

Mathematik für Maschinenbauer, Bauingenieure und Umwelttechniker

Vorlesungsskriptum WS 2001/02 - WS 2002/03
überarbeitet November 2006

R. Verfürth

Fakultät für Mathematik, Ruhr-Universität Bochum

Inhaltsverzeichnis

Vorbemerkung	13
Kapitel I. Zahlen und Vektoren	15
I.1. Mengen und Abbildungen	15
I.1.1. Mengen	15
I.1.2. Mengenoperationen	15
I.1.3. Abbildungen	16
I.2. Die reellen Zahlen	17
I.2.1. Bezeichnungen	17
I.2.2. Ungleichungen	17
I.2.3. Intervalle	19
I.2.4. Schranken	19
I.2.5. Der Betrag	20
I.2.6. Vollständige Induktion	20
I.2.7. Rekursive Definition	22
I.2.8. Binomialkoeffizienten und Binomische Formel	23
I.3. Die Ebene	24
I.3.1. Kartesische Koordinatensysteme	24
I.3.2. Winkel	25
I.3.3. Sinus, Cosinus	25
I.3.4. Drehungen	26
I.4. Vektoren	28
I.4.1. Vektoren	28
I.4.2. Addition von Vektoren	28
I.4.3. Skalare Vielfache von Vektoren	29
I.4.4. Der Betrag	29
I.4.5. Der Winkel zwischen zwei Vektoren	29
I.4.6. Das Skalarprodukt	30
I.4.7. Das Vektorprodukt	31
I.4.8. Das Spatprodukt	32
I.4.9. Koordinatendarstellungen	33
I.5. Geraden und Ebenen	37
I.5.1. Die Parameterdarstellung einer Geraden	37
I.5.2. Die Koordinatengleichungen einer Geraden	38
I.5.3. Die Momentengleichung einer Geraden	38
I.5.4. Abstand Punkt - Gerade	39
I.5.5. Abstand Gerade - Gerade	40

I.5.6.	Die Parameterdarstellung einer Ebene	41
I.5.7.	Parameterfreie Darstellungen einer Ebene	41
I.5.8.	Die Schnittgerade zweier Ebenen	43
I.5.9.	Die Winkel zwischen zwei Ebenen und zwischen einer Ebene und einer Geraden	44
I.6.	Die komplexen Zahlen	45
I.6.1.	Motivation	45
I.6.2.	Definition	45
I.6.3.	Betrag und Konjugation	45
I.6.4.	Addition und Subtraktion	46
I.6.5.	Multiplikation und Division	47
I.6.6.	Wurzeln komplexer Zahlen	49
Kapitel II. Lineare Algebra		53
II.1.	Lineare Gleichungssysteme und Matrizen	53
II.1.1.	Matrizen	53
II.1.2.	Rechenregeln	54
II.1.3.	Lineare Gleichungssysteme und Matrizen	56
II.1.4.	Das Gaußsche Eliminationsverfahren	57
II.2.	Die Matrixmultiplikation	63
II.2.1.	Das Matrixprodukt	63
II.2.2.	Die transponierte Matrix	65
II.2.3.	Invertierbare Matrizen	66
II.2.4.	Das Gaußsche Eliminationsverfahren zur Berechnung der inversen Matrix	66
II.2.5.	Die LR-Zerlegung	68
II.3.	Determinanten	72
II.3.1.	Determinanten von 2×2 und 3×3 Matrizen	72
II.3.2.	Die Determinante einer $n \times n$ Matrix	74
II.3.3.	Rechenregeln für Determinanten	75
II.3.4.	Die Cramersche Regel	76
II.3.5.	Kegelschnitte	77
II.4.	Eigenwerte und Eigenvektoren	78
II.4.1.	Definition	78
II.4.2.	Das charakteristische Polynom	79
II.4.3.	Berechnung der Eigenwerte und Eigenvektoren	80
II.4.4.	Numerische Berechnung der Eigenwerte und Eigenvektoren*	82
II.4.5.	Rechenregeln	87
II.4.6.	Ähnliche Matrizen	88
II.4.7.	Orthogonale Matrizen	89
II.4.8.	Symmetrische Matrizen	91
II.4.9.	Die Schursche Normalform	92
II.4.10.	Hauptvektoren	93
II.5.	Quadratische Formen	96

II.5.1.	Definition	96
II.5.2.	Reduktion auf Normalform	98
II.5.3.	Normalformen der ebenen Quadriken	99
II.5.4.	Normalformen der räumlichen Quadriken	100
II.5.5.	Positiv definite Matrizen	101
II.6.	Vektorräume und lineare Abbildungen	102
II.6.1.	Vektorräume	102
II.6.2.	Unterräume	103
II.6.3.	Linearkombination und lineare Hülle	103
II.6.4.	Lineare Abhängigkeit	104
II.6.5.	Basis und Dimension	105
II.6.6.	Skalarprodukte	107
II.6.7.	Normen	108
II.6.8.	Orthogonalität	109
II.6.9.	Lineare Abbildungen	111
II.6.10.	Matrixdarstellung	111
II.6.11.	Komposition linearer Abbildungen	112
II.6.12.	Basiswechsel	112
Kapitel III.	Stetigkeit	115
III.1.	Folgen	115
III.1.1.	Definition	115
III.1.2.	Rechenregeln	115
III.1.3.	Beschränktheit	115
III.1.4.	Monotonie	116
III.1.5.	Teilfolgen	116
III.2.	Grenzwerte von Folgen	116
III.2.1.	Definition	116
III.2.2.	Rechenregeln	117
III.2.3.	Konvergenzkriterien	118
III.2.4.	Die Exponentialfunktion	121
III.2.5.	Uneigentliche Grenzwerte	122
III.3.	Stetigkeit	123
III.3.1.	Definition	123
III.3.2.	Rechenregeln	123
III.3.3.	Eigenschaften stetiger Funktionen	124
III.3.4.	Einseitige Grenzwerte	125
III.3.5.	Polynome	126
III.3.6.	Rationale Funktionen	128
III.3.7.	Trigonometrische Funktionen	129
III.3.8.	Exponential- und Logarithmusfunktion	130
Kapitel IV.	Differentiation	133
IV.1.	Die Ableitung einer differenzierbaren Funktion	133
IV.1.1.	Definition der Ableitung	133
IV.1.2.	Deutungen der Ableitung	134

IV.1.3.	Stetigkeit ist notwendig für Differenzierbarkeit	135
IV.1.4.	Differentiationsregeln	136
IV.1.5.	Differentiation von Polynomen und rationalen Funktionen	136
IV.1.6.	Differentiation der trigonometrischen Funktionen	137
IV.1.7.	Kettenregel	137
IV.1.8.	Höhere Ableitungen	139
IV.2.	Anwendungen der Differentiation	140
IV.2.1.	Maxima und Minima einer Funktion	140
IV.2.2.	Der Mittelwertsatz	142
IV.2.3.	Wendepunkte	144
IV.2.4.	Die Regeln von de l'Hôpital	146
IV.2.5.	Die Fixpunktiteration	148
IV.2.6.	Das Newtonverfahren	150
IV.3.	Umkehrfunktionen	154
IV.3.1.	Grundlagen	154
IV.3.2.	n -te Wurzel, rationale Exponenten	156
IV.3.3.	Die Umkehrfunktionen der trigonometrischen Funktionen	158
IV.4.	Die Exponential- und Logarithmusfunktion	161
IV.4.1.	Die Exponentialfunktion	161
IV.4.2.	Die Logarithmusfunktion	164
IV.4.3.	Exponential- und Logarithmusfunktionen zu beliebigen Basen	165
IV.4.4.	Die Hyperbelfunktionen	167
Kapitel V.	Integration	169
V.1.	Das bestimmte Integral	169
V.1.1.	Definition	169
V.1.2.	Elementare Integrationsregeln	170
V.1.3.	Differentiation und Integration	171
V.2.	Integrationsregeln	174
V.2.1.	Linearität	174
V.2.2.	Partielle Integration	175
V.2.3.	Substitutionsregel	178
V.2.4.	Symmetrien	182
V.3.	Integration rationaler Funktionen	183
V.3.1.	Partialbruchzerlegung	183
V.3.2.	Integration	187
V.3.3.	Verallgemeinerte rationale Funktionen	189
V.4.	Uneigentliche Integrale	193
V.4.1.	Definition	193
V.4.2.	Ein Konvergenzkriterium	195
V.4.3.	Die Eulersche Gammafunktion	196
V.5.	Ebene Kurven; Längen- und Flächenmessung	197
V.5.1.	Parameterdarstellung	197

V.5.2.	Tangente und Normale	199
V.5.3.	Bogenlänge	200
V.5.4.	Krümmung und Krümmungskreis	202
V.5.5.	Polardarstellung einer ebenen Kurve	204
V.5.6.	Flächeninhalte	206
V.5.7.	Volumina von Rotationskörpern	208
V.5.8.	Mantelflächen von Rotationskörpern	209
V.6.	Numerische Integration*	209
V.6.1.	Quadraturformeln*	209
V.6.2.	Zusammengesetzte Quadraturformeln*	211
V.6.3.	Romberg Verfahren*	213
Kapitel VI. Gewöhnliche Differentialgleichungen I		
	Skalare Gleichungen	217
VI.1.	Einführung	217
VI.1.1.	Beispiele	217
VI.1.2.	Grundbegriffe	218
VI.1.3.	Geometrische Deutung	219
VI.1.4.	Eindeutigkeitsfragen	220
VI.2.	Differentialgleichungen 1. Ordnung	221
VI.2.1.	Trennung der Variablen	221
VI.2.2.	Variation der Konstanten	225
VI.2.3.	Homogene Differentialgleichungen	229
VI.2.4.	Bernoulli-Differentialgleichung	231
VI.2.5.	Ricatti-Differentialgleichung	232
VI.3.	Differentialgleichungen 2. Ordnung	233
VI.3.1.	Lineare Differentialgleichungen mit konstanten Koeffizienten	233
VI.3.2.	Die homogene Gleichung	233
VI.3.3.	Die inhomogene Gleichung	236
VI.3.4.	Differentialgleichungen vom Typ $y'' = f(x, y')$	241
VI.3.5.	Differentialgleichungen vom Typ $y'' = f(y, y')$	242
VI.4.	Numerische Lösung gewöhnlicher Differentialgleichungen*	243
VI.4.1.	Motivation*	243
VI.4.2.	Einschrittverfahren*	246
VI.4.3.	Runge-Kutta Verfahren*	248
VI.4.4.	Stabilität*	251
Kapitel VII. Potenzreihen		255
VII.1.	Reihen	255
VII.1.1.	Definition	255
VII.1.2.	Absolute Konvergenz	256
VII.1.3.	Konvergenzkriterien	256
VII.1.4.	Rechenregeln	259
VII.2.	Potenzreihen	260
VII.2.1.	Definition	260

VII.2.2.	Konvergenzradius	260
VII.2.3.	Bestimmung des Konvergenzradius	261
VII.2.4.	Differentiation und Integration von Potenzreihen	262
VII.2.5.	Potenzreihendarstellung einiger Funktionen	263
VII.2.6.	Die Binomialreihe	264
VII.2.7.	Potenzreihen mit beliebigem Entwicklungspunkt	265
VII.2.8.	Koeffizientenvergleich	266
VII.3.	Taylorreihen	267
VII.3.1.	Die Taylor-Formel	267
VII.3.2.	Die Taylor-Reihe	269
VII.3.3.	Methoden der Potenzreihenentwicklung	270
VII.4.	Anwendungen	271
VII.4.1.	Grenzwertbestimmung	271
VII.4.2.	Näherungsformeln	272
VII.4.3.	Integration	273
VII.4.4.	Lösen von Differentialgleichungen	273
Kapitel VIII. Differentiation von Funktionen in mehreren Variablen		277
VIII.1.	Kurven im \mathbb{R}^n	277
VIII.1.1.	Parameterdarstellungen	277
VIII.1.2.	Das begleitende Dreibein, Krümmung und Torsion	279
VIII.2.	Reellwertige Funktionen mehrerer reeller Veränderlicher	283
VIII.2.1.	Einige topologische Grundbegriffe	283
VIII.2.2.	Reelle Funktionen mehrerer reeller Veränderlicher	284
VIII.2.3.	Grenzwerte und Stetigkeit	285
VIII.2.4.	Partielle Ableitungen, der Gradient	286
VIII.2.5.	Die totale Ableitung und lineare Approximation	289
VIII.2.6.	Einfache Anwendungen	290
VIII.2.7.	Die Richtungsableitung	292
VIII.2.8.	Die Kettenregel	293
VIII.3.	Anwendungen der Differentiation	295
VIII.3.1.	Richtung des stärksten Anstieges	295
VIII.3.2.	Tangenten und Tangentialebene	295
VIII.3.3.	Die Taylor-Formel	297
VIII.3.4.	Die Hesse-Matrix	299
VIII.3.5.	Implizite Funktionen	301
VIII.3.6.	Lokale Extrema	305
VIII.3.7.	Extrema unter Nebenbedingungen	307
VIII.3.8.	Extrema auf kompakten Mengen	312
VIII.4.	Vektorwertige Funktionen	314
VIII.4.1.	Die Differentiation	314
VIII.4.2.	Das Newtonverfahren	316
VIII.4.3.	Die Kettenregel	318
VIII.4.4.	Räumliche Skalaren- und Vektorfelder	319

VIII.4.5.	Gradient, Divergenz, Rotation und Laplace Operator	321
Kapitel IX.	Integration von Funktionen in mehreren Variablen	325
IX.1.	Parameterintegrale	325
IX.1.1.	Übersicht	325
IX.1.2.	Eigentliche Parameterintegrale	325
IX.1.3.	Uneigentliche Parameterintegrale	328
IX.2.	Kurvenintegrale	329
IX.2.1.	Das Kurvenintegral einer skalaren Funktion	329
IX.2.2.	Anwendungen	331
IX.2.3.	Das Kurvenintegral eines Vektorfeldes	333
IX.2.4.	Das Potential eines Gradientenfeldes	335
IX.2.5.	Die praktische Bestimmung einer Stammfunktion	338
IX.3.	Integration über ebene Bereiche	341
IX.3.1.	Der Flächeninhalt	341
IX.3.2.	Das Doppelintegral	343
IX.3.3.	Praktische Berechnung des Doppelintegrals	344
IX.3.4.	Der Satz von Green	348
IX.4.	Integration über Flächen im Raum	351
IX.4.1.	Parameterdarstellungen	351
IX.4.2.	Der Flächeninhalt	355
IX.4.3.	Das Oberflächenintegral einer skalaren Funktion	356
IX.4.4.	Transformationsformel für Gebietsintegrale	360
IX.4.5.	Der Fluss eines Vektorfeldes	362
IX.4.6.	Der Satz von Stokes	363
IX.5.	Integration über dreidimensionale Bereiche	366
IX.5.1.	Das Volumen	366
IX.5.2.	Das Dreifachintegral	367
IX.5.3.	Praktische Berechnung des Dreifachintegrals	367
IX.5.4.	Die Transformationsformel für Volumenintegrale	369
IX.5.5.	Der Satz von Gauß	372
Kapitel X.	Gewöhnliche Differentialgleichungen II	
	Systeme	375
X.1.	Existenz- und Eindeigkeitssätze	375
X.1.1.	Grundbegriffe	375
X.1.2.	Reduktion der Ordnung	376
X.1.3.	Lipschitz-stetige Funktionen	377
X.1.4.	Der Satz von Picard-Lindelöf	378
X.2.	Differentialgleichungen 1. Ordnung	380
X.2.1.	Exakte Differentialgleichungen	380
X.2.2.	Der integrierende Faktor	382
X.3.	Systeme linearer Differentialgleichungen	386
X.3.1.	Grundlegende Eigenschaften	386
X.3.2.	Fundamentalsysteme	386
X.3.3.	Variation der Konstanten	389

X.3.4.	Homogene Systeme mit konstanter Matrix: Der symmetrische Fall	391
X.3.5.	Homogene Systeme mit konstanter Matrix: Der allgemeine Fall	393
X.3.6.	Lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten	398
X.4.	Stabilität	401
X.4.1.	Motivation	401
X.4.2.	Stabilitätsbegriffe	403
X.4.3.	Stabilitätskriterien	404
X.4.4.	Autonome Systeme	405
Kapitel XI. Stochastik I		
	Diskrete Modelle	409
XI.1.	Modelle für Zufallsexperimente	409
XI.1.1.	Endliche Wahrscheinlichkeitsräume	409
XI.1.2.	Urnenmodelle	411
XI.1.3.	Anwendungsbeispiele	412
XI.1.4.	Die hypergeometrische Verteilung	414
XI.1.5.	Multinomialkoeffizienten	415
XI.1.6.	Identitäten für Binomialkoeffizienten	416
XI.2.	Bedingte Wahrscheinlichkeit und Unabhängigkeit	416
XI.2.1.	Definition bedingter Wahrscheinlichkeiten	416
XI.2.2.	Eigenschaften	418
XI.2.3.	Unabhängigkeit	420
XI.2.4.	Produktexperimente	421
XI.2.5.	Binomialverteilung	422
XI.2.6.	Multinomialverteilung	423
XI.2.7.	Geometrische Verteilung	423
XI.2.8.	Negative Binomialverteilung	423
XI.3.	Zufallsvariable, Erwartungswert, Varianz	423
XI.3.1.	Zufallsvariable	423
XI.3.2.	Gemeinsame Verteilung mehrerer Zufallsvariabler	424
XI.3.3.	Unabhängigkeit	426
XI.3.4.	Erwartungswert	426
XI.3.5.	Varianz und Kovarianz	428
XI.3.6.	Das schwache Gesetz der großen Zahlen	432
XI.4.	Grundbegriffe der Schätztheorie	434
XI.4.1.	Motivation	434
XI.4.2.	Der allgemeine Rahmen von Schätzproblemen	435
XI.4.3.	Maximum-Likelihood Schätzer	435
XI.4.4.	Erwartungstreue	436
XI.4.5.	Der mittlere quadratische Fehler	438
XI.5.	Approximationen der Binomialverteilung	439
XI.5.1.	Approximation von $n!$ und $b_{n,p}(k)$	440

XI.5.2.	Der Satz von Moivre-Laplace	443
XI.5.3.	Die Poisson-Approximation	445
XI.6.	Tests	447
XI.6.1.	Motivation	447
XI.6.2.	Grundbegriffe der Testtheorie	449
XI.6.3.	Zurück zur Motivation	450
Kapitel XII. Stochastik II		
	Allgemeine Modelle	453
XII.1.	Wahrscheinlichkeitsmaße mit Dichten	453
XII.1.1.	Ergebnismengen	453
XII.1.2.	σ -Algebren	453
XII.1.3.	Wahrscheinlichkeitsmaße	454
XII.1.4.	Wahrscheinlichkeitsmaße mit Dichten	455
XII.1.5.	Gleichverteilung auf einem Intervall	455
XII.1.6.	Exponentialverteilung	456
XII.1.7.	Normalverteilung	456
XII.1.8.	Produktdichten	456
XII.2.	Zufallsvariable und ihre Momente	457
XII.2.1.	Messbare Funktionen	457
XII.2.2.	Zufallsvariable	457
XII.2.3.	Unabhängigkeit	458
XII.2.4.	Erwartungswert	460
XII.2.5.	Varianz	462
XII.3.	Schätzverfahren	464
XII.3.1.	Maximum-Likelihood Schätzung	464
XII.3.2.	Die Methode der kleinsten Quadrate	466
XII.3.3.	Median	468
XII.4.	Tests	469
XII.4.1.	Vorbemerkungen	469
XII.4.2.	Der t -Test	470
XII.4.3.	Der χ^2 -Test	473
Kapitel XIII. Fourier-Analysis		
XIII.1.	Trigonometrische Polynome und Reihen	477
XIII.1.1.	Periodische Funktionen	477
XIII.1.2.	Trigonometrische Polynome	478
XIII.1.3.	Trigonometrische Reihen	479
XIII.2.	Fourier-Reihen	480
XIII.2.1.	Die Fourier-Reihe einer Funktion	480
XIII.2.2.	Rechenregeln	481
XIII.2.3.	Die Bessel-Ungleichung	482
XIII.2.4.	Konvergenz der Fourier-Reihe	484
XIII.2.5.	Anwendung auf gewöhnliche Differentialgleichungen	486
XIII.3.	Die Fourier-Transformation	487
XIII.3.1.	Definition	487

XIII.3.2.	Rechenregeln	488
XIII.3.3.	Existenz- und Eindeutigkeitssätze	490
Kapitel XIV.	Partielle Differentialgleichungen	493
XIV.1.	Einführung	493
XIV.1.1.	Beispiele	493
XIV.1.2.	Bezeichnungen	497
XIV.1.3.	Lineare partielle Differentialgleichungen zweiter Ordnung	498
XIV.1.4.	Anfangs- und Randbedingungen	499
XIV.2.	Die Wärmeleitungsgleichung	500
XIV.2.1.	Vorbemerkungen	500
XIV.2.2.	Fortsetzung der Randwerte	502
XIV.2.3.	Lösung der homogenen Gleichung mit homogenen Randbedingungen	502
XIV.2.4.	Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen	505
XIV.2.5.	Integraldarstellung	507
XIV.2.6.	Höhere Raumdimensionen	507
XIV.3.	Die Wellengleichung	508
XIV.3.1.	Vorbemerkung	508
XIV.3.2.	Lösung der homogenen Gleichung mit homogenen Randbedingungen	509
XIV.3.3.	Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen	511
XIV.3.4.	Integraldarstellung	512
XIV.3.5.	Höhere Raumdimensionen	513
XIV.3.6.	Dämpfung	514
XIV.4.	Die Poissongleichung	514
XIV.4.1.	Das Eigenwertproblem im Rechteck	514
XIV.4.2.	Die Poissongleichung im Rechteck	516
XIV.4.3.	Das Eigenwertproblem im Kreis	518
XIV.4.4.	Die Poissongleichung im Kreis	520
Zusammenfassung		525
Index		535

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

<http://www.ruhr-uni-bochum.de/num1/skripte/NumIng.pdf>

zur Verfügung.

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

<http://www.ruhr-uni-bochum.de/num1/demo/index.html>

findet man das Java-Applet **Numerics** mitsamt Benutzeranleitung in **html**- bzw. **pdf**-Form, 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 $c \in \mathbb{R}$	$\implies a + c < b + c$

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

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. F\"ur 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

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

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

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

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

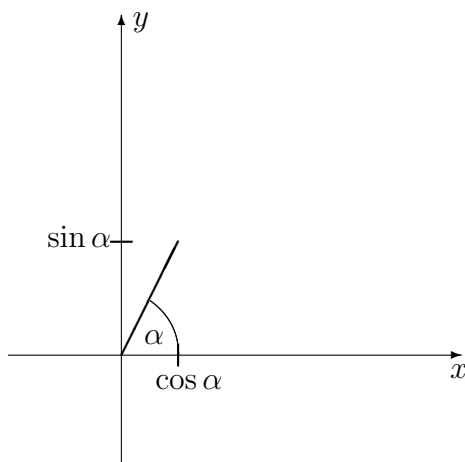


ABBILDUNG I.3.1. Sinus und Cosinus

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):

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

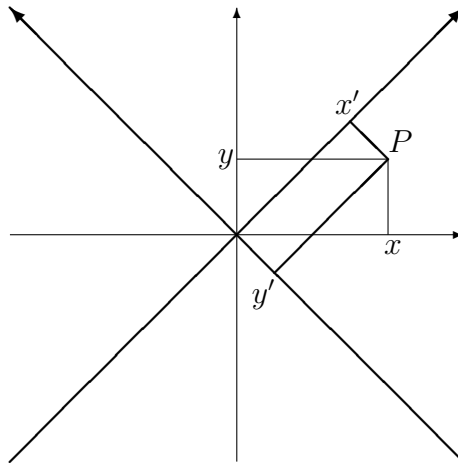


ABBILDUNG I.3.2. Drehung eines kartesischen Koordinatensystems

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

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

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

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

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:

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

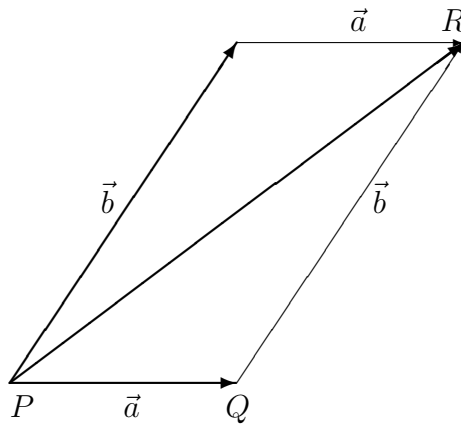


ABBILDUNG I.4.1. Parallelogrammregel

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

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 O rechtwinklig schneidenden Zahlengeraden gleicher Längeneinheit und jeweils dem Nullpunkt im Schnittpunkt O . O 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 $(O; \vec{e}_1, \vec{e}_2, \vec{e}_3)$.

Der Vektor $\vec{a} = \overrightarrow{OA}$ 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

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

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{OA} .

Sind $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$ zwei Punkte, so ist $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$. 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 1.4.9 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 1.4.9 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 1.4.7

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

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 1.5.1 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 1.4.1 (S. 30) 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 1.5.6. 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

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

somit 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 $\vec{0P}$ 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).

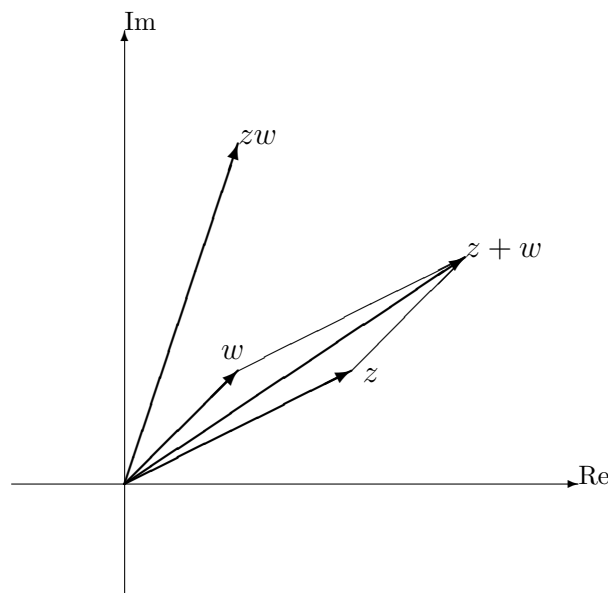


ABBILDUNG I.6.1. Addition und Multiplikation zweier komplexer Zahlen

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 unbekanntnen 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

$$\begin{array}{l}
 r \\
 \\
 \\
 \\
 \\
 \\
 \\
 \\
 m - r
 \end{array}
 \left\{
 \begin{array}{cccccccccc}
 \bullet & * & * & \dots & \dots & \dots & \dots & * & * \\
 0 & \bullet & * & \dots & \dots & \dots & \dots & * & * \\
 0 & 0 & \bullet & \dots & \dots & \dots & \dots & * & * \\
 \vdots & \vdots & & \ddots & \vdots & \vdots & & \vdots & \vdots \\
 \vdots & \vdots & & & \vdots & \vdots & & \vdots & \vdots \\
 \vdots & \vdots & & & \bullet & * & \dots & * & * \\
 0 & & & & & 0 & \dots & 0 & * \\
 \vdots & & & & & \vdots & \dots & \vdots & \vdots \\
 0 & & & & & 0 & \dots & 0 & *
 \end{array}
 \right\}.$$

$\underbrace{\hspace{15em}}_{n+1}$

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 TRANSPONIERTE 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. 61) 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. 58), $O(n^3)$ -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. 58) 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. 58) 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. 61) 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:

$$\begin{array}{cccccc}
 & + & & + & & + \\
 a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\
 & \searrow & & \swarrow & & \swarrow & & \nearrow & \\
 a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\
 & \nearrow & & \swarrow & & \swarrow & & \searrow & \\
 a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \\
 - & & - & & - & & & &
 \end{array} .$$

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. 58) 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. 58) 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. 58) 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

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

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

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

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 1.4.9). 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. 58) 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. 91) 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. 58) 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.

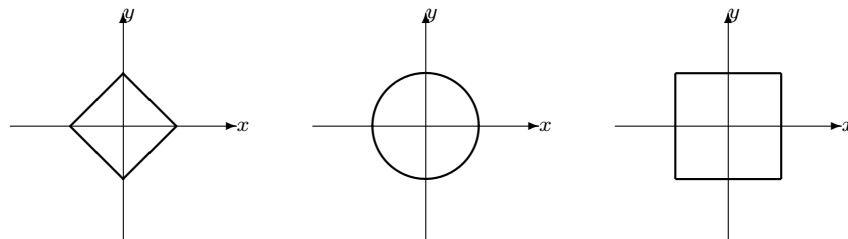


ABBILDUNG II.6.1. Einheitsbälle in \mathbb{R}^2 für die Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ (v.l.n.r)

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

$$(II.6.2) \quad A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}.$$

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

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{n+1}{n} &= 1, \\ \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n &= 0\end{aligned}$$

und

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{2n+1}{n^2} &= \lim_{n \rightarrow \infty} \left(\frac{2}{n} + \frac{1}{n^2}\right) \\ &= 0\end{aligned}$$

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. 22),

$$|q|^n \geq \left(1 + \frac{1}{m}\right)^n \geq 1 + \frac{n}{m}.$$

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

$$\begin{aligned} \frac{\left(1 + \frac{x}{n+1}\right)^{n+1}}{\left(1 + \frac{x}{n}\right)^n} &= \frac{(n+1+x)^{n+1}}{(n+1)^{n+1}} \cdot \frac{n^n}{(n+x)^n} \\ &= \left[\frac{(n+1+x)n}{(n+1)(n+x)} \right]^{n+1} \frac{n+x}{n} \\ &= \left[\frac{(n+1)n + nx}{(n+1)n + (n+1)x} \right]^{n+1} \frac{n+x}{n} \\ &= \left[1 - \frac{x}{(n+1)n + (n+1)x} \right]^{n+1} \frac{n+x}{n}. \end{aligned}$$

Aus der Bernoullischen Ungleichung, Beispiel I.2.7 (S. 22), 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

$$\frac{\left(1 + \frac{x}{n+1}\right)^{n+1}}{\left(1 + \frac{x}{n}\right)^n} \geq 1 \quad \implies \quad \left(1 + \frac{x}{n+1}\right)^{n+1} \geq \left(1 + \frac{x}{n}\right)^n.$$

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. 22), 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. 48), 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:

$$\begin{aligned}
 f^{(0)}(x) &= f(x) \\
 f^{(1)}(x) &= f'(x) \\
 f^{(n)}(x) &= \frac{d}{dx}f^{(n-1)}(x) \quad , n \geq 2.
 \end{aligned}$$

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.

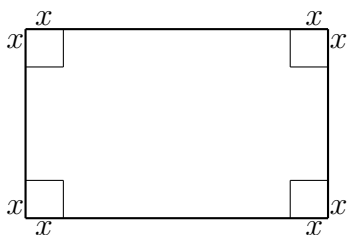


ABBILDUNG IV.2.1. Blechwanne

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:

$$\begin{aligned} 0 = V'(x) &= -2(16 - 2x)x - 2(10 - 2x)x + (10 - 2x)(16 - 2x) \\ &= 12x^2 - 104x + 160 \\ &= 12 \left(x^2 - \frac{26}{3}x + \frac{40}{3} \right) \\ &= 12 \left(\left(x - \frac{13}{3} \right)^2 - \frac{169}{9} + \frac{40}{3} \right) \\ &= 12 \left(\left(x - \frac{13}{3} \right)^2 - \frac{49}{9} \right) \\ &= 12 \left(x - \frac{20}{3} \right) \left(x - 2 \right). \end{aligned}$$

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:

$$f'(x) = 0 \text{ für alle } x \in (a, b) \iff f \text{ ist konstant}$$

bzw. für differenzierbare Funktionen f und g :

$$f'(x) = g'(x) \text{ für alle } x \in (a, b) \iff \text{es gibt eine Konstante } c \text{ mit } f(x) = g(x) + c \text{ für alle } x \in (a, b)$$

BEISPIEL IV.2.4. Wir behaupten, dass jede lineare Pendelbewegung, für 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 für 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

$$\dot{s}(t)^2 + \omega^2 s(t)^2 = c \quad \text{für alle } t.$$

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:

$$\begin{aligned} f'' > 0 &\implies \text{ Die Kurve } y = f(x) \text{ ist von unten konvex} \\ &\hspace{15em} \text{(Linkskrümmung).} \\ f'' < 0 &\implies \text{ Die Kurve } y = f(x) \text{ ist von oben konvex} \\ &\hspace{15em} \text{(Rechtsskrümmung).} \end{aligned}$$

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 :

$$0 = f(x_0) + f'(x_0)(x_1 - x_0) \quad \implies \quad x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

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.

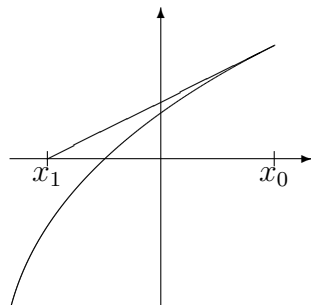


ABBILDUNG IV.2.2. Geometrische Interpretation des Newtonverfahrens

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 ARCCOSCOSINUS 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 $\ell b \leq BL$.

