1}AB - \lambda I)^m \\ &= (B^{-1}(A - \lambda I)B)^m \\ &= B^{-1}(A - \lambda I)^m B \end{aligned}$$

folgt

$$(A - \lambda I)^m B \mathbf{e}_i = B(S - \lambda I)^m \mathbf{e}_i = 0 \quad \text{für } i = 1, \dots, m.$$

Ein Vektor $\mathbf{u} \in \mathbb{C}^n$ mit $\mathbf{u} \neq 0$ und $(A - \lambda I)^m \mathbf{u} = 0$ nennt man einen HAUPTVEKTOR zum EW λ . Wegen $(A - \lambda I)^0 \mathbf{u} = I\mathbf{u} = \mathbf{u} \neq 0$ und $(A - \lambda I)^m \mathbf{u} = 0$ gibt es eine Zahl $\ell \geq 1$ mit $(A - \lambda I)^{\ell-1} \mathbf{u} \neq 0$ und $(A - \lambda I)^\ell \mathbf{u} = 0$. Diese Zahl heißt STUFE des Hauptvektors. Eigenvektoren sind somit Hauptvektoren der Stufe 1.

Wir haben also gezeigt, dass A zum m -fachen EW λ m linear unabhängige Hauptvektoren besitzt, nämlich $B\mathbf{e}_1, \dots, B\mathbf{e}_m$. (Man beachte, dass wir die Spalten der ursprünglichen Matrix B evtl. vertauscht haben!) Man kann zeigen, dass Hauptvektoren zu verschiedenen EW immer linear unabhängig sind. Damit folgt:

Zu jeder Matrix $A \in \mathbb{C}^{n \times n}$ gibt es eine Basis des \mathbb{C}^n , die aus Hauptvektoren von A besteht.

Die Berechnung einer Basis aus Hauptvektoren und einer Schurschen Normalform geschieht mit folgendem Algorithmus:

ALGORITHMUS II.4.14. *Berechnung der Hauptvektoren und der Schurschen Normalform einer Matrix A .*

- (1) *Bestimme alle EW der Matrix A .*
- (2) *Zu jedem EW λ von A bestimme die maximal mögliche Zahl linear unabhängiger EV. Falls diese Zahl kleiner ist als die algebraische Vielfachheit m von λ (geometrische Vielfachheit $<$ algebraische Vielfachheit), ergänze diese Vektoren zu einer Basis von $\text{Kern}(A - \lambda I)^m$.*
- (3) *Setze $B = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ mit den Hauptvektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$ aus Schritt (2) und berechne $S = B^{-1}AB$.*

BEISPIEL II.4.15. Betrachte

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -2 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Das charakteristische Polynom lautet

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -\lambda & 1 & -1 \\ -2 & 3 - \lambda & -1 \\ -1 & 1 & 1 - \lambda \end{vmatrix} \\ &= -\lambda(\lambda - 3)(\lambda - 1) + 1 + 2 + (\lambda - 3) - \lambda - 2(\lambda - 1) \\ &= -\lambda(\lambda - 3)(\lambda - 1) - 2(\lambda - 1) \\ &= -(\lambda - 1)[\lambda^2 - 3\lambda + 2] \\ &= -(\lambda - 1)^2(\lambda - 2). \end{aligned}$$

Also ist 2 ein einfacher EW und 1 ein doppelter EW. Die Bestimmungsgleichung für einen EV zum EW 2 lautet

$$\begin{aligned} &\begin{pmatrix} -2 & 1 & -1 \\ -2 & 1 & -1 \\ -1 & 1 & -1 \end{pmatrix} \mathbf{u} = 0 \\ \implies &\begin{pmatrix} -2 & 1 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \mathbf{u} = 0 \\ &\implies \mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Die Bestimmungsgleichung für einen EV zum EW 1 ist

$$\begin{aligned} &\begin{pmatrix} -1 & 1 & -1 \\ -2 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} \mathbf{v} = 0 \\ \implies &\begin{pmatrix} -1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{v} = 0 \end{aligned}$$

$$\implies \mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Weiter ist

$$\begin{aligned} (A - I)^2 &= \begin{pmatrix} -1 & 1 & -1 \\ -2 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & -1 \\ -2 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}. \end{aligned}$$

Daher lautet die Bestimmungsgleichung für einen Hauptvektor zum EW 1

$$-x_1 + x_2 = 0.$$

Daher ist

$$\mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

ein Hauptvektor zum EW 1, der linear unabhängig vom EV \mathbf{v} ist. $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ ist die gesuchte Basis. Wegen

$$A\mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{w} - \mathbf{v}$$

ergibt sich mit

$$B = (\mathbf{u}, \mathbf{v}, \mathbf{w}) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

die Schursche Normalform von A zu

$$\begin{aligned} B^{-1}AB &= B^{-1}(2\mathbf{u}, \mathbf{v}, \mathbf{w} - \mathbf{v}) \\ &= (2\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 - \mathbf{e}_2) \\ &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

II.5. Quadratische Formen

II.5.1. Definition. Eine Funktion $p : \mathbb{R}^n \rightarrow \mathbb{R}$ der Form

$$p(\mathbf{x}) = \alpha + \mathbf{a}^T \mathbf{x} + \mathbf{x}^T A \mathbf{x}$$

mit $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n \times n}$ symmetrisch heißt QUADRATISCHES POLYNOM in den Variablen x_1, \dots, x_n . Im Fall $\alpha = 0$ und $\mathbf{a} = \mathbf{0}$ heißt p ein HOMOGENES QUADRATISCHES POLYNOM oder auch eine QUADRATISCHE FORM. Ist zusätzlich A eine Diagonalmatrix, so heißt p REIN QUADRATISCH.

Ist p ein quadratisches Polynom, so heißt die Menge

$$\{\mathbf{x} \in \mathbb{R}^n : p(\mathbf{x}) = 0\}$$

eine QUADRIK. Eine Quadrik hat NORMALFORM, wenn $\mathbf{x}^T A \mathbf{x}$ rein quadratisch ist und $\alpha + \mathbf{a}^T \mathbf{x}$ durch keine Transformation der Form $\mathbf{x} = \mathbf{b} + B \mathbf{y}$ verkürzt werden kann.

BEISPIEL II.5.1. Für

$$\alpha = 4, \quad \mathbf{a} = \begin{pmatrix} -2 \\ -4 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

erhalten wir das quadratische Polynom

$$p(\mathbf{x}) = 4 - 2x_1 - 4x_2 + x_1^2 + x_2^2.$$

Quadratische Ergänzung liefert für die zugehörige Quadrik

$$\begin{aligned} 0 &= 4 - 2x_1 + x_1^2 - 4x_2 + x_2^2 \\ &= 4 + (x_1 - 1)^2 - 1 + (x_2 - 2)^2 - 4 \\ &= (x_1 - 1)^2 + (x_2 - 2)^2 - 1. \end{aligned}$$

Dies ist die Gleichung eines Kreises in der (x_1, x_2) -Ebene mit Mittelpunkt $(1, 2)$ und Radius 1. Die Koordinatentransformation $y_1 = x_1 - 1$, $y_2 = x_2 - 2$ (Verschiebung des Ursprungs) liefert die Normalform $y_1^2 + y_2^2 - 1 = 0$.

BEISPIEL II.5.2. Der Spannungstensor eines unter Krafteinwirkung stehenden elastischen Körpers P hat die Form

$$S = \begin{pmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \sigma_y & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \sigma_z \end{pmatrix}$$

mit vom Ort abhängigen Matrixelementen. Er ist symmetrisch. Betrachte in einem festen Punkt P einen Schnitt durch P senkrecht zu dem Normalenvektor \mathbf{n} mit $|\mathbf{n}| = 1$. Dann können wir den Spannungsvektor $S\mathbf{n}$ in die zu \mathbf{n} parallele Normalenspannung \mathbf{t} und die dazu senkrechte Schubspannung \mathbf{s} zerlegen, d.h. $S\mathbf{n} = \mathbf{t} + \mathbf{s}$. Wegen $\mathbf{n} \cdot \mathbf{s} = 0$ und $\mathbf{n} \cdot \mathbf{t} = |\mathbf{t}|$ gilt $\mathbf{n}^T S \mathbf{n} = |\mathbf{t}|$, d.h. der Betrag der Normalenspannung berechnet sich aus der quadratischen Form $q(\mathbf{x}) = \mathbf{x}^T S \mathbf{x}$.

BEISPIEL II.5.3. Ein einfaches Problem der Ausgleichsrechnung ist die Bestimmung einer Geraden $y = ax + b$, die möglichst genau durch die Messpunkte $(x_1, y_1), \dots, (x_n, y_n)$ in dem Sinne verläuft, dass das quadratische Polynom $q(u, v) = (y_1 - ux_1 - v)^2 + \dots + (y_n - ux_n - v)^2$ in $(u, v) = (a, b)$ ein Minimum annimmt.

II.5.2. Reduktion auf Normalform. Die Quadrik zu einem gegebenen quadratischen Polynom

$$p(\mathbf{x}) = \alpha + \mathbf{a}^T \mathbf{x} + \mathbf{x}^T A \mathbf{x}$$

kann durch die folgenden zwei Schritte auf Normalform gebracht werden:

- Bestimme wie in Abschnitt II.4.7 eine Diagonalmatrix D und eine orthogonale Matrix T mit $A = TDT^{-1}$. Die Variablentransformation $\mathbf{x} = T\mathbf{y}$ liefert

$$\begin{aligned} p(\mathbf{y}) &= \alpha + \mathbf{a}^T T\mathbf{y} + \mathbf{y}^T D\mathbf{y} \\ &= \alpha + \beta_1 y_1 + \dots + \beta_n y_n + \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 \end{aligned}$$

mit den EW $\lambda_1, \dots, \lambda_n$ von A . Man beachte, dass jeder EW entsprechend seiner algebraischen Vielfachheit mehrfach aufgeführt wird!

- Die Koordinatentransformation

$$z_i = \begin{cases} y_i & \text{falls } \lambda_i = 0 \\ y_i + \frac{\beta_i}{2\lambda_i} & \text{falls } \lambda_i \neq 0 \end{cases}$$

(d.h. eine quadratische Ergänzung) liefert die Normalform

$$\begin{aligned} p(\mathbf{z}) &= \gamma + \mathbf{c}^T \mathbf{z} + \mathbf{z}^T D\mathbf{z} \\ &= \gamma + \delta_1 z_1 + \dots + \delta_n z_n + \lambda_1 z_1^2 + \dots + \lambda_n z_n^2 \end{aligned}$$

mit

$$\delta_i = \begin{cases} \beta_i & \text{falls } \lambda_i = 0 \\ 0 & \text{falls } \lambda_i \neq 0 \end{cases}$$

und

$$\gamma = \alpha - \frac{1}{4} \sum' \frac{\beta_i^2}{\lambda_i},$$

wobei \sum' andeutet, dass nur die Beiträge mit $\lambda_i \neq 0$ aufsummiert werden.

BEISPIEL II.5.4. Betrachte das quadratische Polynom p mit

$$\alpha = -\frac{49}{48}, \quad \mathbf{a} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Gemäß Beispiel II.4.13 (S. 87) ist

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{pmatrix},$$

$$T = \begin{pmatrix} \frac{\sqrt{10}}{10} & -\frac{3\sqrt{14}}{14} & -\frac{3\sqrt{35}}{35} \\ 0 & -\frac{\sqrt{14}}{7} & \frac{\sqrt{35}}{7} \\ \frac{3\sqrt{10}}{10} & \frac{\sqrt{14}}{14} & \frac{\sqrt{35}}{35} \end{pmatrix}.$$

Damit ergibt sich

$$\mathbf{a}^T T = \left(0, -\frac{\sqrt{14}}{7}, \frac{\sqrt{35}}{7}\right)$$

und

$$\begin{aligned} \gamma &= -\frac{49}{48} - \frac{1}{4} \left[\frac{14}{49 \cdot 3} + \frac{35}{49 \cdot (-4)} \right] \\ &= -\frac{49}{48} - \frac{1}{4} \left[\frac{2}{21} - \frac{5}{28} \right] \\ &= -\frac{49}{48} + \frac{1}{48} \\ &= -1. \end{aligned}$$

Also ist die Normalform der zugehörigen Quadrik

$$x_1^2 + 3x_2^2 - 4x_3^2 - 1 = 0.$$

Dies ist die Gleichung eines einschaligen Hyperboloids.

II.5.3. Normalformen der ebenen Quadriken. Im \mathbb{R}^2 gibt es die folgenden 9 verschiedenen Normalformen der Quadriken. Dabei sind a, b, p von Null verschiedene reelle Zahlen.

Rang $A = 2$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \text{ Ellipse, bzw. Kreis falls } |a| = |b|,$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0 \text{ leere Menge,}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0 \text{ Hyperbel,}$$

$$x^2 + a^2 y^2 = 0 \text{ Punkt,}$$

$$x^2 - a^2 y^2 = 0 \text{ Paar sich schneidender Geraden,}$$

Rang $A = 1$:

$$x^2 - 2py = 0 \text{ Parabel,}$$

$$x^2 - a^2 = 0 \text{ Paar paralleler Geraden,}$$

$$x^2 + a^2 = 0 \text{ leere Menge,}$$

$$x^2 = 0 \text{ Gerade } x = 0.$$

II.5.4. Normalformen der räumlichen Quadriken. Im \mathbb{R}^3 gibt es die folgenden 17 verschiedenen Normalformen der Quadriken. Dabei sind a, b, c, p von Null verschiedene reelle Zahlen.

Rang $A = 3$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0 \text{ Ellipsoid bzw. Kugel, falls } |a| = |b| = |c|,$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} + 1 = 0 \text{ leere Menge,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0 \text{ zweischaliges Hyperboloid,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0 \text{ einschaliges Hyperboloid,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0 \text{ Punkt,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \text{ Kegel,}$$

Rang $A = 2$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2pz = 0 \text{ elliptisches Paraboloid,}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 2pz = 0 \text{ hyperbolisches Paraboloid,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0 \text{ leere Menge,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \text{ elliptischer Zylinder,}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + 1 = 0 \text{ hyperbolischer Zylinder,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0 \text{ Gerade,}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0 \text{ Paar sich schneidender Ebenen,}$$

Rang $A = 1$:

$$x^2 - 2py = 0 \text{ parabolischer Zylinder,}$$

$$x^2 - a^2 = 0 \text{ Paar paralleler Ebenen,}$$

$$x^2 + a^2 = 0 \text{ leere Menge,}$$

$$x^2 = 0 \text{ Ebene } x = 0.$$

II.5.5. Positiv definite Matrizen. Die Bestimmung der Extremwerte einer reellen Funktion in n reellen Veränderlichen ist eng verbunden mit der Frage, wann eine quadratische Form $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ für $\mathbf{x} \neq \mathbf{0}$ nur positive oder nur negative Werte annimmt.

Eine quadratische Form $q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ bzw. die zugehörige symmetrische Matrix heißt **POSITIV DEFINIT** bzw. **NEGATIV DEFINIT**, wenn für alle $\mathbf{x} \neq \mathbf{0}$ gilt $q(\mathbf{x}) > 0$ bzw. $q(\mathbf{x}) < 0$. Die quadratische Form bzw. die zugehörige Matrix heißt **INDEFINIT**, wenn sie sowohl positive als auch negative Werte annimmt. Sie heißt **POSITIV SEMIDEFINIT** bzw. **NEGATIV SEMIDEFINIT**, wenn für alle $\mathbf{x} \neq \mathbf{0}$ gilt $q(\mathbf{x}) \geq 0$ bzw. $q(\mathbf{x}) \leq 0$.

BEISPIEL II.5.5. In der Relativitätstheorie spielt die quadratische Form

$$q(x, y, z, t) = x^2 + y^2 + z^2 - c^2 t^2$$

eine wichtige Rolle. Sie ist indefinit; es gibt „raumartige“ Vektoren $\mathbf{u} \in \mathbb{R}^4$ mit $q(\mathbf{u}) > 0$ und „zeitartige“ Vektoren $\mathbf{v} \in \mathbb{R}^4$ mit $q(\mathbf{v}) < 0$.

Es gilt folgende Charakterisierung:

Eine symmetrische Matrix ist genau dann positiv definit bzw. negativ definit, wenn alle ihre EW positiv bzw. negativ sind.
 Sie ist indefinit genau dann, wenn sie positive und negative EW besitzt.
 Sie ist genau dann positiv semidefinit bzw. negativ semidefinit, wenn alle ihre EW nicht negativ bzw. nicht positiv sind.

Die positive Definitheit kann aus dem Vorzeichen geeigneter Determinanten abgelesen werden:

Eine symmetrische Matrix ist genau dann positiv definit, wenn die Determinanten der n Hauptmatrizen

$$\begin{aligned} H_1 &= a_{11}, \\ H_2 &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \\ &\vdots \\ H_k &= \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}, \\ &\vdots \end{aligned}$$

$$H_n = A$$

alle positiv sind.

BEISPIEL II.5.6. Für

$$A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 6 & -1 \\ 2 & -1 & 4 \end{pmatrix}$$

lauten die Hauptmatrizen und ihre Determinanten

$$\begin{aligned} H_1 &= 5, & \det H_1 &= 5, \\ H_2 &= \begin{pmatrix} 5 & -2 \\ -2 & 6 \end{pmatrix}, & \det H_2 &= 26, \\ H_3 &= A, & \det H_3 &= 83. \end{aligned}$$

Also ist A positiv definit.

BEISPIEL II.5.7. Für

$$A = \begin{pmatrix} 5 & -3 & 9 \\ -3 & 3 & -3 \\ 9 & -3 & 5 \end{pmatrix}$$

lauten die Hauptmatrizen und ihre Determinanten

$$\begin{aligned} H_1 &= 5, & \det H_1 &= 5, \\ H_2 &= \begin{pmatrix} 5 & -3 \\ -3 & 3 \end{pmatrix}, & \det H_2 &= 6, \\ H_3 &= A, & \det H_3 &= -96. \end{aligned}$$

Das charakteristische Polynom ist

$$\chi_A(\lambda) = -(4 + \lambda)(\lambda^2 - 17\lambda + 24).$$

Also sind die EW -4 , $\frac{17-\sqrt{193}}{2}$ und $\frac{17+\sqrt{193}}{2}$. Die Matrix ist indefinit.

II.6. Vektorräume und lineare Abbildungen

II.6.1. Vektorräume. Eine nichtleere Menge V , in der man zu je zwei Elementen $\mathbf{u}, \mathbf{v} \in V$ eine Summe $\mathbf{u} + \mathbf{v} \in V$ und zu jedem Element $\mathbf{u} \in V$ und jeder Zahl $\alpha \in \mathbb{R}$ das α -fache $\alpha\mathbf{u} \in V$ bilden kann, heißt ein (REELLER) VEKTORRAUM, wenn folgende acht Rechengesetze erfüllt sind:

VEKTORRAUMAXIOME:

- (1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ für alle $\mathbf{u}, \mathbf{v} \in V$ (KOMMUTATIVGESETZ).

- (2) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ (ASSOZIATIVGESETZ).
- (3) Es gibt ein $\mathbf{0} \in V$, NULLELEMENT oder NULLVEKTOR genannt, mit $\mathbf{u} + \mathbf{0} = \mathbf{u}$ für alle $\mathbf{u} \in V$.
- (4) Zu jedem $\mathbf{u} \in V$ gibt es ein mit $-\mathbf{u}$ bezeichnetes Element, für das gilt $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
- (5) $1\mathbf{u} = \mathbf{u}$ für alle $\mathbf{u} \in V$.
- (6) $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$ für alle $\mathbf{u} \in V, \alpha, \beta \in \mathbb{R}$.
- (7) $\alpha(\mathbf{u} + \mathbf{v}) = (\alpha\mathbf{u}) + (\alpha\mathbf{v})$ für alle $\mathbf{u}, \mathbf{v} \in V, \alpha \in \mathbb{R}$.
- (8) $(\alpha + \beta)\mathbf{u} = (\alpha\mathbf{u}) + (\beta\mathbf{u})$ für alle $\mathbf{u} \in V, \alpha, \beta \in \mathbb{R}$.

Die Elemente eines Vektorraumes nennt man **VEKTOREN**. Statt $\mathbf{u} + (-\mathbf{v})$ schreibt man $\mathbf{u} - \mathbf{v}$.

BEISPIEL II.6.1. Der dreidimensionale Anschauungsraum ist ein Vektorraum (vgl. Paragraph I.4). $\mathbb{R}^{n \times 1}$, $\mathbb{R}^{1 \times m}$ und $\mathbb{R}^{n \times m}$ sind Vektorräume (vgl. Abschnitt II.1.1).

BEISPIEL II.6.2. Sei $I \subset \mathbb{R}$ ein Intervall und \mathcal{F} die Menge aller Abbildungen $f : I \rightarrow \mathbb{R}$. Für je zwei Abbildungen $f, g : I \rightarrow \mathbb{R}$ und jede Zahl $\alpha \in \mathbb{R}$ können wir die Abbildungen $f + g : I \rightarrow \mathbb{R}$ und $\alpha f : I \rightarrow \mathbb{R}$ definieren durch

$$x \mapsto f(x) + g(x) \quad \text{bzw.} \quad x \mapsto \alpha f(x).$$

Wie man leicht nachrechnet, gelten für diese Vorschriften die obigen Rechenregeln. \mathcal{F} ist also ein Vektorraum.

II.6.2. Unterräume. Eine nicht leere Teilmenge U eines Vektorraumes V heißt **UNTERVEKTORRAUM** oder **UNTERRAUM** von V , wenn für je zwei Vektoren $\mathbf{u}, \mathbf{v} \in U$ und jede Zahl $\alpha \in \mathbb{R}$ gilt $\mathbf{u} + \mathbf{v} \in U$ und $\alpha\mathbf{u} \in U$.

BEISPIEL II.6.3. Die Menge U aller Spaltenvektoren mit n Komponenten, deren letzte Komponente verschwindet, ist ein Unterraum von $\mathbb{R}^{n \times 1}$. Ebenso ist die Menge W aller Spaltenvektoren mit n Komponenten, deren i -te Komponente ($i \in \{1, \dots, n\}$ fest) verschwindet ein Unterraum von $\mathbb{R}^{n \times 1}$.

BEISPIEL II.6.4. Für jede $m \times n$ Matrix A ist

$$\text{Kern } A = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\},$$

die Lösungsmenge des homogenen LGS, ein Unterraum von \mathbb{R}^n .

II.6.3. Linearkombination und lineare Hülle. Jede aus endlich vielen Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ eines Vektorraumes V gebildete Summe der Form $\alpha_1\mathbf{v}_1 + \dots + \alpha_k\mathbf{v}_k$ mit Koeffizienten $\alpha_i \in \mathbb{R}$ heißt eine **LINEARKOMBINATION** der \mathbf{v}_i . Eine solche Linearkombination wird **TRIVIAL** genannt, wenn alle Koeffizienten α_i gleich Null sind. Die Menge aller

Linearkombinationen von gegebenen Vektoren $\mathbf{u}_1, \dots, \mathbf{u}_m \in V$ heißt die LINEARE HÜLLE von $\mathbf{u}_1, \dots, \mathbf{u}_m$ und wird mit $\text{span}(\mathbf{u}_1, \dots, \mathbf{u}_m)$ bezeichnet:

$$\text{span}(\mathbf{u}_1, \dots, \mathbf{u}_m) = \{\alpha_1 \mathbf{u}_1 + \dots + \alpha_m \mathbf{u}_m : \alpha_1, \dots, \alpha_m \in \mathbb{R}\}.$$

Die lineare Hülle gegebener Vektoren ist stets ein Untervektorraum.

BEISPIEL II.6.5. Für

$$\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$$

ist

$$\begin{aligned} \text{span}(\mathbf{u}, \mathbf{v}) &= \{s\mathbf{u} + t\mathbf{v} : s, t \in \mathbb{R}\} \\ &= \left\{ \begin{pmatrix} s \\ 2s + 3t \\ -s + 2t \end{pmatrix} : s, t \in \mathbb{R} \right\}. \end{aligned}$$

Geometrisch ist dies die von \mathbf{u} und \mathbf{v} aufgespannte Ebene durch den Nullpunkt.

II.6.4. Lineare Abhängigkeit. Endlich viele Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ eines Vektorraumes V heißen LINEAR ABHÄNGIG, wenn es Zahlen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ gibt, die nicht alle gleich Null sind, mit

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}.$$

Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ heißen LINEAR UNABHÄNGIG, wenn sie nicht linear abhängig sind, d.h. wenn aus $\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$ folgt $\alpha_1 = \dots = \alpha_k = 0$.

BEISPIEL II.6.6. (1) Die Vektoren

$$\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} -4 \\ 0 \\ 5 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} -3 \\ 2 \\ 8 \end{pmatrix}$$

sind linear abhängig, denn $\mathbf{u} + \mathbf{v} - \mathbf{w} = \mathbf{0}$.

(2) Betrachte die Vektoren

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -3 \\ 0 \\ 12 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 4 \\ 8 \\ 0 \end{pmatrix}.$$

Die Bedingung $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_3 \mathbf{u}_3 = \mathbf{0}$ führt auf das homogene LGS

$$\begin{aligned} \alpha_1 - 3\alpha_2 + 4\alpha_3 &= 0 \\ 2\alpha_1 + 8\alpha_3 &= 0 \\ -2\alpha_1 + 12\alpha_2 &= 0. \end{aligned}$$

Das Gaußsche Eliminationsverfahren II.1.3 (S. 54) liefert

$$\begin{array}{ccc|ccc} 1 & -3 & 4 & & 1 & -3 & 4 & & 1 & -3 & 4 \\ 2 & 0 & 8 & \rightarrow & 0 & 6 & 0 & \rightarrow & 0 & 6 & 0 \\ -2 & 12 & 0 & & 0 & 6 & 8 & & 0 & 0 & 8 \end{array}$$

Also besitzt das LGS nur die Nulllösung; die Vektoren \mathbf{u}_1 , \mathbf{u}_2 , \mathbf{u}_3 sind linear unabhängig.

Betrachte n Vektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$ des $\mathbb{R}^{m \times 1}$. Dann ist die Bestimmung von Zahlen $\alpha_1, \dots, \alpha_n$ mit $\alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n = \mathbf{0}$ äquivalent zum Lösen des homogenen LGS $A\mathbf{x} = \mathbf{0}$ mit $A = (\mathbf{u}_1, \dots, \mathbf{u}_n)$. Wegen Abschnitt II.1.4 erhalten wir somit:

$$\begin{array}{c} \mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbb{R}^{m \times 1} \text{ sind linear unabhängig} \\ \iff \text{Rang}(\mathbf{u}_1, \dots, \mathbf{u}_n) = n. \end{array}$$

Hieraus folgt insbesondere:

Ist $n > m$, so sind je n Vektoren des \mathbb{R}^m stets linear abhängig.

II.6.5. Basis und Dimension. Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ eines Vektorraumes V heißen eine BASIS von V , wenn gilt:

- $\mathbf{v}_1, \dots, \mathbf{v}_n$ sind linear unabhängig und
- $V = \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$.

Man kann die zweite Bedingung auch so formulieren:

- Für jeden Vektor $\mathbf{v} \in V$ sind die Vektoren $\mathbf{v}, \mathbf{v}_1, \dots, \mathbf{v}_n$ linear abhängig.

Die Bedeutung einer Basis ergibt sich aus:

Ist $\mathbf{v}_1, \dots, \mathbf{v}_n$ eine Basis von V , dann gibt es zu jedem Vektor $\mathbf{u} \in V$ genau ein n -Tupel $\alpha_1, \dots, \alpha_n$ reeller Zahlen mit $\mathbf{u} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n$, d.h. die Darstellung von \mathbf{u} als Linearkombination der \mathbf{v}_i ist eindeutig.

Um dies einzusehen, beachte man, dass aus $V = \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$ die Existenz mindestens eines n -Tupels $\alpha_1, \dots, \alpha_n$ mit den gewünschten Eigenschaften folgt. Ist β_1, \dots, β_n ein zweites derartiges n -Tupel folgt

$$\begin{aligned} & (\alpha_1 - \beta_1)\mathbf{v}_1 + \dots + (\alpha_n - \beta_n)\mathbf{v}_n \\ &= \underbrace{(\alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n)}_{=\mathbf{u}} - \underbrace{(\beta_1 \mathbf{v}_1 + \dots + \beta_n \mathbf{v}_n)}_{=\mathbf{u}} \\ &= \mathbf{0}. \end{aligned}$$

Da $\mathbf{v}_1, \dots, \mathbf{v}_n$ linear unabhängig sind, folgt $\alpha_1 = \beta_1, \dots, \alpha_n = \beta_n$.

Eine weitere wichtige Eigenschaft von Basen ist:

Ist $\mathbf{v}_1, \dots, \mathbf{v}_n$ eine Basis von V und $m > n$, so sind je m Vektoren aus V linear abhängig.

Um dies einzusehen, wähle m Vektoren $\mathbf{w}_1, \dots, \mathbf{w}_m \in V$. Zu jedem \mathbf{w}_i gibt es Zahlen $\alpha_{i1}, \dots, \alpha_{in}$ mit

$$\mathbf{w}_i = \alpha_{i1}\mathbf{v}_1 + \dots + \alpha_{in}\mathbf{v}_n.$$

Gemäß Abschnitt II.1.4 hat die Matrix $A = (\alpha_{ji})_{n \times m}$ höchstens den Rang n (beachte $m > n$). Also hat das homogene LGS $A\mathbf{x} = \mathbf{0}$ eine nicht-triviale Lösung $\lambda_1, \dots, \lambda_m$. Für diese Zahlen gilt dann

$$\begin{aligned} & \lambda_1\mathbf{w}_1 + \dots + \lambda_m\mathbf{w}_m \\ &= (\alpha_{11}\lambda_1 + \dots + \alpha_{1m}\lambda_m)\mathbf{v}_1 + \dots + (\alpha_{n1}\lambda_1 + \dots + \alpha_{nm}\lambda_m)\mathbf{v}_n \\ &= \mathbf{0}. \end{aligned}$$

Aus dem soeben Gezeigten folgt:

Sind $\mathbf{v}_1, \dots, \mathbf{v}_n$ und $\mathbf{w}_1, \dots, \mathbf{w}_m$ zwei Basen desselben Vektorraumes V , so ist $m = n$.

Gibt es überhaupt eine Basis, so ist also die Zahl der Basiselemente eindeutig festgelegt. Daher ist die folgende Definition sinnvoll:

Ein Vektorraum V heißt **ENDLICH DIMENSIONAL**, wenn es endlich viele Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_d$ in V gibt, die eine Basis bilden. Die Zahl d heißt dann die **DIMENSION** von V und wird mit $\dim V$ bezeichnet.

BEISPIEL II.6.7. (1) Die Vektoren $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ aus Beispiel II.6.6 bilden eine Basis des \mathbb{R}^3 .

(2) Es ist $\dim \mathbb{R}^{n \times 1} = n$, $\dim \mathbb{R}^{1 \times m} = m$, $\dim \mathbb{R}^{m \times n} = mn$.

(3) Eine Basis des $\mathbb{R}^{m \times n}$ wird aus den Matrizen A_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$, gebildet, deren Elemente alle verschwinden mit Ausnahme des i -ten Elements in der j -ten Spalte, das gleich 1 ist.

(4) Der Vektorraum \mathcal{F} aus Beispiel II.6.2 ist nicht endlich dimensional.

BEISPIEL II.6.8. Sei A eine $n \times n$ Matrix mit $r = \text{Rang } A < n$. Gemäß Beispiel II.6.4 ist Kern A ein Vektorraum. Aus Abschnitt II.1.4 folgt $\dim \text{Kern } A = n - r$.

Endlich dimensionale Vektorräume haben folgende wichtige Eigenschaft:

In einem n -dimensionalen Vektorraum V gilt:

- Je n linear unabhängige Vektoren bilden eine Basis.
- Je $n + 1$ Vektoren sind linear abhängig.

II.6.6. Skalarprodukte. Eine Abbildung $(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ heißt ein SKALARPRODUKT auf dem Vektorraum V , wenn folgende Eigenschaften erfüllt sind:

SKALARPRODUKTAXIOME:

- (1) $(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})$ für alle $\mathbf{u}, \mathbf{v} \in V$ (SYMMETRIE).
- (2) $(\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}) = \alpha(\mathbf{u}, \mathbf{w}) + \beta(\mathbf{v}, \mathbf{w})$ für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \alpha, \beta \in \mathbb{R}$ (LINEARITÄT).
- (3) $(\mathbf{u}, \mathbf{u}) \geq 0$ für alle $\mathbf{u} \in V$ und
 $(\mathbf{u}, \mathbf{u}) = 0 \iff \mathbf{u} = \mathbf{0}$ (POSITIVITÄT).

BEISPIEL II.6.9. Durch

$$(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{v} \text{ für } \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n \times 1}$$

wird ein Skalarprodukt, das EUKLIDISCHE SKALARPRODUKT, auf \mathbb{R}^n definiert.

BEISPIEL II.6.10. Sei A eine symmetrische, positiv definite $n \times n$ Matrix. Definiere die Abbildung $(\cdot, \cdot)_A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ durch

$$(\mathbf{u}, \mathbf{v})_A = \mathbf{u}^T A \mathbf{v}.$$

Wegen der Symmetrie von A ist diese Abbildung symmetrisch. Wegen der Rechenregeln für die Matrixmultiplikation ist sie linear. Da A positiv definit ist, ist sie auch positiv. $(\cdot, \cdot)_A$ ist also ein Skalarprodukt auf \mathbb{R}^n .

Seien V ein Vektorraum, (\cdot, \cdot) ein Skalarprodukt auf V und $\mathbf{u}, \mathbf{v} \in V$ zwei Vektoren. Nehme zunächst an, dass $\mathbf{v} \neq \mathbf{0}$ ist. Wegen der Positivität ist dann

$$\alpha = \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})}$$

definiert. Aus den Eigenschaften des Skalarproduktes folgt

$$\begin{aligned} 0 &\leq (\mathbf{u} - \alpha\mathbf{v}, \mathbf{u} - \alpha\mathbf{v}) \\ &= (\mathbf{u}, \mathbf{u}) - 2\alpha(\mathbf{u}, \mathbf{v}) + \alpha^2(\mathbf{v}, \mathbf{v}) \\ &= (\mathbf{u}, \mathbf{u}) - \frac{(\mathbf{u}, \mathbf{v})^2}{(\mathbf{v}, \mathbf{v})} \end{aligned}$$

also

$$(\mathbf{u}, \mathbf{v})^2 \leq (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}).$$

Sei nun $\mathbf{v} = \mathbf{0}$. Dann folgt aus der Linearität und Symmetrie

$$\begin{aligned}(\mathbf{u}, \mathbf{0}) &= (\mathbf{u}, \mathbf{u} - \mathbf{u}) \\ &= (\mathbf{u}, \mathbf{u}) - (\mathbf{u}, \mathbf{u}) \\ &= 0\end{aligned}$$

also

$$\begin{aligned}(\mathbf{u}, \mathbf{v}) &= 0 \\ &\leq (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}).\end{aligned}$$

Dies beweist:

CAUCHY-SCHWARZSCHE UNGLEICHUNG: Sei (\cdot, \cdot) ein Skalarprodukt auf dem Vektorraum V . Dann gilt für alle $\mathbf{u}, \mathbf{v} \in V$

$$(\mathbf{u}, \mathbf{v})^2 \leq (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}).$$

II.6.7. Normen. Eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ heißt eine NORM auf dem Vektorraum V , wenn folgende Eigenschaften erfüllt sind:

NORMAXIOME:

- (1) $\|\mathbf{u}\| \geq 0$ für alle $\mathbf{u} \in V$ und $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$ (POSITIVITÄT).
- (2) $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$ für alle $\mathbf{u} \in V, \alpha \in \mathbb{R}$ (HOMOGENITÄT).
- (3) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ für alle $\mathbf{u}, \mathbf{v} \in V$ (DREIECKSUNGLEICHUNG).

BEISPIEL II.6.11. Der Betrag von Vektoren des dreidimensionalen Anschauungsraumes ist eine Norm (vgl. Abschnitt I.4.4).

Aus den Skalarproduktaxiomen und der Cauchy-Schwarzschen Ungleichung folgt:

Jedes Skalarprodukt (\cdot, \cdot) definiert durch

$$\|\mathbf{u}\| = \sqrt{(u, u)}$$

eine Norm.

BEISPIEL II.6.12. (1) Durch

$$\|\mathbf{u}\|_2 = \sqrt{\mathbf{u}^T \mathbf{u}}$$

wird eine Norm auf $\mathbb{R}^{n \times 1}$ definiert, die EUKLIDISCHE NORM.

(2) Für jede symmetrische, positiv definite $n \times n$ Matrix A wird durch

$$\|\mathbf{u}\|_A = \sqrt{\mathbf{u}^T A \mathbf{u}}$$

eine Norm auf $\mathbb{R}^{n \times 1}$ definiert.

BEISPIEL II.6.13. Durch

$$\|\mathbf{u}\|_1 = |u_1| + \dots + |u_n|$$

und

$$\|\mathbf{u}\|_\infty = \max\{|u_i| : 1 \leq i \leq n\}$$

werden zwei Normen auf \mathbb{R}^n definiert. Man kann zeigen, dass sie von keinem Skalarprodukt definiert werden.

In Abbildung II.6.1 sind die Einheitsbälle $\{\mathbf{u} \in V : \|\mathbf{u}\| \leq 1\}$ in \mathbb{R}^2 für die Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ skizziert.

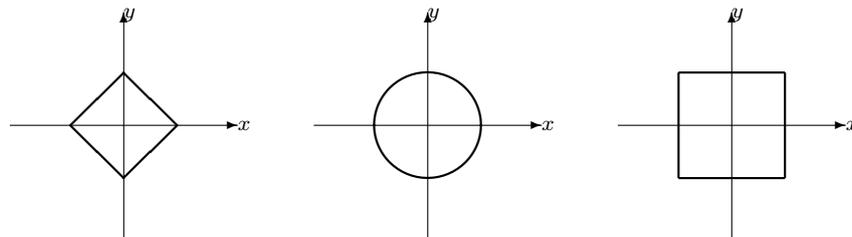


ABBILDUNG II.6.1. Einheitsbälle in \mathbb{R}^2 für die Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ (v.l.n.r)

II.6.8. Orthogonalität. Sei (\cdot, \cdot) ein Skalarprodukt auf dem Vektorraum V . Zwei Vektoren $\mathbf{u}, \mathbf{v} \in V$ heißen ORTHOGONAL (bzgl. (\cdot, \cdot)), wenn $(\mathbf{u}, \mathbf{v}) = 0$ ist. Ein Vektor $\mathbf{w} \in V$ heißt NORMIERT, wenn $(\mathbf{w}, \mathbf{w}) = 1$ ist. Ein System von endlich vielen Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ heißt ein ORTHONORMALSYSTEM, wenn die Vektoren normiert und paarweise orthogonal sind.

BEISPIEL II.6.14. (1) Die Einheitsvektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

bilden ein Orthonormalsystem bzgl. des Euklidischen Skalarproduktes $(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{v}$ auf dem \mathbb{R}^2 .

(2) Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

ist symmetrisch, positiv definit. Bzgl. des Skalarproduktes $(\mathbf{u}, \mathbf{v})_A = \mathbf{u}^T A \mathbf{v}$ sind die Vektoren $\mathbf{e}_1, \mathbf{e}_2$ nicht orthogonal.

Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ eines Orthonormalsystems sind immer linear unabhängig. Denn ist $\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k = 0$, folgt für jedes $i \in \{1, \dots, k\}$

$$\begin{aligned} 0 &= (\alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k, \mathbf{v}_i) \\ &= \alpha_1 (\mathbf{v}_1, \mathbf{v}_i) + \dots + \alpha_k (\mathbf{v}_k, \mathbf{v}_i) \\ &= \alpha_i (\mathbf{v}_i, \mathbf{v}_i) \\ &= \alpha_i. \end{aligned}$$

Umgekehrt kann mit dem SCHMIDTSCHEN ORTHOGONALISIERUNGSVERFAHREN aus einem System $\mathbf{v}_1, \dots, \mathbf{v}_k$ linear unabhängiger Vektoren ein Orthonormalsystem $\mathbf{w}_1, \dots, \mathbf{w}_k$ mit

$$\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \text{span}(\mathbf{w}_1, \dots, \mathbf{w}_k)$$

erzeugt werden. Dieses Verfahren verläuft wie folgt:

(1) Setze

$$\mathbf{w}_1 = \frac{1}{\sqrt{(\mathbf{v}_1, \mathbf{v}_1)}} \mathbf{v}_1.$$

(2) Für $i = 2, \dots, k$ setze

$$\tilde{\mathbf{w}}_i = \mathbf{v}_i - (\mathbf{v}_i, \mathbf{w}_1) \mathbf{w}_1 - \dots - (\mathbf{v}_i, \mathbf{w}_{i-1}) \mathbf{w}_{i-1}$$

und

$$\mathbf{w}_i = \frac{1}{\sqrt{(\tilde{\mathbf{w}}_i, \tilde{\mathbf{w}}_i)}} \tilde{\mathbf{w}}_i.$$

BEISPIEL II.6.15. Für die Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -3 \\ 0 \\ 12 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 4 \\ 8 \\ 0 \end{pmatrix}$$

und das Euklidische Skalarprodukt liefert das Schmidtsche Orthogonalisierungsverfahren:

$$\begin{aligned} \mathbf{w}_1 &= \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}, \\ \tilde{\mathbf{w}}_2 &= \mathbf{v}_2 + 9\mathbf{w}_1 \\ &= \begin{pmatrix} 0 \\ 6 \\ 6 \end{pmatrix}, \\ \mathbf{w}_2 &= \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \end{aligned}$$

$$\begin{aligned}\tilde{\mathbf{w}}_3 &= \mathbf{v}_3 - \frac{20}{3}\mathbf{w}_1 - \frac{8}{\sqrt{2}}\mathbf{w}_2 \\ &= \begin{pmatrix} \frac{16}{9} \\ -\frac{4}{9} \\ \frac{4}{9} \end{pmatrix}, \\ \mathbf{w}_3 &= \begin{pmatrix} \frac{4}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}.\end{aligned}$$

II.6.9. Lineare Abbildungen. Eine Abbildung $L : V \rightarrow W$ eines Vektorraumes V in einen Vektorraum W heißt LINEAR, wenn für alle $\mathbf{v}_1, \mathbf{v}_2 \in V$ und $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt

$$L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1L(\mathbf{v}_1) + \alpha_2L(\mathbf{v}_2).$$

BEISPIEL II.6.16. (1) Die Abbildung $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$, die jedem Vektor seine i -te Komponente zuordnet, $i \in \{1, \dots, n\}$ fest, ist linear.

(2) Ist A eine $m \times n$ Matrix, so ist die Abbildung $\mathbf{x} \mapsto A\mathbf{x}$ eine lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m .

II.6.10. Matrixdarstellung. Wir betrachten zwei endlich dimensionale Vektorräume V und W mit Basen $\mathbf{v}_1, \dots, \mathbf{v}_n$ bzw. $\mathbf{w}_1, \dots, \mathbf{w}_m$ und eine lineare Abbildung $L : V \rightarrow W$. Für jedes $i \in \{1, \dots, n\}$ gibt es dann Zahlen $\alpha_{1i}, \dots, \alpha_{mi} \in \mathbb{R}$ mit

$$(II.6.1) \quad L(\mathbf{v}_i) = \alpha_{1i}\mathbf{w}_1 + \dots + \alpha_{mi}\mathbf{w}_m.$$

Setze

$$(II.6.2) \quad A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}.$$

Ist $\mathbf{v} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$ ein beliebiger Vektor aus V , so hat $L(\mathbf{v})$ die Darstellung $y_1\mathbf{w}_1 + \dots + y_m\mathbf{w}_m$ mit eindeutig definierten Koeffizienten. Aus der Linearität von L folgt:

$$\begin{aligned}& y_1\mathbf{w}_1 + \dots + y_m\mathbf{w}_m \\ &= L(\mathbf{v}) \\ &= x_1L(\mathbf{v}_1) + \dots + x_nL(\mathbf{v}_n) \\ &= x_1(\alpha_{11}\mathbf{w}_1 + \dots + \alpha_{m1}\mathbf{w}_m) + \dots + x_n(\alpha_{1n}\mathbf{w}_1 + \dots + \alpha_{mn}\mathbf{w}_m) \\ &= (\alpha_{11}x_1 + \dots + \alpha_{1n}x_n)\mathbf{w}_1 + \dots + (\alpha_{m1}x_1 + \dots + \alpha_{mn}x_n)\mathbf{w}_m.\end{aligned}$$

Da die Koeffizienten eindeutig bestimmt sind, folgt

$$y_j = \alpha_{j1}x_1 + \dots + \alpha_{jn}x_n, \quad j = 1, \dots, m,$$

d.h.

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Dies beweist:

Der Koeffizientenvektor von $L(\mathbf{v})$ bzgl. der Basis $\mathbf{w}_1, \dots, \mathbf{w}_m$ ist das Produkt des Koeffizientenvektor von \mathbf{v} bzgl. der Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ mit der Matrix A aus (II.6.1), (II.6.2).

In diesem Sinne werden also lineare Abbildungen zwischen endlich dimensionalen Räumen durch Matrizen dargestellt. Die Darstellung hängt von den gewählten Basen ab.

II.6.11. Komposition linearer Abbildungen. Wir betrachten drei endlich dimensionale Vektorräume U , V und W und zwei lineare Abbildungen $L_1 : U \rightarrow V$ und $L_2 : V \rightarrow W$. Dann ist ihre Komposition $L_2 \circ L_1 : U \rightarrow W$ auch eine lineare Abbildung. Wählen wir in jedem der drei Vektorräume eine Basis aus, so werden L_1 , L_2 und $L_2 \circ L_1$ bzgl. dieser Basen durch Matrizen A , B und C dargestellt. Aus dem vorigen Abschnitt und der Definition des Matrixproduktes aus Abschnitt II.2.1 folgt dann mit etwas Rechnen $C = BA$, d.h.:

Der Komposition linearer Abbildungen entspricht die Multiplikation der zugehörigen Matrizen.

II.6.12. Basiswechsel. Wir betrachten einen endlich dimensionalen Vektorraum V und eine lineare Abbildung $L : V \rightarrow V$. Wir wählen zwei Basen $\mathbf{v}_1, \dots, \mathbf{v}_k$ und $\mathbf{w}_1, \dots, \mathbf{w}_k$ von V . Die Abbildung L werde bzgl. dieser Basen durch die Matrizen A und B dargestellt, d.h.

$$\begin{aligned} A &\text{ entspricht } L : \text{span}(\{\mathbf{v}_i\}) \rightarrow \text{span}(\{\mathbf{v}_i\}), \\ B &\text{ entspricht } L : \text{span}(\{\mathbf{w}_j\}) \rightarrow \text{span}(\{\mathbf{w}_j\}). \end{aligned}$$

Welche Beziehung besteht zwischen A und B ?

Um diese Frage zu beantworten, betrachten wir die identische Abbildung $J : V \rightarrow V$ mit $J(\mathbf{v}) = \mathbf{v}$ für alle $\mathbf{v} \in V$. Allerdings versehen wir V im Urbild mit der \mathbf{v} -Basis und im Bild mit der \mathbf{w} -Basis. Bzgl. dieser Basen wird J durch eine Matrix C dargestellt.

J ist bijektiv und besitzt eine Inverse J^{-1} . Aus dem vorigen Abschnitt folgt, dass J^{-1} durch C^{-1} dargestellt wird.

Wir haben daher das folgende Diagramm

$$\begin{array}{ccc} \text{span}(\{\mathbf{v}_i\}) & \xrightarrow{L,A} & \text{span}(\{\mathbf{v}_i\}) \\ J,C \downarrow & & \downarrow J^{-1},C^{-1} \\ \text{span}(\{\mathbf{w}_j\}) & \xrightarrow{L,B} & \text{span}(\{\mathbf{w}_j\}) \end{array}$$

Aus diesem Diagramm und dem vorigen Abschnitt folgt:

$$A = C^{-1}BC, \quad B = CAC^{-1}.$$

Wegen Abschnitt II.4.5 können wir dies auch so formulieren:

Ähnliche Matrizen stellen dieselbe lineare Abbildung bzgl. verschiedener Basen dar.

BEISPIEL II.6.17. Die lineare Abbildung $L : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ werde bzgl. der Basis, die aus den üblichen Einheitsvektoren besteht, durch die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$$

dargestellt. Gemäß Beispiel II.6.6 (2) ist

$$\mathbb{R}^3 = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 12 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ 0 \end{pmatrix} \right\}.$$

Bzgl. dieser Basis werde L durch B dargestellt. Der Wechsel von der Standardbasis in diese neue Basis wird durch die Matrix

$$C = \begin{pmatrix} -2 & 1 & -\frac{1}{2} \\ -\frac{1}{3} & \frac{1}{6} & 0 \\ \frac{1}{2} & -\frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

dargestellt. Damit folgt

$$\begin{aligned} B = CAC^{-1} &= \begin{pmatrix} -2 & 1 & -\frac{1}{2} \\ -\frac{1}{3} & \frac{1}{6} & 0 \\ \frac{1}{2} & -\frac{1}{8} & \frac{1}{8} \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} \begin{pmatrix} 1 & -3 & 4 \\ 2 & 0 & 8 \\ -2 & 12 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 1 & -\frac{1}{2} \\ -\frac{1}{3} & \frac{1}{6} & 0 \\ \frac{1}{2} & -\frac{1}{8} & \frac{1}{8} \end{pmatrix} \begin{pmatrix} -1 & 33 & 20 \\ 2 & 60 & 56 \\ 3 & 99 & 92 \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} \frac{5}{2} & -\frac{111}{2} & -30 \\ \frac{2}{3} & -1 & \frac{8}{3} \\ -\frac{3}{8} & \frac{171}{8} & \frac{29}{2} \end{pmatrix}.$$

KAPITEL III

Stetigkeit

III.1. Folgen

III.1.1. Definition. Eine ZAHLENFOLGE oder kurz FOLGE ist eine Abbildung $\mathbb{N} \rightarrow \mathbb{R}$, die jeder natürlichen Zahl n eine reelle Zahl x_n zuordnet. Eine Folge wird mit $(x_n)_{n \in \mathbb{N}}$ oder x_0, x_1, \dots bezeichnet. Die Zahlen x_n heißen die GLIEDER der Folge.

BEISPIEL III.1.1. (1) $x_n = a$ für alle $n \in \mathbb{N}$; a, a, \dots (konstante Folge).

(2) $x_n = x_0 + nd$; $x_0, x_0 + d, x_0 + 2d, \dots$ (arithmetische Folge).