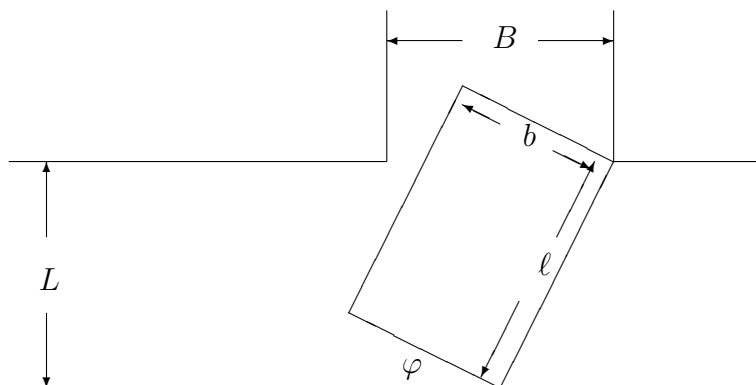


ABBILDUNG IV.3.1. Verkanteter Schrank

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. 126), 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.

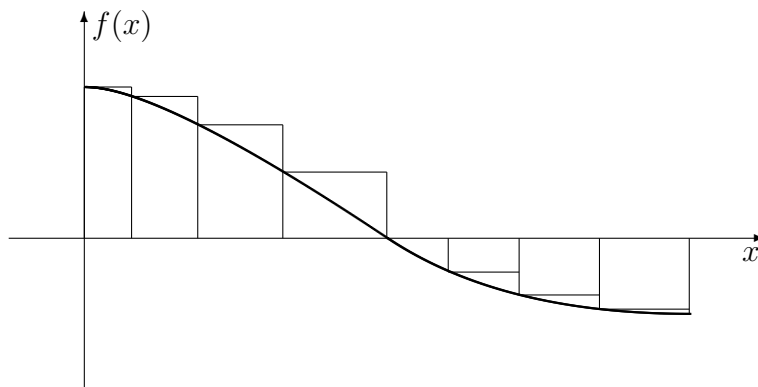


ABBILDUNG V.1.1. Z_n mit $\eta_i = x_{i-1}$ für alle i

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

$$\begin{aligned} \int_a^b u'(x)v(x)dx &= u(x)v(x) \Big|_a^b - \int_a^b u(x)v'(x)dx \\ &= (u(b)v(b) - u(a)v(a)) - \int_a^b u(x)v'(x)dx. \end{aligned}$$

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. 48), 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. 173)

$$\begin{aligned} \int \frac{t^2}{t^2 - 1} dt &= \int \frac{t^2 - 1 + 1}{t^2 - 1} dt \\ &= \int 1 + \frac{1}{t^2 - 1} dt \\ &= t + \int \frac{1}{t^2 - 1} dt \\ &= t + \frac{1}{2} \ln \left| \frac{t - 1}{t + 1} \right| + c. \end{aligned}$$

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

$$\begin{array}{r}
 x^8 \quad -2x^7 \quad +4x^6 \quad -6x^5 \quad +6x^4 \quad -6x^3 \quad +4x^2 \quad -2x \quad +1 \quad : \\
 x^8 \quad -2x^7 \quad +x^6 \\
 \hline
 3x^6 \quad -6x^5 \quad +6x^4 = \\
 3x^6 \quad -6x^5 \quad +3x^4 \\
 \hline
 3x^4 \quad -6x^3 \quad +4x^2 \\
 3x^4 \quad -6x^3 \quad +3x^2 \\
 \hline
 x^2 \quad -2x \quad +1 \\
 x^2 \quad -2x \quad +1 \\
 \hline
 0
 \end{array}$$

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

$$\begin{array}{r}
2x^5 \quad -7x^4 \quad +3x^3 \quad +14x^2 \quad -16x \quad +7 \quad : \quad x^4 - 5x^3 + 9x^2 - 7x + 2 \\
2x^5 \quad -10x^4 \quad +18x^3 \quad -14x^2 \quad +4x \quad \quad \quad = 2x + 3. \\
\hline
\quad \quad 3x^4 \quad -15x^3 \quad +28x^2 \quad -20x \quad +7 \\
\quad \quad 3x^4 \quad -15x^3 \quad +27x^2 \quad -21x \quad +6 \\
\hline
\quad \quad \quad \quad \quad \quad \quad x^2 \quad +x \quad +1
\end{array}$$

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

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1.$$

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

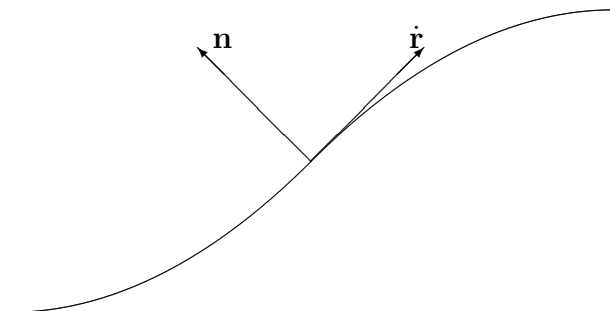
$$\mathbf{n}(t) = \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix}$$

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

$$x = x(t),$$

$$y = y(t)$$

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

$$x = x(t),$$

$$y = y(t)$$

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

$$\begin{aligned} F &= \frac{1}{2} \int_0^{2\pi} [(R \cos^3(t) + 2)3R \sin^2(t) \cos(t) \\ &\quad - (R \sin^3(t) + 3)3R \cos^2(t)(-\sin(t))]dt \\ &= \frac{1}{2} \int_0^{2\pi} [3R^2 \cos^4(t) \sin^2(t) + 3R^2 \sin^4(t) \cos^2(t) \\ &\quad + 6R \sin^2(t) \cos(t) + 9R \cos^2(t) \sin(t)]dt \\ &= \frac{1}{2} \int_0^{2\pi} 3R^2 \cos^2(t) \sin^2(t)dt + \underbrace{\frac{1}{2} \int_0^{2\pi} 6R \sin^2(t) \cos(t)dt}_{=0} \\ &\quad + \underbrace{\frac{1}{2} \int_0^{2\pi} 9R \cos^2(t) \sin(t)dt}_{=0} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \int_0^{2\pi} \frac{3}{4} R^2 \sin^2(2t) dt \quad \left| z = 2t \right. \\
&= \int_0^{4\pi} \frac{3}{16} R^2 \sin^2(z) dz \\
&= \frac{3}{8} R^2 \int_0^{2\pi} \sin^2(z) dz \quad \left| \text{Beispiel V.2.2(3)} \text{ (S. 175)} \right. \\
&= \frac{3}{8} \pi R^2 \quad \left| R = 4r \right. \\
&= 6\pi r^2.
\end{aligned}$$

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

$$\begin{aligned}
V &= \pi \int_{-r}^r (y_1^2 - y_2^2) dx \\
&= \pi \int_{-r}^r 4R\sqrt{r^2 - x^2} dx \quad \left| x = r \cos(t) \right. \\
&= 4\pi R \int_0^\pi \sqrt{r^2 - r^2 \cos^2(t)} r \sin(t) dt \\
&= 4\pi R r^2 \int_0^\pi \sin^2(t) dt \quad \left| \text{Beispiel V.2.2(3)} \text{ (S. 175)} \right.
\end{aligned}$$

$$= 2\pi^2 r^2 R.$$

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 Eindeigkeitssatz 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 Eindeigkeit 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. 218) 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

$$\Phi(x, y, h; f) = \frac{1}{h} \left[[Id - \frac{h}{2} f(t+h, \cdot)]^{-1} [y + \frac{h}{2} f(t, y)] - y \right],$$

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.

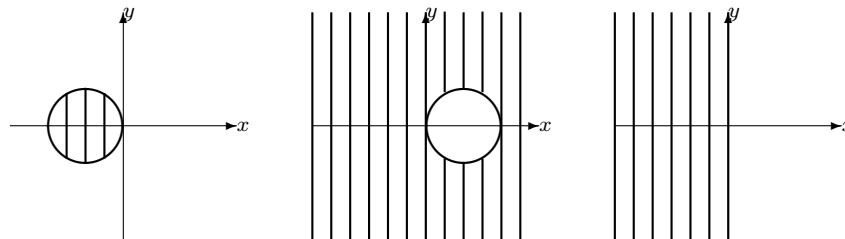


ABBILDUNG VI.4.1. Stabilitätsgebiete des expliziten und impliziten Eulerverfahrens und des Crank-Nicolson-Verfahrens (v.l.n.r.)

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.

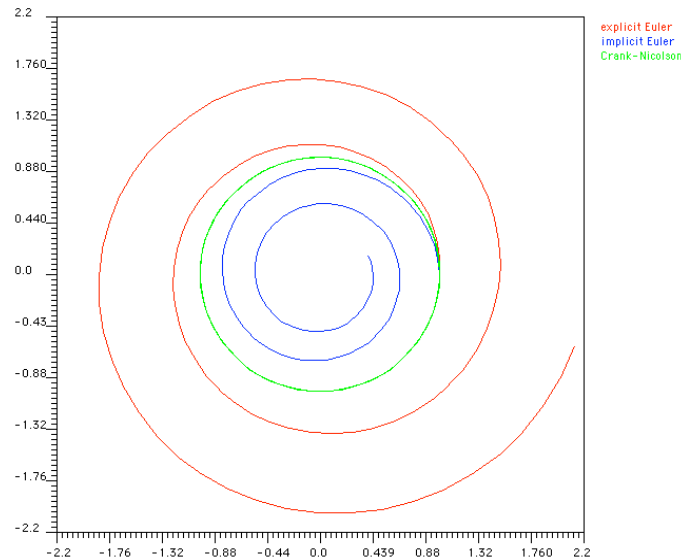


ABBILDUNG VI.4.2. Lösungskurven der beiden Eulerverfahren und des Crank-Nicolson-Verfahrens für das AWP aus Beispiel VI.4.2

KAPITEL VII

Potenzreihen

VII.1. Reihen

VII.1.1. Definition. Jeder Zahlenfolge $(a_n)_{n \in \mathbb{N}}$ ordnen wir eine neue Zahlenfolge $(s_n)_{n \in \mathbb{N}}$ zu durch die Vorschrift

$$s_n = \sum_{k=0}^n a_k = a_0 + \dots + a_n.$$

Diese neue Zahlenfolge nennen wir eine REIHE und bezeichnen sie mit $\sum_{k \in \mathbb{N}} a_k$. Die Zahlen a_0, a_1, \dots heißen die GLIEDER der Reihe; die Zahlen s_0, s_1, \dots heißen ihre PARTIALSUMMEN.

$$\sum_{k \in \mathbb{N}} a_k \text{ ist eine Abkürzung für } \left(\sum_{k=0}^n a_k \right)_{n \in \mathbb{N}}.$$

Die Reihe $\sum_{k \in \mathbb{N}} a_k$ heißt KONVERGENT, wenn die Folge $(s_n)_{n \in \mathbb{N}}$ ihrer Partialsummen konvergiert. Ist $s = \lim_{n \rightarrow \infty} s_n$, so schreiben wir $s = \sum_{k=0}^{\infty} a_k$.

$$s = \sum_{k=0}^{\infty} a_k \text{ ist eine Abkürzung für } s = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k.$$

BEISPIEL VII.1.1. Die GEOMETISCHE REIHE $\sum_{k \in \mathbb{N}} q^k$ besitzt die Glieder $a_k = q^k$ und die Partialsummen (vgl. Beispiel 1.2.6 (S. 22))

$$s_n = \sum_{k=0}^n q^k = \begin{cases} \frac{1-q^{n+1}}{1-q} & \text{falls } q \neq 1, \\ n+1 & \text{falls } q = 1. \end{cases}$$

Wegen der Beispiele III.2.4 (S. 118) und III.2.9 (S. 120) ist die geometrische Reihe genau dann konvergent, wenn $|q| < 1$ ist. In diesem Fall ist

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad \text{falls } |q| < 1.$$

BEISPIEL VII.1.2. Die HARMONISCHE REIHE $\sum_{k \in \mathbb{N}^*} \frac{1}{k}$ hat die Glieder $a_k = \frac{1}{k}$. Die harmonische Reihe ist nicht konvergent.

DENN: Da alle Glieder positiv sind, ist die Folge der Partialsummen monoton wachsend. Für $m \in \mathbb{N}^*$ ist

$$\begin{aligned} s_{2^m} &= \sum_{k=1}^{2^m} \frac{1}{k} \\ &= 1 + \sum_{\ell=1}^m \left(\sum_{k=2^{\ell-1}+1}^{2^\ell} \underbrace{\frac{1}{k}}_{\geq 2^{-\ell}} \right) \\ &\geq \sum_{\ell=1}^m \left(\sum_{k=2^{\ell-1}+1}^{2^\ell} 2^{-\ell} \right) \quad \text{je } 2^\ell - 2^{\ell-1} \text{ Summanden} \\ &= \sum_{\ell=1}^m \underbrace{(2^\ell - 2^{\ell-1})}_{=\frac{1}{2}} 2^{-\ell} \\ &= \frac{m}{2}. \end{aligned}$$

Also ist die Folge der Partialsummen nicht beschränkt und damit auch nicht konvergent.

BEISPIEL VII.1.3. Die ALTERNIERENDE HARMONISCHE REIHE $\sum_{k \in \mathbb{N}^*} (-1)^{k-1} \frac{1}{k}$ hat die Glieder $a_k = (-1)^{k-1} \frac{1}{k}$. Wegen des Leibniz-Kriteriums (s.u.) ist sie konvergent.

VII.1.2. Absolute Konvergenz. Die Reihe $\sum_{k \in \mathbb{N}} a_k$ heißt ABSOLUT KONVERGENT, wenn die Reihe $\sum_{k \in \mathbb{N}} |a_k|$ konvergiert. Eine konvergente Reihe, die nicht absolut konvergiert, heißt BEDINGT KONVERGENT.

BEISPIEL VII.1.4. Die alternierende harmonische Reihe ist bedingt konvergent.

BEISPIEL VII.1.5. Die Reihe $\sum_{k \in \mathbb{N}} (-1)^k 2^{-k}$ ist konvergent und absolut konvergent.

Es gilt folgender Zusammenhang zwischen Konvergenz und absoluter Konvergenz:

Jede absolut konvergente Reihe ist konvergent.

VII.1.3. Konvergenzkriterien. Wir stellen im Folgenden einige Kriterien für die Konvergenz von Reihen zusammen.

Die Glieder einer konvergenten Reihe bilden eine Nullfolge:

$$\sum_{k \in \mathbb{N}} a_k \text{ konvergent} \implies (a_n)_{n \in \mathbb{N}} \text{ ist Nullfolge.}$$

BEMERKUNG VII.1.6. Die harmonische Reihe zeigt, dass die Umkehrung dieser Folgerung nicht gilt.

BEISPIEL VII.1.7. $\sum_{k \in \mathbb{N}^*} \frac{k^k}{k!}$ ist nicht konvergent, denn für alle $k \in \mathbb{N}^*$ ist $\frac{k^k}{k!} \geq 1$.

Ein wichtiges Konvergenzkriterium für alternierende Reihen ist:

LEIBNIZ-KRITERIUM: Für jede monoton fallende Nullfolge $(a_n)_{n \in \mathbb{N}}$ konvergiert die zugehörige alternierende Reihe

$$\sum_{k \in \mathbb{N}} (-1)^k a_k.$$

Ein wichtiges Vergleichskriterium für Reihen ist:

MAJORANTENKRITERIUM: Besteht für die Reihenglieder die Abschätzung

$$0 \leq |a_k| \leq b_k \text{ für alle } k \geq k_0,$$

dann gilt

$$\sum_{k \in \mathbb{N}} b_k \text{ konvergent} \implies \sum_{k \in \mathbb{N}} a_k \text{ absolut konvergent}$$

$$\sum_{k \in \mathbb{N}} |a_k| \text{ nicht konvergent} \implies \sum_{k \in \mathbb{N}} b_k \text{ nicht konvergent.}$$

BEISPIEL VII.1.8. Die Reihe $\sum_{k \in \mathbb{N}^*} \frac{1}{k^\alpha}$ ist genau dann konvergent, wenn $\alpha > 1$ ist.

DENN: Ist $\alpha \leq 1$, gilt $\frac{1}{k} \leq \frac{1}{k^\alpha}$ für alle $k \in \mathbb{N}^*$, und aus dem Majorantenkriterium und Beispiel VII.1.2 folgt, dass $\sum_{k \in \mathbb{N}^*} \frac{1}{k^\alpha}$ nicht konvergiert. Sei nun $\alpha > 1$. Die Folge der Partialsummen ist monoton wachsend, da die Glieder der Reihe allesamt positiv sind. Wir müssen daher nur zeigen, dass die Folge der Partialsummen beschränkt ist. Dazu gehen wir ähnlich wie in Beispiel VII.1.2 vor. Für jedes $m \in \mathbb{N}^*$ erhalten wir

$$s_{2^m} = \sum_{k=1}^{2^m} \frac{1}{k^\alpha}$$

$$\begin{aligned}
&= 1 + \sum_{\ell=1}^m \left(\sum_{k=2^{\ell-1}+1}^{2^\ell} \underbrace{\frac{1}{k^\alpha}}_{\leq 2^{-\alpha\ell+\alpha}} \right) \\
&\leq 1 + \sum_{\ell=1}^m \left(\sum_{k=2^{\ell-1}+1}^{2^\ell} 2^{-\alpha\ell+\alpha} \right) \quad \text{je } 2^\ell - 2^{\ell-1} \text{ Summanden} \\
&= 1 + \sum_{\ell=1}^m (2^\ell - 2^{\ell-1}) 2^{-\alpha\ell+\alpha} \quad 2^\ell - 2^{\ell-1} = \frac{1}{2} \cdot 2^\ell \\
&= 1 + \sum_{\ell=1}^m \frac{1}{2} 2^{-(\alpha-1)\ell+\alpha} \\
&= 1 + 2^{\alpha-1} \sum_{\ell=0}^m (2^{-(\alpha-1)})^\ell \quad \text{Beispiel VII.1.1 mit } q = 2^{-(\alpha-1)} \\
&= 1 + \frac{2^{\alpha-1}}{1 - 2^{-(\alpha-1)}}.
\end{aligned}$$

Hieraus folgt die Behauptung.

Ein anderes wichtiges Konvergenzkriterium ist:

QUOTIENTENKRITERIUM: Ist $a_k \neq 0$ für alle $k \geq k_0$ und konvergiert die Folge $(\frac{a_{k+1}}{a_k})_{k \geq k_0}$, so gilt:

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| < 1 \implies \sum_{k \in \mathbb{N}} a_k \text{ konvergiert absolut.}$$

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| > 1 \implies \sum_{k \in \mathbb{N}} a_k \text{ konvergiert nicht.}$$

BEMERKUNG VII.1.9. An den Beispielen der harmonischen Reihe und der alternierenden harmonischen Reihe erkennt man, dass im Fall $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = 1$ keine Aussage möglich ist.

BEISPIEL VII.1.10. Die Reihen

$$\begin{aligned}
&\sum_{k \in \mathbb{N}} \frac{x^k}{k!}, \\
&\sum_{k \in \mathbb{N}} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \\
&\sum_{k \in \mathbb{N}} (-1)^k \frac{x^{2k}}{(2k)!}
\end{aligned}$$

konvergieren für jedes $x \in \mathbb{R}$ absolut.

DENN: Für $x = 0$ sind die Glieder bis auf dasjenige zu $k = 0$ alle gleich Null, und die Behauptung ist offensichtlich. Für $x \neq 0$ gilt jeweils

$$\lim_{k \rightarrow \infty} \left| \frac{\frac{x^{k+1}}{(k+1)!}}{\frac{x^k}{k!}} \right| = \lim_{k \rightarrow \infty} \frac{|x|}{k+1} = 0,$$

$$\lim_{k \rightarrow \infty} \left| \frac{(-1)^{k+1} \frac{x^{2k+3}}{(2k+3)!}}{(-1)^k \frac{x^{2k+1}}{(2k+1)!}} \right| = \lim_{k \rightarrow \infty} \frac{|x|^2}{(2k+2)(2k+3)} = 0,$$

$$\lim_{k \rightarrow \infty} \left| \frac{(-1)^{k+1} \frac{x^{2k+2}}{(2k+2)!}}{(-1)^k \frac{x^{2k}}{(2k)!}} \right| = \lim_{k \rightarrow \infty} \frac{|x|^2}{(2k+1)(2k+2)} = 0,$$

und die Behauptung folgt aus dem Quotientenkriterium.

VII.1.4. Rechenregeln. Konvergente Reihen können gliedweise addiert, subtrahiert und mit einem konstanten Faktor multipliziert werden:

$$\begin{aligned} \sum_{k=0}^{\infty} a_k = a, \quad \sum_{k=0}^{\infty} b_k = b \\ \implies \sum_{k=0}^{\infty} (a_k \pm b_k) = a \pm b, \\ \sum_{k=0}^{\infty} (ca_k) = ca. \end{aligned}$$

WARNUNG: Elementare Manipulationen, die bei endlichen Summen den Summenwert nicht ändern, sind bei Reihen *nicht uneingeschränkt* erlaubt. Z.B. können nicht beliebig Klammern gesetzt, Klammern fortgelassen oder Summanden vertauscht werden. Diese Manipulationen sind *nur* bei absolut konvergenten Reihen erlaubt.

CAUCHY-PRODUKT: Für *absolut* konvergente Reihen $\sum_{k \in \mathbb{N}} a_k$ und $\sum_{k \in \mathbb{N}} b_k$ gilt die Produktformel

$$\left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right) = \sum_{k=0}^{\infty} \left(\sum_{\ell=0}^k a_{\ell} b_{k-\ell} \right).$$

VII.2. Potenzreihen

VII.2.1. Definition. Eine POTENZREIHE ist eine Reihe der Form

$$\sum_{k \in \mathbb{N}} a_k x^k$$

mit $x \in \mathbb{R}$ (variabel!) und $a_0, a_1, \dots \in \mathbb{R}$ (fest!). Die Zahlen a_0, a_1, \dots heißen die KOEFFIZIENTEN der Potenzreihe.

BEISPIEL VII.2.1. Die Potenzreihe $\sum_{k \in \mathbb{N}} x^k$ hat die Koeffizienten $a_k = 1$ ($k \in \mathbb{N}$).

Die Potenzreihe $\sum_{k \in \mathbb{N}} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$ hat die Koeffizienten $a_0 = 0, a_1 = 1, a_2 = 0, a_3 = -\frac{1}{6}, \dots$ (allgemein: $a_{2k} = 0, a_{2k+1} = \frac{(-1)^k}{(2k+1)!}$).

VII.2.2. Konvergenzradius. An einer Potenzreihe $\sum_{k \in \mathbb{N}} a_k x^k$ interessieren zu allererst die Menge

$$M = \left\{ x \in \mathbb{R} : \sum_{k \in \mathbb{N}} a_k x^k \text{ ist konvergent} \right\}$$

und die Zahl

$$R = \begin{cases} \sup\{|x| : x \in M\} & \text{falls } M \text{ beschränkt ist,} \\ \infty & \text{falls } M \text{ unbeschränkt ist.} \end{cases}$$

Man nennt R den KONVERGENZRADIUS der Potenzreihe. Es gibt drei Möglichkeiten:

$$R = 0, \quad 0 < R < \infty \quad \text{oder} \quad R = \infty.$$

BEISPIEL VII.2.2. Die Potenzreihe $\sum_{k \in \mathbb{N}} x^k$ hat den Konvergenzradius 1. Denn gemäß Beispiel VII.1.1 (S. 255) ist sie genau dann konvergent, wenn $|x| < 1$ ist.

BEISPIEL VII.2.3. Die drei Reihen aus Beispiel VII.1.10 (S. 258) sind für alle $x \in \mathbb{R}$ konvergent und haben daher den Konvergenzradius ∞ .

BEISPIEL VII.2.4. Die Potenzreihe $\sum_{k \in \mathbb{N}} k! x^k$ ist für $x = 0$ konvergent, da alle Glieder bis auf das Erste gleich Null sind. Für $x \neq 0$ ist sie nicht konvergent, da dann $(k! x^k)_{k \in \mathbb{N}}$ keine Nullfolge ist. Also hat diese Reihe den Konvergenzradius 0.

Für eine Potenzreihe $\sum_{k \in \mathbb{N}} a_k x^k$ mit Konvergenzradius R gilt:

- (1) $R = 0 \iff$ Die Reihe konvergiert nur für $x = 0$.
- (2) Ist $R > 0$, konvergiert die Reihe $\sum_{k \in \mathbb{N}} a_k x^k$ für alle $x \in \mathbb{R}$ mit $|x| < R$.
- (3) Für alle $x \in \mathbb{R}$ mit $|x| > R$ ist die Reihe $\sum_{k \in \mathbb{N}} a_k x^k$ nicht konvergent.

VII.2.3. Bestimmung des Konvergenzradius. Der Konvergenzradius R einer Potenzreihe $\sum_{k \in \mathbb{N}} a_k x^k$ kann häufig mit einem der folgenden drei „Tricks“ berechnet werden:

- (1) Es gibt ein $k_0 \in \mathbb{N}$ mit $a_k \neq 0$ für alle $k \geq k_0$ und $\lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right|$ existiert oder ist ∞ , dann ist

$$R = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right|.$$

- (2) In regelmäßigen Abständen sind einer oder mehrere Koeffizienten gleich Null, d.h. $\sum_{k \in \mathbb{N}} a_k x^k = \sum_{n \in \mathbb{N}} b_n x^{r+\ell n}$ mit $r \geq 0$, $\ell \geq 1$ und $b_n \neq 0$ für alle $n \in \mathbb{N}$. Falls $\lim_{n \rightarrow \infty} \left| \frac{b_n}{b_{n+1}} \right|$ existiert oder gleich ∞ ist, gilt

$$R = \sqrt[\ell]{\lim_{n \rightarrow \infty} \left| \frac{b_n}{b_{n+1}} \right|}.$$

(Dabei ist $\sqrt[\ell]{\infty} = \infty$ zu setzen!)

- (3) $R = \sup B$ mit $B = \{r \geq 0 : (|a_k| r^k)_{k \in \mathbb{N}} \text{ ist beschränkt}\}$.

BEISPIEL VII.2.5. Für die Potenzreihe $\sum_{k \in \mathbb{N}} \frac{k^k}{k!} x^k$ erhalten wir

$$\begin{aligned} \left| \frac{a_k}{a_{k+1}} \right| &= \frac{\frac{k^k}{k!}}{\frac{(k+1)^{k+1}}{(k+1)!}} \\ &= \frac{k^k}{(k+1)^{k+1}} (k+1) \\ &= \frac{1}{\left(\frac{k+1}{k}\right)^k} \\ &= \frac{1}{\left(1 + \frac{1}{k}\right)^k} \\ &\xrightarrow{k \rightarrow \infty} \frac{1}{e}. \end{aligned}$$

Also hat sie den Konvergenzradius $\frac{1}{e}$.

BEISPIEL VII.2.6. Die Potenzreihen

$$\sum_{k \in \mathbb{N}} x^k, \quad \sum_{k \in \mathbb{N}^*} \frac{1}{k} x^k, \quad \sum_{k \in \mathbb{N}^*} \frac{1}{k^2} x^k$$

haben alle den Konvergenzradius 1 (es ist jeweils $\lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| = 1$). Sie haben aber ein sehr unterschiedliches Verhalten in den Randpunkten $x = -1$ und $x = 1$:

$$\sum_{k \in \mathbb{N}} x^k \quad \text{ist für } x = -1 \text{ und } x = 1 \text{ nicht konvergent,}$$

$\sum_{k \in \mathbb{N}} \frac{1}{k} x^k$ ist für $x = 1$ nicht konvergent

und für $x = -1$ bedingt konvergent,

$\sum_{k \in \mathbb{N}} \frac{1}{k^2} x^k$ ist für $x = -1$ und $x = 1$ absolut konvergent.

BEISPIEL VII.2.7. Für die Potenzreihe $\sum_{k \in \mathbb{N}^*} (-1)^{k+1} \frac{1}{k^5 k} x^{2k+1}$ erhalten wir mit unserem zweiten „Trick“ den Konvergenzradius

$$R = \sqrt{\lim_{n \rightarrow \infty} \frac{(n+1)5^{n+1}}{n5^n}} = \sqrt{5}.$$

BEISPIEL VII.2.8. Die Koeffizientenfolge der Potenzreihe $\sum_{k \in \mathbb{N}} x^{k^2}$ weist immer größer werdende Lücken auf. Wir können daher die ersten beiden „Tricks“ nicht anwenden. Wir versuchen also den dritten. Es ist $(|a_k| r^k)_{k \in \mathbb{N}} = (r^{k^2})_{k \in \mathbb{N}}$. Diese Folge ist offensichtlich genau dann beschränkt, wenn $r \leq 1$ ist. Also ist der Konvergenzradius 1.

VII.2.4. Differentiation und Integration von Potenzreihen.

Im Folgenden ist stets $\sum_{k \in \mathbb{N}} a_k x^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Die Vorschrift

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

definiert dann eine Funktion $f : (-R, R) \rightarrow \mathbb{R}$. Wir sagen, die Funktion f wird durch die Potenzreihe dargestellt.

DIFFERENTIATION VON POTENZREIHEN: Eine durch eine Potenzreihe dargestellte Funktion f ist im offenen Konvergenzintervall $(-R, R)$ beliebig oft differenzierbar. Die Ableitungen können durch gliedweise Differentiation bestimmt werden, d.h. für alle $x \in (-R, R)$, und alle $n \in \mathbb{N}$ ist

$$f^{(n)}(x) = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} a_k x^{k-n}.$$

BEISPIEL VII.2.9. Für die geometrische Reihe $\sum_{k \in \mathbb{N}} x^k$ aus Beispiel VII.1.2 (S. 256) erhalten wir für alle $x \in (-1, 1)$:

$$\begin{aligned} \frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k, \\ \frac{1}{(1-x)^2} &= \frac{d}{dx} \left(\frac{1}{1-x} \right) \\ &= \sum_{k=1}^{\infty} k x^{k-1}, \end{aligned}$$

$$\begin{aligned}\frac{2}{(1-x)^3} &= \frac{d^2}{dx^2} \left(\frac{1}{1-x} \right) \\ &= \sum_{k=2}^{\infty} k(k-1)x^{k-2}.\end{aligned}$$

INTEGRATION VON POTENZREIHEN: Potenzreihen können im offenen Konvergenzintervall $(-R, R)$ gliedweise integriert werden, d.h.

$$\int \left\{ \sum_{k=0}^{\infty} a_k x^k \right\} dx = \sum_{k=0}^{\infty} \frac{1}{k+1} a_k x^{k+1}.$$

Insbesondere ist für alle $-R < \alpha < \beta < R$

$$\int_{\alpha}^{\beta} \left\{ \sum_{k=0}^{\infty} a_k x^k \right\} dx = \sum_{k=0}^{\infty} \frac{1}{k+1} a_k (\beta^{k+1} - \alpha^{k+1}).$$

BEISPIEL VII.2.10. Durch Integration der geometrischen Reihe $\sum_{k \in \mathbb{N}} x^k$ ergibt sich für alle $x \in (-1, 1)$:

$$-\ln(1-x) = \int_0^x \frac{1}{1-t} dt = \sum_{k=0}^{\infty} \frac{1}{k+1} x^{k+1}.$$

VII.2.5. Potenzreihendarstellung einiger Funktionen. Als erste Anwendung der Sätze des vorigen Abschnittes ergibt sich:

$$\begin{aligned}\exp(x) &= \sum_{k=0}^{\infty} \frac{1}{k!} x^k && \text{für alle } x \in \mathbb{R} \\ \sin(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} && \text{für alle } x \in \mathbb{R} \\ \cos(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} && \text{für alle } x \in \mathbb{R} \\ \ln(1+x) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} && \text{für alle } x \in (-1, 1) \\ \arctan(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} && \text{für alle } x \in (-1, 1).\end{aligned}$$

Zum Nachweis dieser Identitäten gehen wir wie folgt vor:
Wegen

$$\lim_{k \rightarrow \infty} \frac{\frac{1}{k!}}{\frac{1}{(k+1)!}} = \lim_{k \rightarrow \infty} (k+1) = \infty$$

hat die erste Potenzreihe den Konvergenzradius ∞ und stellt damit eine auf ganz \mathbb{R} differenzierbare Funktion f dar. Gliedweises Differenzieren liefert für alle $x \in \mathbb{R}$

$$\begin{aligned} f'(x) &= \sum_{k=1}^{\infty} \frac{1}{k!} k x^{k-1} \\ &= \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1} \\ &= f(x). \end{aligned}$$

Also ist (vgl. Abschnitt IV.4.1 (S. 161)) $f(x) = c \exp(x)$. Wegen $f(0) = 1$ ist $c = 1$.

Die zweite Potenzreihe hat ebenfalls den Konvergenzradius ∞ und stellt eine auf ganz \mathbb{R} beliebig oft differenzierbare Funktion g dar. Indem wir die Potenzreihe zweimal gliedweise differenzieren, erhalten wir

$$g''(x) = -g(x)$$

für alle $x \in \mathbb{R}$. Aus Beispiel IV.2.4 (S. 143) folgt

$$g(x) = g(0) \cos(x) + g'(0) \sin(x) = \sin(x).$$

Die Argumentation bei der dritten Potenzreihe ist völlig analog zu derjenigen bei der zweiten.

Die Aussagen für die vierte und fünfte Potenzreihe erhalten wir durch Integration der Identitäten

$$\begin{aligned} \frac{1}{1+x} &= \sum_{k=0}^{\infty} (-1)^k x^k, \\ \frac{1}{1+x^2} &= \sum_{k=0}^{\infty} (-1)^k x^{2k}. \end{aligned}$$

VII.2.6. Die Binomialreihe. Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ definieren wir den Binomialkoeffizienten $\binom{\alpha}{k}$ durch

$$\binom{\alpha}{k} = \begin{cases} \frac{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)}{k!} & \text{falls } k \geq 1 \\ 1 & \text{falls } k = 0. \end{cases}$$

Dies ist offensichtlich eine Verallgemeinerung des Binomialkoeffizienten $\binom{n}{k}$ aus Abschnitt I.2.8 (S. 23). Die Potenzreihe $\sum_{k \in \mathbb{N}} \binom{\alpha}{k} x^k$ mit $\alpha \in \mathbb{R}$

heißt BINOMIALREIHE. Wegen

$$\begin{aligned} \lim_{k \rightarrow \infty} \left| \frac{\binom{\alpha}{k}}{\binom{\alpha}{k+1}} \right| &= \lim_{k \rightarrow \infty} \left| \frac{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)}{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1) \cdot (\alpha-k)} \frac{(k+1)!}{k!} \right| \\ &= \lim_{k \rightarrow \infty} \frac{k+1}{|\alpha-k|} \\ &= 1 \end{aligned}$$

hat sie den Konvergenzradius 1 und stellt damit auf dem Intervall $(-1, 1)$ eine beliebig oft differenzierbare Funktion g dar. Durch gliedweise Differentiation erhalten wir

$$\begin{aligned} (1+x)g'(x) &= \sum_{k=1}^{\infty} (1+x) \binom{\alpha}{k} k x^{k-1} \\ &= \underbrace{\sum_{k=1}^{\infty} \binom{\alpha}{k} k x^{k-1}}_{=\sum_{\ell=0}^{\infty} \binom{\alpha}{\ell+1} (\ell+1) x^{\ell}} + \sum_{k=1}^{\infty} \binom{\alpha}{k} k x^k \\ &= \binom{\alpha}{1} + \sum_{k=1}^{\infty} \underbrace{\left[\binom{\alpha}{k+1} (k+1) + \binom{\alpha}{k} k \right]}_{=\alpha \binom{\alpha}{k}} x^k \\ &= \alpha \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \\ &= \alpha g(x). \end{aligned}$$

Für die Funktion

$$h(x) = \frac{g(x)}{(1+x)^{\alpha}}$$

ergibt sich daher $h'(x) = 0$, also $h(x) = h(0) = 1$. Insgesamt erhalten wir:

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad \text{für alle } x \in (-1, 1), \alpha \in \mathbb{R}.$$

VII.2.7. Potenzreihen mit beliebigem Entwicklungspunkt.

Eine Reihe der Form

$$\sum_{k \in \mathbb{N}} a_k (x - x_0)^k$$

mit $a_0, a_1, \dots \in \mathbb{R}$ und $x_0 \in \mathbb{R}$ heißt POTENZREIHE MIT ENTWICKLUNGSPUNKT x_0 . Alle früher betrachteten Potenzreihen hatten den Entwicklungspunkt $x_0 = 0$. Durch die Transformation $z = x - x_0$ geht eine Potenzreihe mit Entwicklungspunkt x_0 in eine Potenzreihe

$\sum_{k \in \mathbb{N}} a_k z^k$ vom bisher betrachteten Typ über. Hieran erkennt man, dass sich alle bisherigen Aussagen übertragen. Insbesondere bezeichnet man als Konvergenzradius der Potenzreihe $\sum_{k \in \mathbb{N}} a_k (x - x_0)^k$ den Konvergenzradius R der Potenzreihe $\sum_{k \in \mathbb{N}} a_k z^k$. Wegen

$$|x - x_0| < R \iff x_0 - R < x < x_0 + R$$

gilt

$x \in (x_0 - R, x_0 + R) \implies \sum_{k \in \mathbb{N}} a_k (x - x_0)^k \text{ konvergiert}$ $x \notin [x_0 - R, x_0 + R] \implies \sum_{k \in \mathbb{N}} a_k (x - x_0)^k \text{ konvergiert nicht.}$

BEISPIEL VII.2.11. Wegen $\exp(x) = \exp(x - x_0) \exp(x_0)$ wird die Exponentialfunktion durch die Potenzreihe

$$\sum_{k \in \mathbb{N}} \frac{e^{x_0}}{k!} (x - x_0)^k$$

mit Entwicklungspunkt x_0 dargestellt.

VII.2.8. Koeffizientenvergleich. Zwei Potenzreihen

$$\sum_{k \in \mathbb{N}} a_k x^k \quad \text{und} \quad \sum_{k \in \mathbb{N}} b_k x^k$$

stellen genau dann dieselbe Funktion dar, wenn alle ihre Koeffizienten übereinstimmen, d.h. $a_k = b_k$ für alle $k \in \mathbb{N}$.

BEISPIEL VII.2.12. Aus Abschnitt VII.2.6 und dem Cauchy-Produkt aus Abschnitt VII.1.4 (S. 259) folgt für $\alpha, \beta \in \mathbb{R}$ und $x \in (-1, 1)$

$$\begin{aligned} \sum_{k=0}^{\infty} \binom{\alpha + \beta}{k} x^k &= (1 + x)^{\alpha + \beta} \\ &= (1 + x)^{\alpha} (1 + x)^{\beta} \\ &= \left\{ \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \right\} \left\{ \sum_{k=0}^{\infty} \binom{\beta}{k} x^k \right\} \\ &= \sum_{k=0}^{\infty} \left\{ \sum_{\ell=0}^k \binom{\alpha}{\ell} \binom{\beta}{k - \ell} \right\} x^k. \end{aligned}$$

Also gilt für alle $\alpha, \beta \in \mathbb{R}$ und alle $k \in \mathbb{N}$

$$\binom{\alpha + \beta}{k} = \sum_{\ell=0}^k \binom{\alpha}{\ell} \binom{\beta}{k - \ell}.$$

VII.3. Taylorreihen

VII.3.1. Die Taylor-Formel. Wir betrachten eine beliebig oft differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ auf einem Intervall I und zwei verschiedene Punkte x, x_0 in I . Aus dem Hauptsatz der Differential- und Integralrechnung aus Abschnitt V.1.3 (S. 171) folgt

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt.$$

Auf das Integral wenden wir partielle Integration an mit

$$u(t) = -(x - t), \quad u'(t) = 1, \quad v(t) = f'(t)$$

und erhalten

$$\begin{aligned} f(x) &= f(x_0) - (x - t)f'(t) \Big|_{t=x_0}^{t=x} + \int_{x_0}^x f''(t)(x - t) dt \\ &= f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^x f''(t)(x - t) dt. \end{aligned}$$

Erneute partielle Integration mit

$$u(t) = -\frac{1}{2}(x - t)^2, \quad u'(t) = x - t, \quad v(t) = f''(t)$$

ergibt

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 \\ &\quad + \int_{x_0}^x f'''(t) \frac{1}{2}(x - t)^2 dt. \end{aligned}$$

Führen wir diese Prozedur fort, erhalten wir nach n -Schritten

$$f(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k + \int_{x_0}^x f^{(n+1)}(t) \frac{1}{n!} (x - t)^n dt.$$

Nach dem Mittelwertsatz der Integralrechnung aus Abschnitt V.1.2 (S. 170) gibt es ein η zwischen x und x_0 mit

$$\begin{aligned} \int_{x_0}^x \frac{1}{n!} f^{(n+1)}(t)(x - t)^n dt &= f^{(n+1)}(\eta) \int_{x_0}^x \frac{1}{n!} (x - t)^n dt \\ &= \frac{1}{(n+1)!} f^{(n+1)}(\eta)(x - x_0)^{n+1}. \end{aligned}$$

Das Polynom

$$T_n(f; x_0)(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$$

heißt das n -TE TAYLOR-POLYNOM VON f MIT ENTWICKLUNGSPUNKT x_0 .

Unsere obigen Überlegungen beweisen:

TAYLOR-FORMEL: Für jede auf dem Intervall I $n + 1$ -mal differenzierbare Funktion f und jeden Punkt $x_0 \in I$ gilt

$$f(x) = T_n(f; x_0)(x) + R_{n+1}(f; x_0, x)$$

für alle $x \in I$ mit

$$R_{n+1}(f; x_0, x) = \int_{x_0}^x \frac{1}{n!} f^{(n+1)}(t)(x-t)^n dt$$

und

$$R_{n+1}(f; x_0, x) = \frac{1}{(n+1)!} f^{(n+1)}(\eta)(x-x_0)^{n+1}$$

für ein η zwischen x_0 und x .

Wir können die Taylor-Formel wie folgt deuten: Die Funktion f und das Taylor-Polynom $T_n(f; x_0)$ beschreiben zwei Kurven. Beide Kurven gehen durch den Punkt $(x_0, f(x_0))$. In diesem Punkt haben sie die gleiche Steigung, die gleiche Krümmung und alle Eigenschaften gemeinsam, die durch die Ableitungen bis zur Ordnung n in diesem Punkt bestimmt werden. Das Restglied $R_{n+1}(f; x_0, x)$ erlaubt die Abschätzung des Fehlers $f(x) - T_n(f; x_0)(x)$ in einem von x_0 verschiedenen Punkt.

BEISPIEL VII.3.1. Wegen $\exp^{(n)} = \exp$ ist

$$T_n(\exp; 0)(x) = \sum_{k=0}^n \frac{1}{k!} x^k$$

und

$$R_{n+1}(\exp; 0, x) = \frac{\exp(\eta)}{(n+1)!} x^{n+1}.$$

Für $|x| \leq 1$ ergibt sich hieraus z.B.

$$\left| e^x - \sum_{k=0}^n \frac{1}{k!} x^k \right| \leq \frac{e}{(n+1)!} |x|^{n+1}.$$

Wollen wir also z.B. $e^{0.1}$ mit einem Fehler von $\pm 10^{-6}$ berechnen, müssen wir n so groß wählen, dass

$$\frac{e}{(n+1)!} 10^{-(n+1)} \leq 10^{-6}$$

ist. Wegen

$$\frac{e}{5!} 10^{-5} \leq \frac{3}{120} 10^{-5} = 2.5 \cdot 10^{-7}$$

ist dies sicherlich für $n = 4$ erfüllt und wir erhalten

$$e^{0.1} = 1 + 0.1 + \frac{1}{2} 0.01 + \frac{1}{6} 0.001 + \frac{1}{24} 0.0001 \pm 10^{-6}.$$

Eine Folgerung aus der Taylor-Formel ist:

Jede auf einem Intervall $n+1$ -mal differenzierbare Funktion f mit verschwindender $(n+1)$ -ter Ableitung ist ein Polynom vom Grad $\leq n$.

Eine weitere Konsequenz ist der folgende EXTREMWERT-TEST:

Die Funktion f sei auf dem Intervall I n -mal differenzierbar und für $x_0 \in I$ gelte

$$f'(x_0) = \dots = f^{(n-1)}(x_0) = 0, f^{(n)}(x_0) \neq 0.$$

Dann gilt:

- (1) x_0 ist eine Extremalstelle von $f \iff n$ ist gerade.
- (2) n gerade und $f^{(n)}(x_0) < 0 \implies x_0$ ist lokales Maximum.
 n gerade und $f^{(n)}(x_0) > 0 \implies x_0$ ist lokales Minimum.

BEISPIEL VII.3.2. Für

$$f(x) = x^4 - 8x^3 + 24x^2 - 32x + 19$$

gilt

$$f'(2) = f''(2) = f'''(2) = 0$$

und

$$f^{(4)}(2) = 24.$$

Also ist $x_0 = 2$ ein lokales Minimum von f .

VII.3.2. Die Taylor-Reihe. Die Funktion f sei auf dem Intervall I beliebig oft differenzierbar und x_0 sei ein Punkt aus I .

Die Potenzreihe

$$T(f; x_0)(x) = \sum_{k \in \mathbb{N}} \frac{1}{k!} f^{(k)}(x_0) (x - x_0)^k$$

heißt die TAYLOR-REIHE VON f MIT ENTWICKLUNGSPUNKT x_0 .

Gibt es ein $R > 0$ mit

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0) (x - x_0)^k \quad \text{für alle } |x - x_0| < R,$$

so sagt man, „ f lässt sich in x_0 als Taylor-Reihe darstellen“ oder „ f lässt sich in x_0 in eine Taylor-Reihe entwickeln“.

Wegen Abschnitt VII.2.8 (S. 266) gilt:

Eine Funktion f lässt sich genau dann als eine Potenzreihe mit Entwicklungspunkt x_0 darstellen, wenn sich f in x_0 in eine Taylor-Reihe entwickeln lässt.

WARNUNG: Selbst wenn die Taylor-Reihe von f existiert und einen positiven Konvergenzradius hat, muss sie *nicht* die Funktion f darstellen.

Es gilt:

Die Funktion f sei auf dem Intervall I beliebig oft differenzierbar und x_0 sei ein Punkt aus I . Dann wird f durch seine Taylor-Reihe mit Entwicklungspunkt x_0 dargestellt, wenn das Restglied $R_{n+1}(f; x_0, x)$ aus der Taylor-Formel für $n \rightarrow \infty$ gegen Null konvergiert. Dies ist sicherlich dann der Fall, wenn es Konstanten $A, B \in \mathbb{R}$ gibt, so dass für alle $x \in I$ gilt $|f^{(n)}(x)| \leq AB^n$.

VII.3.3. Methoden der Potenzreihenentwicklung. Es gibt verschiedene Möglichkeiten eine gegebene, beliebig oft differenzierbare Funktion in eine Potenzreihe zu entwickeln. Die erste Möglichkeit ist, die Taylor-Reihe zu bestimmen und die Restglieder $R_{n+1}(f; x_0, x)$ abzuschätzen (vgl. die Taylor-Formel). Dies ist jedoch häufig mühselig. Die zweite Möglichkeit ist, bekannte Potenzreihen zu differenzieren oder zu integrieren.

BEISPIEL VII.3.3. Gemäß Abschnitt VII.2.6 (S. 264) ist für $|t| < 1$

$$\frac{1}{\sqrt{1+t}} = \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} t^k.$$

Die Transformation $t = -x^2$ liefert für alle $|x| < 1$

$$\frac{1}{\sqrt{1-x^2}} = \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} x^{2k}.$$

Gliedweise Integration dieser Reihe liefert

$$\arcsin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \binom{-\frac{1}{2}}{k} x^{2k+1}.$$

Eine weitere Möglichkeit ist, die Funktion als Summe oder Produkt bekannter Potenzreihen darzustellen.

BEISPIEL VII.3.4. Aus der bekannten Potenzreihenentwicklung der Exponentialfunktion ergibt sich

$$\begin{aligned}\cosh(x) &= \frac{1}{2}(e^x + e^{-x}) \\ &= \frac{1}{2} \sum_{k=0}^{\infty} \frac{1}{k!} x^k + \frac{1}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} x^k \\ &= \sum_{\ell=0}^{\infty} \frac{1}{(2\ell)!} x^{2\ell}.\end{aligned}$$

Analog folgt

$$\begin{aligned}\sinh(x) &= \frac{1}{2}(e^x - e^{-x}) \\ &= \sum_{\ell=0}^{\infty} \frac{1}{(2\ell+1)!} x^{2\ell+1}.\end{aligned}$$

BEISPIEL VII.3.5. Aus den bekannten Potenzreihen für $\cos(x)$ und $\frac{1}{1-x}$ und dem Cauchy-Produkt aus Abschnitt VII.1.4 (S. 259) folgt für $|x| < 1$

$$\begin{aligned}\frac{\cos(x)}{1-x} &= \left\{ \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \right\} \left\{ \sum_{k=0}^{\infty} x^k \right\} \\ &= 1 + x + \left(1 - \frac{1}{2!}\right)x^2 + \left(1 - \frac{1}{2!}\right)x^3 + \left(1 - \frac{1}{2!} + \frac{1}{4!}\right)x^4 + \dots \\ &= 1 + x + \frac{1}{2}x^2 + \frac{1}{2}x^3 + \frac{3}{4}x^4 + \dots\end{aligned}$$

VII.4. Anwendungen

VII.4.1. Grenzwertbestimmung. Häufig können Grenzwerte der Form $\frac{0}{0}$ mit Hilfe bekannter Potenzreihenentwicklungen leichter bestimmt werden als mit der Regel von de l'Hôpital aus Abschnitt IV.2.4 (S. 146). Zudem gibt dieser Zugang einen besseren Einblick in das Konvergenzverhalten.

BEISPIEL VII.4.1. Definiere für $x \neq 0$

$$f(x) = \frac{e^{-x^2} - 1}{x}.$$

Wie verhält sich diese Funktion für $x \rightarrow 0$?

Aus der Potenzreihenentwicklung der Exponentialfunktion erhalten wir mit der Substitution $t = -x^2$:

$$f(x) = \frac{1}{x} \left[\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} x^{2k} - 1 \right]$$

$$\begin{aligned}
&= \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} x^{2k-1} \\
&= -x + \frac{1}{2} x^3 - \frac{1}{6} x^5 + \dots
\end{aligned}$$

Also ist

$$\lim_{x \rightarrow 0} f(x) = 0.$$

Mit $f(0) = 0$ erhalten wir dann eine auf ganz \mathbb{R} definierte Funktion, die beliebig oft differenzierbar ist. Insbesondere können wir in der Reihenentwicklung die Werte der Ableitungen im Nullpunkt ablesen:

$$f'(0) = -1, \quad f''(0) = 0, \quad f^{(3)}(0) = 3, \quad \dots$$

VII.4.2. Näherungsformeln. Mit den Taylor-Formeln erhält man oft leicht berechenbare Näherungsformeln für komplizierte Funktionen.

BEISPIEL VII.4.2. Für den Fall mit Luftwiderstand gilt das Weg-Zeit-Gesetz

$$s(t) = \frac{v_0^2}{g} \ln\left(\cosh\left(\frac{gt}{v_0}\right)\right)$$

mit der Grenzggeschwindigkeit

$$v_0 = \lim_{t \rightarrow \infty} v(t).$$

Für die ersten vier Ableitungen in $t_0 = 0$ erhalten wir

$$\begin{aligned}
\dot{s}(t) &= v_0 \frac{\sinh\left(\frac{gt}{v_0}\right)}{\cosh\left(\frac{gt}{v_0}\right)} \\
\Rightarrow \dot{s}(0) &= 0 \\
\ddot{s}(t) &= g \frac{1}{\cosh^2\left(\frac{gt}{v_0}\right)} \\
\Rightarrow \ddot{s}(0) &= g \\
s^{(3)}(t) &= -2 \frac{g^2}{v_0} \frac{\sinh\left(\frac{gt}{v_0}\right)}{\cosh^3\left(\frac{gt}{v_0}\right)} \\
\Rightarrow s^{(3)}(0) &= 0 \\
s^{(4)}(t) &= -2 \frac{g^3}{v_0^2} \left[\frac{1}{\cosh^2\left(\frac{gt}{v_0}\right)} - 3 \frac{\sinh^2\left(\frac{gt}{v_0}\right)}{\cosh^4\left(\frac{gt}{v_0}\right)} \right] \\
\Rightarrow s^{(4)}(0) &= -2 \frac{g^3}{v_0^2}.
\end{aligned}$$

Also gilt für kleine t die Näherungsformel

$$\begin{aligned} s(t) &\approx T_4(s; 0)(t) \\ &= \frac{1}{2}gt^2 - \frac{1}{12} \frac{g^3}{v_0^2} t^4. \end{aligned}$$

VII.4.3. Integration. Durch gliedweise Integration von Potenzreihendarstellungen kann man häufig Integrale berechnen, die nicht geschlossen dargestellt werden können.

BEISPIEL VII.4.3. Die Gaußsche Fehlerfunktion ist definiert durch

$$\Phi(x) = \int_0^x e^{-t^2} dt.$$

Dieses Integral ist nicht geschlossen darstellbar. Durch Integration der Potenzreihe von e^{-t^2} erhält man die Reihendarstellung

$$\begin{aligned} \Phi(x) &= \int_0^x \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} t^{2k} dt \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(2k+1)} x^{2k+1}. \end{aligned}$$

Diese Darstellung erlaubt die Berechnung der Funktionswerte bis zu jeder gewünschten Genauigkeit.

VII.4.4. Lösen von Differentialgleichungen. Oft kann man Differentialgleichungen durch einen Potenzreihenansatz lösen.

BEISPIEL VII.4.4. Wir betrachten die Auslenkung $s(t)$ aus der Ruhelage eines Pendels, dessen Rückstellkraft proportional zum Quadrat der Zeit ist. Zur Zeit $t = 0$ soll sich das Pendel in der Ruhelage befinden und die Geschwindigkeit 1 haben. Wir müssen somit das ANFANGSWERTPROBLEM

$$\begin{aligned} \ddot{s}(t) + t^2 s(t) &= 0 \\ s(0) &= 0 \\ \dot{s}(0) &= 1 \end{aligned}$$

lösen. Hierzu machen wir den Potenzreihenansatz

$$s(t) = \sum_{k \in \mathbb{N}^*} a_k t^k$$

mit unbekanntem Koeffizienten a_k (dabei ist $s(0) = 0$ mit $a_0 = 0$ schon ausgenutzt!). Falls diese Reihe einen positiven Konvergenzradius hat, können wir sie in die Differentialgleichung einsetzen, gliedweise differenzieren und erhalten

$$0 = \ddot{s}(t) + t^2 s(t)$$

$$\begin{aligned}
&= \frac{d^2}{dt^2} \left(\sum_{k=1}^{\infty} a_k t^k \right) + t^2 \sum_{k=1}^{\infty} a_k t^k \\
&= \sum_{k=2}^{\infty} k(k-1) a_k t^{k-2} + \sum_{k=1}^{\infty} a_k t^{k+2} \\
&= 2a_2 + 6a_3 t + 12a_4 t^2 + \underbrace{\sum_{k=5}^{\infty} k(k-1) a_k t^{k-2}}_{\ell=k-2} + \underbrace{\sum_{k=1}^{\infty} a_k t^{k+2}}_{\ell=k+2} \\
&= 2a_2 + 6a_3 t + 12a_4 t^2 + \sum_{\ell=3}^{\infty} t^{\ell} [(\ell+2)(\ell+1)a_{\ell+2} + a_{\ell-2}].
\end{aligned}$$

Die Bedingung $\dot{s}(0) = 1$ und Koeffizientenvergleich ergeben

$$\begin{aligned}
a_1 &= 1, \\
a_2 &= 0, \\
a_3 &= 0, \\
a_4 &= 0, \\
a_{\ell+2} &= -\frac{1}{(\ell+1)(\ell+2)} a_{\ell-2} \quad \ell \geq 3.
\end{aligned}$$

Hieraus folgt für alle $k \in \mathbb{N}$

$$\begin{aligned}
a_{2k} &= 0, \\
a_{4k+3} &= 0, \\
a_{4k+1} &= (-1)^k \prod_{\ell=1}^k \frac{1}{4\ell(4\ell+1)}.
\end{aligned}$$

Also ist

$$\begin{aligned}
s(t) &= \sum_{k=0}^{\infty} (-1)^k \left[\prod_{\ell=1}^k \frac{1}{\ell(4\ell+1)} \right] t^{4k+1} \\
&= t - \frac{1}{20} t^5 + \frac{1}{1440} t^9 - \frac{1}{224640} t^{13} + \dots
\end{aligned}$$

Dies ist zunächst einmal eine rein formale Lösung. Um nachzuweisen, dass die Lösung des Anfangswertproblems tatsächlich durch diese Potenzreihe dargestellt wird, müssen wir zum Abschluss noch nachprüfen, ob sie einen positiven Konvergenzradius hat. Mit dem zweiten „Trick“ aus Abschnitt VII.2.3 (S. 261) erhalten wir für den Konvergenzradius R

$$R = \sqrt[4]{\lim_{k \rightarrow \infty} \left| \frac{a_{4k+1}}{a_{4k+5}} \right|}$$

$$\begin{aligned} &= \sqrt[4]{\lim_{k \rightarrow \infty} (4k + 5)(4k + 4)} \\ &= \infty. \end{aligned}$$

Also stellt die berechnete Potenzreihe auf ganz \mathbb{R} die Lösung des Anfangswertproblems dar.

KAPITEL VIII

Differentiation von Funktionen in mehreren Variablen

In der mathematischen Behandlung von Naturvorgängen treten häufig Funktionen auf, die von mehreren Variablen x_1, \dots, x_n abhängen (etwa von den Ortskoordinaten x, y, z , der Zeit t , dem Druck p, \dots) und deren Wertebereich mehrdimensional ist. Hierfür ist es hilfreich, die Variablen x_1, \dots, x_n zu einem Spaltenvektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

zusammenzufassen. Im Folgenden untersuchen wir daher Funktionen $\mathbf{f} : D \rightarrow \mathbb{R}^m$ mit $D \subset \mathbb{R}^n$, die jedem Spaltenvektor $\mathbf{x} = (x_1, \dots, x_n)^T \in D$ einen mit $\mathbf{f}(\mathbf{x})$ oder $\mathbf{f}(x_1, \dots, x_n)$ bezeichneten Spaltenvektor im \mathbb{R}^m zuordnen. Hierfür schreiben wir $\mathbf{f} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$ bzw. $D \ni \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$. \mathbb{R}^n heißt URBILDRAUM, D DEFINITIONSBEREICH, \mathbb{R}^m BILDRAUM und $\mathbf{f}(D) = \{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in D\}$ das BILD der Funktion \mathbf{f} . Zuerst behandeln wir die Sonderfälle $n = 1$, d.h. $\mathbf{f} : \mathbb{R} \supset I \rightarrow \mathbb{R}^m$, und $m = 1$, d.h. $\mathbf{f} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$. Der allgemeine Fall ($m > 1, n > 1$) lässt sich hierauf zurückführen.

VIII.1. Kurven im \mathbb{R}^n

VIII.1.1. Parameterdarstellungen. In der Praxis hat man häufig das Problem, ein geometrisches Gebilde „Kurve“ oder „Weg“ analytisch, d.h. durch Funktionen darzustellen. In der Ebene sind bereits einige Darstellungsarten bekannt (s. Abschnitt V.5 (S. 197)):

- die evtl. nur abschnittsweise explizite Darstellung $y = f(x)$ als Graph einer Funktion,
- die implizite Darstellung $F(x, y) = 0$,
- die Parameterdarstellung.

Für die allgemeine Behandlung eignet sich am besten die Parameterdarstellung mit der Vorstellung, dass die Kurve die Bahn eines bewegten Punktes darstellt, der zur Zeit t den Ortsvektor $\mathbf{x}(t)$ besitzt.

Wir betrachten daher vektorwertige, auf einem Intervall $I \subset \mathbb{R}$ erklärte Funktionen $\mathbf{x} : I \rightarrow \mathbb{R}^n$. Jede derartige Funktion besteht aus

n Komponentenfunktionen $x_i : I \rightarrow \mathbb{R}$, $1 \leq i \leq n$, d.h.

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, t \in I.$$

Der Grenzwertbegriff wird komponentenweise erklärt:

$$\lim_{t \rightarrow t_0} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \iff \lim_{t \rightarrow t_0} x_i(t) = c_i \quad (1 \leq i \leq n).$$

Dementsprechend heißt $\mathbf{x} : I \rightarrow \mathbb{R}^n$ in $t_0 \in I$ bzw. auf I stetig bzw. differenzierbar, wenn *alle* Komponentenfunktionen in $t_0 \in I$ bzw. auf I stetig bzw. differenzierbar sind. Die Ableitung wird komponentenweise berechnet:

$$\dot{\mathbf{x}}(t) = \frac{d}{dt} \mathbf{x}(t) = \lim_{h \rightarrow 0} \frac{1}{h} [\mathbf{x}(t+h) - \mathbf{x}(t)] = \begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}$$

mit $\dot{x}_i(t) = \frac{d}{dt} x_i(t)$. Man nennt $\dot{\mathbf{x}}(t)$ den TANGENTIALVEKTOR von \mathbf{x} an der Stelle t (vgl. Abschnitt V.5.2 (S. 199) für den Fall $n = 2$).

Mit $\mathbf{x}, \mathbf{y} : I \rightarrow \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$ und $g : I \rightarrow \mathbb{R}$ gelten folgende Rechenregeln:

$$\begin{aligned} \frac{d}{dt} [\alpha \mathbf{x}(t) + \beta \mathbf{y}(t)] &= \alpha \dot{\mathbf{x}}(t) + \beta \dot{\mathbf{y}}(t) \\ \frac{d}{dt} [\mathbf{x}(t) \cdot \mathbf{y}(t)] &= \dot{\mathbf{x}}(t) \cdot \mathbf{y}(t) + \mathbf{x}(t) \cdot \dot{\mathbf{y}}(t) \\ \frac{d}{dt} [\mathbf{x}(t) \times \mathbf{y}(t)] &= \dot{\mathbf{x}}(t) \times \mathbf{y}(t) + \mathbf{x}(t) \times \dot{\mathbf{y}}(t) \quad \text{falls } n = 3 \\ \frac{d}{dt} [g(t) \mathbf{x}(t)] &= \dot{g}(t) \mathbf{x}(t) + g(t) \dot{\mathbf{x}}(t). \end{aligned}$$

DEFINITION VIII.1.1. Seien $G \subset \mathbb{R}^n$ und $[a, b] \subset \mathbb{R}$. Wir nennen jede stetig differenzierbare Funktion $\mathbf{x} : [a, b] \rightarrow G$ ein KURVENSTÜCK in G mit ANFANGSPUNKT $\mathbf{x}(a)$, ENDPUNKT $\mathbf{x}(b)$ und SPUR $\{\mathbf{x}(t) : a \leq t \leq b\}$. $\dot{\mathbf{x}}(a)$ und $\dot{\mathbf{x}}(b)$ sind als einseitige Ableitungen zu verstehen. Ein Kurvenstück heißt REGULÄR, wenn $\dot{\mathbf{x}}(t) \neq \mathbf{0}$ ist für alle $t \in [a, b]$.

BEMERKUNG VIII.1.2. (1) Unter einer KURVE versteht man eine Kette von aneinanderhängenden Kurvenstücken (vgl. Definition IX.2.2

(S. 330)). Wir beschränken uns im Augenblick aber auf Kurvenstücke.
 (2) Statt „Spur eines Kurvenstückes“ sagt man häufig auch kurz „Kurvenstück“.

Für ein reguläres Kurvenstück $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n$ existiert für jedes $t \in [a, b]$ die Zahl

$$s(t) = \int_a^t |\dot{\mathbf{x}}(\tau)| d\tau = \int_a^t \sqrt{\dot{x}_1(\tau)^2 + \dots + \dot{x}_n(\tau)^2} d\tau.$$

Sie heißt **BOGENLÄNGE** des Kurvenstückes. $L(\mathbf{x}) = s(b)$ heißt die **LAENGE** des Kurvenstückes. Wegen des Hauptsatzes der Differential- und Integralrechnung aus Abschnitt V.1.3 (S. 171) ist

$$\dot{s}(t) = |\dot{\mathbf{x}}(t)|.$$

Man nennt

$$\begin{aligned} ds &= |\dot{\mathbf{x}}(t)| dt && \text{das (SKALARE) BOGENELEMENT} \\ d\mathbf{x} &= \dot{\mathbf{x}}(t) dt && \text{das VEKTORIELLE BOGENELEMENT.} \end{aligned}$$

ds ist in erster Näherung die im Zeitintervall $[t, t + dt]$ zurückgelegte Weglänge.

VIII.1.2. Das begleitende Dreibein, Krümmung und Torsion. Wir betrachten ein reguläres Kurvenstück $\mathbf{x} : I \rightarrow \mathbb{R}^3$ im Raum, d.h. $n = 3$. Die Bewegung eines Massenpunktes entlang eines solchen Kurvenstückes und die Auswirkung von Kräften und Momenten in den einzelnen Kurvenpunkten wird am besten in einem der Geometrie angepassten begleitenden kartesischen Koordinatensystem $(\mathbf{x}(t) : \mathbf{T}, \mathbf{N}, \mathbf{B})$ beschrieben, in dem sich auch die Basisvektoren $\mathbf{T}, \mathbf{N}, \mathbf{B}$ mit dem Parameter t verändern. Dieses Koordinatensystem heißt **BEGLEITENDES DREIBEIN** und wird wie folgt definiert. Es ist

$$\mathbf{T}(t) = \frac{1}{|\dot{\mathbf{x}}(t)|} \dot{\mathbf{x}}(t)$$

$\mathbf{T}(t)$ heißt **TANGENTENVEKTOR** des Kurvenstücks an der Stelle t . Weiter ist

$$\mathbf{N}(t) = \frac{1}{|\dot{\mathbf{T}}(t)|} \dot{\mathbf{T}}(t)$$

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t)$$

$\mathbf{N}(t)$ heißt HAUPTNORMALENVEKTOR, $\mathbf{B}(t)$ heißt BINORMALENVEKTOR.