(3) $x_n = x_0 q^n$; $x_0, x_0 q, x_0 q^2, \dots$ (geometrische Folge).

(4) $x_n = (n + 1)^2$; $1, 4, 9, \dots$ (Folge der Quadratzahlen).

(5) $x_0 = 2$, $x_{n+1} = \frac{1}{2}(x_n + \frac{2}{x_n})$ für alle n (rekursiv definierte Folge).

(6) $x_0 = 1, x_1 = 1, x_{n+1} = x_n + x_{n-1}$ für alle $n \geq 1$ (Folge der Fibonacci Zahlen, rekursiv definiert).

BEMERKUNG III.1.2. Analog zu obiger Definition bezeichnet man Abbildungen von \mathbb{N} in eine Menge A auch als Folgen. So gibt es Folgen komplexer Zahlen ($A = \mathbb{C}$), Folgen von Vektoren ($A = \mathbb{R}^k$) usw..

III.1.2. Rechenregeln. Folgen werden gliedweise addiert, subtrahiert, multipliziert und dividiert. Bei der Division müssen dabei natürlich die Nenner alle von Null verschieden sein.

BEISPIEL III.1.3. Für die Folgen $(x_n)_{n \in \mathbb{N}} = ((n + 1)^2)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}} = (2^n)_{n \in \mathbb{N}}$ erhalten wir

$$(x_n + y_n)_{n \in \mathbb{N}} = ((n + 1)^2 + 2^n)_{n \in \mathbb{N}} \quad (\text{Summenfolge})$$

$$(x_n y_n)_{n \in \mathbb{N}} = ((n + 1)^2 2^n)_{n \in \mathbb{N}} \quad (\text{Produktfolge})$$

$$\left(\frac{x_n}{y_n} \right)_{n \in \mathbb{N}} = ((n + 1)^2 2^{-n})_{n \in \mathbb{N}} \quad (\text{Quotientenfolge}).$$

III.1.3. Beschränktheit. Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt BESCHRÄNKT, wenn es eine reelle Zahl K gibt, sodass für alle Folgenglieder gilt $|x_n| \leq K$.

BEISPIEL III.1.4. (1) $(x_n)_{n \in \mathbb{N}} = \left(\left(-\frac{1}{2} \right)^n \right)_{n \in \mathbb{N}}$: Für alle $n \in \mathbb{N}$ gilt

$$|x_n| = \left| \left(-\frac{1}{2} \right)^n \right| = \frac{1}{2^n} \leq 1.$$

Die Folge ist beschränkt ($K = 1$).

(2) $(x_n)_{n \in \mathbb{N}} = ((n+1)^2)_{n \in \mathbb{N}}$: Die Folge ist nicht beschränkt.

(3) Wähle eine Zahl $a < 0$ und setze für $n \in \mathbb{N}$

$$x_n = \left(1 + \frac{a}{n}\right)^n.$$

Für alle $n \geq |a|$ gilt

$$0 \leq 1 + \frac{a}{n} \leq 1$$

und damit $0 \leq x_n \leq 1$. Da nur endlich viele natürliche Zahlen $\leq |a|$ sind, existiert

$$K_0 = \max \left\{ \left|1 + \frac{a}{n}\right|^n : n \leq |a| \right\}$$

Mit $K = \max\{1, K_0\}$ folgt $|x_n| \leq K$ für alle $n \in \mathbb{N}$. Die Folge ist also beschränkt.

III.1.4. Monotonie. Eine Folge heißt MONOTON WACHSEND bzw. MONOTON FALLEND, wenn für alle $n \in \mathbb{N}$ gilt $x_n \leq x_{n+1}$ bzw. $x_n \geq x_{n+1}$. Eine Folge heißt MONOTON, wenn sie monoton wachsend oder monoton fallend ist.

BEISPIEL III.1.5. (1) $(2^n)_{n \in \mathbb{N}}$ ist monoton wachsend.

(2) $\left(\left(\frac{1}{2}\right)^n\right)_{n \in \mathbb{N}}$ ist monoton fallend.

(3) $\left(\left(-\frac{1}{2}\right)^n\right)_{n \in \mathbb{N}}$ ist nicht monoton.

III.1.5. Teilfolgen. Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge und $n_0 < n_1 < n_2 < \dots$ eine monoton wachsende Indexfolge, so heißt $x_{n_0}, x_{n_1}, x_{n_2}, \dots$ bzw. $(x_{n_k})_{k \in \mathbb{N}}$ eine TEILFOLGE der Folge $(x_n)_{n \in \mathbb{N}}$.

BEISPIEL III.1.6. $\left(\left(\frac{1}{2}\right)^{2k}\right)_{k \in \mathbb{N}}$ und $\left(-\left(\frac{1}{2}\right)^{2k+1}\right)_{k \in \mathbb{N}}$ sind Teilfolgen der Folge $\left(\left(-\frac{1}{2}\right)^n\right)_{n \in \mathbb{N}}$. Sie entsprechen den Indexfolgen der geraden bzw. ungeraden Indizes.

III.2. Grenzwerte von Folgen

III.2.1. Definition. Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt KONVERGENT, wenn es eine Zahl x^* gibt, sodass für jedes noch so kleine $\varepsilon > 0$ ein Index n_ε existiert derart, dass für alle $n \geq n_\varepsilon$ gilt $|x_n - x^*| \leq \varepsilon$.

Mit anderen Worten: Für jedes noch so kleine $\varepsilon > 0$ haben – mit Ausnahme von endlich vielen – alle Folgenglieder einen Abstand $\leq \varepsilon$ von x^* . Die Zahl x^* ist eindeutig bestimmt, sie heißt GRENZWERT oder LIMES der Folge. Wir schreiben dann

$$x_n \xrightarrow[n \rightarrow \infty]{} x^* \quad \text{oder} \quad x^* = \lim_{n \rightarrow \infty} x_n.$$

Konvergente Folgen mit Grenzwert Null nennt man NULLFOLGEN.

BEISPIEL III.2.1. (1) Es ist

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} = 0.$$

DENN: Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $n_\varepsilon \geq \frac{1}{\varepsilon}$. Für alle $n \geq n_\varepsilon$ gilt

$$\begin{aligned} n+1 &\geq n_\varepsilon + 1 \geq \frac{1}{\varepsilon} + 1 > \frac{1}{\varepsilon} \\ \implies 0 &< \frac{1}{n+1} \leq \frac{1}{n_\varepsilon + 1} < \varepsilon. \end{aligned}$$

(2) Es ist

$$\lim_{n \rightarrow \infty} \left(-\frac{1}{2}\right)^n = 0.$$

DENN: Sei $\varepsilon > 0$. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $2^{n_\varepsilon} \geq \frac{1}{\varepsilon}$. Um dies einzusehen, betrachte die Dualzahldarstellung von $\frac{1}{\varepsilon}$ und setze n_ε gleich der Zahl der Vorkommastellen plus 1. Für alle $n \geq n_\varepsilon$ gilt

$$\begin{aligned} 2^n &\geq 2^{n_\varepsilon} \geq \frac{1}{\varepsilon} \\ \implies 0 &< \left| \left(-\frac{1}{2}\right)^n \right| = 2^{-n} \leq 2^{-n_\varepsilon} \leq \varepsilon. \end{aligned}$$

(3) Die Folge $((-1)^n)_{n \in \mathbb{N}}$ konvergiert nicht. DENN: Ist $a \in \mathbb{R}$ beliebig, setze $\varepsilon = \frac{1}{2} \max\{|a-1|, |a+1|\}$. Dann haben alle Folgenglieder mit geradem oder ungeradem Index einen Abstand $> \varepsilon$ von a , je nachdem ob $a \leq 0$ oder $a > 0$ ist.

III.2.2. Rechenregeln. Seien $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen. Dann gilt:

$\begin{aligned} \lim_{n \rightarrow \infty} (x_n + y_n) &= \lim_{n \rightarrow \infty} x_n + \lim_{n \rightarrow \infty} y_n \\ \lim_{n \rightarrow \infty} (x_n - y_n) &= \lim_{n \rightarrow \infty} x_n - \lim_{n \rightarrow \infty} y_n \\ \lim_{n \rightarrow \infty} (x_n \cdot y_n) &= \lim_{n \rightarrow \infty} x_n \cdot \lim_{n \rightarrow \infty} y_n \\ \lim_{n \rightarrow \infty} \left(\frac{x_n}{y_n}\right) &= \frac{\lim_{n \rightarrow \infty} x_n}{\lim_{n \rightarrow \infty} y_n} && \text{falls } \lim_{n \rightarrow \infty} y_n \neq 0 \\ \lim_{n \rightarrow \infty} x_n &= \left \lim_{n \rightarrow \infty} x_n \right \end{aligned}$
--

BEISPIEL III.2.2. (1) Es gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{n+1}{n} &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right) \\ &= \lim_{n \rightarrow \infty} 1 + \lim_{n \rightarrow \infty} \frac{1}{n} \end{aligned}$$

$$= 1.$$

(2) Aus

$$\lim_{n \rightarrow \infty} \frac{n+1}{n} = 1,$$

$$\lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n = 0$$

und

$$\lim_{n \rightarrow \infty} \frac{2n+1}{n^2} = \lim_{n \rightarrow \infty} \left(\frac{2}{n} + \frac{1}{n^2}\right) = 0$$

folgt

$$\lim_{n \rightarrow \infty} \left(\frac{2n+1}{n^2} + \left(\frac{1}{2}\right)^n \frac{n+1}{n}\right) = 0.$$

(3) Es gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{6n^3 + 2n^2 + 1}{7n^3 + 3n + 5} &= \lim_{n \rightarrow \infty} \frac{n^3(6 + \frac{2}{n} + \frac{1}{n^3})}{n^3(7 + \frac{3}{n^2} + \frac{5}{n^3})} \\ &= \lim_{n \rightarrow \infty} \frac{6 + \frac{2}{n} + \frac{1}{n^3}}{7 + \frac{3}{n^2} + \frac{5}{n^3}} \\ &= \frac{6}{7}. \end{aligned}$$

BEMERKUNG III.2.3. Aus der Konvergenz der Summenfolge $(x_n + y_n)_{n \in \mathbb{N}}$ folgt *nicht* die Konvergenz der Folgen $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$. Ebenso folgt aus der Konvergenz von $(|x_n|)_{n \in \mathbb{N}}$ *nicht* die Konvergenz von $(x_n)_{n \in \mathbb{N}}$. Als Gegenbeispiel betrachte man $x_n = (-1)^n$ und $y_n = (-1)^{n+1}$.

III.2.3. Konvergenzkriterien. 1. Eine konvergente Folge ist immer beschränkt. Mit anderen Worten: Ist eine Folge nicht beschränkt, kann sie nicht konvergieren.

BEISPIEL III.2.4. Die Folge $(q^n)_{n \in \mathbb{N}}$ mit $|q| > 1$ ist nicht konvergent. DENN: Wegen

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right) = 1 < |q|$$

gibt es ein $m \in \mathbb{N}^*$ mit $1 + \frac{1}{m} \leq |q|$. Für $n \in \mathbb{N}$ folgt dann aus der Bernoulli Ungleichung, Beispiel I.2.7 (S. 18),

$$|q|^n \geq \left(1 + \frac{1}{m}\right)^n \geq 1 + \frac{n}{m}.$$

Also ist die Folge $(q^n)_{n \in \mathbb{N}}$ nicht beschränkt und damit auch nicht konvergent.

BEMERKUNG III.2.5. Aus der Beschränktheit einer Folge folgt *nicht* ihre Konvergenz; Beispiel: $((-1)^n)_{n \in \mathbb{N}}$ ist beschränkt, aber nicht konvergent.

2. Jede Teilfolge einer konvergenten Folge ist konvergent.

BEISPIEL III.2.6. Aus

$$\lim_{n \rightarrow \infty} 2^{-n} = 0$$

folgt

$$\lim_{n \rightarrow \infty} 2^{-n^2} = 0.$$

3. Jede monotone und beschränkte Folge ist konvergent.

BEISPIEL III.2.7. Wir betrachten die Folge aus Beispiel III.1.1 (5)

$$\begin{aligned} x_0 &= 2 \\ x_{n+1} &= \frac{1}{2} \left(x_n + \frac{2}{x_n} \right) \end{aligned}$$

und behaupten, dass sie gegen $\sqrt{2}$ konvergiert. Der Beweis dieser Behauptung erfolgt in drei Schritten.

1. SCHRITT: Für alle n gilt $x_n \geq \sqrt{2}$.

Beweis durch Induktion:

Induktionsanfang $n = 0$: $x_0 = 2 > \sqrt{2}$.

Induktionsschritt $n \rightarrow n + 1$:

$$\begin{aligned} x_{n+1} &= \frac{1}{2} \left(x_n + \frac{2}{x_n} \right) \\ &= \frac{1}{2} \left(x_n - 2\sqrt{2} + \frac{2}{x_n} \right) + \sqrt{2} \\ &= \frac{1}{2} \left(\sqrt{x_n} - \frac{\sqrt{2}}{\sqrt{x_n}} \right)^2 + \sqrt{2} \\ &\geq \sqrt{2}. \end{aligned}$$

2. SCHRITT: Die Folge ist monoton fallend.

Denn

$$\begin{aligned} x_n - x_{n+1} &= x_n - \frac{1}{2} \left(x_n + \frac{2}{x_n} \right) \\ &= \frac{1}{2} \left(x_n - \frac{2}{x_n} \right) \\ &= \frac{x_n^2 - 2}{2x_n} \\ &\geq 0 \quad \text{wegen Schritt 1.} \end{aligned}$$

3. SCHRITT: Wegen Schritt 1 und Schritt 2 gilt $\sqrt{2} \leq x_n \leq 2$ für alle n . Die Folge ist also beschränkt und monoton und damit konvergent. Sei

$$X = \lim_{n \rightarrow \infty} x_n.$$

Wegen Schritt 1 ist $X \geq \sqrt{2} > 0$. Wegen der Rechenregeln für Limeses gilt

$$\begin{aligned} X &= \lim_{n \rightarrow \infty} x_n \\ &= \lim_{n \rightarrow \infty} x_{n+1} \\ &= \lim_{n \rightarrow \infty} \frac{1}{2} \left(x_n + \frac{2}{x_n} \right) \\ &= \frac{1}{2} \left(X + \frac{2}{X} \right). \end{aligned}$$

Auflösen nach X liefert $X^2 = 2$ und – wegen $X > 0$ – damit

$$\lim_{n \rightarrow \infty} x_n = \sqrt{2}.$$

4. Sind $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen mit $x_n \leq y_n$ für alle bis auf endlich viele n , so ist

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

BEMERKUNG III.2.8. Die Limesbildung erhält strikte Ungleichungen *nicht*; Beispiel: $x_n = 0$, $y_n = \frac{1}{n+1}$.

5. Zu der Folge $(x_n)_{n \in \mathbb{N}}$ gebe es zwei konvergente Folgen $(y_n)_{n \in \mathbb{N}}$ und $(z_n)_{n \in \mathbb{N}}$ mit:

- $y_n \leq x_n \leq z_n$ für alle bis auf endlich viele n und
- $\lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} z_n$.

Dann ist $(x_n)_{n \in \mathbb{N}}$ konvergent und

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n.$$

BEISPIEL III.2.9. Sei $|q| < 1$. Dann ist

$$\lim_{n \rightarrow \infty} q^n = 0.$$

Zum Beweis betrachte zunächst den Fall $0 \leq q < 1$. Dann gilt für alle $n \in \mathbb{N}$: $0 \leq q^n < 1$ und $q^{n+1} \leq q^n$. Also ist die Folge beschränkt und monoton und damit konvergent. Sei Q ihr Grenzwert. Wegen der Rechenregeln für Limeses folgt

$$Q^2 = \left(\lim_{n \rightarrow \infty} q^n \right)^2 = \lim_{n \rightarrow \infty} q^{2n} = Q.$$

Also ist $Q = 0$ oder $Q = 1$. Da die Folge monoton fallend und $0 \leq q < 1$ ist, ist der Fall $Q = 1$ ausgeschlossen. Also ist

$$\lim_{n \rightarrow \infty} q^n = 0.$$

Sei nun $-1 < q < 0$. Dann gilt für alle n : $-|q|^n \leq q^n \leq |q|^n$. Wegen

$$\lim_{n \rightarrow \infty} |q|^n = 0$$

folgt auch in diesem Fall

$$\lim_{n \rightarrow \infty} q^n = 0.$$

III.2.4. Die Exponentialfunktion. Wir behaupten, dass für jedes $x \in \mathbb{R}$ die Folge $((1 + \frac{x}{n})^n)_{n \in \mathbb{N}^*}$ konvergiert. Der Beweis erfolgt in drei Schritten.

1. SCHRITT: Für $x = 0$ handelt es sich offensichtlich um die konstante Folge 1, die natürlich konvergiert.

2. SCHRITT: Sei $x < 0$. Gemäß Beispiel III.1.4 (3) ist die Folge beschränkt. Weiter gibt es eine natürliche Zahl n_x mit $n_x > |x|$. Für alle $n \geq n_x$ gilt

$$\begin{aligned} \frac{(1 + \frac{x}{n+1})^{n+1}}{(1 + \frac{x}{n})^n} &= \frac{(n+1+x)^{n+1}}{(n+1)^{n+1}} \cdot \frac{n^n}{(n+x)^n} \\ &= \left[\frac{(n+1+x)n}{(n+1)(n+x)} \right]^{n+1} \frac{n+x}{n} \\ &= \left[\frac{(n+1)n + nx}{(n+1)n + (n+1)x} \right]^{n+1} \frac{n+x}{n} \\ &= \left[1 - \frac{x}{(n+1)n + (n+1)x} \right]^{n+1} \frac{n+x}{n}. \end{aligned}$$

Aus der Bernoullischen Ungleichung, Beispiel I.2.7 (S. 18), folgt

$$\begin{aligned} \left[1 - \frac{x}{(n+1)n + (n+1)x} \right]^{n+1} &\geq 1 - \frac{(n+1)x}{(n+1)n + (n+1)x} \\ &= 1 - \frac{x}{n+x} \\ &= \frac{n}{n+x}. \end{aligned}$$

Setzen wir dies in obige Gleichung ein, erhalten wir

$$\frac{(1 + \frac{x}{n+1})^{n+1}}{(1 + \frac{x}{n})^n} \geq 1 \quad \implies \quad \left(1 + \frac{x}{n+1}\right)^{n+1} \geq \left(1 + \frac{x}{n}\right)^n.$$

Also ist die Folge ab dem Index n_x monoton wachsend. Da die Folge bis auf endlich viele Indizes monoton und beschränkt ist, ist sie konvergent.

3. SCHRITT: Sei nun $x > 0$. Setze zur Abkürzung

$$\begin{aligned} x_n &= \left(1 + \frac{x}{n}\right)^n, \\ y_n &= \left(1 - \frac{x}{n}\right)^n, \\ z_n &= x_n y_n \end{aligned}$$

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

Wegen Schritt 2 ist die Folge $(y_n)_{n \in \mathbb{N}}$ konvergent. Aus der Bernoulli'schen Ungleichung, Beispiel I.2.7 (S. 18), folgt für alle $n \in \mathbb{N}^*$

$$1 \geq z_n \geq 1 - \frac{nx^2}{n^2} = 1 - \frac{x^2}{n}.$$

Hieraus folgt, dass die Folge $(z_n)_{n \in \mathbb{N}}$ gegen 1 konvergiert. Da wegen des 2. Schrittes der Grenzwert von $(y_n)_{n \in \mathbb{N}}$ größer als 0 ist, folgt die Konvergenz der Folge $(x_n)_{n \in \mathbb{N}} = \left(\frac{z_n}{y_n}\right)_{n \in \mathbb{N}}$.

Die Vorschrift

$$x \mapsto \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$$

definiert eine Funktion $\mathbb{R} \rightarrow \mathbb{R}$. Sie heißt EXPONENTIALFUNKTION und wird mit $\exp(x)$ bezeichnet:

$$\exp(x) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Die Zahl $e = \exp(1)$ heißt EULERSCHE ZAHL. Es ist $e \approx 2.71828$. Statt $\exp(x)$ schreibt man auch e^x . In späteren Abschnitten werden wir sehen, dass die Interpretation „Zahl e hoch x “ dieser Schreibweise in der Tat gerechtfertigt ist.

Aus den obigen Überlegungen können wir folgende Eigenschaften der Exponentialfunktion ablesen:

$$\begin{aligned} \exp(0) &= 1 \\ \exp(x) &> 0 \quad \text{für alle } x \in \mathbb{R} \\ \exp(x) &< 1 \quad \text{für alle } x < 0 \\ \exp(x) &> 1 \quad \text{für alle } x > 0 \\ \exp(x) \cdot \exp(-x) &= 1 \quad \text{für alle } x \in \mathbb{R} \\ x < y &\Rightarrow \exp(x) < \exp(y) \quad \text{MONOTONIE} \end{aligned}$$

III.2.5. Uneigentliche Grenzwerte. Die Folge $(x_n)_{n \in \mathbb{N}}$ hat den GRENZWERT ∞ bzw. $-\infty$, wenn es zu *jeder* Zahl $K > 0$ bzw. $K < 0$ einen Index n_K gibt mit $x_n > K$ bzw. $x_n < K$ für alle $n \geq n_K$. Wir schreiben in diesem Fall

$$\lim_{n \rightarrow \infty} x_n = \infty \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} x_n = -\infty.$$

BEISPIEL III.2.10. Sei $q > 1$. Dann ist

$$\lim_{n \rightarrow \infty} q^n = \infty.$$

III.3. Stetigkeit

III.3.1. Definition. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion. f heißt **STETIG IM PUNKT** $\bar{x} \in I$, wenn für *jede* konvergente Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in I$ für alle n und $\lim_{n \rightarrow \infty} x_n = \bar{x}$ gilt:
 $\lim_{n \rightarrow \infty} f(x_n) = f(\bar{x})$.

f heißt **STETIG AUF** I oder kurz **STETIG**, wenn f in *jedem* Punkt $\bar{x} \in I$ stetig ist.

BEISPIEL III.3.1. (1) Die Funktion $f(x) = x^2$ ist stetig auf \mathbb{R} .
DENN: Ist $\bar{x} \in \mathbb{R}$ und $(x_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit $\lim_{n \rightarrow \infty} x_n = \bar{x}$, so folgt aus Abschnitt III.2.2

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_n^2 = \left(\lim_{n \rightarrow \infty} x_n \right)^2 = \bar{x}^2 = f(\bar{x}).$$

(2) Die Funktion $f(x) = |x|$ ist stetig auf \mathbb{R} .

DENN: Konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen \bar{x} , so konvergiert $(|x_n|)_{n \in \mathbb{N}}$ gemäß Abschnitt III.2.2 gegen $|\bar{x}|$.

(3) Die Funktion $f(x) = \sqrt{x}$ ist auf $[0, \infty)$ stetig.

DENN: Betrachte zunächst den Punkt $\bar{x} = 0$. Sei $(x_n)_{n \in \mathbb{N}}$ eine beliebige Nullfolge mit Gliedern in $[0, \infty)$. Zu jedem $\varepsilon > 0$ gibt es dann ein $n_\varepsilon \in \mathbb{N}$ mit $0 \leq x_n \leq \varepsilon^2$ für alle $n \geq n_\varepsilon$. Für diese n gilt dann $0 \leq \sqrt{x_n} \leq \varepsilon$. Also ist $\lim_{n \rightarrow \infty} \sqrt{x_n} = 0$ und f somit in 0 stetig.

Sei nun $\bar{x} > 0$ und $(x_n)_{n \in \mathbb{N}}$ eine Folge mit Gliedern in $[0, \infty)$, die gegen \bar{x} konvergiert. Für jedes $n \in \mathbb{N}$ gilt dann

$$\begin{aligned} 0 &\leq |\sqrt{x_n} - \sqrt{\bar{x}}| \\ &= \left| \frac{(\sqrt{x_n} - \sqrt{\bar{x}})(\sqrt{x_n} + \sqrt{\bar{x}})}{\sqrt{x_n} + \sqrt{\bar{x}}} \right| \\ &= \frac{|x_n - \bar{x}|}{\sqrt{x_n} + \sqrt{\bar{x}}} \\ &\leq \frac{|x_n - \bar{x}|}{\sqrt{\bar{x}}} \\ &\xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Also ist $\lim_{n \rightarrow \infty} \sqrt{x_n} = \sqrt{\bar{x}}$ und damit f in \bar{x} stetig.

III.3.2. Rechenregeln. Es gilt:

Die Summe, die Differenz und das Produkt stetiger Funktionen sind stetig. Der Quotient stetiger Funktionen ist in allen Punkten stetig, in denen der Nenner nicht Null ist. Die Komposition stetiger Funktionen ist ebenfalls stetig.

BEISPIEL III.3.2. Aus der Stetigkeit der konstanten Funktion 1 und der Stetigkeit der Funktionen aus Beispiel III.3.1 folgt z.B.:

$$\begin{aligned}
 |x| + x^2 & \text{ ist stetig auf } \mathbb{R}, \\
 (1 + x^2)\sqrt{x} & \text{ ist stetig auf } [0, \infty) \\
 \frac{1}{x} & \text{ ist stetig auf } \mathbb{R} \setminus \{0\} \\
 \frac{x^4 - x^2 + 1}{x^2 + 1} & \text{ ist stetig auf } \mathbb{R} \text{ (Nenner ist nie 0)} \\
 \sqrt{|x|} & \text{ ist stetig auf } \mathbb{R} \\
 \sqrt{\frac{x^4 + x^2 + 2}{2x^2 + 3}} & \text{ ist stetig auf } \mathbb{R} \\
 & \text{(Radikand ist stets positiv und Nenner ist nie 0)}
 \end{aligned}$$

III.3.3. Eigenschaften stetiger Funktionen. Eine erste wichtige Eigenschaft stetiger Funktionen ist:

Jede auf einem beschränkten, abgeschlossenen Intervall $[a, b]$ stetige Funktion f ist beschränkt, d.h. es gibt eine Zahl $K \geq 0$ mit $|f(x)| \leq K$ für alle $x \in [a, b]$.

BEMERKUNG III.3.3. Diese Aussage gilt nicht für offene, halboffene oder unbeschränkte Intervalle. Beispiel: $\frac{1}{x}$ ist stetig auf $(0, 1]$, aber unbeschränkt; x^2 ist stetig auf \mathbb{R} , aber unbeschränkt.

Die folgende Eigenschaft ist eine Verschärfung der ersten:

Jede auf einem beschränkten, abgeschlossenen Intervall $[a, b]$ stetige Funktion f nimmt dort ihr Maximum und Minimum an, d.h. es gibt Zahlen x_m und x_M in $[a, b]$ mit

$$f(x_m) \leq f(x) \leq f(x_M) \quad \text{für alle } x \in [a, b].$$

BEMERKUNG III.3.4. Diese Aussage gilt nicht für offene, halboffene oder unbeschränkte Intervalle. Beispiel: x^2 ist stetig auf $(0, 1)$. Es ist $\inf\{x^2 : x \in (0, 1)\} = 0$, $\sup\{x^2 : x \in (0, 1)\} = 1$, aber für alle $x \in (0, 1)$ gilt $0 < x^2 < 1$.

Die folgende Eigenschaft ist für Anwendungen besonders wichtig:

ZWISCHENWERTSATZ: Eine auf einem Intervall I stetige Funktion f nimmt jeden Wert zwischen zwei beliebigen Funktionswerten an, d.h. sind $x, y \in I$ mit $f(x) < f(y)$

und u mit $f(x) \leq u \leq f(y)$ beliebig, so gibt es ein $z \in I$ mit $f(z) = u$.

Eine Anwendung des Zwischenwertsatzes ist:

Die Funktion f sei auf dem Intervall I stetig. Es gebe zwei Punkte $a, b \in I$ mit $f(a)f(b) \leq 0$. Dann besitzt f eine Nullstelle in I , d.h. es gibt ein $x \in I$ mit $f(x) = 0$.

BEISPIEL III.3.5. Betrachte das Polynom $p(x) = 4x^4 - 5x^2 - 10$. Wegen der Rechenregeln für stetige Funktionen ist p stetig. Es ist $p(0) = -10 < 0$ und $p(2) = 34 > 0$. Also hat p eine Nullstelle zwischen 0 und 2. Weiter ist $p(1) = -11 < 0$. Also hat p eine Nullstelle zwischen 1 und 2. Durch fortgesetztes Halbieren des Intervalles können wir die Nullstelle beliebig genau einschließen.

Es gilt folgende Verschärfung der Stetigkeit:

Jede auf einem beschränkten, abgeschlossenen Intervall $[a, b]$ stetige Funktion f ist GLEICHMÄSSIG STETIG, d.h. zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, das nur von ε abhängt, so dass für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.

BEMERKUNG III.3.6. Diese Aussage gilt nicht für offene, halboffene oder unbeschränkte Intervalle.

III.3.4. Einseitige Grenzwerte. Seien $a, c \in \mathbb{R} \cup \{-\infty, \infty\}$. Die Funktion f hat in a den LINKSSEITIGEN GRENZWERT c , kurz

$$\lim_{x \rightarrow a^-} f(x) = c,$$

wenn es ein $b \in \mathbb{R}$ mit $b < a$ gibt, so dass f auf (b, a) definiert ist und für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit Gliedern in (b, a) und Grenzwert a gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = c.$$

Analog hat f in a den RECHTSSEITIGEN GRENZWERT c , kurz

$$\lim_{x \rightarrow a^+} f(x) = c$$

wenn es ein $b \in \mathbb{R}$ mit $b > a$ gibt, so dass f auf (a, b) definiert ist und für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit Gliedern in (a, b) und Grenzwert a gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = c.$$

Ist $a = -\infty$, ist nur der rechtsseitige Grenzwert definiert, und wir lassen das Symbol $+$ fort. Analog ist für $a = \infty$ nur der linksseitige Grenzwert definiert, und wir lassen das Symbol $-$ fort.

BEISPIEL III.3.7. (1) Es ist

$$\begin{aligned}\lim_{x \rightarrow 0^+} \frac{1}{x} &= \infty, \\ \lim_{x \rightarrow 0^-} \frac{1}{x} &= -\infty, \\ \lim_{x \rightarrow \infty} \frac{1}{x} &= 0, \\ \lim_{x \rightarrow -\infty} \frac{1}{x} &= 0.\end{aligned}$$

(2) Die Vorzeichenfunktion $\operatorname{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ ist definiert durch

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{falls } x > 0, \\ 0 & \text{falls } x = 0, \\ -1 & \text{falls } x < 0. \end{cases}$$

Sie ist auf $\mathbb{R} \setminus \{0\}$ stetig. Es ist

$$\begin{aligned}\lim_{x \rightarrow 0^+} \operatorname{sgn}(x) &= 1, \\ \lim_{x \rightarrow 0^-} \operatorname{sgn}(x) &= -1.\end{aligned}$$

(3) Die Entier-Funktion $[\cdot] : \mathbb{R} \rightarrow \mathbb{R}$ ist definiert durch

$$[x] = \max\{n \in \mathbb{Z} : n \leq x\}.$$

Sie ist auf $\mathbb{R} \setminus \mathbb{Z}$ stetig. Für $a \in \mathbb{Z}$ ist

$$\begin{aligned}\lim_{x \rightarrow a^+} [x] &= a, \\ \lim_{x \rightarrow a^-} [x] &= a - 1.\end{aligned}$$

Es besteht folgender Zusammenhang zwischen einseitigen Grenzwerten und Stetigkeit:

Die Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann im Punkt a des Intervalls I stetig, wenn die links- und rechtsseitigen Grenzwerte von f in a existieren, endlich sind und übereinstimmen.

III.3.5. Polynome. Ein POLYNOM ist eine Funktion der Form

$$x \mapsto a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

mit $n \in \mathbb{N}$ und $a_0, \dots, a_n \in \mathbb{R}$. Das Polynom hat den GRAD n , falls $a_n \neq 0$ ist. Aus den Rechenregeln für stetige Funktionen folgt:

Jedes Polynom ist stetig auf \mathbb{R} .

Ist $u \in \mathbb{R}$ eine Nullstelle des Polynoms p , so kann p durch den Linearfaktor $x - u$ „dividiert“ werden:

Sei $u \in \mathbb{R}$ eine Nullstelle des Polynoms

$$p = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$

Dann gibt es ein Polynom

$$q = b_{n-1} x^{n-1} + \dots + b_1 x + b_0$$

mit

$$p(x) = (x - u)q(x) \quad \text{für alle } x \in \mathbb{R}.$$

Für die Koeffizienten der beiden Polynome gilt die Beziehung

$$b_{n-1} = a_n$$

$$b_k = a_{k+1} + u b_{k+1} \quad k = n-2, \dots, 0.$$

Bei der Rechnung per Hand wertet man obige Rekursionsformel wie beim „Treppenrechnen“ der schriftlichen Division von Zahlen aus.

BEISPIEL III.3.8. Durch Probieren finden wir, dass das Polynom $x^4 - 3x^3 + 3x^2 - 3x + 2$ die Nullstelle $x = 1$ hat. Also können wir es durch $x - 1$ dividieren und erhalten

$$\begin{array}{r} x^4 - 3x^3 + 3x^2 - 3x + 2 : x - 1 = x^3 - 2x^2 + x - 2. \\ \underline{x^4 - x^3} \\ -2x^3 + 3x^2 \\ \underline{-2x^3 + 2x^2} \\ x^2 - 3x \\ \underline{x^2 - x} \\ -2x + 2 \\ \underline{-2x + 2} \\ 0 \end{array}$$

Durch weiteres Probieren finden wir, dass das Polynom $x^3 - 2x^2 + x - 2$ die Nullstelle $x = 2$ hat. Also können wir es durch $x - 2$ dividieren und erhalten

$$\begin{array}{r} x^3 - 2x^2 + x - 2 : x - 2 = x^2 + 1. \\ \underline{x^3 - 2x^2} \\ x - 2 \\ \underline{x - 2} \\ 0 \end{array}$$

Insgesamt ergibt sich somit

$$x^4 - 3x^3 + 3x^2 - 3x + 2 = (x - 1)(x - 2)(x^2 + 1).$$

Wie das Beispiel $x^2 + 1$ zeigt, besitzt nicht jedes Polynom eine Nullstelle in \mathbb{R} . Es gilt jedoch:

Jedes Polynom mit ungeradem Grad besitzt mindestens eine Nullstelle in \mathbb{R} .

Um dies einzusehen, betrachte ein Polynom

$$p = \sum_{\ell=0}^{2k+1} a_{\ell} x^{\ell}$$

mit ungeradem Grad, d.h. $a_{2k+1} \neq 0$. Aus den Rechenregeln für Limes folgt

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{-2k-1} p(n) &= \lim_{n \rightarrow \infty} \left(a_{2k+1} + a_{2k} \frac{1}{n} + \dots + a_0 \frac{1}{n^{2k+1}} \right) \\ &= a_{2k+1}. \end{aligned}$$

Also gibt es ein $a > 0$ mit $\operatorname{sgn} p(a) = \operatorname{sgn} a_{2k+1}$. Ebenso folgt

$$\begin{aligned} \lim_{n \rightarrow \infty} (-n)^{-2k-1} p(-n) &= \lim_{n \rightarrow \infty} \left(a_{2k+1} - a_{2k} \frac{1}{n} + \dots + a_0 \frac{(-1)^{2k+1}}{n^{2k+1}} \right) \\ &= a_{2k+1}. \end{aligned}$$

Wegen $(-n)^{-2k-1} < 0$ gibt es daher ein $b < 0$ mit $\operatorname{sgn} p(b) = -\operatorname{sgn} a_{2k+1}$. Also ist $p(a)p(b) < 0$, und die Behauptung folgt aus dem Zwischenwertsatz.

Betrachtet man auch Nullstellen in \mathbb{C} , so gilt:

Jedes Polynom p vom Grad $n \geq 1$ kann in komplexe Linearfaktoren zerlegt werden, d.h. es gibt ein $r \in \mathbb{N}^*$, komplexe Zahlen $\lambda_1, \dots, \lambda_r$ und positive natürliche Zahlen m_1, \dots, m_r mit $m_1 + \dots + m_r = n$ und

$$p(x) = a_n (x - \lambda_1)^{m_1} \cdot \dots \cdot (x - \lambda_r)^{m_r} \quad \text{für alle } x \in \mathbb{R}.$$

Die Zahl m_i heißt die **VIELFACHHEIT** der Nullstelle λ_i . Ist $\lambda \in \mathbb{C}$ eine Nullstelle von p mit $\operatorname{Im} \lambda \neq 0$, so ist $\bar{\lambda}$ auch eine Nullstelle von p ; λ und $\bar{\lambda}$ haben die gleichen Vielfachheiten.

III.3.6. Rationale Funktionen. Eine **RATIONALE FUNKTION** hat die Form $\frac{p}{q}$ mit teilerfremden Polynomen p und q . p bzw. q heißen das **ZÄHLERPOLYNOM** bzw. das **NENNERPOLYNOM**. Aus den Rechenregeln für stetige Funktionen folgt:

Eine rationale Funktion ist stetig auf $\mathbb{R} \setminus \{u_1, \dots, u_s\}$, wobei u_1, \dots, u_s die reellen Nullstellen des Nennerpolynoms sind (sofern vorhanden).

BEMERKUNG III.3.9. Die reellen Nullstellen des Nennerpolynoms einer rationalen Funktion heißen auch **POLE** oder **POLSTELLEN** der Funktion.

BEISPIEL III.3.10. (1) $f(x) = \frac{x+1}{x^2+1}$ ist eine rationale Funktion. Das Zählerpolynom ist $x+1$; das Nennerpolynom ist x^2+1 . Da das Nennerpolynom keine reelle Nullstelle hat, ist f auf ganz \mathbb{R} stetig.

(2) $g(x) = \frac{x^2+5x-3}{x^3-2x^2-x+2}$ ist eine rationale Funktion. Das Nennerpolynom x^3-2x^2-x+2 hat gemäß Beispiel III.3.8 die Nullstellen -1 , 1 und 2 . Daher ist g auf $\mathbb{R} \setminus \{-1, 1, 2\}$ stetig.

III.3.7. Trigonometrische Funktionen. In Abschnitt I.3.3 haben wir die trigonometrischen Funktionen \sin und \cos aufgrund geometrischer Überlegungen eingeführt. Aus diesen folgt, dass alle Nullstellen von \sin bzw. \cos von der Form $k\pi$ bzw. $(2k+1)\frac{\pi}{2}$ mit $k \in \mathbb{Z}$. Daher können wir die trigonometrischen Funktionen $\tan = \frac{\sin}{\cos}$, genannt **TANGENS**, auf $\mathbb{R} \setminus \{(2k+1)\frac{\pi}{2} : k \in \mathbb{Z}\}$ und $\cot = \frac{\cos}{\sin}$, genannt **COTANGENS**, auf $\mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$ definieren.

Die Funktionen \sin , \cos , \tan und \cot sind auf ihren Definitionsbereichen stetig.

Der **BEWEIS** erfolgt in vier Schritten.

1. **SCHRITT:** Der Sinus ist stetig in 0.

Aus Abschnitt I.3.3 folgt für alle $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$

$$|\sin(x)| = \sin(|x|).$$

Da die Sekante kürzer ist als die Länge des zugehörigen Kreisbogens, gilt für diese x zusätzlich

$$\sin(|x|) \leq |x|.$$

Daher gilt für jede Nullfolge $(x_n)_{n \in \mathbb{N}}$

$$0 \leq |\sin(x_n)| \leq |x_n| \xrightarrow[n \rightarrow \infty]{} 0 \implies \sin(x_n) \xrightarrow[n \rightarrow \infty]{} 0.$$

Also ist der Sinus in 0 stetig.

2. **SCHRITT:** Der Cosinus ist stetig in 0.

Für $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ ist $\cos(x) = \sqrt{1 - \sin^2(x)}$. Damit folgt die Stetigkeit des Cosinus in 0 aus Schritt 1 und den Rechenregeln für stetige Funktionen.

3. **SCHRITT:** Sinus und Cosinus sind stetig auf \mathbb{R} .

Betrachte einen beliebigen Punkt $x \in \mathbb{R}$ und irgendeine Folge $(x_n)_{n \in \mathbb{N}}$ mit Grenzwert x . Dann ist $(x_n - x)_{n \in \mathbb{N}}$ eine Nullfolge. Aus den Additionstheoremen, Beispiel I.6.3 (S. 44), und den Schritten 1 und 2 folgt

$$\begin{aligned} \cos(x_n) &= \cos(x + (x_n - x)) \\ &= \cos(x) \cos(x_n - x) - \sin(x) \sin(x_n - x) \end{aligned}$$

$$\begin{aligned}
& \xrightarrow{n \rightarrow \infty} \cos(x) \cdot 1 - \sin(x) \cdot 0 \\
& = \cos(x) \\
\sin(x_n) &= \sin(x + (x_n - x)) \\
&= \sin(x) \cos(x_n - x) + \cos(x) \sin(x_n - x) \\
& \xrightarrow{n \rightarrow \infty} \sin(x) \cdot 1 + \cos(x) \cdot 0 \\
& = \sin(x).
\end{aligned}$$

Dies beweist die behauptete Stetigkeit von Cosinus und Sinus.

4. SCHRITT: Die Stetigkeit von Tangens und Cotangens folgt aus der Stetigkeit von Sinus und Cosinus und den Rechenregeln für stetige Funktionen.

III.3.8. Exponential- und Logarithmusfunktion. Wir erinnern an die Definition der Exponentialfunktion aus Abschnitt III.2.4

$$\exp(x) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Es gilt:

Die Exponentialfunktion ist stetig auf \mathbb{R} .

Der BEWEIS erfolgt in drei Schritten.

1. SCHRITT: $\exp(x+y) \leq \exp(x)\exp(y)$, $\exp(x-y) \geq \exp(x)\exp(-y)$ für alle $x \geq 0$, $y \geq 0$.

Für alle $n \in \mathbb{N}^*$ ist

$$\begin{aligned}
\left(1 + \frac{x}{n}\right)^n \left(1 + \frac{y}{n}\right)^n &= \left(1 + \frac{x}{n} + \frac{y}{n} + \frac{xy}{n}\right)^n \\
&\geq \left(1 + \frac{x+y}{n}\right)^n \\
\left(1 + \frac{x}{n}\right)^n \left(1 - \frac{y}{n}\right)^n &= \left(1 + \frac{x}{n} - \frac{y}{n} - \frac{xy}{n}\right)^n \\
&\leq \left(1 + \frac{x-y}{n}\right)^n.
\end{aligned}$$

Hieraus folgt die Behauptung durch Grenzübergang $n \rightarrow \infty$.

2. SCHRITT: $\lim_{n \rightarrow \infty} \exp\left(-\frac{1}{n+1}\right) = \lim_{n \rightarrow \infty} \exp\left(\frac{1}{n+1}\right) = 1$.

Aus Abschnitt III.2.4 folgt, dass die Folge $\left(\left(1 + \frac{a}{n}\right)^n\right)_{n \in \mathbb{N}}$ für $a \geq -1$ monoton wachsend ist. Daher gilt für solche a : $\exp(a) \geq 1 + a$. Damit folgt mit der Monotonie der Exponentialfunktion

$$1 = \exp(0) \geq \exp\left(-\frac{1}{n+1}\right) \geq 1 - \frac{1}{n+1} \xrightarrow{n \rightarrow \infty} 1.$$

Dies beweist die erste Grenzwertaussage. Die zweite folgt aus der ersten und der Beziehung $\exp(x)\exp(-x) = 1$ aus Abschnitt III.2.4.

3. SCHRITT: Die Exponentialfunktion ist stetig auf \mathbb{R} .

Sei dazu $x \in \mathbb{R}$ beliebig. Wegen der Monotonie und Positivität der Exponentialfunktion muss man zum Nachweis der Stetigkeit in x nur zeigen, dass

$$\lim_{n \rightarrow \infty} \exp\left(x + \frac{1}{n+1}\right) = \lim_{n \rightarrow \infty} \exp\left(x - \frac{1}{n+1}\right) = \exp(x)$$

ist. Dies folgt aber sofort aus den ersten beiden Schritten und den Rechenregeln für Limes.

Weiter gilt:

$$\exp : \mathbb{R} \rightarrow (0, \infty) \text{ ist bijektiv.}$$

Zum BEWEIS dieser Aussage beachten wir, dass die Exponentialfunktion wegen ihrer Monotonie (vgl. Abschnitt III.2.4) injektiv ist.

Wir müssen also nur noch die Surjektivität beweisen, d.h. dass es zu jedem $a > 0$ ein $x \in \mathbb{R}$ gibt mit $\exp(x) = a$.

Für $a = 1$ leistet $x = 0$ das Gewünschte.

Sei nun $a > 1$. Dann ist

$$\exp(0) = 1 < a < a + 1 \leq \exp(a).$$

Wegen des Zwischenwertsatzes gibt es also ein $x \in (0, a)$ mit $\exp(x) = a$.

Sei schließlich $0 < a < 1$. Dann ist $\frac{1}{a} > 1$, und es gibt ein $x \in \mathbb{R}$ mit $\exp(x) = \frac{1}{a}$. Dann ist $\exp(-x) = a$.

Da die Exponentialfunktion bijektiv ist, besitzt sie eine Umkehrfunktion. Diese heißt LOGARITHMUSFUNKTION und wird mit \ln bezeichnet. Es ist $\ln : (0, \infty) \rightarrow \mathbb{R}$ bijektiv. Wegen der Monotonie der Exponentialfunktion ist die Logarithmusfunktion auch monoton:

$$0 < x < y \implies \ln(x) < \ln(y).$$

Weitere Eigenschaften der Logarithmusfunktion werden wir in Abschnitt IV.4.2 kennen lernen.

KAPITEL IV

Differentiation

IV.1. Die Ableitung einer differenzierbaren Funktion

IV.1.1. Definition der Ableitung. Sei $I \subset \mathbb{R}$ ein offenes Intervall, $f : I \rightarrow \mathbb{R}$ eine Funktion und $\bar{x} \in I$. Die Funktion f heißt im Punkt \bar{x} DIFFERENZIERBAR, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x})}{h}$$

existiert. In diesem Fall heißt er die ABLEITUNG von f in \bar{x} und wird mit $f'(\bar{x})$ bezeichnet:

$$f'(\bar{x}) = \lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x})}{h}.$$

Mit anderen Worten: Die Funktion f ist in \bar{x} differenzierbar, wenn es eine Zahl c , die die Ableitung von f in \bar{x} genannt und mit $f'(\bar{x})$ bezeichnet wird, gibt, sodass für *jede* Nullfolge $(h_n)_{n \in \mathbb{N}}$ gilt

$$\lim_{n \rightarrow \infty} \frac{f(\bar{x} + h_n) - f(\bar{x})}{h_n} = c.$$

Ist f in jedem Punkt von I differenzierbar, so heißt f auf I DIFFERENZIERBAR. In diesem Fall ist $x \mapsto f'(x)$ eine auf I definierte Funktion, die man die Ableitung von f nennt und die man mit f' bezeichnet.

Andere Schreibweisen für $f'(x)$ sind $\frac{df(x)}{dx}$ und $\frac{d}{dx}f(x)$. Den Übergang von f zu f' nennt man ABLEITEN oder DIFFERENZIEREN.

BEISPIEL IV.1.1. (1) Die konstante Funktion $f(x) = c$ ist differenzierbar mit $f'(x) = 0$.

(2) $f(x) = ax + b$ ist differenzierbar mit $f'(x) = a$.

(3) $f(x) = x^n$ mit $n \in \mathbb{N}^*$ ist differenzierbar mit $f'(x) = nx^{n-1}$.

(4) $f(x) = \sqrt{x}$ ist differenzierbar mit $f'(x) = \frac{1}{2\sqrt{x}}$ für $x > 0$.

BEWEIS. Sei jeweils x beliebig. Dann gilt

$$(1) \quad \frac{f(x+h) - f(x)}{h} = \frac{c - c}{h} = 0 \quad \text{für alle } h \neq 0.$$

$$(2) \quad \frac{f(x+h) - f(x)}{h} = \frac{a(x+h) + b - ax - b}{h}$$

$$\begin{aligned}
&= \frac{ah}{h} \\
&= a \quad \text{für alle } h \neq 0. \\
(3) \quad \frac{f(x+h) - f(x)}{h} &= \frac{(x+h)^n - x^n}{h} \\
&= \frac{1}{h} \left[nx^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \dots + \binom{n}{n}h^n \right] \\
&= nx^{n-1} + \binom{n}{2}x^{n-2}h + \dots + \binom{n}{n}h^{n-1} \\
&\xrightarrow{h \rightarrow 0} nx^{n-1}. \\
(4) \quad \frac{f(x+h) - f(x)}{h} &= \frac{\sqrt{x+h} - \sqrt{x}}{h} \\
&= \frac{(\sqrt{x+h} - \sqrt{x})(\sqrt{x+h} + \sqrt{x})}{h(\sqrt{x+h} + \sqrt{x})} \\
&= \frac{x+h-x}{h(\sqrt{x+h} + \sqrt{x})} \\
&= \frac{1}{\sqrt{x+h} + \sqrt{x}} \\
&\xrightarrow{h \rightarrow 0} \frac{1}{2\sqrt{x}}.
\end{aligned}$$

□

IV.1.2. Deutungen der Ableitung. In einem kartesischen Koordinatensystem betrachten wir die Kurve $y = f(x)$ in der Nähe des Punktes $(\bar{x}, f(\bar{x}))$. Die Sekante durch die Punkte $(\bar{x}, f(\bar{x}))$ und $(\bar{x} + h, f(\bar{x} + h))$ mit $h \neq 0$ hat die Steigung $\frac{f(\bar{x}+h) - f(\bar{x})}{h}$. Für $h \rightarrow 0$ strebt diese Steigung gegen die Steigung der Tangente an die Kurve $y = f(x)$ im Punkt $(\bar{x}, f(\bar{x}))$ (sofern der Grenzwert existiert). Dies liefert die folgende GEOMETRISCHE DEUTUNG der Ableitung:

Die TANGENTE an den Graphen $y = f(x)$ im Punkt $(\bar{x}, f(\bar{x}))$ hat die Gleichung

$$y = f(\bar{x}) + (x - \bar{x})f'(\bar{x}).$$

Wir suchen nun nach einer Geraden $g(x) = m(x - \bar{x}) + f(\bar{x})$, die die Funktion f in der Nähe von \bar{x} im folgenden Sinne best möglich approximieren soll: Der relative Fehler $\frac{g(x) - f(x)}{x - \bar{x}}$ soll für $x \rightarrow \bar{x}$ verschwinden. Einsetzen der Gleichung für g liefert $m = f'(\bar{x})$. Dies ergibt die folgende ANALYTISCHE DEUTUNG der Ableitung:

$f(x) \approx f(\bar{x}) + (x - \bar{x})f'(\bar{x})$ ist die BESTE LINEARE APPROXIMATION an f in der Nähe von \bar{x} .

BEISPIEL IV.1.2. Für $f(x) = \sqrt{x}$ ergibt sich in der Nähe von $\bar{x} > 0$

$$\sqrt{x} \approx \sqrt{\bar{x}} + \frac{1}{2\sqrt{\bar{x}}}(x - \bar{x}).$$

Für $x = 2$ und $\bar{x} = 1.96$ erhalten wir z.B.

$$\begin{aligned} \sqrt{2} &\approx \sqrt{1.96} + \frac{1}{2\sqrt{1.96}}(2 - 1.96) \\ &= 1.4 + \frac{0.04}{1.4} \\ &= 1.4142857\dots \end{aligned}$$

Diese Näherung für $\sqrt{2}$ ist auf vier Stellen genau, d.h. der relative Fehler ist kleiner als 0.01%.

Sei $s = s(t)$ die von einem Massenpunkt bei geradliniger Bewegung bis zur Zeit t zurückgelegte Strecke. Der Differenzenquotient

$$\frac{\Delta s}{\Delta t} = \frac{s(t + \Delta t) - s(t)}{\Delta t}$$

gibt den Betrag der mittleren Geschwindigkeit im Zeitintervall $[t, t + \Delta t]$ an. Im Grenzübergang $\Delta t \rightarrow 0$ erhält man den Betrag der momentanen Geschwindigkeit zur Zeit t . Dies liefert die folgende PHYSIKALISCHE DEUTUNG der Ableitung:

$s'(t)$ ist der Betrag der GESCHWINDIGKEIT der Bewegung zur Zeit t .

BEMERKUNG IV.1.3. Für die Geschwindigkeit benutzt man häufig auch die Bezeichnung $\dot{s}(t)$ statt $s'(t)$.

IV.1.3. Stetigkeit ist notwendig für Differenzierbarkeit. Ist f im Punkt \bar{x} differenzierbar, so ist

$$\lim_{h \rightarrow 0} [f(\bar{x} + h) - f(\bar{x}) - hf'(\bar{x})] = 0.$$

Wegen

$$\lim_{h \rightarrow 0} hf'(\bar{x}) = 0$$

folgt

$$f(\bar{x}) = \lim_{h \rightarrow 0} f(\bar{x} + h).$$

Also ist f in \bar{x} stetig:

f in \bar{x} differenzierbar $\implies f$ in \bar{x} stetig

BEMERKUNG IV.1.4. Die Umkehrung dieser Aussage ist FALSCH.

BEISPIEL IV.1.5. Betrachte die Funktion $f(x) = |x|$. Sie ist auf \mathbb{R} stetig. Es ist

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{f(\frac{1}{n}) - f(0)}{\frac{1}{n}} &= \lim_{n \rightarrow \infty} \frac{|\frac{1}{n}|}{\frac{1}{n}} \\ &= 1 \\ \lim_{n \rightarrow \infty} \frac{f(-\frac{1}{n}) - f(0)}{-\frac{1}{n}} &= \lim_{n \rightarrow \infty} \frac{|-\frac{1}{n}|}{-\frac{1}{n}} \\ &= -1.\end{aligned}$$

Also ist f in 0 nicht differenzierbar.

IV.1.4. Differentiationsregeln. Die Funktionen f und g seien auf dem Intervall I differenzierbar; c sei eine reelle Zahl. Es gelten folgende Regeln:

$$\begin{aligned}[f(x) + g(x)]' &= f'(x) + g'(x) \\ [cf(x)]' &= cf'(x) \\ [f(x)g(x)]' &= f'(x)g(x) + f(x)g'(x) \quad (\text{PRODUKTREGEL}) \\ \left[\frac{f(x)}{g(x)}\right]' &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \quad \text{falls } g(x) \neq 0 \\ & \quad \quad \quad (\text{QUOTIENTENREGEL}) \\ \left[\frac{1}{g(x)}\right]' &= -\frac{g'(x)}{g(x)^2} \quad \text{falls } g(x) \neq 0.\end{aligned}$$

IV.1.5. Differentiation von Polynomen und rationalen Funktionen. Jedes Polynom ist differenzierbar, die Ableitung ist wieder ein Polynom:

$$\begin{aligned}p(x) &= a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \\ \implies p'(x) &= n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + a_1.\end{aligned}$$

Jede rationale Funktion ist auf ihrem Definitionsbereich differenzierbar, die Differentiation erfolgt mit der Quotientenregel, die Ableitung ist wieder eine rationale Funktion.

BEISPIEL IV.1.6. Für das Polynom

$$p(x) = 4x^5 + 3x^3 - 2x^2 + 1$$

erhalten wir

$$p'(x) = 20x^4 + 9x^2 - 4x.$$

Für die rationale Funktion

$$r(x) = \frac{2x^3 + x - 1}{x^2 + 1}$$

liefert die Quotientenregel

$$\begin{aligned} r'(x) &= \frac{(6x^2 + 1)(x^2 + 1) - (2x^3 + x - 1)2x}{(x^2 + 1)^2} \\ &= \frac{4x^4 + 5x^2 + 2x + 1}{x^4 + 2x^2 + 1}. \end{aligned}$$

IV.1.6. Differentiation der trigonometrischen Funktionen.

Die trigonometrischen Funktionen sind auf ihrem Definitionsbereich differenzierbar. Es gilt:

$$\begin{aligned} \sin'(x) &= \cos(x) \\ \cos'(x) &= -\sin(x) \\ \tan'(x) &= \frac{1}{\cos^2(x)} \quad \text{für } x \neq (2k+1)\frac{\pi}{2}, k \in \mathbb{Z} \\ \cot'(x) &= -\frac{1}{\sin^2(x)} \quad \text{für } x \neq k\pi, k \in \mathbb{Z}. \end{aligned}$$

BEISPIEL IV.1.7. Aus den Rechenregeln für Ableitungen folgt

$$\begin{aligned} [(x^3 + 5 \sin(x)) \cos(x)]' &= (3x^2 + 5 \cos(x)) \cos(x) \\ &\quad + (x^3 + 5 \sin(x))(-\sin(x)) \\ &= 3x^2 \cos(x) + 5 \cos^2(x) \\ &\quad - x^3 \sin(x) - 5 \sin^2(x). \end{aligned}$$

IV.1.7. Kettenregel. Die Komposition zweier differenzierbarer Funktionen ist ebenfalls differenzierbar und es gilt die KETTENREGEL:

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

BEISPIEL IV.1.8. (1) Für

$$h(x) = (x^3 + 2x^2 + 1)^3$$

gilt $h = f \circ g$ mit

$$f(x) = x^3$$

$$g(x) = x^3 + 2x^2 + 1.$$

Damit folgt

$$\begin{aligned} f'(x) &= 3x^2 \\ g'(x) &= 3x^2 + 4x \\ h'(x) &= 3(g(x))^2 g'(x) \\ &= 3(x^3 + 2x^2 + 1)^2(3x^2 + 4x). \end{aligned}$$

(2) Für

$$h(x) = \cos(5x^2 + 2)$$

gilt $h = f \circ g$ mit

$$\begin{aligned} f(x) &= \cos(x) \\ g(x) &= 5x^2 + 2. \end{aligned}$$

Damit folgt

$$\begin{aligned} f'(x) &= -\sin(x) \\ g'(x) &= 10x \\ h'(x) &= -\sin(g(x)) g'(x) \\ &= -\sin(5x^2 + 2)10x. \end{aligned}$$

(3) Die Funktion

$$h(x) = [\sin((x^4 + 2x)^2)]^5$$

kann als Komposition von 4 Funktionen geschrieben werden:

$$h = f_4 \circ f_3 \circ f_2 \circ f_1$$

mit

$$\begin{aligned} f_1(x) &= x^4 + 2x \\ f_2(x) &= x^2 \\ f_3(x) &= \sin(x) \\ f_4(x) &= x^5. \end{aligned}$$

Für die Differentiation „bröseln“ wir diese Verschachtelung auf und setzen

$$\begin{aligned} g_1 &= f_2 \circ f_1 \\ g_2 &= f_3 \circ g_1 \end{aligned}$$

Dann ist

$$g_2 = f_3 \circ f_2 \circ f_1$$

und

$$h = f_4 \circ g_2.$$

Damit folgt aus der Kettenregel:

$$\begin{aligned}
 h'(x) &= f'_4(g_2(x))g'_2(x) & \left| \begin{array}{l} f'_4(x) = 5x^4 \\ g_2(x) = \sin((x^4 + 2x)^2) \end{array} \right. \\
 &= 5 [g_2(x)]^4 g'_2(x) \\
 &= 5 [\sin((x^4 + 2x)^2)]^4 g'_2(x), \\
 g'_2(x) &= f'_3(g_1(x))g'_1(x) & \left| \begin{array}{l} f'_3(x) = \cos(x) \\ g_1(x) = (x^4 + 2x)^2 \end{array} \right. \\
 &= \cos(g_1(x))g'_1(x) \\
 &= \cos((x^4 + 2x)^2)g'_1(x), \\
 g'_1(x) &= f'_2(f_1(x))f'_1(x) & \left| \begin{array}{l} f'_2(x) = 2x \\ f_1(x) = x^4 + 2x \end{array} \right. \\
 &= 2 (f_1(x)) f'_1(x) & \left| \begin{array}{l} f'_1(x) = 4x^3 + 2 \end{array} \right. \\
 &= 2(x^4 + 2x)f'_1(x) \\
 &= 2(x^4 + 2x)(4x^3 + 2).
 \end{aligned}$$

Setzen wir alle Teilschritte zusammen, erhalten wir schließlich

$$h'(x) = 5 [\sin((x^4 + 2x)^2)]^4 \cos((x^4 + 2x)^2) 2(x^4 + 2x)(4x^3 + 2).$$

IV.1.8. Höhere Ableitungen. Die Ableitung der Ableitung von f bezeichnen wir, falls sie existiert, mit f'' , und für $\frac{d}{dx}(\frac{d}{dx}f(x))$ schreiben wir $\frac{d^2}{dx^2}f(x)$.

Allgemein definieren wir die HÖHEREN ABLEITUNGEN rekursiv:

$$\begin{array}{l}
 f^{(0)}(x) = f(x) \\
 f^{(1)}(x) = f'(x) \\
 f^{(n)}(x) = \frac{d}{dx}f^{(n-1)}(x) \quad , n \geq 2.
 \end{array}$$

BEISPIEL IV.1.9. Für

$$p(x) = 3x^6 + 2x^4 - 7x + 1$$

erhalten wir

$$\begin{aligned}
 p'(x) &= 18x^5 + 8x^3 - 7 \\
 p''(x) &= 90x^4 + 24x^2 \\
 p^{(3)}(x) &= 360x^3 + 48x \\
 p^{(4)}(x) &= 1080x^2 + 48 \\
 p^{(5)}(x) &= 2160x
 \end{aligned}$$

$$p^{(6)}(x) = 2160$$

$$p^{(k)}(x) = 0 \quad \text{für alle } k \geq 7.$$

Allgemein gilt:

Die n -te Ableitung eines Polynoms vom Grad k mit $k < n$ ist Null.

BEISPIEL IV.1.10. Ein gemäß

$$x(t) = A \cos(\omega t + \alpha)$$

schwingender Oszillator hat die Geschwindigkeit

$$\begin{aligned} \dot{x}(t) &= x'(t) \\ &= -A\omega \sin(\omega t + \alpha) \end{aligned}$$

und die Beschleunigung

$$\begin{aligned} \ddot{x}(t) &= x''(t) \\ &= -A\omega^2 \cos(\omega t + \alpha) \\ &= -\omega^2 x(t). \end{aligned}$$

IV.2. Anwendungen der Differentiation

IV.2.1. Maxima und Minima einer Funktion. Die Funktion f sei auf dem Intervall $I \subset \mathbb{R}$ definiert. Sie hat in dem Punkt $\bar{x} \in I$ ein GLOBALES oder ABSOLUTES MAXIMUM, wenn für alle $x \in I$ gilt $f(x) \leq f(\bar{x})$. Sie hat in \bar{x} ein LOKALES MAXIMUM, wenn es ein Intervall J gibt mit $J \subset I$ und $\bar{x} \in J$, sodass die Einschränkung von f auf J in \bar{x} ein globales Maximum hat.

Analog hat f in \bar{x} ein GLOBALES MINIMUM bzw. LOKALES MINIMUM, wenn für alle $x \in I$ bzw. für alle $x \in J$ gilt $f(x) \geq f(\bar{x})$.

Jedes Minimum oder Maximum heißt auch ein EXTREMUM.

BEISPIEL IV.2.1. (1) Die Funktion $f(x) = (1 - x^2)^2$ hat auf dem Intervall $[-1, 1]$ in $\bar{x} = 0$ ein globales Maximum.

DENN: Es ist $f(0) = 1$ und für alle $x \in [-1, 1]$ gilt

$$\begin{aligned} x \in [-1, 1] &\implies |x| \leq 1 \\ &\implies x^2 \leq 1 \\ &\implies 0 \leq 1 - x^2 \leq 1 \\ &\implies f(x) \leq 1. \end{aligned}$$

(2) Die Funktion $f(x) = (1 - x^2)^2$ hat auf dem Intervall $[-2, 2]$ in $\bar{x} = 0$ ein lokales Maximum. Dieses ist aber wegen $f(2) = 9 > 1 = f(0)$ kein globales Maximum.

Die Funktion f habe in dem inneren Punkt \bar{x} von I ein lokales Maximum und sei in \bar{x} differenzierbar. Für hinreichend kleines $h > 0$ gilt dann

$$\frac{f(\bar{x} + h) - f(\bar{x})}{h} \leq 0 \quad \Longrightarrow \quad \lim_{h \rightarrow 0^+} \frac{f(\bar{x} + h) - f(\bar{x})}{h} \leq 0.$$

Analog gilt für hinreichend kleines $h < 0$

$$\frac{f(\bar{x} + h) - f(\bar{x})}{h} \geq 0 \quad \Longrightarrow \quad \lim_{h \rightarrow 0^-} \frac{f(\bar{x} + h) - f(\bar{x})}{h} \geq 0.$$

Da f in \bar{x} differenzierbar ist, stimmen die beiden einseitigen Grenzwerte überein, und wir erhalten $f'(\bar{x}) = 0$. Ganz analog können wir argumentieren, wenn f in \bar{x} ein lokales Minimum hat; es ändern sich nur die Vorzeichen in den jeweiligen Ungleichungen. Insgesamt erhalten wir das folgende Kriterium:

\bar{x} innerer Punkt, f in \bar{x} differenzierbar und \bar{x} lokales Extremum von $f \implies f'(\bar{x}) = 0$.

BEMERKUNG IV.2.2. Die Umkehrung dieser Aussage gilt nicht. Beispiel: $f(x) = x^3$ hat in $\bar{x} = 0$ kein Extremum, aber es ist $f'(0) = 0$

Aus obigem Kriterium folgt:

Die Kandidaten für Extremalstellen von $f : I \rightarrow \mathbb{R}$ sind:

- die Endpunkte des Intervalls I ,
- die Punkte aus I , in denen f nicht differenzierbar ist, und
- die inneren Punkte $x \in I$ mit $f'(x) = 0$ (STATIONÄRE PUNKTE).

Werden sämtliche Extrema von $f : I \rightarrow \mathbb{R}$ gesucht, müssen alle diese Kandidaten bestimmt und die entsprechenden Funktionswerte berechnet werden. Der kleinste bzw. größte dieser Funktionswerte liefert dann das globale Minimum bzw. Maximum.

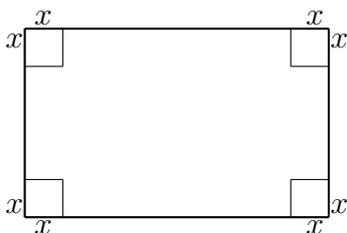


ABBILDUNG IV.2.1. Blechwanne

BEISPIEL IV.2.3. Aus einer rechtwinkligen Blechplatte mit Kantenlängen 16cm und 10cm soll eine quaderförmige, oben offene Wanne mit maximalem Volumen geformt werden. Dies geschieht, indem an den vier Ecken der Blechplatte jeweils ein Quadrat der Kantenlänge x ausgeschnitten wird und die entstehenden rechteckigen Stücke hochgebogen werden (vgl. Abbildung IV.2.1). Offensichtlich ist nur der Bereich $0 \leq x \leq 5$ (in cm) sinnvoll. Das Volumen des entstehenden Quaders ist

$$V(x) = (10 - 2x)(16 - 2x)x \text{ cm}^3.$$

Da V differenzierbar ist und $V(0) = V(5) = 0$ gilt, muss das maximale Volumen an einem stationären Punkt angenommen werden. (Beachte: Wegen Abschnitt III.3.3 nimmt V sein Maximum an.) Wir erhalten folgende Bestimmungsgleichung für einen stationären Punkt:

$$\begin{aligned} 0 = V'(x) &= -2(16 - 2x)x - 2(10 - 2x)x + (10 - 2x)(16 - 2x) \\ &= 12x^2 - 104x + 160 \\ &= 12 \left(x^2 - \frac{26}{3}x + \frac{40}{3} \right) \\ &= 12 \left(\left(x - \frac{13}{3} \right)^2 - \frac{169}{9} + \frac{40}{3} \right) \\ &= 12 \left(\left(x - \frac{13}{3} \right)^2 - \frac{49}{9} \right) \\ &= 12 \left(x - \frac{20}{3} \right) (x - 2). \end{aligned}$$

Also ist $\bar{x} = 2$ der einzige stationäre Punkt im Intervall $[0, 5]$. Das maximale Volumen ist

$$V(2) = 144 \text{ cm}^3.$$

IV.2.2. Der Mittelwertsatz. Die Funktion f sei auf $[a, b]$ stetig und auf (a, b) differenzierbar. Die Funktion

$$F(x) = f(x) - (x - b) \frac{f(b) - f(a)}{b - a}$$

ist dann auf $[a, b]$ stetig und nimmt gemäß Abschnitt III.3.3 dort ihr Maximum und Minimum an. Da $F(a) = f(b) = F(b)$ und F auf (a, b) differenzierbar ist, muss F einen stationären Punkt $\bar{x} \in (a, b)$ besitzen. Für diesen Punkt gilt

$$0 = F'(\bar{x}) = f'(\bar{x}) - \frac{f(b) - f(a)}{b - a}.$$

Dies beweist:

MITTELWERTSATZ: Die Funktion f sei auf $[a, b]$ stetig und auf (a, b) differenzierbar. Dann gibt es ein $\bar{x} \in (a, b)$ mit

$$f'(\bar{x}) = \frac{f(b) - f(a)}{b - a}.$$

Mit anderen Worten:

In mindestens einem Punkt hat die Tangente an die Kurve $y = f(x)$ die gleiche Steigung wie die Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$.

Ist speziell $f(a) = f(b)$ erhalten wir:

SATZ VON ROLLE: Ist f auf $[a, b]$ stetig und auf (a, b) differenzierbar und ist $f(a) = f(b)$, so gibt es ein $\bar{x} \in (a, b)$ mit $f'(\bar{x}) = 0$.

Eine wichtige Konsequenz des Mittelwertsatzes ist:

$$f'(x) = 0 \text{ f\"ur alle } x \in (a, b) \iff f \text{ ist konstant}$$

bzw. f\"ur differenzierbare Funktionen f und g :

$$f'(x) = g'(x) \text{ f\"ur alle } x \in (a, b) \iff \text{es gibt eine Konstante } c \text{ mit } f(x) = g(x) + c \text{ f\"ur alle } x \in (a, b)$$

BEISPIEL IV.2.4. Wir behaupten, dass jede lineare Pendelbewegung, f\"ur die das Hookesche Gesetz gilt, eine harmonische Schwingung ist. Um dies einzusehen, bezeichnen wir mit $s(t)$ die Auslenkung zur Zeit t . Dann ist $\ddot{s}(t)$ die Beschleunigung zur Zeit t . Wegen des Hookeschen Gesetzes gilt

$$\ddot{s}(t) = -\omega^2 s(t) \quad \text{bzw.} \quad \ddot{s}(t) + \omega^2 s(t) = 0$$

mit einer Konstanten ω . Multiplizieren wir die rechte Gleichung mit $\dot{s}(t)$, erhalten wir wegen der Produktregel f\"ur die Differentiation

$$\begin{aligned} 0 &= \dot{s}(t)\ddot{s}(t) + \omega^2 \dot{s}(t)s(t) \\ &= \frac{1}{2} \frac{d}{dt} [\dot{s}(t)^2 + \omega^2 s(t)^2]. \end{aligned}$$

Also gibt es eine Konstante c mit

$$\dot{s}(t)^2 + \omega^2 s(t)^2 = c \quad \text{f\"ur alle } t.$$

Insbesondere ist $c = \dot{s}(0)^2 + \omega^2 s(0)^2$. Ist also $\dot{s}(0) = s(0) = 0$, so folgt

$$\dot{s}(t)^2 + \omega^2 s(t)^2 = 0 \quad \text{für alle } t$$

und somit

$$\dot{s}(t) = s(t) = 0 \quad \text{für alle } t.$$

Seien also $\dot{s}(0)$ und $s(0)$ nicht beide gleich Null. Betrachte die Funktion

$$s_0(t) = s(t) - s(0) \cos(\omega t) - \frac{1}{\omega} \dot{s}(0) \sin(\omega t).$$

Dann ist

$$\begin{aligned} \dot{s}_0(t) &= \dot{s}(t) + \omega s(0) \sin(\omega t) - \dot{s}(0) \cos(\omega t) \\ \ddot{s}_0(t) &= \ddot{s}(t) + \omega^2 s(0) \cos(\omega t) + \omega \dot{s}(0) \sin(\omega t) \\ &= -\omega^2 s_0(t) \end{aligned}$$

und

$$s_0(0) = \dot{s}_0(0) = 0.$$

Also ist $s_0(t) = 0$ für alle t und damit

$$s(t) = s(0) \cos(\omega t) + \frac{1}{\omega} \dot{s}(0) \sin(\omega t) \quad \text{für alle } t.$$

Dies beweist die Behauptung.

Eine weitere Anwendung des Mittelwertsatzes ist:

EXTREMWERTEST: Die Funktion f sei auf (a, b) zweimal differenzierbar und habe in $\bar{x} \in (a, b)$ einen stationären Punkt. Dann gilt

$$\begin{aligned} f''(\bar{x}) < 0 &\implies f \text{ hat in } \bar{x} \text{ ein lokales Maximum,} \\ f''(\bar{x}) > 0 &\implies f \text{ hat in } \bar{x} \text{ ein lokales Minimum.} \end{aligned}$$

IV.2.3. Wendepunkte. An dem Vorzeichen der zweiten Ableitung f'' kann man das Krümmungsverhalten der Kurve $y = f(x)$ ablesen:

$$\begin{aligned} f'' > 0 &\implies \text{ Die Kurve } y = f(x) \text{ ist von unten konvex} \\ &\qquad\qquad\qquad \text{(Linkskrümmung).} \\ f'' < 0 &\implies \text{ Die Kurve } y = f(x) \text{ ist von oben konvex} \\ &\qquad\qquad\qquad \text{(Rechtsskrümmung).} \end{aligned}$$

Diejenigen Punkte, in denen die Kurve $y = f(x)$ ihr Krümmungsverhalten ändert, nennt man **WENDEPUNKTE** von f .

Kandidaten für Wendepunkte von $f : I \rightarrow \mathbb{R}$ sind

- die Punkte, in denen f'' nicht existiert, und
- die Punkte, in denen f'' verschwindet.

Dies führt auf folgenden WENDEPUNKTTTEST:

$$f''(\bar{x}) = 0 \text{ und } f'''(\bar{x}) \neq 0 \implies f \text{ hat in } \bar{x} \text{ einen Wendepunkt.}$$

BEISPIEL IV.2.5. Die kritische Temperatur T_0 , oberhalb derer man ein Gas nicht mehr verflüssigen kann, wird näherungsweise mit der Zustandsgleichung nach van der Waals

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT$$

so bestimmt, dass die Kurve

$$p = p(V) = \frac{RT_0}{V - b} - \frac{a}{V^2}$$

einen Wendepunkt mit horizontaler Tangente hat. Aus

$$p'(V) = p''(V) = 0$$

ergeben sich die Bedingungen

$$\begin{aligned} -\frac{RT_0}{(V - b)^2} + \frac{2a}{V^3} &= 0 \\ \frac{2RT_0}{(V - b)^3} - \frac{6a}{V^4} &= 0. \end{aligned}$$

Multiplizieren wir die zweite Gleichung mit $\frac{V}{3}$ und addieren das Ergebnis zur ersten Gleichung, erhalten wir

$$\begin{aligned} 0 &= -\frac{RT_0}{(V - b)^2} + \frac{2RT_0V}{3(V - b)^3} \\ &= \frac{-3RT_0(V - b) + 2RT_0V}{3(V - b)^3} \\ &= \frac{-RT_0V + 3RT_0b}{3(V - b)^3}. \end{aligned}$$

Hieraus ergibt sich das kritische Volumen zu $V_0 = 3b$. Einsetzen in die erste Gleichung ergibt die kritische Temperatur

$$\begin{aligned} T_0 &= \frac{2a}{(3b)^3} \cdot \frac{(3b - b)^2}{R} \\ &= \frac{8a}{27bR}. \end{aligned}$$

Da

$$\begin{aligned} p'''(V_0) &= -\frac{6RT_0}{(V_0 - b)^4} + \frac{24a}{V_0^5} \\ &= -\frac{48a}{27b} \cdot \frac{1}{16b^4} + \frac{24a}{243b^5} \\ &= -\frac{a}{81b^5} \end{aligned}$$

ist, handelt es sich tatsächlich um einen Wendepunkt.

IV.2.4. Die Regeln von de l'Hôpital. Wir wollen Grenzwerte der Form $\lim_{x \rightarrow \bar{x}} \frac{f(x)}{g(x)}$ für den Fall bestimmen, dass $f(\bar{x}) = g(\bar{x}) = 0$ oder $\lim_{x \rightarrow \bar{x}} f(x) = \lim_{x \rightarrow \bar{x}} g(x) = \pm\infty$ ist. Hierfür benötigen wir folgende Konsequenz des Mittelwertsatzes:

Die Funktionen f und g seien auf $[a, b]$ stetig und auf (a, b) differenzierbar. Für alle $x \in (a, b)$ sei $g'(x) \neq 0$. Dann gibt es ein $\bar{x} \in (a, b)$ mit

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

BEWEIS. Wegen $g'(x) \neq 0$ für alle $x \in (a, b)$ folgt aus dem Satz von Rolle $g(b) \neq g(a)$. Daher ist die Funktion

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

auf $[a, b]$ stetig und auf (a, b) differenzierbar. Wegen

$$F(b) = \frac{f(a)g(b) - f(b)g(a)}{g(b) - g(a)} = F(a)$$

gibt es nach dem Satz von Rolle ein $\bar{x} \in (a, b)$ mit $F'(\bar{x}) = 0$. Wegen

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(x)$$

folgt hieraus die Behauptung. □

Halten wir b fest und lassen a gegen b streben, strebt \bar{x} von links gegen b . Dies führt auf:

REGELN VON DE L'HÔPITAL: Die Funktionen f und g seien auf $[a, b]$ stetig und auf (a, b) differenzierbar. Für alle $x \in (a, b)$ sei $g'(x) \neq 0$. Weiter gelte:

- $f(x) \rightarrow 0$ und $g(x) \rightarrow 0$ für $x \rightarrow b^-$ bzw. $f(x) \rightarrow \pm\infty$ und $g(x) \rightarrow \pm\infty$ für $x \rightarrow b^-$,

- $\frac{f'(x)}{g'(x)} \rightarrow L$ für $x \rightarrow b-$ für ein $L \in \mathbb{R} \cup \{-\infty, \infty\}$.

Dann gilt

$$\lim_{x \rightarrow b-} \frac{f(x)}{g(x)} = \lim_{x \rightarrow b-} \frac{f'(x)}{g'(x)}.$$

Analoge Aussagen gelten für die Grenzwerte $x \rightarrow a+$, $x \rightarrow -\infty$ und $x \rightarrow \infty$.

BEISPIEL IV.2.6. Betrachte

$$f(x) = x^3 - x^2 - 5x - 3$$

$$g(x) = 3x^2 - 7x - 6.$$

Es ist

$$f(3) = g(3) = 0$$

und

$$g'(3) = 6 \cdot 3 - 7 = 11 \neq 0.$$

Daher liefert die Regel von de l'Hôpital

$$\begin{aligned} \lim_{x \rightarrow 3} \frac{x^3 - x^2 - 5x - 3}{3x^2 - 7x - 6} &= \lim_{x \rightarrow 3} \frac{3x^2 - 2x - 5}{6x - 7} \\ &= \frac{16}{11}. \end{aligned}$$

BEISPIEL IV.2.7. Die Funktion

$$\frac{1}{x} - \frac{1}{\sin(x)}$$

ist auf $(-\frac{\pi}{2}, \frac{\pi}{2}) \setminus \{0\}$ definiert und stetig. Wie verhält sie sich für $x \rightarrow 0$?
Für $x \neq 0$ ist

$$\frac{1}{x} - \frac{1}{\sin(x)} = \frac{\sin(x) - x}{x \sin(x)}.$$

Setze

$$f(x) = \sin(x) - x$$

$$g(x) = x \sin(x).$$

Beide Funktionen sind differenzierbar und haben eine Nullstelle in 0.
Weiter ist

$$g'(x) = \sin(x) + x \cos(x)$$

in der Nähe von Null ungleich Null. Also liefert die Regel von de l'Hôpital

$$\begin{aligned} \lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{\sin(x)} \right) &= \lim_{x \rightarrow 0} \frac{\sin(x) - x}{x \sin(x)} \\ &= \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x) + x \cos(x)}. \end{aligned}$$

Aber $\cos(x) - 1$ und $\sin(x) + x \cos(x)$ verschwinden beide für $x = 0$. Wir müssen die Regel von de l'Hôpital also nochmals anwenden. Dies liefert

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x) + x \cos(x)} &= \lim_{x \rightarrow 0} \frac{-\sin(x)}{2 \cos(x) - x \sin(x)} \\ &= \frac{0}{2} \\ &= 0. \end{aligned}$$

Also ist

$$\lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{\sin(x)} \right) = 0.$$

IV.2.5. Die Fixpunktiteration. Wir betrachten eine auf dem Intervall $[a, b]$ stetige Funktion f und suchen eine Lösung der Gleichung

$$x = f(x).$$

Jede Lösung dieser Gleichung heißt ein **FIXPUNKT** von f .

Die Funktion f habe zusätzlich folgende Eigenschaften:

- $a \leq f(x) \leq b$ für alle $x \in [a, b]$,
- f ist auf (a, b) differenzierbar und

$$K = \sup_{a < x < b} |f'(x)| < 1.$$

Dann gilt:

- f besitzt genau einen Fixpunkt \bar{x} in $[a, b]$.
- Für jeden Startwert $x_0 \in [a, b]$ konvergiert die Folge $(x_n)_{n \in \mathbb{N}}$ mit

$$x_{n+1} = f(x_n) \quad n = 0, 1, \dots$$

(FIXPUNKTITERATION)

gegen \bar{x} .

- Es gelten die Fehlerabschätzungen

$$|x_n - \bar{x}| \leq \frac{K^n}{1 - K} |x_1 - x_0|$$

(A PRIORI),

$$|x_n - \bar{x}| \leq \frac{K}{1 - K} |x_n - x_{n-1}|$$

(A POSTERIORI).

BEWEIS. Betrachte die Funktion

$$g(x) = f(x) - x.$$

Sie ist stetig, und wegen $a \leq f(x) \leq b$ für alle $x \in [a, b]$ gilt $g(a) \geq 0$, $g(b) \leq 0$. Nach dem Zwischenwertsatz hat g eine Nullstelle \bar{x} in $[a, b]$. Diese ist konstruktionsgemäß ein Fixpunkt von f .

Wir nehmen nun an, f habe zwei verschiedene Fixpunkte y und z in $[a, b]$. Nach dem Mittelwertsatz gibt es dann ein $u \in (a, b)$ mit

$$f'(u) = \frac{f(y) - f(z)}{y - z} = \frac{y - z}{y - z} = 1.$$

Dies ist ein Widerspruch zur zweiten Eigenschaft von f . Also hat f genau einen Fixpunkt.

Betrachte nun die Folge $(x_n)_{n \in \mathbb{N}}$ mit einem beliebigen $x_0 \in [a, b]$. Wegen des Mittelwertsatzes gibt es für jedes $n \in \mathbb{N}^*$ ein $y_n \in (a, b)$ mit

$$\underbrace{x_n}_{=f(x_{n-1})} - \underbrace{\bar{x}}_{=f(\bar{x})} = f(x_{n-1}) - f(\bar{x}) = f'(y_n)(x_{n-1} - \bar{x}).$$

Wegen der zweiten Eigenschaft von f folgt

$$|x_n - \bar{x}| \leq K|x_{n-1} - \bar{x}| \leq K^2|x_{n-2} - \bar{x}| \leq \dots \leq K^n|x_0 - \bar{x}|.$$

Da $K < 1$ ist, konvergiert $(K^n)_{n \in \mathbb{N}}$ gegen Null. Also konvergiert die Folge $(x_n)_{n \in \mathbb{N}}$ gegen den Fixpunkt \bar{x} .

Aus der soeben bewiesenen Abschätzung folgt mit der Dreiecksungleichung

$$|x_n - \bar{x}| \leq K|x_{n-1} - \bar{x}| \leq K|x_n - \bar{x}| + K|x_n - x_{n-1}|.$$

Wegen $K < 1$ liefert dies

$$|x_n - \bar{x}| \leq \frac{K}{1 - K}|x_n - x_{n-1}|.$$

Dies beweist die a posteriori Fehlerabschätzung.

Wegen des Mittelwertsatzes gibt es schließlich zu jedem $n \geq 2$ ein $z_n \in (a, b)$ mit

$$\underbrace{x_n}_{=f(x_{n-1})} - \underbrace{x_{n-1}}_{=f(x_{n-2})} = f(x_{n-1}) - f(x_{n-2}) = f'(z_n)(x_{n-1} - x_{n-2}).$$

Also ist

$$|x_n - x_{n-1}| \leq K|x_{n-1} - x_{n-2}| \leq \dots \leq K^{n-1}|x_1 - x_0|.$$

Setzen wir dies in die a posteriori Fehlerabschätzung ein, erhalten wir schließlich die a priori Fehlerabschätzung. \square

BEMERKUNG IV.2.8. Aus dem obigen Beweis folgt, dass die a posteriori Fehlerabschätzung genauer ist als die a priori Fehlerabschätzung. Andererseits erlaubt die a priori Fehlerabschätzung schon vor der Berechnung der Folgenglieder eine Aussage, wie viele Schritte höchstens benötigt werden, um den Fixpunkt mit einer vorgegebenen Genauigkeit zu berechnen.

BEISPIEL IV.2.9. Wegen $\cos(0) = 1$ und $0 < \cos(\frac{1}{2}) < 1$ hat die Funktion

$$f(x) = \cos\left(\frac{1}{2}x\right)$$

einen Fixpunkt im Intervall $[0, 1]$. Sie ist differenzierbar, und für alle $x \in [0, 1]$ gilt

$$|f'(x)| = \frac{1}{2} \sin\left(\frac{1}{2}x\right) \leq \frac{1}{2} \sin\left(\frac{1}{2}\right) \leq \frac{1}{2}.$$

Also sind die Voraussetzungen mit $K = \frac{1}{2}$ erfüllt. Für den Startwert $x_0 = 0.8$ erhalten wir die Ergebnisse von Tabelle IV.2.1. Insbesondere ist der relative Fehler von x_5 kleiner als 0.07 %.

TABELLE IV.2.1. Fixpunktiteration für $f(x) = \cos(\frac{1}{2}x)$

n	x_n	$ x_n - \bar{x} \leq$
0	0.800000	
1	0.895817	0.09600
2	0.901355	0.00560
3	0.900152	0.00120
4	0.900414	0.00030
5	0.900357	0.00006

IV.2.6. Das Newtonverfahren. Das Newtonverfahren und seine Varianten sind die verbreitetsten und effizientesten Verfahren zur Nullstellenberechnung bei differenzierbaren Funktionen.

Zur Beschreibung des Verfahrens nehmen wir an, dass wir eine Näherung x_0 für die gesuchte Nullstelle \bar{x} der Funktion f geraten haben. Dann wird f in der Nähe von x_0 durch die Tangente an die Kurve $y = f(x)$ im Punkt $(x_0, f(x_0))$ approximiert. Der Schnittpunkt x_1 dieser Tangente mit der x -Achse sollte daher eine passable Näherung für die gesuchte Nullstelle sein. (Dies gilt natürlich nur, sofern die Tangente nicht waagrecht ist!) Die Gleichung der Tangente ist

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Also lautet die Bestimmungsgleichung für x_1 :

$$0 = f(x_0) + f'(x_0)(x_1 - x_0) \implies x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Wenn wir diese Vorschrift iterieren erhalten wir das

NEWTONVERFAHREN:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad n = 0, 1, \dots$$

Die Vorgehensweise des Newtonverfahrens wird in Abbildung IV.2.2 verdeutlicht.

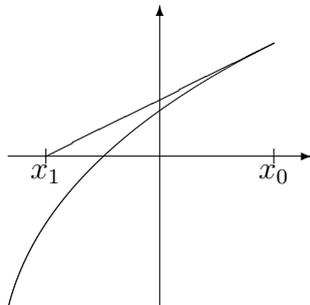


ABBILDUNG IV.2.2. Geometrische Interpretation des Newtonverfahrens

Das Newtonverfahren bricht natürlich zusammen, wenn die Tangente waagrecht ist, d.h. wenn $f'(x_i) = 0$ wird für ein i . Bei ungünstiger Wahl des Startwertes kann es auch divergieren, d.h. $|x_i| \rightarrow \infty$ für $i \rightarrow \infty$, oder in einen KESSEL geraten, d.h. es gibt ein $m \in \mathbb{N}^*$ mit $x_i = x_{i+m}$ für alle i .

BEISPIEL IV.2.10. Betrachte das Polynom $f(x) = x^4 - 3x^2 - 2$. Dann lautet das Newtonverfahren mit $x_0 = 1$,

$$\begin{aligned} x_1 &= 1 - \frac{-4}{-2} \\ &= -1, \\ x_2 &= -1 - \frac{-4}{2} \\ &= 1 \\ &= x_0. \end{aligned}$$

Das Newtonverfahren gerät also in einen Kessel.

BEISPIEL IV.2.11. Das Newtonverfahren angewandt auf die Funktion $f(x) = \arctan(x)$ liefert für den Startwert $x_0 = 2$ die in Tabelle IV.2.2 angegebenen Ergebnisse. Offensichtlich divergiert das Verfahren.

TABELLE IV.2.2. Newtonverfahren für $f(x) = \arctan(x)$

n	x_n	$f(x_n)$
0	2.0	1.107148
1	-3.535743	-1.295169
2	13.950959	1.499239
3	-279.344066	-1.567216
4	122016.998918	1.570788

Ist aber der Startwert hinreichend nahe bei der Nullstelle, konvergiert das Newtonverfahren QUADRATISCH, d.h. es gibt eine Konstante c mit

$$|x_{n+1} - \bar{x}| \leq c|x_n - \bar{x}|^2$$

für alle hinreichend großen n .

Numerisch äußert sich die quadratische Konvergenz in einer Verdoppelung der korrekten Nachkommastellen mit jedem Iterationsschritt.

Zum Nachweis der quadratischen Konvergenz nutzen wir $f(\bar{x}) = 0$ aus und wenden den Mittelwertsatz zweimal an: Zunächst erhalten wir aus der Iterationsvorschrift

$$\begin{aligned} x_{n+1} - \bar{x} &= x_n - \frac{f(x_n)}{f'(x_n)} - \bar{x} \\ &= x_n - \bar{x} - \frac{f(x_n) - f(\bar{x})}{f'(x_n)} \quad \text{wegen } f(\bar{x}) = 0 \\ &= \frac{1}{f'(x_n)} [f'(x_n)(x_n - \bar{x}) - f(x_n) + f(\bar{x})]. \end{aligned}$$

Nach dem Mittelwertsatz gibt es ein y_n zwischen x_n und \bar{x} mit $f(x_n) - f(\bar{x}) = f'(y_n)(x_n - \bar{x})$. Dies liefert

$$\begin{aligned} x_{n+1} - \bar{x} &= \frac{1}{f'(x_n)} [f'(x_n)(x_n - \bar{x}) - f'(y_n)(x_n - \bar{x})] \\ &= \frac{1}{f'(x_n)} [f'(x_n) - f'(y_n)](x_n - \bar{x}). \end{aligned}$$

Wiederum nach dem Mittelwertsatz gibt es ein z_n zwischen x_n und y_n mit $f'(x_n) = f'(y_n) + f''(z_n)(x_n - y_n)$ (Beachte: Wir setzen hier voraus, dass f zweimal differenzierbar ist!) Wegen $|x_n - y_n| \leq |x_n - \bar{x}|$ erhalten wir somit

$$\begin{aligned} |x_{n+1} - \bar{x}| &= \left| \frac{f''(z_n)}{f'(x_n)} (x_n - y_n)(x_n - \bar{x}) \right| \\ &\leq \frac{|f''(z_n)|}{|f'(x_n)|} |x_n - \bar{x}|^2. \end{aligned}$$

Dies beweist die quadratische Konvergenz mit

$$c = \frac{\sup_{a < x < b} |f''(x)|}{\inf_{a < x < b} |f'(x)|},$$

wenn das Intervall $[a, b]$ so gewählt ist, dass die Nullstelle \bar{x} und alle Folgenglieder x_n in ihm liegen.

BEISPIEL IV.2.12. Wir betrachten das Polynom

$$p(x) = x^3 + x^2 + 2x + 1.$$

Wegen $p(-1) = -1$ und $p(0) = 1$ hat es eine Nullstelle im Intervall $(-1, 0)$. Für die Ableitung gilt in $(-1, 0)$

$$p'(x) = 3x^2 + 2x + 2 > 3x^2 > 0.$$

Also ist das Newtonverfahren durchführbar. Für den Startwert $x_0 = -0.5$ erhalten wir die in Tabelle IV.2.3 angegebenen Ergebnisse.

TABELLE IV.2.3. Newtonverfahren für $f(x) = x^3 + x^2 + 2x + 1$

i	x_i
0	-0.500000
1	-0.571429
2	-0.569841
3	-0.569840

BEISPIEL IV.2.13. Seien $k \geq 2$ und $a > 0$. Dann ist $\sqrt[k]{a}$ die Nullstelle des Polynoms $x^k - a$. Die Iterationsvorschrift des Newtonverfahrens lautet in diesem Fall

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^k - a}{kx_n^{k-1}} \\ &= \frac{1}{k} \left[(k-1)x_n + \frac{a}{x_n^{k-1}} \right]. \end{aligned}$$

Dieses Verfahren, das die Berechnung von Wurzeln mit den Grundrechenarten erlaubt, war schon den alten Griechen bekannt und wird als VERFAHREN VON HERON bezeichnet. Für $k = 2$, $a = 2$ und $x_0 = 2$ erhalten wir z.B. die in Tabelle IV.2.4 angegebenen Näherungswerte für $\sqrt{2}$.

TABELLE IV.2.4. Newtonverfahren zur Berechnung von $\sqrt{2}$

i	x_i	$f(x_i)$
0	2	2
1	1.5	0.25
2	1.41 $\bar{6}$	0.0069 $\bar{4}$
3	1.41421568	0.00000601

BEISPIEL IV.2.14. Betrachte die Funktion $\frac{1}{x} - a$ mit $a > 0$. Die Iterationsvorschrift des Newtonverfahrens lautet in diesem Fall

$$x_{n+1} = x_n - \frac{\frac{1}{x_n} - a}{-\frac{1}{x_n^2}}$$

$$\begin{aligned}
 &= 2x_n - ax_n^2 \\
 &= x_n(2 - ax_n).
 \end{aligned}$$

Dieses Verfahren zur divisionsfreien Berechnung von Reziproken war schon den alten Ägyptern bekannt.

IV.3. Umkehrfunktionen

IV.3.1. Grundlagen. Die Funktion f sei auf dem Intervall I stetig. Wegen des Zwischenwertsatzes ist das Bild J von I ebenfalls ein Intervall. Falls f auf I injektiv ist, ist dann $f : I \rightarrow J$ bijektiv und besitzt damit eine Umkehrfunktion $f^{-1} : J \rightarrow I$.

Wie kann man leicht nachprüfen, ob f injektiv ist? Nehmen wir dazu an, dass f auch differenzierbar ist. Dann liefert der Mittelwertsatz folgendes, leicht nachprüfbares INJEKTIVITÄTSKRITERIUM:

$$f'(x) \neq 0 \text{ für alle } x \in I \implies f \text{ ist injektiv.}$$

Mit etwas Technik kann man zeigen, dass die Umkehrfunktion f^{-1} dann auch differenzierbar ist.

Wie lautet die Ableitung von f^{-1} ? Für jedes $x \in I$ gilt konstruktionsgemäß $f^{-1}(f(x)) = x$. Wir können diese Identität mit der Kettenregel differenzieren und erhalten

$$1 = \frac{d}{dx}x = \frac{d}{dx}f^{-1}(f(x)) = (f^{-1})'(f(x)) \cdot f'(x).$$

Also ist notwendigerweise

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}.$$

Zusammenfassend erhalten wir:

SATZ ÜBER DIE UMKEHRFUNKTION: Die Funktion f sei auf dem Intervall I differenzierbar, und es gelte $f'(x) \neq 0$ für alle $x \in I$. Dann besitzt f eine Umkehrfunktion $f^{-1} : J \rightarrow I$ mit $J = f(I)$. Die Umkehrfunktion ist differenzierbar auf J , und für alle $y \in J$ gilt

$$(f^{-1})'(y) = \frac{1}{f'(x)} \quad \text{mit } y = f(x).$$

BEISPIEL IV.3.1. Betrachte die Funktion

$$f(x) = x^5 + x$$

auf \mathbb{R} . Für alle $x \in \mathbb{R}$ ist

$$f'(x) = 5x^4 + 1 \geq 1 > 0.$$

Also besitzt f eine Umkehrfunktion. Wegen

$$\lim_{x \rightarrow -\infty} f(x) = -\infty$$

und

$$\lim_{x \rightarrow \infty} f(x) = \infty$$

ist $J = \mathbb{R}$. Schreiben wir zur Abkürzung g statt f^{-1} , erhalten wir für jedes $y \in \mathbb{R}$

$$g'(y) = \frac{1}{5x^4 + 1} \quad \text{mit } y = f(x)$$

bzw. – wegen $x = g(y)$ –

$$g'(y) = \frac{1}{5g(y)^4 + 1}.$$

Aus dieser Formel können wir weitere Ableitungen von g berechnen, z.B.

$$\begin{aligned} g''(y) &= \frac{-20g'(y)g(y)^3}{(5g(y)^4 + 1)^2} \\ &= \frac{-20g(y)^3}{(5g(y)^4 + 1)^3} \\ g'''(y) &= \frac{-60g'(y)g(y)^2}{(5g(y)^4 + 1)^3} + \frac{60g(y)^3 \cdot 20g(y)^3 g'(y)}{(5g(y)^4 + 1)^4} \\ &= \frac{-60g(y)^2}{(5g(y)^4 + 1)^4} + \frac{1200g(y)^6}{(5g(y)^4 + 1)^5} \\ &= \frac{60g(y)^2 (15g(y)^4 - 1)}{(5g(y)^4 + 1)^5}. \end{aligned}$$

Betrachten wir z.B. $y = 2$, erhalten wir wegen $f(1) = 2$

$$\begin{aligned} g(2) &= 1 \\ g'(2) &= \frac{1}{6} \\ g''(2) &= -\frac{20}{6^3} \\ &= -\frac{5}{54} \\ g'''(2) &= \frac{840}{6^5} \\ &= \frac{35}{324}. \end{aligned}$$

Wie wir in Abschnitt VII.3.1 sehen werden, folgt hieraus die Näherungsdarstellung

$$g(y) \approx 1 + \frac{1}{6}(y-2) - \frac{5}{108}(y-2)^2 + \frac{35}{1944}(y-2)^3$$

für $y \approx 2$.

IV.3.2. n -te Wurzel, rationale Exponenten. Für $n \in \mathbb{N}^*$ betrachten wir die Funktion

$$f_n(x) = x^n$$

auf \mathbb{R} . Ist n gerade, ist $f_n(-1) = f_n(1)$. Die Funktion ist also für gerades n nicht injektiv auf \mathbb{R} . Schränken wir sie aber auf das Intervall $(0, \infty)$ ein, erhalten wir

$$f'_n(x) = nx^{n-1} > 0$$

für alle $x > 0$. Die Funktion f_n ist also auf $(0, \infty)$ injektiv. Wegen

$$\lim_{x \rightarrow 0} f_n(x) = 0$$

und

$$\lim_{x \rightarrow \infty} f_n(x) = \infty$$

ist das Bild von $(0, \infty)$ unter f_n gleich $(0, \infty)$. Also ist $f_n : (0, \infty) \rightarrow (0, \infty)$ bijektiv. Die Umkehrfunktion heißt n -TE WURZEL und wird mit $\sqrt[n]{y}$ bezeichnet, d.h.

$$x = \sqrt[n]{y} \iff y = x^n.$$

Sie ist auf $(0, \infty)$ differenzierbar, und für die Ableitung gilt

$$\frac{d}{dy} \sqrt[n]{y} = \frac{1}{nx^{n-1}} \quad \text{mit } y = x^n$$

d.h.

$$\frac{d}{dy} \sqrt[n]{y} = \frac{1}{n (\sqrt[n]{y})^{n-1}} \quad \text{für } y > 0 \text{ und } n \text{ gerade.}$$

Betrachte nun ungerade Exponenten n . Dann ist f_n auf ganz \mathbb{R} injektiv, hat ganz \mathbb{R} als Bildbereich und besitzt eine Umkehrfunktion auf ganz \mathbb{R} . Diese heißt wieder n -TE WURZEL und wird auch mit $\sqrt[n]{y}$ bezeichnet. Sie ist für $y \neq 0$ differenzierbar, und es gilt wieder

$$\frac{d}{dy} \sqrt[n]{y} = \frac{1}{n (\sqrt[n]{y})^{n-1}} \quad \text{für } y \neq 0 \text{ und } n \text{ ungerade.}$$

Mit Hilfe der n -ten Wurzel können wir leicht Potenzen mit rationalem Exponenten definieren:

$$y^{\frac{m}{n}} = (\sqrt[n]{y})^m \quad \text{für } y > 0, \text{ falls } n \text{ gerade,} \\ \text{bzw. } y \text{ beliebig, falls } n \text{ ungerade.}$$

Für die Ableitung von $y^{\frac{m}{n}}$ erhalten wir mit der Kettenregel

$$\begin{aligned} \frac{d}{dy} y^{\frac{m}{n}} &= \frac{d}{dy} (\sqrt[n]{y})^m \\ &= m (\sqrt[n]{y})^{m-1} \frac{d}{dy} \sqrt[n]{y} \\ &= \frac{m (\sqrt[n]{y})^{m-1}}{n (\sqrt[n]{y})^{n-1}} \\ &= \frac{m}{n} y^{\frac{m-n}{n}} \\ &= \frac{m}{n} y^{\frac{m}{n}-1} \end{aligned}$$

also:

$$\frac{d}{dy} y^{\frac{m}{n}} = \frac{m}{n} y^{\frac{m}{n}-1}.$$

BEISPIEL IV.3.2. Im Punkt $B = (0, h)$, $h > 0$, wird eine Lampe angebracht. Ihre Leuchtstärke im Punkt $A = (a, 0)$, $a > 0$, wird durch die Funktion

$$g(h) = h(a^2 + h^2)^{-\frac{3}{2}}$$

beschrieben. Wie ist h zu wählen, damit bei festem a die Leuchtstärke maximal wird?