Die Parameterdarstellung der Kurventangente im Punkt $\mathbf{x}(t_0)$ lautet

$$\mathbf{y}(\lambda) = \mathbf{x}(t_0) + \lambda \mathbf{T}(t_0), \quad \lambda \in \mathbb{R}.$$

Die von $\mathbf{N}(t_0)$ und $\mathbf{T}(t_0)$ im Punkt $\mathbf{x}(t_0)$ aufgespannte Ebene heißt die SCHMIEGEEBENE der Kurve im Punkt t_0 . Ihre Parameterdarstellung lautet

$$\mathbf{y}(\lambda, \mu) = \mathbf{x}(t_0) + \lambda \mathbf{T}(t_0) + \mu \mathbf{N}(t_0), \quad \lambda, \mu \in \mathbb{R}.$$

Die Änderungsrate

$$\frac{1}{\Delta s} \Delta \mathbf{T} = \frac{1}{s(t_1) - s(t)} [\mathbf{T}(t_1) - \mathbf{T}(t)]$$

des Tangentenvektors bezogen auf die Bogenlänge beschreibt anschaulich das mittlere Krümmungsverhalten der Kurve im Parameterintervall $[t, t_1]$ bzw. $[t_1, t]$. Mit der Regel von le l'Hôpital ergibt sich für den Grenzwert $t_1 \rightarrow t$:

$$\lim_{t_1 \rightarrow t} \frac{1}{\Delta s} \Delta \mathbf{T} = \frac{1}{\dot{s}(t)} \dot{\mathbf{T}}(t).$$

Wegen $\dot{s}(t) = |\dot{\mathbf{x}}(t)|$ nennt man daher

$$\frac{1}{|\dot{\mathbf{x}}(t)|} \dot{\mathbf{T}}(t) \quad \text{den KRÜMMUNGSVEKTOR und}$$

$$\kappa(t) = \frac{|\dot{\mathbf{T}}(t)|}{|\dot{\mathbf{x}}(t)|} \quad \text{die KRÜMMUNG}$$

der Kurve an der Stelle t .

Für die Darstellung von $\dot{\mathbf{x}}$ und $\ddot{\mathbf{x}}$ ergibt sich im begleitenden Dreibein:

$$\dot{\mathbf{x}}(t) = \dot{s}(t) \mathbf{T}(t)$$

$$\ddot{\mathbf{x}}(t) = \ddot{s}(t) \mathbf{T}(t) + \dot{s}(t)^2 \kappa(t) \mathbf{N}(t).$$

Diese Formel hat folgende KINEMATISCHE DEUTUNG:

Stellt \mathbf{x} die Bewegung eines Massenpunktes dar, so ist dessen Geschwindigkeitsvektor $\dot{\mathbf{x}} = \dot{s} \mathbf{T}$ tangential zur Bahn in Durchlaufrichtung und dem Betrag nach gleich der Momentangeschwindigkeit \dot{s} . Der Beschleunigungsvektor $\ddot{\mathbf{x}}$ hat $\ddot{s} \mathbf{T}$ als Tangential- und $\dot{s}^2 \kappa \mathbf{N}$ als Normalkomponente. Daher kann $\frac{1}{\kappa}$ als Radius eines Krümmungskreises in der Schmiegeebene gedeutet werden.

Wegen $\mathbf{T} \times \mathbf{T} = 0$ und $\mathbf{T} \times \mathbf{N} = \mathbf{B}$ ergibt sich aus obiger Formel

$$\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t) = \dot{s}(t)^3 \kappa(t) \mathbf{B}(t).$$

Wegen $|\mathbf{B}| = 1$ folgt

$$\begin{aligned} \kappa(t) &= \frac{|\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t)|}{|\dot{\mathbf{x}}(t)|^3}. \\ \mathbf{B}(t) &= \frac{1}{|\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t)|} \dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t). \end{aligned}$$

Wegen dieser Formel bestimmt man das begleitende Dreibein $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ am besten in der Reihenfolge

$$\begin{aligned} \mathbf{T} &= \frac{1}{|\dot{\mathbf{x}}|} \dot{\mathbf{x}}, \\ \mathbf{B} &= \frac{1}{|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}|} \dot{\mathbf{x}} \times \ddot{\mathbf{x}}, \\ \mathbf{N} &= \mathbf{B} \times \mathbf{T}. \end{aligned}$$

Für eine ebene Kurve, d.h. $x_3(t) = 0$ für alle t , liefern die hergeleiteten Formeln bis aufs Vorzeichen die Formeln für die Krümmung aus Abschnitt V.5.4 (S. 202).

Die Änderungsrate des Binormalenvektors bezogen auf die Bogenlänge beschreibt das Herauswinden der Kurve aus der Schmiegeebene. Daher nennt man

$$\frac{1}{|\dot{\mathbf{x}}(t)|} \dot{\mathbf{B}}(t) \quad \text{den TORSIONSVEKTOR.}$$

Wegen $\mathbf{B}(t) \cdot \mathbf{B}(t) = |\mathbf{B}(t)|^2 = 1$ ist $\mathbf{B}(t) \cdot \dot{\mathbf{B}}(t) = 0$, d.h. $\dot{\mathbf{B}}$ ist orthogonal zu \mathbf{B} . Wegen

$$\begin{aligned} \dot{\mathbf{B}} &= \frac{d}{dt}(\mathbf{T} \times \mathbf{N}) \\ &= \dot{\mathbf{T}} \times \mathbf{N} + \mathbf{T} \times \dot{\mathbf{N}} \\ &= \mathbf{T} \times \dot{\mathbf{N}} \end{aligned}$$

ist $\dot{\mathbf{B}}$ auch orthogonal zu \mathbf{T} , also parallel zu \mathbf{N} . Daher gibt es eine skalare Funktion $\tau(t)$ mit

$$\frac{1}{|\dot{\mathbf{x}}(t)|} \dot{\mathbf{B}}(t) = -\tau(t) \mathbf{N}(t).$$

Diese heißt TORSION der Kurve. Es ist

$$\tau(t) = \frac{\det(\dot{\mathbf{x}}(t), \ddot{\mathbf{x}}(t), \ddot{\mathbf{x}}(t))}{|\dot{\mathbf{x}}(t) \times \ddot{\mathbf{x}}(t)|^2}.$$

BEISPIEL VIII.1.3. Die neutrale Faser einer Schraubenfeder wird dargestellt durch

$$\mathbf{x}(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ ht \end{pmatrix}, \quad 0 \leq t \leq 2\pi n.$$

Dabei ist r der Radius, $2\pi h$ die Ganghöhe und n die Windungszahl. Es ist

$$\begin{aligned} \dot{\mathbf{x}} &= \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix}, \\ \ddot{\mathbf{x}}(t) &= \begin{pmatrix} -r \cos t \\ -r \sin t \\ 0 \end{pmatrix}, \\ \ddot{\mathbf{x}}(t) &= \begin{pmatrix} r \sin t \\ -r \cos t \\ 0 \end{pmatrix}. \end{aligned}$$

Mit der Abkürzung

$$R = |\dot{\mathbf{x}}(t)| = \sqrt{r^2 + h^2}$$

folgt hieraus für das Begleitende Dreibein

$$\begin{aligned} \mathbf{T}(t) &= \frac{1}{R} \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix}, \\ \mathbf{N}(t) &= \begin{pmatrix} -\cos t \\ -\sin t \\ 0 \end{pmatrix}, \\ \mathbf{B}(t) &= \frac{1}{R} \begin{pmatrix} h \sin t \\ -h \cos t \\ r \end{pmatrix} \end{aligned}$$

und für Krümmung und Torsion

$$\begin{aligned} \kappa(t) &= \frac{r}{r^2 + h^2}, \\ \tau(t) &= \frac{h}{r^2 + h^2}. \end{aligned}$$

Eine Belastung $\mathbf{F} = -F\mathbf{e}_3$ mit der Schraubenachse als Wirkungslinie erzeugt in jedem Punkt $\mathbf{x}(t)$ der Schraubenfeder das Moment

$$\mathbf{m}(t) = -\mathbf{x} \times \mathbf{F} = rF \begin{pmatrix} \sin t \\ -\cos t \\ 0 \end{pmatrix}.$$

Es besitzt im begleitenden Dreibein $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ die Zerlegung

$$\begin{aligned} \mathbf{m} &= (\mathbf{m} \cdot \mathbf{T})\mathbf{T} + (\mathbf{m} \cdot \mathbf{N})\mathbf{N} + (\mathbf{m} \cdot \mathbf{B})\mathbf{B} \\ &= -\frac{Fr^2}{R}\mathbf{T} + \frac{Frh}{R}\mathbf{B}. \end{aligned}$$

VIII.2. Reellwertige Funktionen mehrerer reeller Veränderlicher

VIII.2.1. Einige topologische Grundbegriffe. Der Betrag eines Vektors $\mathbf{x} \in \mathbb{R}^n$ ist gemäß Abschnitt II.6.7 (S. 108) definiert durch

$$|\mathbf{x}| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Dementsprechend ist $|\mathbf{x} - \mathbf{y}|$ der Abstand zweier Punkte $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Zu $\mathbf{a} \in \mathbb{R}^n$ und $r > 0$ heißt die Menge

$$B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| < r\}$$

der (OFFENE) BALL UM \mathbf{a} MIT RADIUS r .

BEISPIEL VIII.2.1. Für $n = 1$ ist $B_r(a)$ das offene Intervall $(a - r, a + r)$. Für $n = 2$ ist $B_r(\mathbf{a})$ das Innere des Kreises mit Mittelpunkt \mathbf{a} und Radius r . Für $n = 3$ ist $B_r(\mathbf{a})$ das Innere der Kugel mit Mittelpunkt \mathbf{a} und Radius r .

DEFINITION VIII.2.2. Sei $D \subset \mathbb{R}^n$ eine nichtleere Menge.

- (1) Ein Punkt $\mathbf{a} \in D$ heißt INNERER PUNKT von D , wenn es ein $r > 0$ und einen Ball $B_r(\mathbf{a})$ um \mathbf{a} mit Radius r gibt, der ganz in D liegt, d.h. $B_r(\mathbf{a}) \subset D$.
- (2) Die Menge D heißt OFFEN, wenn jeder Punkt von D ein innerer Punkt ist.
- (3) Ein Punkt $\mathbf{b} \in \mathbb{R}^n$ heißt RANDPUNKT von D , wenn *jeder* Ball $B_r(\mathbf{b})$ um \mathbf{b} mindestens einen Punkt von D und mindestens einen nicht zu D gehörenden Punkt enthält. Die Menge aller Randpunkte von D heißt RAND von D und wird mit ∂D bezeichnet.
- (4) Eine Menge D heißt ABGESCHLOSSEN, wenn sie alle ihre Randpunkte enthält.

BEISPIEL VIII.2.3. (1) Die Menge $K = \{(x, y) : x^2 + y^2 < 1\}$ ist offen mit Rand $\partial K = \{(x, y) : x^2 + y^2 = 1\}$.

(2) Die Menge $R = \{(x, y) : |x| \leq 1, |y| \leq 1\}$ ist abgeschlossen. Die Menge ihrer inneren Punkte ist $\{(x, y) : |x| < 1, |y| < 1\}$, ihr Rand ist

$\partial R = \{(x, y) : |x| = 1 \text{ oder } |y| = 1\}$.

(3) Die Menge $S = \{(x, y) : x^2 + y^2 < 1, x \geq 0, y \geq 0\}$ ist weder offen, noch abgeschlossen. Ihr Rand ist $\partial S = \{(x, y) : x^2 + y^2 = 1, x \geq 0, y \geq 0\} \cup \{(x, 0) : 0 \leq x \leq 1\} \cup \{(0, y) : 0 \leq y \leq 1\}$, die Menge ihrer inneren Punkte ist $\{(x, y) : x^2 + y^2 < 1, x > 0, y > 0\}$.

DEFINITION VIII.2.4. (1) Eine Menge $D \subset \mathbb{R}^n$ heißt **BESCHRÄNKT**, wenn es ein $K > 0$ gibt mit $D \subset B_K(\mathbf{0})$, d.h. wenn für alle $\mathbf{x} \in D$ gilt $|\mathbf{x}| < K$.

(2) Eine Menge die beschränkt und abgeschlossen ist, nennen wir **KOMPAKT**.

BEISPIEL VIII.2.5. (1) Die Menge R aus Beispiel VIII.2.3 (2) ist kompakt.

(2) Die Mengen K und S aus Beispiel VIII.2.3 (1) und (3) sind beschränkt, aber nicht kompakt.

(3) Die Menge $H = \{(x, y) : x \geq 0\}$ ist abgeschlossen, aber nicht beschränkt.

VIII.2.2. Reelle Funktionen mehrerer reeller Veränderlicher. Unter einer reellen Funktion von n reellen Veränderlichen versteht man eine auf einer nichtleeren Teilmenge D des \mathbb{R}^n definierte Abbildung $f : D \rightarrow \mathbb{R}$ mit Werten in \mathbb{R} . Sie ordnet jedem Punkt $\mathbf{x} = (x_1, \dots, x_n)^T \in D$ des Definitionsbereiches eine reelle Zahl $z = f(\mathbf{x}) = f(x_1, \dots, x_n)$ zu. Für Funktionen in 2 bzw. 3 reellen Veränderlichen schreibt man häufig auch $f(x, y)$ bzw. $f(x, y, z)$.

Zur Untersuchung und Darstellung einer Funktion $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ ist es häufig vorteilhaft, sie auf speziellen Mengen zu betrachten:

- den NIVEAUMENGEN $N_c = \{\mathbf{x} \in D : f(\mathbf{x}) = c\}$ mit $c \in \mathbb{R}$,
- den zu den Koordinatenachsen parallelen Geraden durch D ; man erhält dann die Funktionen einer reellen Veränderlichen $x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ mit $(a_1, \dots, a_n) \in D$ konstant.

Im Fall $n = 2$ bzw. $n = 3$ bezeichnen die Niveaumengen Kurven in der Ebene bzw. Flächen im Raum.

BEISPIEL VIII.2.6. (1) Die Gerade $2x + 3y = 1$ ist die Niveaumenge der Funktion $f(x, y) = 2x + 3y$ zum Wert $c = 1$.

(2) Der Kreis $x^2 + y^2 = r^2$ ist die Niveaumenge der Funktion $f(x, y) = x^2 + y^2$ zum Wert $c = r^2$.

(3) Die Ebene $4x - 2y + 3z = 4$ ist die Niveaumenge der Funktion $f(x, y, z) = 4x - 2y + 3z$ zum Wert $c = 4$.

(4) Die Niveauflächen eines quadratischen Polynoms in drei Veränderlichen sind Quadriken $\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = 0$ im \mathbb{R}^3 .

Wie im Fall $n = 1$ kann man jeder Funktion $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ ihren **GRAPHEN**

$$\Gamma_f = \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in D\} \subset \mathbb{R}^{n+1}$$

zuordnen. Er ist die Niveaumenge zum Wert $c = 0$ der Funktion $F(x_1, \dots, x_{n+1}) = x_{n+1} - f(x_1, \dots, x_n)$.

VIII.2.3. Grenzwerte und Stetigkeit. Wir erinnern an die Definition der Stetigkeit einer Funktion einer reellen Veränderlichen aus Abschnitt III.3.1 (S. 123).

DEFINITION VIII.2.7. Sei $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ und $\mathbf{a} \in D \cup \partial D$.

- (1) f hat in \mathbf{a} den GRENZWERT $c \in \mathbb{R}$, kurz $f(\mathbf{x}) \rightarrow c$ für $\mathbf{x} \rightarrow \mathbf{a}$ oder $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = c$, wenn es zu jedem noch so kleinen $\varepsilon > 0$ einen Ball $B_\delta(\mathbf{a})$ um \mathbf{a} mit Radius δ gibt, sodass $|f(\mathbf{x}) - c| < \varepsilon$ ist für alle $\mathbf{x} \in D \cap B_\delta(\mathbf{a})$.
- (2) f heißt STETIG in $\mathbf{a} \in D$, wenn $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$ ist.
- (3) f heißt STETIG auf D , wenn f in jedem Punkt $\mathbf{a} \in D$ stetig ist.

Für stetige Funktionen mehrerer reeller Veränderlicher gelten die gleichen Rechenregeln wie in Abschnitt III.3.2 (S. 123).

BEISPIEL VIII.2.8. Die Projektionen $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $p_i(\mathbf{x}) = x_i$ sind stetig. Aufgrund der Rechenregeln für stetige Funktionen ist jedes Polynom

$$p(\mathbf{x}) = \sum_{k_i \leq m} a_{k_1 \dots k_n} x_1^{k_1} \cdot \dots \cdot x_n^{k_n}$$

in n reellen Veränderlichen stetig. Ebenso ist jede rationale Funktion

$$f(\mathbf{x}) = \frac{p(\mathbf{x})}{q(\mathbf{x})}$$

in allen Punkten \mathbf{x}_0 stetig, in denen $q(\mathbf{x}_0) \neq 0$ ist.

BEISPIEL VIII.2.9. Betrachte $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = \begin{cases} \frac{2xy^2}{x^2+y^4} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0). \end{cases}$$

Es ist

$$\begin{aligned} \lim_{x \rightarrow 0} f(x, 0) &= 0 \\ &= f(0, 0) \end{aligned}$$

$$\begin{aligned} \lim_{y \rightarrow 0} f(0, y) &= 0 \\ &= f(0, 0) \end{aligned}$$

aber

$$\lim_{t \rightarrow 0} f(t^2, t) = \lim_{t \rightarrow 0} \frac{2t^2 t^2}{t^4 + t^4}$$

$$= 1 \\ \neq f(0, 0).$$

Daher ist f nicht stetig in $\mathbf{0}$.

Analog zu Abschnitt III.3.3 (S. 124) gilt

Jede auf einer *kompakten* Menge $D \subset \mathbb{R}^n$ stetige Funktion $f : D \rightarrow \mathbb{R}$ nimmt dort ihr Minimum und Maximum an, d.h. es gibt ein $\mathbf{x}_{\min} \in D$ und ein $\mathbf{x}_{\max} \in D$ mit

$$f(\mathbf{x}_{\min}) \leq f(\mathbf{x}) \leq f(\mathbf{x}_{\max})$$

für alle $\mathbf{x} \in D$.

BEISPIEL VIII.2.10. Die Funktion $f(x, y) = e^{xy}$ ist auf \mathbb{R}^2 stetig. Wegen

$$\lim_{t \rightarrow -\infty} e^t = 0 \quad \text{und} \quad \lim_{t \rightarrow \infty} e^t = \infty$$

nimmt sie auf \mathbb{R}^2 weder Minimum noch Maximum an. Dagegen besitzt sie auf der kompakten Menge R aus Beispiel VIII.2.3 (2) ein Minimum und ein Maximum. Mit Hilfe der Methoden der Abschnitt VIII.3.6 (S. 305) und VIII.3.7 (S. 307) erhalten wir

$$\max_{\mathbf{x} \in R} f(x) = e, \quad \min_{\mathbf{x} \in R} f(\mathbf{x}) = e^{-1}.$$

VIII.2.4. Partielle Ableitungen, der Gradient. Sei $D \subset \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ und $\mathbf{a} \in D$. Existiert die Ableitung der Funktion $x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ an der Stelle $x_i = a_i$, so nennt man sie die PARTIELLE ABLEITUNG VON f NACH x_i IM PUNKT \mathbf{a} und bezeichnet sie mit $\frac{\partial f}{\partial x_i}(\mathbf{a})$ oder $f_{x_i}(\mathbf{a})$:

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\mathbf{a}) &= f_{x_i}(\mathbf{a}) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) \\ &\quad - f(a_1, \dots, a_n)] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [f(\mathbf{a} + h\mathbf{e}_i) - f(\mathbf{a})]. \end{aligned}$$

BEISPIEL VIII.2.11. Für $f(x, y) = x^2y^3 + y \ln x$ auf $D = \{(x, y) \in \mathbb{R}^2 : x > 0\}$ folgt

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= f_x(x, y) = 2xy^3 + \frac{y}{x}, \\ \frac{\partial f}{\partial y}(x, y) &= f_y(x, y) = 3x^2y^2 + \ln x. \end{aligned}$$

Die Funktion f heißt PARTIELL DIFFERENZIERBAR, wenn alle partiellen Ableitungen existieren. Sie heißt STETIG PARTIELL DIFFERENZIERBAR, wenn alle partiellen Ableitungen existieren und stetig sind. Die Funktion aus Beispiel VIII.2.11 ist stetig partiell differenzierbar.

Fasst man die partiellen Ableitungen einer Funktion f zu einem Vektor zusammen, erhält man den GRADIENTEN von f an der Stelle \mathbf{a} :

$$\text{grad } f(\mathbf{a}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{a}) \end{pmatrix}.$$

BEISPIEL VIII.2.12. Für $f(x, y, z) = e^{x+2y} + x^2yz^3$ erhalten wir

$$\text{grad } f(x, y, z) = \begin{pmatrix} e^{x+2y} + 2xyz^3 \\ 2e^{x+2y} + x^2z^3 \\ 3x^2yz^2 \end{pmatrix}.$$

Jede partielle Ableitung $\frac{\partial f}{\partial x_i}$ ist eine reelle Funktion von n reellen Veränderlichen und kann ggf. erneut partiell abgeleitet werden. Dies definiert die partiellen Ableitungen höherer Ordnung:

$$\begin{aligned} f_{x_i x_i} &= \frac{\partial^2 f}{\partial x_i^2} \\ &= \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right), \\ f_{x_i x_j} &= \frac{\partial^2 f}{\partial x_i \partial x_j} \\ &= \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right), \text{ usw.} \end{aligned}$$

BEISPIEL VIII.2.13. (1) Für die Funktion aus Beispiel VIII.2.11 erhalten wir

$$\begin{aligned} f_{xx}(x, y) &= \frac{\partial^2 f}{\partial x^2}(x, y) \\ &= 2y^3 - \frac{y}{x^2}, \\ f_{yy}(x, y) &= \frac{\partial^2 f}{\partial y^2}(x, y) \\ &= 6x^2y, \\ f_{xy}(x, y) &= \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ &= 6xy^2 + \frac{1}{x}, \end{aligned}$$

$$\begin{aligned} f_{yx}(x, y) &= \frac{\partial^2 f}{\partial y \partial x}(x, y) \\ &= 6xy^2 + \frac{1}{x}. \end{aligned}$$

(2) Definiere $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ durch

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

Da für alle $x, y \in \mathbb{R}$ gilt $f(x, 0) = 0$ und $f(0, y) = 0$, ist $f_x(0, 0) = f_y(0, 0) = 0$. Für $(x, y) \neq (0, 0)$ erhalten wir

$$\begin{aligned} f_x(x, y) &= y \left(\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^2}{(x^2 + y^2)^2} \right), \\ f_y(x, y) &= x \left(\frac{x^2 - y^2}{x^2 + y^2} - \frac{4x^2 y^2}{(x^2 + y^2)^2} \right). \end{aligned}$$

Hieraus folgt

$$\begin{aligned} f_{xy}(0, 0) &= \lim_{y \rightarrow 0} \frac{f_x(0, y) - f_x(0, 0)}{y} \\ &= \lim_{y \rightarrow 0} \frac{-y}{y} \\ &= -1 \end{aligned}$$

und

$$\begin{aligned} f_{yx}(0, 0) &= \lim_{x \rightarrow 0} \frac{f_y(x, 0) - f_y(0, 0)}{x} \\ &= \lim_{x \rightarrow 0} \frac{x}{x} \\ &= 1. \end{aligned}$$

Also ist

$$f_{xy}(0, 0) \neq f_{yx}(0, 0).$$

Wie Beispiel VIII.2.13 (2) zeigt, kann es bei höheren partiellen Ableitungen einen Unterschied machen, in welcher Reihenfolge man die Variablen „abarbeitet“. Es gilt jedoch:

VERTAUSCHBARKEITSKRITERIUM FÜR PARTIELLE ABLEITUNGEN: Existieren alle partiellen Ableitungen zweiter Ordnung und sind *stetig*, so ist

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)$$

für alle $1 \leq i, j \leq n$. Eine analoge Aussage gilt für partielle Ableitungen höherer Ordnung.

Wir führen folgende Abkürzung ein ($k \geq 1$):

$$C^0(D, \mathbb{R}) = \{f : D \rightarrow \mathbb{R} : f \text{ ist stetig}\},$$

$$C^k(D, \mathbb{R}) = \{f : D \rightarrow \mathbb{R} : \text{alle partiellen Ableitungen der Ordnung } \leq k \text{ existieren und sind stetig}\}.$$

VIII.2.5. Die totale Ableitung und lineare Approximation.

Sei wieder $D \subset \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ und $\mathbf{a} \in D$.

Die Funktion f heißt in Punkt \mathbf{a} TOTAL DIFFERENZIERBAR bzw. LINEAR APPROXIMIERBAR, wenn es einen Vektor $\mathbf{g} \in \mathbb{R}^n$ gibt mit

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{1}{|\mathbf{x} - \mathbf{a}|} \{f(\mathbf{x}) - f(\mathbf{a}) - \mathbf{g} \cdot (\mathbf{x} - \mathbf{a})\} = 0.$$

Ist f im Punkt \mathbf{a} total differenzierbar, wird f in der Nähe von \mathbf{a} durch die lineare Funktion

$$g_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + \mathbf{g} \cdot (\mathbf{x} - \mathbf{a})$$

approximiert. Für den Fehler

$$R_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{x}) - g_{\mathbf{a}}(\mathbf{x})$$

dieser Approximation gilt

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{1}{|\mathbf{x} - \mathbf{a}|} R_{\mathbf{a}}(\mathbf{x}) = 0.$$

Man sagt hierfür auch

$$R_{\mathbf{a}}(\mathbf{x}) = o(|\mathbf{x} - \mathbf{a}|).$$

Die Niveaumenge

$$\{\mathbf{x} \in D : f(\mathbf{x}) = f(\mathbf{a})\}$$

wird also in der Nähe von \mathbf{a} durch die Menge

$$\{\mathbf{x} \in D : \mathbf{g}_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{a})\} = \{\mathbf{x} \in D : \mathbf{g} \cdot (\mathbf{x} - \mathbf{a}) = 0\}$$

approximiert. Für $n = 2$ ist dies eine Gerade durch \mathbf{a} senkrecht zu \mathbf{g} ; für $n = 3$ handelt es sich um eine Ebene durch \mathbf{a} senkrecht zu \mathbf{g} .

Es gilt folgender Zusammenhang zwischen partieller und totaler Differenzierbarkeit:

Die Funktion f sei in $\mathbf{a} \in D$ total differenzierbar. Dann gilt:

- (1) f ist in \mathbf{a} stetig.
- (2) Für alle $\mathbf{v} \in \mathbb{R}^n$ ist

$$\lim_{t \rightarrow 0} \frac{1}{t} [f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})] = \mathbf{g} \cdot \mathbf{v}.$$

(3) f ist in \mathbf{a} partiell differenzierbar und $\mathbf{g} = \text{grad } f(\mathbf{a})$.

Teil (3) dieses Ergebnisses besagt:

„ f in \mathbf{a} total differenzierbar $\Rightarrow f$ in \mathbf{a} partiell differenzierbar.“

Die Umkehrung dieser Aussage ist im allgemeinen falsch. Es gilt jedoch:

Jede Funktion $f \in C^1(D, \mathbb{R})$ ist auf D total differenzierbar.

Hieraus ergibt sich:

TEST FÜR DIE TOTALE DIFFERENZIERBARKEIT: Berechne die partiellen Ableitungen und prüfe, ob sie stetig sind.

BEISPIEL VIII.2.14. Für die Funktion $f(x, y) = x^4 + 2x^3y^2 + y$ und den Punkt $\mathbf{a} = (1, 1)$ erhalten wir

$$\begin{aligned} f(1, 1) &= 4, \\ \text{grad } f(x, y) &= \begin{pmatrix} 4x^3 + 6x^2y^2 \\ 4x^3y + 1 \end{pmatrix}, \\ \text{grad } f(1, 1) &= \begin{pmatrix} 10 \\ 5 \end{pmatrix}. \end{aligned}$$

Die partiellen Ableitungen sind Polynome und somit stetig. Daher ist f total differenzierbar. Die lineare Approximation im Punkt $\mathbf{a} = (1, 1)$ lautet

$$g_{\mathbf{a}}(x, y) = 4 + 10(x - 1) + 5(y - 1).$$

Die Niveaumenge $\{(x, y) : f(x, y) = 4\}$ wird in der Nähe von $(1, 1)$ durch die Gerade $10(x - 1) + 5(y - 1) = 0$, d.h. $10x + 5y = 15$ approximiert. Der Graph von f ist die Fläche $z = x^4 + 2x^3y^2 + y$. Der Punkt $(1, 1, 4)$ liegt auf dieser Fläche. Die Gleichung der Tangentialebene an diese Fläche in diesem Punkt ist

$$z = g_{\mathbf{a}}(x, y)$$

d.h.

$$z = 4 + 10(x - 1) + 5(y - 1).$$

VIII.2.6. Einfache Anwendungen. Für einfache Fehler- und Näherungsrechnungen wird eine Funktion f mehrerer Veränderlicher in der Nähe eines Punktes \mathbf{a} durch ihre lineare Approximation ersetzt.

BEISPIEL VIII.2.15. $f(x, y) = x^y = e^{y \ln(x)}$ wird in $(1, 3)$ wegen

$$f(1, 3) = 1,$$

$$\begin{aligned}\text{grad } f(x, y) &= \begin{pmatrix} yx^{y-1} \\ \ln(x)x^y \end{pmatrix}, \\ \text{grad } f(1, 3) &= \begin{pmatrix} 3 \\ 0 \end{pmatrix}\end{aligned}$$

durch $1+3(x-1)$ linear approximiert. Damit erhält man z.B. für $1.02^{3.01}$ den Näherungswert 1.06. Der genaue Wert ist $1.061418168\dots$

Werden statt der wahren Werte $\mathbf{x} = (x_1, \dots, x_n)$ die Näherungswerte $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$ gemessen, dann belasten die Messfehler $|\Delta x_i| = |x_i - x_{0i}|$ den Funktionswert mit dem Fehler

$$\begin{aligned}|\Delta f(\mathbf{x})| &= |f(\mathbf{x}) - f(\mathbf{x}_0)| \\ &\approx |\text{grad } f(\mathbf{x}_0) \cdot (x - x_0)| \\ &= \left| \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0)(x_i - x_{0i}) \right| \\ &\leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \right| |x_i - x_{0i}|.\end{aligned}$$

Sind also für die einzelnen Messungen Fehlerschranken $|\Delta x_i| \leq S_i$ bekannt, so erhält man eine ungefähre Fehlerschranke S für die Funktionswerte ($|\Delta f(\mathbf{x})| \leq S$) durch

$$S \approx \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \right| S_i.$$

BEISPIEL VIII.2.16. Für den Elastizitätsmodul E eines Stabes mit quadratischem Querschnitt a , Länge ℓ , Ausbiegung h und Belastung F gilt

$$E = E(F, h, \ell, a) = 4F\ell^3h^{-1}a^{-4}.$$

Misst man $\ell = 50$ cm auf 1%, $a = 2$ cm auf 1%, $h = 2$ mm auf 3% und $F = 130$ N auf 0.5% genau, so ergibt die Fehlerrechnung bei linearer Approximation

$$\begin{aligned}\left| \frac{\Delta E}{E} \right| &\approx \frac{1}{|E|} \left| \frac{\partial E}{\partial F} \Delta F + \frac{\partial E}{\partial h} \Delta h + \frac{\partial E}{\partial \ell} \Delta \ell + \frac{\partial E}{\partial a} \Delta a \right| \\ &= \frac{1}{|4F\ell^3h^{-1}a^{-4}|} |4\ell^3h^{-1}a^{-4} \Delta F - 4F\ell^3h^{-2}a^{-4} \Delta h \\ &\quad + 12F\ell^2h^{-1}a^{-4} \Delta \ell - 16F\ell^3h^{-2}a^{-5} \Delta a| \\ &\leq \left| \frac{\Delta F}{F} \right| + \left| \frac{\Delta h}{h} \right| + 3 \left| \frac{\Delta \ell}{\ell} \right| + 4 \left| \frac{\Delta a}{a} \right| \\ &= 0.105\end{aligned}$$

und somit

$$\begin{aligned} E &= 203125 \frac{\text{N}}{\text{mm}^2} \pm 10.5\% \\ &= 203125 \pm 21328.125 \frac{\text{N}}{\text{mm}^2}. \end{aligned}$$

VIII.2.7. Die Richtungsableitung. Die partiellen Ableitungen einer Funktion f geben die Änderung der Funktionswerte entlang der Koordinatenachsen an. Es ist naheliegend, die entsprechenden Änderungen entlang beliebiger Richtungen zu betrachten. Ist daher $\mathbf{v} \in \mathbb{R}^n$ ein beliebiger *Einheitsvektor*, d.h. $|\mathbf{v}| = 1$, so nennen wir den Grenzwert

$$\partial_{\mathbf{v}} f(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{1}{t} [f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})]$$

– sofern er existiert – die **RICHTUNGSABTEILUNG VON f IN RICHTUNG \mathbf{v} AN DER STELLE \mathbf{a} .**

Es gilt folgender Zusammenhang mit totaler und partieller Differenzierbarkeit.