Wegen

$$\lim_{h \rightarrow 0} g(h) = 0$$

und

$$\lim_{h \rightarrow \infty} g(h) = 0$$

und $g(1) > 0$ muss es mindestens eine Extremalstelle im Intervall $(0, \infty)$ geben. Diese muss ein stationärer Punkt sein. Für die Ableitung von g erhalten wir mit obiger Formel, der Produkt- und der Kettenregel

$$\begin{aligned} g'(h) &= (a^2 + h^2)^{-\frac{3}{2}} - \frac{3}{2} h (a^2 + h^2)^{-\frac{5}{2}} 2h \\ &= (a^2 + h^2 - 3h^2) (a^2 + h^2)^{-\frac{5}{2}} \end{aligned}$$

$$= (a^2 - 2h^2)(a^2 + h^2)^{-\frac{5}{2}}.$$

Also ist die einzige Extremalstelle

$$h = \frac{a}{\sqrt{2}}.$$

Die maximale Leuchtstärke ist

$$\begin{aligned} g\left(\frac{a}{\sqrt{2}}\right) &= \frac{a}{\sqrt{2}} \left(a^2 + \frac{1}{2}a^2\right)^{-\frac{3}{2}} \\ &= \frac{a}{\sqrt{2}} \left(\frac{3}{2}a^2\right)^{-\frac{3}{2}} \\ &= \frac{2}{\sqrt{3}^3} \frac{1}{a^2}. \end{aligned}$$

IV.3.3. Die Umkehrfunktionen der trigonometrischen Funktionen. Betrachte zunächst die Cosinus-Funktion. Sie ist auf dem Intervall $[0, \pi]$ monoton fallend und damit injektiv. Der Wertebereich ist $[-1, 1]$. Also besitzt $\cos : [0, \pi] \rightarrow [-1, 1]$ eine Umkehrfunktion. Diese heißt ARCUSCOSINUS und wird mit \arccos bezeichnet:

$$\begin{aligned} \arccos : [-1, 1] &\longrightarrow [0, \pi] \\ x = \arccos(y) &\iff y = \cos(x). \end{aligned}$$

Für beliebiges $x \in [0, \pi]$ ist

$$\cos'(x) = -\sin(x).$$

Da der Sinus auf $[0, \pi]$ nicht negativ ist, gilt für $x \in [0, \pi]$

$$\sin(x) = \sqrt{1 - \cos^2(x)}$$

und somit

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

Also gilt für die Ableitung des Arcuscosinus

$$\frac{d}{dy} \arccos(y) = -\frac{1}{\sqrt{1 - y^2}} \quad \text{für } -1 < y < 1.$$

Betrachte nun die Sinus-Funktion. Sie ist auf dem Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ monoton wachsend und damit injektiv. Der Wertebereich ist $[-1, 1]$. Die Umkehrfunktion von $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ heißt ARCUSSINUS und wird mit \arcsin bezeichnet:

$$\begin{aligned} \arcsin : [-1, 1] &\longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ x = \arcsin(y) &\iff y = \sin(x). \end{aligned}$$

Wegen

$$\sin'(x) = \cos(x)$$

und

$$\cos(x) = \sqrt{1 - \sin^2(x)}$$

für $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ gilt für die Ableitung

$$\frac{d}{dy} \arcsin(y) = \frac{1}{\sqrt{1 - y^2}} \quad \text{für } -1 < y < 1.$$

Die Tangens-Funktion ist auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ definiert durch

$$\tan(x) = \frac{\sin(x)}{\cos(x)}.$$

Der Wertebereich ist \mathbb{R} . Für die Ableitung gilt nach der Quotientenregel

$$\begin{aligned} \frac{d}{dx} \tan(x) &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} \\ &= \frac{1}{\cos^2(x)} \\ &= 1 + \tan^2(x). \end{aligned}$$

Also besitzt die Tangens Funktion eine Umkehrfunktion $\arctan : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ genannt **ARCUSTANGENS**:

$$\begin{aligned} \arctan : \mathbb{R} &\longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ x = \arctan(y) &\iff y = \tan(x). \end{aligned}$$

Für die Ableitung gilt

$$\frac{d}{dy} \arctan(y) = \frac{1}{1 + y^2} \quad \text{für } y \in \mathbb{R}.$$

Die Cotangens-Funktion schließlich ist auf $(0, \pi)$ definiert durch

$$\cot(x) = \frac{\cos(x)}{\sin(x)}.$$

Der Wertebereich ist \mathbb{R} . Für die Ableitung gilt nach der Quotientenregel

$$\begin{aligned}\frac{d}{dx} \cot(x) &= -\frac{\cos^2(x) + \sin^2(x)}{\sin^2(x)} \\ &= -\frac{1}{\sin^2(x)} \\ &= -1 - \cot^2(x).\end{aligned}$$

Also besitzt die Cotangens Funktion eine Umkehrfunktion $\operatorname{arccot} : \mathbb{R} \rightarrow (0, \pi)$ genannt ARCUSCOTANGENS:

$$\begin{aligned}\operatorname{arccot} : \mathbb{R} &\longrightarrow (0, \pi) \\ x = \operatorname{arccot}(y) &\iff y = \cot(x).\end{aligned}$$

Für die Ableitung gilt

$$\frac{d}{dy} \operatorname{arccot}(y) = -\frac{1}{1+y^2} \quad \text{für } y \in \mathbb{R}.$$

BEISPIEL IV.3.3. Ein Schrank der Länge ℓ und der Breite b soll von einem Gang der Breite L durch eine dazu rechtwinklig gelegene Tür der Breite B transportiert werden. Dabei ist $0 < b < L < B < \ell$. Wie groß dürfen b und ℓ höchstens sein?

Der Winkel φ , bei dem der Schrank verkantet, ist $\varphi = \arccos(\frac{L}{\ell})$ (vgl. Abbildung IV.3.1). Der Schrank passt durch die Tür so lange $b \leq B \cos(\varphi)$ ist. Dies bedeutet $b \leq B \frac{L}{\ell}$ oder $\ell b \leq BL$.

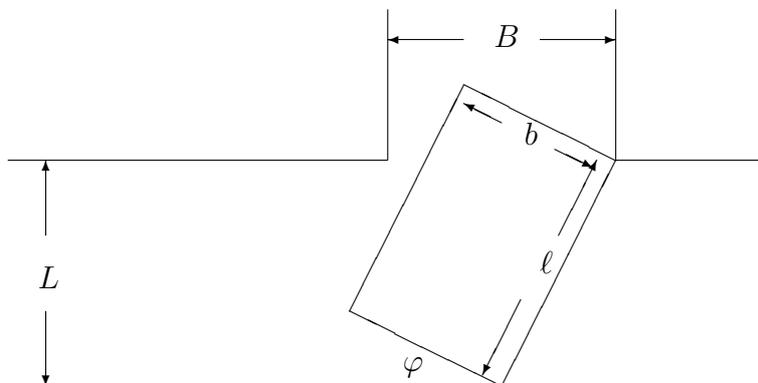


ABBILDUNG IV.3.1. Verkanteter Schrank

IV.4. Die Exponential- und Logarithmusfunktion

IV.4.1. Die Exponentialfunktion. Wir wollen die Differenzierbarkeit der Exponentialfunktion nachweisen. Seien dazu $x \in \mathbb{R}$ und $h > 0$ beliebig. Wegen des Mittelwertsatzes gibt es zu jedem $n \in \mathbb{N}^*$ ein $\eta_n \in (0, h)$ mit

$$\frac{\left(1 + \frac{x+h}{n}\right)^n - \left(1 + \frac{x}{n}\right)^n}{h} = \left(1 + \frac{\eta_n}{n}\right)^{n-1}.$$

Da die linke Seite dieser Gleichung für $n \rightarrow \infty$ konvergiert, gilt Gleiches für die rechte Seite, und wir erhalten

$$\lim_{n \rightarrow \infty} \left(1 + \frac{\eta_n}{n}\right)^{n-1} = \frac{\exp(x+h) - \exp(x)}{h}.$$

Da die Folge $(\eta_n)_{n \in \mathbb{N}}$ beschränkt ist, ist

$$\lim_{n \rightarrow \infty} \left(1 + \frac{\eta_n}{n}\right) = 1$$

und damit

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(1 + \frac{\eta_n}{n}\right)^n &= \lim_{n \rightarrow \infty} \left(1 + \frac{\eta_n}{n}\right)^{n-1} \\ &= \frac{\exp(x+h) - \exp(x)}{h}. \end{aligned}$$

Für alle hinreichend großen $n \in \mathbb{N}^*$ gilt andererseits

$$\begin{aligned} \left(1 + \frac{x}{n}\right)^n &\leq \left(1 + \frac{\eta_n}{n}\right)^n \\ &\leq \left(1 + \frac{x+h}{n}\right)^n. \end{aligned}$$

Grenzübergang $n \rightarrow \infty$ liefert

$$\begin{aligned} \exp(x) &= \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \\ &\leq \lim_{n \rightarrow \infty} \left(1 + \frac{\eta_n}{n}\right)^n \\ &\leq \lim_{n \rightarrow \infty} \left(1 + \frac{x+h}{n}\right)^n \\ &\leq \exp(x+h). \end{aligned}$$

Also ist

$$\begin{aligned} \exp(x) &\leq \frac{\exp(x+h) - \exp(x)}{h} \\ &\leq \exp(x+h). \end{aligned}$$

Da die Exponentialfunktion stetig ist, folgt

$$\lim_{h \rightarrow 0^+} \frac{\exp(x+h) - \exp(x)}{h} = \exp(x).$$

Ganz analog erhalten wir

$$\lim_{h \rightarrow 0^-} \frac{\exp(x+h) - \exp(x)}{h} = \exp(x).$$

Dies beweist:

Die Exponentialfunktion ist auf \mathbb{R} differenzierbar und erfüllt die DIFFERENTIALGLEICHUNG

$$\exp' = \exp.$$

Als nächstes wollen wir zeigen, dass jede Lösung der Differentialgleichung

$$y' = y$$

ein Vielfaches der Exponentialfunktion ist. Sei dazu $y : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar mit $y' = y$. Setze $u(x) = y(x) \exp(-x)$. Aus der Produkt- und Kettenregel folgt

$$\begin{aligned} u'(x) &= y'(x) \exp(-x) - y(x) \exp'(-x) \\ &= y(x) \exp(-x) - y(x) \exp(-x) \\ &= 0. \end{aligned}$$

Also ist u konstant, d.h. es gibt ein $c \in \mathbb{R}$ mit

$$u(x) = c \quad \text{bzw.} \quad y(x) = c \exp(x)$$

für alle x .

Ist $y : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit $y' = y$, so gibt es ein $c \in \mathbb{R}$ mit $y(x) = c \exp(x)$ für alle $x \in \mathbb{R}$.

Sei nun $z \in \mathbb{R}$ beliebig, aber fest gewählt. Betrachte die Funktion $y(x) = \exp(x+z)$. Diese Funktion ist differenzierbar und erfüllt $y' = y$. Also ist $y(x) = \exp(x+z) = c \exp(x)$ für alle $x \in \mathbb{R}$ für ein $c \in \mathbb{R}$. Einsetzen von $x = 0$ liefert $\exp(z) = y(0) = c \exp(0) = c$. Dies beweist:

Für alle $x, z \in \mathbb{R}$ gilt

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

Aus dieser Funktionalgleichung für die Exponentialfunktion ergibt sich für alle $m \in \mathbb{N}$, $n \in \mathbb{N}^*$ die Beziehung

$$\exp\left(\frac{m}{n}\right) = \exp(1)^{\frac{m}{n}}.$$

Aus diesem Grund ist die Schreibweise $\exp(x) = e^x$ mit der Interpretation „ e hoch x “ gerechtfertigt.

Als nächstes untersuchen wir das Verhalten der Exponentialfunktion für $x \rightarrow \pm\infty$. Seien dazu $x > 0$ und $n \in \mathbb{N}^*$ beliebig. Aus der Binomischen Formel folgt

$$\begin{aligned} \frac{1}{x^n} \exp(x) &\geq \frac{1}{x^n} \left(1 + \frac{x}{n+1}\right)^{n+1} \\ &= \frac{1}{x^n} \sum_{k=0}^{n+1} \underbrace{\binom{n+1}{k}}_{\geq 0} \frac{x^k}{(n+1)^k} \\ &\geq \frac{1}{x^n} \frac{x^{n+1}}{(n+1)^{n+1}} \\ &= \frac{x}{(n+1)^{n+1}} \\ &\xrightarrow{x \rightarrow \infty} \infty. \end{aligned}$$

Hieraus folgt:

$$\begin{aligned} \lim_{x \rightarrow \infty} \exp(x) &= \infty, \\ \lim_{x \rightarrow -\infty} \exp(x) &= 0, \\ \lim_{x \rightarrow \infty} x^{-n} \exp(x) &= \infty \quad \text{für alle } n \in \mathbb{N}^*. \end{aligned}$$

Die letzte Beziehung kann man auch so formulieren:

Die Exponentialfunktion wächst für $x \rightarrow \infty$ schneller als jede Potenz von x .

BEISPIEL IV.4.1 (RADIOAKTIVER ZERFALL). Aufgrund von Beobachtungen geht man davon aus, dass für $N(t)$ Atome einer radioaktiven Substanz die Anzahl ΔN der Zerfälle in einem kleinen Zeitintervall Δt gegeben ist durch $\Delta N = -kN\Delta t$. Unter der Annahme, dass N differenzierbar ist, ergibt sich hieraus die Differentialgleichung

$$N' = -kN,$$

wobei $k > 0$ die vom Material abhängige Zerfallskonstante ist. Also ist $N(t) = N(0)e^{-kt}$. Die HALBWERTSZEIT T ist durch $N(T) = \frac{1}{2}N(0)$ definiert. Mit der Logarithmusfunktion \ln aus Abschnitt III.3.8 folgt

$$T = \frac{1}{k} \ln(2).$$

BEISPIEL IV.4.2 (NEWTONSCHES ABKÜHLUNGSGESETZ). Eine Substanz der Temperatur $T(t)$ wird in ein Medium der Temperatur

$U(t)$ eingetaucht. Für die Temperaturdifferenz $P(t) = T(t) - U(t)$ fand I. Newton in zahlreichen Versuchen die Beziehung

$$P' = \lambda P,$$

d.h. die Abkühlungsgeschwindigkeit ist proportional zur Temperaturdifferenz. Also ist

$$P(t) = P(0)e^{\lambda t}$$

bzw.

$$T(t) = (T(0) - U(0))e^{\lambda t} + U(t)$$

IV.4.2. Die Logarithmusfunktion. In Abschnitt III.3.8 habe wir die Logarithmusfunktion $\ln : (0, \infty) \rightarrow \mathbb{R}$ als Umkehrfunktion der Exponentialfunktion eingeführt. Wegen $\exp'(x) = \exp(x) > 0$ für alle $x \in \mathbb{R}$ ist gemäß Abschnitt IV.3.1 die Logarithmusfunktion differenzierbar und erfüllt

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

Aus den Beziehungen $\exp(x + y) = \exp(x) \cdot \exp(y)$ und $\exp(x) \cdot \exp(-x) = 1$ folgt weiter

$$\begin{aligned} \ln(xy) &= \ln(x) + \ln(y) \\ \ln\left(\frac{x}{y}\right) &= \ln(x) - \ln(y) \quad \text{für } x, y > 0. \end{aligned}$$

Mit Hilfe der Regel von de l'Hôpital erhält man:

$$\lim_{x \rightarrow \infty} \frac{\ln(x)}{\sqrt[n]{x}} = 0 \quad \text{für jedes } n \in \mathbb{N}^*.$$

Dies kann man auch so formulieren:

Die Logarithmusfunktion wächst für $x \rightarrow \infty$ langsamer als jede Wurzel von x .

BEISPIEL IV.4.3. Ein Reiz der Intensität R verursacht eine Empfindung der Intensität $E(R)$. Versuche ergeben, dass die Änderung ΔE der Empfindung proportional ist zur relativen Änderung der Reizintensität, d.h. $\Delta E = a \frac{\Delta R}{R}$. Dies führt auf die Differentialgleichung

$$\frac{d}{dR} E(R) = \frac{a}{R}.$$

Also ist $E(R) = a \ln(R) + c$ mit einer Konstanten c . Als Reizschwelle bezeichnet man den Wert R_0 , für den gilt $E(R_0) = 0$ und $E(R) > 0$ für $R > R_0$. Damit ergibt sich $c = -a \ln(R_0)$ und

$$E(R) = a \ln(R) - a \ln(R_0) = a \ln\left(\frac{R}{R_0}\right).$$

Wegen des langsamen Wachstums der Logarithmusfunktion ist bei hohen Intensitäten eine ganz erhebliche Änderung von R nötig, um eine merkbare Änderung von E hervorzurufen.

IV.4.3. Exponential- und Logarithmusfunktionen zu beliebigen Basen. Sei $a > 0$. Die Beziehung

$$(e^x)^r = \exp(x)^r = \exp(rx) = e^{rx}$$

mit $x = \ln(a)$ ergibt

$$a^r = e^{r \ln(a)}$$

für jede rationale Zahl r . Daher definiert man die Potenzen a^x für beliebiges $x \in \mathbb{R}$ durch

$$a^x = e^{x \ln(a)} \quad \text{für alle } a > 0, x \in \mathbb{R}.$$

Man nennt $x \rightarrow a^x$ die EXPONENTIALFUNKTION ZUR BASIS a .

Aus den Eigenschaften der exp-Funktion folgen die Eigenschaften der Funktion a^x :

$$\begin{aligned} a^x a^y &= a^{x+y}, \\ (ab)^x &= a^x b^x, \\ (a^x)^y &= a^{xy}, \\ \ln(a^x) &= x \ln(a), \\ \frac{d}{dx}(a^x) &= a^x \ln(a). \end{aligned}$$

Für jedes $\alpha \in \mathbb{R}$ ist die Potenzfunktion $x^\alpha = e^{\alpha \ln(x)}$ für alle $x > 0$ definiert. Aus der Kettenregel folgt

$$\frac{d}{dx} x^\alpha = \alpha x^{\alpha-1}.$$

BEISPIEL IV.4.4 (POISSONSCHES GASGESETZ). Experimente zeigen, dass zwischen dem Druck p und dem Volumen V eines Gases bei reversiblen, adiabatischen Prozessen der Zusammenhang

$$\frac{\Delta p}{p} + \gamma \frac{\Delta V}{V} = 0$$

besteht, wobei γ eine vom jeweiligen Gas abhängige Konstante ist. Nach Division durch Δt geht dies über in

$$\frac{1}{p} \frac{\Delta p}{\Delta t} + \gamma \frac{1}{V} \frac{\Delta V}{\Delta t} = 0,$$

was im Grenzübergang $\Delta t \rightarrow 0$ auf die Differentialgleichung

$$\frac{p'}{p} + \gamma \frac{V'}{V} = 0$$

führt. Wegen der Kettenregel bedeutet dies

$$\frac{d}{dt}(\ln(p) + \gamma \ln(V)) = 0.$$

Also ist

$$\ln(p) + \gamma \ln(V) = c$$

bzw.

$$\ln(pV^\gamma) = c$$

und

$$pV^\gamma = C$$

mit

$$C = e^c.$$

BEISPIEL IV.4.5. Die ALLGEMEINE HÖHENFORMEL

$$p = p_0 \left(1 - \frac{\gamma - 1}{\gamma} \frac{\rho_0}{p_0} gh \right)^{\frac{\gamma}{\gamma-1}}$$

beschreibt den Zusammenhang zwischen dem Druck p und der Höhe h in einem gasgefüllten Behälter. Uns interessiert der Grenzübergang $\gamma \rightarrow 1$, d.h. der Übergang zu einem idealen Gas. Mit $x = \frac{\gamma}{\gamma-1}$ und $a = -\frac{\rho_0}{p_0} gh$ wird dieser Grenzübergang durch

$$p = p_0 \lim_{x \rightarrow \infty} \left(1 + \frac{a}{x} \right)^x = p_0 \lim_{x \rightarrow \infty} e^{x \ln(1 + \frac{a}{x})}$$

beschrieben. Mit der Regel von de l'Hôpital erhalten wir

$$\begin{aligned} \lim_{x \rightarrow \infty} x \ln\left(1 + \frac{a}{x}\right) &= \lim_{x \rightarrow \infty} \frac{\ln\left(1 + \frac{a}{x}\right)}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{1 + \frac{a}{x}} \cdot \frac{-a}{x^2}}{-\frac{1}{x^2}} \\ &= \lim_{x \rightarrow \infty} \frac{a}{1 + \frac{a}{x}} \\ &= a. \end{aligned}$$

Daher geht die allgemeine Höhenformel für $\gamma \rightarrow 1$ in die BAROMETRISCHE HÖHENFORMEL

$$p = p_0 e^{-\frac{\rho_0}{p_0} gh}$$

über.

Zu jeder Basis $a > 0$ ist die Funktion a^x umkehrbar. Denn für $y > 0$ hat die Gleichung $y = a^x = e^{x \ln(a)}$ die eindeutige Lösung $x = \frac{\ln(y)}{\ln(a)}$. Diese Umkehrfunktion wird mit \log_a bezeichnet und heißt LOGARITHMUSFUNKTION ZUR BASIS a . Die explizite Darstellung ist

$$\log_a(z) = \frac{\ln(z)}{\ln(a)}.$$

Aus den Rechenregeln für den üblichen Logarithmus ergeben sich folgende Regeln:

$$\begin{aligned} \log_a(xy) &= \log_a(x) + \log_a(y), \\ \log_a\left(\frac{x}{y}\right) &= \log_a(x) - \log_a(y), \\ \frac{d}{dx} \log_a(x) &= \frac{1}{x \ln(a)}. \end{aligned}$$

IV.4.4. Die Hyperbelfunktionen. Die Hyperbelfunktionen sind definiert durch:

$$\begin{aligned} \sinh(x) &= \frac{e^x - e^{-x}}{2} && \text{SINUS HYPERBOLICUS} \\ \cosh(x) &= \frac{e^x + e^{-x}}{2} && \text{COSINUS HYPERBOLICUS} \\ \tanh(x) &= \frac{\sinh(x)}{\cosh(x)} \\ &= \frac{e^x - e^{-x}}{e^x + e^{-x}} && \text{TANGENS HYPERBOLICUS} \end{aligned}$$

Es gelten folgende, zu den trigonometrischen Funktionen analoge Rechenregeln:

$$\begin{aligned} \sinh(-x) &= -\sinh(x), \\ \cosh(-x) &= \cosh(x), \\ \sinh(x+y) &= \sinh(x) \cosh(y) + \cosh(x) \sinh(y), \end{aligned}$$

$$\begin{aligned}\cosh(x+y) &= \cosh(x)\cosh(y) + \sinh(x)\sinh(y), \\ \cosh^2(x) - \sinh^2(x) &= 1.\end{aligned}$$

Für die Ableitungen ergibt sich:

$$\begin{aligned}\sinh' &= \cosh, \\ \cosh' &= \sinh, \\ \tanh' &= \frac{1}{\cosh^2}.\end{aligned}$$

Da die Ableitung des Sinus hyperbolicus auf ganz \mathbb{R} positiv ist, besitzt er eine Umkehrfunktion auf \mathbb{R} . Diese wird mit $\operatorname{arsinh} : \mathbb{R} \rightarrow \mathbb{R}$ bezeichnet und heißt AREA SINUS HYPERBOLICUS. Aus dem gleichen Grund besitzt der Cosinus hyperbolicus eine Umkehrfunktion auf $[0, \infty)$. Diese wird mit $\operatorname{arcosh} : [1, \infty) \rightarrow [0, \infty)$ bezeichnet und heißt AREA COSINUS HYPERBOLICUS. Durch Auflösen der Gleichungen $y = \frac{1}{2}(e^x - e^{-x})$ und $y = \frac{1}{2}(e^x + e^{-x})$ nach x erhält man explizite Darstellungen dieser Funktionen:

$$\begin{aligned}\operatorname{arsinh}(y) &= \ln\left(y + \sqrt{y^2 + 1}\right) && \text{für } -\infty < y < \infty, \\ \operatorname{arcosh}(y) &= \ln\left(y + \sqrt{y^2 - 1}\right) && \text{für } 1 \leq y < \infty.\end{aligned}$$

Da die Ableitung des Tangens hyperbolicus auf \mathbb{R} positiv ist, besitzt er eine Umkehrfunktion auf \mathbb{R} . Diese wird mit $\operatorname{artanh} : (-1, 1) \rightarrow \mathbb{R}$ bezeichnet und hat die explizite Darstellung

$$\operatorname{artanh}(y) = \frac{1}{2} \ln\left(\frac{1+y}{1-y}\right) \quad \text{für } -1 < y < 1.$$

Für die Ableitungen dieser drei Umkehrfunktionen ergibt sich aus Abschnitt IV.3.1:

$$\begin{aligned}\frac{d}{dx} \operatorname{arsinh}(x) &= \frac{1}{\sqrt{x^2 + 1}} && \text{für } x \in \mathbb{R}, \\ \frac{d}{dx} \operatorname{arcosh}(x) &= \frac{1}{\sqrt{x^2 - 1}} && \text{für } x > 1, \\ \frac{d}{dx} \operatorname{artanh}(x) &= \frac{1}{1 - x^2} && \text{für } -1 < x < 1.\end{aligned}$$

KAPITEL V

Integration

V.1. Das bestimmte Integral

V.1.1. Definition. Sofern nicht anders vermerkt, betrachten wir in diesem Kapitel stets Funktionen f , die auf einem abgeschlossenen Intervall $[a, b]$ definiert, beschränkt und STÜCKWEISE STETIG sind. Dabei bedeutet Letzteres, dass in jedem Punkt $x \in [a, b]$ die links- und rechtsseitigen Grenzwerte von f existieren und endlich sind (mit der offensichtlichen Modifikation für $x = a$ und $x = b$). Beispiele für derartige Funktionen sind stetige Funktionen oder die Vorzeichen- und Entier-Funktion aus Beispiel III.3.7(2,3) (S. 122), die nicht stetig sind.

Wir wollen f eine Zahl zuordnen, die die Fläche zwischen der x -Achse und der Kurve $y = f(x)$ misst. Dabei sollen Kurvenstücke oberhalb der x -Achse eine positive Fläche und solche unterhalb der x -Achse eine negative Fläche haben. Um dieses Ziel zu erreichen, wählen wir $n \in \mathbb{N}^*$, $n + 1$ Punkte x_0, x_1, \dots, x_n mit $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$ sowie Punkte η_1, \dots, η_n mit $x_{i-1} \leq \eta_i \leq x_i$ für $1 \leq i \leq n$ und setzen

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

Z_n ist die Summe der Flächen der Rechtecke mit Breite $x_i - x_{i-1}$ und Höhe $f(\eta_i)$ (vgl. Abbildung V.1.1). Dabei wird die Fläche positiv oder negativ genommen, je nachdem ob das Rechteck oberhalb oder unterhalb der x -Achse liegt.

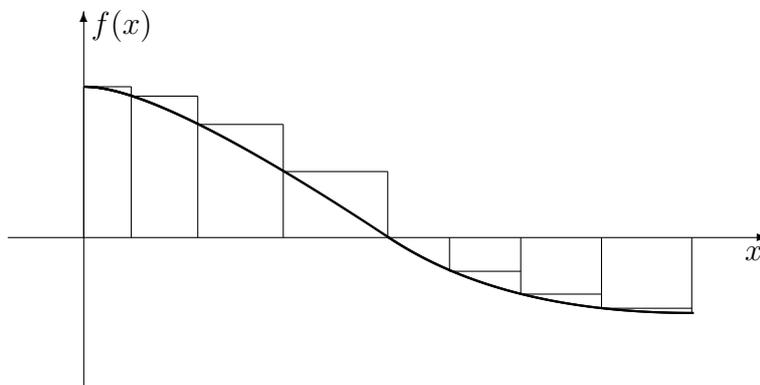


ABBILDUNG V.1.1. Z_n mit $\eta_i = x_{i-1}$ für alle i

Die Zahl Z_n hängt natürlich von der Wahl von n und der Punkte x_0, \dots, x_n und η_1, \dots, η_n ab. Man kann aber zeigen, dass die Folge $(Z_n)_{n \in \mathbb{N}}$ konvergiert, wenn die „Feinheit“ $\max_{1 \leq i \leq n} (x_i - x_{i-1})$ gegen Null strebt und dass der Grenzwert *nicht* von der jeweiligen Wahl der Punkte x_0, \dots, x_n und η_1, \dots, η_n abhängt. Wir nennen diesen Grenzwert das (BESTIMMTE) INTEGRAL von f über $[a, b]$ und bezeichnen ihn mit $\int_a^b f(x) dx$:

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\eta_i)(x_i - x_{i-1}).$$

BEMERKUNG V.1.1. Das Integral ist eine Zahl; die „Integrationsvariable“ kann beliebig bezeichnet werden:

$$\int_a^b f(x) dx = \int_a^b f(t) dt = \int_a^b f(u) du.$$

Zur Vermeidung von Fallunterscheidungen setzt man

$$\int_a^a f(x) dx = 0,$$

$$\int_b^a f(x) dx = - \int_a^b f(x) dx \quad \text{falls } b \geq a.$$

BEISPIEL V.1.2. Es ist

$$\int_a^b c dx = c(b - a),$$

$$\int_a^b x dx = \frac{1}{2}(b^2 - a^2).$$

V.1.2. Elementare Integrationsregeln. Das bestimmte Integral ist linear, monoton und additiv:

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx$$

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

falls $f(x) \leq g(x)$ für alle $x \in [a, b]$

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

falls $a \leq c \leq b$.

Aus der Monotonie des Integrals ergeben sich folgende hilfreiche Abschätzungen:

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

$$(b-a) \min_{a \leq x \leq b} f(x) \leq \int_a^b f(x) dx \leq (b-a) \max_{a \leq x \leq b} f(x).$$

Für die Integration gibt es ein Analogon zum Mittelwertsatz der Differentialrechnung, Abschnitt IV.2.2:

MITTELWERTSATZ DER INTEGRALRECHNUNG: Die Funktionen f und g seien auf $[a, b]$ stetig und es gelte $g(x) \geq 0$ für alle $x \in [a, b]$. Dann gibt es ein $\eta \in [a, b]$ mit

$$\int_a^b f(x)g(x)dx = f(\eta) \int_a^b g(x)dx.$$

BEWEIS. Seien m und M das Minimum und das Maximum von f auf $[a, b]$. Dann ist

$$mg(x) \leq f(x)g(x) \leq Mg(x)$$

für alle $x \in [a, b]$ und daher

$$m \int_a^b g(x)dx \leq \int_a^b f(x)g(x)dx \leq M \int_a^b g(x)dx.$$

Nach dem Zwischenwertsatz muss die stetige Funktion

$$y \mapsto f(y) \int_a^b g(x)dx$$

den Wert $\int_a^b f(x)g(x)dx$ annehmen. □

V.1.3. Differentiation und Integration. Die bisherige Definition des bestimmten Integrals ist für praktische Rechnungen ungeeignet. Dieses Manko wird durch den Begriff der Stammfunktion und den Hauptsatz der Differential- und Integralrechnung behoben.

Eine Funktion F heißt **STAMMFUNKTION** der Funktion f auf dem Intervall I , wenn F auf I differenzierbar ist und für alle $x \in I$ gilt

$$F'(x) = f(x).$$

Aus Abschnitt IV.2.2 folgt:

Sind F und G zwei Stammfunktionen der Funktion f , so ist die Differenz $F - G$ konstant.

Der folgende fundamentale Satz stellt einen Zusammenhang her zwischen Stammfunktion und Integral:

HAUPTSATZ DER DIFFERENTIAL- UND INTEGRALRECHNUNG: Die Funktion f sei auf dem Intervall I beschränkt und stetig; a, b seien zwei beliebige Zahlen aus I .

(1) Die durch

$$F_a(x) = \int_a^x f(t)dt$$

für alle $x \in I$ definierte Funktion F_a ist eine Stammfunktion von f .

(2) Ist F irgendeine Stammfunktion von f , so ist

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

BEWEIS. AD (1): Für jedes $x \in I$ und jedes hinreichend kleine $h \neq 0$ gibt es gemäß dem Mittelwertsatz der Integralrechnung ein η_h zwischen x und $x + h$ mit

$$\begin{aligned} F_a(x+h) - F_a(x) &= \int_a^{x+h} f(t)dt - \int_a^x f(t)dt \\ &= \int_x^{x+h} f(t)dt \\ &= hf(\eta_h). \end{aligned}$$

Da $\eta_h \xrightarrow{h \rightarrow 0} x$ gilt und f stetig ist, folgt

$$\begin{aligned} F'_a(x) &= \lim_{h \rightarrow 0} \frac{F_a(x+h) - F_a(x)}{h} \\ &= \lim_{h \rightarrow 0} f(\eta_h) \\ &= f(x). \end{aligned}$$

AD (2): Wegen Teil (1) gibt es ein $c \in \mathbb{R}$ mit $F(x) = F_a(x) + c$. Damit folgt

$$\begin{aligned} F(b) - F(a) &= F_a(b) - \underbrace{F_a(a)}_{=0} \\ &= F_a(b) \\ &= \int_a^b f(t)dt. \end{aligned}$$

Dies beweist die Behauptung. \square

Wegen des Hauptsatzes der Differential- und Integralrechnung geht man bei der Berechnung von $\int_a^b f(x)dx$ wie folgt vor:

- Bestimme eine Stammfunktion F von f .
- Berechne $F(b) - F(a) = \int_a^b f(x)dx$.

Die Menge *aller* Stammfunktionen von f wird mit $\int f(x)dx$ bezeichnet und heißt UNBESTIMMTES INTEGRAL von f . Es gilt also definitionsgemäß

$$\int f(x)dx = F(x) + c \iff F'(x) = f(x).$$

Tabelle V.1.1 enthält eine Liste der Stammfunktionen einiger elementarer Funktionen.

TABELLE V.1.1. Stammfunktionen einiger elementarer Funktionen

$f(x)$	$\int f(x)dx$	Bemerkung
x^n	$\frac{1}{n+1}x^{n+1} + c$	$n \neq -1$
$\frac{1}{x}$	$\ln x + c$	$x \neq 0$
$\sin(x)$	$-\cos(x) + c$	
$\cos(x)$	$\sin(x) + c$	
$\frac{1}{\cos^2(x)}$	$\tan(x) + c$	$x \neq (2k+1)\frac{\pi}{2}, k \in \mathbb{K}$
$-\frac{1}{\sin^2(x)}$	$\cot(x) + c$	$x \neq k\pi, k \in \mathbb{K}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x) + c$	$ x < 1$
$\frac{1}{1+x^2}$	$\arctan(x) + c$	
$\frac{1}{1-x^2}$	$\frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) + c$	$ x < 1$
e^{ax}	$\frac{1}{a}e^{ax} + c$	$a \neq 0$
$\sinh(x)$	$\cosh(x) + c$	
$\cosh(x)$	$\sinh(x) + c$	
$\frac{1}{\sqrt{1+x^2}}$	$\ln(x + \sqrt{1+x^2}) + c$	
$\frac{1}{\sqrt{x^2-1}}$	$\ln x + \sqrt{x^2-1} + c$	$ x > 1$

BEISPIEL V.1.3. Die von der Kraft $K(x) = \gamma m M x^{-2}$ längs $[a, b]$, $0 < a < b$, geleistete Arbeit beträgt

$$\begin{aligned} A &= \int_a^b K(x) dx \\ &= \int_a^b \gamma m M x^{-2} dx \\ &= \gamma m M \left(-\frac{1}{x} \right) \Big|_a^b \\ &= \gamma m M \left(\frac{1}{a} - \frac{1}{b} \right). \end{aligned}$$

BEISPIEL V.1.4. Ein biegeweiches Seil nimmt bei einer spezifischen Längenbelastung $q(x)$ die Form einer SEILKURVE

$$y(x) = c \int_0^x \left(\int_0^u q(t) dt \right) du$$

an. Für eine HÄNGEBRÜCKE mit konstantem $q(x) = q_0$ erhält man

$$\begin{aligned} y(x) &= c \int_0^x \left(\int_0^u q_0 dt \right) du \\ &= c \int_0^x q_0 u du \\ &= \frac{1}{2} c q_0 x^2. \end{aligned}$$

Für die KETTENLINIE (Seil unter Eigenlast) mit $q(x) = \cosh(x)$ erhält man

$$\begin{aligned} y(x) &= c \int_0^x \left(\int_0^u \cosh(t) dt \right) du \\ &= c \int_0^x \sinh(u) du \quad \text{wegen } \sinh(0) = 0 \\ &= c(\cosh(x) - 1) \quad \text{wegen } \cosh(0) = 1. \end{aligned}$$

V.2. Integrationsregeln

V.2.1. Linearität. Das unbestimmte Integral ist linear:

$$\int (\alpha f(x) + \beta g(x)) dx = \alpha \int f(x) dx + \beta \int g(x) dx$$

BEISPIEL V.2.1. Es ist

$$\int (a_n x^n + \dots + a_1 x + a_0) dx = \frac{1}{n+1} a_n x^{n+1} + \dots + \frac{1}{2} a_1 x^2 + a_0 x + c.$$

V.2.2. Partielle Integration. Die Funktionen u und v seien auf dem Intervall I differenzierbar mit stetigen Ableitungen. Nach der Produktregel ist

$$(uv)' = u'v + uv'.$$

Wegen des Hauptsatzes der Differential- und Integralrechnung folgt hieraus

$$u(x)v(x) = \int u'(x)v(x)dx + \int u(x)v'(x)dx$$

bzw.

$$\int u'(x)v(x)dx = u(x)v(x) - \int u(x)v'(x)dx.$$

Für das bestimmte Integral bedeutet dies

$$\begin{aligned} \int_a^b u'(x)v(x)dx &= u(x)v(x) \Big|_a^b - \int_a^b u(x)v'(x)dx \\ &= \left(u(b)v(b) - u(a)v(a) \right) - \int_a^b u(x)v'(x)dx. \end{aligned}$$

BEISPIEL V.2.2. (1) Mit $u'(x) = e^x$, $v(x) = x$ und $u(x) = e^x$, $v'(x) = 1$ ergibt sich

$$\begin{aligned} \int xe^x dx &= xe^x - \int e^x dx \\ &= xe^x - e^x + c \\ &= (x-1)e^x + c. \end{aligned}$$

(2) Mit $u'(x) = \sin(x)$, $v(x) = x^2$ und $u(x) = -\cos(x)$, $v'(x) = 2x$ ergibt sich

$$\int x^2 \sin(x)dx = -x^2 \cos(x) + 2 \int x \cos(x)dx.$$

Das Integral auf der rechten Seite berechnen wir mit dem Ansatz $u'(x) = \cos(x)$, $v(x) = x$, $u(x) = \sin(x)$, $v'(x) = 1$ und erhalten

$$\begin{aligned} \int x \cos(x)dx &= x \sin(x) - \int \sin(x)dx \\ &= x \sin(x) + \cos(x) + c. \end{aligned}$$

Insgesamt ergibt sich (mit einer anderen Konstanten c !)

$$\int x^2 \sin(x)dx = -x^2 \cos(x) + 2x \sin(x) + 2 \cos(x) + c.$$

(3) Für $u'(x) = \sin(x)$, $v(x) = \sin(x)$, $u(x) = -\cos(x)$, $v'(x) = \cos(x)$ erhalten wir

$$\begin{aligned}\int \sin^2(x) dx &= \int \sin(x) \sin(x) dx \\ &= -\cos(x) \sin(x) + \int \underbrace{\cos(x) \cos(x)}_{=\cos^2(x)=1-\sin^2(x)} dx \\ &= -\cos(x) \sin(x) + \int (1 - \sin^2(x)) dx \\ &= -\cos(x) \sin(x) + x + c - \int \sin^2(x) dx.\end{aligned}$$

Auflösen nach $\int \sin^2(x) dx$ liefert (mit einer anderen Konstanten c !)

$$\int \sin^2(x) dx = -\frac{1}{2} \cos(x) \sin(x) + \frac{1}{2} x + c.$$

(4) Wegen $\cos^2(x) + \sin^2(x) = 1$ für alle x folgt aus Teil (3)

$$\begin{aligned}\int \cos^2(x) dx &= \int (1 - \sin^2(x)) dx \\ &= x - \int \sin^2(x) dx \\ &= \frac{1}{2} x + \frac{1}{2} \cos(x) \sin(x) + c.\end{aligned}$$

BEISPIEL V.2.3. Mit dem Ansatz $u'(x) = 1$, $u(x) = x$ erhält man

$$\int v(x) dx = xv(x) - \int xv'(x) dx.$$

Dies liefert z.B.

$$\begin{aligned}\int \ln(x) dx &= x \ln(x) - \int x \frac{1}{x} dx \\ &= x \ln(x) - x + c\end{aligned}$$

und

$$\begin{aligned}\int \arcsin(x) dx &= x \arcsin(x) - \int x \frac{1}{\sqrt{1-x^2}} dx \\ &= x \arcsin(x) + \sqrt{1-x^2} + c.\end{aligned}$$

BEISPIEL V.2.4. Bei Integralen der Form

$$\begin{aligned}S_n &= \int_a^b [\sin(x)]^n dx, \\ C_n &= \int_a^b [\cos(x)]^n dx,\end{aligned}$$

$$\begin{aligned}
 A_n &= \int_a^b x^n \sin(x) dx, \\
 B_n &= \int_a^b x^n \cos(x) dx, \\
 E_n &= \int_a^b x^n e^x dx, \\
 L_n &= \int_a^b [\ln(x)]^n dx
 \end{aligned}$$

liefert die partielle Integration jeweils eine Rekursionsformel.

Wir erläutern dies an zwei Beispielen.

Mit $u'(x) = \sin(x)$, $v(x) = [\sin(x)]^{n-1}$ und $u(x) = -\cos(x)$, $v'(x) = (n-1)[\sin(x)]^{n-2} \cos(x)$ erhalten wir für $n \geq 2$ und $a = 0$, $b = \frac{\pi}{2}$

$$\begin{aligned}
 S_n &= \int_0^{\frac{\pi}{2}} [\sin(x)]^n dx \\
 &= \underbrace{-\cos(x) [\sin(x)]^{n-1}}_{=0} \Big|_0^{\frac{\pi}{2}} + (n-1) \int_0^{\frac{\pi}{2}} [\sin(x)]^{n-2} \cos^2(x) dx \\
 &= (n-1) \int_0^{\frac{\pi}{2}} [\sin(x)]^{n-2} [1 - \sin^2(x)] dx \\
 &= (n-1)S_{n-2} - (n-1)S_n.
 \end{aligned}$$

Also gilt

$$S_n = \frac{n-1}{n} S_{n-2}.$$

Wegen $S_0 = \frac{\pi}{2}$ und $S_1 = \cos(0) - \cos(\frac{\pi}{2}) = 1$ folgt hieraus

$$S_n = \begin{cases} \frac{n-1}{n} \cdot \frac{n-3}{n-2} \cdot \dots \cdot \frac{1}{2} \cdot \frac{\pi}{2} & \text{falls } n \text{ gerade,} \\ \frac{n-1}{n} \cdot \frac{n-3}{n-2} \cdot \dots \cdot \frac{2}{3} \cdot 1 & \text{falls } n \text{ ungerade.} \end{cases}$$

Als zweites Beispiel betrachten wir

$$\begin{aligned}
 A_n &= \int_0^{\frac{\pi}{2}} x^n \sin(x) dx, \\
 B_n &= \int_0^{\frac{\pi}{2}} x^n \cos(x) dx.
 \end{aligned}$$

Mit $u'(x) = \sin(x)$, $v(x) = x^n$, $u(x) = -\cos(x)$, $v'(x) = nx^{n-1}$ erhalten wir für $n \geq 1$

$$\begin{aligned}
 A_n &= \underbrace{-\cos(x)x^n}_{=0} \Big|_0^{\frac{\pi}{2}} + n \int_0^{\frac{\pi}{2}} x^{n-1} \cos(x) dx \\
 &= nB_{n-1}.
 \end{aligned}$$

Analog erhalten wir mit $u'(x) = \cos(x)$, $v(x) = x^n$

$$\begin{aligned} B_n &= \underbrace{\sin(x)x^n \Big|_0^{\frac{\pi}{2}}}_{=\left(\frac{\pi}{2}\right)^n} - n \int_0^{\frac{\pi}{2}} x^{n-1} \sin(x) dx \\ &= \left(\frac{\pi}{2}\right)^n - nA_{n-1}. \end{aligned}$$

Wir setzen nun die beiden Formeln ineinander ein:

$$\begin{aligned} A_n &= nB_{n-1} \\ &= n \left[\left(\frac{\pi}{2}\right)^{n-1} - (n-1)A_{n-2} \right] \\ &= n \left(\frac{\pi}{2}\right)^{n-1} - n(n-1)A_{n-2} \\ B_n &= \left(\frac{\pi}{2}\right)^n - nA_{n-1} \\ &= \left(\frac{\pi}{2}\right)^n - n(n-1)B_{n-2}. \end{aligned}$$

Zusammen mit $A_0 = B_0 = 1$ können hieraus alle A_n und B_n rekursiv berechnet werden.

V.2.3. Substitutionsregel. Sei F eine Stammfunktion für f und g differenzierbar. Nach der Kettenregel ist dann

$$\begin{aligned} \frac{d}{dx} F(g(x)) &= F'(g(x)) \cdot g'(x) \\ &= f(g(x))g'(x). \end{aligned}$$

Dies bedeutet für das unbestimmte Integral

$$\int f(g(x))g'(x)dx = \int f(t)dt \quad \text{mit } t = g(x)$$

und für das bestimmte Integral

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt$$

Die Substitutionsregel wird in zwei Versionen angewandt:

1. VERSION DER SUBSTITUTIONSREGEL: Berechnung von $\int f(g(x))g'(x)dx$.

1. SCHRITT: Substitution

$$g(x) = t \text{ und } g'(x)dx = dt.$$

2. SCHRITT: Berechnung der Stammfunktion

$$\int f(t)dt = F(t) + c.$$

3. SCHRITT: Rücksubstitution

$$t = g(x) \text{ und } F(t) = F(g(x)).$$

INSGESAMT:

$$\begin{aligned} \int f(g(x))g'(x)dx &= \int f(t)dt \\ &= F(t) + c \\ &= F(g(x)) + c. \end{aligned}$$

BEISPIEL V.2.5. (1) Sei g eine stetige, differenzierbare Funktion ohne Nullstellen. Dann ist

$$\begin{aligned} \int \frac{g'(x)}{g(x)}dx &= \int \frac{1}{t}dt \quad |t = g(x), dt = g'(x)dx \\ &= \ln(|t|) + c \\ &= \ln(|g(x)|) + c. \end{aligned}$$

(2) Mit $t = g(x) = \ln(x)$, $dt = g'(x)dx = \frac{1}{x}dx$ ergibt sich

$$\begin{aligned} \int \frac{[\ln(x)]^2}{x}dx &= \int t^2 dt \\ &= \frac{1}{3}t^3 + c \\ &= \frac{1}{3}[\ln(x)]^3 + c. \end{aligned}$$

(3) Mit $t = g(x) = \sin(x)$, $dt = \cos(x)dx$ erhalten wir

$$\begin{aligned} \int e^{\sin(x)} \cos(x)dx &= \int e^t dt \\ &= e^t + c \\ &= e^{\sin(x)} + c. \end{aligned}$$

BEISPIEL V.2.6. Für $k \in \mathbb{Z}^*$ und $\varphi \in \mathbb{R}$ ergibt sich

$$\begin{aligned} \int \sin(kx + \varphi)dx &= \frac{1}{k} \int \sin(kx + \varphi)kdx \\ &= \frac{1}{k} \int \sin(t)dt \quad |t = kx + \varphi \\ &= -\frac{1}{k} \cos(t) + c \\ &= -\frac{1}{k} \cos(kx + \varphi) + c. \end{aligned}$$

Also ist

$$\int_0^{2\pi} \sin(kx + \varphi) dx = 0 \quad \text{für } k \in \mathbb{Z}^*, \varphi \in \mathbb{R}$$

mit den beiden Spezialfällen ($\varphi = 0$ bzw. $\varphi = -\pi$)

$$\int_0^{2\pi} \sin(kx) dx = 0, \quad \int_0^{2\pi} \cos(kx) dx = 0 \quad \text{für } k \in \mathbb{Z}^*.$$

Nach den Additionstheoremen, Beispiel I.6.3 (S. 44), ist

$$\begin{aligned} \sin(mx) \sin(nx) &= \frac{1}{2} [\cos((m-n)x) - \cos((m+n)x)], \\ \cos(mx) \cos(nx) &= \frac{1}{2} [\cos((m+n)x) + \cos((m-n)x)], \\ \sin(mx) \cos(nx) &= \frac{1}{2} [\sin((m+n)x) + \sin((m-n)x)]. \end{aligned}$$

Hieraus ergeben sich folgende ORTHOGONALITÄTSBEZIEHUNGEN für die Sinus- und Cosinusfunktionen:

$$\begin{aligned} \int_0^{2\pi} \sin(mx) \sin(nx) dx &= \begin{cases} 0 & \text{falls } m = n = 0, \\ 0 & \text{falls } m \neq n, \\ \pi & \text{falls } m = n \neq 0, \end{cases} \\ \int_0^{2\pi} \cos(mx) \cos(nx) dx &= \begin{cases} 2\pi & \text{falls } m = n = 0, \\ 0 & \text{falls } m \neq n, \\ \pi & \text{falls } m = n \neq 0, \end{cases} \\ \int_0^{2\pi} \sin(mx) \cos(nx) dx &= 0 \quad \text{für alle } m, n. \end{aligned}$$

2. VERSION DER SUBSTITUTIONSREGEL: Berechnung von $\int f(x) dx$.

1. SCHRITT: Substitution

$$x = g(t) \text{ und } dx = g'(t) dt$$

mit einer geeigneten, umkehrbaren Funktion g .