Die Funktion $f : D \rightarrow \mathbb{R}$ sei im Punkt \mathbf{a} total differenzierbar. Dann existieren alle Richtungsableitungen $\partial_{\mathbf{v}} f(\mathbf{a})$ von f in \mathbf{a} und es ist

$$\partial_{\mathbf{v}} f(\mathbf{a}) = \text{grad } f(\mathbf{a}) \cdot \mathbf{v} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{a}) v_i.$$

BEISPIEL VIII.2.17. (1) Ein Massenpunkt durchquere mit der konstanten Geschwindigkeit $\mathbf{v} = (2, 4, 7)^T$ einen Raum mit der Temperaturverteilung $T = f(x, y, z) = 2xy + 4yz$. Er erfährt dann im Punkt \mathbf{a} die momentane Temperaturänderung

$$|\mathbf{v}| \partial_{\frac{\mathbf{v}}{|\mathbf{v}|}} f(\mathbf{a}) = |\mathbf{v}| \text{grad } f(\mathbf{a}) \cdot \frac{\mathbf{v}}{|\mathbf{v}|} = \text{grad } f(\mathbf{a}) \cdot \mathbf{v}.$$

Für $\mathbf{a} = (1, 1, 1)$ ergibt sich insbesondere der Wert

$$\text{grad } f(1, 1, 1) \cdot \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix} = \begin{pmatrix} 2 \\ 2+4 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix} = 56.$$

(2) Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Für die zugehörige quadratische Form q gilt

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^T A \mathbf{x} \\ \text{grad } q(\mathbf{x}) &= 2A \mathbf{x} \end{aligned}$$

$$\begin{aligned}\partial_{\mathbf{v}}q(\mathbf{x}) &= 2(A\mathbf{x}) \cdot \mathbf{v} \\ &= 2\mathbf{x}^T A\mathbf{v}.\end{aligned}$$

VIII.2.8. Die Kettenregel. Seien $D \subset \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ partiell differenzierbar und $\mathbf{x} : [a, b] \rightarrow D$ die Parameterdarstellung eines Kurvenstückes in D . Dann wird durch $[a, b] \ni t \mapsto f(\mathbf{x}(t))$ eine Funktion $\varphi : [a, b] \rightarrow \mathbb{R}$ definiert. Diese Funktion ist differenzierbar und es gilt die KETTENREGEL

$$\begin{aligned}\frac{d}{dt}f(\mathbf{x}(t)) &= \text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}(t))\dot{x}_i(t).\end{aligned}$$

BEISPIEL VIII.2.18. Für die Funktion $f(x, y, z) = x^2y^3z$ und die Schraubenlinie $\mathbf{x}(t) = (r \cos t, r \sin t, ht)^T$ erhalten wir:

$$\begin{aligned}\text{grad } f(x, y, z) &= \begin{pmatrix} 2xy^3z \\ 3x^2y^2z \\ x^2y^3 \end{pmatrix}, \\ \dot{\mathbf{x}}(t) &= \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix}, \\ \frac{d}{dt}f(\mathbf{x}(t)) &= \begin{pmatrix} 2r^4ht \cos t \sin^3 t \\ 3r^4ht \cos^2 t \sin^2 t \\ r^5 \cos^2 t \sin^3 t \end{pmatrix} \cdot \begin{pmatrix} -r \sin t \\ r \cos t \\ h \end{pmatrix} \\ &= -2r^5ht \cos t \sin^4 t + 3r^5ht \cos^3 t \sin^2 t \\ &\quad + r^5h \cos^2 t \sin^3 t \\ &= r^5h \cos t \sin^2 t [-2t \sin^2 t + 3t \cos^2 t + \cos t \sin t].\end{aligned}$$

Die Kettenregel benötigt man immer dann, wenn neue Variablen eingeführt werden und die partiellen Ableitungen bzgl. der neuen Variablen zu berechnen sind.

BEISPIEL VIII.2.19 (EBENE POLARKOORDINATEN). Durch $x = r \cos \varphi$, $y = r \sin \varphi$ wird die Funktion $f(x, y)$ transformiert in

$$F(r, \varphi) = f(r \cos \varphi, r \sin \varphi).$$

Falls $f \in C^2$ ist, ergibt sich für die partiellen Ableitungen von F mit der Kettenregel

$$\begin{aligned}F_r &= f_x \cdot \cos \varphi + f_y \cdot \sin \varphi \\ F_\varphi &= f_x \cdot (-r \sin \varphi) + f_y \cdot (r \cos \varphi) \\ F_{rr} &= f_{xx} \cdot \cos^2 \varphi + 2f_{xy} \cdot \cos \varphi \sin \varphi + f_{yy} \cdot \sin^2 \varphi\end{aligned}$$

$$\begin{aligned}
F_{r\varphi} &= f_{xx} \cdot (-r \sin \varphi \cos \varphi) + f_{xy} \cdot (-r \sin^2 \varphi + r \cos^2 \varphi) \\
&\quad + f_{yy} \cdot (r \cos \varphi \sin \varphi) + f_x \cdot (-\sin \varphi) + f_y \cdot (\cos \varphi) \\
F_{\varphi\varphi} &= f_{xx} \cdot (r^2 \sin^2 \varphi) + f_{xy} \cdot (-2r^2 \cos \varphi \sin \varphi) + f_{yy} \cdot (r^2 \cos^2 \varphi) \\
&\quad + f_x(-r \cos \varphi) + f_y(-r \sin \varphi).
\end{aligned}$$

Löst man diese Beziehungen nach der partiellen Ableitung von f auf, erhält man für $r \neq 0$

$$\begin{aligned}
f_x &= F_r \cdot \cos \varphi - \frac{1}{r} F_\varphi \sin \varphi \\
f_y &= F_r \cdot \sin \varphi + \frac{1}{r} F_\varphi \cos \varphi.
\end{aligned}$$

Hieraus ergibt sich für den LAPLACE-OPERATOR (vgl. Abschnitt VIII.4.4 (S. 319)):

$$\begin{aligned}
\Delta f &= f_{xx} + f_{yy} \\
&= F_{rr} + \frac{1}{r} F_r + \frac{1}{r^2} F_{\varphi\varphi} \\
&= \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial F}{\partial r} \right) + \frac{1}{r^2} F_{\varphi\varphi}.
\end{aligned}$$

BEISPIEL VIII.2.20 (RÄUMLICHE KUGELKOORDINATEN). Mit dem Ansatz $x = r \cos \varphi \sin \theta$, $y = r \sin \varphi \sin \theta$, $z = r \cos \theta$ mit $r > 0$, $0 \leq \varphi < 2\pi$, $0 \leq \theta < \pi$ erhält man für die Funktion $f(x, y, z)$ die transformierte Funktion

$$F(r, \varphi, \theta) = f(r \cos \varphi \sin \theta, r \sin \varphi \sin \theta, r \cos \theta).$$

Anwenden der Kettenregel liefert

$$\begin{aligned}
F_r &= f_x \cdot (\cos \varphi \sin \theta) + f_y \cdot (\sin \varphi \sin \theta) + f_z \cdot (\cos \theta) \\
F_\varphi &= f_x \cdot (-r \sin \varphi \sin \theta) + f_y \cdot (r \cos \varphi \sin \theta) \\
F_\theta &= f_x \cdot (r \cos \varphi \cos \theta) + f_y \cdot (r \sin \varphi \cos \theta) + f_z \cdot (-r \sin \theta).
\end{aligned}$$

Hieraus ergibt sich für den LAPLACE-OPERATOR (vgl. Abschnitt VIII.4.4 (S. 319)):

$$\begin{aligned}
\Delta f &= f_{xx} + f_{yy} + f_{zz} \\
&= F_{rr} + \frac{2}{r} F_r + \frac{1}{r^2 \sin^2 \theta} F_{\varphi\varphi} + \frac{1}{r^2} F_{\theta\theta} + \frac{1}{r^2} (\cot \theta) \cdot F_\theta \\
&= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial F}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} F_{\varphi\varphi} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial F}{\partial \theta} \right).
\end{aligned}$$

VIII.3. Anwendungen der Differentiation

VIII.3.1. Richtung des stärksten Anstieges. Für eine differenzierbare Funktion $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ ist der Anstieg in einem Punkt $\mathbf{a} \in D$ in Richtung eines Einheitsvektors $\mathbf{v} \in \mathbb{R}^n, |\mathbf{v}| = 1$, gemäß Abschnitt VIII.2.7 (S. 292) gegeben durch die Richtungsableitung

$$\partial_{\mathbf{v}} f(\mathbf{a}) = \text{grad } f(\mathbf{a}) \cdot \mathbf{v} = |\text{grad } f(\mathbf{a})| \cos \alpha.$$

Dabei ist α der Winkel zwischen den Vektoren \mathbf{v} und $\text{grad } f(\mathbf{a})$. Also ist der Anstieg maximal, wenn $\cos \alpha$ maximal ist, d.h. wenn $\alpha = 0$ ist.

Das bedeutet:

$\text{grad } f(\mathbf{a})$ = Richtung des maximalen Anstieges der Funktion f im Punkt \mathbf{a} .

Da f genau dann ansteigt, wenn $-f$ abfällt, gilt:

$-\text{grad } f(\mathbf{a})$ = Richtung des maximalen Abfallens der Funktion f im Punkt \mathbf{a} .

Diese Beobachtung führt auf das GRADIENTENVERFAHREN zur Bestimmung eines Maximums (bzw. Minimums) einer Funktion f :

Ausgehend von einer Startnäherung \mathbf{x}_0 berechne $\text{grad } f(\mathbf{x}_0)$ und setze $\mathbf{x}_1 = \mathbf{x}_0 + h \text{grad } f(\mathbf{x}_0)$ (bzw. $\mathbf{x}_1 = \mathbf{x}_0 - h \text{grad } f(\mathbf{x}_0)$) mit einer problemangepassten Schrittweite $h > 0$. Falls $f(\mathbf{x}_1) > f(\mathbf{x}_0)$ (bzw. $f(\mathbf{x}_1) < f(\mathbf{x}_0)$) ist, setze das Verfahren mit \mathbf{x}_1 an Stelle von \mathbf{x}_0 fort. Andernfalls halbiere h und probiere den entsprechenden neuen Wert von \mathbf{x}_1 aus.

VIII.3.2. Tangenten und Tangentialebene. Ist $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ eine differenzierbare Funktion, definiert die Gleichung $f(\mathbf{x}) = c$ mit konstantem $c \in \mathbb{R}$ eine Hyperfläche in D . Für jede Kurve $\mathbf{x}(t)$ auf dieser Hyperfläche gilt $f(\mathbf{x}(t)) = c$ für alle t . Durch Differentiation folgt mit der Kettenregel

$$\text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) = 0 \quad \text{für alle } t.$$

Wir werten diese Formel nun für die wichtigen Spezialfälle $n = 2$ und $n = 3$ aus.

„ $n = 2$ “: In diesem Fall beschreibt $f(\mathbf{x}) = c$ eine Kurve. Ist (x_0, y_0) ein Punkt auf dieser Kurve, besagt obige Gleichung, dass $\text{grad } f(x_0, y_0)$ senkrecht ist zur Tangente an die Kurve im Punkt (x_0, y_0) . Also lautet die NORMALENGLEICHUNG DER TANGENTE AN DIE NIVEAUKURVE $f(x, y) = c$ IM PUNKT (x_0, y_0) :

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

„ $n = 3$ “: In diesem Fall beschreibt $f(\mathbf{x}) = c$ eine Fläche. Ist $\mathbf{x}_0 = (x_0, y_0, z_0)$ ein Punkt auf dieser Fläche, besagt obige Gleichung, dass grad $f(\mathbf{x}_0)$ senkrecht steht zu allen Tangenten an die Fläche im Punkt \mathbf{x}_0 . Alle diese Tangenten spannen eine Ebene auf, die sog. TANGENTIALEBENE. Daher lautet die NORMALENGLEICHUNG DER TANGENTIALEBENE AN DIE NIVEAUFLÄCHE $f(x, y, z) = c$ IM PUNKT (x_0, y_0, z_0) :

$$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0.$$

BEISPIEL VIII.3.1. (1) Betrachte einen Kreis mit Mittelpunkt (u, v) und Radius r . Dieser entspricht der Niveaukurve

$$f(x, y) = (x - u)^2 + (y - v)^2 = r^2.$$

Ist daher (x_0, y_0) ein Punkt auf dem Kreis, lautet die Normalengleichung der Tangente in diesem Punkt

$$2(x_0 - u)(x - x_0) + 2(y_0 - v)(y - y_0) = 0.$$

Dividieren wir diese Gleichung durch 2, schreiben

$$\begin{aligned} x - x_0 &= x - u + u - x_0 \\ y - y_0 &= y - v + v - y_0 \end{aligned}$$

und nutzen die Kreisgleichung für (x_0, y_0) aus, erhalten wir die bekannte Darstellung

$$(x_0 - u)(x - u) + (y_0 - v)(y - v) = r^2.$$

(2) Das Paraboloid $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2pz = 0$ mit $p \neq 0$ besitzt im Flächenpunkt (x_0, y_0, z_0) die Tangentialebene

$$\frac{2x_0(x - x_0)}{a^2} + \frac{2y_0(y - y_0)}{b^2} - 2p(z - z_0) = 0.$$

Dividieren wir diese Gleichung durch 2 und nutzen die Beziehung

$$\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} - 2pz_0 = 0$$

aus, erhalten wir die Darstellung

$$\frac{x_0x}{a^2} + \frac{y_0y}{b^2} - p(z + z_0) = 0.$$

VIII.3.3. Die Taylor-Formel. Sei $D \subset \mathbb{R}^n$ eine offene Menge, \mathbf{a} ein Punkt in D , $f : D \rightarrow \mathbb{R}$ eine hinreichend oft differenzierbare Funktion und $\mathbf{v} \in \mathbb{R}^n$ mit $|\mathbf{v}| = 1$ ein Einheitsvektor. Dann gibt es ein $\varepsilon > 0$, so dass die Funktion

$$\varphi(t) = f(\mathbf{a} + t\mathbf{v})$$

auf $(-\varepsilon, \varepsilon)$ definiert und differenzierbar ist. Für diese Funktion gilt die Taylor-Formel aus Abschnitt VII.3.1 (S. 267)

$$\begin{aligned} \varphi(t) &= \varphi(0) + \dot{\varphi}(0)t + \frac{1}{2}\ddot{\varphi}(0)t^2 \\ &+ \dots + \frac{1}{k!}\varphi^{(k)}(0)t^k + \frac{1}{(k+1)!}\varphi^{(k+1)}(\eta)t^{k+1} \end{aligned}$$

mit einem geeigneten $\eta \in (0, 1)$.

Wir wollen diese Formel durch die Funktion f und deren Ableitungen ausdrücken.

Konstruktionsgemäß ist

$$\begin{aligned} \varphi(t) &= f(\mathbf{a} + t\mathbf{v}), \\ \varphi(0) &= f(\mathbf{a}). \end{aligned}$$

Wegen der Kettenregel ist

$$\begin{aligned} \dot{\varphi}(t) &= \text{grad } f(\mathbf{a} + t\mathbf{v}) \cdot \mathbf{v}, \\ \dot{\varphi}(0) &= \text{grad } f(\mathbf{a}) \cdot \mathbf{v} \\ &= \partial_{\mathbf{v}} f(\mathbf{a}). \end{aligned}$$

Wir wollen auch die höheren Ableitungen $\ddot{\varphi}, \dots, \varphi^{(k)}$ durch Richtungsableitungen von f in Richtung \mathbf{v} ausdrücken. Dazu führen wir den Differentialoperator

$$\partial_{\mathbf{v}} = v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} + \dots + v_n \frac{\partial}{\partial x_n}$$

ein und bezeichnen mit $\partial_{\mathbf{v}}^k$ dessen k -malige Ausführung. Dann ist

$$\dot{\varphi} = \partial_{\mathbf{v}} f, \quad \ddot{\varphi} = \partial_{\mathbf{v}}^2 f, \quad \dots, \quad \varphi^{(k)} = \partial_{\mathbf{v}}^k f.$$

BEISPIEL VIII.3.2. Für

$$f(x, y) = x^3 y^2$$

und

$$\mathbf{v} = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

erhalten wir

$$\partial_{\mathbf{v}} f = \frac{1}{5}(9x^2 y^2 + 8x^3 y),$$

$$\begin{aligned}\partial_{\mathbf{v}}^2 f &= \frac{1}{25} \{3(18xy^2 + 24x^2y) + 4(18x^2y + 8x^3)\} \\ &= \frac{1}{25} \{54xy^2 + 144x^2y + 32x^3\} \\ \partial_{\mathbf{v}}^3 f &= \frac{1}{125} \{3(54y^2 + 288xy + 96x^2) + 4(108xy + 144x^2)\} \\ &= \frac{1}{125} \{162y^2 + 1296xy + 864x^2\}.\end{aligned}$$

Drücken wir in der Taylor-Formel für φ alle Ableitungen durch $\partial_{\mathbf{v}}$ aus, erhalten wir die TAYLOR-FORMEL IN n VERÄNDERLICHEN:

$$\begin{aligned}f(\mathbf{a} + t\mathbf{v}) &= f(\mathbf{a}) + \partial_{\mathbf{v}}f(\mathbf{a})t + \frac{1}{2}\partial_{\mathbf{v}}^2f(\mathbf{a})t^2 + \dots + \frac{1}{k!}\partial_{\mathbf{v}}^kf(\mathbf{a})t^k \\ &\quad + \frac{1}{(k+1)!}\partial_{\mathbf{v}}^{k+1}f(\mathbf{a} + \eta\mathbf{v})t^{k+1} \quad \text{mit einem } \eta \in (0, 1).\end{aligned}$$

Sind \mathbf{x} und \mathbf{x}_0 zwei verschiedene Punkte in D derart, dass die Strecke von \mathbf{x}_0 nach \mathbf{x} ganz in D verläuft, können wir in obiger Formel $\mathbf{a} = \mathbf{x}_0$, $\mathbf{v} = \frac{1}{|\mathbf{x} - \mathbf{x}_0|}(\mathbf{x} - \mathbf{x}_0)$ und $t = |\mathbf{x} - \mathbf{x}_0|$ wählen. Wir erhalten dann

$$\begin{aligned}f(\mathbf{x}) &= f(\mathbf{x}_0) + \partial_{\mathbf{v}}f(\mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0| + \frac{1}{2}\partial_{\mathbf{v}}^2f(\mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0|^2 \\ &\quad + \dots + \frac{1}{k!}\partial_{\mathbf{v}}^kf(\mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0|^k \\ &\quad + \frac{1}{(k+1)!}\partial_{\mathbf{v}}^{k+1}f(\mathbf{x}_0 + \eta(\mathbf{x} - \mathbf{x}_0))|\mathbf{x} - \mathbf{x}_0|^{k+1} \\ &\quad \text{mit } \mathbf{v} = \frac{1}{|\mathbf{x} - \mathbf{x}_0|}(\mathbf{x} - \mathbf{x}_0) \text{ und } \eta \in (0, 1).\end{aligned}$$

Der Ausdruck

$$\begin{aligned}T_k(f; \mathbf{x}_0)(\mathbf{x}) &= f(\mathbf{x}_0) + \partial_{\mathbf{v}}f(\mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0| + \frac{1}{2}\partial_{\mathbf{v}}^2f(\mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0|^2 \\ &\quad + \dots + \frac{1}{k!}\partial_{\mathbf{v}}^kf(\mathbf{x}_0)|\mathbf{x} - \mathbf{x}_0|^k \\ &\quad \text{mit } \mathbf{v} = \frac{1}{|\mathbf{x} - \mathbf{x}_0|}(\mathbf{x} - \mathbf{x}_0)\end{aligned}$$

ist ein Polynom vom Grade k in \mathbf{x} und heißt k -tes TAYLOR-POLYNOM von f mit Entwicklungspunkt \mathbf{x}_0 .

BEISPIEL VIII.3.3. Wir wollen für die Funktion f aus Beispiel VIII.3.2 das Taylor-Polynom $T_2(f; \mathbf{x}_0)$ zum Entwicklungspunkt $\mathbf{x}_0 =$

(1, 2) bestimmen. Dazu betrachten wir zunächst einen beliebigen Einheitsvektor \mathbf{v} . Dann ist

$$\begin{aligned}\partial_{\mathbf{v}} f &= v_1 f_x + v_2 f_y \\ &= v_1 3x^2 y^2 + v_2 2x^3 y, \\ \partial_{\mathbf{v}^2} f &= v_1 [v_1 f_x + v_2 f_y]_x + v_2 [v_1 f_x + v_2 f_y]_y \\ &= v_1 [v_1 6xy^2 + v_2 6x^2 y] + v_2 [v_1 6x^2 y + v_2 2x^3] \\ &= v_1^2 6xy^2 + v_1 v_2 12x^2 y + v_2^2 2x^3.\end{aligned}$$

Damit ergibt sich

$$\begin{aligned}\partial_{\mathbf{v}} f(\mathbf{x}_0) &= 12v_1 + 4v_2, \\ \partial_{\mathbf{v}^2} f(\mathbf{x}_0) &= 24v_1^2 + 24v_1 v_2 + 2v_2^2.\end{aligned}$$

Sei nun $\mathbf{x} \neq \mathbf{x}_0$ ein beliebiger Punkt. Mit dem speziellen Einheitsvektor $\mathbf{v} = \frac{1}{|\mathbf{x} - \mathbf{x}_0|}(\mathbf{x} - \mathbf{x}_0)$ erhalten wir aus obiger Formel

$$\begin{aligned}|\mathbf{x} - \mathbf{x}_0| \partial_{\mathbf{v}} f(\mathbf{x}_0) &= 12(x - 1) + 4(y - 2), \\ |\mathbf{x} - \mathbf{x}_0|^2 \partial_{\mathbf{v}^2} f(\mathbf{x}_0) &= 24(x - 1)^2 + 24(x - 1)(y - 2) + 2(y - 2)^2.\end{aligned}$$

Wegen $f(\mathbf{x}_0) = 4$ ergibt sich damit

$$\begin{aligned}T_2(f; \mathbf{x}_0)(\mathbf{x}) &= 4 + 12(x - 1) + 4(y - 2) \\ &\quad + 12(x - 1)^2 + 12(x - 1)(y - 2) + (y - 2)^2.\end{aligned}$$

VIII.3.4. Die Hesse-Matrix. Für eine zweimal differenzierbare Funktion f heißt die symmetrische Matrix

$$\begin{aligned}H_f(\mathbf{a}) &= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) \right)_{1 \leq i, j \leq n} \\ &= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{pmatrix}\end{aligned}$$

die HESSE-MATRIX von f an der Stelle \mathbf{a} .

Mit Hilfe der Hesse-Matrix können wir das 2-te Taylor-Polynom von f besonders leicht darstellen

$$\begin{aligned}T_2(f; \mathbf{x}_0)(\mathbf{x}) &= f(\mathbf{x}_0) + \text{grad } f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \\ &\quad + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).\end{aligned}$$

BEISPIEL VIII.3.4. Für die Funktion $f(x, y) = x^3y^2$ aus Beispiel VIII.3.2 ist

$$\begin{aligned}\text{grad } f(x, y) &= \begin{pmatrix} 3x^2y^2 \\ 2x^3y \end{pmatrix}, \\ H_f(x, y) &= \begin{pmatrix} 6xy^2 & 6x^2y \\ 6x^2y & 2x^3 \end{pmatrix}.\end{aligned}$$

Damit ergibt sich für den Entwicklungspunkt $\mathbf{x}_0 = (1, 2)^T$ aus Beispiel VIII.3.3

$$\begin{aligned}T_2(f; \mathbf{x}_0)(\mathbf{x}) &= 4 + \begin{pmatrix} 12 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} \\ &\quad + \frac{1}{2}(x-1, y-2) \begin{pmatrix} 24 & 12 \\ 12 & 2 \end{pmatrix} \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} \\ &= 4 + 12(x-1) + 4(y-2) \\ &\quad + 12(x-1)^2 + 12(x-1)(y-2) + (y-2)^2.\end{aligned}$$

Aus der Taylor-Formel folgt

$$f(\mathbf{x}) = T_2(f; \mathbf{x}_0)(\mathbf{x}) + R_2(\mathbf{x}; \mathbf{x}_0)$$

mit

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{1}{|\mathbf{x} - \mathbf{x}_0|^2} R_2(\mathbf{x}; \mathbf{x}_0) = 0.$$

D.h., die Funktion f wird durch das Polynom $T_2(f; \mathbf{x}_0)(\mathbf{x})$ in der Nähe von \mathbf{x}_0 mit einem Fehler $o(|\mathbf{x} - \mathbf{x}_0|^2)$ approximiert. Zusammen mit der Darstellung von $T_2(f; \mathbf{x}_0)(\mathbf{x})$ hat dies in den Fällen $n = 2$ und $n = 3$ folgende anschauliche Deutung:

„ $n = 2$ “: Der Graph $z = f(x, y)$ wird in der Nähe des Flächenpunktes $(x_0, y_0, f(x_0, y_0))$ approximiert durch die SCHMIEGEQUADRIK

$$\begin{aligned}z &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &\quad + \frac{1}{2}f_{xx}(x_0, y_0)(x - x_0)^2 + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) \\ &\quad + \frac{1}{2}f_{yy}(x_0, y_0)(y - y_0)^2.\end{aligned}$$

Falls $H_f(x_0, y_0) \neq 0$ ist, ist diese ein Paraboloid oder ein parabolischer Zylinder. Man nennt den Flächenpunkt über (x_0, y_0) FLACH, wenn die Schmiegequadrik eine Ebene ist, d.h. wenn $H_f(x_0, y_0) = 0$ ist. Man nennt den Flächenpunkt ELLIPTISCH bzw. HYPERBOLISCH bzw. PARABOLISCH, wenn die Schmiegequadrik ein elliptisches Paraboloid bzw. ein hyperbolisches Paraboloid bzw. ein parabolischer Zylinder ist.

BEISPIEL VIII.3.5. Der AFFENSATTEL $z = x^3 - 3xy^2$ hat im Punkt $(x, y) = (0, 0)$ einen Flachpunkt mit der (x, y) -Ebene als Tangentialebene. Die Hesse-Matrix ist

$$H_f = \begin{pmatrix} 6x & -6y \\ -6y & -6x \end{pmatrix}.$$

Ihre Determinante ist $-36(x^2 + y^2)$ und ist außerhalb von $(0, 0)$ stets negativ. Daher ist jeder Punkt der Fläche mit $(x, y) \neq (0, 0)$ hyperbolisch.

„ $n = 3$ “: Die Niveaufläche $f(x, y, z) = c$ wird in der Nähe eines Flächenpunktes (x_0, y_0, z_0) approximiert durch die Quadrik

$$\begin{aligned} & \text{grad } f(x_0, y_0, z_0) \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \\ & + \frac{1}{2} \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix}^T H_f(x_0, y_0, z_0) \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0. \end{aligned}$$

BEISPIEL VIII.3.6. Betrachte die Niveaufläche

$$x \cos y + y \cos z + z \cos x = 2$$

und den Flächenpunkt $(0, 0, 2)$. Es ist

$$\begin{aligned} f(x, y, z) &= x \cos y + y \cos z + z \cos x, \\ \text{grad } f(x, y, z) &= \begin{pmatrix} \cos y - z \sin x \\ -x \sin y + \cos z \\ -y \sin z + \cos x \end{pmatrix}, \\ \text{grad } f(0, 0, 2) &= \begin{pmatrix} 1 \\ \cos 2 \\ 1 \end{pmatrix}, \\ H_f(x, y, z) &= \begin{pmatrix} -z \cos x & -\sin y & -\sin x \\ -\sin y & -x \cos y & -\sin z \\ -\sin x & -\sin z & -y \cos z \end{pmatrix}, \\ H_f(0, 0, 2) &= \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & -\sin 2 \\ 0 & -\sin 2 & 0 \end{pmatrix}. \end{aligned}$$

Daher lautet die Schmiegequadrik

$$x + y \cos 2 + z - 2 - x^2 - y(z - 2) \sin 2 = 0.$$

Dies ist ein zweischaliges Hyperboloid.

VIII.3.5. Implizite Funktionen. Sei D eine offene Teilmenge des \mathbb{R}^2 und $f : D \rightarrow \mathbb{R}$ eine hinreichend oft differenzierbare Funktion. Wir betrachten die Niveaufläche $f(x, y) = 0$ und fragen uns, ob sie in der Nähe eines Punktes (x_0, y_0) mit $f(x_0, y_0) = 0$ als Graph einer Funktion g der Variablen x darstellbar ist. Wenn dies der Fall ist, muss $g(x_0) = y_0$ sein und für alle x in der Nähe von x_0 muss gelten

$$f(x, g(x)) = 0.$$

Falls g differenzierbar ist, folgt hieraus mit der Kettenregel

$$f_x(x, g(x)) + f_y(x, g(x))g'(x) = 0.$$

Diese Gleichung sollte zumindest für $x = x_0$ nach der Steigung $g'(x_0)$ auflösbar sein. Das ist aber dann der Fall, wenn

$$f_y(x_0, g(x_0)) = f_y(x_0, y_0) \neq 0$$

ist. In der Tat kann man zeigen, dass $f_y(x_0, y_0) \neq 0$ tatsächlich die Darstellbarkeit der Niveaukurve als Graph garantiert.

SATZ ÜBER IMPLIZITE FUNKTIONEN: Die Funktion f sei auf der offenen Menge $D \subset \mathbb{R}^2$ differenzierbar. Für $(x_0, y_0) \in D$ gelte

$$f(x_0, y_0) = 0 \quad \text{und} \quad f_y(x_0, y_0) \neq 0.$$

Dann gibt es ein offenes Intervall I mit $x_0 \in I$ und eine differenzierbare Funktion $g : I \rightarrow \mathbb{R}$ mit

- (1) $g(x_0) = y_0$,
- (2) $(x, g(x)) \in D$ für alle $x \in I$,
- (3) $f(x, g(x)) = 0$ für alle $x \in I$.
- (4) $f_y(x, g(x)) \neq 0$ für alle $x \in I$.

Für die Ableitung von g gilt

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))}.$$

Falls $f \in C^k$ ist, $k \geq 2$, ist g k -mal differenzierbar. Die Ableitungen von g erhält man durch Differenzieren obiger Gleichung.

BEISPIEL VIII.3.7. Wir betrachten die nichtlineare Gleichung

$$(*) \quad e^y + y^3 + x^3 + x^2 - 1 = 0$$

oder äquivalent

$$e^y + y^3 = 1 - x^2 - x^3.$$

Um die Lösbarkeit dieser Gleichung zu untersuchen, betrachten wir die Funktion

$$h(y) = e^y + y^3.$$

Wegen

$$h'(y) = e^y + 3y^2 > 0$$

ist sie monoton wachsend. Weiter ist

$$\lim_{y \rightarrow -\infty} h(y) = -\infty, \quad \lim_{y \rightarrow \infty} h(y) = \infty.$$

Daher ist $h : \mathbb{R} \rightarrow \mathbb{R}$ bijektiv und zu jedem $x \in \mathbb{R}$ gibt es genau ein $y \in \mathbb{R}$ mit

$$1 - x^2 - x^3 = h(y) = e^y + y^3.$$

Also definiert die Gleichung (*) implizit eine Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$e^{g(x)} + g(x)^3 + x^3 + x^2 - 1 = 0 \quad \text{für alle } x.$$

Alle Versuche, diese Gleichung nach g aufzulösen, scheitern. Wir können aber den Satz über implizite Funktionen anwenden:

Es ist

$$\begin{aligned} f(x, y) &= e^y + y^3 + x^3 + x^2 - 1, \\ f_x(x, y) &= 3x^2 + 2x, \\ f_y(x, y) &= e^y + 3y^2 > 0 \quad \text{für alle } y. \end{aligned}$$

Damit folgt

$$g'(x) = -\frac{3x^2 + 2x}{e^{g(x)} + 3g(x)^2} \quad \text{für alle } x.$$

Durch Differenzieren dieser Gleichung können wir Ableitungen beliebiger Ordnung von g berechnen, z.B.

$$\begin{aligned} g''(x) &= -\frac{6x + 2}{e^{g(x)} + 3g(x)^2} + \frac{3x^2 + 2x}{(e^{g(x)} + 3g(x)^2)^2} \cdot (g'(x)e^{g(x)} + 6g'(x)g(x)) \\ &= -\frac{6x + 2}{e^{g(x)} + 3g(x)^2} + \frac{3x^2 + 2x}{(e^{g(x)} + 3g(x)^2)^2} \cdot (e^{g(x)} + 6g(x))g'(x) \\ &= -\frac{6x + 2}{e^{g(x)} + 3g(x)^2} - (e^{g(x)} + 6g(x)) \cdot \frac{(3x^2 + 2x)^2}{(e^{g(x)} + 3g(x)^2)^3}. \end{aligned}$$

Für $x = 0$ erhalten wir wegen $e^0 = 1$ z.B.

$$\begin{aligned} g(0) &= 0 \\ g'(0) &= 0 \\ g''(0) &= -2 \end{aligned}$$

und

$$g(x) = -x^2 + o(x^2).$$

Mit Hilfe des Satzes über implizite Funktionen können auch die Extrema von implizit definierten Funktionen bestimmt werden:

Differenzieren wir die Identität für g' im Satz über implizite Funktionen ein weiteres Mal, erhalten wir

$$\begin{aligned} g''(x) &= -\frac{f_{xx}(x, g(x)) + f_{xy}(x, g(x))g'(x)}{f_y(x, g(x))} \\ &\quad + \frac{f_x(x, g(x))}{f_y(x, g(x))^2} (f_{xy}(x, g(x)) + f_{yy}(x, g(x))g'(x)). \end{aligned}$$

Insbesondere ist $g'(x_0) = 0$, wenn $f_x(x_0, y_0) = 0$ ist. In diesem Fall vereinfacht sich der Ausdruck für $g''(x_0)$ zu

$$g''(x_0) = -\frac{f_{xx}(x_0, y_0)}{f_y(x_0, y_0)}.$$

Damit folgt

Die durch $f(x, y) = 0$ implizit definierte Funktion $y = g(x)$ hat an der Stelle (x_0, y_0) eine horizontale Tangente, wenn gilt

$$f(x_0, y_0) = 0, \quad f_x(x_0, y_0) = 0, \quad f_y(x_0, y_0) \neq 0.$$

x_0 ist ein lokales Maximum von g , wenn gilt

$$\frac{f_{xx}(x_0, y_0)}{f_y(x_0, y_0)} > 0.$$

x_0 ist ein lokales Minimum von g , wenn gilt

$$\frac{f_{xx}(x_0, y_0)}{f_y(x_0, y_0)} < 0.$$

Der Satz über implizite Funktionen gilt allgemein für Funktionen $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ in der Umgebung eines Punktes $\mathbf{a} = (a_1, \dots, a_n)^T$ mit $f(\mathbf{a}) = 0$ und $\frac{\partial f}{\partial x_n}(\mathbf{a}) \neq 0$. In diesem Fall gibt es eine offene Menge $U \subset \mathbb{R}^{n-1}$ mit $(a_1, \dots, a_{n-1})^T \in U$ und eine differenzierbare Funktion $g : U \rightarrow \mathbb{R}$ mit

$$f(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})) = 0 \quad \text{für alle } (x_1, \dots, x_{n-1}) \in U.$$

Für die partiellen Ableitungen von g gilt

$$\frac{\partial}{\partial x_i} g(x_1, \dots, x_{n-1}) = -\frac{f_{x_i}(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))}{f_{x_n}(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))}.$$

BEISPIEL VIII.3.8. Durch

$$f(x, y, z) = x \cos y + y \cos z + z \cos x - 2 = 0$$

ist implizit eine Funktion $z = g(x, y)$ erklärt. Es ist $g(0, 0) = 2$ und

$$\begin{aligned} g_x(0, 0) &= -\frac{f_x(0, 0, 2)}{f_z(0, 0, 2)} \\ &= -\frac{\cos y - z \sin x}{-y \sin z + \cos x} \Big|_{(0,0,2)} \\ &= -1, \\ g_y(0, 0) &= -\frac{f_y(0, 0, 2)}{f_z(0, 0, 2)} \\ &= -\frac{\cos z - x \sin y}{-y \sin z + \cos x} \Big|_{(0,0,2)} \end{aligned}$$

$$= -\cos 2.$$

VIII.3.6. Lokale Extrema. Wir betrachten eine Funktion $f : D \rightarrow \mathbb{R}$ in n Variablen.