2. SCHRITT: Berechnung der Stammfunktion

$$\int f(g(t)) g'(t) dt = H(t) + c.$$

3. SCHRITT: Auflösen von $x = g(t)$ nach t , d.h. $t = h(x)$ und einsetzen.

INSGESAMT:

$$\int f(x) dx = H(h(x)) + c.$$

BEISPIEL V.2.7. Wir wollen

$$F(x) = \int \frac{e^{3x}}{e^{2x} - 1} dx$$

berechnen

1. SCHRITT: Die Substitution $x = \ln(t)$, $dx = \frac{1}{t} dt$ führt auf

$$\begin{aligned} F(\ln(t)) &= \int \frac{t^3}{t^2 - 1} \frac{1}{t} dt \\ &= \int \frac{t^2}{t^2 - 1} dt. \end{aligned}$$

2. SCHRITT: Es ist gemäß Tabelle V.1.1 (S. 169)

$$\begin{aligned} \int \frac{t^2}{t^2 - 1} dt &= \int \frac{t^2 - 1 + 1}{t^2 - 1} dt \\ &= \int 1 + \frac{1}{t^2 - 1} dt \\ &= t + \int \frac{1}{t^2 - 1} dt \\ &= t + \frac{1}{2} \ln \left| \frac{t - 1}{t + 1} \right| + c. \end{aligned}$$

3. SCHRITT: Die Rücksubstitution $t = e^x$ ergibt

$$F(x) = e^x + \frac{1}{2} \ln \left| \frac{e^x - 1}{e^x + 1} \right| + c.$$

BEMERKUNG V.2.8. Sind p und q zwei stetige Funktionen, so führt die Substitution $x = \ln t$, $dx = \frac{1}{t} dt$ auf

$$\int \frac{p(e^x)}{q(e^x)} dx = \int \frac{p(t)}{tq(t)} dt.$$

In Beispiel V.2.7 ist $p(t) = t^3$ und $q(t) = t^2 - 1$.

BEISPIEL V.2.9. Wir wollen

$$F(x) = \int \sqrt{1 - x^2} dx$$

auf $(-1, 1)$ berechnen.

1. SCHRITT: Die Substitution $x = \sin(t)$, $dx = \cos(t) dt$ führt auf

$$\begin{aligned} F(\sin(t)) &= \int \sqrt{1 - \sin^2(t)} \cos(t) dt \\ &= \int \cos^2(t) dt \end{aligned}$$

auf $(-\frac{\pi}{2}, \frac{\pi}{2})$.

2. SCHRITT: Gemäß Beispiel V.2.2 (4) ist

$$\int \cos^2(t) dt = \frac{1}{2} t + \frac{1}{2} \cos(t) \sin(t) + c.$$

3. SCHRITT: Die Rücksubstitution $t = \arcsin(x)$ ergibt wegen $\cos(t) = \sqrt{1 - \sin^2(t)}$ auf $(-\frac{\pi}{2}, \frac{\pi}{2})$

$$F(x) = \frac{1}{2} \arcsin(x) + \frac{1}{2} x \sqrt{1 - x^2} + c.$$

BEMERKUNG V.2.10. Bei der Berechnung eines bestimmten Integrals sind im ersten Schritt auch die Integrationsgrenzen zu substituieren. Setzt man im zweiten Schritt die neuen Grenzen ein, entfällt dann der dritte Schritt der Rücksubstitution.

BEISPIEL V.2.11. Wir wollen

$$I = \int_0^1 \frac{1}{(1+x^2)^2} dx$$

berechnen.

Die Substitution $x = \tan(t)$, $dx = \frac{1}{\cos^2(t)} dt$ und die neuen Integrationsgrenzen 0 und $\frac{\pi}{4}$ (wegen $\tan(0) = 0$ und $\tan(\frac{\pi}{4}) = 1$) liefern

$$\begin{aligned} I &= \int_0^{\frac{\pi}{4}} \frac{1}{\left(1 + \frac{\sin^2(t)}{\cos^2(t)}\right)^2} \cdot \frac{1}{\cos^2(t)} dt \\ &= \int_0^{\frac{\pi}{4}} \frac{1}{\left(\frac{\sin^2(t) + \cos^2(t)}{\cos^2(t)}\right)^2} \cdot \frac{1}{\cos^2(t)} dt \\ &= \int_0^{\frac{\pi}{4}} \cos^2(t) dt \\ &= \left(\frac{1}{2}t + \frac{1}{2}\sin(t)\cos(t)\right) \Big|_0^{\frac{\pi}{4}} \quad \text{wegen Beispiel V.2.2 (4)} \\ &= \frac{\pi}{8} + \frac{1}{2} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \\ &= \frac{\pi}{8} + \frac{1}{4}. \end{aligned}$$

V.2.4. Symmetrien. Ist f eine GERADE Funktion, d.h. $f(-x) = f(x)$ für alle x , so ist für alle $a > 0$

$$\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx.$$

Ist g eine UNGERADE Funktion, d.h. $g(-x) = -g(x)$ für alle x , so ist für alle $a > 0$

$$\int_{-a}^a g(x) dx = 0.$$

BEISPIEL V.2.12. Es ist

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(t) dt = 2 \int_0^{\frac{\pi}{2}} \cos(t) dt$$

$$= 2,$$

$$\int_{-5}^5 \sinh(x) dx = 0.$$

V.3. Integration rationaler Funktionen

V.3.1. Partialbruchzerlegung. Wir betrachten rationale Funktionen der Form $\frac{p(x)}{q(x)}$. Dabei sind p und q zwei Polynome mit Grad $p < \text{Grad } q$, die keine gemeinsamen Nullstellen besitzen sollen. Die Partialbruchzerlegung dieser Funktionen erfolgt in den folgenden drei Schritten:

- 1. SCHRITT: Man erstellt eine Produktdarstellung des Nennerpolynoms der Form

$$q(x) = c(x - b_1)^{k_1} \cdot \dots \cdot (x - b_r)^{k_r} \cdot q_1(x)^{\ell_1} \cdot \dots \cdot q_s(x)^{\ell_s}.$$

Dabei sind b_1, \dots, b_r die paarweise verschiedenen reellen Nullstellen von q , k_1, \dots, k_r ihre Vielfachheiten und q_1, \dots, q_s paarweise verschiedene quadratische Polynome *ohne* reelle Nullstellen. (Es ist $q_i(x) = x^2 - 2 \operatorname{Re} \lambda_i x + |\lambda_i|^2$, wobei λ_i eine komplexe Nullstelle der Vielfachheit ℓ_i von q ist.)

- 2. SCHRITT: Die zu jedem Linearfaktor $x - b$ von q , $b \in \{b_1, \dots, b_r\}$, der Vielfachheit k und zu jedem quadratischen Faktor $Q \in \{q_1, \dots, q_s\}$ der Vielfachheit ℓ gebildeten Funktionen der Form

$$\frac{A_1}{x - b}, \quad \frac{A_2}{(x - b)^2}, \quad \dots, \quad \frac{A_k}{(x - b)^k}$$

$$\frac{B_1 x + C_1}{Q(x)}, \quad \frac{B_2 x + C_2}{Q(x)^2}, \quad \dots, \quad \frac{B_\ell x + C_\ell}{Q(x)^\ell}$$

mit reellen Koeffizienten A_i, B_j, C_j nennt man PARTIALBRÜCHE von $\frac{p}{q}$. Man stellt nun $\frac{p}{q}$ als Summe all dieser Partialbrüche mit unbekanntem Koeffizienten dar.

- 3. SCHRITT: Man multipliziert die Ansatzgleichung für $\frac{p}{q}$ mit q und erhält auf beiden Seiten ein Polynom. Nun ergeben sich Bestimmungsgleichungen für die Koeffizienten A_i, B_j, C_j des Partialbruchansatzes durch Gleichsetzen gleicher x -Potenzen auf der linken und rechten Seite oder durch Einsetzen spezieller Werte für x wie z.B. b_1, \dots, b_r .

BEMERKUNG V.3.1. Für die praktische Rechnung empfiehlt es sich, die beiden Vorgehensweisen in Schritt 3 (Koeffizientenvergleich und Einsetzen spezieller x -Werte) zu kombinieren. Ist z.B. $x - b$ ein Linearfaktor der Vielfachheit k von q , erhält man den Koeffizienten A_k des Partialbruches $\frac{A_k}{(x-b)^k}$ durch Einsetzen des Wertes $x = b$ in Schritt 3.

BEISPIEL V.3.2. Betrachte

$$\frac{x^2 + x + 1}{x^4 - 5x^3 + 9x^2 - 7x + 2}.$$

Durch Probieren finden wir die Nullstellen 1 und 2 des Nennerpolynoms q . Wir können daher q durch

$$(x - 1)(x - 2) = x^2 - 3x + 1$$

dividieren und erhalten

$$\begin{array}{r} x^4 - 5x^3 + 9x^2 - 7x + 2 : x^2 - 3x + 2 = x^2 - 2x + 1. \\ \underline{x^4 - 3x^3 + 2x^2} \\ -2x^3 + 7x^2 - 7x \\ \underline{-2x^3 + 6x^2 - 4x} \\ x^2 - 3x + 2 \\ \underline{x^2 - 3x + 2} \\ 0 \end{array}$$

Da offensichtlich

$$x^2 - 2x + 1 = (x - 1)^2$$

ist, gilt

$$\begin{aligned} q(x) &= x^4 - 5x^3 + 9x^2 - 7x + 2 \\ &= (x - 2)(x - 1)^3. \end{aligned}$$

Wie man leicht nachprüft sind 1 und 2 keine Nullstellen des Zählerpolynoms. Der Partialbruchansatz lautet also

$$\begin{aligned} \frac{x^2 + x + 1}{x^4 - 5x^3 + 9x^2 - 7x + 2} &= \frac{x^2 + x + 1}{(x - 2)(x - 1)^3} \\ &= \frac{A}{x - 2} + \frac{B}{x - 1} + \frac{C}{(x - 1)^2} + \frac{D}{(x - 1)^3}. \end{aligned}$$

Multiplikation beider Seiten mit q liefert

$$\begin{aligned} x^2 + x + 1 &= A(x - 1)^3 + B(x - 2)(x - 1)^2 \\ &\quad + C(x - 2)(x - 1) + D(x - 2). \end{aligned}$$

Einsetzen von $x = 1$ ergibt

$$-D = 3;$$

Einsetzen von $x = 2$ liefert

$$A = 7.$$

Wir setzen diese beiden Werte in obige Gleichung ein und erhalten nach Ausmultiplizieren

$$\begin{aligned} x^2 + x + 1 &= \underbrace{7}_{=A} (x - 1)^3 + B(x - 2)(x - 1)^2 \\ &\quad + C(x - 2)(x - 1) \underbrace{-3}_{=D} (x - 2) \end{aligned}$$

$$\begin{aligned}
 &= 7(x^3 - 3x^2 + 3x - 1) + B \underbrace{(x - 2)(x^2 - 2x + 1)}_{=x^3 - 4x^2 + 5x - 2} \\
 &\quad + C(x^2 - 3x + 2) - 3x + 6 \\
 &= (7 + B)x^3 + (-21 - 4B + C)x^2 \\
 &\quad + (21 + 5B - 3C - 3)x + (-7 - 2B + 2C + 6).
 \end{aligned}$$

Da auf der linken Seite kein x^3 -Term auftritt, muss

$$B = -7$$

sein. Setzen wir dies in den x^2 -Term auf der rechten Seite ein und vergleichen mit dem entsprechenden Term auf der linken Seite, erhalten wir

$$\begin{aligned}
 1 &= -21 - 4(-7) + C \\
 &= 7 + C \\
 \implies C &= -6.
 \end{aligned}$$

Damit lautet die Partialbruchzerlegung unserer Funktion

$$\frac{x^2 + x + 1}{x^4 - 5x^3 + 9x^2 - 7x + 2} = \frac{7}{x - 2} - \frac{7}{x - 1} - \frac{6}{(x - 1)^2} - \frac{3}{(x - 1)^3}.$$

BEISPIEL V.3.3. Betrachte

$$\frac{x + 1}{x^8 - 2x^7 + 4x^6 - 6x^5 + 6x^4 - 6x^3 + 4x^2 - 2x + 1}.$$

Durch Probieren finden wir, dass das Nennerpolynom und seine Ableitung eine Nullstelle bei $x = 1$ haben. Also können wir es durch

$$(x - 1)^2 = x^2 - 2x + 1$$

dividieren und erhalten

$$\begin{array}{r}
 x^8 \quad -2x^7 \quad +4x^6 \quad -6x^5 \quad +6x^4 \quad -6x^3 \quad +4x^2 \quad -2x \quad +1 \quad : \\
 \underline{x^8 \quad -2x^7 \quad +x^6} \\
 \quad -6x^5 \quad +6x^4 \quad = \\
 \quad \underline{3x^6 \quad -6x^5 \quad +6x^4} \quad +1. \\
 \quad \underline{3x^6 \quad -6x^5 \quad +3x^4} \\
 \quad \quad -6x^3 \quad +4x^2 \\
 \quad \underline{3x^4 \quad -6x^3 \quad +3x^2} \\
 \quad \quad -2x \quad +1 \\
 \quad \underline{x^2 \quad -2x \quad +1} \\
 \quad 0
 \end{array}$$

Offensichtlich ist

$$x^6 + 3x^4 + 3x^2 + 1 = (x^2 + 1)^3.$$

Da $x^2 + 1$ keine reellen Nullstellen hat, lautet der Partialbruchansatz

$$\frac{x + 1}{x^8 - 2x^7 + 4x^6 - 6x^5 + 6x^4 - 6x^3 + 4x^2 - 2x + 1}$$

$$= \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{Cx+D}{x^2+1} + \frac{Ex+F}{(x^2+1)^2} + \frac{Gx+H}{(x^2+1)^3}.$$

Wir multiplizieren diese Gleichung mit dem Nennerpolynom und erhalten

$$\begin{aligned} x+1 &= A(x-1)(x^2+1)^3 + B(x^2+1)^3 \\ &\quad + (Cx+D)(x-1)^2(x^2+1)^2 \\ &\quad + (Ex+F)(x-1)^2(x^2+1) \\ &\quad + (Gx+H)(x-1)^2. \end{aligned}$$

Einsetzen von $x = 1$ liefert

$$B = \frac{1}{4}.$$

Wir setzen dies in obige Gleichung ein und schaffen den B -Term auf die linke Seite

$$\begin{aligned} &-\frac{1}{4}x^6 - \frac{3}{4}x^4 - \frac{3}{4}x^2 + x + \frac{3}{4} \\ &= x+1 - \frac{1}{4}(x^2+1)^3 \\ &= A(x-1)(x^2+1)^3 + (Cx+D)(x-1)^2(x^2+1)^2 \\ &\quad + (Ex+F)(x-1)^2(x^2+1) + (Gx+H)(x-1)^2. \end{aligned}$$

Da die linke Seite dieser Gleichung bei $x = 1$ verschwindet, können wir die ganze Gleichung durch $x - 1$ dividieren und erhalten

$$\begin{aligned} &-\frac{1}{4}x^5 - \frac{1}{4}x^4 - x^3 - x^2 - \frac{7}{4}x - \frac{3}{4} \\ &= A(x^2+1)^3 + (Cx+D)(x-1)(x^2+1)^2 \\ &\quad + (Ex+F)(x-1)(x^2+1) + (Gx+H)(x-1). \end{aligned}$$

Einsetzen von $x = 1$ liefert

$$A = -\frac{5}{8}.$$

Wir setzen diesen Wert in obige Gleichung ein und multiplizieren den Rest aus

$$\begin{aligned} &-\frac{1}{4}x^5 - \frac{1}{4}x^4 - x^3 - x^2 - \frac{7}{4}x - \frac{3}{4} \\ &= -\frac{5}{8}(x^6 + 3x^4 + 3x^2 + 1) + (Cx+D) \underbrace{(x-1)(x^2+1)^2}_{=x^5-x^4+2x^3-2x^2+x-1} \\ &\quad + (Ex+F)(x^3-x^2+x-1) + (Gx+H)(x-1) \\ &= x^6\left(-\frac{5}{8}+C\right) + x^5(-C+D) + x^4\left(-\frac{15}{8}+2C-D+E\right) \\ &\quad + x^3(-2C+2D-E+F) + x^2\left(-\frac{15}{8}+C-2D+E-F+G\right) \end{aligned}$$

$$+ x(-C + D - E + F - G + H) - \frac{5}{8} - D - F - H.$$

Koeffizientenvergleich liefert die Bedingungen

$$\begin{aligned} x^6 : & & C &= \frac{5}{8} \\ x^5 : & & -C + D &= -\frac{1}{4} \implies D = \frac{3}{8} \\ x^4 : & & 2C - D + E &= -\frac{1}{4} + \frac{15}{8} \implies E = \frac{6}{8} \\ x^3 : & & -2C + 2D - E + F &= -1 \implies F = \frac{2}{8} \\ x^2 : & & C - 2D + E - F + G &= -1 + \frac{15}{8} \implies G = \frac{4}{8} \\ x : & & -C + D - E + F - G + H &= -\frac{7}{4} \implies H = -\frac{4}{8}. \end{aligned}$$

Also lautet die Partialbruchzerlegung

$$\begin{aligned} & \frac{x+1}{x^8 - 2x^7 + 4x^6 - 6x^5 + 6x^4 - 6x^3 + 4x^2 - 2x + 1} \\ &= \frac{1}{8} \left(-\frac{5}{x-1} + \frac{2}{(x-1)^2} + \frac{5x+3}{x^2+1} + \frac{6x+2}{(x^2+1)^2} + \frac{4x-4}{(x^2+1)^3} \right). \end{aligned}$$

V.3.2. Integration. Die Integration einer rationalen Funktion $f(x) = \frac{p(x)}{q(x)}$ erfolgt in drei Schritten:

- 1. SCHRITT: Falls der Zählergrad größer oder gleich dem Nennergrad ist, dividiere p mit Rest durch q . Dies liefert zwei Polynome r und s mit $p = s \cdot q + r$. Daher ist $f = s + \frac{r}{q}$.
- 2. SCHRITT: Bestimme die Partialbruchzerlegung von $\frac{r}{q}$ wie im vorigen Abschnitt.
- 3. SCHRITT: Integriere das Polynom s und die Partialbrüche von $\frac{r}{q}$.

Für den 3. Schritt sind folgende Formeln, die man durch Differentiation nachprüft, hilfreich. Dabei ist $\alpha^2 - 4\beta < 0$ und $k > 1$.

$$\begin{aligned} \int \frac{1}{x-a} dx &= \ln|x-a| + c \\ \int \frac{1}{(x-a)^k} dx &= -\frac{1}{k-1} \frac{1}{(x-a)^{k-1}} + c \\ \int \frac{1}{x^2 + \alpha x + \beta} dx &= \frac{2}{\sqrt{4\beta - \alpha^2}} \arctan \left(\frac{2x + \alpha}{\sqrt{4\beta - \alpha^2}} \right) + c \\ \int \frac{ax + b}{x^2 + \alpha x + \beta} dx &= \frac{a}{2} \ln|x^2 + \alpha x + \beta| \end{aligned}$$

$$\begin{aligned}
& + \left(b - \frac{a\alpha}{2}\right) \int \frac{1}{x^2 + \alpha x + \beta} dx \\
\int \frac{1}{(x^2 + \alpha x + \beta)^k} dx &= \frac{2x + \alpha}{(k-1)(4\beta - \alpha^2)(x^2 + \alpha x + \beta)^{k-1}} \\
& + \frac{2(2k-3)}{(k-1)(4\beta - \alpha^2)} \int \frac{1}{(x^2 + \alpha x + \beta)^{k-1}} dx \\
\int \frac{ax + b}{(x^2 + \alpha x + \beta)^k} dx &= -\frac{a}{2} \frac{1}{k-1} \frac{1}{(x^2 + \alpha x + \beta)^{k-1}} \\
& + \left(b - \frac{a\alpha}{2}\right) \int \frac{1}{(x^2 + \alpha x + \beta)^k} dx.
\end{aligned}$$

BEISPIEL V.3.4. Betrachte

$$f(x) = \frac{2x^5 - 7x^4 + 3x^3 + 14x^2 - 16x + 7}{x^4 - 5x^3 + 9x^2 - 7x + 2}.$$

Die Division des Zählerpolynoms durch das Nennerpolynom ergibt

$$\begin{array}{r}
2x^5 \quad -7x^4 \quad +3x^3 \quad +14x^2 \quad -16x \quad +7 \quad : \quad x^4 - 5x^3 + 9x^2 - 7x + 2 \\
2x^5 \quad -10x^4 \quad +18x^3 \quad -14x^2 \quad +4x \quad \quad \quad = 2x + 3. \\
\hline
\quad \quad 3x^4 \quad -15x^3 \quad +28x^2 \quad -20x \quad +7 \\
\quad \quad 3x^4 \quad -15x^3 \quad +27x^2 \quad -21x \quad +6 \\
\hline
\quad \quad \quad \quad \quad x^2 \quad \quad +x \quad \quad +1
\end{array}$$

Also ist

$$f(x) = 2x + 3 + \frac{x^2 + x + 1}{x^4 - 5x^3 + 9x^2 - 7x + 2}.$$

Die Partialbruchzerlegung des rationalen Teils haben wir in Beispiel V.3.2 bestimmt, und wir erhalten

$$\begin{aligned}
\int f(x) dx &= \int \left(2x + 3 + \frac{7}{x-2} - \frac{7}{x-1} \right. \\
&\quad \left. - \frac{6}{(x-1)^2} - \frac{3}{(x-1)^3} \right) dx \\
&= x^2 + 3x + 7 \ln|x-2| - 7 \ln|x-1| \\
&\quad + \frac{6}{x-1} + \frac{3}{2(x-1)^2} + c \\
&= x^2 + 3x + 7 \ln \left| \frac{x-2}{x-1} \right| + \frac{12x-9}{2(x-1)^2} + c.
\end{aligned}$$

BEISPIEL V.3.5. Wir betrachten die Funktion aus Beispiel V.3.3. Zuerst behandeln wir die beiden Partialbrüche

$$\frac{6x+2}{(x^2+1)^2} \quad \text{und} \quad \frac{4x-4}{(x^2+1)^3}.$$

Für diese erhalten wir

$$\begin{aligned} \int \frac{6x+2}{(x^2+1)^2} dx &= 3 \int \frac{2x}{(x^2+1)^2} dx + 2 \int \frac{1}{(x^2+1)^2} dx \\ &= -\frac{3}{x^2+1} + \frac{4x}{4(x^2+1)} + \frac{4}{4} \int \frac{1}{(x^2+1)} dx \\ &= -\frac{3}{x^2+1} + \frac{x}{x^2+1} + \arctan(x) + c \\ &= \frac{x-3}{x^2+1} + \arctan(x) + c \end{aligned}$$

und

$$\begin{aligned} \int \frac{4x-4}{(x^2+1)^3} dx &= 2 \int \frac{2x}{(x^2+1)^3} dx - 4 \int \frac{1}{(x^2+1)^3} dx \\ &= -\frac{2}{2(x^2+1)^2} - \frac{8x}{2 \cdot 4 \cdot (x^2+1)^2} \\ &\quad - \frac{8 \cdot 3}{2 \cdot 4} \int \frac{1}{(x^2+1)^2} dx \\ &= -\frac{x+1}{(x^2+1)^2} - \frac{3 \cdot 2x}{4(x^2+1)} - \frac{3 \cdot 2}{4} \int \frac{1}{(x^2+1)} dx \\ &= -\frac{x+1}{(x^2+1)^2} - \frac{3x}{2(x^2+1)} - \frac{3}{2} \arctan(x) + c. \end{aligned}$$

Damit folgt insgesamt

$$\begin{aligned} &\int \frac{x+1}{x^8 - 2x^7 + 4x^6 - 6x^5 + 6x^4 - 6x^3 + 4x^2 - 2x + 1} dx \\ &= \int \frac{1}{8} \left(-\frac{5}{x-1} + \frac{2}{(x-1)^2} + \frac{5x+3}{x^2+1} + \frac{6x+2}{(x^2+1)^2} + \frac{4x-4}{(x^2+1)^3} \right) dx \\ &= \frac{1}{8} \left(-5 \ln|x-1| - \frac{2}{x-1} + \frac{5}{2} \ln(x^2+1) + 3 \int \frac{1}{x^2+1} dx \right. \\ &\quad \left. + \frac{x-3}{x^2+1} + \arctan(x) \right. \\ &\quad \left. - \frac{x+1}{(x^2+1)^2} - \frac{3x}{2(x^2+1)} - \frac{3}{2} \arctan(x) \right) + c \\ &= -\frac{5}{8} \ln|x-1| - \frac{1}{4(x-1)} + \frac{5}{16} \ln(x^2+1) + \frac{5}{8} \arctan(x) \\ &\quad - \frac{x+6}{16(x^2+1)} - \frac{x+1}{8(x^2+1)^2} + c. \end{aligned}$$

V.3.3. Verallgemeinerte rationale Funktionen. Ein Integral der Form $\int r(e^{\alpha x}) dx$ mit einer rationalen Funktion r kann durch die

Substitution

$$\begin{aligned}t &= e^{\alpha x}, \\x &= \frac{1}{\alpha} \ln t, \\dx &= \frac{1}{\alpha t} dt\end{aligned}$$

auf ein Integral der Form $\int r(t) \frac{1}{\alpha t} dt$ mit rationalem Integrandem transformiert werden.

BEISPIEL V.3.6. Die Substitution $t = e^x$ liefert

$$\begin{aligned}\int \frac{1}{e^x + e^{-x}} dx &= \int \frac{1}{t + \frac{1}{t}} \frac{1}{t} dt \\&= \int \frac{1}{t^2 + 1} dt \\&= \arctan(t) + c \\&= \arctan(e^x) + c.\end{aligned}$$

Im Folgenden bezeichnen wir mit $r(x, y)$ einen Ausdruck, der aus x , y und Konstanten allein durch Addition, Subtraktion, Multiplikation und Division gebildet wird, z.B.

$$r(x, y) = \frac{4xy + y^2}{3x^3 - 2y}.$$

Viele Integranden sind dann von der Form $r(x, f(x))$, d.h. die Variable y in $r(x, y)$ ist durch eine Funktion $f(x)$ von x ersetzt. Solche Integrale können häufig durch eine geeignete Substitution auf das einer rationalen Funktion transformiert werden. Wir betrachten hierzu im Folgenden einige Beispiele.

Bei einem Integral der Form

$$\int r \left(x, \sqrt[k]{\frac{ax + b}{cx + e}} \right) dx$$

mit $ae - bc \neq 0$ hilft die Substitution

$$\begin{aligned}t &= \sqrt[k]{\frac{ax + b}{cx + e}}, \\x &= \frac{et^k - b}{a - ct^k}, \\dx &= k(ae - bc) \frac{t^{k-1}}{(a - ct^k)^2} dt\end{aligned}$$

weiter.

BEISPIEL V.3.7. Es ist

$$\begin{aligned} \int \frac{1 - \sqrt{x}}{x + \sqrt{x}} dx &= \int \frac{1 - t}{t^2 + t} 2t dt \quad \left| \begin{array}{l} t = \sqrt{x}, \\ x = t^2, \\ dx = 2t dt \end{array} \right. \\ &= 2 \int \frac{1 - t}{1 + t} dt \\ &= 2 \int \left(-1 + \frac{2}{1 + t} \right) dt \\ &= 4 \ln |1 + t| - 2t + c \\ &= 4 \ln |1 + \sqrt{x}| - 2\sqrt{x} + c. \end{aligned}$$

Bei einem Integral der Form

$$\int r(\sin(x), \cos(x)) dx$$

hilft die Substitution

$$\begin{aligned} x &= 2 \arctan(t), \\ dx &= \frac{2}{1 + t^2} dt \end{aligned}$$

mit

$$\begin{aligned} \sin(x) &= \frac{2t}{1 + t^2}, \\ \cos(x) &= \frac{1 - t^2}{1 + t^2} \end{aligned}$$

weiter.

BEISPIEL V.3.8. Es ist

$$\begin{aligned} \int \frac{1}{\cos(x)} dx &= \int \frac{1 + t^2}{1 - t^2} \frac{2}{1 + t^2} dt \quad \left| \begin{array}{l} x = 2 \arctan(t) \end{array} \right. \\ &= \int \frac{2}{1 - t^2} dt \\ &= \ln \left| \frac{1 + t}{1 - t} \right| + c \\ &= \ln \left| \frac{1 + \arctan(\frac{x}{2})}{1 - \arctan(\frac{x}{2})} \right| + c. \end{aligned}$$

Integrale der Form

$$\int r(x, \sqrt{ax^2 + bx + c}) dx$$

mit $a \neq 0$ lassen sich mit einer Substitution

$$u = \alpha x + \beta,$$

$$dx = \frac{1}{\alpha} du$$

in eine der folgenden Normalformen umwandeln:

$$\int r(u, \sqrt{u^2 + 1}) du,$$

$$\int r(u, \sqrt{u^2 - 1}) du,$$

$$\int r(u, \sqrt{1 - u^2}) du.$$

Diese können durch die Substitutionen

$$u = \sinh(t)$$

bzw.

$$u = \cosh(t)$$

bzw.

$$u = \sin(t)$$

auf eine der bereits behandelten Formen transformiert werden.

BEISPIEL V.3.9. Es ist

$$\begin{aligned} \int \sqrt{4x^2 + 12x + 5} dx &= \int \sqrt{(2x + 3)^2 - 4} dx \\ &= 2 \int \sqrt{\left(\frac{2x + 3}{2}\right)^2 - 1} dx \quad \left| u = \frac{2x + 3}{2} \right. \\ &= 2 \int \sqrt{u^2 - 1} du \quad \left| u = \cosh(t) \right. \\ &= 2 \int \underbrace{\sqrt{\cosh^2(t) - 1}}_{=\sinh(t)} \sinh(t) dt \\ &= 2 \int \sinh^2(t) dt \\ &= \frac{1}{2} \int (e^t - e^{-t})^2 dt \quad \left| z = e^t \right. \\ &= \frac{1}{2} \int \left(z - \frac{1}{z}\right)^2 \frac{1}{z} dz \\ &= \frac{1}{2} \int \left(z - \frac{2}{z} + \frac{1}{z^3}\right) dz \\ &= \frac{1}{4} z^2 - \ln |z| - \frac{1}{4z^2} + c \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4}e^{2t} - t - \frac{1}{4}e^{-2t} + c \\
&= \frac{1}{4}(e^t + e^{-t})(e^t - e^{-t}) - t + c \\
&= \cosh(t) \sinh(t) - t + c \\
&= \cosh(t) \sqrt{\cosh^2(t) - 1} - t + c \\
&\quad \left| \begin{array}{l} e^t = u + \sqrt{u^2 - 1}, \quad u = \cosh(t) \end{array} \right. \\
&= u\sqrt{u^2 - 1} - \ln \left| u + \sqrt{u^2 - 1} \right| + c \\
&= \left(\frac{2x+3}{2} \right) \sqrt{\left(\frac{2x+3}{2} \right)^2 - 1} \\
&\quad - \ln \left| \frac{2x+3}{2} + \sqrt{\left(\frac{2x+3}{2} \right)^2 - 1} \right| + c.
\end{aligned}$$

V.4. Uneigentliche Integrale

V.4.1. Definition. Wir wollen den Integralbegriff auf nicht beschränkte Integrationsbereiche und unbeschränkte Funktionen ausdehnen. Sei dazu $b \in \mathbb{R} \cup \{\infty\}$, $a \in \mathbb{R}$ und $f : [a, b) \rightarrow \mathbb{R}$ auf jedem Teilintervall $[a, c] \subset [a, b)$ stückweise stetig und beschränkt. Die Funktion f heißt auf $[a, b)$ (UNEIGENTLICH) INTEGRIERBAR, wenn der Grenzwert $\lim_{c \rightarrow b^-} \int_a^c f(x) dx$ existiert. In diesem Fall heißt dieser Grenzwert das UNEIGENTLICHE INTEGRAL von f und wird mit $\int_a^b f(x) dx$ bezeichnet:

$$\int_a^b f(x) dx = \lim_{c \rightarrow b^-} \int_a^c f(x) dx.$$

Analog geht man bei Singularitäten am linken Rand vor. Sei dazu $\alpha \in \mathbb{R} \cup \{-\infty\}$, $\beta \in \mathbb{R}$ und $f : (\alpha, \beta] \rightarrow \mathbb{R}$. Dann ist das uneigentliche Integral definiert durch

$$\int_\alpha^\beta f(x) dx = \lim_{c \rightarrow \alpha^+} \int_c^\beta f(x) dx,$$

sofern der Grenzwert existiert.

BEISPIEL V.4.1. Sei $\alpha \in \mathbb{R} \setminus \{1\}$. Für $c > 1$ ist

$$\begin{aligned} \int_1^c \frac{1}{x^\alpha} dx &= -\frac{1}{\alpha-1} \frac{1}{x^{\alpha-1}} \Big|_1^c \\ &= \frac{1}{\alpha-1} \left[1 - \frac{1}{c^{\alpha-1}} \right]. \end{aligned}$$

Dieser Ausdruck konvergiert für $c \rightarrow \infty$, falls $\alpha > 1$ ist. Also ist $x^{-\alpha}$ für $\alpha > 1$ auf $[1, \infty)$ uneigentlich integrierbar und

$$\int_1^\infty \frac{1}{x^\alpha} dx = \frac{1}{\alpha-1} \quad \text{für } \alpha > 1.$$

Analog ist für $0 < c < 1$

$$\int_c^1 \frac{1}{x^\alpha} dx = \frac{1}{\alpha-1} \left[\frac{1}{c^{\alpha-1}} - 1 \right].$$

Dieser Ausdruck konvergiert für $c \rightarrow 0$, falls $\alpha < 1$ ist. Also ist $x^{-\alpha}$ für $\alpha < 1$ auf $(0, 1]$ uneigentlich integrierbar und

$$\int_0^1 \frac{1}{x^\alpha} dx = \frac{1}{1-\alpha} \quad \text{für } \alpha < 1.$$

Sei nun $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ und $f : (a, b) \rightarrow \mathbb{R}$. f ist auf (a, b) uneigentlich integrierbar, wenn für ein $c \in (a, b)$ die beiden uneigentlichen Integrale $\int_a^c f(x) dx$ und $\int_c^b f(x) dx$ existieren. Die Summe dieser beiden uneigentlichen Integrale ist dann das uneigentliche Integral $\int_a^b f(x) dx$:

$$\int_a^b f(x) dx = \lim_{\alpha \rightarrow a^+} \int_\alpha^c f(x) dx + \lim_{\beta \rightarrow b^-} \int_c^\beta f(x) dx.$$

BEMERKUNG V.4.2. Die beiden Grenzwerte in obiger Formel müssen *unabhängig* voneinander bestimmt werden!

BEISPIEL V.4.3. Es ist

$$\begin{aligned} \lim_{\beta \rightarrow \infty} \int_0^\beta \frac{1}{1+x^2} dx &= \lim_{\beta \rightarrow \infty} [\arctan \beta - \arctan 0] \\ &= \frac{\pi}{2} \\ \lim_{\alpha \rightarrow -\infty} \int_\alpha^0 \frac{1}{1+x^2} dx &= \lim_{\alpha \rightarrow -\infty} [\arctan 0 - \arctan \alpha] \end{aligned}$$

$$= \frac{\pi}{2}.$$

Daher ist $\frac{1}{1+x^2}$ auf $\mathbb{R} = (-\infty, \infty)$ uneigentlich integrierbar und

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \pi.$$

V.4.2. Ein Konvergenzkriterium. Die Existenz eines uneigentlichen Integrals kann häufig durch einen Vergleich mit einer Funktion der Form $x^{-\alpha}$ gezeigt werden.

KONVERGENZKRITERIUM FÜR UNEIGENTLICHE INTEGRALE: Es seien f auf $[a, \infty)$ und g auf $(0, b]$ stetig und $\alpha, K \in \mathbb{R}$. Dann gilt:

$$|f(x)| \leq \frac{K}{x^\alpha} \text{ mit } a \leq x < \infty, 1 < \alpha \implies \int_a^\infty f(x) dx \text{ existiert,}$$

$$|g(x)| \leq \frac{K}{x^\alpha} \text{ mit } 0 < x \leq b, 0 < \alpha < 1 \implies \int_0^b g(x) dx \text{ existiert.}$$

BEWEIS. Die Funktion

$$F : c \mapsto \int_a^c |f(x)| dx$$

ist monoton wachsend und durch $\int_a^\infty \frac{K}{x^\alpha} dx$ nach oben beschränkt. Daher ist sie für $c \rightarrow \infty$ konvergent, und $\int_a^\infty |f(x)| dx$ existiert. Analog erhält man die Existenz von $\int_a^\infty f(x) + |f(x)| dx$. Daher existiert $\int_a^\infty f(x) dx$.

Genauso argumentiert man für $\int_0^b g(x) dx$. □

BEISPIEL V.4.4. Existiert $\int_0^\infty \frac{\sin x}{x}$?

Aus der Regel von de l'Hôpital folgt

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\sin x}{x} &= \lim_{x \rightarrow 0} \frac{\cos x}{1} \\ &= 1. \end{aligned}$$

Also ist $\frac{\sin x}{x}$ auf $[0, \infty)$ stetig und auf $[0, 1]$ integrierbar. Für $c > 1$ liefert partielle Integration

$$\begin{aligned} \int_1^c \frac{\sin x}{x} dx &= -\frac{\cos x}{x} \Big|_1^c - \int_1^c \frac{\cos x}{x^2} dx \\ &= \cos(1) - \frac{\cos(c)}{c} - \int_1^c \frac{\cos x}{x^2} dx. \end{aligned}$$

Da die Cosinus-Funktion beschränkt ist, ist

$$\lim_{c \rightarrow \infty} \frac{\cos(c)}{c} = 0.$$

Wegen des Konvergenzkriteriums existiert $\int_1^{\infty} \frac{\cos x}{x^2} dx$. Also existiert

$$\int_0^{\infty} \frac{\sin x}{x}.$$

Mit Hilfe von Sätzen aus der Funktionentheorie kann man

$$\int_0^{\infty} \frac{\sin x}{x} = \frac{\pi}{2}$$

zeigen.

V.4.3. Die Eulersche Gammafunktion. Die EULERSCHE GAMMAFUNKTION ist definiert durch

$$\Gamma : (0, \infty) \rightarrow \mathbb{R}$$

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt.$$

Die Konvergenz des uneigentlichen Integrals wird wie folgt nachgewiesen:

- 1. SCHRITT: Das Integral $\int_0^1 e^{-t} t^{x-1} dt$ ist an der unteren Grenze nur für $0 < x < 1$ eigentlich. In diesem Fall folgt die Existenz des Integrals aus dem Konvergenzkriterium und der Abschätzung $0 \leq e^{-t} t^{x-1} \leq t^{x-1} = t^{-\alpha}$ mit $\alpha = 1 - x < 1$.
- 2. SCHRITT: Wegen

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{t^2 t^{x-1}}{e^t} &= \lim_{t \rightarrow \infty} \frac{t^{x+1}}{e^t} \\ &= 0 \end{aligned}$$

gilt für hinreichend großes t die Abschätzung $e^{-t} t^{x-1} \leq t^{-2}$. Damit folgt die Existenz von $\int_1^{\infty} e^{-t} t^{x-1} dt$ aus dem Konvergenzkriterium.

Wegen

$$\begin{aligned} \int_0^{\infty} e^{-t} dt &= \lim_{c \rightarrow \infty} \int_0^c e^{-t} dt \\ &= \lim_{c \rightarrow \infty} [1 - e^{-c}] \\ &= 1 \end{aligned}$$

ist

$$\Gamma(1) = 1.$$

Mit partieller Integration $u' = e^{-t}$, $v = t^{x-1}$ erhalten wir für $0 < \alpha < \beta$

$$\int_{\alpha}^{\beta} e^{-t} t^{x-1} dt = -e^{-t} t^{x-1} \Big|_{\alpha}^{\beta} + \int_{\alpha}^{\beta} (x-1) e^{-t} t^{x-2} dt.$$

Durch Grenzübergang folgt hieraus die FUNKTIONALGLEICHUNG

$$\Gamma(x+1) = x\Gamma(x) \quad \text{für alle } x > 0.$$

Insbesondere ergibt sich

$$\Gamma(n+1) = n! \quad \text{für alle } n \in \mathbb{N}.$$

V.5. Ebene Kurven; Längen- und Flächenmessung

V.5.1. Parameterdarstellung. Die allgemeinste und für unsere Zwecke günstigste Darstellung einer ebenen Kurve erfolgt mit zwei differenzierbaren Funktionen $x(t)$, $y(t)$ auf einem Intervall $a \leq t \leq b$. In einem festen kartesischen Koordinatensystem durchläuft der sich mit t stetig ändernde Punkt

$$P(t) = (x(t), y(t)) \quad , a \leq t \leq b,$$

eine Kurve. Die vektorwertige Funktion

$$\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

bzw. die beiden Gleichungen

$$\begin{aligned} x &= x(t), \\ y &= y(t) \end{aligned}$$

heißen eine PARAMETERDARSTELLUNG dieser Kurve; t und $[a, b]$ heißen der PARAMETER bzw. das PARAMETERINTERVALL. Man sagt, die Kurve werde mit wachsendem t in positiver Richtung durchlaufen.

Deutet man t als Zeit, so beschreibt eine Parameterdarstellung die Bewegung eines Massenpunktes auf der Kurve. Da auf einer Kurve ganz verschiedene Bewegungen stattfinden können, besitzt ein und dieselbe Kurve verschiedene Parameterdarstellungen.

BEISPIEL V.5.1. Der GRAPH $y = f(x)$, $a \leq x \leq b$, einer Funktion f besitzt die Parameterdarstellung

$$\begin{aligned} x &= t, \\ y &= f(t) \end{aligned}$$

mit $a \leq t \leq b$.

BEISPIEL V.5.2. Die GERADE durch die Punkte (x_1, y_1) und (x_2, y_2) besitzt die Parameterdarstellung

$$\begin{aligned}x &= x_1 + t(x_2 - x_1), \\y &= y_1 + t(y_2 - y_1)\end{aligned}$$

mit $t \in \mathbb{R}$.

BEISPIEL V.5.3. Der KREIS um (x_0, y_0) mit Radius r besitzt die Parameterdarstellung

$$\begin{aligned}x &= x_0 + r \cos(t), \\y &= y_0 + r \sin(t)\end{aligned}$$

mit $0 \leq t \leq 2\pi$.

BEISPIEL V.5.4. Die ELLIPSE

$$\left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}$$

hat die Parameterdarstellung

$$\begin{aligned}x &= a \cos(t), \\y &= b \sin(t)\end{aligned}$$

mit $0 \leq t \leq 2\pi$.

BEISPIEL V.5.5 (ZYKLOIDEN). Rollt ein Kreis K mit Radius r , ohne zu gleiten, auf der x -Achse, dann beschreibt ein mit K fest verbundener Punkt P , der vom Kreismittelpunkt den Abstand a hat, eine ZYKLOIDE oder Radkurve. Diese besitzt die Parameterdarstellung

$$\begin{aligned}x &= rt - a \sin(t), \\y &= r - a \cos(t)\end{aligned}$$

mit $t \in \mathbb{R}$. Dabei beschreibt der Parameter t den Rollwinkel. Rollt der Kreis mit dem Punkt P nicht auf der x -Achse, sondern auf dem Kreis $x^2 + y^2 = \rho^2$ dann ist die Bahn von P eine EPIZYKLOIDE mit der Parameterdarstellung

$$\begin{aligned}x &= (\rho + r) \cos(t) - a \cos\left(\frac{\rho + r}{r}t\right), \\y &= (\rho + r) \sin(t) - a \sin\left(\frac{\rho + r}{r}t\right).\end{aligned}$$

Im Sonderfall $\rho = r = a$ entsteht die KARDOIDE oder Herzlinie

$$\begin{aligned}x &= a(2 \cos(t) - \cos(2t)), \\y &= a(2 \sin(t) - \sin(2t)).\end{aligned}$$

Beim Abrollen im Innern des Kreises entsteht die HYPOZYKLOIDE

$$x = (\rho - r) \cos(t) + a \cos\left(\frac{\rho - r}{r}t\right),$$

$$y = (\rho - r) \sin(t) - a \sin\left(\frac{\rho - r}{r}t\right)$$

Diese geht für den Sonderfall $a = r = \frac{\rho}{4}$ in die ASTROIDE über. Aus den Additionstheoremen ergibt sich dann

$$\begin{aligned}x &= \rho \cos^3(t), \\y &= \rho \sin^3(t).\end{aligned}$$

V.5.2. Tangente und Normale. Zu jeder Parameterdarstellung

$$\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

einer Kurve K definieren wir

$$\begin{aligned}\dot{\mathbf{r}}(t) &= \lim_{h \rightarrow 0} \frac{1}{h} [\mathbf{r}(t+h) - \mathbf{r}(t)] \\ &= \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix},\end{aligned}$$

wobei der Punkt die Ableitung nach t bezeichnet. Als Grenzwert von Sekantenvektoren ist $\dot{\mathbf{r}}(t)$ parallel zur Tangente an K im Punkt $\mathbf{r}(t)$.

Beschreibt $\mathbf{r}(t)$ die Bewegung eines Massenpunktes auf der Kurve, ist $\dot{\mathbf{r}}(t)$ die Geschwindigkeit des Punktes.

Ist in einem Kurvenpunkt $\mathbf{r}(t)$ der Vektor $\dot{\mathbf{r}}(t)$ vom Nullvektor verschieden, dann weist er in die positive Tangentenrichtung (positiv bzgl. des Durchlaufsinnes der Kurve). Der hieraus durch Drehung um 90° im Gegenuhrzeigersinn entstandene Vektor

$$\mathbf{n}(t) = \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix}$$

gibt die positive Normalenrichtung in diesem Punkt an (vgl. Abbildung V.5.1). Daher besitzen Tangente und Normale im Kurvenpunkt $\mathbf{r}(t)$ die Parameterdarstellung

<p>TANGENTE : $x = x(t) + \lambda \dot{x}(t),$ $y = y(t) + \lambda \dot{y}(t)$</p> <p>NORMALE : $x = x(t) - \lambda \dot{y}(t),$ $y = y(t) + \lambda \dot{x}(t).$</p>

Dabei ist λ der Geradenparameter und t der Kurvenparameter.

BEMERKUNG V.5.6. Die Vektoren $\dot{\mathbf{r}}$ und \mathbf{n} sind nicht normiert. Benötigt man normierte Tangenten- bzw. Normalenvektoren, muss man $\dot{\mathbf{r}}$ mit $\frac{1}{|\dot{\mathbf{r}}|}$ und \mathbf{n} mit $\frac{1}{|\mathbf{n}|} = \frac{1}{|\dot{\mathbf{r}}|}$ multiplizieren.

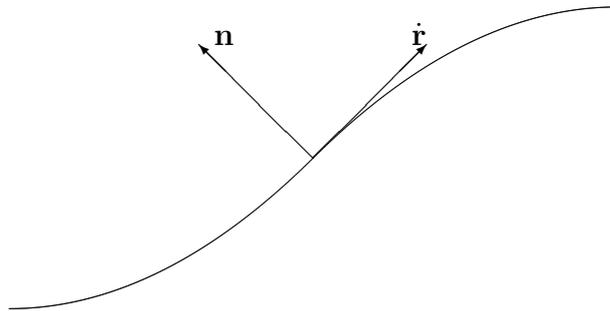


ABBILDUNG V.5.1. Tangente $\dot{\mathbf{r}}$ und Normale \mathbf{n}

V.5.3. Bogenlänge. Die Parameterdarstellung

$$\begin{aligned}x &= x(t), \\y &= y(t)\end{aligned}$$

der Kurve K heißt REGULÄR, wenn $\dot{x}(t)$ und $\dot{y}(t)$ stetig sind und

$$\dot{x}(t)^2 + \dot{y}(t)^2 \neq 0$$

ist für alle t .

Wir wollen der Kurve nun eine Länge zuordnen. Dazu unterteilen wir das Parameterintervall durch $a = t_0 < t_1 < \dots < t_{n-1} < t_n = b$. Dann ist

$$|\mathbf{r}(t_i) - \mathbf{r}(t_{i-1})| = \{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2\}^{\frac{1}{2}}$$

die Länge der Sekante zwischen den Punkten $\mathbf{r}(t_{i-1})$ und $\mathbf{r}(t_i)$. Die Summe

$\sum_{i=1}^n |\mathbf{r}(t_i) - \mathbf{r}(t_{i-1})|$ ist eine Näherung für die Länge der Kurve. Wir lassen nun die „Feinheit“ $\max_{1 \leq i \leq n} |t_i - t_{i-1}|$ der Unterteilung gegen Null

streben. Dann konvergiert obige Summe gegen $\int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt$.

Wir nennen daher dieses Integral die Länge der Kurve K .

Eine Kurve K mit der regulären Parameterdarstellung

$$\begin{aligned}x &= x(t), \\y &= y(t)\end{aligned}$$

mit $a \leq t \leq b$ hat die LÄNGE

$$\int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt.$$

BEMERKUNG V.5.7. Man kann zeigen, dass die so definierte Länge nur von der Kurve und nicht von der speziellen Parameterdarstellung abhängt.

Betrachtet man nur den Kurvenabschnitt, der dem Parameterbereich $[a, t_0]$, $a \leq t_0 \leq b$, entspricht, erhält man analog die Länge des entsprechenden Kurvenstückes. Diese Länge heißt die Bogenlänge und wird mit $s(t_0)$ bezeichnet.