Ein Punkt $\mathbf{a} \in D$ heißt ein **LOKALES MAXIMUM** (bzw. **LOKALES MINIMUM**) von f , wenn es ein $r > 0$ gibt, sodass für alle $\mathbf{x} \in B_r(\mathbf{a}) \cap D$ gilt

$$f(\mathbf{a}) \geq f(\mathbf{x}) \quad (\text{bzw. } f(\mathbf{a}) \leq f(\mathbf{x})).$$

Gilt die entsprechende Ungleichung für alle $\mathbf{x} \in D$, heißt \mathbf{a} ein **GLOBALES MAXIMUM** bzw. (**GLOBALES MINIMUM**). Ein **LOKALES EXTREMUM** ist ein lokales Minimum oder Maximum.

Eine Hauptaufgabe der angewandten Analysis ist die Bestimmung solcher lokalen Extrema. Je nachdem, ob sie im Innern von D oder auf dem Rand von D liegen, hat man unterschiedliche Charakterisierungen.

Wir beschäftigen uns in diesem Abschnitt mit lokalen Extrema im Inneren von D . Sei \mathbf{a} ein solches Extremum und $\mathbf{v} \in \mathbb{R}^n$ mit $|\mathbf{v}| = 1$ ein beliebiger Einheitsvektor. Dann hat die Funktion $\varphi(t) = f(\mathbf{a} + t\mathbf{v})$ in $t = 0$ ein lokales Extremum. Gemäß Abschnitt IV.2.1 (S. 140) muss also gelten

$$0 = \varphi'(0) = \text{grad } f(\mathbf{a}) \cdot \mathbf{v}.$$

Da \mathbf{v} ein beliebiger Einheitsvektor ist, muss also $\text{grad } f(\mathbf{a})$ zu allen Einheitsvektoren senkrecht sein. Das ist nur möglich, wenn $\text{grad } f(\mathbf{a}) = 0$ ist. Damit haben wir folgende **NOTWENDIGE CHARAKTERISIERUNG LOKALER EXTREMA IM INNEREN VON D** :

$$\mathbf{a} \text{ ist lokale Extremalstelle von } f \implies \text{grad } f(\mathbf{a}) = 0.$$

Punkte \mathbf{a} mit $\text{grad } f(\mathbf{a}) = 0$ nennt man auch **STATIONÄRE PUNKTE** von f . Es gilt also:

Jede lokale Extremalstelle im Innern von D ist ein stationärer Punkt.

MERKREGEL: Zur Bestimmung der lokalen Extrema sind zunächst alle stationären Punkte zu bestimmen.

Sei nun \mathbf{a} ein stationärer Punkt von f . Dann ist in der Nähe von \mathbf{a}

$$f(\mathbf{x}) = f(\mathbf{a}) + \frac{1}{2}(\mathbf{x} - \mathbf{a})^T H_f(\mathbf{a})(\mathbf{x} - \mathbf{a}) + o(|\mathbf{x} - \mathbf{a}|^2).$$

Daher ist \mathbf{a} genau dann ein Minimum bzw. Maximum von f , wenn es ein Minimum bzw. Maximum der Quadrik $\frac{1}{2}(\mathbf{x} - \mathbf{a})^T H_f(\mathbf{a})(\mathbf{x} - \mathbf{a})$ ist. Damit haben wir folgenden **EXTREMSTELLEN-TEST**:

\mathbf{a} sei ein stationärer Punkt der C^2 -Funktion f . Dann gilt:

- $H_f(\mathbf{a})$ ist positiv definit $\implies \mathbf{a}$ ist lokales Minimum.
- $H_f(\mathbf{a})$ ist negativ definit $\implies \mathbf{a}$ ist lokales Maximum.
- $H_f(\mathbf{a})$ ist indefinit $\implies \mathbf{a}$ ist kein lokales Extremum.

MERKREGEL: Berechne für jeden stationären Punkt die Hesse-Matrix und prüfe, ob sie positiv definit, negativ definit oder indefinit ist.

Im Falle von zwei Veränderlichen, d.h. $n = 2$, lautet der Extremwerttest:

\mathbf{a} sei stationärer Punkt von $f : \mathbb{R}^2 \supset D \rightarrow \mathbb{R}$, d.h. $f_x(\mathbf{a}) = f_y(\mathbf{a}) = 0$. Dann gilt

- $\det H_f(\mathbf{a}) > 0$ und $f_{xx}(\mathbf{a}) > 0 \implies \mathbf{a}$ ist lokales Minimum.
- $\det H_f(\mathbf{a}) > 0$ und $f_{xx}(\mathbf{a}) < 0 \implies \mathbf{a}$ ist lokales Maximum.
- $\det H_f(\mathbf{a}) < 0 \implies \mathbf{a}$ ist kein Extremum.

BEISPIEL VIII.3.9. Wir betrachten die Funktion

$$f(x, y) = 2x^4 + y^4 - 2x^2 - 2y^2.$$

Es ist

$$\text{grad } f(x, y) = \begin{pmatrix} 8x^3 - 4x \\ 4y^3 - 4y \end{pmatrix}.$$

Also sind die stationären Punkte von f bestimmt durch die Lösungen des nichtlinearen Gleichungssystems

$$4x(2x^2 - 1) = 0$$

$$4y(y^2 - 1) = 0.$$

Da ein Produkt $\alpha\beta$ genau dann Null ist, wenn mindestens ein Faktor gleich Null ist, hat f neun stationäre Punkte

$$\begin{array}{lll} \mathbf{a}_1 = (0, 0), & \mathbf{a}_2 = (0, 1), & \mathbf{a}_3 = (0, -1), \\ \mathbf{a}_4 = \left(\frac{1}{\sqrt{2}}, 0\right), & \mathbf{a}_5 = \left(\frac{1}{\sqrt{2}}, 1\right), & \mathbf{a}_6 = \left(\frac{1}{\sqrt{2}}, -1\right), \\ \mathbf{a}_7 = \left(-\frac{1}{\sqrt{2}}, 0\right), & \mathbf{a}_8 = \left(-\frac{1}{\sqrt{2}}, 1\right), & \mathbf{a}_9 = \left(-\frac{1}{\sqrt{2}}, -1\right). \end{array}$$

Für die Hesse-Matrix erhalten wir

$$H_f(x, y) = \begin{pmatrix} 24x^2 - 4 & 0 \\ 0 & 12y^2 - 4 \end{pmatrix},$$

$$\det H_f(x, y) = 16(6x^2 - 1)(3y^2 - 1).$$

Dies ergibt folgende Tabelle

	\mathbf{a}_1	\mathbf{a}_2	\mathbf{a}_3	\mathbf{a}_4	\mathbf{a}_5	\mathbf{a}_6	\mathbf{a}_7	\mathbf{a}_8	\mathbf{a}_9
$f_{xx}(\mathbf{a}_i)$	-4	-4	-4	8	8	8	8	8	8
$\det H_f(\mathbf{a}_i)$	16	-32	-32	-32	64	64	-32	64	64

Also hat f bei \mathbf{a}_1 ein lokales Maximum und bei \mathbf{a}_5 , \mathbf{a}_6 , \mathbf{a}_8 und \mathbf{a}_9 ein lokales Minimum; \mathbf{a}_2 , \mathbf{a}_3 , \mathbf{a}_4 und \mathbf{a}_7 sind keine Extrema.

VIII.3.7. Extrema unter Nebenbedingungen. Wir betrachten nun Extrempunkte auf dem Rand des Definitionsbereiches D der Funktion f . Häufig ist der Rand lokal als Schnitt von Niveauflächen gegeben. Diesen Fall wollen wir untersuchen. Dazu betrachten wir der Einfachheit halber zunächst den Fall nur einer Niveaufläche. Dazu seien f und g zwei genügend oft differenzierbare Funktionen in n Veränderlichen. Wir suchen die Extremwerte von f auf der Niveaufläche

$$M = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}.$$

1. METHODE: AUFLÖSEN DER NEBENBEDINGUNG. Man löst die Bedingung $g(\mathbf{x}) = 0$ nach einer Variablen auf, d.h. man stellt M als Graph einer Funktion in $n - 1$ Variablen dar. Lösen wir z.B. nach der Variablen x_n auf, ist M von der Form $\{(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1}))\}$ mit einer geeigneten Funktion h . Wir können diese Darstellung in f einsetzen und erhalten auf diese Weise eine Funktion in $n - 1$ Variablen

$$F(x_1, \dots, x_{n-1}) = f(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1})).$$

Die Extrempunkte von F können wir nun wie im vorigen Paragraphen bestimmen.

BEISPIEL VIII.3.10. Wir wollen die Extrema von

$$f(x, y) = e^{xy}$$

auf dem Kreisrand

$$x^2 + y^2 = 1$$

bestimmen, d.h.

$$g(x, y) = x^2 + y^2 - 1.$$

Wegen $f(-x, -y) = f(x, y)$ können wir uns auf den oberen Halbkreis beschränken. Dort ist

$$y = \sqrt{1 - x^2} \quad \text{mit} \quad -1 \leq x \leq 1.$$

Setzen wir dies in f ein, müssen wir die Extrema von

$$F(x) = e^{x\sqrt{1-x^2}}$$

auf $[-1, 1]$ bestimmen. Kandidaten sind zunächst die Randpunkte $x = \pm 1$; es ist

$$F(\pm 1) = 1.$$

Weitere Kandidaten sind die stationären Punkte von F . Wegen

$$\begin{aligned} F'(x) &= e^{x\sqrt{1-x^2}} \left\{ \sqrt{1-x^2} - x^2 \frac{1}{\sqrt{1-x^2}} \right\} \\ &= \frac{1}{\sqrt{1-x^2}} e^{x\sqrt{1-x^2}} (1-2x^2) \end{aligned}$$

sind dies die Punkte $\pm \frac{1}{\sqrt{2}}$. Wegen

$$\begin{aligned} F\left(\frac{1}{\sqrt{2}}\right) &= e^{\frac{1}{2}}, \\ F\left(-\frac{1}{\sqrt{2}}\right) &= e^{-\frac{1}{2}} \end{aligned}$$

ist $\frac{1}{\sqrt{2}}$ ein Maximum und $-\frac{1}{\sqrt{2}}$ ein Minimum. Wegen der Symmetrie $f(-x, -y) = f(x, y)$ folgt insgesamt:

f hat Maxima in $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ und $(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.

f hat Minima in $(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ und $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.

2. METHODE: PARAMETRISIERUNG DER NEBENBEDINGUNG. Betrachte als Beispiel den Spezialfall $n = 2$. Falls wir eine Parameterdarstellung $\mathbf{x}(t)$, $t \in I$, der Kurve M kennen, können wir diese in f einsetzen. Wir erhalten so eine Funktion einer Veränderlichen

$$\varphi(t) = f(\mathbf{x}(t)),$$

deren Extrema wir in Abschnitt IV.2.1 (S. 140) bestimmen können.

BEISPIEL VIII.3.11. In Beispiel VIII.3.10 ist

$$\mathbf{x}(t) = (\cos t, \sin t), \quad 0 \leq t \leq 2\pi,$$

eine Parameterdarstellung. Wir erhalten

$$\varphi(t) = e^{\cos t \sin t} = e^{\frac{1}{2} \sin(2t)}.$$

Wegen

$$\varphi'(t) = \cos(2t) e^{\frac{1}{2} \sin(2t)}$$

sind die stationären Punkte von φ gegeben durch $t = \frac{\pi}{4}$, $t = \frac{3\pi}{4}$, $t = \frac{5\pi}{4}$ und $t = \frac{7\pi}{4}$, wobei $\frac{\pi}{4}$ und $\frac{5\pi}{4}$ Maxima und $\frac{3\pi}{4}$ und $\frac{7\pi}{4}$ Minima sind.

3. METHODE: LAGRANGE-MULTIPLIKATOREN. Dieser Methode liegt folgender Satz zugrunde:

Sei $\mathbf{a} \in \mathbb{R}^n$ eine Extremalstelle von f unter der Nebenbedingung $g(\mathbf{x}) = 0$. Dann gibt es eine Zahl λ , genannt LAGRANGE-MULTIPLIKATOR, so dass (\mathbf{a}, λ) ein stationärer Punkt der LAGRANGE-FUNKTION

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$$

ist, d.h.

$$\text{grad } f(\mathbf{a}) + \lambda \text{grad } g(\mathbf{a}) = 0$$

und

$$g(\mathbf{a}) = 0.$$

MERKREGEL: Stelle die Lagrange Funktion

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$$

auf und bestimme alle ihre stationären Punkte.

BEISPIEL VIII.3.12. Die Lagrange-Funktion zu Beispiel [VIII.3.10](#) lautet

$$L(\mathbf{x}, \lambda) = e^{xy} + \lambda(x^2 + y^2 - 1).$$

Die stationären Punkte sind die Lösungen des nichtlinearen Gleichungssystems

$$\begin{aligned} L_x(x, y, \lambda) = 0 &\iff ye^{xy} + 2\lambda x = 0 \\ L_y(x, y, \lambda) = 0 &\iff xe^{xy} + 2\lambda y = 0 \\ L_\lambda(x, y, \lambda) = 0 &\iff x^2 + y^2 - 1 = 0. \end{aligned}$$

Wir multiplizieren die erste Gleichung mit x und die zweite Gleichung mit y , addieren die beiden Ergebnisse und nützen die dritte Gleichung aus. So erhalten wir

$$\begin{aligned} 0 &= xye^{xy} + 2\lambda x^2 + xye^{xy} + 2\lambda y^2 \\ &= 2xye^{xy} + 2\lambda \underbrace{(x^2 + y^2)}_{=1} \\ &= 2xye^{xy} + 2\lambda. \end{aligned}$$

Also ist

$$\lambda = -xye^{xy}.$$

Setzen wir dies in die ersten beiden Gleichungen ein, erhalten wir

$$\begin{aligned} 0 &= ye^{xy} - 2x^2ye^{xy} = y(1 - 2x^2)e^{xy} \\ 0 &= xe^{xy} - 2xy^2e^{xy} = x(1 - 2y^2)e^{xy}. \end{aligned}$$

Die Lösungen dieses Gleichungssystems sind

$$\begin{aligned} \mathbf{a}_1 &= (0, 0), \\ \mathbf{a}_2 &= \left(0, \frac{1}{\sqrt{2}}\right), & \mathbf{a}_3 &= \left(0, -\frac{1}{\sqrt{2}}\right), \\ \mathbf{a}_4 &= \left(\frac{1}{\sqrt{2}}, 0\right), & \mathbf{a}_5 &= \left(-\frac{1}{\sqrt{2}}, 0\right), \\ \mathbf{a}_6 &= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), & \mathbf{a}_7 &= \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \\ \mathbf{a}_8 &= \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), & \mathbf{a}_9 &= \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right). \end{aligned}$$

Da die Punkte $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ und \mathbf{a}_5 die Nebenbedingung nicht erfüllen, bleiben nur die Kandidaten $\mathbf{a}_6, \mathbf{a}_7, \mathbf{a}_8$ und \mathbf{a}_9 . Durch Einsetzen in f und Vergleichen der Funktionswerte sehen wir, dass \mathbf{a}_6 und \mathbf{a}_7 Maxima und \mathbf{a}_8 und \mathbf{a}_9 Minima sind.

Wir betrachten nun den allgemeinen Fall des Durchschnittes mehrerer Niveauflächen

$$g_1(\mathbf{x}) = 0, \dots, g_k(\mathbf{x}) = 0.$$

In diesem Fall stellt man eine Lagrange-Funktion auf und bestimmt deren stationäre Punkte. Die Lagrange-Funktion hat jetzt die Form

$$L(\mathbf{x}, \lambda_1, \dots, \lambda_k) = f(\mathbf{x}) + \sum_{i=1}^k \lambda_i g_i(\mathbf{x}).$$

BEISPIEL VIII.3.13. Sei S die Kugeloberfläche

$$x^2 + y^2 + z^2 = 1$$

und E die Ebene

$$x + y + z = 0.$$

Wir suchen die Extrema von

$$f(x, y, z) = e^{xyz}$$

auf $S \cap E$. Die Lagrange-Funktion lautet mit λ, μ an Stelle von λ_1, λ_2 :

$$L(x, y, z, \lambda, \mu) = e^{xyz} + \lambda(x^2 + y^2 + z^2 - 1) + \mu(x + y + z).$$

Die Bedingungen für einen stationären Punkt sind

$$\begin{aligned} yze^{xyz} + 2\lambda x + \mu &= 0 \\ xze^{xyz} + 2\lambda y + \mu &= 0 \\ xye^{xyz} + 2\lambda z + \mu &= 0 \\ x^2 + y^2 + z^2 - 1 &= 0 \\ x + y + z &= 0. \end{aligned}$$

Multiplizieren wir die erste Gleichung mit x , die zweite mit y und die dritte mit z , addieren die Ergebnisse und beachten die vierte und fünfte Gleichung, erhalten wir

$$\begin{aligned} 0 &= 3xyze^{xyz} + 2\lambda \underbrace{(x^2 + y^2 + z^2)}_{=1} + \mu \underbrace{(x + y + z)}_{=0} \\ &= 3xyze^{xyz} + 2\lambda \end{aligned}$$

also

$$\lambda = -\frac{3}{2}xyze^{xyz}.$$

Setzen wir dies in die ersten drei Gleichungen ein, erhalten wir

$$\begin{aligned} e^{xyz}yz(1 - 3x^2) + \mu &= 0 \\ (**) \quad e^{xyz}xz(1 - 3y^2) + \mu &= 0 \\ e^{xyz}xy(1 - 3z^2) + \mu &= 0. \end{aligned}$$

Subtrahieren wir die erste von der zweiten und dritten Gleichung erhalten wir

$$\begin{aligned} e^{xyz}z(y - x)(1 + 3xy) &= 0 \\ e^{xyz}y(z - x)(1 + 3xz) &= 0. \end{aligned}$$

Man beachte, dass die zweite Gleichung aus der ersten durch Vertauschen von y und z hervorgeht. Analog erhält man durch Vertauschen von x und z die Gleichung

$$e^{xyz}x(y - z)(1 + 3yz) = 0.$$

Wir betrachten zuerst den Fall $x = 0$. Einsetzen in die Gleichungen (***) liefert $\mu = 0$ und $yz = 0$. Also ist auch $y = 0$ oder $z = 0$. Dies ist aber nicht mit den beiden Nebenbedingungen vereinbar. Analog führen die Annahmen $y = 0$ und $z = 0$ zum Widerspruch.

Als nächstes betrachten wir den Fall $x = y$. Diese Annahme führt wegen der beiden Nebenbedingungen auf die Lösungen

$$\mathbf{a}_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad \mathbf{a}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Analog erhält man für die Fälle $y = z$ und $x = z$ die Lösungen

$$\begin{aligned} \mathbf{a}_3 &= \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, & \mathbf{a}_4 &= \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \\ \mathbf{a}_5 &= \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, & \mathbf{a}_6 &= \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}. \end{aligned}$$

Schließlich müssen wir noch den Fall $1 + 3xy = 0$ betrachten. Da offensichtlich $x \neq 0$ sein muss, können wir diese Gleichung nach y auflösen.

Setzen wir das Ergebnis in die erste Nebenbedingung ein und berücksichtigen die zweite Nebenbedingung, erhalten wir

$$\begin{aligned} 1 &= x^2 + \underbrace{y^2}_{=\frac{1}{9x^2}} + \underbrace{z^2}_{=-x-\underbrace{y}_{=-\frac{1}{3x}}} \\ &= x^2 + \frac{1}{9x^2} + \left(\frac{1-3x^2}{3x}\right)^2 \\ &= \frac{1}{9x^2}(18x^4 - 6x^2 + 2) \end{aligned}$$

also

$$0 = 18x^4 - 15x^2 + 2$$

mit den Lösungen

$$x = \pm \frac{1}{\sqrt{6}}$$

und

$$x = \pm \frac{2}{\sqrt{6}}.$$

Dies sind die bereits bekannten Lösungen $\mathbf{a}_3, \dots, \mathbf{a}_6$. Analog sieht man ein, dass die Fälle $1 + 3yz = 0$ und $1 + 3xz = 0$ zusätzlich nur noch die bereits bekannten Lösungen \mathbf{a}_1 und \mathbf{a}_2 liefern.

Wegen

$$f(\mathbf{a}_1) = f(\mathbf{a}_3) = f(\mathbf{a}_5) = e^{-\frac{1}{3\sqrt{6}}},$$

$$f(\mathbf{a}_2) = f(\mathbf{a}_4) = f(\mathbf{a}_6) = e^{\frac{1}{3\sqrt{6}}}$$

sind daher $\mathbf{a}_1, \mathbf{a}_3$ und \mathbf{a}_5 Minima und $\mathbf{a}_2, \mathbf{a}_4$ und \mathbf{a}_6 Maxima.

VIII.3.8. Extrema auf kompakten Mengen. Jede stetige Funktion $f : D \rightarrow \mathbb{R}$ auf einer *kompakten* Menge $D \subset \mathbb{R}^n$ nimmt ihr Maximum und Minimum an (vgl. Abschnitt VIII.2.3 (S. 285)). Die Extremstellen bestimmt man durch eine Kombination der Methoden der beiden vorigen Paragraphen.

BEISPIEL VIII.3.14. Gesucht sind die Extrema von

$$f(x, y) = 3x^2 - 2xy + y^2$$

auf der Kreisscheibe

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Wir suchen zunächst Extremwerte im Innern von D und gehen dazu wie in Abschnitt VIII.3.6 vor. Die Bedingungen für einen stationären Punkt sind

$$6x - 2y = 0$$

$$-2x + 2y = 0.$$

Wegen

$$\det \begin{pmatrix} 6 & -2 \\ -2 & 2 \end{pmatrix} = 8$$

hat dieses LGS nur die Lösung $x = 0, y = 0$. Da

$$H_f(0, 0) = \begin{pmatrix} 6 & -2 \\ -2 & 2 \end{pmatrix}$$

positiv definit ist, ist dieses ein lokales Minimum; der zugehörige Funktionswert ist 0. Wir wissen aber noch nicht, ob es sich um ein globales Minimum handelt!

Nun betrachten wir den Rand von D und gehen dazu wie in Abschnitt VIII.3.7 vor. Die Lagrange-Funktion lautet

$$L(x, y, \lambda) = 3x^2 - 2xy + y^2 + \lambda(x^2 + y^2 - 1).$$

Die Bedingungen für einen stationären Punkt sind

$$\begin{aligned} 6x - 2y + 2\lambda x &= 0 & \text{d.h.} & & -3x + y &= \lambda x \\ -2x + 2y + 2\lambda y &= 0 & \text{d.h.} & & x - y &= \lambda y \\ x^2 + y^2 - 1 &= 0. \end{aligned}$$

Also müssen wir die Eigenwerte λ der Matrix

$$\begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$$

und die zugehörigen normierten Eigenvektoren bestimmen.

Das charakteristische Polynom lautet:

$$\begin{aligned} \det \begin{pmatrix} -3 - z & 1 \\ 1 & -1 - z \end{pmatrix} &= (3 + z)(1 + z) - 1 \\ &= 2 + 4z + z^2 \\ &= (z + 2)^2 - 2 \end{aligned}$$

und hat die Nullstellen $-2 \pm \sqrt{2}$. Die Gleichungssysteme für die Eigenvektoren lauten für $\lambda = -2 + \sqrt{2}$

$$\begin{aligned} (-1 - \sqrt{2})x + y &= 0 & \implies y &= (1 + \sqrt{2})x \\ x + (1 - \sqrt{2})y &= 0 & \implies x &= \pm \frac{1}{\sqrt{4 + 2\sqrt{2}}} \end{aligned}$$

und für $\lambda = -2 - \sqrt{2}$

$$\begin{aligned} (-1 + \sqrt{2})x + y &= 0 & \implies y &= (1 - \sqrt{2})x \\ x + (1 + \sqrt{2})y &= 0 & \implies x &= \pm \frac{1}{\sqrt{4 - 2\sqrt{2}}}. \end{aligned}$$

Also sind unsere Kandidaten für die Extremalstellen

$$\begin{aligned} \mathbf{a}_1 &= \frac{1}{\sqrt{4+2\sqrt{2}}} \begin{pmatrix} 1 \\ 1+\sqrt{2} \end{pmatrix}, & \mathbf{a}_2 &= -\mathbf{a}_1 \\ \mathbf{a}_3 &= \frac{1}{\sqrt{4-2\sqrt{2}}} \begin{pmatrix} 1 \\ 1-\sqrt{2} \end{pmatrix}, & \mathbf{a}_4 &= -\mathbf{a}_3. \end{aligned}$$

Die zugehörigen Funktionswerte sind

$$\begin{aligned} f(\mathbf{a}_2) &= f(\mathbf{a}_1) = \frac{4}{4+2\sqrt{2}} \\ f(\mathbf{a}_4) &= f(\mathbf{a}_3) = \frac{4}{4-2\sqrt{2}}. \end{aligned}$$

Also sind \mathbf{a}_3 und \mathbf{a}_4 globale Maxima und $(0,0)$ das globale Minimum von f .

VIII.4. Vektorwertige Funktionen

VIII.4.1. Die Differentiation. Wir betrachten Funktionen \mathbf{f} , die Punkten $\mathbf{x} \in D \subset \mathbb{R}^n$ Vektoren $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$ zuordnen:

$$\begin{aligned} \mathbf{f} : \mathbb{R}^n \supset D &\rightarrow \mathbb{R}^m \\ \mathbf{f}(\mathbf{x}) &= \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} \\ &= \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}. \end{aligned}$$

Die Begriffe „Grenzwert“, „Stetigkeit“, „partielle Ableitung“ sind für solche vektorwertigen Funktionen komponentenweise definiert:

- GRENZWERT:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \begin{pmatrix} \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_1(\mathbf{x}) \\ \vdots \\ \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_m(\mathbf{x}) \end{pmatrix}.$$

- STETIGKEIT: \mathbf{f} ist stetig in $\mathbf{x}_0 \iff$ alle Komponentenfunktionen f_1, \dots, f_m sind stetig in \mathbf{x}_0 .
- PARTIELLE ABLEITUNGEN:

$$\frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_i}(\mathbf{x}_0) \end{pmatrix} \quad 1 \leq i \leq n.$$

Die vektorwertige Funktion $\mathbf{f} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$ heißt im Punkt $\mathbf{x}_0 \in D$ TOTAL DIFFERENZIERBAR oder LINEAR APPROXIMIERBAR, wenn es eine $m \times n$ Matrix A gibt mit

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{1}{|\mathbf{x} - \mathbf{x}_0|} \{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - A(\mathbf{x} - \mathbf{x}_0)\} = 0.$$

\mathbf{f} ist genau dann linear approximierbar, wenn alle Komponentenfunktionen f_1, \dots, f_m linear approximierbar sind. In diesem Fall ist die obige Matrix A eindeutig bestimmt. Sie heißt JACOBI-MATRIX von \mathbf{f} an der Stelle \mathbf{x}_0 und wird mit $D\mathbf{f}(\mathbf{x}_0)$ bezeichnet. Ihre Zeilenvektoren sind die transponierten Gradienten der Komponentenfunktionen f_1, \dots, f_m :

$$\begin{aligned} D\mathbf{f}(\mathbf{x}_0) &= \begin{pmatrix} (\text{grad } f_1(\mathbf{x}_0))^T \\ \vdots \\ (\text{grad } f_m(\mathbf{x}_0))^T \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}. \end{aligned}$$

BEISPIEL VIII.4.1. Jede lineare Abbildung $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$ und $A \in \mathbb{R}^{m \times n}$ ist total differenzierbar. Für alle \mathbf{x} ist $D\mathbf{f}(\mathbf{x}) = A$.

BEISPIEL VIII.4.2 (POLARKOORDINATEN). Der Streifen

$$D = \{(r, \varphi) : r \geq 0, 0 \leq \varphi < 2\pi\}$$

der kartesischen (r, φ) -Ebene wird mittels

$$\mathbf{f}(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

auf die kartesische (x, y) -Ebene abgebildet, sodass zu jedem Punkt $(x, y) \neq (0, 0)$ genau ein Punkt $(r, \varphi) \in D$ gehört. Die Jacobi-Matrix dieser Abbildung ist

$$D\mathbf{f}(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Für die Jacobi-Matrix gelten folgende Rechenregeln ($\mathbf{v}, \mathbf{w} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\alpha, \beta \in \mathbb{R}$):

$$\begin{aligned} D(\alpha\mathbf{v} + \beta\mathbf{w})(\mathbf{x}) &= \alpha D\mathbf{v}(\mathbf{x}) + \beta D\mathbf{w}(\mathbf{x}) \\ D(f\mathbf{v})(\mathbf{x}) &= f(\mathbf{x})D\mathbf{v}(\mathbf{x}) + \mathbf{v}(\mathbf{x})(\text{grad } f(\mathbf{x}))^T \\ D(\mathbf{v} \cdot \mathbf{w})(\mathbf{x}) &= \mathbf{v}(\mathbf{x})^T D\mathbf{w}(\mathbf{x}) + \mathbf{w}(\mathbf{x})^T D\mathbf{v}(\mathbf{x}) \\ D(\mathbf{v} \times \mathbf{w})(\mathbf{x}) &= \mathbf{v}(\mathbf{x}) \times D\mathbf{w}(\mathbf{x}) - \mathbf{w}(\mathbf{x}) \times D\mathbf{v}(\mathbf{x}) \text{ falls } m = 3. \end{aligned}$$

ACHTUNG: Der Ausdruck $\mathbf{v}(\mathbf{x})(\text{grad } f(\mathbf{x}))^T$ in der zweiten Formel ist kein Skalarprodukt, sondern eine $m \times n$ Matrix. Die Vektorprodukte auf der rechten Seite der vierten Gleichung sind spaltenweise zu nehmen.

BEISPIEL VIII.4.3 (QUASILINEARE ABBILDUNGEN). Wir betrachten eine skalare Funktion $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ und eine $m \times n$ Matrix A . Mit ihnen definieren wir eine vektorwertige Funktion $\mathbf{f} : D \rightarrow \mathbb{R}^m$ durch $\mathbf{f}(\mathbf{x}) = f(\mathbf{x})A\mathbf{x}$. Eine solche Funktion nennt man QUASILINEAR. Für die Jacobi-Matrix ergibt sich

$$D\mathbf{f}(\mathbf{x}) = f(\mathbf{x})A + (A\mathbf{x})(\text{grad } f(\mathbf{x}))^T.$$

VIII.4.2. Das Newtonverfahren. Das Newtonverfahren mit seinen Modifikationen ist das wichtigste Verfahren zur numerischen Lösung nichtlinearer Gleichungssysteme mit n Gleichungen und n Unbekannten. Ein solches Gleichungssystem kann stets in die Form gebracht werden

$$(*) \quad \mathbf{f}(\mathbf{x}) = 0$$

mit einer Funktion $\mathbf{f} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^n$. Die Idee des Verfahrens ist die gleiche wie bei Gleichungen in einer Unbekannten (vgl. Abschnitt IV.2.6 (S. 150)):

Wir nehmen an, dass wir eine Näherungslösung \mathbf{x}_0 für (*) „geraten“ haben. Falls \mathbf{f} differenzierbar ist, wird \mathbf{f} in der Nähe von \mathbf{x}_0 durch die lineare Funktion $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ approximiert. Wir ersetzen \mathbf{f} in (*) durch diese Approximation und erhalten das lineare Gleichungssystem

$$\mathbf{f}(\mathbf{x}_0) + D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = 0$$

für den unbekanntem Vektor \mathbf{x} . Es ist genau dann lösbar, wenn die Jacobi-Matrix $D\mathbf{f}(\mathbf{x}_0)$ invertierbar ist. In diesem Fall lautet die Lösung des linearen Gleichungssystems

$$\mathbf{x}_1 = \mathbf{x}_0 - D\mathbf{f}(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_0).$$

Wir nehmen \mathbf{x}_1 als neue, hoffentlich bessere Näherung für die Lösung von (*) und wiederholen das Verfahren mit \mathbf{x}_1 an Stelle von \mathbf{x}_0 .