Eine Kurve K mit der regulären Parameterdarstellung

$$x = x(t),$$

$$y = y(t)$$

mit $a \leq t \leq b$ hat die BOGENLÄNGE

$$s(t) = \int_a^t \sqrt{\dot{x}(\tau)^2 + \dot{y}(\tau)^2} d\tau \quad \text{für } a \leq t \leq b.$$

Es ist

$$\dot{s}(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}.$$

BEISPIEL V.5.8. Für den Graphen $y = f(x)$, $a \leq x \leq b$, ergibt sich die Länge

$$\int_a^b \sqrt{1 + f'(x)^2} dx.$$

Speziell ergibt sich für die Parabel $y = x^2$, $0 \leq x \leq b$, die Länge

$$\int_0^b \sqrt{1 + (2x)^2} dx = \frac{1}{4} \left[2b\sqrt{1 + 4b^2} + \ln(2b + \sqrt{1 + 4b^2}) \right].$$

BEISPIEL V.5.9. Für den Zykloidenbogen

$$x = r(t - \sin(t)),$$

$$y = r(1 - \cos(t)),$$

$0 \leq t \leq 2\pi$, ergibt sich die Länge

$$\begin{aligned} & \int_0^{2\pi} \sqrt{r^2(1 - \cos(t))^2 + r^2 \sin^2(t)} dt \\ &= \int_0^{2\pi} \left\{ 2r^2 - 2r^2 \underbrace{\cos(t)}_{=\cos^2(\frac{t}{2}) - \sin^2(\frac{t}{2})} \right\}^{\frac{1}{2}} dt \\ &= \int_0^{2\pi} \sqrt{2r^2(1 - \cos^2(\frac{t}{2}) + \sin^2(\frac{t}{2}))} dt \\ &= \int_0^{2\pi} 2r \sin(\frac{t}{2}) dt \\ &= -4r \cos(\frac{t}{2}) \Big|_0^{2\pi} \\ &= 8r. \end{aligned}$$

V.5.4. Krümmung und Krümmungskreis. Es sei

$$x = x(t),$$

$$y = y(t),$$

$a \leq t \leq b$, eine zweimal differenzierbare, reguläre Parameterdarstellung einer Kurve K . Mit $\varphi(t)$ bezeichnen wir den positiv gemessenen Winkel zwischen der positiven x -Achse und der Richtung des Tangentenvektors $\dot{\mathbf{r}}(t)$. Weiter ist

$$s(t) = \int_a^t \sqrt{\dot{x}(\tau)^2 + \dot{y}(\tau)^2} d\tau$$

die Bogenlänge. Die Änderung $\Delta\varphi$ der Tangentenrichtung über dem Intervall $[t, t + \Delta t]$ bezogen auf die Änderung der Bogenlänge, also $\frac{\Delta\varphi}{\Delta s}$, ist offenbar ein gutes Maß für die durchschnittliche Krümmung der Kurve über diesem Intervall. Daher nennt man

$$\kappa = \lim_{\Delta s \rightarrow 0} \frac{\Delta\varphi}{\Delta s} = \lim_{\Delta t \rightarrow 0} \frac{\Delta\varphi(t)}{\Delta t} \frac{\Delta t}{s(t)} = \frac{\dot{\varphi}(t)}{\dot{s}(t)}$$

die KRÜMMUNG der Kurve im Punkt $\mathbf{r}(t)$.

Wir wollen die Krümmung κ durch die Funktionen $x(t)$, $y(t)$ der Parameterdarstellung ausdrücken. Es ist

$$\dot{s}(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}.$$

Aus

$$\sin \varphi(t) = \frac{\dot{y}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}$$

$$\cos \varphi(t) = \frac{\dot{x}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}$$

folgt durch Differenzieren

$$\begin{aligned} \dot{\varphi}(t) \cos \varphi(t) &= \frac{\ddot{y}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}} - \dot{y}(t) \frac{\dot{x}(t)\ddot{x}(t) + \dot{y}(t)\ddot{y}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}^3} \\ -\dot{\varphi}(t) \sin \varphi(t) &= \frac{\ddot{x}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}} - \dot{x}(t) \frac{\dot{x}(t)\ddot{x}(t) + \dot{y}(t)\ddot{y}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}^3}. \end{aligned}$$

Multiplizieren wir die erste Gleichung mit $\cos \varphi(t)$ und die zweite mit $-\sin \varphi(t)$ und addieren beide Ergebnisse, erhalten wir

$$\begin{aligned} \dot{\varphi}(t) &= \dot{\varphi}(t) [\cos^2 \varphi(t) + \sin^2 \varphi(t)] \\ &= \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{\dot{x}(t)^2 + \dot{y}(t)^2}. \end{aligned}$$

Insgesamt folgt:

Die KRÜMMUNG einer Kurve mit der zweimal differenzierbaren, regulären Parameterdarstellung

$$x = x(t),$$

$$y = y(t)$$

ist gegeben durch

$$\kappa(t) = \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{\frac{3}{2}}}.$$

BEISPIEL V.5.10. Die Krümmung des Graphen $y = f(x)$ ist

$$\kappa(x) = \frac{f''(x)}{(1 + f'(x)^2)^{\frac{3}{2}}}.$$

Speziell ergibt sich für die Parabel $y = x^2$ die Krümmung

$$\kappa(x) = \frac{2}{(1 + 4x^2)^{\frac{3}{2}}}.$$

BEISPIEL V.5.11. Für die Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

ergibt sich aus der Parameterdarstellung von Beispiel V.5.4 die Krümmung

$$\begin{aligned} \kappa(t) &= \frac{-a \sin(t)(-b \sin(t)) - b \cos(t)(-a \cos(t))}{(a^2 \cos^2(t) + b^2 \sin^2(t))^{\frac{3}{2}}} \\ &= \frac{ab}{(a^2 \cos^2(t) + b^2 \sin^2(t))^{\frac{3}{2}}}. \end{aligned}$$

Insbesondere ergibt sich für einen Kreis mit Radius r , d.h. $a = b = r$, die konstante Krümmung

$$\kappa = \frac{1}{r}.$$

BEMERKUNG V.5.12. Man kann zeigen, dass Kreise dadurch eindeutig charakterisiert sind, dass sie eine konstante, von Null verschiedene Krümmung haben. Die konstante Krümmung ist das Reziproke des Kreisradius. Also ist die Krümmung umso größer, je kleiner der Kreisradius ist.

Einen Kreis, der durch den Punkt $P = \mathbf{r}(t)$ der Kurve K geht und dort dieselbe Krümmung und dieselbe Tangentenrichtung hat, nennt man den KRÜMMUNGSKREIS an die Kurve K im Punkt P . Wegen Beispiel V.5.11 hat er den Radius $\frac{1}{|\kappa(t)|}$. Sein Mittelpunkt liegt auf der

Normalen im Abstand $\frac{1}{|\kappa(t)|}$ von P . Unter Berücksichtigung der Orientierung bzw. des Vorzeichens von $\kappa(t)$ ergibt sich für den Mittelpunkt (x_M, y_M) des Krümmungskreises

$$\begin{aligned} x_M &= x(t) - \frac{1}{\kappa(t)} \frac{\dot{y}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}} \\ &= x(t) - \dot{y}(t) \frac{\dot{x}(t)^2 + \dot{y}(t)^2}{\dot{x}(t)\dot{y}(t) - \ddot{x}(t)\dot{y}(t)} \\ y_M &= y(t) + \frac{1}{\kappa(t)} \frac{\dot{x}(t)}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}} \\ &= y(t) + \dot{x}(t) \frac{\dot{x}(t)^2 + \dot{y}(t)^2}{\dot{x}(t)\dot{y}(t) - \ddot{x}(t)\dot{y}(t)}. \end{aligned}$$

Durchläuft t das Parameterintervall, beschreibt (x_M, y_M) eine Kurve, die sogenannte EVOLUTE der gegebenen Kurve.

BEISPIEL V.5.13. Die Evolute der Parabel $y = x^2$ hat wegen

$$\begin{aligned} x(t) &= t, \\ y(t) &= t^2 \end{aligned}$$

die Parameterdarstellung

$$\begin{aligned} x_M &= t - 2t \frac{1 + 4t^2}{2} \\ &= -4t^3 \\ y_M &= t^2 + \frac{1 + 4t^2}{2} \\ &= 3t^2 + \frac{1}{2}. \end{aligned}$$

Durch Elimination von t ergibt sich hieraus die explizite Darstellung

$$y_M = 3 \left(\frac{x_M}{4} \right)^{\frac{2}{3}} + \frac{1}{2}.$$

Diese Kurve heißt NEILSCHE PARABEL.

V.5.5. Polardarstellung einer ebenen Kurve. Jeder Punkt $P = (x, y)$ der Ebene ist eindeutig bestimmt durch seinen Abstand r zum Ursprung und den Winkel φ der Drehung, die $(r, 0)$ nach (x, y) abbildet. Das Paar (r, φ) heißt POLARKOORDINATEN des Punktes P . Die Umrechnung zwischen kartesischen Koordinaten und Polarkoordinaten erfolgt mit den Beziehungen

$x = r \cos \varphi \quad , \quad y = r \sin \varphi \quad (0 \leq \varphi < 2\pi)$
$r = \sqrt{x^2 + y^2} \quad , \quad \varphi = \begin{cases} \arccos(\frac{x}{r}) & \text{falls } y \geq 0 \text{ und } r > 0 \\ 2\pi - \arccos(\frac{x}{r}) & \text{falls } y < 0 \text{ und } r > 0 \\ \text{undefiniert} & \text{falls } r = 0. \end{cases}$

Die Spitze eines um den Nullpunkt kreisenden Zeigers, der mit dem Drehwinkel φ auch seine Länge ändert, $r = r(\varphi)$, liegt auf einer Kurve. Die Polarkoordinaten der Kurvenpunkte sind $(r(\varphi), \varphi)$. Die Vorschrift $r = r(\varphi)$, $\alpha \leq \varphi \leq \beta$, heißt **POLARDARSTELLUNG** der Kurve; die Achse, von der aus φ gemessen wird, heißt die **POLARACHSE**. Ist die Polarachse die positive x -Achse, ergibt sich aus der Polardarstellung der Kurve die Parameterdarstellung

$$\begin{aligned}x &= r(\varphi) \cos(\varphi), \\y &= r(\varphi) \sin(\varphi)\end{aligned}$$

mit dem Winkel φ als Parameter. Hieraus kann man Tangenten, Normalen, Krümmung, Länge usw. berechnen. Insbesondere ergibt sich:

Die Länge einer Kurve mit Polardarstellung $r = r(\varphi)$, $\alpha \leq \varphi \leq \beta$, ist

$$\int_{\alpha}^{\beta} \sqrt{r(\varphi)^2 + \left(\frac{d}{d\varphi}r(\varphi)\right)^2} d\varphi$$

BEISPIEL V.5.14. Die **ARCHIMEDISCHE SPIRALE** hat die Polardarstellung $r = a\varphi$ mit $a > 0$ und $0 \leq \varphi < \infty$. Die Länge nach einem Umlauf, d.h. $0 \leq \varphi \leq 2\pi$, ist

$$\begin{aligned}\int_0^{2\pi} \sqrt{a^2\varphi^2 + a^2} d\varphi &= a \int_0^{2\pi} \sqrt{\varphi^2 + 1} d\varphi \\ &= \frac{a}{2} \left[2\pi\sqrt{1 + 4\pi^2} + \ln(2\pi + \sqrt{1 + 4\pi^2}) \right].\end{aligned}$$

BEISPIEL V.5.15. Durch $r = a(1 + \cos \varphi)$, $a > 0$, $0 \leq \varphi \leq 2\pi$, wird eine Epizykloide beschrieben. Der Umfang ist

$$\begin{aligned}& \int_0^{2\pi} \sqrt{a^2(1 + \cos \varphi)^2 + a^2 \sin^2 \varphi} d\varphi \\ &= a \int_0^{2\pi} \sqrt{2 + 2 \cos \varphi} d\varphi \quad \left| \cos \varphi = \cos^2\left(\frac{\varphi}{2}\right) - \sin^2\left(\frac{\varphi}{2}\right) \right. \\ &= a \int_0^{2\pi} \sqrt{2\left(1 + \cos^2\left(\frac{\varphi}{2}\right) - \sin^2\left(\frac{\varphi}{2}\right)\right)} d\varphi \\ &= 2a \int_0^{2\pi} \left| \cos\left(\frac{\varphi}{2}\right) \right| d\varphi \quad \left| t = \frac{\varphi}{2} \right. \\ &= 4a \int_0^{\pi} |\cos(t)| dt \quad \left| \cos\left(\frac{\pi}{2} + t\right) = -\cos\left(\frac{\pi}{2} - t\right) \right. \\ &= 8a \int_0^{\frac{\pi}{2}} \cos(t) dt \\ &= 8a.\end{aligned}$$

V.5.6. Flächeninhalte. Sind $f, g : [a, b] \rightarrow \mathbb{R}$ zwei stetige Funktionen, dann ist der Inhalt F der durch die vier Kurven $y = f(x)$, $y = g(x)$, $x = a$ und $x = b$ berandeten Fläche gegeben durch

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

BEISPIEL V.5.16. Die Fläche zwischen den beiden Parabeln $4y = x^2$ und $y^2 = 4x$ über dem Intervall $[0, 4]$ ist

$$\begin{aligned} \int_0^4 \underbrace{\left| 2\sqrt{x} - \frac{1}{4}x^2 \right|}_{\geq 0 \text{ auf } [0,4]} dx &= \left(2 \cdot \frac{2}{3}x^{\frac{3}{2}} - \frac{1}{4} \cdot \frac{1}{3}x^3 \right) \Big|_0^4 \\ &= \frac{16}{3}. \end{aligned}$$

Ist $r : [\alpha, \beta] \rightarrow [0, \infty)$ stetig, dann beträgt der Flächeninhalt F der Fläche, die in Polarkoordinaten von den drei Kurven $r = r(\varphi)$, $\varphi = \alpha$ und $\varphi = \beta$ begrenzt wird,

$$F = \frac{1}{2} \int_{\alpha}^{\beta} r(\varphi)^2 d\varphi.$$

Wir betrachten eine Kurve K mit der Eigenschaft, dass jeder vom Ursprung ausgehende Strahl die Kurve K höchstens einmal schneidet. Die Strahlenstücke zwischen dem Ursprung und der Kurve überstreichen eine Fläche, die zu K gehörende SEKTORFLÄCHE. Ist

$$\begin{aligned} x &= x(t), \\ y &= y(t), \end{aligned}$$

$a \leq t \leq b$, eine Parameterdarstellung von K , kann zu jedem Parameterwert t durch

$$\varphi(t) = \arctan\left(\frac{y(t)}{x(t)}\right)$$

eindeutig ein Polarwinkel festgelegt werden. Die Transformation $t \mapsto \varphi(t)$ ist also bijektiv. Wegen

$$\begin{aligned} \frac{d\varphi}{dt} &= \frac{1}{1 + \left(\frac{y(t)}{x(t)}\right)^2} \frac{\dot{y}(t)x(t) - y(t)\dot{x}(t)}{x(t)^2} \\ &= \frac{\dot{y}(t)x(t) - y(t)\dot{x}(t)}{x(t)^2 + y(t)^2} \\ &= \frac{\dot{y}(t)x(t) - y(t)\dot{x}(t)}{r(\varphi(t))^2} \end{aligned}$$

gilt für den Flächeninhalt der Sektorfläche die

LEIBNIZSCHE SEKTORFORMEL:

$$F = \frac{1}{2} \int_a^b [\dot{y}(t)x(t) - y(t)\dot{x}(t)]dt.$$

BEISPIEL V.5.17. Für die Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

ergibt sich wegen

$$\begin{aligned} x &= a \cos(t), \\ y &= b \sin(t) \end{aligned}$$

der Flächeninhalt

$$\begin{aligned} F &= \frac{1}{2} \int_0^{2\pi} [b \cos(t)a \cos(t) - b \sin(t)(-a \sin(t))]dt \\ &= \frac{1}{2} \int_0^{2\pi} abdt \\ &= ab\pi. \end{aligned}$$

BEISPIEL V.5.18. Die von der Astroide, Beispiel V.5.5,

$$\begin{aligned} x &= R \cos^3(t) + 2, \\ y &= R \sin^3(t) + 3, \end{aligned}$$

$0 \leq t \leq 2\pi$, begrenzte Fläche hat den Inhalt

$$\begin{aligned} F &= \frac{1}{2} \int_0^{2\pi} [(R \cos^3(t) + 2)3R \sin^2(t) \cos(t) \\ &\quad - (R \sin^3(t) + 3)3R \cos^2(t)(-\sin(t))]dt \\ &= \frac{1}{2} \int_0^{2\pi} [3R^2 \cos^4(t) \sin^2(t) + 3R^2 \sin^4(t) \cos^2(t) \\ &\quad + 6R \sin^2(t) \cos(t) + 9R \cos^2(t) \sin(t)]dt \\ &= \frac{1}{2} \int_0^{2\pi} 3R^2 \cos^2(t) \sin^2(t)dt + \underbrace{\frac{1}{2} \int_0^{2\pi} 6R \sin^2(t) \cos(t)dt}_{=0} \\ &\quad + \underbrace{\frac{1}{2} \int_0^{2\pi} 9R \cos^2(t) \sin(t)dt}_{=0} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \int_0^{2\pi} \frac{3}{4} R^2 \sin^2(2t) dt \quad | z = 2t \\
&= \int_0^{4\pi} \frac{3}{16} R^2 \sin^2(z) dz \\
&= \frac{3}{8} R^2 \int_0^{2\pi} \sin^2(z) dz \quad | \text{Beispiel V.2.2(3) (S. 171)} \\
&= \frac{3}{8} \pi R^2 \quad | R = 4r \\
&= 6\pi r^2.
\end{aligned}$$

Das ist das 6-fache des Inhaltes des kleinen, im Innern abrollenden Kreises.

V.5.7. Volumina von Rotationskörpern. Ein durch Drehung der Kurve $y = f(x)$, $a \leq x \leq b$, um die x -Achse erzeugter Rotationskörper hat das Volumen

$$V = \pi \int_a^b f(x)^2 dx.$$

BEISPIEL V.5.19. Nach Wahl eines passenden Koordinatensystems entsteht ein Kreiskegel durch Rotation der Kurve $y = \frac{r}{h}x$, $0 \leq x \leq h$, um die x -Achse. Das Volumen ist

$$\begin{aligned}
V &= \pi \int_0^h \left(\frac{r}{h}x\right)^2 dx \\
&= \pi \frac{r^2}{h^2} \frac{1}{3} h^3 \\
&= \frac{1}{3} \pi r^2 h.
\end{aligned}$$

BEISPIEL V.5.20. Man erhält einen TORUS (Reifen) mit Radien $0 < r < R$, indem man aus dem Rotationskörper, der durch Drehung von $y_1 = R + \sqrt{r^2 - x^2}$, $-r \leq x \leq r$, um die x -Achse entsteht, den Rotationskörper ausschneidet, der durch Drehung von $y_2 = R - \sqrt{r^2 - x^2}$, $-r \leq x \leq r$, um die x -Achse entsteht. Damit ist das Volumen des Torus

$$\begin{aligned}
V &= \pi \int_{-r}^r (y_1^2 - y_2^2) dx \\
&= \pi \int_{-r}^r 4R\sqrt{r^2 - x^2} dx \quad | x = r \cos(t) \\
&= 4\pi R \int_0^\pi \sqrt{r^2 - r^2 \cos^2(t)} r \sin(t) dt \\
&= 4\pi R r^2 \int_0^\pi \sin^2(t) dt \quad | \text{Beispiel V.2.2(3) (S. 171)}
\end{aligned}$$

$$= 2\pi^2 r^2 R.$$

V.5.8. Mantelflächen von Rotationskörpern. Die Mantelfläche eines durch Drehung der Kurve $y = f(x)$, $a \leq x \leq b$, um die x -Achse erzeugten Rotationskörpers ist

$$M = 2\pi \int_a^b f(x) \sqrt{1 + f'(x)^2} dx.$$

BEISPIEL V.5.21. Die Mantelfläche des Kegels aus Beispiel V.5.19 ist

$$\begin{aligned} M &= 2\pi \int_0^h \frac{r}{h} x \sqrt{1 + \left(\frac{r}{h}\right)^2} dx \\ &= 2\pi \frac{r}{h} \sqrt{1 + \left(\frac{r}{h}\right)^2} \frac{1}{2} h^2 \\ &= \pi r \sqrt{r^2 + h^2}. \end{aligned}$$

BEISPIEL V.5.22. Die Mantelfläche des Torus aus Beispiel V.5.20 ist die Summe der Mantelflächen der beiden zu y_1 und y_2 gehörenden Körper. Wegen $y_2' = -y_1'$ erhalten wir

$$\begin{aligned} M &= 2\pi \int_{-r}^r (y_1 + y_2) \sqrt{1 + (y_1')^2} dx \\ &= 2\pi \int_{-r}^r 2R \sqrt{1 + (y_1')^2} dx \quad \left| y_1' = -\frac{x}{\sqrt{r^2 - x^2}} \right. \\ &= 4\pi R \int_{-r}^r \frac{r}{\sqrt{r^2 - x^2}} dx \\ &= 4\pi R \int_{-r}^r \frac{1}{\sqrt{1 - \left(\frac{x}{r}\right)^2}} dx \quad \left| t = \frac{x}{r} \right. \\ &= 4\pi R r \int_{-1}^1 \frac{1}{\sqrt{1 - t^2}} dt \\ &= 8\pi R r \int_0^1 \frac{1}{\sqrt{1 - t^2}} dt \\ &= 8\pi R r \arcsin(t) \Big|_0^1 \\ &= 4\pi^2 R r. \end{aligned}$$

V.6. Numerische Integration*

V.6.1. Quadraturformeln*. Viele bestimmte Integrale kann man nicht explizit bestimmen. Man muss sie numerisch näherungsweise berechnen. Hierzu benutzt man QUADRATURFORMELN; das sind Ausdrücke der Form

$$Q(f) = \sum_{k=0}^n c_k f(x_k).$$

Die Zahlen c_0, c_1, \dots, c_n heißen die **GEWICHTE** der Quadraturformel; $x_0, x_1, \dots, x_n \in [a, b]$ sind die **KNOTEN** ($[a, b]$ ist der Integrationsbereich). Die Güte einer Quadraturformel wird durch ihre **ORDNUNG** gemessen. Dies ist die maximale Zahl K , sodass alle Polynome vom Grad $\leq K$ durch die Quadraturformel exakt integriert werden.

Die wichtigsten Quadraturformeln sind:

$$\begin{aligned} (b-a)f\left(\frac{a+b}{2}\right) & \text{ MITTELPUNKTSREGEL,} \\ & \text{Ordnung 1} \\ \frac{b-a}{2}[f(a) + f(b)] & \text{ TRAPEZREGEL,} \\ & \text{Ordnung 1} \\ \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] & \text{ SIMPSONREGEL,} \\ & \text{Ordnung 3.} \end{aligned}$$

Geometrisch beschreibt die Trapezregel die Fläche des Trapezes mit Grundseite $[a, b]$ und Höhen $f(a)$ und $f(b)$.

Bei fest vorgegebenen Knoten (z.B. ÄQUIDISTANT: $x_k = a + \frac{b-a}{n}k$, $0 \leq k \leq n$) kann man die Gewichte so bestimmen, dass man eine Quadraturformel der Ordnung n erhält. Die entsprechenden Gewichte sind gegeben durch

$$c_k = \int_a^b \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x - x_i}{x_k - x_i} \quad (0 \leq k \leq n).$$

Die entsprechenden Quadraturformeln heißen **NEWTON-COTES-FORMELN**. Die Trapezregel ist ein Beispiel für eine Newton-Cotes-Formel.

Lässt man auch die Wahl der Knoten frei, kann man Quadraturformeln der Ordnung $2n + 1$ konstruieren. Derartige Formeln heißen **GAUSS-FORMELN**. Die Mittelpunktsregel ist ein Beispiel für eine Gauß-Formel. Weitere Beispiele sind:

$$\frac{b-a}{2} \left[f\left(\frac{3+\sqrt{3}}{6}a + \frac{3-\sqrt{3}}{6}b\right) + f\left(\frac{3-\sqrt{3}}{6}a + \frac{3+\sqrt{3}}{6}b\right) \right]$$

Ordnung 3,

$$\frac{b-a}{18} \left[5f\left(\frac{5+\sqrt{15}}{10}a + \frac{5-\sqrt{15}}{10}b\right) + 8f\left(\frac{a+b}{2}\right) + 5f\left(\frac{5-\sqrt{15}}{10}a + \frac{5+\sqrt{15}}{10}b\right) \right]$$

Ordnung 5.

Für beliebiges n sind die Knoten der entsprechenden Gauß-Formel die auf das Intervall $[a, b]$ transformierten Nullstellen des $(n+1)$ -ten Legendre-Polynoms; die Gewichte sind durch obige Formel für Newton-Cotes-Formeln gegeben. Knoten und Gewichte können numerisch stabil und effizient durch Lösen eines geeigneten Eigenwertproblems bestimmt werden.

V.6.2. Zusammengesetzte Quadraturformeln*. Man kann offensichtlich versuchen, ein zu berechnendes Integral immer genauer numerisch anzunähern, indem man Quadraturformeln mit immer mehr Knoten und entsprechend höherer Ordnung benutzt. Man kann jedoch zeigen, dass dieser einfache Weg nicht funktioniert: Egal wie man die Folge der Knoten wählt, es gibt immer eine Funktion, dessen Integral mit wachsender Knotenzahl *immer schlechter* approximiert wird.

Einen Ausweg aus dieser Zwickmühle bieten ZUSAMMENGESETZTE QUADRATURFORMELN. Die Idee ist einfach: Das Integrationsintervall $[a, b]$ wird in viele gleich große Intervalle unterteilt, auf jedem Teilintervall wird eine feste Quadraturformel, z.B. die Trapezregel, angewandt und die Ergebnisse werden aufaddiert. Zur Verdeutlichung wähle ein $n \geq 2$ und setze $h = \frac{b-a}{n}$, dann lauten die ZUSAMMENGESETZTE MITTELPUNKTSREGEL, TRAPEZREGEL und SIMPSONREGEL:

$$h \sum_{i=1}^n f\left(a + \frac{2i-1}{2}h\right),$$

$$\frac{h}{2} \left[f(a) + 2 \sum_{i=1}^{n-1} f(a+ih) + f(b) \right],$$

$$\frac{h}{6} \left[f(a) + 2 \sum_{i=1}^{n-1} f(a+ih) + 4 \sum_{i=1}^n f\left(a + \frac{2i-1}{2}h\right) + f(b) \right].$$

Der Fehler dieser drei Näherungen zu dem Integral $\int_a^b f(x)dx$ verhält sich wie ch^2 für die Mittelpunkts- und Trapezregel und wie ch^4 für die Simpsonregel.

Generell gilt:

Basiert die zusammengesetzte Quadraturformel auf einer Formel der Ordnung K , so hat sie bei hinreichend oft differenzierbaren Integranden f einen Fehler von ch^{K+1} .

Dabei ist c eine Konstante, die nur von der betreffenden Quadraturformel und dem Integranden f abhängt. Ein Fehler von ch^2 bzw. ch^4 bedeutet, dass sich der Fehler bei Halbieren von h um den Faktor 4 bzw. 16 reduziert.

Die folgenden Java-Programme realisieren die zusammengesetzte Mittelpunkts-, Trapez- und Simpsonregel sowie die zusammengesetzten Gauß-Formeln der Ordnungen 3, 5 und 7. Dabei ist `fct` eine abstrakte Klasse, die für eine stetige Funktion den Funktionswert `f` liefert.

```
// midpoint rule
private double midpoint( int nn ) {
    double integral = 0.0;
    double h = (rightBoundary - leftBoundary)/nn;
    double x = leftBoundary + h/2;
    for( int i = 0; i < nn; i++ ) {
        integral += fct.f( x );
        x += h;
    }
    return integral*h;
} // end of midpoint
// trapezoidal rule
private double trapezoidal( int nn ) {
    double integral;
    double h = (rightBoundary - leftBoundary)/nn;
    integral = (fct.f( leftBoundary ) + fct.f( rightBoundary ))/2;
    double x = leftBoundary + h;
    for( int i = 1; i < nn; i++ ) {
        integral += fct.f( x );
        x += h;
    }
    return integral*h;
} // end of trapezoidal
// simpson rule
private double simpson( int nn ) {
    return (trapezoidal( nn ) + 2.0*midpoint( nn ))/3.0;
} // end of simpson
// 2 point Gauss rule
private double Gauss2( int nn ) {
    double[] x = new double[2];
    double[] w = new double[2];
    double integral = 0.0;
    double a = leftBoundary;
    double h = (rightBoundary - leftBoundary)/nn;
    w[0] = w[1] = 0.5;
    x[0] = (3.0 - Math.sqrt(3.0))/6.0;
    x[1] = 1.0 - x[0];
    x[0] *= h,    x[1] *= h;
    for( int i = 0; i < nn; i++ ) {
        integral += w[0]*fct.f( a + x[0] ) + w[1]*fct.f( a + x[1] );
        a += h;
    }
}
```

```

    }
    return integral*h;
} // end of Gauss2
// 3 point Gauss rule
private double Gauss3( int nn ) {
    double[] x = new double[3];
    double[] w = new double[3];
    double integral = 0.0;
    double a = leftBoundary;
    double h = (rightBoundary - leftBoundary)/nn;
    w[0] = w[2] = 5.0/18.0,    w[1] = 4.0/9.0;
    x[0] = (5.0 - Math.sqrt(15.0))/10.0;
    x[2] = 1.0 - x[0],    x[1] = 0.5;
    x[0] *= h,    x[1] *= h,    x[2] *= h;
    for( int i = 0; i < nn; i++ ) {
        integral += w[0]*fct.f( a + x[0] ) + w[1]*fct.f( a + x[1] )
            + w[2]*fct.f( a + x[2] );
        a += h;
    }
    return integral*h;
} // end of Gauss3
// 4 point Gauss rule
private double Gauss4( int nn ) {
    double[] x = new double[4];
    double[] w = new double[4];
    double integral = 0.0;
    double a = leftBoundary;
    double h = (rightBoundary - leftBoundary)/nn;
    w[0] = w[3] = (18.0 - Math.sqrt(30.0))/72.0;
    w[1] = w[2] = 0.5 - w[0];
    x[0] = (1.0 - Math.sqrt( (15.0 + 2.0*Math.sqrt(30.0))/35.0 ))/2.0;
    x[1] = (1.0 - Math.sqrt( (15.0 - 2.0*Math.sqrt(30.0))/35.0 ))/2.0;
    x[2] = 1.0 - x[1],    x[3] = 1.0 - x[0];
    x[0] *= h,    x[1] *= h,    x[2] *= h,    x[3] *= h;
    for( int i = 0; i < nn; i++ ) {
        integral += w[0]*fct.f( a + x[0] ) + w[1]*fct.f( a + x[1] );
        integral += w[2]*fct.f( a + x[2] ) + w[3]*fct.f( a + x[3] );
        a += h;
    }
    return integral*h;
} // end of Gauss4

```

BEISPIEL V.6.1. Tabelle V.6.1 zeigt die Ergebnisse der zusammengesetzten Mittelpunkts-, Trapez- und Simpsonregel für das Integral

$$\int_0^1 \frac{3}{2} \sqrt{x} dx = 1.$$

Für die zusammengesetzten Gauß-Formeln mit 2, 3 und 4 Knoten (Ordnungen 3, 5 und 7) sind die entsprechenden Ergebnisse in Tabelle V.6.2 wiedergegeben.

V.6.3. Romberg Verfahren*. Es soll $\int_a^b f(x)dx$ näherungsweise berechnet werden. Wir bezeichnen dazu mit $T_{0,k}$, $k = 0, 1, \dots, N$, das Ergebnis der zusammengesetzten Trapezregel für dieses Integral zu $h =$

TABELLE V.6.1. Mittelpunkts-, Trapez- und Simpsonregel angewandt auf das Integral $\int_0^1 \frac{3}{2}\sqrt{x}$

h	Mittelpunkt	Trapez	Simpson
1	1.06066	0.75000	0.95711
$\frac{1}{2}$	1.02452	0.99530	0.98479
$\frac{1}{4}$	1.00947	0.96493	0.99462
$\frac{1}{8}$	1.00355	0.98720	0.99810
$\frac{1}{16}$	1.00131	0.99537	0.99933
$\frac{1}{32}$	1.00047	0.99834	0.99976
$\frac{1}{64}$	1.00017	0.99941	0.99992
$\frac{1}{128}$	1.00006	0.99979	0.99997

TABELLE V.6.2. Gaußsche Formeln mit 2, 3 und 4 Knoten angewandt auf das Integral $\int_0^1 \frac{3}{2}\sqrt{x}$

h	2 Knoten	3 Knoten	4 Knoten
1	1.01083	1.00377	1.00174
$\frac{1}{2}$	1.00386	1.00133	1.00062
$\frac{1}{4}$	1.00137	1.00047	1.00022
$\frac{1}{8}$	1.00048	1.00017	1.00008
$\frac{1}{16}$	1.00017	1.00006	1.00003
$\frac{1}{32}$	1.00006	1.00002	1.00001
$\frac{1}{64}$	1.00002	1.00001	1.00000
$\frac{1}{128}$	1.00001	1.00000	1.00000

$(b-a)2^{-k}$, d.h. zur Unterteilung in 2^k Teilintervalle. Wir wählen ein $K \in \{1, \dots, N\}$ und berechnen nun rekursiv

ROMBERG-VERFAHREN:

$$T_{i+1,k} = \frac{4^{i+1}T_{i,k+1} - T_{i,k}}{4^{i+1} - 1}$$

$$i = 0, 1, \dots, K-1, \quad k = 0, 1, \dots, N-i-1.$$

Jedes $T_{i,k}$ ist eine Näherung für $\int_a^b f(x)dx$ mit einem Fehler $c_i \left(\frac{b-a}{2^k}\right)^{i+1}$, d.h.

$$\int_a^b f(x)dx - T_{i,k} = O(2^{-k(i+1)}).$$

Dabei hängen die Faktoren c_i nur von i und dem Integranden f ab. Der Übergang von $T_{i,k}$ zu $T_{i,k+1}$ bedeutet also eine Reduktion des Fehlers um den Faktor 2^{i+1} .

Das folgende Java-Programm realisiert das Romberg-Verfahren. Die Methode `trapezoidal` ist die im vorigen Abschnitt angegebene zusammengesetzte Trapezregel. Die Variable `numberOfIntervals` gibt die ursprüngliche Zahl der Teilintervalle für die zusammengesetzte Trapezregel an.

```
// Romberg scheme
private void romberg() {
    double factor = 4.0;
    int ni = numberOfIntervals;
    for( int level = 0; level <= numberOfRefinements; level++ ) {
        quad[0][level] = trapezoidal( ni );
        ni *=2;
    }
    for( int column = 1; column <= numberOfRefinements; column++ ) {
        for( int level = 0; level <= numberOfRefinements - column;
            level++ )
            quad[column][level] = (factor*quad[column-1][level+1]
                                   - quad[column-1][level])
                                   /(factor - 1.0);
        factor *= 4.0;
    }
} // end of romberg
```

BEISPIEL V.6.2. Tabelle V.6.3 zeigt die Ergebnisse des Romberg-Verfahrens angewandt auf das Integral $\int_0^1 \frac{3}{2}\sqrt{x} = 1$.

TABELLE V.6.3. Romberg-Verfahren für $\int_0^1 \frac{3}{2}\sqrt{x} = 1$

k	$T_{0,k}$	$T_{1,k}$	$T_{2,k}$	$T_{3,k}$	$T_{4,k}$
0	0.750000	0.957107	0.986635	0.995411	0.998389
1	0.905330	0.984789	0.995274	0.998378	0.999431
2	0.964925	0.994619	0.998329	0.999426	0.999799
3	0.987195	0.998097	0.999409	0.999797	0.999929
4	0.995372	0.999327	0.999791	0.999928	
5	0.998338	0.999762	0.999926		
6	0.999406	0.999916			
7	0.999788				

KAPITEL VI

Gewöhnliche Differentialgleichungen I Skalare Gleichungen

VI.1. Einführung

VI.1.1. Beispiele. Zur Motivation beginnen wir mit einigen vertrauten Beispielen und interpretieren sie als gewöhnliche Differentialgleichungen.

BEISPIEL VI.1.1. Wir suchen eine Stammfunktion y zu einer gegebenen Funktion f auf einem Intervall I , d.h. gesucht ist eine differenzierbare Funktion $y : I \rightarrow \mathbb{R}$ mit

$$y'(x) = f(x) \quad \text{für alle } x \in I.$$

Dies ist eine skalare gewöhnliche Differentialgleichung erster Ordnung. Da jede stetige Funktion eine Stammfunktion besitzt, hat die Differentialgleichung mindestens eine Lösung. Da je zwei Stammfunktionen von f sich um eine additive Konstante unterscheiden, besitzt die Differentialgleichung unendlich viele Lösungen, die Lösungsmenge ist einparametrig.

Wir wählen nun einen beliebigen Punkt $x_0 \in I$ und eine beliebige reelle Zahl y_0 aus. Dann gibt es genau eine Stammfunktion y von f mit $y(x_0) = y_0$:

$$\begin{aligned} y'(x) &= f(x) \quad \text{für alle } x \in I \\ y(x_0) &= y_0. \end{aligned}$$

Dies ist ein Anfangswertproblem. Seine eindeutige Lösung ist

$$y(x) = y_0 + \int_{x_0}^x f(t) dt.$$

BEISPIEL VI.1.2. Zu einer gegebenen stetigen Funktion f auf einem Intervall I suchen wir jetzt eine Funktion y , deren n -te Ableitung, $n \geq 1$, gleich f ist:

$$y^{(n)}(x) = f(x) \quad \text{für alle } x \in I.$$

Dies ist eine skalare gewöhnliche Differentialgleichung n -ter Ordnung. Wir können alle ihre Lösungen durch n -malige unbestimmte Integration von f bestimmen. Die Lösungsmenge ist also n -parametrig.

Wählen wir wieder einen Punkt $x_0 \in I$ aus und schreiben die Werte

von $y(x_0), \dots, y^{(n-1)}(x_0)$ vor, erhalten wir ein Anfangswertproblem:

$$\begin{aligned} y^{(n)}(x) &= f(x) \quad \text{für alle } x \in I \\ y(x_0) &= y_0 \\ &\vdots \\ y^{(n-1)}(x_0) &= y_{n-1}. \end{aligned}$$

Es hat wieder eine eindeutige Lösung.

BEISPIEL VI.1.3. Wir betrachten eine lineare Pendelbewegung. Nach dem Hookeschen Gesetz ist die rücktreibende Kraft proportional zur Auslenkung $s(t)$. Nach den Newtonschen Bewegungsgesetzen ist diese Kraft andererseits proportional zur Beschleunigung $\ddot{s}(t)$. Daher erfüllt die Auslenkung die gewöhnliche Differentialgleichung

$$\ddot{s}(t) + \omega^2 s(t) = 0 \quad \text{für alle } t \in \mathbb{R}$$

mit $\omega \in \mathbb{R}$. Die Auslenkung ist für alle Zeiten eindeutig bestimmt durch die Angabe der Anfangsauslenkung $s(t_0)$ und der Anfangsgeschwindigkeit $\dot{s}(t_0)$ zu einer beliebig gewählten Anfangszeit t_0 .

BEISPIEL VI.1.4. Wir betrachten eine gedämpfte Federschwingung. Die Auslenkung $s(t)$ genügt aufgrund der Newtonschen Kraftgesetze der gewöhnlichen Differentialgleichung

$$\ddot{s}(t) + r\dot{s}(t) + \omega^2 s(t) = 0 \quad \text{für alle } t \in \mathbb{R}.$$

Dabei ist $\omega \in \mathbb{R}$ wie in Beispiel VI.1.3, und $r > 0$ beschreibt die Reibung.

BEISPIEL VI.1.5. Ein Fahrzeug der Masse $m > 0$ werde durch die konstante Kraft $K > 0$ beschleunigt. Der Luftwiderstand sei proportional zum Quadrat der Geschwindigkeit. Dann führen die Newtonschen Kraftgesetze auf die gewöhnliche Differentialgleichung

$$m\ddot{x}(t) = K - r\dot{x}(t)^2 \quad \text{für alle } t \in \mathbb{R}$$

für die zurückgelegte Strecke $x(t)$. Dabei ist $r > 0$.

VI.1.2. Grundbegriffe. Gegeben seien ein Intervall $I \subset \mathbb{R}$, eine natürliche Zahl $n \geq 1$ und eine Funktion $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}$.

Eine Bestimmungsgleichung der Form

$$(VI.1.1) \quad y^{(n)} = f(x, y, \dots, y^{(n-1)}),$$

in der neben der Variablen x und der gesuchten Funktion $y : I \rightarrow \mathbb{R}$ auch deren Ableitungen bis und mit der Ordnung n vorkommen, heißt eine (EXPLIZITE, SKALARE) GEWÖHNLICHE DIFFERENTIALGLEICHUNG n -TER ORDNUNG oder kurz GDGL n -TER ORDNUNG.

(Die Gleichung heißt explizit, weil die höchste Ableitung $y^{(n)}$ nur auf der linken Seite der Gleichung auftritt. Sie heißt skalar, weil die gesuchte Funktion $y : I \rightarrow \mathbb{R}$ reell- und nicht vektorwertig ist.)

Im Fall $n = 1$ sprechen wir einfach von einer GEWÖHNLICHEN DIFFERENTIALGLEICHUNG oder kurz gDGL. Eine gDgl (beliebiger Ordnung) heißt AUTONOM, wenn die Funktion f nicht von der Variablen x abhängt, d.h. die abhängige Variable tritt nicht explizit auf.

Eine auf einem Intervall J n -mal differenzierbare Funktion $y : J \rightarrow \mathbb{R}$ heißt eine LÖSUNG der gDgl (VI.1.1), wenn $J \subset I$ ist und wenn für alle $x \in J$ gilt

$$y^{(n)}(x) = f(x, y(x), \dots, y^{(n-1)}(x)).$$

Man beachte, dass der Definitionsbereich J einer Lösung y von (VI.1.1) eine echte Teilmenge des Intervalls I sein kann.

Wie die Beispiele des vorigen Abschnittes zeigen, kann die Lösungsmenge einer gDgl mehrparametrische Lösungsscharen enthalten. Man fasst jede r -parametrische Lösungsschar als *eine* Lösung mit r freien Parametern auf. Eine einzelne Lösung, die keine frei wählbaren Parameter enthält, wird als SPEZIELLE oder PARTIKULÄRE LÖSUNG bezeichnet. Häufig erhält man eine partikuläre Lösung durch Festlegen der Parameter einer parameterabhängigen Lösung. Eine partikuläre Lösung, die keiner parameterabhängigen Lösungsschar angehört, nennt man eine SINGULÄRE LÖSUNG. Eine parameterabhängige Lösung einer gDgl n -ter Ordnung heißt ALLGEMEIN, wenn sie n frei wählbare Parameter enthält. Sie heißt VOLLSTÄNDIG, wenn durch Variation ihrer Parameter *alle* Lösungen der gDgl erfasst werden.

Unter einem ANFANGSWERTPROBLEM, kurz AWP, versteht man eine gDgl n -ter Ordnung zusammen mit n Bestimmungsgleichungen für die Ableitungen $y, \dots, y^{(n-1)}$ in *einem* Punkt $x_0 \in I$:

$$\begin{aligned} y^{(n)} &= f(x, y, \dots, y^{(n-1)}) \\ y(x_0) &= y_0 \\ &\vdots \\ &\vdots \\ y^{(n-1)}(x_0) &= y_{n-1}. \end{aligned}$$

Ein AWP heißt LOKAL LÖSBAR, wenn es positive Zahlen ε_1 und ε_2 gibt, so dass es auf dem Intervall $(x_0 - \varepsilon_1, x_0 + \varepsilon_2)$ eine Lösung y besitzt, d.h. y ist auf $(x_0 - \varepsilon_1, x_0 + \varepsilon_2)$ n -mal differenzierbar, löst die gDgl und erfüllt die Anfangsbedingungen. Dieses y heißt eine LOKALE LÖSUNG. Ein AWP heißt SACHGEMÄSS GESTELLT, englisch PROPERLY POSED, wenn folgende drei Bedingungen erfüllt sind:

- Existenz einer lokalen Lösung,
- Eindeutigkeit der lokalen Lösung,
- stetige Abhängigkeit der lokalen Lösung von den Anfangswerten.

VI.1.3. Geometrische Deutung. Wir betrachten die gDgl

$$y'(x) = f(x, y)$$

mit $f : I \times J \rightarrow \mathbb{R}$ und zwei Intervallen I, J in \mathbb{R} . Durch jeden Punkt $(x, y) \in I \times J$ zeichnen wir ein kurzes Geradenstück mit der Steigung $f(x, y)$. Dies liefert das sogenannte RICHTUNGSFELD der gDgl. Eine Lösung der gDgl ist dann eine Kurve, die in jedem Kurvenpunkt tangential zum Richtungsfeld verläuft.

VI.1.4. Eindeutigkeitsfragen. Das folgende Beispiel zeigt, dass nicht jedes AWP sachgemäß gestellt ist.

BEISPIEL VI.1.6. Wir betrachten das AWP

$$\begin{aligned} y' &= \sqrt{|y|} \\ y(0) &= 0. \end{aligned}$$

Offensichtlich ist die konstante Funktion $u = 0$ eine Lösung des AWP. Für die Funktion

$$v(x) = \begin{cases} \frac{1}{4}x^2 & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

erhalten wir andererseits

$$\sqrt{|v(x)|} = \begin{cases} \frac{1}{2}x & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

und

$$v'(x) = \begin{cases} \frac{1}{2}x & \text{für } x \geq 0 \\ 0 & \text{für } x < 0. \end{cases}$$

Also ist v auch eine Lösung des AWP, und das AWP ist nicht sachgemäß gestellt.

Es gibt sogar unendlich viele Lösungen: Für beliebige Zahlen $a < 0$, $b > 0$ definieren wir

$$v_{a,b}(x) = \begin{cases} \frac{1}{4}(x-b)^2 & \text{für } x > b \\ 0 & \text{für } a \leq x \leq b \\ -\frac{1}{4}(x-a)^2 & \text{für } x < a. \end{cases}$$

Wie man leicht nachrechnet, ist jede dieser Funktionen eine Lösung des AWP.

Wir werden in Kapitel X einen allgemeinen Existenz- und Eindeutigkeitssatz für Systeme gewöhnlicher Differentialgleichungen angeben, der als Spezialfall skalare gDgl umfasst. Wir können ihn an dieser Stelle nicht angeben, da wir dazu Kenntnisse über Funktionen mehrerer Veränderlicher benötigen, die wir erst in Kapitel VIII kennen lernen werden. Daher werden wir in den folgenden Abschnitten bei den dort betrachteten gDgl stets separat auf die Frage der Existenz und Eindeutigkeit eingehen.

VI.2. Differentialgleichungen 1. Ordnung**VI.2.1. Trennung der Variablen.** Wir betrachten die gDgl

$$y' = f(x)g(y).$$

Dabei sind $f : I \rightarrow \mathbb{R}$ und $g : J \rightarrow \mathbb{R}$ zwei stetige Funktionen auf Intervallen I, J in \mathbb{R} .

Ist η eine Nullstelle von g , d.h. $g(\eta) = 0$, so ist die konstante Funktion $y(x) = \eta$ für alle $x \in I$ offensichtlich eine Lösung der gDgl.

Ist $g(\eta) \neq 0$, so gibt es wegen der Stetigkeit von g ein offenes Intervall J' mit $J' \subset J$, $\eta \in J'$ und $g(z) \neq 0$ für alle $z \in J'$. Die Funktion $\frac{1}{g}$ ist auf J' stetig und besitzt daher eine Stammfunktion G

$$G = \int \frac{1}{g(z)} dz.$$

Aus der Kettenregel folgt für jede differenzierbare Funktion $y : I \rightarrow J'$

$$\begin{aligned} \frac{d}{dx} G(y(x)) &= G'(y(x))y'(x) \\ &= \frac{y'(x)}{g(y(x))}. \end{aligned}$$

Ist also y eine Lösung unserer gDgl, folgt

$$\frac{d}{dx} G(y(x)) = f(x)$$

und $G \circ y$ muss eine Stammfunktion von f sein. Bezeichnen wir mit

$$F = \int f(x) dx$$

irgendeine Stammfunktion von f , gilt daher

$$G(y) = F(x) + c$$

mit $c \in \mathbb{R}$. Da die Ableitung $\frac{1}{g}$ von G nicht verschwindet, kann gemäß Abschnitt IV.3 diese Gleichung nach y aufgelöst werden. Die so gewonnene Funktion löst unsere gDgl.

Zusammenfassend erhalten wir:

Bestimmung der allgemeinen Lösung der gDgl

$$y' = f(x)g(y) :$$

- Bestimme alle Nullstellen η von g . Die konstante Funktion

$$y(x) = \eta \quad \text{für alle } x \in I$$

ist eine Lösung der gDgl.

- Bestimme alle Teilintervalle J' von J , auf denen g nicht verschwindet. Für diese Teilintervalle berechne

$$G = \int \frac{1}{g(z)} dz$$

$$F = \int f(x) dx$$

und löse die Gleichung

$$G(y) = F(x) + c$$

mit $c \in \mathbb{R}$ nach y auf.

Für das entsprechende AWP ergibt sich folgende Vorgehensweise:

Lösen des AWP

$$y' = f(x)g(y)$$

$$y(x_0) = y_0 :$$

- Ist $g(y_0) = 0$, so ist

$$y(x) = y_0 \quad \text{für alle } x \in I$$

Lösung des AWP.

- Ist $g(y_0) \neq 0$, bestimme

$$G(y) = \int_{y_0}^y \frac{1}{g(z)} dz$$

$$F(x) = \int_{x_0}^x f(s) ds$$

und löse die Gleichung

$$G(y) = F(x)$$

nach y auf.

BEISPIEL VI.2.1. Wir greifen Beispiel VI.1.5 (S. 214) auf und bezeichnen mit $v(t) = \dot{x}(t)$ die Geschwindigkeit des Fahrzeugs. Dann gilt für v die gDgl

$$m\dot{v} = K - rv^2$$

und somit nach Division durch m

$$\dot{v} = \frac{K}{m} - \frac{r}{m}v^2.$$

Diese ist vom betrachteten Typ mit

$$f(x) = 1,$$

$$\begin{aligned} g(y) &= \frac{K}{m} - \frac{r}{m}y^2 \\ &= \frac{r}{m} \left(\sqrt{\frac{K}{r}} - y \right) \left(\sqrt{\frac{K}{r}} + y \right). \end{aligned}$$

Wegen

$$g\left(\pm\sqrt{\frac{K}{r}}\right) = 0$$

sind

$$v(t) = \sqrt{\frac{K}{r}}$$

und

$$v(t) = -\sqrt{\frac{K}{r}}$$

zwei Lösungen der gDgl. Da v hier der Betrag der Geschwindigkeit ist, ist die zweite Lösung physikalisch sinnlos.