BEMERKUNG VIII.4.4. Selbstverständlich wird bei der Berechnung \mathbf{x}_1 die Matrix $D\mathbf{f}(\mathbf{x}_0)$ nicht invertiert. Stattdessen wird das lineare Gleichungssystem

$$D\mathbf{f}(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0) = -\mathbf{f}(\mathbf{x}_0)$$

mit dem Gaußschen Eliminationsverfahren oder einem seiner Verwandten gelöst (vgl. Abschnitte II.1.4 (S. 57) und II.2.5 (S. 68)).

Zusammenfassend lautet das NEWTONVERFAHREN:

- (1) Gegeben sei eine Näherungslösung \mathbf{x}_0 für das Gleichungssystem

$$\mathbf{f}(\mathbf{x}) = 0.$$

- (2) Für $k = 0, 1, \dots$ führe folgende Schritte aus:
- (a) Bestimme die Jacobi-Matrix $D\mathbf{f}(\mathbf{x}_k)$.
 - (b) Falls $D\mathbf{f}(\mathbf{x}_k)$ singularär ist, breche das Verfahren ab.
 - (c) Andernfalls berechne die Lösung \mathbf{z} des linearen Gleichungssystems

$$D\mathbf{f}(\mathbf{x}_k)\mathbf{z} = -\mathbf{f}(\mathbf{x}_k).$$

- (d) Setze

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{z}.$$

Wie in einer Dimension, $n = 1$, konvergiert das Newton-Verfahren nicht für jeden Startwert. Falls aber \mathbf{x}_0 hinreichend nahe bei einer Lösung \mathbf{x}^* des Gleichungssystems liegt, \mathbf{f} zweimal stetig differenzierbar ist und $D\mathbf{f}(\mathbf{x}^*)$ invertierbar ist, kann man zeigen, dass das Newtonverfahren quadratisch konvergiert. D.h., es gibt eine Konstante $c > 0$ mit

$$|\mathbf{x}^* - \mathbf{x}_{k+1}| \leq c|\mathbf{x}^* - \mathbf{x}_k|^2$$

für alle k .

BEISPIEL VIII.4.5. Zur Darstellung der Zahnflanken von Stirnradgetrieben verwendet man die Kreisevolvente γ mit der Parameterdarstellung

$$\mathbf{x}(t) = \begin{pmatrix} r \sin t - rt \cos t \\ r \cos t + rt \sin t - r \end{pmatrix}.$$

Soll γ durch die Punkte $(0, 0)$ und (a, b) gehen, führt dies auf die beiden folgenden nichtlinearen Gleichungen für r und t :

$$\begin{aligned} f_1(r, t) &= r \sin t - rt \cos t - a &= 0 \\ f_2(r, t) &= r \cos t + rt \sin t - r - b &= 0. \end{aligned}$$

Die Jacobi-Matrix von \mathbf{f} im Punkt (r_0, t_0) lautet

$$\begin{pmatrix} \sin t_0 - t_0 \cos t_0 & r_0 t_0 \sin t_0 \\ \cos t_0 + t_0 \sin t_0 - 1 & r_0 t_0 \cos t_0 \end{pmatrix}.$$

Die Determinante der Matrix ist $r_0 t_0 (\sin t_0 - t_0)$. Sie ist ungleich Null sofern $r_0 t_0 \neq 0$ ist.

Für $(a, b) = (1, 1)$ und die Startwerte $(r_0, t_0) = (2, 1.2)$ liefert das Newtonverfahren dann z.B. folgende Werte:

$$r_0 = 2 \qquad t_0 = 1.2$$

$$\begin{array}{ll} r_1 = 2.12598 & t_0 = 1.17449 \\ r_2 = 2.12891 & t_0 = 1.17504 \\ r_3 = 2.12891 & t_0 = 1.17504. \end{array}$$

VIII.4.3. Die Kettenregel. Betrachte zwei offene Mengen $D \subset \mathbb{R}^n$ und $G \subset \mathbb{R}^m$ und zwei Funktionen $\mathbf{f} : D \rightarrow \mathbb{R}^m$ und $\mathbf{g} : G \rightarrow \mathbb{R}^k$ derart, dass $\mathbf{f}(D) \subset G$ ist. Dann ist die Komposition $\mathbf{g} \circ \mathbf{f} : D \rightarrow \mathbb{R}^k$ wohldefiniert. Sind \mathbf{f} und \mathbf{g} differenzierbar, ist auch $\mathbf{g} \circ \mathbf{f}$ differenzierbar und es gilt die KETTENREGEL:

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}_0) = D\mathbf{g}(\mathbf{f}(\mathbf{x}_0))D\mathbf{f}(\mathbf{x}_0).$$

BEISPIEL VIII.4.6. Sei $\mathbf{f} : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ wie in Beispiel VIII.4.2 und $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definiert durch

$$\mathbf{g}(x, y) = \begin{pmatrix} x^2 y^3 \\ e^{x+y} \\ \sin(xy) \end{pmatrix}.$$

Es ist

$$D\mathbf{g}(x, y) = \begin{pmatrix} 2xy^3 & 3x^2y^2 \\ e^{x+y} & e^{x+y} \\ y \cos(xy) & x \cos(xy) \end{pmatrix}.$$

Damit liefert die Kettenregel

$$\begin{aligned} & D(\mathbf{g} \circ \mathbf{f})(r, \varphi) \\ &= \begin{pmatrix} 2r^4 \cos \varphi \sin^3 \varphi & 3r^4 \cos^2 \varphi \sin^2 \varphi \\ e^{r(\cos \varphi + \sin \varphi)} & e^{r(\cos \varphi + \sin \varphi)} \\ r \sin \varphi \cos(r^2 \sin \varphi \cos \varphi) & r \cos \varphi \cos(r^2 \sin \varphi \cos \varphi) \end{pmatrix} \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} \\ &= \begin{pmatrix} 5r^4 \cos^2 \varphi \sin^3 \varphi & r^5 [-2 \cos \varphi \sin^4 \varphi + 3 \cos^3 \varphi \sin^2 \varphi] \\ e^{r(\cos \varphi + \sin \varphi)} (\cos \varphi + \sin \varphi) & r e^{r(\cos \varphi + \sin \varphi)} (\cos \varphi - \sin \varphi) \\ 2r \sin \varphi \cos \varphi \cos(r^2 \sin \varphi \cos \varphi) & r^2 (\cos^2 \varphi - \sin^2 \varphi) \cos(r^2 \sin \varphi \cos \varphi) \end{pmatrix}. \end{aligned}$$

BEISPIEL VIII.4.7 (BASISWECHSEL). Wir betrachten eine Funktion $\mathbf{f} : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$. Im \mathbb{R}^n führen wir ein neues kartesisches Koordinatensystem $(P; \mathbf{b}_1, \dots, \mathbf{b}_n)$ ein und wählen im \mathbb{R}^m eine neue Basis $(\mathbf{w}_1, \dots, \mathbf{w}_m)$. Durch diesen doppelten Basiswechsel im Urbild- und Bildraum geht \mathbf{f} in eine andere Funktion \mathbf{g} über. Wir wollen diese Funktion bestimmen und ihre Jacobi-Matrix $D\mathbf{g}$ durch die Jacobi-Matrix $D\mathbf{f}$ von \mathbf{f} ausdrücken.

Einem Punkt $\mathbf{x} \in \mathbb{R}^n$ entspricht im neuen Koordinatensystem der Vektor \mathbf{y} , der durch $\mathbf{x} = B\mathbf{y} + \mathbf{p}$ bestimmt ist, wobei B die orthogonale Matrix $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ und $\mathbf{p} = \overrightarrow{OP}$ ist. Ein Vektor $\mathbf{v} \in \mathbb{R}^m$ hat bezüglich der Basis $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ den Koordinatenvektor \mathbf{w} , der durch $\mathbf{v} = W\mathbf{w}$ mit der invertierbaren Matrix $W = (\mathbf{w}_1, \dots, \mathbf{w}_m)$ bestimmt ist. Setzt man $\mathbf{v} = \mathbf{f}(\mathbf{x})$ und $\mathbf{x} = B\mathbf{y} + \mathbf{p}$, so ist $\mathbf{g}(\mathbf{y})$ implizit definiert durch

$$\mathbf{v} = \mathbf{f}(\mathbf{x}) = \mathbf{f}(B\mathbf{y} + \mathbf{p}) = W\mathbf{w} = W\mathbf{g}(\mathbf{y}).$$

Lösen wir diese Gleichung nach $\mathbf{g}(\mathbf{y})$ auf, erhalten wir

$$\mathbf{g}(\mathbf{y}) = W^{-1}\mathbf{f}(B\mathbf{y} + \mathbf{p}).$$

Damit folgt aus der Kettenregel

$$D\mathbf{g}(\mathbf{y}_0) = W^{-1}D\mathbf{f}(\mathbf{x}_0)B \quad \text{mit } \mathbf{x}_0 = B\mathbf{y}_0 + \mathbf{p}.$$

Mit diesen beiden Identitäten lassen sich leicht Eigenschaften einer Funktion untersuchen, die unabhängig sind von der speziellen Wahl des Koordinatensystems im Urbild- und Bildraum.

VIII.4.4. Räumliche Skalaren- und Vektorfelder. Der dreidimensionale Anschauungsraum wird nach Wahl eines kartesischen Koordinatensystem $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ durch den \mathbb{R}^3 dargestellt. Dementsprechend wird jeder räumliche Bereich durch eine Menge $D \subset \mathbb{R}^3$ dargestellt. Ein SKALARENFELD (oder Belegungsfunktion) $f: \mathbb{R}^3 \supset D \rightarrow \mathbb{R}$ ordnet jedem Punkt $\mathbf{x} \in D$ einen Zahlenwert zu. Ein (räumliches) VEKTORFELD $\mathbf{v}: \mathbb{R}^3 \supset D \rightarrow \mathbb{R}^3$ ordnet jedem Punkt $\mathbf{x} \in D$ einen Vektor $\mathbf{v}(\mathbf{x})$ zu. Man stellt sich dabei vor, dass jedem Punkt \mathbf{x} der entsprechende Vektor $\mathbf{v}(\mathbf{x})$ angeheftet ist. Eine Kurve in D heißt FELDLINIE des Vektorfeldes \mathbf{v} , wenn der Vektor $\mathbf{v}(\mathbf{x})$ in jedem Kurvenpunkt \mathbf{x} parallel zur Kurventangente ist. Man spricht von einem C^r -Skalarenfeld f bzw. C^r -Vektorfeld \mathbf{v} , wenn f bzw. \mathbf{v} r -mal stetig differenzierbar sind.

BEISPIEL VIII.4.8 (STARRE DREHUNG). Eine Rechtsdrehung mit konstanter Winkelgeschwindigkeit ω um eine Achse durch den Nullpunkt mit Richtungsvektor $\mathbf{a} = (a_1, a_2, a_3)^T$ mit $|\mathbf{a}| = 1$ wird dargestellt durch

$$\mathbf{x}(t) = \cos(\omega t)\mathbf{x}_0 + (1 - \cos(\omega t))(\mathbf{x}_0 \cdot \mathbf{a})\mathbf{a} + \sin(\omega t)\mathbf{a} \times \mathbf{x}_0.$$

Dabei bezeichnet \mathbf{x}_0 die Lage zur Zeit $t = 0$. Für die Geschwindigkeit $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$ ergibt sich wegen

$$\begin{aligned} \mathbf{a} \times (\mathbf{a} \times \mathbf{x}_0) &= (\mathbf{a} \cdot \mathbf{x}_0)\mathbf{a} - (\mathbf{a} \cdot \mathbf{a})\mathbf{x}_0 \\ &= (\mathbf{a} \cdot \mathbf{x}_0)\mathbf{a} - \mathbf{x}_0 \end{aligned}$$

und $\mathbf{a} \times \mathbf{a} = 0$ (vgl. Abschnitt 1.4.7 (S. 31)) die Beziehung

$$\begin{aligned} \mathbf{v}(t) &= \omega\{-\sin(\omega t)\mathbf{x}_0 + \sin(\omega t)(\mathbf{x}_0 \cdot \mathbf{a})\mathbf{a} + \cos(\omega t)\mathbf{a} \times \mathbf{x}_0\} \\ &= \omega\mathbf{a} \times \mathbf{x}(t). \end{aligned}$$

Damit hat das Geschwindigkeitsfeld einer gleichförmigen Drehbewegung die Darstellung

$$\mathbf{v}(\mathbf{x}) = \omega\mathbf{a} \times \mathbf{x}.$$

Das Vektorfeld $\mathbf{v} : \mathbf{x} \mapsto \omega \mathbf{a} \times \mathbf{x}$ ist linear und hat die konstante Jacobi-Matrix

$$\begin{aligned} D\mathbf{v}(\mathbf{x}) &= (\mathbf{v}(\mathbf{e}_1), \mathbf{v}(\mathbf{e}_2), \mathbf{v}(\mathbf{e}_3)) \\ &= \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \omega. \end{aligned}$$

Offensichtlich ist $D\mathbf{v}$ schiefsymmetrisch und hat verschwindende Spur, d.h. $\operatorname{div} \mathbf{v}(\mathbf{x}) = 0$ für alle $\mathbf{x} \in \mathbb{R}^3$ (vgl. Abschnitt VIII.4.5).

BEISPIEL VIII.4.9 (ZENTRALES KRAFTFELD). Eine Punktmasse M im Ursprung zieht nach I. NEWTON die Punktmasse m in $X = (x_1, x_2, x_3)$ mit der GRAVITATIONS-KRAFT

$$\mathbf{K}(\mathbf{x}) = \frac{c}{|\mathbf{x}|^3} \mathbf{x}, \quad \mathbf{x} \neq 0$$

an, wobei $\mathbf{x} = \overrightarrow{OX}$ und $c = -\gamma m M$ mit $\gamma > 0$ ist. Das zentrale Kraftfeld ist quasilinear. Daher ergibt sich für $\mathbf{x} \neq 0$ mit der 3×3 Einheitsmatrix \mathbb{I}

$$\begin{aligned} D\mathbf{K}(\mathbf{x}) &= \mathbf{x}(\operatorname{grad}(\frac{c}{|\mathbf{x}|^3}))^T + \frac{c}{|\mathbf{x}|^3} \mathbb{I} \\ &= \frac{c}{|\mathbf{x}|^5} \{ \mathbf{x} \cdot \mathbf{x} \mathbb{I} - 3\mathbf{x}\mathbf{x}^T \} \\ &= \frac{c}{|\mathbf{x}|^5} \begin{pmatrix} x_2^2 + x_3^2 - 2x_1^2 & -3x_1x_2 & -3x_1x_3 \\ -3x_1x_2 & x_1^2 + x_3^2 - 2x_2^2 & -3x_2x_3 \\ -3x_1x_3 & -3x_2x_3 & x_1^2 + x_2^2 - 2x_3^2 \end{pmatrix}. \end{aligned}$$

$D\mathbf{K}$ ist symmetrisch mit verschwindender Spur, d.h. $\operatorname{div} \mathbf{K}(\mathbf{x}) = 0$ für alle $\mathbf{x} \neq 0$ (vgl. Abschnitt VIII.4.5).

BEISPIEL VIII.4.10 (LAMINARE ROHRSTRÖMUNG). Eine zähe Flüssigkeit wird durch ein zur x_2 -Achse koaxiales Rohr vom Radius r mit geringer Geschwindigkeit gepresst, sodass eine laminare Strömung entsteht. Das Geschwindigkeitsfeld wurde 1850 unabhängig voneinander von dem deutschen Ingenieur G. HAGEN und dem französischen Arzt J. L. M. POISEUILLE bestimmt zu

$$\mathbf{v}(\mathbf{x}) = c \begin{pmatrix} 0 \\ r^2 - x_1^2 - x_3^2 \\ 0 \end{pmatrix} \quad \text{mit } x_1^2 + x_3^2 \leq r^2, \quad c > 0.$$

Für die Jacobi-Matrix ergibt sich

$$D\mathbf{v}(\mathbf{x}) = \begin{pmatrix} 0 & 0 & 0 \\ -2x_1 & 0 & -2x_3 \\ 0 & 0 & 0 \end{pmatrix} c.$$

Sie ist weder symmetrisch noch schiefsymmetrisch, hat aber verschwindende Spur, d.h. $\operatorname{div} \mathbf{v}(\mathbf{x}) = 0$ für alle \mathbf{x} (vgl. Abschnitt VIII.4.5).

VIII.4.5. Gradient, Divergenz, Rotation und Laplace Operator. Im Folgenden ist stets $D \subset \mathbb{R}^3$ eine offene Menge und $f : D \rightarrow \mathbb{R}$ bzw. $\mathbf{v} : D \rightarrow \mathbb{R}^3$ ein hinreichend oft differenzierbares Skalaren- bzw. Vektorfeld. Hierfür können wir die folgenden Differentialoperatoren definieren:

GRADIENT:

$$\text{grad } f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{pmatrix}$$

LAPLACE-OPERATOR (skalar):

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2}$$

DIVERGENZ:

$$\text{div } \mathbf{v} = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3}$$

ROTATION:

$$\text{rot } \mathbf{v} = \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}$$

LAPLACE-OPERATOR (vektoriell):

$$\Delta \mathbf{v} = \begin{pmatrix} \Delta v_1 \\ \Delta v_2 \\ \Delta v_3 \end{pmatrix}$$

Mittels Volumen- und Oberflächenintegralen (vgl. Abschnitte [IX.4.5](#) (S. 362), [IX.4.6](#) (S. 363) und [IX.5.5](#) (S. 372)) kann man zeigen, dass $\text{div } \mathbf{v}$ die QUELLDICHTE von \mathbf{v} und $\text{rot } \mathbf{v}$ die WIRBELDICHTE von \mathbf{v} beschreiben.

BEISPIEL VIII.4.11. Für die Vektorfelder der Beispiele [VIII.4.8](#) – [VIII.4.10](#) gilt:

- starre Drehung:

$$\mathbf{v}(\mathbf{x}) = \omega \mathbf{a} \times \mathbf{x}$$

$$\text{div } \mathbf{v}(\mathbf{x}) = 0$$

$$\text{rot } \mathbf{v}(\mathbf{x}) = 2\omega \mathbf{a}.$$

- zentrales Kraftfeld:

$$\mathbf{v}(\mathbf{x}) = \frac{c}{|\mathbf{x}|^3} \mathbf{x}$$

$$\text{div } \mathbf{v}(\mathbf{x}) = 0$$

$$\operatorname{rot} \mathbf{v}(\mathbf{x}) = 0.$$

- Rohrströmung:

$$\mathbf{v}(\mathbf{x}) = c \begin{pmatrix} 0 \\ r^2 - x_1^2 - x_3^2 \\ 0 \end{pmatrix}$$

$$\operatorname{div} \mathbf{v}(\mathbf{x}) = 0$$

$$\operatorname{rot} \mathbf{v}(\mathbf{x}) = 2c \begin{pmatrix} x_3 \\ 0 \\ -x_1 \end{pmatrix}$$

BEISPIEL VIII.4.12. Für das Vektorfeld

$$\mathbf{v}(\mathbf{x}) = \begin{pmatrix} 2xy \\ x^2 + 3z^2 \\ 9yz^2 \end{pmatrix}$$

erhalten wir

$$D\mathbf{v}(\mathbf{x}) = \begin{pmatrix} 2y & 2x & 0 \\ 2x & 0 & 6z \\ 0 & 9z^2 & 18yz \end{pmatrix}$$

$$\operatorname{div} \mathbf{v}(\mathbf{x}) = 2y + 18yz$$

$$\begin{aligned} \operatorname{rot} \mathbf{v}(\mathbf{x}) &= \begin{pmatrix} 9z^2 - 6z \\ 0 \\ 2x - 2x \end{pmatrix} \\ &= \begin{pmatrix} 9z^2 - 6z \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Es besteht der folgende Zusammenhang zwischen der Jacobi-Matrix und den Operatoren Gradient, Divergenz und Rotation:

$$\begin{aligned} \operatorname{grad} f(\mathbf{x}_0) &= Df(\mathbf{x}_0)^T \\ \operatorname{div} \mathbf{v}(\mathbf{x}_0) &= \operatorname{Spur} D\mathbf{v}(\mathbf{x}_0) \\ \operatorname{rot} \mathbf{v}(\mathbf{x}_0) \times (\mathbf{x} - \mathbf{x}_0) &= [D\mathbf{v}(\mathbf{x}_0) - D\mathbf{v}(\mathbf{x}_0)^T](\mathbf{x} - \mathbf{x}_0). \end{aligned}$$

BEMERKUNG VIII.4.13 (KOORDINATENINVARIANZ). Der Wert von $\operatorname{div} \mathbf{v}(\mathbf{x})$ ist unabhängig vom Koordinatensystem, in dem man \mathbf{v} darstellt und die partiellen Ableitungen berechnet. Dasselbe gilt für die Länge und Richtung von $\operatorname{grad} f(\mathbf{x})$ und $\operatorname{rot} \mathbf{v}(\mathbf{x})$.

Es gelten folgende Rechenregeln:

$$\begin{aligned}
\operatorname{rot}(\operatorname{grad} f) &= 0 \quad (\text{Gradientenfeld ist wirbelfrei}) \\
\operatorname{div}(\operatorname{rot} \mathbf{v}) &= 0 \quad (\text{Feld der Rotation ist quelfrei}) \\
\operatorname{div}(\operatorname{grad} f) &= \Delta f \\
\operatorname{div}(f \mathbf{v}) &= (\operatorname{grad} f) \cdot \mathbf{v} + f \operatorname{div} \mathbf{v} \\
\operatorname{rot}(f \mathbf{v}) &= (\operatorname{grad} f) \times \mathbf{v} + f \operatorname{rot} \mathbf{v} \\
\operatorname{rot}(\operatorname{rot} \mathbf{v}) &= \operatorname{grad}(\operatorname{div} \mathbf{v}) - \Delta \mathbf{v}.
\end{aligned}$$

Mit dem symbolischen Operator

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix},$$

dem sog. NABLA-OPERATOR, lassen sich die vier hier betrachteten Felder formal als Produkte schreiben:

$$\begin{aligned}
\operatorname{grad} f &= \nabla f, \\
\Delta f &= \nabla \cdot (\nabla f), \\
\operatorname{div} \mathbf{v} &= \nabla \cdot \mathbf{v} \\
\operatorname{rot} \mathbf{v} &= \nabla \times \mathbf{v}.
\end{aligned}$$

ACHTUNG: Beim Umgang mit dem Nabla-Operator ist Vorsicht geboten, da man mit ihm nicht immer wie mit einem Vektor rechnen kann.

BEISPIEL VIII.4.14. Für die Vektorfelder

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x^2 \\ y^2 \\ -xz \end{pmatrix}$$

und

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$

erhalten wir

$$\begin{aligned}
(\mathbf{f} \times \mathbf{g})(\mathbf{x}) &= \begin{pmatrix} x^2 z \\ xyz \\ x^3 + y^3 \end{pmatrix}, \\
(\mathbf{f} \cdot \mathbf{g})(\mathbf{x}) &= -x^2 y + xy^2
\end{aligned}$$

und damit

$$\nabla(\mathbf{f} \cdot \mathbf{g}) = \begin{pmatrix} -2xy + y^2 \\ -x^2 + 2xy \\ 0 \end{pmatrix}$$

$$\begin{aligned}\nabla \cdot (\mathbf{f} \times \mathbf{g}) &= 2xz + xz + 0 \\ &= 3xz\end{aligned}$$

$$\begin{aligned}\nabla \times (\mathbf{f} \times \mathbf{g}) &= \begin{pmatrix} 3y^2 - xy \\ x^2 - 3x^2 \\ yz - 0 \end{pmatrix} \\ &= \begin{pmatrix} 3y^2 - xy \\ -2x^2 \\ yz \end{pmatrix}.\end{aligned}$$

KAPITEL IX

Integration von Funktionen in mehreren Variablen

IX.1. Parameterintegrale

IX.1.1. Übersicht. Viele wichtige Funktionen der Analysis haben die Form

$$F(x) = \int_c^d f(x, y) dy \quad , a \leq x \leq b.$$

Da der Integrand von dem Parameter x abhängt, spricht man von einem **PARAMETERINTEGRAL**. Beispiele sind:

- die **EULERSCHE GAMMAFUNKTION**

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad x > 0,$$

- die **BESSELFUNKTIONEN**

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) dt \quad n \in \mathbb{Z}, x \in \mathbb{R},$$

- die **FOURIER-TRANSFORMIERTE**

$$\mathcal{F}(f)(\omega) = \int_{-\infty}^\infty e^{-i\omega t} f(t) dt \quad \omega \in \mathbb{R}.$$

Wenn alle Zahlen a, b, c, d endlich sind und der Integrand f auf dem Bereich $a \leq x \leq b, c \leq y \leq d$ stetig ist, spricht man von einem **EIGENTLICHEN PARAMETERINTEGRAL**, sonst von einem **UNEIGENTLICHEN PARAMETERINTEGRAL**. Für beide Fälle wollen wir im Folgenden die Integration und die Differentiation bzgl. der Variablen x untersuchen.

IX.1.2. Eigentliche Parameterintegrale. Wir betrachten ein abgeschlossenes und beschränktes Rechteck $D = [a, b] \times [c, d] = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$ in \mathbb{R}^2 und eine stetige Funktion $f : D \rightarrow \mathbb{R}$. Dann hat die durch

$$F(x) = \int_c^d f(x, y) dy$$

definierte Funktion $F : [a, b] \rightarrow \mathbb{R}$ folgende Eigenschaften:

- F ist stetig auf $[a, b]$.
- F ist integrierbar und

$$\int_a^b F(x) dx = \int_a^b \left\{ \int_c^d f(x, y) dy \right\} dx$$

$$= \int_c^d \left\{ \int_a^b f(x, y) dx \right\} dy.$$

- Ist f zusätzlich nach der Variablen x partiell differenzierbar und ist f_x auf D stetig, so ist F differenzierbar und

$$F'(x) = \frac{d}{dx} \int_c^d f(x, y) dy = \int_c^d f_x(x, y) dy.$$

BEISPIEL IX.1.1. Für

$$F(x) = \int_1^\pi \frac{\sin(tx)}{t} dt$$

erhalten wir

$$F'(x) = \int_1^\pi \cos(tx) dt,$$

$$F''(x) = - \int_1^\pi t \sin(tx) dt.$$

BEISPIEL IX.1.2. Für die BESSELFUNKTION

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) dt, \quad n \in \mathbb{Z}$$

erhalten wir

$$J'_n(x) = - \frac{1}{\pi} \int_0^\pi \sin(t) \sin(x \sin t - nt) dt,$$

$$J''_n(x) = - \frac{1}{\pi} \int_0^\pi \sin^2(t) \cos(x \sin t - nt) dt.$$

Den Ausdruck für J'_n können wir mittels partieller Integration umformen

$$\begin{aligned} J'_n(x) &= \frac{1}{\pi} \int_0^\pi \underbrace{(-\sin(t))}_{= \frac{d}{dt} \cos(t)} \sin(x \sin t - nt) dt \\ &= \frac{1}{\pi} \cos(t) \sin(x \sin t - nt) \Big|_{t=0}^{t=\pi} \\ &\quad - \frac{1}{\pi} \int_0^\pi \cos(t) \cos(x \sin t - nt) [x \cos(t) - n] dt \\ &= \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) [-x \cos^2(t) + n \cos(t)] dt. \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} &x^2 J''_n(x) + x J'_n(x) + (x^2 - n^2) J_n(x) \\ &= \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) [-x^2 \sin^2 t - x^2 \cos^2 t + nx \cos t + x^2 - n^2] dt \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) n [x \cos t - n] dt \\
&= \frac{n}{\pi} \sin(x \sin t - nt) \Big|_{t=0}^{t=\pi} \\
&= 0.
\end{aligned}$$

Also löst J_n die gewöhnliche Differentialgleichung

$$x^2 J_n'' + x J_n' + (x^2 - n^2) J_n = 0.$$

Häufig hängen die Integrationsgrenzen auch von x ab:

$$F(x) = \int_{g(x)}^{h(x)} f(x, y) dy.$$

Wenn g und h differenzierbar und f nach x partiell differenzierbar sind, kann man die Ableitung von F in diesem Fall wie folgt bestimmen: Definiere

$$G(x, u, v) = \int_u^v f(x, y) dy.$$

Dann ist

$$F(x) = G(x, g(x), h(x)).$$

Daher folgt mit der Kettenregel

$$\begin{aligned}
F'(x) &= G_x(x, g(x), h(x)) + G_u(x, g(x), h(x))g'(x) \\
&\quad + G_v(x, g(x), h(x))h'(x) \\
&= \int_{g(x)}^{h(x)} f_x(x, y) dy - f(x, g(x))g'(x) + f(x, h(x))h'(x).
\end{aligned}$$

Also insgesamt

$$\begin{aligned}
&\frac{d}{dx} \int_{g(x)}^{h(x)} f(x, y) dy \\
&= \int_{g(x)}^{h(x)} f_x(x, y) dy - f(x, g(x))g'(x) + f(x, h(x))h'(x).
\end{aligned}$$

BEISPIEL IX.1.3. Für

$$x(t) = \frac{1}{k} \int_0^t f(u) \sin(k(t-u)) du$$

erhalten wir mit obiger Formel und $g(t) = 0$, $h(t) = t$

$$\begin{aligned}
\dot{x}(t) &= \frac{1}{k} \int_0^t k f(u) \cos(k(t-u)) du + \frac{1}{k} f(t) \sin(k(t-t)) \\
&= \int_0^t f(u) \cos(k(t-u)) du
\end{aligned}$$

$$\begin{aligned}\ddot{x}(t) &= - \int_0^t k f(u) \sin(k(t-u)) du + f(t) \cos(k(t-t)) \\ &= -k^2 x(t) + f(t).\end{aligned}$$

Also ist x eine Lösung der Schwingungsgleichung

$$\ddot{x} + k^2 x = f.$$

IX.1.3. Uneigentliche Parameterintegrale. Die Resultate der vorigen Abschnittes können nicht ohne Weiteres auf uneigentliche Parameterintegrale, bei denen z.B. $d = \infty$ ist, übertragen werden. Man benötigt weitere Zusatzbedingungen. Der Einfachheit halber beschränken wir uns auf einen Spezialfall.

Sei $d = \infty$ und $D = [a, b] \times [c, d) = \{(x, y) : a \leq x \leq b, c \leq y < d\}$. Die Funktion f sei in D stetig und nach der Variablen x partiell differenzierbar mit stetiger Ableitung. Weiter gebe es zwei Funktionen $g, h : [c, d) \rightarrow \mathbb{R}$ mit den folgenden Eigenschaften:

- $|f(x, y)| \leq g(y)$ für alle $(x, y) \in D$,
- $|f_x(x, y)| \leq h(y)$ für alle $(x, y) \in D$,
- die uneigentlichen Integrale $\int_c^d g(y) dy$ und $\int_c^d h(y) dy$ existieren.