Auf $(-\sqrt{\frac{K}{r}}, \sqrt{\frac{K}{r}})$ erhalten wir

$$\begin{aligned} \int \frac{1}{g(z)} dz &= \frac{m}{r} \int \frac{1}{\left(\sqrt{\frac{K}{r}} - z\right) \left(\sqrt{\frac{K}{r}} + z\right)} dz \\ &= \frac{m}{r} \frac{1}{2\sqrt{\frac{K}{r}}} \int \frac{1}{\left(\sqrt{\frac{K}{r}} - z\right)} + \frac{1}{\left(\sqrt{\frac{K}{r}} + z\right)} dz \\ &= \frac{m}{2\sqrt{Kr}} \left[\ln\left(\sqrt{\frac{K}{r}} + z\right) - \ln\left(\sqrt{\frac{K}{r}} - z\right) \right] \\ &= \frac{m}{2\sqrt{Kr}} \ln \left(\frac{\sqrt{\frac{K}{r}} + z}{\sqrt{\frac{K}{r}} - z} \right) \\ &= \frac{m}{2\sqrt{Kr}} \ln \left(\left| \frac{\sqrt{\frac{K}{r}} + z}{\sqrt{\frac{K}{r}} - z} \right| \right). \end{aligned}$$

Analog ergibt sich auf $(-\infty, -\sqrt{\frac{K}{r}})$ und auf $(\sqrt{\frac{K}{r}}, \infty)$

$$\int \frac{1}{g(z)} dz = \frac{m}{2\sqrt{Kr}} \ln \left(\left| \frac{\sqrt{\frac{K}{r}} + z}{\sqrt{\frac{K}{r}} - z} \right| \right).$$

Also lautet die Bestimmungsgleichung für v

$$\frac{m}{2\sqrt{Kr}} \ln \left(\left| \frac{\sqrt{\frac{K}{r}} + v}{\sqrt{\frac{K}{r}} - v} \right| \right) = \int 1 dt$$

$$= t + c.$$

Bezeichne mit $v_0 = v(0)$ die Anfangsgeschwindigkeit zur Anfangszeit $t_0 = 0$. Ist $v_0 = \sqrt{\frac{K}{r}}$, ergibt sich die Lösung

$$v(t) = \sqrt{\frac{K}{r}}$$

des AWP. Ist $v_0 \neq \sqrt{\frac{K}{r}}$ und $v_0 \geq 0$ (der Fall $v_0 < 0$ ist physikalisch sinnlos), erhalten wir durch Einsetzen von $t = 0$ und $v(0) = v_0$ in die Bestimmungsgleichung die Beziehung

$$c = \frac{m}{2\sqrt{Kr}} \ln \left(\left| \frac{\sqrt{\frac{K}{r}} + v_0}{\sqrt{\frac{K}{r}} - v_0} \right| \right).$$

Damit ergibt sich die Bestimmungsgleichung

$$\frac{m}{2\sqrt{Kr}} \ln \left(\left| \frac{\sqrt{\frac{K}{r}} + v}{\sqrt{\frac{K}{r}} - v} \right| \right) = t + \frac{m}{2\sqrt{Kr}} \ln \left(\left| \frac{\sqrt{\frac{K}{r}} + v_0}{\sqrt{\frac{K}{r}} - v_0} \right| \right)$$

für die Lösung v des AWP. Beim Auflösen dieser Gleichung nach v müssen wir die Fälle $0 \leq v_0 < \sqrt{\frac{K}{r}}$ und $v_0 > \sqrt{\frac{K}{r}}$ unterscheiden. Aus Stetigkeitsgründen gilt dann für hinreichend kleines $t \neq 0$ ebenso $0 \leq v(t) < \sqrt{\frac{K}{r}}$ bzw. $v(t) > \sqrt{\frac{K}{r}}$. Mit etwas Rechnung erhalten wir in beiden Fällen die Lösung

$$v(t) = \sqrt{\frac{K}{r}} \frac{(1 + \sqrt{\frac{r}{K}} v_0) e^{\frac{2\sqrt{Kr}}{m} t} - 1 + \sqrt{\frac{r}{K}} v_0}{(1 + \sqrt{\frac{r}{K}} v_0) e^{\frac{2\sqrt{Kr}}{m} t} + 1 - \sqrt{\frac{r}{K}} v_0}.$$

Für $t \rightarrow \infty$ ergibt sich die Grenzgeschwindigkeit

$$v_\infty = \lim_{t \rightarrow \infty} v(t)$$

$$= \sqrt{\frac{K}{r}}.$$

BEISPIEL VI.2.2. Ein zylindrischer Wasserbehälter mit Radius R entleert sich unter Einfluss der Schwerkraft durch ein kreisförmiges Loch mit Radius ρ im Boden. Nach dem Gesetz von Toricelli gilt für die Ausströmgeschwindigkeit

$$v(t) = -\sqrt{2gh(t)},$$

wobei $h(t)$ die Füllhöhe ist. Außerdem gilt die Kontinuitätsgleichung

$$R\dot{h} = \rho v.$$

Insgesamt erhalten wir das AWP

$$\begin{aligned}\dot{h} &= -\frac{\rho}{R}\sqrt{2gh} \\ h(0) &= h_0.\end{aligned}$$

Zur Abkürzung setzen wir

$$k = \frac{\rho}{R}\sqrt{2g}.$$

Dann ist

$$\begin{aligned}f(x) &= 1 \\ g(y) &= -k\sqrt{y}.\end{aligned}$$

Wir erhalten die Bestimmungsgleichung (sofern $h > 0$!)

$$\begin{aligned}\frac{2}{k}(\sqrt{h_0} - \sqrt{h}) &= -\int_{h_0}^h \frac{1}{k\sqrt{z}} dz \\ &= \int_0^t 1 ds \\ &= t.\end{aligned}$$

Lösen wir diese Gleichung nach $h(t)$ auf und beachten, dass sie nur so lange gilt, wie $h(t) > 0$ ist, erhalten wir die Lösung

$$h(t) = \begin{cases} (\sqrt{h_0} - \frac{kt}{2})^2 & \text{für } 0 \leq t < \frac{2\sqrt{h_0}}{k} \\ 0 & \text{für } t \geq \frac{2\sqrt{h_0}}{k}. \end{cases}$$

Man beachte: Das AWP

$$\begin{aligned}\dot{h} &= -k\sqrt{h} \\ h(t_0) &= 0\end{aligned}$$

ist für $t < t_0$ nicht eindeutig lösbar. Man kann bei einem leeren Behälter nicht die ursprüngliche Füllhöhe bestimmen.

VI.2.2. Variation der Konstanten. Wir betrachten jetzt gDgl der Form

$$y' + a(x)y = f(x)$$

mit auf einem Intervall I stetigen Funktionen a und f . Eine solche gDgl heißt LINEARE GDGL 1. ORDNUNG. Sie heißt HOMOGEN, wenn $f = 0$ ist; sonst heißt sie INHOMOGEN.

Wir machen folgende Beobachtungen:

- Sind y_1 und y_2 zwei Lösungen der inhomogenen gDgl, so ist $y = y_1 - y_2$ eine Lösung der homogenen gDgl.

- Ist y_h eine Lösung der homogenen gDgl und y_p eine Lösung der inhomogenen gDgl, so ist $y = y_h + y_p$ auch eine Lösung der inhomogenen gDgl.
- Sind y_1 und y_2 zwei Lösungen der homogenen gDgl und sind α, β beliebige reelle Zahlen, so ist $y = \alpha y_1 + \beta y_2$ auch eine Lösung der homogenen gDgl.

Diese Eigenschaften beruhen auf der Linearität der Abbildung $y \mapsto y' + a(x)y$.

Wegen der zweiten Beobachtung suchen wir zunächst die allgemeine Lösung der homogenen gDgl und versuchen dann, eine partikuläre Lösung der inhomogenen gDgl zu finden.

Die homogene lineare gDgl

$$y' + a(x)y = 0$$

ist vom im vorigen Abschnitt betrachteten Typ mit (in der dortigen Notation!)

$$f(x) = -a(x)$$

$$g(y) = y.$$

Bezeichnen wir mit

$$A = \int a(x)dx$$

irgendeine Stammfunktion von a , erhalten wir die Bestimmungsgleichung

$$\begin{aligned} \ln(|y|) &= \int \frac{1}{y} dy \\ &= - \int a(x) dx + c \\ &= -A(x) + c \end{aligned}$$

und damit

$$\begin{aligned} |y| &= e^{-A(x)} \underbrace{e^c}_{=\tilde{c}} \\ &= \tilde{c} e^{-A(x)}, \end{aligned}$$

woraus folgt

$$y = \hat{c} e^{-A(x)} \quad \text{mit } \hat{c} \in \mathbb{R}.$$

Insgesamt erhalten wir:

Die allgemeine Lösung der gDgl

$$y' + a(x)y = 0$$

lautet

$$y = ce^{-A(x)}$$

mit $c \in \mathbb{R}$ und

$$A = \int a(x)dx.$$

Zur Bestimmung einer partikulären Lösung der inhomogenen gDgl machen wir den Ansatz

$$y(x) = c(x)e^{-A(x)}$$

mit einer unbekannten Funktion $c(x)$. Dieser wird VARIATION DER KONSTANTEN genannt. Setzen wir diesen Ansatz in die inhomogene gDgl ein, erhalten wir wegen $A' = a$ und der Produktregel

$$\begin{aligned} f(x) &= y' + a(x)y \\ &= (c(x)e^{-A(x)})' + a(x)c(x)e^{-A(x)} \\ &= c'(x)e^{-A(x)} - c(x)\underbrace{A'(x)}_{=a(x)}e^{-A(x)} + a(x)c(x)e^{-A(x)} \\ &= c'(x)e^{-A(x)} \end{aligned}$$

und somit

$$c'(x) = e^{A(x)}f(x),$$

woraus folgt

$$c(x) = \int e^{A(x)}f(x)dx.$$

Insgesamt folgt:

Die allgemeine Lösung der gDgl

$$y' + a(x)y = f(x)$$

lautet

$$y(x) = e^{-A(x)} \left\{ c + \int f(x)e^{A(x)}dx \right\}$$

mit $c \in \mathbb{R}$ und

$$A(x) = \int a(x)dx.$$

BEISPIEL VI.2.3. Für die gDgl

$$y' + \frac{1}{x}y = x^3$$

erhalten wir auf dem Intervall $(0, \infty)$

$$\begin{aligned} A(x) &= \int \frac{1}{x} dx \\ &= \ln(x) \end{aligned}$$

und

$$\begin{aligned} y(x) &= \underbrace{e^{-\ln(x)}}_{=\frac{1}{x}} \left\{ c + \int \underbrace{x^3 e^{\ln(x)}}_{=x^4} dx \right\} \\ &= \frac{1}{x} \left\{ c + \int x^4 dx \right\} \\ &= \frac{1}{x} \left\{ c + \frac{1}{5} x^5 \right\} \\ &= \frac{c}{x} + \frac{x^4}{5}. \end{aligned}$$

BEISPIEL VI.2.4. Für die gDgl

$$y' + 2y = 3e^{5x} + x^3 - 1$$

ist

$$\begin{aligned} A(x) &= \int 2 dx \\ &= 2x \end{aligned}$$

und

$$\begin{aligned} y(x) &= e^{-2x} \left\{ c + \int e^{2x} [3e^{5x} + x^3 - 1] dx \right\} \\ &= ce^{-2x} + e^{-2x} \int [3e^{7x} + x^3 e^{2x} - e^{2x}] dx. \end{aligned}$$

Offensichtlich ist

$$\begin{aligned} \int 3e^{7x} dx &= \frac{3}{7} e^{7x}, \\ \int e^{2x} dx &= \frac{1}{2} e^{2x}. \end{aligned}$$

Mehrfache partielle Integration ergibt

$$\begin{aligned} \int x^3 e^{2x} dx &= x^3 \frac{1}{2} e^{2x} - \int \frac{3}{2} x^2 e^{2x} dx \\ &= \frac{1}{2} x^3 e^{2x} - \frac{3}{4} x^2 e^{2x} + \int \frac{3}{2} x e^{2x} dx \\ &= \frac{1}{2} x^3 e^{2x} - \frac{3}{4} x^2 e^{2x} + \frac{3}{4} x e^{2x} - \int \frac{3}{4} e^{2x} dx \end{aligned}$$

$$= \frac{1}{2}x^3e^{2x} - \frac{3}{4}x^2e^{2x} + \frac{3}{4}xe^{2x} - \frac{3}{8}e^{2x}.$$

Damit ergibt sich für die allgemeine Lösung der gDgl

$$y(x) = ce^{-2x} + \frac{3}{7}e^{5x} + \frac{1}{2}x^3 - \frac{3}{4}x^2 + \frac{3}{4}x - \frac{7}{8}.$$

VI.2.3. Homogene Differentialgleichungen. Dies sind Differentialgleichungen der Form

$$y' = f\left(\frac{y}{x}\right) \quad \text{für } x \neq 0$$

mit einer stetigen Funktion f . Ist $y(x)$ eine Lösung und $\alpha \in \mathbb{R} \setminus \{0\}$, so ist $\alpha y\left(\frac{x}{\alpha}\right)$ auch eine Lösung der gDgl. Die Lösungsmenge ist also invariant unter der Ähnlichkeitstransformation $x \mapsto \alpha x, y \mapsto \alpha y$. Daher spricht man auch von **ÄHNLICHKEITS-DIFFERENTIALGLEICHUNGEN**.

Zur Bestimmung der Lösung machen wir den Ansatz

$$v(x) = \frac{y(x)}{x} \quad \Longleftrightarrow \quad y(x) = xv(x).$$

Einsetzen in die gDgl liefert

$$\begin{aligned} f(v) &= f\left(\frac{y}{x}\right) \\ &= y' \\ &= (xv)' \\ &= xv' + v \end{aligned}$$

und somit

$$v' = \frac{1}{x}[f(v) - v].$$

Dies ist eine gDgl vom im ersten Abschnitt betrachteten Typ. Insgesamt erhalten wir:

Lösung der **HOMOGENEN DIFFERENTIALGLEICHUNG**

$$y' = f\left(\frac{y}{x}\right)$$

für $x \neq 0$:

- Bestimme alle Nullstellen η von $f(v) - v$. Die Funktion

$$y(x) = \eta x$$

ist jeweils eine Lösung der homogenen Differentialgleichung.

- Bestimme die allgemeine Lösung v der gDgl

$$v' = \frac{1}{x}[f(v) - v].$$

Die Funktion

$$y(x) = xv(x)$$

ist dann jeweils eine Lösung der homogenen Differentialgleichung.

BEISPIEL VI.2.5. Wir betrachten die gDgl

$$y' = \frac{y}{x} - \sqrt{1 - \frac{y}{x}} \quad \text{für } x \neq 0, \frac{y}{x} \leq 1.$$

Es ist

$$\begin{aligned} f(v) - v &= v - \sqrt{1 - v} - v \\ &= -\sqrt{1 - v}. \end{aligned}$$

Also ist

$$y(x) = x$$

eine Lösung der gDgl.

Für $v \neq 1$ erhalten wir gemäß Abschnitt VI.2.1 die Bestimmungsgleichung

$$\begin{aligned} \ln(|x|) + c &= \int \frac{1}{x} dx \\ &= \int \frac{1}{f(v) - v} dv \\ &= \int \frac{1}{-\sqrt{1 - v}} dv \\ &= 2\sqrt{1 - v} \end{aligned}$$

und damit

$$v(x) = 1 - \frac{1}{4}[\ln(|x|) + c]^2,$$

woraus folgt

$$y(x) = x - \frac{x}{4}[\ln(|x|) + c]^2.$$

VI.2.4. Bernoulli-Differentialgleichung. Sie ist von der Form

$$y' + a(x)y = b(x)y^\alpha$$

mit stetigen Funktionen a und b und einer Zahl $\alpha \notin \{0, 1\}$. Zur Bestimmung einer Lösung multiplizieren wir die gDgl zunächst mit $(1 - \alpha)y^{-\alpha}$ und erhalten

$$(1 - \alpha)y^{-\alpha}y' + a(x)(1 - \alpha)y^{1-\alpha} = b(x)(1 - \alpha).$$

Wegen

$$\frac{d}{dx}y(x)^{1-\alpha} = (1 - \alpha)y(x)^{-\alpha}y'(x)$$

machen wir den Ansatz

$$z(x) = y(x)^{1-\alpha}$$

und erhalten für z die gDgl

$$z' + a(x)(1 - \alpha)z = b(x)(1 - \alpha).$$

Diese gDgl kann wie in Abschnitt VI.2.2 gelöst werden.

BEISPIEL VI.2.6. Die gDgl

$$\ddot{x} + \omega^2x + r(\dot{x})^2 = 0$$

modelliert einen gedämpften harmonischen Oszillator, dessen Bremskraft proportional zum Quadrat der Geschwindigkeit ist. Wir wollen die Geschwindigkeit als Funktion des Ortes darstellen $\dot{x} = v(x)$. Durch Differentiation folgt mit der Kettenregel

$$\begin{aligned}\ddot{x} &= \frac{d}{dt}v(x) \\ &= v'(x)\dot{x} \\ &= v'(x)v(x).\end{aligned}$$

Einsetzen in die gDgl für x ergibt

$$v'v + \omega^2x + rv^2 = 0.$$

Unter der Annahme, dass v nie Null wird, können wir die gDgl durch v dividieren und erhalten

$$v' = -\omega^2\frac{x}{v} - rv.$$

Mit dem Ansatz

$$z(x) = v(x)^2$$

erhalten wir die lineare gDgl

$$z' + 2rz = -2\omega^2x$$

für z . Sie hat die allgemeine Lösung

$$z(x) = e^{-2rx} \left\{ c - \int e^{2rx} 2\omega^2x dx \right\}$$

$$\begin{aligned}
&= e^{-2rx} \left\{ c - \frac{\omega^2}{r} x e^{2rx} + \int \frac{\omega^2}{r} e^{2rx} dx \right\} \\
&= e^{-2rx} \left\{ c - \frac{\omega^2}{r} x e^{2rx} + \frac{\omega^2}{2r^2} e^{2rx} \right\} \\
&= c e^{-2rx} - \frac{\omega^2}{r} x + \frac{\omega^2}{2r^2}.
\end{aligned}$$

VI.2.5. Ricatti-Differentialgleichung. Sie hat die Form

$$y' = a(x)y + b(x)y^2 + f(x)$$

mit stetigen Funktionen a , b und f . Für diese gDgl gibt es keine allgemeinen Lösungsverfahren. Kennt man allerdings eine partikuläre Lösung $y_p(x)$ der Ricatti-Gleichung, führt der Ansatz

$$v(x) = \frac{1}{y(x) - y_p(x)} \iff y(x) = y_p(x) + \frac{1}{v(x)}$$

weiter. Denn Einsetzen in die Ricatti-Gleichung liefert die gDgl

$$\begin{aligned}
0 &= -y' + a(x)y + b(x)y^2 + f(x) \\
&= -y_p' + \frac{v'}{v^2} + a(x)y_p + \frac{a(x)}{v} + b(x)y_p^2 + 2b(x)\frac{y_p}{v} + \frac{b(x)}{v^2} + f(x) \\
&= \underbrace{-y_p' + a(x)y_p + b(x)y_p^2 + f(x)}_{=0} + \frac{v'}{v^2} + \frac{a(x)}{v} + 2b(x)\frac{y_p}{v} + \frac{b(x)}{v^2}
\end{aligned}$$

und damit

$$v' + [a(x) + 2b(x)y_p(x)]v + b(x) = 0.$$

Diese lineare gDgl kann wie in Abschnitt VI.2.2 gelöst werden.

Zur Bestimmung einer partikulären Lösung y_p hilft manchmal die Beobachtung, dass der Ansatz

$$z(x) = e^{-\int b(x)y(x)dx}$$

auf die homogene lineare gDgl 2. Ordnung

$$z'' - \left(\frac{b'}{b} + a\right)z' + bfz = 0$$

für z führt.

BEISPIEL VI.2.7. Betrachte die Ricatti-Gleichung

$$y' = 4x^2y - xy^2 + 4.$$

Zur Bestimmung einer partikulären Lösung machen wir den Ansatz

$$y_p(x) = \gamma x^\alpha.$$

Einsetzen in die gDgl ergibt

$$\begin{aligned}
\gamma \alpha x^{\alpha-1} &= y_p' \\
&= 4x^2y_p - xy_p^2 + 4
\end{aligned}$$

$$= \underbrace{4\gamma x^{\alpha+2} - \gamma^2 x^{2\alpha+1}}_{=0 \text{ für } \alpha=1, \gamma=4} + 4.$$

Wir haben also Glück und erhalten die partikuläre Lösung

$$y_p(x) = 4x.$$

Die Bestimmungsgleichung für v lautet dann

$$v' + [4x^2 - 2x \cdot 4x]v - x = 0$$

und somit

$$v' - 4x^2v = x.$$

Gemäß Abschnitt VI.2.2 ergibt sich hierfür die allgemeine Lösung

$$v(x) = e^{\frac{4}{3}x^3} \left\{ c + \int e^{-\frac{4}{3}x^3} x dx \right\}.$$

VI.3. Differentialgleichungen 2. Ordnung

VI.3.1. Lineare Differentialgleichungen mit konstanten Koeffizienten. Diese haben die Form

$$y'' + ay' + by = f(x)$$

mit reellen Zahlen a und b und einer stetigen Funktion f auf einem Intervall I . Die gDgl heißt HOMOGEN, falls $f = 0$ ist; ansonsten heißt sie INHOMOGEN.

Da die Vorschrift $y \mapsto y'' + ay' + by$ linear ist, gelten die gleichen Beobachtungen wie in Abschnitt VI.2.2. Daher bestimmt man die allgemeine Lösung der inhomogenen gDgl in zwei Schritten:

- Bestimme die allgemeine Lösung der homogenen gDgl.
- Bestimme eine partikuläre Lösung der inhomogenen gDgl.

Die allgemeine Lösung der inhomogenen gDgl ist dann die Summe der allgemeinen Lösung aus dem ersten Schritt und der partikulären Lösung aus dem zweiten Schritt.

VI.3.2. Die homogene Gleichung. Wir suchen die allgemeine Lösung der gDgl

$$y'' + ay' + by = 0$$

mit konstanten Koeffizienten $a, b \in \mathbb{R}$. Dazu ist es hilfreich, vorübergehend auch komplexe Lösungen, d.h. Funktionen $y : I \rightarrow \mathbb{C}$, zuzulassen. Für $\lambda = \alpha + i\omega \in \mathbb{C}$ betrachten wir insbesondere die Funktion

$$\begin{aligned} y_\lambda(x) &= e^{\lambda x} \\ &= e^{\alpha x} e^{i\omega x} \\ &= e^{\alpha x} \{ \cos(\omega x) + i \sin(\omega x) \}. \end{aligned}$$

Dann ist

$$y'_\lambda(x) = \lambda y_\lambda(x)$$

und daher

$$y_\lambda'' + ay_\lambda' + by_\lambda = (\lambda^2 + a\lambda + b)y_\lambda.$$

Da die Funktion y_λ keine Nullstellen hat, ist dieser Ausdruck genau dann gleich Null, wenn λ eine Nullstelle des Polynoms

$$p(z) = z^2 + az + b$$

ist. Dieses Polynom heißt das CHARAKTERISTISCHE POLYNOM der gDgl.

Wir müssen nun drei Fälle unterscheiden:

- 1. FALL: p HAT ZWEI VERSCHIEDENE REELLE NULLSTELLEN s_1, s_2 : Dann sind

$$\begin{aligned} y_1(x) &= e^{s_1 x} \\ y_2(x) &= e^{s_2 x} \end{aligned}$$

zwei linear unabhängige Lösungen der gDgl. Die allgemeine Lösung lautet

$$y(x) = c_1 e^{s_1 x} + c_2 e^{s_2 x}.$$

- 2. FALL: p HAT ZWEI KOMPLEXE NULLSTELLEN $\lambda_{1,2} = \alpha \pm i\omega$: Dann sind

$$\begin{aligned} Y_1(x) &= e^{\lambda_1 x} \\ Y_2(x) &= e^{\lambda_2 x} \end{aligned}$$

zwei linear unabhängige *komplexe* Lösungen der gDgl. Damit sind

$$\begin{aligned} y_1(x) &= \frac{1}{2} [Y_1(x) + Y_2(x)] \\ &= e^{\alpha x} \cos(\omega x) \\ y_2(x) &= \frac{1}{2i} [Y_1(x) - Y_2(x)] \\ &= e^{\alpha x} \sin(\omega x) \end{aligned}$$

zwei linear unabhängige *reelle* Lösungen der gDgl. Die allgemeine Lösung lautet

$$y(x) = e^{\alpha x} [c_1 \cos(\omega x) + c_2 \sin(\omega x)].$$

- 3. FALL: p HAT EINE DOPPELTE REELLE NULLSTELLE s : Dann ist $p(s) = p'(s) = 0$. Unser Ansatz liefert uns zunächst nur die Lösung

$$y_1(x) = e^{sx}.$$

Uns fehlt eine zweite, linear unabhängige Lösung. Um eine Vorstellung über deren mögliche Form zu erhalten, machen wir

folgendes Gedankenexperiment: Falls p die zwei reellen Nullstellen s und $s + \varepsilon$ hat, ist

$$y_\varepsilon(x) = \frac{e^{(s+\varepsilon)x} - e^{sx}}{\varepsilon}$$

eine Lösung der gDgl. Für $\varepsilon \rightarrow 0$ konvergiert dies gegen

$$y_2(x) = xe^{sx}.$$

Da die Lösungen der gDgl stetig von den Nullstellen von p und diese wiederum stetig von den Koeffizienten a, b abhängen, ist diese Funktion ein „heißer Kandidat“ für die fehlende Lösung der gDgl. Wir setzen daher y_2 in die gDgl ein und erhalten

$$\begin{aligned} & y_2'' + ay_2' + by_2 \\ &= xs^2e^{sx} + 2se^{sx} + a[xse^{sx} + e^{sx}] + bxe^{sx} \\ &= xp(s)e^{sx} + p'(s)e^{sx} \\ &= 0. \end{aligned}$$

Also ist y_2 in der Tat die gesuchte zweite Lösung. Die allgemeine Lösung lautet in diesem Fall daher

$$y(x) = e^{sx}[c_1 + c_2x].$$

Zusammenfassend erhalten wir:

Lösung der homogenen linearen gDgl

$$y'' + ay' + by = 0 :$$

- FALL $a^2 - 4b > 0$: Die allgemeine Lösung ist

$$y(x) = c_1e^{s_1x} + c_2e^{s_2x}$$

mit

$$s_{1,2} = \frac{1}{2}[-a \pm \sqrt{a^2 - 4b}].$$

- FALL $a^2 - 4b < 0$: Die allgemeine Lösung ist

$$y(x) = e^{\alpha x}[c_1 \cos(\omega x) + c_2 \sin(\omega x)]$$

mit

$$\alpha = -\frac{a}{2}$$

$$\omega = \frac{1}{2}\sqrt{4b - a^2}.$$

- FALL $a^2 - 4b = 0$: Die allgemeine Lösung ist

$$y(x) = e^{sx}[c_1 + c_2x].$$

mit

$$s = -\frac{a}{2}.$$

BEISPIEL VI.3.1. Betrachte die gDgl

$$y'' - 6y' + 5y = 0.$$

Das charakteristische Polynom ist

$$\begin{aligned} p(z) &= z^2 - 6z + 5 \\ &= (z - 1)(z - 5). \end{aligned}$$

Die allgemeine Lösung der gDgl lautet

$$y(x) = c_1 e^x + c_2 e^{5x}.$$

BEISPIEL VI.3.2. Betrachte die gDgl

$$y'' - 4y' + 4y = 0.$$

Das charakteristische Polynom ist

$$\begin{aligned} p(z) &= z^2 - 4z + 4 \\ &= (z - 2)^2. \end{aligned}$$

Die allgemeine Lösung der gDgl lautet

$$y(x) = e^{2x}[c_1 + c_2 x].$$

BEISPIEL VI.3.3. Betrachte die gDgl

$$y'' - 6y' + 34y = 0.$$

Das charakteristische Polynom ist

$$\begin{aligned} p(z) &= z^2 - 6z + 34 \\ &= (z - 3)^2 + 25 \end{aligned}$$

mit den komplexen Nullstellen $3 \pm 5i$. Die allgemeine Lösung der gDgl lautet

$$y(x) = e^{3x}[c_1 \cos(5x) + c_2 \sin(5x)].$$

VI.3.3. Die inhomogene Gleichung. Wir suchen eine partikuläre Lösung der inhomogenen gDgl

$$y'' + ay' + by = f(x)$$

mit reellen Zahlen a , b und einer stetigen Funktion f . Die allgemeine Lösung der homogenen Gleichung sei

$$y(x) = c_1 y_1(x) + c_2 y_2(x)$$

mit y_1 und y_2 wie im vorigen Abschnitt. Wir machen nun den Ansatz (VARIATION DER KONSTANTEN)

$$y_p(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$$

mit unbekannt Funktionen $c_1(x)$ und $c_2(x)$. Setzen wir y_p in die linke Seite der gDgl ein, erhalten wir

$$\begin{aligned} y_p'' + ay_p' + by_p &= c_1'' y_1 + 2c_1' y_1' + c_1 y_1'' + a[c_1' y_1 + c_1 y_1'] + bc_1 y_1 \\ &\quad + c_2'' y_2 + 2c_2' y_2' + c_2 y_2'' + a[c_2' y_2 + c_2 y_2'] + bc_2 y_2 \end{aligned}$$

$$\begin{aligned}
&= c_1 \underbrace{[y_1'' + ay_1' + by_1]}_{=0} + c_2 \underbrace{[y_2'' + ay_2' + by_2]}_{=0} \\
&\quad + a[c_1' y_1 + c_2' y_2] + c_1' y_1' + c_2' y_2' \\
&\quad + \underbrace{[c_1'' y_1 + c_1' y_1' + c_2'' y_2 + c_2' y_2']}_{=[c_1' y_1 + c_2' y_2]'} \\
&= c_1' y_1' + c_2' y_2' + a[c_1' y_1 + c_2' y_2] + [c_1' y_1 + c_2' y_2]'.
\end{aligned}$$

Hieran erkennen wir, dass y_p eine Lösung der inhomogenen gDgl ist, falls die unbekannt Funktionen c_1 und c_2 so bestimmt werden können, dass gilt

$$\begin{aligned}
c_1' y_1 + c_2' y_2 &= 0 \\
c_1' y_1' + c_2' y_2' &= f.
\end{aligned}$$

Dies ist ein lineares Gleichungssystem für die Unbekannten c_1' und c_2' mit der Koeffizientenmatrix

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}.$$

Es ist eindeutig lösbar, wenn die Determinante

$$W = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$$

stets ungleich Null ist. W heißt die WRONSKI-DETERMINANTE der gDgl.

Wir nehmen für einen Augenblick an, dass $W(x) \neq 0$ ist für alle x . Dann folgt aus der Cramerschen Regel aus Abschnitt II.3.4

$$\begin{aligned}
c_1' &= -\frac{1}{W} y_2 f \\
c_2' &= \frac{1}{W} y_1 f.
\end{aligned}$$

Da die rechten Seiten dieser Gleichungen stetige Funktionen sind, erhalten wir die gesuchten Funktionen c_1 und c_2 zu

$$\begin{aligned}
c_1 &= -\int \frac{y_2(x) f(x)}{W(x)} dx \\
c_2 &= \int \frac{y_1(x) f(x)}{W(x)} dx.
\end{aligned}$$

Wir müssen noch nachweisen, dass die Wronski Determinante tatsächlich niemals verschwindet. Dazu leiten wir zunächst eine gDgl für W her:

$$\begin{aligned}
W' &= [y_1 y_2' - y_1' y_2]' \\
&= y_1' y_2' + y_1 y_2'' - y_1'' y_2 - y_1' y_2' \\
&= y_1 y_2'' - y_1'' y_2
\end{aligned}$$

$$\begin{aligned}
&= y_1[-ay'_2 - by_2] - [-ay'_1 - by_1]y_2 \\
&= -ay_1y'_2 - by_1y_2 + ay'_1y_2 + by_1y_2 \\
&= -a[y_1y'_2 - y'_1y_2] \\
&= -aW.
\end{aligned}$$

Gemäß Abschnitt VI.2.1 ist daher

$$W(x) = W(0)e^{-ax} \quad \text{für alle } x \in \mathbb{R}.$$

Wir müssen also nur noch nachweisen, dass in allen drei möglichen Fällen $W(0) \neq 0$ ist:

- FALL $a^2 - 4b > 0$:

$$\begin{aligned}
W(0) &= \pm \det \begin{pmatrix} 1 & 1 \\ \frac{1}{2}[-a - \sqrt{a^2 - 4b}] & \frac{1}{2}[-a + \sqrt{a^2 - 4b}] \end{pmatrix} \\
&= \pm \sqrt{a^2 - 4b} \\
&\neq 0.
\end{aligned}$$

- FALL $a^2 - 4b < 0$:

$$\begin{aligned}
W(0) &= \pm \det \begin{pmatrix} 1 & 0 \\ -\frac{a}{2} & \frac{1}{2}\sqrt{4b - a^2} \end{pmatrix} \\
&= \pm \frac{1}{2}\sqrt{4b - a^2} \\
&\neq 0.
\end{aligned}$$

- FALL $a^2 - 4b = 0$:

$$\begin{aligned}
W(0) &= \pm \det \begin{pmatrix} 1 & 0 \\ -\frac{a}{2} & 1 \end{pmatrix} \\
&= \pm 1.
\end{aligned}$$

Insgesamt erhalten wir:

Eine partikuläre Lösung der inhomogenen linearen gDgl

$$y'' + ay' + by = f(x)$$

ist

$$\begin{aligned}
y_p(x) &= -y_1(x) \int \frac{1}{W_0} e^{ax} y_2(x) f(x) dx \\
&\quad + y_2(x) \int \frac{1}{W_0} e^{ax} y_1(x) f(x) dx
\end{aligned}$$

mit y_1 und y_2 aus Abschnitt VI.3.2 und

$$W_0 = \det \begin{pmatrix} y_1(0) & y_2(0) \\ y'_1(0) & y'_2(0) \end{pmatrix}.$$

BEISPIEL VI.3.4. Betrachte die gDgl

$$y'' - 6y' + 5y = x.$$

Laut Beispiel VI.3.1 ist

$$y_1(x) = e^x$$

und

$$y_2(x) = e^{5x}.$$

Wegen

$$\begin{aligned} W_0 &= \sqrt{36 - 4 \cdot 5} \\ &= 4 \end{aligned}$$

ist

$$\begin{aligned} y_p(x) &= -\frac{1}{4}e^x \underbrace{\int e^{-6x} e^{5x} x dx}_{=-e^{-x}(x+1)} + \frac{1}{4}e^{5x} \underbrace{\int e^{-6x} e^x x dx}_{=-e^{-5x}(\frac{1}{5}x + \frac{1}{25})} \\ &= \frac{1}{4} \left[x + 1 - \frac{1}{5}x - \frac{1}{25} \right] \\ &= \frac{1}{5}x + \frac{6}{25}. \end{aligned}$$

BEISPIEL VI.3.5. Betrachte die gDgl

$$y'' - 4y' + 4y = x^2.$$

Laut Beispiel VI.3.2 ist

$$y_1(x) = e^{2x}$$

und

$$y_2(x) = xe^{2x}.$$

Wegen

$$W_0 = 1$$

ist

$$\begin{aligned} y_p(x) &= -e^{2x} \underbrace{\int e^{-4x} x e^{2x} x^2 dx}_{=-e^{-2x}[\frac{1}{2}x^3 + \frac{3}{4}x^2 + \frac{3}{4}x + \frac{3}{8}]} + x e^{2x} \underbrace{\int e^{-4x} e^{2x} x^2 dx}_{=-e^{-2x}[\frac{1}{2}x^2 + \frac{1}{2}x + \frac{1}{4}]} \\ &= \frac{1}{4}x^2 + \frac{1}{2}x + \frac{3}{8}. \end{aligned}$$

BEISPIEL VI.3.6. Betrachte die gDgl

$$y'' - 6y' + 34y = \sin(2x).$$

Laut Beispiel VI.3.3 ist

$$y_1(x) = e^{3x} \cos(5x)$$

und

$$y_2(x) = e^{3x} \sin(5x).$$

Wegen

$$\begin{aligned} W_0 &= \frac{1}{2} \sqrt{4 \cdot 34 - 36} \\ &= 5 \end{aligned}$$

ist

$$\begin{aligned} y_p(x) &= -\frac{1}{5} e^{3x} \cos(5x) \int e^{-6x} e^{3x} \sin(5x) \sin(2x) dx \\ &\quad + \frac{1}{5} e^{3x} \sin(5x) \int e^{-6x} e^{3x} \cos(5x) \sin(2x) dx. \end{aligned}$$

Wegen

$$\begin{aligned} \int e^{ax} \sin(\omega x) dx &= \frac{1}{a^2 + \omega^2} e^{ax} [a \sin(\omega x) - \omega \cos(\omega x)] \\ \int e^{ax} \cos(\omega x) dx &= \frac{1}{a^2 + \omega^2} e^{ax} [a \cos(\omega x) + \omega \sin(\omega x)] \end{aligned}$$

und

$$\begin{aligned} \sin(5x) \sin(2x) &= \frac{1}{2} [\cos(3x) - \cos(7x)] \\ \cos(5x) \sin(2x) &= \frac{1}{2} [\sin(7x) - \sin(3x)] \end{aligned}$$

erhalten wir

$$\begin{aligned} y_p(x) &= -\frac{1}{5} e^{3x} \cos(5x) \frac{1}{2} \int (e^{-3x} \cos(3x) - e^{-3x} \cos(7x)) dx \\ &\quad + \frac{1}{5} e^{3x} \sin(5x) \frac{1}{2} \int (e^{-3x} \sin(7x) - e^{-3x} \sin(3x)) dx \\ &= \frac{1}{10} \left\{ -\cos(5x) \left[-\frac{3}{18} \cos(3x) + \frac{3}{18} \sin(3x) \right. \right. \\ &\quad \left. \left. + \frac{3}{58} \cos(7x) - \frac{7}{58} \sin(7x) \right] \right. \\ &\quad \left. + \sin(5x) \left[-\frac{3}{58} \sin(7x) - \frac{7}{58} \cos(7x) \right] \right\} \end{aligned}$$

$$\begin{aligned}
& \left. + \frac{3}{18} \sin(3x) + \frac{3}{18} \cos(3x) \right\} \\
= & \frac{1}{10} \left\{ \frac{3}{18} \underbrace{[\cos(5x) \cos(3x) + \sin(5x) \sin(3x)]}_{=\cos(2x)} \right. \\
& + \frac{3}{18} \underbrace{[\sin(5x) \cos(3x) - \cos(5x) \sin(3x)]}_{=\sin(2x)} \\
& - \frac{3}{58} \underbrace{[\cos(7x) \cos(5x) + \sin(7x) \sin(5x)]}_{=\cos(2x)} \\
& \left. + \frac{7}{58} \underbrace{[\sin(7x) \cos(5x) - \sin(5x) \cos(7x)]}_{=\sin(2x)} \right\} \\
= & \frac{1}{10} \left\{ \frac{10}{87} \cos(2x) + \frac{25}{87} \sin(2x) \right\} \\
= & \frac{1}{174} \{2 \cos(2x) + 5 \sin(2x)\} \\
= & \frac{\sqrt{29}}{174} \left\{ \underbrace{\frac{2}{\sqrt{29}} \cos(2x)}_{=\sin(\varphi)} + \underbrace{\frac{5}{\sqrt{29}} \sin(2x)}_{=\cos(\varphi)} \right\} \\
= & \frac{\sqrt{29}}{174} \sin(2x + \varphi)
\end{aligned}$$

mit

$$\tan \varphi = \frac{2}{5}.$$

Die partikuläre Lösung schwingt also phasenverschoben mit der gleichen Frequenz wie die Anregung f .

VI.3.4. Differentialgleichungen vom Typ $y'' = f(x, y')$. Der Ansatz $u = y'$ führt auf die gDgl 1. Ordnung

$$u' = f(x, u).$$

Diese kann mit den Methoden von Paragraph VI.2 gelöst werden. Integration von u liefert dann die Lösung y .

BEISPIEL VI.3.7. Betrachte die gDgl

$$y'' = 2xy'.$$

Für $u = y'$ ergibt sich

$$u' = 2xu.$$

Gemäß Abschnitt VI.2.1 erhalten wir

$$\ln u = \int \frac{1}{u} du$$

$$\begin{aligned}
 &= \int 2x dx \\
 &= x^2 + \tilde{c}
 \end{aligned}$$

und somit

$$\begin{aligned}
 u &= e^{x^2 + \tilde{c}} \\
 &= ce^{x^2}.
 \end{aligned}$$

Damit ergibt sich

$$y(x) = c_2 + \int c_1 e^{x^2} dx$$

mit $c_1, c_2 \in \mathbb{R}$.

VI.3.5. Differentialgleichungen vom Typ $y'' = f(y, y')$. Wir bestimmen y' als Funktion von y , d.h.

$$y' = v(y).$$

Mit der Kettenregel ergibt sich

$$\begin{aligned}
 y'' &= (v(y))' \\
 &= \frac{dv}{dy} y' \\
 &= v \frac{dv}{dy}
 \end{aligned}$$

und damit

$$\frac{dv}{dy} = \frac{1}{v} f(y, v(y)).$$

Diese gDgl 1. Ordnung kann mit den Methoden von Paragraph VI.2 gelöst werden. Anschließend wird y mit den gleichen Methoden aus der gDgl

$$y' = v(y)$$

bestimmt.

BEISPIEL VI.3.8. Betrachte die gDgl

$$y'' = -\frac{(y')^2}{5y}.$$

Offensichtlich ist jede konstante Funktion eine Lösung. Die gDgl für v lautet

$$\begin{aligned}
 \frac{dv}{dy} &= \frac{1}{v} \left[-\frac{v^2}{5y} \right] \\
 &= -\frac{v}{5y}.
 \end{aligned}$$

Mit Trennung der Variablen ergibt sich

$$\begin{aligned}\ln v &= \int \frac{1}{v} dv \\ &= - \int \frac{1}{5y} dy \\ &= -\frac{1}{5} \ln y + c.\end{aligned}$$

und somit

$$\begin{aligned}v &= e^{-\frac{1}{5} \ln y + c} \\ &= c_1 y^{-\frac{1}{5}}.\end{aligned}$$

Damit lautet die gDgl für y :

$$y' = c_1 y^{-\frac{1}{5}}.$$

Nochmalige Trennung der Variablen liefert

$$\begin{aligned}\frac{5}{6} y^{\frac{6}{5}} &= \int y^{\frac{1}{5}} dy \\ &= \int c_1 dx \\ &= c_1 x + c_2.\end{aligned}$$

Also ist

$$y = \left[\frac{6}{5} (c_1 x + c_2) \right]^{\frac{5}{6}}.$$

VI.4. Numerische Lösung gewöhnlicher Differentialgleichungen*

VI.4.1. Motivation*. Viele der in der Praxis auftretenden gewöhnlichen Differentialgleichungen können nicht analytisch gelöst werden. Stattdessen muss man ihre Lösung numerisch approximieren. In diesem Paragraphen wollen wir einen kurzen Überblick über die hierfür gebräuchlichsten Verfahren und die bei ihrer Anwendung zu berücksichtigenden Aspekte geben.

Für die Motivation beschränken wir uns auf den einfachsten Fall und betrachten das Anfangswertproblem

$$\begin{aligned}y' &= f(t, y(t)) \\ y(t_0) &= y_0\end{aligned}$$

mit einer hinreichend glatten Funktion f . Für die numerische Lösung führen wir Gitterpunkte $t_0 < t_1 < \dots$ ein und bezeichnen die numerische Approximation für $y(t_i)$ mit η_i oder $\eta(t_i, h_i)$. Dabei bezeichnet

$h_i = t_i - t_{i-1}$ die Schrittweite. Im Folgenden betrachten wir der Einfachheit halber nur äquidistante Gitterpunkte, d.h. $h_i = h$ für alle i . In der Praxis muss man aber veränderliche Schrittweiten h_i vorsehen und diese mittels Schrittweitenkontrolle zusammen mit der Lösung adaptiv bestimmen.

Angenommen wir kennen die Lösung y des AWP im Punkte t . Wegen $y'(t) = f(t, y(t))$ gibt $f(t, y(t))$ die Steigung der Tangente an die Lösungskurve im Punkt t an. Daher sollte $y(t) + hf(t, y(t))$ eine passable Näherung an $y(t+h)$ sein. Diese Überlegung führt auf:

EXPLIZITES EULERVERFAHREN:

$$\begin{aligned}\eta_0 &= y_0 \\ \eta_{i+1} &= \eta_i + hf(t_i, \eta_i) \\ t_{i+1} &= t_i + h.\end{aligned}$$

Genauso gut hätten wir eine Näherung η_{i+1} für $y(t+h)$ aus der Bedingung herleiten können, dass die Tangente durch die Lösungskurve im Punkte $t_{i+1} = t_i + h$ durch den Punkt (t_i, y_i) gehen soll, d.h. $y_i = \eta_{i+1} - hf(t_i + h, \eta_{i+1})$. Dies führt auf:

IMPLIZITES EULERVERFAHREN:

$$\begin{aligned}\eta_0 &= y_0 \\ \eta_{i+1} &= \eta_i + hf(t_{i+1}, \eta_{i+1}) \\ t_{i+1} &= t_i + h.\end{aligned}$$

Im Gegensatz zum expliziten Eulerverfahren muss man beim impliziten Eulerverfahren in jedem Schritt eine nicht lineare Gleichung der Form $u = z + hf(t, u)$ mit bekannten Größen z und t lösen. Falls f (als Funktion von u bei festem t) differenzierbar ist, folgt aus dem Satz über die Umkehrfunktion, Abschnitt IV.3.1, dass diese Gleichung für hinreichend kleines h eine Lösung in der Nähe von z hat. Diese Lösung kann mit wenigen Iterationen des Newtonverfahrens aus Abschnitt IV.2.6 berechnet werden.

Ein drittes Verfahren erhalten wir durch folgende Überlegung: Für die exakte Lösung des AWP gilt

$$\begin{aligned}y(t+h) &= y(t) + \int_t^{t+h} y'(s) ds \\ &= y(t) + \int_t^{t+h} f(s, y(s)) ds.\end{aligned}$$

Das Integral wird beim expliziten Eulerverfahren durch $hf(t, y(t))$ approximiert und beim impliziten Eulerverfahren durch $hf(t+h, y(t+h))$.

Ebenso könnten wir es aber auch durch die Trapezregel

$$\int_t^{t+h} f(s, y(s)) ds \approx \frac{h}{2} [f(t, y(t)) + f(t+h, y(t+h))]$$

aus Abschnitt V.6.1 annähern. Dies führt auf:

TRAPEZREGEL oder
VERFAHREN VON CRANK-NICOLSON:

$$\begin{aligned} \eta_0 &= y_0 \\ \eta_{i+1} &= \eta_i + \frac{h}{2} [f(t_i, \eta_i) + f(t_{i+1}, \eta_{i+1})] \\ t_{i+1} &= t_i + h. \end{aligned}$$

Wie beim impliziten Eulerverfahren ist in jedem Schritt eine nicht lineare Gleichung zu lösen. Wieder ist dies für hinreichend kleines h möglich und kann mit wenigen Schritten des Newtonverfahrens aus Abschnitt IV.2.6 geschehen.

Das folgende Java-Programm realisiert die drei beschriebenen Verfahren. Dabei steuert der Parameter `theta` die Verfahrenswahl:

$$\text{theta} = \begin{cases} 0 & \text{explizites Eulerverfahren,} \\ 1 & \text{implizites Eulerverfahren,} \\ \frac{1}{2} & \text{Crank-Nicolson-Verfahren.} \end{cases}$$

Die Methode `gaussElimination` löst ein lineares Gleichungssystem mit dem Gaußschen Eliminationsverfahren und ist in Abschnitt II.1.4 wiedergegeben. `force` ist eine abstrakte Klasse, die eine differenzierbare Funktion mehrerer Veränderlicher realisiert und unter `f` bzw. `df` den Funktionswert bzw. die Ableitung bereitstellt.

```
// linear single step method
public void lssm( double theta ) throws LinearAlgebraException {
    double[] ff = new double[dim];
    double alpha = -theta*dt;
    double beta = (1.0 - theta)*dt;
    t = initialTime;
    copy(eta, x0);
    for( int step = 1; step <= steps; step++ ) {
        ff = add(eta, force.f(t, eta), 1.0, beta);
        t += dt;
        eta = ivpNewton(alpha, t, ff);
    }
} // end of lssm

// Newton method for the solution of equations of the form z + a*f(s,z) = g
public double[] ivpNewton(double a, double s, double[] g)
    throws LinearAlgebraException {
    double[] y = new double[dim]; // solution
    double[] dy = new double[dim]; // increment
    double[][] jf = new double[dim][dim]; // Jacobian
    double resf; // norm of f
```

```

double resy; // norm of dy
copy(y, g);
int iter = 0;
b = add(g, force.f(s, y), 1.0, -a);
accumulate(b, y, -1.0);
resf = Math.sqrt( innerProduct(b, b)/dim );
resy = 1.0;
while( iter < MAXIT && resf > TOL && resy > TOL ) {
    jf = force.df(s, y);
    for( int i = 0; i < dim; i++ )
        for( int j = 0; j < dim; j++ ) {
            a[i][j] = a*jf[i][j];
            if( j == i ) a[i][j] += 1.0;
        }
    gaussElimination();
    copy(dy, x);
    resy = Math.sqrt( innerProduct(dy,dy)/dim );
    accumulate(y, dy, 1.0);
    b = add(g, force.f(s, y), 1.0, -a);
    accumulate(b, y, -1.0);
    resf = Math.sqrt( innerProduct(b, b)/dim );
    iter++;
}
return y;
} // end of ivpNewton
// copy vector v to vector u
public void copy(double[] u, double[] v) {
    for( int i = 0; i < dim; i++ )
        u[i] = v[i];
} // end of copy
// add s times vector v to vector u
public void accumulate(double[] u, double[] v, double s) {
    for( int i = 0; i < dim; i++ )
        u[i] += s*v[i];
} // end of accumulate
// return s*u+t*v
public double[] add(double[] u, double[] v, double s, double t) {
    double[] w = new double[dim];
    for( int i = 0; i < dim; i++ )
        w[i] = s*u[i] + t*v[i];
    return w;
} // end of add
// inner product of two vectors
public double innerProduct(double[] u, double[] v) {
    double prod = 0;
    for( int i = 0; i < dim; i++ )
        prod += u[i]*v[i];
    return prod;
} // end of inner product

```

VI.4.2. Einschrittverfahren*. Bei den drei Verfahren des vorigen Abschnittes wird die Näherung η_{i+1} für $y(t_{i+1})$ ausschließlich durch die letzte Näherung η_i für $y(t_i)$ bestimmt. Daher spricht man von einem EINSCHRITTVERFAHREN (kurz ESV). Die allgemeine Definition eines ESV lautet:

EINSCHRITTVERFAHREN:

$$\begin{aligned}\eta_0 &= y_0 \\ \eta_{i+1} &= \eta_i + h\Phi(t_i, \eta_i, h; f) \\ t_{i+1} &= t_i + h.\end{aligned}$$

Die Funktion Φ heißt die VERFAHRENSFUNKTION des ESV.

Für das explizite Eulerverfahren ist offensichtlich

$$\Phi(t, y, h; f) = f(t, y).$$

Die Verfahrensfunktionen für das implizite Eulerverfahren und die Trapezregel sind komplizierter. Für hinreichend kleines h erhalten wir für das implizite Eulerverfahren

$$\Phi(t, y, h; f) = \frac{1}{h} \left[[Id - hf(t+h, \cdot)]^{-1} y - y \right]$$

und für die Trapezregel

$$\Phi(x, y, h; f) = \frac{1}{h} \left[[Id - \frac{h}{2} f(t+h, \cdot)]^{-1} \left[y + \frac{h}{2} f(t, y) \right] - y \right],$$

wobei g^{-1} die Umkehrfunktion der Funktion g bezeichnet. Man beachte, dass diese Darstellung der Verfahrensfunktion nur für die theoretische Analyse benötigt wird. Für die praktische Rechnung benutzt man die Darstellung der Verfahren aus dem vorigen Abschnitt.

Die Qualität eines ESV wird durch den LOKALEN VERFAHRENSFEHLER gemessen. Dieser ist definiert durch

$$\tau(x, y, h) = \frac{1}{h} \{ z(x+h) - y \} - \Phi(x, y, h; f) \quad , h > 0.$$

Dabei ist z die Lösung des AWP

$$\begin{aligned}z' &= f(t, z(t)), \\ z(x) &= y.\end{aligned}$$

Ein ESV hat die ORDNUNG p , falls für alle hinreichend glatten Funktionen f der Ausdruck $h^{-p}\tau(x, y, h)$ für $h \rightarrow 0$ beschränkt bleibt. Hierfür schreibt man auch $\tau(x, y, h) = O(h^p)$.

Der lokale Verfahrensfehler beschreibt die Fehlerfortpflanzung in einem einzelnen Schritt eines ESV. Man kann aber zeigen, dass er auch den globalen Fehler, d.h. $|y(t_i) - \eta_i|$ für alle i , beschreibt:

Hat das ESV die Ordnung p , gilt für alle $t \neq t_0$ und alle $n \in \mathbb{N}^*$

$$\begin{aligned} & \max_{0 \leq i \leq n} \left| y\left(t_0 + i \frac{t-t_0}{n}\right) - \eta\left(t_0 + i \frac{t-t_0}{n}, \frac{t-t_0}{n}\right) \right| \\ & \leq c(t, f) \left(\frac{|t-t_0|}{n} \right)^p. \end{aligned}$$

Für das explizite Eulerverfahren erhalten wir durch Taylorentwicklung (vgl. Abschnitt VII.3.1)

$$\begin{aligned} \tau(x, y, h) &= \frac{1}{h} \{z(x+h) - y\} - f(x, y) \\ &= z'(x) + \frac{1}{2} h z''(x) + 0(h^2) - \underbrace{f(x, y)}_{=z'(x)} \\ &= \frac{1}{2} h z''(x) + 0(h^2). \end{aligned}$$

Also hat das explizite Eulerverfahren die Ordnung 1. Mit etwas mehr Aufwand kann man zeigen, dass das implizite Eulerverfahren ebenfalls die Ordnung 1 hat und dass das Verfahren von Crank-Nicolson die Ordnung 2 hat.

VI.4.3. Runge-Kutta Verfahren*. Viele der für die Praxis relevanten Einschrittverfahren gehören dieser Verfahrensklasse an. Die allgemeine Form lautet:

RUNGE-KUTTA-VERFAHREN:

$$\begin{aligned} \eta_0 &= y_0 \\ \eta_{i,j} &= \eta_i + h \sum_{k=1}^r a_{jk} f(t_i + c_k h, \eta_{i,k}), \quad 1 \leq j \leq r \\ \eta_{i+1} &= \eta_i + h \sum_{k=1}^r b_k f(t_i + c_k h, \eta_{i,k}) \\ t_{i+1} &= t_i + h. \end{aligned}$$

Dabei ist $0 \leq c_1 \leq \dots \leq c_r \leq 1$. Die Zahl r heißt STUFE des Runge-Kutta-Verfahrens. Das Verfahren heißt EXPLIZIT, wenn $a_{jk} = 0$ ist für alle $k \geq j$; ansonsten heißt es IMPLIZIT. Implizite Verfahren sind aufwändiger als explizite Verfahren, da bei ihnen in jedem Schritt nicht lineare Gleichungssysteme gelöst werden müssen. Dafür haben sie bessere Stabilitätseigenschaften und sind insgesamt den expliziten Verfahren überlegen (s. Abschnitt VI.4.4).

Der Übersichtlichkeit halber fasst man die Zahlen c_k , a_{jk} , b_k in einem Schema der Form

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_r & a_{r1} & a_{r2} & \dots & a_{rr} \\ \hline & b_1 & b_2 & \dots & b_r \end{array}$$

zusammen.

Die Verfahren aus Abschnitt VI.4.1 sind alle Runge-Kutta-Verfahren. Dem expliziten Eulerverfahren entspricht das Schema

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array} \quad r = 1.$$

Dem impliziten Eulerverfahren entspricht das Schema

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array} \quad r = 1.$$

Dem Verfahren von Crank-Nicolson entspricht schließlich das Schema

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \quad r = 2.$$

Das sog. KLASSISCHE RUNGE-KUTTA-VERFAHREN ist schließlich gegeben durch das Schema

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} \end{array} \quad r = 4.$$

Die entsprechende Verfahrensvorschrift lautet:

$$\begin{aligned} \eta_0 &= y_0 \\ \eta_{i,1} &= \eta_i \\ \eta_{i,2} &= \eta_i + \frac{h}{2} f(t_i, \eta_{i,1}) \\ \eta_{i,3} &= \eta_i + \frac{h}{2} f\left(t_i + \frac{h}{2}, \eta_{i,2}\right) \\ \eta_{i,4} &= \eta_i + h f\left(t_i + \frac{h}{2}, \eta_{i,3}\right) \\ \eta_{i+1} &= \eta_i + \frac{h}{6} \left\{ f(t_i, \eta_{i,1}) + 2f\left(t_i + \frac{h}{2}, \eta_{i,2}\right) \right. \\ &\quad \left. + 2f\left(t_i + \frac{h}{2}, \eta_{i,3}\right) + f(t_i + h, \eta_{i,4}) \right\} \end{aligned}$$

$$t_{i+1} = t_i + h.$$

Es hat die Stufe 4 und die Ordnung 4. Als explizites Verfahren hat es aber keine besonders guten Stabilitätseigenschaften.

Wegen ihrer hohen Ordnung und guten Stabilitätseigenschaften sind die STARK DIAGONAL IMPLIZITEN RUNGE-KUTTA-VERFAHREN (kurz SDIRK-VERFAHREN) für die praktische Rechnung besonders gut geeignet. Die einfachsten Verfahren dieser Klasse haben die Stufe 2 und die Ordnung 3 bzw. die Stufe 5 und die Ordnung 4. Sie sind bestimmt durch die Schemata

$$\begin{array}{c|cc} \frac{3+\sqrt{3}}{6} & \frac{3+\sqrt{3}}{6} & 0 \\ \frac{3-\sqrt{3}}{6} & -\frac{\sqrt{3}}{3} & \frac{3+\sqrt{3}}{6} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \quad \text{SDIRK2.}$$

und

$$\begin{array}{c|ccccc} \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{3}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ \frac{11}{20} & \frac{17}{50} & -\frac{1}{25} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & \frac{371}{1360} & -\frac{137}{2720} & \frac{15}{544} & \frac{1}{4} & 0 \\ 1 & \frac{25}{24} & -\frac{49}{48} & \frac{125}{16} & -\frac{85}{12} & \frac{1}{4} \\ \hline & \frac{25}{24} & -\frac{49}{48} & \frac{125}{16} & -\frac{85}{12} & \frac{1}{4} \end{array} \quad \text{SDIRK5.}$$

Damit lautet die Verfahrensvorschrift z. B. für das SDIRK2-Verfahren

$$\begin{aligned} \eta_0 &= y_0 \\ \eta_{i,1} &= \eta_i + \frac{3 + \sqrt{3}}{6} h f\left(t_i + \frac{3 + \sqrt{3}}{6} h, \eta_{i,1}\right) \\ \eta_{i,2} &= \eta_i - \frac{\sqrt{3}}{3} h f\left(t_i + \frac{3 + \sqrt{3}}{6} h, \eta_{i,1}\right) \\ &\quad + \frac{3 + \sqrt{3}}{6} h f\left(t_i + \frac{3 - \sqrt{3}}{6} h, \eta_{i,2}\right) \\ \eta_{i+1} &= \eta_i + \frac{h}{2} \left\{ f\left(t_i + \frac{3 + \sqrt{3}}{6} h, \eta_{i,1}\right) + f\left(t_i + \frac{3 - \sqrt{3}}{6} h, \eta_{i,2}\right) \right\} \\ t_{i+1} &= t_i + h. \end{aligned}$$

Das folgende Java-Programm realisiert ein allgemeines SDIRK-Verfahren. Dabei stellt die Klasse `RKParameters` die nötigen Parameter bereit. Insbesondere liefert `rk.d` die als konstant angenommene Diagonale der Matrix $(a_{ij})_{1 \leq i, j \leq r}$. Die Methoden `accumulate`, `copy` und `ivpNewton` sind bereits oben wiedergegeben. `force` hat die gleiche Bedeutung wie oben.

```
// Runge-Kutta method
public void sdirk( int type ) throws LinearAlgebraException {
    RKParameters rk = new RKParameters(type);
```

```

// auxiliary quantities for runge-kutta levels
double[] [] rkf = new double[rk.l][dim]; // function values
double[] [] rketa = new double[rk.l][dim]; // intermediate eta's
double[] tt = new double[rk.l]; // intermediate times
// end of runge-kutta declarations
double[] ff = new double[dim];
double alpha = -rk.d*dt;
t = initialTime;
copy(eta, x0);
for( int step = 1; step <= steps; step++ ) {
  for( int j = 0; j < rk.l; j++ )
    tt[j] = t + rk.c[j]*dt;
  int jj = 0;
  for( int j = 0; j < rk.l; j++ ) {
    copy(ff, eta);
    for( int k = 0; k < j; k++ ) {
      accumulate(ff, rkf[k], rk.a[jj]*dt);
      jj++;
    }
    rketa[j] = ivpNewton(alpha, tt[j], ff);
    rkf[j] = force.f(tt[j], rketa[j]);
  }
  t += dt;
  for( int j = 0; j < rk.l; j++ )
    accumulate(eta, rkf[j], rk.b[j]*dt);
}
} // end of sdirk

```

VI.4.4. Stabilität*. Das explizite und das implizite Eulerverfahren haben beide die Ordnung 1; für $h \rightarrow 0$ verhält sich der Fehler bei beiden Verfahren gleich. Das implizite Verfahren ist aber aufwändiger als das explizite Verfahren, da in jedem Schritt eine nicht lineare Gleichung gelöst werden muss. Macht sich dieser erhöhte Aufwand denn überhaupt bezahlt?

Um diese Frage zu beantworten, betrachten wir das AWP

$$y' = -\lambda y$$

$$y(0) = 1$$

mit $\lambda \gg 1$. Die exakte Lösung ist $y(t) = e^{-\lambda t}$ und klingt sehr schnell ab.

Wir wenden auf dieses AWP zuerst das explizite Eulerverfahren mit $\eta_0 = y_0 = 1$ an und erhalten

$$\eta_{i+1} = \eta_i - h\lambda\eta_i$$

$$\implies \eta_i = (1 - h\lambda)^i \quad \text{für alle } i.$$

Die η_i konvergieren für $i \rightarrow \infty$ offensichtlich genau dann gegen Null, wenn gilt

$$|1 - h\lambda| < 1 \quad \iff \quad h < \frac{2}{\lambda}.$$

Also gilt:

Die numerische Lösung des expliziten Verfahrens hat nur dann das gleiche qualitative Verhalten wie die exakte Lösung, wenn die Schrittweite h hinreichend klein ist.

Wir betrachten nun das implizite Eulerverfahren mit $\eta_0 = y_0 = 1$ und erhalten

$$\begin{aligned} \eta_{i+1} &= \eta_i - h\lambda\eta_{i+1} \\ \implies \eta_i &= \left(\frac{1}{1+h\lambda} \right)^i \quad \text{für alle } i. \end{aligned}$$

Wegen $\lambda > 0$ konvergieren die η_i für $i \rightarrow \infty$ jetzt für jede Schrittweite h gegen Null, d.h.:

Das implizite Verfahren liefert für jede Schrittweite eine numerische Lösung, die das gleiche qualitative Verhalten hat wie die exakte Lösung.

Der Mehraufwand für das implizite Verfahren macht sich also bezahlt!

Dieses Phänomen nennt man STABILITÄT. Um es genauer zu beschreiben, wenden wir ein beliebiges Runge-Kutta-Verfahren auf das obige AWP an. Mit ein wenig Rechnung sieht man dann, dass die numerische Lösung die Form hat

$$\eta_i = g(h\lambda)^i$$

mit

$$g(z) = 1 + zb^T(I - zA)^{-1}e$$

und

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix}, \quad e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Die Menge

$$S = \{ \mu \in \mathbb{C} : I - \mu A \text{ ist regulär und } |g(\mu)| = |1 + \mu b^T(I - \mu A)^{-1}e| \leq 1 \}$$

heißt das STABILITÄTSGEBIET des Verfahrens.

BEISPIEL VI.4.1. Für das explizite Eulerverfahren ist

$$g(z) = z + 1$$

und damit

$$S = \{ \mu \in \mathbb{C} : |\mu + 1| \leq 1 \}.$$

Für das implizite Eulerverfahren ist

$$g(z) = \frac{1}{1-z}$$

und damit

$$\begin{aligned} S &= \{\mu \in \mathbb{C} : \frac{1}{|1-\mu|} \leq 1\} \\ &= \{\mu \in \mathbb{C} : |1-\mu| \geq 1\}. \end{aligned}$$

Für das Verfahren von Crank-Nicolson erhalten wir schließlich

$$g(z) = \frac{2+z}{2-z}$$

und

$$\begin{aligned} S &= \{\mu \in \mathbb{C} : |\frac{2+\mu}{2-\mu}| \leq 1\} \\ &= \{\mu \in \mathbb{C} : \operatorname{Re} \mu \leq 0\}. \end{aligned}$$

Abbildung VI.4.1 zeigt die Stabilitätsgebiete dieser drei Verfahren.

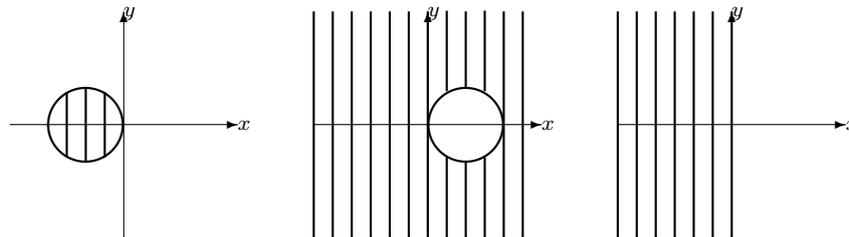


ABBILDUNG VI.4.1. Stabilitätsgebiete des expliziten und impliziten Eulerverfahrens und des Crank-Nicolson-Verfahrens (v.l.n.r.)

Je größer das Stabilitätsgebiet ist, um so größere Schrittweiten können verwendet werden. Im Idealfall ist $S \supset \{z \in \mathbb{C} : \operatorname{Re} z \leq 0\}$. Verfahren, die dieses Kriterium erfüllen, nennt man A-STABIL. Bei solchen Verfahren erhält man bei obigem AWP für jede Schrittweite eine qualitativ richtige numerische Lösung. Das implizite Eulerverfahren, das Crank-Nicolson-Verfahren und die in Abschnitt VI.4.3 genannten SDIRK-Verfahren sind A-stabil.

Bei expliziten Runge-Kutta-Verfahren ist die Funktion g ein Polynom. Da für jedes nicht konstante Polynom p gilt

$$\lim_{x \rightarrow -\infty} |p(x)| = \infty,$$

können explizite Runge-Kutta-Verfahren niemals A-stabil sein. Für obiges AWP liefern sie nur dann eine qualitativ richtige numerische Lösung, wenn die Schrittweite h hinreichend klein ist.

BEISPIEL VI.4.2. Wir führen jeweils 100 Schritte der beiden Eulerverfahren und des Crank-Nicolson-Verfahrens für das AWP

$$\mathbf{y}' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{y}$$

$$\mathbf{y}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

mit $h = 0.13$ aus. Die exakte Lösung lautet $\mathbf{y}(t) = (\cos t, \sin t)$ mit $0 \leq t \leq 13$. Abbildung VI.4.2 zeigt die entsprechenden Lösungskurven. Wir erkennen, dass das explizite Eulerverfahren explodiert. Dies ist eine Folge seiner fehlenden Stabilität. Das implizite Eulerverfahren dagegen dämpft die Lösung zu stark. Das Crank-Nicolson-Verfahren schließlich liefert eine Lösungskurve, die auf der exakten Lösungskurve, dem Kreis um den Ursprung mit Radius 1, liegt.

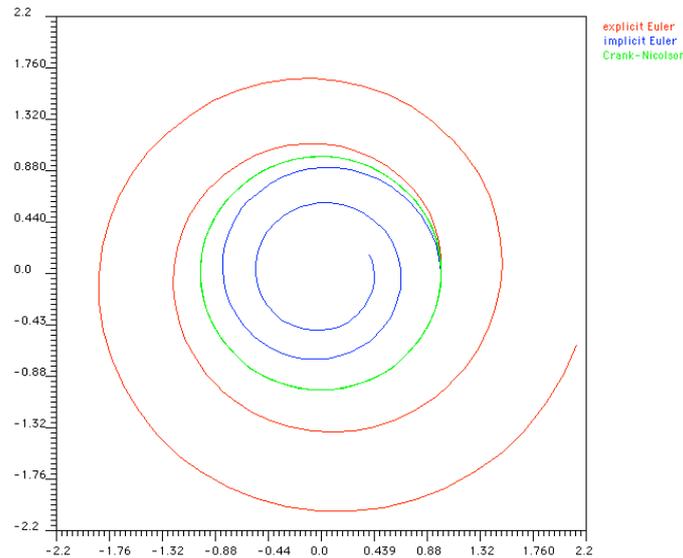


ABBILDUNG VI.4.2. Lösungskurven der beiden Eulerverfahren und des Crank-Nicolson-Verfahrens für das AWP aus Beispiel VI.4.2

Zusammenfassung

I Zahlen und Vektoren

1. Mengen und Abbildungen

Mengen; Elemente; leere Menge; Teilmengen; Gleichheit von Mengen; Mengenoperationen; disjunkte Mengen; Abbildungen, Funktionen; Definitions- und Wertebereich; injektiv, surjektiv, bijektiv; Umkehrabbildung; Komposition; Gleichheit von Abbildungen

2. Die reellen Zahlen

\mathbb{N} , \mathbb{N}^* , \mathbb{Z} , \mathbb{Q} , \mathbb{R} ; Ungleichungen und Eigenschaften; Intervalle; $\pm\infty$; beschränkte Mengen; Infimum und Supremum; Betrag; Eigenschaften des Betrages, Dreiecksungleichung; Induktionsprinzip; geometrische Summenformel; Bernoulli-Ungleichung; rekursive Definition; Potenzen; Fakultäten; Binomialkoeffizienten; eine n -elementige Menge hat $\binom{n}{k}$ k -elementige Teilmengen; Eigenschaften der Binomialkoeffizienten

3. Die Ebene

Kartesische Koordinatensysteme; Achsen; Ursprung; Graph einer Funktion; Gerade, Kreis, Ellipse, Hyperbel, Parabel; Winkel; Bogenmaß; Drehungen; Umrechnung zwischen gedrehten Koordinatensystemen

4. Vektoren

Vektoren; Addition; Parallelogrammregel; skalare Vielfache; Betrag; Dreiecksungleichung; Winkel; Skalarprodukt; Orthogonalität; Cauchy-Schwarzsche Ungleichung; orthogonale Zerlegung; Vektorprodukt; Eigenschaften; Spatprodukt; Eigenschaften; Volumen von Tetraedern; Koordinatendarstellung; Ortsvektor; Darstellung von Skalar-, Vektor- und Spatprodukt; Fläche von Dreiecken; nichtorthogonale Zerlegung

5. Geraden und Ebenen

Darstellungen einer Geraden: Punkt-Richtungs-Gleichung, Zwei-Punkte-Gleichung, Momentengleichung; Koordinatendarstellungen; Abstände zwischen Punkte und Geraden und zwischen Geraden; Darstellungen einer Ebene: Parameterdarstellung, Determinantenform, Normalengleichung, Hesse-Normalform; Koordinatendarstellungen; Schnittgerade zweier Ebenen; Abstände von Punkten, Geraden und Ebenen zu Ebenen; Winkel zwischen Ebenen und Geraden

6. Die komplexen Zahlen

Definition; Addition, Subtraktion; Multiplikation, Division; Betrag; konjugiert komplexe Zahl; geometrische Interpretation; Rechenregeln; Additionstheoreme für Sinus und Cosinus; Wurzeln; eine komplexe Zahl hat genau n verschiedene n -te Wurzeln; quadratische Gleichungen

II Lineare Algebra

1. Lineare Gleichungssysteme und Matrizen

Matrizen, Zeilen- und Spaltenvektoren; Rechenregeln; Nullmatrix und Nullvektor; Basen des \mathbb{R}^n und $\mathbb{R}^{1 \times n}$; lineare Gleichungssysteme und Matrizen; Koeffizienten, Absolutglieder, Koeffizientenmatrix, rechte Seite; Lösungsvektor; homogene und inhomogene LGS; ein homogenes LGS besitzt stets die triviale Lösung; ein inhomogenes LGS besitzt entweder genau eine, keine oder unendlich viele Lösungen; Gaußsches Eliminationsverfahren; Rang einer Matrix

2. Die Matrixmultiplikation

Produkt von Zeilen- und Spaltenvektoren; Produkt von Matrizen; Matrixprodukt ist nicht kommutativ; Einheitsmatrix; Rechenregeln für das Produkt; transponierte Matrix; Rechenregeln für die Transposition; symmetrische und schiefsymmetrische Matrizen; invertierbare und singuläre Matrizen; Rechenregeln für inverse Matrizen; Gaußsches Eliminationsverfahren zur Berechnung der inversen Matrix; Charakterisierung inverser Matrizen; Aufwand des Gaußschen Eliminationsverfahrens; Idee der LR-Zerlegung; Durchführung; Aufwand

3. Determinanten

Determinante einer 2×2 und 3×3 Matrix; Formel von Sarrus; Determinante einer $n \times n$ Matrix; A invertierbar $\iff \det A \neq 0$; Rechenregeln: Determinante kann nach beliebiger Zeile oder Spalte entwickelt werden, Determinante ist linear in jeder Zeile und Spalte, Determinante ist alternierend, Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente, $\det A^T = \det A$, $\det(AB) = (\det A)(\det B)$; Cramersche Regel zur Lösung eines LGS; Aufwand $n!$; Darstellung von Kegelschnitten durch Determinanten

4. Eigenwerte und Eigenvektoren

λ Eigenwert von $A \iff \exists \mathbf{b} \neq 0$ mit $A\mathbf{b} = \lambda\mathbf{b}$, \mathbf{b} heißt Eigenvektor; charakteristisches Polynom $\chi_A(\lambda) = \det(A - \lambda I)$; EW sind die Nullstellen des charakteristischen Polynoms; algebraische und geometrische Vielfachheit eines EW; Spur einer Matrix; Spur A ist die Summe der EW; $\det A$ ist das Produkt der EW; Berechnung der EW und EV: Stelle das charakteristische Polynom auf, bestimme alle seine Nullstellen, löse für jeden EV λ das homogene LGS $(A - \lambda I)\mathbf{x} = \mathbf{0}$; Numerische Verfahren: Potenzmethode, Rayleigh-Quotient, inverse Iteration von Wielandt, inverse Rayleigh-Quotient, Berechnung eines EV; Rechenregeln; ähnliche Matrizen; ähnliche Matrizen haben die gleichen EW; diagonalisierbare Matrizen; nicht jede Matrix ist diagonalisierbar; orthogonale Matrizen; A orthogonal $\iff A^T = A^{-1}$; A orthogonal $\implies |\det A| = 1$; Drehmatrizen; Householder-Matrizen; symmetrische Matrizen; eine symmetrische $n \times n$ Matrix hat n reelle EW und n zugehörige EV; symmetrische Matrizen sind diagonalisierbar; Schursche Normalform; Hauptvektoren; jede $n \times n$ Matrix besitzt n linear unabhängige Hauptvektoren; Berechnung der Hauptvektoren und der Schurschen Normalform einer Matrix

5. Quadratische Formen

Quadratische Polynome; quadratische Formen; Quadriken; Normalformen; Bestimmung der Normalform; Normalformen ebener und

räumlicher Quadriken; positiv definite, negativ definite, indefinite, positiv semidefinite und negativ semidefinite Matrizen und quadratische Formen; A positiv definit \iff alle EW sind positiv; A positiv definit \iff die Determinanten aller Hauptmatrizen sind positiv

6. Vektorräume und lineare Abbildungen

Vektorraumaxiome; Beispiele; Unterräume; Linearkombination; lineare Hülle; linear abhängig, linear unabhängig; $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbb{R}^m$ linear unabhängig \iff $\text{Rang}(\mathbf{u}_1, \dots, \mathbf{u}_n) = n$; je $n > m$ Vektoren in \mathbb{R}^m sind stets linear abhängig; Basis; endlich dimensionale Vektorräume; Dimension; Eigenschaften von Basen und endlich dimensionale Vektorräume; nicht jeder Vektorraum ist endlich dimensional; Skalarprodukt; Cauchy-Schwarzsche Ungleichung; Norm; Dreiecksungleichung; orthogonale und normierte Vektoren; Orthonormalsysteme; Schmidtsches Orthogonalisierungsverfahren; lineare Abbildungen; Matrixdarstellung linearer Abbildungen; Matrixdarstellung hängt von gewählten Basen ab; Komposition linearer Abbildungen entspricht der Multiplikation der zugehörigen Matrizen; ähnliche Matrizen stellen dieselbe lineare Abbildung bzgl. verschiedener Basen dar

III Stetigkeit

1. Folgen

Folgen, Folgenglieder; Beispiele; Summe, Produkt, Quotient von Folgen; beschränkte Folgen; monoton wachsende und monoton fallende Folgen; Teilfolgen

2. Grenzwerte von Folgen

Konvergenz von Folgen; Definition des Grenzwertes; Nullfolgen; Summen, Produkte, Quotienten konvergenter Folgen sind konvergent, der Grenzwert ist Summe, Produkt, Quotient der Grenzwerte; konvergent \Rightarrow beschränkt; beschränkt $\not\Rightarrow$ konvergent; Teilfolgen konvergenter Folgen sind konvergent; monoton und beschränkt \Rightarrow konvergent; Ungleichungen bleiben unter Konvergenz erhalten; Exponentialfunktion; Eulersche Zahl; Eigenschaften der Exponentialfunktion; Konvergenz gegen $\pm\infty$

3. Stetigkeit

Stetigkeit; Beispiele; Summe, Differenz, Produkt, Komposition stetiger Funktionen ist stetig; Quotient stetiger Funktionen ist in allen Punkten stetig, in denen der Nenner nicht verschwindet; stetige Funktionen auf beschränkten, abgeschlossenen Intervallen sind beschränkt und nehmen Maximum und Minimum an; Aussage gilt nicht für offene, halboffene oder unbeschränkte Intervalle; Zwischenwertsatz und Anwendungen; stetige Funktionen auf beschränkten, abgeschlossenen Intervallen sind gleichmäßig stetig; Aussage gilt nicht für offene, halboffene oder unbeschränkte Intervalle; links- und rechtsseitige Grenzwerte; stetig \iff links- und rechtsseitiger Grenzwert existieren, sind endlich und stimmen überein; Polynome; Zerlegung in Linearfaktoren; Polynome ungeraden Grades haben mindestens eine reelle Nullstelle; Polynome vom Grad n haben unter Berücksichtigung der Vielfachheit n komplexe Nullstellen; rationale Funktionen; rationale Funktionen sind in allen Punkten

stetig, in denen das Nennerpolynom nicht verschwindet; Definition von \tan , \cot ; die trigonometrischen Funktionen sind auf ihrem Definitionsbereich stetig; die Exponentialfunktion ist stetig und bijektiv von \mathbb{R} nach $(0, \infty)$; Logarithmusfunktion $\ln : (0, \infty) \rightarrow \mathbb{R}$ ist die Umkehrfunktion der Exponentialfunktion; \ln ist monoton

IV Differentiation

1. Die Ableitung einer differenzierbaren Funktion

Definition der Ableitung; geometrische, analytische und physikalische Deutung der Ableitung; differenzierbar \implies stetig, Umkehrung ist falsch; Differentiationsregeln, Produkt- und Quotientenregel; Polynome und rationale Funktionen sind auf ihrem Definitionsbereich differenzierbar, die Ableitung ist wieder ein Polynom bzw. eine rationale Funktion; trigonometrische Funktionen sind auf ihrem Definitionsbereich differenzierbar; Kettenregel; rekursive Definition höherer Ableitungen; p Polynom vom Grad $k \implies p^{(n)} = 0$ für $n > k$

2. Anwendungen der Differentiation

Maxima und Minima; \bar{x} innerer Punkt, f differenzierbar in \bar{x} , \bar{x} Extremum $\implies f'(\bar{x}) = 0$; Mittelwertsatz; Satz von Rolle; $f'(x) = 0$ für alle $x \implies f$ ist konstant; Extremwerttest; Wendepunkte; Wendepunkttest; Regeln von de l'Hôpital; Fixpunkte; Fixpunktiteration, a priori und a posteriori Fehlerabschätzung; Newtonverfahren; geometrische Interpretation; quadratische Konvergenz; Kessel oder Divergenz bei ungünstigem Startwert; Verfahren von Heron; divisionsfreie Berechnung von Reziproken

3. Umkehrfunktionen

Satz über die Umkehrfunktion; Ableitung der Umkehrfunktion; n -te Wurzel; rationale Exponenten; \arccos , \arcsin , \arctan , arccot

4. Die Exponential- und Logarithmusfunktion

$\exp' = \exp$; $\exp(x + y) = \exp(x)\exp(y)$; die Exponentialfunktion wächst schneller als jede Potenz; $\ln'(x) = \frac{1}{x}$; die Logarithmusfunktion wächst langsamer als jede Wurzel; a^x ; \log_a ; \sinh , \cosh , \tanh , arsinh , arcosh , artanh

V Integration

1. Das bestimmte Integral

Definition des bestimmten Integrals; Integral ist linear, monoton und additiv; Mittelwertsatz der Integralrechnung; Stammfunktionen; Hauptsatz der Differential- und Integralrechnung; unbestimmtes Integral

2. Integrationsregeln

Linearität; partielle Integration; 1. Version der Substitutionsregel; Orthogonalitätsbeziehungen der Sinus- und Cosinus-Funktionen; 2. Version der Substitutionsregel; Integration gerader und ungerader Funktionen über symmetrische Intervalle

3. Integration rationaler Funktionen

Partialbruchzerlegung; Integration der Partialbrüche; Transformation von Integralen verallgemeinerter rationaler Funktionen auf Integrale rationaler Funktionen

4. Uneigentliche Integrale

Definition; $\int_1^\infty x^{-\alpha} dx$ existiert für $\alpha > 1$; $\int_0^1 x^{-\alpha} dx$ existiert für $\alpha < 1$; Konvergenzkriterium durch Vergleich mit $x^{-\alpha}$; Eulersche

Gammafunktion $\Gamma(x)$; Eigenschaften: $\Gamma(1) = 1$, $\Gamma(x + 1) = x\Gamma(x)$, $\Gamma(n + 1) = n!$

5. Ebene Kurven; Längen- und Flächenmessung

Parameterdarstellung einer ebenen Kurve; Parameterdarstellung von Graphen, Geraden, Kreisen, Ellipsen, Zykloiden; Tangente; Normale; reguläre Parameterdarstellung; Länge einer Kurve; Bogenlänge; Krümmung; Kurve hat konstante Krümmung $\kappa \neq 0 \iff$ Kurve ist ein Kreis mit Radius $\frac{1}{|\kappa|}$; Krümmungskreis; Polardarstellung einer ebenen Kurve; Länge einer Kurve in Polardarstellung; Flächeninhalt einer durch zwei Graphen berandeten Fläche; Sektorflächen; Leibnizsche Sektorformel; Volumen von Rotationskörpern; Mantelfläche von Rotationskörpern

6. Numerische Integration

Quadraturformeln; Knoten, Gewichte und Ordnung; Mittelpunkts-, Trapez- und Simpsonregel; Newton-Cotes-Formeln; Gauß-Formeln; zusammengesetzte Quadraturformeln; Fehler zusammengesetzter Quadraturformeln; Romberg-Verfahren; Fehler des Romberg-Verfahrens

VI Gewöhnliche Differentialgleichungen I: Skalare Gleichungen

1. Einführung

Gewöhnliche Differentialgleichung; Ordnung; autonom; Lösung einer gewöhnlichen Differentialgleichung; allgemeine, partikuläre und singuläre Lösung; Anfangswertproblem; lokale Lösung eines Anfangswertproblems; sachgemäß gestellte Probleme; Richtungsfeld; nicht jede gewöhnliche Differentialgleichung ist sachgemäß gestellt

2. Differentialgleichungen 1. Ordnung

Trennung der Variablen: Bestimmen der allgemeinen Lösung der gDgl, Lösen des AWP; lineare Differentialgleichungen 1. Ordnung; homogene und inhomogene lineare Differentialgleichungen; allgemeine Lösung der homogenen gDgl; Bestimmung einer partikulären Lösung der inhomogenen gDgl mit der Methode der Variation der Konstanten; homogene Differentialgleichungen; Bestimmung der allgemeinen Lösung; Bernoulli-Differentialgleichung; Ricatti-Differentialgleichung

3. Differentialgleichungen 2. Ordnung

Lineare Differentialgleichungen mit konstanten Koeffizienten; homogene und inhomogene gDgl; Bestimmung der allgemeinen Lösung der homogenen gDgl, charakteristisches Polynom; Bestimmung einer partikulären Lösung der inhomogenen gDgl, Wronski-Determinante; Differentialgleichungen der Form $y'' = f(x, y')$; Differentialgleichungen der Form $y'' = f(y, y')$

4. Numerische Lösung gewöhnlicher Differentialgleichungen

Explizites und implizites Eulerverfahren; Verfahren von Crank-Nicolson; Einschrittverfahren; Verfahrensfunktion; lokaler Verfahrensfehler; Ordnung; Runge-Kutta-Verfahren; klassisches Runge-Kutta-Verfahren; SDIRK-Verfahren; Stabilität; explizite Verfahren haben mangelnde Stabilität; A-stabil; implizite Eulerverfahren, Verfahren von Crank-Nicolson und SDIRK-Verfahren sind A-stabil

Index

- $\chi_A(\lambda)$, 75
- \angle , 26
- \cap , 11
- \circ , 12
- \cup , 11
- \emptyset , 11
- \perp , 26
- $[\cdot, \cdot, \cdot]$, 28
- \subset , 11
- \times , 12
- \times , 27
- \rightarrow , 12
- A^{-1} , 62
- arccos, 154
- arccot, 156
- arcosh, 164
- arcsin, 154
- arctan, 155
- arsinh, 164
- artanh, 164
- A^T , 61
- $|\cdot|$, 16, 25, 41
- \mathbb{C} , 41
- $\mathbb{C}^{m \times n}$, 50
- cos, 21
- cosh, 163
- cot, 125
- det, 68–70
- f' , 129
- dim, 102
- e , 118
- η_i , 239
- $\eta(t_i, h_i)$, 239
- exp, 118
- $\Gamma(x)$, 192
- $[a, b]$, 15
- $[a, b)$, 15
- Im, 41
- I_n , 61
- inf, 15
- $-\infty$, 15
- ∞ , 15
- \int , 169
- \int_a^b , 166
- $(a, b]$, 15
- (a, b) , 15
- Kern, 99
- $\lim_{n \rightarrow \infty}$, 112
- $\lim_{x \rightarrow a^-}$, 121
- $\lim_{x \rightarrow a^+}$, 121
- ln, 127
- \log_a , 163
- \mathbb{N} , 13
- \mathbb{N}^* , 13
- $n!$, 19
- $\binom{n}{k}$, 19
- $\sqrt[n]{z}$, 45
- \mathbb{Q} , 13
- \mathbb{R} , 13
- Rang, 53
- Re, 41
- $\mathbb{R}^{m \times n}$, 50
- \mathbb{R}^n , 50
- sin, 21
- sinh, 163
- span, 100
- Spur, 76
- sup, 15
- tan, 125
- tanh, 163
- \mathbb{Z} , 13
- a posteriori Fehlerabschätzung, 144
- a priori Fehlerabschätzung, 144
- A-stabil, 249
- Abbildung, 12
- abgeschlossenes Intervall, 15
- ableiten, 129
- Ableitung, 129
- Absolutglied, 52
- Abstand, 38

- Additionstheoreme, 44
- ähnliche Matrizen, 84
- Ähnlichkeits-Differentialgleichung, 225
- äquidistante Knoten, 206
- algebraische Vielfachheit, 75
- allgemeine Höhenformel, 162
- allgemeine Lösung, 215
- alternierend, 71
- Anfangswertproblem, 215
- Archimedische Spirale, 201
- Arcuscosinus, 154
- Arcuscotangens, 156
- Arcussinus, 154
- Arcustangens, 155
- Area cosinus hyperbolicus, 164
- Area sinus hyperbolicus, 164
- Argumente, 12
- Assoziativgesetz, 25, 99
- Astroide, 195
- autonom, 215
- AWP, 215

- barometrische Höhenformel, 163
- Basis, 101
- Bernoulli-Differentialgleichung, 227
- Bernoulli-Ungleichung, 18
- beschränkte Folge, 111
- beschränkte Menge, 15
- beste lineare Approximation, 131
- bestimmte Integral, 166
- Betrag, 16, 25, 41
- bijektiv, 12
- Bild, 12
- Binomialkoeffizient, 19
- Binomische Formel, 20
- Bogenlänge, 197
- Bogenmaß, 21

- Cauchy-Schwarzsche Ungleichung, 26, 104
- charakteristisches Polynom, 75, 230
- Cosinus, 21
- Cosinus hyperbolicus, 163
- Cotangens, 125
- Cramersche Regel, 73

- Definitionsbereich, 12
- Determinante, 68–70
- Determinantenform, 37
- diagonalisierbare Matrix, 84
- Differentialgleichung, 158
- differenzierbar, 129
- differenzieren, 129
- Dimension, 102
- disjunkt, 12
- Distributivgesetz, 26
- Drehmoment, 31
- Drei-Punkte-Gleichung, 37
- Dreiecksungleichung, 16, 25, 104
- Durchschnitt, 11

- Eigenvektor, 74
- Eigenwert, 74
- Einheitsmatrix, 61
- Einschrittverfahren, 242
- Elemente, 11
- Elemente einer Matrix, 49
- Ellipse, 21
- endlich dimensional, 102
- Epizykloide, 194
- ESV, 242
- Euklidische Norm, 105
- Euklidisches Skalarprodukt, 103
- Eulersche Gammafunktion, 192
- Eulersche Zahl, 118
- EV, 74
- Evolute, 200
- EW, 74
- explizites Eulerverfahren, 240
- explizites Runge-Kutta-Verfahren, 244
- Exponentialfunktion, 118
- Extremum, 136
- Extremwerttest, 140

- Fakultät, 19
- Fixpunkt, 144
- Fixpunktiteration, 144
- Flächeninhalt eines Dreieckes, 31
- Folge, 111
- Formel von Sarrus, 69
- 5-Punkte-Gleichung für den allgemeinen Kegelschnitt, 74
- Funktion, 12
- Funktionalgleichung, 193

- ganze Zahlen, 13
- Gaußsches Eliminationsverfahren, 54, 62
- Gauß-Formeln, 206
- gDgl, 214
- geometrische Summenformel, 18
- geometrische Vielfachheit, 75
- Gerade, 21, 194
- Geschwindigkeit, 131

- Gewichte einer Quadraturformel, 206
 gewöhnliche Differentialgleichung,
 214
 gleich, 11, 12
 gleiche Matrizen, 50
 gleichmäßig stetig, 121
 Glieder einer Folge, 111
 globales Maximum, 136
 globales Minimum, 136
 Grad, 122
 Graph, 20, 193
 Grenzwert, 112
- Hängebrücke, 170
 halboffenes Intervall, 15
 Halbwertszeit, 159
 Hauptsatz der Differential- und
 Integralrechnung, 168
 Hauptvektor, 90
 Hesse-Normalform, 38
 höhere Ableitungen, 135
 homogene Differentialgleichung, 225
 homogene lineare gDgl, 221, 229
 homogenes LGS, 52
 homogenes quadratisches Polynom,
 92
 Homogenität, 104
 Householder-Matrix, 86
 Hyperbel, 21
 Hypozykloide, 194
- Imaginärteil, 41
 implizites Eulerverfahren, 240
 implizites Runge-Kutta-Verfahren,
 244
 indefinit, 97
 Induktionsanfang, 16
 Induktionsprinzip, 16
 Induktionsschritt, 17
 Infimum, 15
 inhomogene lineare gDgl, 221, 229
 inhomogenes LGS, 52
 injektiv, 12
 Injektivitätskriterium, 150
 Inklusion, 11
 Integral, 166
 integrierbar, 189
 Inverse, 62
 inverse Iteration von Wielandt, 80
 inverse Matrix, 62
 inverse
 Rayleigh-Quotienten-Iteration,
 82
- invertierbare Matrix, 62
 irrationale Zahlen, 13
- kartesische Basis, 29
 kartesisches Koordinatensystem, 20,
 29
 Kessel, 147
 Kettenlinie, 170
 Kettenregel, 133
 klassisches Runge-Kutta-Verfahren,
 245
 Knoten einer Quadraturformel, 206
 Koeffizient, 52
 Koeffizientenmatrix, 52
 Kommutativgesetz, 24, 26, 98
 Komplement, 12
 komplexe Matrix, 49
 komplexe Zahl, 41
 Komposition, 12
 konjugiert komplexe Zahl, 42
 konvergente Folge, 112
 Konvergenzkriterium für
 uneigentliche Integrale, 191
 Koordinatenebene, 29
 Koordinatengleichungen, 34
 Kreis, 21, 194
 Krümmung, 198
 Krümmungskreis, 199
- Länge, 196
 leere Menge, 11
 Leibnizsche Sektorformel, 203
 LGS, 52
 Limes, 112
 linear abhängig, 100
 linear unabhängig, 100
 lineare Abbildung, 107
 lineare gDgl, 221
 lineare Hülle, 100
 lineares Gleichungssystem, 52
 Linearität, 103
 Linearkombination, 99
 linksseitiger Grenzwert, 121
 Lösung einer gDgl, 215
 Logarithmusfunktion, 127
 lokal lösbar, 215
 lokale Lösung, 215
 lokaler Verfahrensfehler, 243
 lokales Maximum, 136
 lokales Minimum, 136
 LR-Zerlegung, 65
- Matrix, 49

- Maximum, 136
- Menge, 11
- Minimum, 136
- minus unendlich, 15
- Mittelpunktsregel, 206
- Mittelwertsatz, 139
- Mittelwertsatz der Integralrechnung, 167
- Momentengleichung, 35
- monoton fallende Folge, 112
- monoton wachsende Folge, 112
- monotone Folge, 112

- nach oben beschränkte Menge, 15
- nach unten beschränkte Menge, 15
- natürliche Basis, 51
- natürliche Zahlen, 13
- negativ definit, 97
- negativ semidefinit, 97
- Neilsche Parabel, 200
- Nennerpolynom, 124
- Newton-Cotes-Formeln, 206
- Newtonsches Abkühlungsgesetz, 159
- Newtonverfahren, 146
- nichtorthogonale Zerlegung, 32
- Norm, 104
- Normale, 195
- Normalengleichung, 38
- Normalenvektor, 38
- Normalform einer Quadrik, 93
- Normaxiome, 104
- normiert, 105
- normierte Vektoren, 85
- n -te Wurzel, 45, 152
- Nullelement, 99
- Nullfolge, 112
- Nullmatrix, 51
- Nullpunkt, 20
- Nullvektor, 24, 51, 99

- obere Dreiecksmatrix, 65
- obere Schranke, 15
- offenes Intervall, 15
- Ordnung, 243
- Ordnung einer Quadraturformel, 206
- orthogonal, 26, 105
- orthogonale Matrix, 85
- orthogonale Vektoren, 85
- orthogonale Zerlegung, 26
- Orthogonalitätsbeziehungen, 176
- Orthogonalitätstest, 26
- Orthonormalsystem, 105
- Ortsvektor, 29

- Parabel, 21
- Parallelepiped, 28
- Parallelfäche, 28
- Parameter, 33, 193
- Parameterdarstellung, 33, 37, 193
- Parameterintervall, 193
- Partialbruch, 179
- partikuläre Lösung, 215
- Pivotsuche, 54
- plus unendlich, 15
- Poissonsches Gasgesetz, 161
- Pol, 125
- Polarachse, 201
- Polardarstellung, 201
- Polarkoordinaten, 200
- Polstelle, 125
- Polynom, 75, 122
- positiv definit, 97
- positiv semidefinit, 97
- Positivität, 103, 104
- Potenz, 19
- Potenzmethode, 78
- Produkt, 12
- Produktregel, 132
- properly posed, 215
- Punkt-Richtungs-Gleichung, 33

- quadratische Form, 92
- quadratische Konvergenz, 148
- quadratisches Polynom, 92
- Quadraturformel, 205
- Quadrik, 93
- Quotientenregel, 132

- radioaktiver Zerfall, 159
- Rang, 53
- rationale Funktion, 124
- rationale Zahlen, 13
- Rayleigh-Quotienten-Iteration, 79
- Realteil, 41
- rechte Seite, 52
- rechtsseitiger Grenzwert, 121
- reelle Matrix, 49
- reelle Zahlen, 13
- Regeln von de l'Hôpital, 142
- reguläre Matrix, 62
- reguläre Parameterdarstellung, 196
- rein quadratisch, 92
- rekursive Definition, 18
- Ricatti-Differentialgleichung, 228
- Richtungsfeld, 216
- Romberg-Verfahren, 210
- Runge-Kutta-Verfahren, 244

- sachgemäß gestellt, 215
- Satz über die Umkehrfunktion, 150
- Satz von Rolle, 139
- schiefsymmetrische Matrix, 62
- Schmidtsches
 - Orthogonalisierungsverfahren, 106
- Schursche Normalform, 89
- SDIRK-Verfahren, 246
- Seilkurve, 170
- Sektorfläche, 202
- senkrecht, 26
- Simpsonregel, 206
- singuläre Lösung, 215
- singuläre Matrix, 62
- Sinus, 21
- Sinus hyperbolicus, 163
- Skalarprodukt, 26, 103
- Skalarproduktaxiome, 103
- Spaltenvektor, 49
- Spat, 28
- Spatprodukt, 28
- spezielle Lösung, 215
- Stabilität, 248
- Stabilitätsgebiet, 248
- Stammfunktion, 167
- stark diagonal implizite
 - Runge-Kutta-Verfahren, 246
- stationärer Punkt, 137
- stetig, 119
- stückweise stetig, 165
- Stufe eines Hauptvektors, 90
- Stufe eines Runge-Kutta-Verfahrens, 244
- Substitutionsregel, 174, 176
- Supremum, 15
- surjektiv, 12
- Symmetrie, 103
- symmetrische Matrix, 62

- Tangens, 125
- Tangens hyperbolicus, 163
- Tangente, 130, 195
- Teilfolge, 112
- Teilmenge, 11
- Torus, 204
- Transformationsformeln für
 - Drehungen, 22
- transponierte Matrix, 61
- Trapezregel, 206, 241
- triviale Linearkombination, 99
- triviale Lösung, 52

- Umkehrabbildung, 12
- unbestimmtes Integral, 169
- uneigentlich integrierbar, 189
- uneigentliches Integral, 189
- untere Dreiecksmatrix, 65
- untere Schranke, 15
- Unterraum, 99
- Untervektorraum, 99
- Ursprung, 20, 29

- Variation der Konstanten, 223, 232
- Vektor, 24, 99
- Vektorprodukt, 27
- Vektorraum, 98
- Vektorraumaxiome, 98
- Vereinigung, 11
- Verfahren von Crank-Nicolson, 241
- Verfahren von Heron, 149
- Verfahrensfunktion, 243
- Vielfachheit, 124
- vollständige Induktion, 16
- vollständige Lösung, 215
- Vollständigkeitsaxiom, 15

- Wendepunkt, 140
- Wendepunkttest, 141
- Wertebereich, 12
- Winkel, 25
- Wronski-Determinante, 233

- x -Achse, 20, 29

- y -Achse, 20, 29

- z -Achse, 29
- Zählerpolynom, 124
- Zahlenfolge, 111
- Zahlengrade, 13
- Zeilenvektor, 49
- zusammengesetzte
 - Mittelpunktsregel, 207
- zusammengesetzte Quadraturformel, 207
- zusammengesetzte Simpsonregel, 207
- zusammengesetzte Trapezregel, 207
- Zwischenwertsatz, 120
- Zykloide, 194