Dann existiert für jedes $x \in [a, b]$ das uneigentliche Integral

$$F(x) = \int_c^d f(x, y) dy.$$

Die dadurch definierte Funktion F ist differenzierbar mit

$$F'(x) = \int_c^d f_x(x, y) dy.$$

BEISPIEL IX.1.4. Wir betrachten

$$F(x) = \int_0^\infty e^{-t^2} \cos(xt) dt.$$

Es ist $f(x, t) = e^{-t^2} \cos(xt)$ und $f_x(x, t) = -te^{-t^2} \sin(xt)$. Daher gilt für alle x, t

$$|f(x, t)| \leq e^{-t^2}, \quad |f_x(x, t)| \leq te^{-t^2}.$$

Die uneigentlichen Integrale

$$\int_0^\infty e^{-t^2} dt, \quad \int_0^\infty te^{-t^2} dt$$

existieren wegen Abschnitt V.4.2 (S. 195). Daher ist F differenzierbar und erfüllt

$$F'(x) = - \int_0^\infty te^{-t^2} \sin(xt) dt.$$

Mittels partieller Integration folgt für $R > 0$

$$\begin{aligned} & \frac{1}{2} \int_0^R \underbrace{(-2te^{-t^2})}_{=\frac{d}{dt}e^{-t^2}} \sin(xt) dt \\ &= \frac{1}{2} e^{-t^2} \sin(xt) \Big|_0^R - \frac{1}{2} \int_0^R e^{-t^2} x \cos(xt) dt \\ &= \frac{1}{2} e^{-R^2} \sin(xR) - \frac{x}{2} \int_0^R e^{-t^2} \cos(xt) dt \\ &\xrightarrow{R \rightarrow \infty} -\frac{x}{2} F(x). \end{aligned}$$

Also erfüllt F die gewöhnliche Differentialgleichung

$$F' = -\frac{x}{2} F.$$

BEISPIEL IX.1.5. Mit obiger Vorgehensweise können die Ableitungen der Gammafunktion berechnet werden:

$$\begin{aligned} \Gamma(x) &= \int_0^\infty e^{-t} t^{x-1} dt, \\ \Gamma'(x) &= \int_0^\infty e^{-t} t^{x-1} \ln t dt, \\ \Gamma''(x) &= \int_0^\infty e^{-t} t^{x-1} (\ln t)^2 dt. \end{aligned}$$

IX.2. Kurvenintegrale

IX.2.1. Das Kurvenintegral einer skalaren Funktion. Wir betrachten eine offene Menge $D \subset \mathbb{R}^n$, eine skalare Funktion $f : D \rightarrow \mathbb{R}$ auf D und ein stetig differenzierbares Kurvenstück $\mathbf{w} : [a, b] \rightarrow D$ in D . Wir fassen f als eine Belegungsfunktion auf \mathbf{w} auf und wollen den Gesamtwert der Belegung von \mathbf{w} berechnen. Dazu unterteilen wir das Kurvenstück, indem wir Punkte $a = t_0 < t_1 < \dots < t_n = b$ im Parameterintervall und in jedem Teilintervall $[t_{i-1}, t_i]$ einen Punkt η_i wählen. Dann ist der Wert der Belegung näherungsweise

$$\sum_{i=1}^n f(\mathbf{w}(\eta_i)) \underbrace{|\mathbf{w}(t_i) - \mathbf{w}(t_{i-1})|}_{\approx |\dot{\mathbf{w}}(\eta_i)|(t_i - t_{i-1})} \approx \sum_{i=1}^n f(\mathbf{w}(\eta_i)) |\dot{\mathbf{w}}(\eta_i)| (t_i - t_{i-1}).$$

Bei immer feiner werdender Zerlegung strebt dies gegen

$$\int_a^b f(\mathbf{w}(t)) |\dot{\mathbf{w}}(t)| dt.$$

Daher definieren wir:

Das KURVENINTEGRAL von f längs \mathbf{w} ist

$$\int_{\mathbf{w}} f ds = \int_a^b f(\mathbf{w}(t)) |\dot{\mathbf{w}}(t)| dt.$$

Die Berechnung eines Kurvenintegrals erfolgt in drei Schritten:

- Bestimme eine Parametrisierung $\mathbf{w} : [a, b] \rightarrow \mathbb{R}^n$ des Kurvenstückes.
- Bestimme das Bogenelement $ds = |\dot{\mathbf{w}}(t)| dt$.
- Berechne das Integral

$$\int_a^b f(\mathbf{w}(t)) |\dot{\mathbf{w}}(t)| dt.$$

BEISPIEL IX.2.1. Wir betrachten eine Schraubenfeder mit Radius 2, Ganghöhe π , Gesamthöhe 2π und Massendichte $\rho(x, y, z) = x^2 y^2 + z^2$. Zu bestimmen ist die Gesamtmasse M . Dies bedeutet die Berechnung eines Kurvenintegrals mit ρ als Belegungsfunktion und der Schraubenfeder als Kurvenstück. Eine Parametrisierung ist

$$\mathbf{w}(t) = \left(2 \cos t, 2 \sin t, \frac{t}{2} \right)^T, \quad 0 \leq t \leq 4\pi.$$

Das Bogenelement ist

$$ds = \left| \left(-2 \sin t, 2 \cos t, \frac{1}{2} \right)^T \right| dt = \sqrt{4 + \frac{1}{4}} dt = \frac{\sqrt{17}}{2} dt.$$

Damit erhalten wir

$$\begin{aligned} M &= \int_0^{4\pi} \rho(\mathbf{w}(t)) |\dot{\mathbf{w}}(t)| dt \\ &= \int_0^{4\pi} \left[4 \sin^2 t \cdot 4 \cos^2 t + \frac{1}{4} t^2 \right] \frac{\sqrt{17}}{2} dt \\ &= \frac{\sqrt{17}}{2} \left\{ \underbrace{16 \int_0^{4\pi} \sin^2 t \cos^2 t dt}_{=\frac{1}{2}\pi} + \frac{1}{4} \underbrace{\int_0^{4\pi} t^2 dt}_{=\frac{64}{3}\pi^3} \right\} \\ &= \frac{\sqrt{17}}{2} \left\{ 8\pi + \frac{16}{3}\pi^3 \right\}. \end{aligned}$$

Es ist zweckmäßig, den Begriff des Kurvenintegrals auch auf Kurven auszudehnen, die aus aneinanderhängenden Kurvenstücken bestehen.

DEFINITION IX.2.2. (1) Unter einer KURVE \mathbf{w} in $D \subset \mathbb{R}^n$ verstehen wir eine endliche Folge $\mathbf{w}_1, \dots, \mathbf{w}_k$ von stetig differenzierbaren Kurvenstücken $\mathbf{w}_i : [a_i, b_i] \rightarrow D$ mit $\mathbf{w}_i(b_i) = \mathbf{w}_{i+1}(a_{i+1})$ für $i = 1, \dots, k-1$. (2) Das KURVENINTEGRAL einer Funktion $f : D \rightarrow \mathbb{R}$ entlang einer Kurve \mathbf{w} , die aus den Kurvenstücken $\mathbf{w}_1, \dots, \mathbf{w}_k$ besteht, ist die Summe der Kurvenintegrale von f längs der Kurvenstücke

$$\int_{\mathbf{w}} f ds = \int_{\mathbf{w}_1} f ds + \dots + \int_{\mathbf{w}_k} f ds.$$

BEISPIEL IX.2.3. Die Kurve \mathbf{w} führe vom Ursprung zum Punkt $(2, 1, 3)$ entlang achsenparalleler Strecken und zwar zunächst parallel zur x -, dann zur y - und zuletzt zur z -Achse. Dann besteht \mathbf{w} aus drei Kurvenstücken $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ mit

$$\begin{aligned} \mathbf{w}_1 : [0, 1] &\rightarrow \mathbb{R}^3 & , t &\mapsto (2t, 0, 0)^T, \\ \mathbf{w}_2 : [0, 1] &\rightarrow \mathbb{R}^3 & , t &\mapsto (2, t, 0)^T, \\ \mathbf{w}_3 : [0, 1] &\rightarrow \mathbb{R}^3 & , t &\mapsto (2, 1, 3t)^T. \end{aligned}$$

Für die Funktion $f(\mathbf{x}) = |\mathbf{x}|^2$ erhalten wir dann das Kurvenintegral

$$\begin{aligned} \int_{\mathbf{w}} f ds &= \int_{\mathbf{w}_1} f ds + \int_{\mathbf{w}_2} f ds + \int_{\mathbf{w}_3} f ds \\ &= \int_0^1 4t^2 \cdot 2dt + \int_0^1 (4 + t^2) \cdot 1dt + \int_0^1 (4 + 1 + 9t^2) \cdot 3dt \\ &= \frac{8}{3} + \left(4 + \frac{1}{3}\right) + (15 + 9) \\ &= 31. \end{aligned}$$

Für Kurvenintegrale gelten die üblichen Rechenregeln ($f, g : D \rightarrow \mathbb{R}, \alpha \in \mathbb{R}$):

$$\begin{aligned} \int_{\mathbf{w}} \alpha f ds &= \alpha \int_{\mathbf{w}} f ds \\ \int_{\mathbf{w}} (f + g) ds &= \int_{\mathbf{w}} f ds + \int_{\mathbf{w}} g ds \end{aligned}$$

Außerdem gilt der MITTELWERTSATZ

$$\int_{\mathbf{w}} f ds = f(\tilde{\mathbf{x}})L,$$

wobei $\tilde{\mathbf{x}}$ ein geeigneter Punkt auf der Kurve und L die Länge der Kurve ist.

IX.2.2. Anwendungen. Stellt $\rho(x, y, z)$ die Massendichte einer Kurve \mathbf{w} dar, so beträgt die Gesamtmasse

$$M = \int_{\mathbf{w}} \rho ds.$$

Man nennt

$$dm = \rho(x, y, z) ds$$

daher auch das Massenelement.

Die k -ten Momente bezüglich der Koordinatenebenen eines in (x, y, z) befindlichen Massenpunktes der Masse dm sind $x^k dm$, $y^k dm$ und $z^k dm$. Dementsprechend sind die k -ten Momente einer Kurve \mathbf{w} mit der Massendichte $\rho(x, y, z)$ gegeben durch

$$M_{x,k} = \int_{\mathbf{w}} x^k \rho ds,$$

$$M_{y,k} = \int_{\mathbf{w}} y^k \rho ds,$$

$$M_{z,k} = \int_{\mathbf{w}} z^k \rho ds.$$

Für $k = 1$ sind dies die STATISCHEN MOMENTE und für $k = 2$ die TRÄGHEITSMOMENTE. Der Punkt $S = (x_S, y_S, z_S)$, in dem eine Punktmasse der Größe M die selben statischen Momente wie die Kurve \mathbf{w} besitzt, heißt MASSENMITTELPUNKT oder SCHWERPUNKT. Damit folgt

$$x_S = \frac{1}{M} \int_{\mathbf{w}} x \rho ds,$$

$$y_S = \frac{1}{M} \int_{\mathbf{w}} y \rho ds,$$

$$z_S = \frac{1}{M} \int_{\mathbf{w}} z \rho ds$$

mit

$$M = \int_{\mathbf{w}} \rho ds.$$

Im Falle einer homogenen Massenverteilung, d.h. $\rho(x, y, z) = \text{const}$, stimmt der Massenmittelpunkt mit dem geometrischen Schwerpunkt überein. In diesem Fall ist in obiger Formel M durch die Länge L der Kurve und ρ durch 1 zu ersetzen.

BEISPIEL IX.2.4. Betrachte die Schraubenlinie

$$\mathbf{w}(t) = (r \cos t, r \sin t, ht)^T, \quad 0 \leq t \leq 2\pi n.$$

Mit

$$R = \sqrt{r^2 + h^2}$$

ist ihre Länge

$$\begin{aligned} L &= \int_0^{2\pi n} \sqrt{r^2 + h^2} dt \\ &= 2\pi n R. \end{aligned}$$

Wegen

$$\int_0^{2\pi n} \cos t dt = \int_0^{2\pi n} \sin t dt = 0$$

sind die x - und die y - Koordinate des Schwerpunktes gleich Null. Mit

$$H = 2\pi n h$$

ergibt sich für die z -Koordinate

$$\begin{aligned} z_S &= \frac{1}{L} \int_0^{2\pi n} h t \sqrt{r^2 + h^2} dt \\ &= \frac{1}{2\pi n R} \frac{1}{2} h R (2\pi n)^2 \\ &= \pi n h \\ &= \frac{H}{2}. \end{aligned}$$

Also ist $(0, 0, \frac{H}{2})$ der Schwerpunkt. Er liegt *nicht* auf \mathbf{w} .

IX.2.3. Das Kurvenintegral eines Vektorfeldes. Wir betrachten eine offene Menge $D \subset \mathbb{R}^n$, ein stetig differenzierbares Kurvenstück $\mathbf{w} : [a, b] \rightarrow D$ in D und ein Vektorfeld $\mathbf{v} : D \rightarrow \mathbb{R}^n$ auf D . Wir wollen \mathbf{v} längs \mathbf{w} integrieren. Dazu überlegen wir uns, dass hierfür in einem beliebigen Kurvenpunkt $\mathbf{w}(t)$ nur die Projektion von \mathbf{v} auf die Kurve eine Rolle spielt. Dies ist aber das Skalarprodukt von $\mathbf{v}(\mathbf{w}(t))$ mit dem Einheitstangentenvektor $\frac{1}{|\dot{\mathbf{w}}(t)|} \dot{\mathbf{w}}(t)$. Daher entspricht die Integration von \mathbf{v} längs \mathbf{w} dem Kurvenintegral der skalaren Belegung $\mathbf{v}(\mathbf{w}(t)) \cdot \dot{\mathbf{w}}(t) \frac{1}{|\dot{\mathbf{w}}(t)|}$. Deshalb definieren wir:

Das KURVENINTEGRAL des Vektorfeldes \mathbf{v} längs des Kurvenstückes \mathbf{w} ist

$$\int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = \int_a^b \mathbf{v}(\mathbf{w}(t)) \cdot \dot{\mathbf{w}}(t) dt.$$

Das Integral läng einer Kurve, die aus den Kurvenstücken $\mathbf{w}_1, \dots, \mathbf{w}_k$ besteht, ist analog definiert als die Summe der Integrale längs der einzelnen Kurvenstücke.

Es gelten die üblichen Rechenregeln ($\mathbf{u}, \mathbf{v} : D \rightarrow \mathbb{R}^n, \alpha \in \mathbb{R}$):

$$\begin{aligned} \int_{\mathbf{w}} \alpha \mathbf{u} \cdot d\mathbf{x} &= \alpha \int_{\mathbf{w}} \mathbf{u} \cdot d\mathbf{x} \\ \int_{\mathbf{w}} (\mathbf{u} + \mathbf{v}) \cdot d\mathbf{x} &= \int_{\mathbf{w}} \mathbf{u} \cdot d\mathbf{x} + \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} \end{aligned}$$

BEISPIEL IX.2.5. Wir betrachten das Vektorfeld

$$\mathbf{v}(x, y, z) = (x^2y, x - z, xyz)^T$$

und als Kurve den Graphen von $y = x^3$, $0 \leq x \leq 2$, in der Ebene $z = 2$. Eine Parametrisierung ist

$$\mathbf{w}(t) = (t, t^3, 2)^T, \quad 0 \leq t \leq 2.$$

Damit ergibt sich

$$\begin{aligned} \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} &= \int_0^2 \begin{pmatrix} t^5 \\ t - 2 \\ 2t^4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3t^2 \\ 0 \end{pmatrix} dt \\ &= \int_0^2 [t^5 + 3t^2(t - 2)] dt \\ &= \frac{64}{6} + \frac{3 \cdot 16}{4} - \frac{6 \cdot 8}{3} \\ &= \frac{20}{3}. \end{aligned}$$

BEISPIEL IX.2.6. Wir wollen das Integral des Wirbels

$$\mathbf{v}(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \end{pmatrix}, \quad (x, y) \neq (0, 0)$$

längs des einmal positiv durchlaufenen Einheitskreises berechnen. Eine Parametrisierung ist

$$\mathbf{w}(t) = (\cos t, \sin t)^T, \quad 0 \leq t \leq 2\pi.$$

Damit ergibt sich

$$\begin{aligned} \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} &= \int_0^{2\pi} \underbrace{\frac{1}{\cos^2 t + \sin^2 t}}_{=1} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt \\ &= \int_0^{2\pi} \underbrace{\cos^2 t + \sin^2 t}_{=1} dt \\ &= 2\pi. \end{aligned}$$

Eine Kurve \mathbf{w} heißt GESCHLOSSEN, wenn ihr Anfangs- und Endpunkt übereinstimmen, d.h. $\mathbf{w}(a) = \mathbf{w}(b)$ mit $\mathbf{w} : [a, b] \rightarrow \mathbb{R}^n$. Ist \mathbf{w} geschlossen, schreiben wir

$$\oint_{\mathbf{w}} f ds \quad \text{bzw.} \quad \oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}$$

statt $\int_{\mathbf{w}} f ds$ bzw. $\int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x}$.

IX.2.4. Das Potential eines Gradientenfeldes. Sei $D \subset \mathbb{R}^n$ offen. Wir nennen ein Vektorfeld $\mathbf{v} : D \rightarrow \mathbb{R}^n$ ein GRADIENTENFELD, wenn es eine skalare Funktion $f : D \rightarrow \mathbb{R}$ gibt mit $\mathbf{v} = \text{grad } f$. In diesem Fall heißt f eine STAMMFUNKTION von \mathbf{v} und $U = -f$ ein POTENTIAL von \mathbf{v} .

BEMERKUNG IX.2.7. In einem konservativen Kraftfeld stellt das Potential die potentielle Energie dar. Das Vorzeichen ist so gewählt, dass $\mathbf{v} = -\text{grad } U$ in Richtung des stärksten Potentialabfalls zeigt.

Im Mittelpunkt dieses Abschnittes steht die Frage: „Wann ist ein gegebenes Vektorfeld ein Gradientenfeld und wie kann man gegebenenfalls eine Stammfunktion bestimmen?“

Dazu überlegen wir uns zunächst einige Konsequenzen aus der Eigenschaft, Gradientenfeld zu sein. Sei also $\mathbf{v} = \text{grad } f$ ein Gradientenfeld und \mathbf{w} ein Kurvenstück in D . Dann folgt mit der Kettenregel

$$\begin{aligned} \int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} &= \int_a^b \mathbf{v}(\mathbf{w}(t)) \cdot \dot{\mathbf{w}}(t) dt \\ &= \int_a^b \text{grad } f(\mathbf{w}(t)) \cdot \dot{\mathbf{w}}(t) dt \\ &= \int_a^b \frac{d}{dt} [f(\mathbf{w}(t))] dt \\ &= f(\mathbf{w}(b)) - f(\mathbf{w}(a)). \end{aligned}$$

Hieraus folgt unmittelbar:

Sei \mathbf{v} ein Gradientenfeld und f eine Stammfunktion von \mathbf{v} . Dann gilt:

- (1) Das Kurvenintegral hängt nur von den Werten der Stammfunktion im Anfangs- und Endpunkt der Kurve ab, d.h.

$$\int_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = f(\mathbf{w}(b)) - f(\mathbf{w}(a)).$$

- (2) Sind \mathbf{w}_1 und \mathbf{w}_2 zwei Wege mit gleichem Anfangs- und Endpunkt, so ist

$$\int_{\mathbf{w}_1} \mathbf{v} \cdot d\mathbf{x} = \int_{\mathbf{w}_2} \mathbf{v} \cdot d\mathbf{x},$$

d.h. das Kurvenintegral ist WEGUNABHÄNGIG.

- (3) Ist \mathbf{w} ein geschlossener Weg, so verschwindet das Kurvenintegral

$$\oint_{\mathbf{w}} \mathbf{v} \cdot d\mathbf{x} = 0.$$

BEISPIEL IX.2.8. Die ZENTRALE KRAFT

$$\mathbf{K} = \frac{c}{|\mathbf{x}|^3} \mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^3 \setminus \{0\}$$

besitzt das Potential

$$U = \frac{c}{|\mathbf{x}|}.$$

Ist \mathbf{w} ein Kurvenstück, das nicht durch den Ursprung geht, ist die längs \mathbf{w} geleistete Arbeit gegeben durch

$$\begin{aligned} A &= \int_{\mathbf{w}} \mathbf{K} \cdot d\mathbf{x} \\ &= U(\mathbf{w}(a)) - U(\mathbf{w}(b)) \\ &= \frac{c}{|\mathbf{w}(a)|} - \frac{c}{|\mathbf{w}(b)|}. \end{aligned}$$

BEISPIEL IX.2.9. Betrachte das elektrische Feld

$$\mathbf{E}(x, y, z) = \begin{pmatrix} 2xy + z^3 \\ x^2 + 3z \\ 3z^2x + 3y \end{pmatrix}$$

mit dem Potential

$$U(x, y, z) = -(x^2y + xz^3 + 3zy).$$

Der Spannungsabfall zwischen den Punkten $P = (1, 1, 1)$ und $Q = (3, 4, 5)$ längs eines Weges von P nach Q hängt nicht von dem Weg ab und ist gegeben durch

$$\int_{\mathbf{w}} \mathbf{E} \cdot d\mathbf{x} = U(P) - U(Q) = 466.$$

Sei $\mathbf{v} = \text{grad } f$ ein Gradientenfeld. Falls \mathbf{v} stetig differenzierbar ist, ist f zweimal stetig differenzierbar und es gilt $f_{x_i x_j} = f_{x_j x_i}$ für alle $i \neq j$. Dies bedeutet:

Ist \mathbf{v} ein stetig differenzierbares Gradientenfeld, so ist die Jacobi-Matrix $D\mathbf{v}(\mathbf{x})$ in jedem Punkt $\mathbf{x} \in D$ symmetrisch.

Der Wirbel \mathbf{v} aus Beispiel IX.2.6 (S. 334) ist kein Gradientenfeld, da das Integral längs des Einheitskreises nicht verschwindet. Es ist aber

$$\begin{aligned} \frac{\partial v_1}{\partial y} &= -\frac{1}{x^2 + y^2} + \frac{2y^2}{(x^2 + y^2)^2} \\ &= \frac{y^2 - x^2}{(x^2 + y^2)^2}, \\ \frac{\partial v_2}{\partial x} &= \frac{1}{x^2 + y^2} - \frac{2x^2}{(x^2 + y^2)^2} \end{aligned}$$

$$= \frac{y^2 - x^2}{(x^2 + y^2)^2},$$

d.h. die Jacobi-Matrix $D\mathbf{v}$ ist symmetrisch. Dieses Beispiel zeigt, dass die Umkehrung obigen Ergebnisses nicht für jede offene Menge D gilt. Für eine möglichst einfache und gleichzeitig hinreichend allgemeine Charakterisierung von Mengen, für die die Umkehrung obiger Aussage gilt, benötigen wir folgende Definition.

DEFINITION IX.2.10. Eine offene Menge $D \subset \mathbb{R}^n$ heißt STERNFÖRMIG, wenn es einen Punkt $\mathbf{x}_0 \in D$ gibt, so dass für *jeden* Punkt $\mathbf{x} \in D$ die Verbindungsstrecke von \mathbf{x}_0 nach \mathbf{x} ganz in D liegt.

BEMERKUNG IX.2.11. Eine andere häufig betrachtete Klasse von Mengen sind die KONVEXEN Mengen. Sie sind dadurch charakterisiert, dass für jedes Paar von Punkten aus der Menge auch die gesamte Verbindungsstrecke in der Menge verläuft. Konvexe Mengen sind immer sternförmig, die Umkehrung gilt aber nicht (vgl. Abb. IX.2.1).

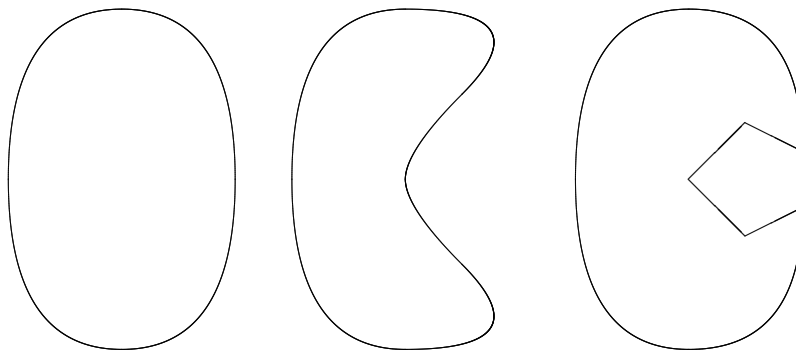


ABBILDUNG IX.2.1. Beispiele einer konvexen Menge (links), einer sternförmigen nicht konvexen Menge (Mitte) und einer nicht sternförmigen Menge (rechts)

BEISPIEL IX.2.12. \mathbb{R}^n und $B_r(0) \subset \mathbb{R}^n$ mit $n \geq 2$ und $r > 0$ sind sternförmig. $\mathbb{R}^n \setminus \{0\}$ und $B_r(0) \setminus \{0\} \subset \mathbb{R}^n$ sind nicht sternförmig.

Das Hauptergebnis dieses Abschnittes lautet:

SATZ VON POINCARÉ: Die Menge $D \subset \mathbb{R}^n$ sei sternförmig. Dann ist ein stetig differenzierbares Vektorfeld $\mathbf{v} : D \rightarrow \mathbb{R}^n$ genau dann ein Gradientenfeld, wenn die Jacobi-Matrix $D\mathbf{v}(\mathbf{x})$ in jedem Punkt $\mathbf{x} \in D$ symmetrisch ist.

Den Satz von Poincaré lautet kann man in Kurzform als Merkgel formulieren:

MERKREGEL: D sternförmig und $D\mathbf{v}(\mathbf{x})$ symmetrisch für alle $\mathbf{x} \in D \implies \mathbf{v}$ ist Gradientenfeld

bzw. in drei Dimensionen:

MERKREGEL: $D \subset \mathbb{R}^3$ sternförmig und $\operatorname{rot} \mathbf{v}(\mathbf{x}) = 0$ für alle $\mathbf{x} \in D \implies \mathbf{v}$ ist Gradientenfeld

IX.2.5. Die praktische Bestimmung einer Stammfunktion.

Wir beschränken uns der Übersichtlichkeit halber auf den praktisch wichtigen Fall von drei Dimensionen, d.h. $n = 3$. Es gibt hier im wesentlichen zwei Vorgehensweisen:

- mittels Kurvenintegralen,
- mittels Ansatz.

Die **Berechnung mit Kurvenintegralen** erfolgt in drei Schritten:

1. SCHRITT: Gibt es ein $\mathbf{x} \in D$ mit $\operatorname{rot} \mathbf{v}(\mathbf{x}) \neq 0$? Falls ja, ist \mathbf{v} kein Gradientenfeld, und wir sind fertig. Falls nein, gehe zu Schritt 2.
2. SCHRITT: Ist D sternförmig? Falls nein, führt die Methode nicht weiter. Falls ja, bestimme einen Punkt $\mathbf{x}_0 \in D$, so dass für jeden Punkt $\mathbf{x} \in D$ die Strecke von \mathbf{x}_0 nach \mathbf{x} in D verläuft.
3. SCHRITT: Mit dem Punkt \mathbf{x}_0 aus Schritt 2 berechne für jedes $\mathbf{x} \in D$

$$f(\mathbf{x}) = \int_0^1 \mathbf{v}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \cdot (\mathbf{x} - \mathbf{x}_0) dt.$$

Es ist $\mathbf{v} = \operatorname{grad} f$.

BEISPIEL IX.2.13. Wir betrachten

$$\mathbf{v}(x, y, z) = \begin{pmatrix} y^2 \cos x \\ 2y \sin x + e^{2z} \\ 2ye^{2z} \end{pmatrix}$$

auf $D = \mathbb{R}^3$. Es ist $\operatorname{rot} \mathbf{v} = 0$ auf D , und D ist sternförmig. Wir wählen $\mathbf{x}_0 = 0$. Dann folgt für beliebiges $\mathbf{x} \in \mathbb{R}^3$

$$\begin{aligned} f(\mathbf{x}) &= \int_0^1 \begin{pmatrix} t^2 y^2 \cos(tx) \\ 2ty \sin(tx) + e^{2tz} \\ 2tye^{2tz} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} dt \\ &= \int_0^1 \left\{ \underbrace{t^2 y^2 x \cos(tx)}_{= \frac{d}{dt} \sin(tx)} + 2ty^2 \sin(tx) + ye^{2tz} + \underbrace{2tyze^{2tz}}_{= yt \frac{d}{dt} (e^{2tz})} \right\} dt \end{aligned}$$

$$\begin{aligned}
&= t^2 y^2 \sin(tx) \Big|_{t=0}^{t=1} - \int_0^1 2ty^2 \sin(tx) dt + \int_0^1 2ty^2 \sin(tx) dt \\
&\quad + \int_0^1 ye^{2tz} dt + yte^{2tz} \Big|_{t=0}^{t=1} - \int_0^1 ye^{2tz} dt \\
&= y^2 \sin x + ye^{2z}.
\end{aligned}$$

Die **Berechnung mittels Ansatz** erfolgt in sechs Schritten:

1. SCHRITT: Ansatz: $v_1 = f_x$, $v_2 = f_y$, $v_3 = f_z$.
2. SCHRITT: Unbestimmte Integration bzgl. x (bei fest gehaltenem y , z) der Gleichung $f_x = v_1$:

$$f(x, y, z) = \int v_1(x, y, z) dx + c(y, z).$$

3. SCHRITT. Leite die Identität aus Schritt 2 partiell nach y ab und setze das Ergebnis in die Gleichung $v_2 = f_y$ ein. Dies liefert eine Gleichung für c_y :

$$\frac{\partial}{\partial y} \int v_1(x, y, z) dx + c_y(y, z) = f_y = v_2.$$

Tritt in dieser Gleichung noch die Variable x auf, ist \mathbf{v} kein Gradientenfeld, und wir sind fertig. Andernfalls gehe zu Schritt 4.

4. SCHRITT: Unbestimmte Integration bzgl. y der Gleichung für c_y :

$$c(y, z) = \int h(y, z) dy + d(z)$$

mit

$$h(y, z) = v_2(x, y, z) - \int \frac{\partial}{\partial y} v_1(x, y, z) dx.$$

5. SCHRITT: Setze c aus Schritt 4 in f aus Schritt 2 ein, leite das Ganze partiell nach z ab und setze das Ergebnis in die Gleichung $v_3 = f_z$ ein. Dies liefert eine Gleichung für $d'(z)$:

$$\frac{\partial}{\partial z} \int v_1(x, y, z) dx + \frac{\partial}{\partial z} \int h(y, z) dy + d'(z) = f_z = v_3.$$

Tritt in dieser Gleichung noch die Variable y auf, ist \mathbf{v} kein Gradientenfeld, und wir sind fertig. Andernfalls gehe zu Schritt 6.

6. SCHRITT: Bestimme $d(z)$ durch unbestimmte Integration bzgl. z der Identität für $d'(z)$ aus Schritt 5 und setze das Ergebnis in die Gleichungen für c und f ein.

BEISPIEL IX.2.14. Wir betrachten das Vektorfeld \mathbf{v} aus Beispiel IX.2.13. Schritt 2 liefert

$$\begin{aligned} f(x, y, z) &= \int y^2 \cos x dx + c(y, z) \\ &= y^2 \sin x + c(y, z). \end{aligned}$$

Die Gleichung aus Schritt 3 lautet

$$\begin{aligned} 2y \sin x + e^{2z} &= v_2(x, y, z) \\ &= \frac{\partial}{\partial y} [y^2 \sin x + c(y, z)] \\ &= 2y \sin x + c_y(y, z) \\ \implies c_y(y, z) &= e^{2z}. \end{aligned}$$

Die Gleichung hängt nicht von x ab. Die Funktion h aus Schritt 4 ist

$$\begin{aligned} h(y, z) &= 2y \sin x + e^{2z} - \int 2y \cos x dx \\ &= e^{2z} \end{aligned}$$

und wir erhalten

$$\begin{aligned} c(y, z) &= \int e^{2z} dy + d(z) \\ &= ye^{2z} + d(z) \\ \implies f(x, y, z) &= y^2 \sin x + ye^{2z} + d(z). \end{aligned}$$

Die Gleichung aus Schritt 5 lautet

$$\begin{aligned} 2ye^{2z} &= v_3(x, y, z) \\ &= \frac{\partial}{\partial z} [y^2 \sin x + ye^{2z} + d(z)] \\ &= 2ye^{2z} + d'(z) \\ \implies d'(z) &= 0. \end{aligned}$$

Sie hängt nicht von y ab. Schritt 6 liefert $d(z) = \gamma$ mit beliebiger Konstante γ . Daher ist

$$f(x, y, z) = y^2 \sin x + ye^{2z} + \gamma.$$

Man beachte, dass in Beispiel IX.2.13 $\gamma = 0$ war, da wir durch die Wahl von $\mathbf{x}_0 = 0$ die Festlegung $f(0, 0, 0) = 0$ getroffen hatten.

BEISPIEL IX.2.15 (ZENTRALFELDER). Dies sind Vektorfelder der Form

$$\mathbf{v}(\mathbf{x}) = \varphi(|\mathbf{x} - \mathbf{z}|)(\mathbf{x} - \mathbf{z})$$

auf $D = \mathbb{R}^3 \setminus \{z\}$ mit einem festen Punkt $\mathbf{z} \in \mathbb{R}^3$. (Man beachte, dass der Wirbel aus Beispiel IX.2.6 (S. 334) nicht unter dieses Beispiel fällt, da er in \mathbb{R}^2 definiert ist.) Da $D = \mathbb{R}^3 \setminus \{z\}$ nicht sternförmig ist, können wir den Satz von Poincaré nicht anwenden. Dennoch ist ein solches

Zentralfeld immer ein Gradientenfeld. Zur Bestimmung einer Stammfunktion machen wir den Ansatz

$$f(\mathbf{x}) = \psi(|\mathbf{x} - \mathbf{z}|).$$

Mit der Kettenregel folgt für $\mathbf{x} \neq \mathbf{z}$

$$\text{grad } f(\mathbf{x}) = \psi'(|\mathbf{x} - \mathbf{z}|) \frac{1}{|\mathbf{x} - \mathbf{z}|} (\mathbf{x} - \mathbf{z}).$$

Ein Vergleich mit \mathbf{v} führt auf die Identität

$$\varphi(|\mathbf{x} - \mathbf{z}|) = \frac{1}{|\mathbf{x} - \mathbf{z}|} \psi'(|\mathbf{x} - \mathbf{z}|) \quad \text{für } \mathbf{x} \neq \mathbf{z}$$

bzw.

$$\psi'(s) = s\varphi(s) \quad \text{für } s > 0.$$

Also ist

$$\psi(s) = \int_0^s t\varphi(t)dt + c \quad \text{für } s > 0$$

und

$$\mathbf{v}(\mathbf{x}) = \int_0^{|\mathbf{x}-\mathbf{z}|} t\varphi(t)dt + c.$$

So erhalten wir z.B. für $n \geq 3$ und das Feld

$$\mathbf{v}(\mathbf{x}) = \frac{1}{|\mathbf{x}|^n} \mathbf{x}, \quad \mathbf{x} \neq 0$$

die Stammfunktion

$$\begin{aligned} f(\mathbf{x}) &= \int_0^{|\mathbf{x}|} t \frac{1}{t^n} dt + c \\ &= -\frac{1}{n-2} \frac{1}{|\mathbf{x}|^{n-2}} + c. \end{aligned}$$

IX.3. Integration über ebene Bereiche

IX.3.1. Der Flächeninhalt. Wir wollen einer beschränkten Menge $M \subset \mathbb{R}^2$ einen Flächeninhalt $F(M)$ zuordnen. Falls M ein Rechteck mit Kantenlängen ℓ_x und ℓ_y ist, ist natürlich $F(M) = \ell_x \ell_y$. Falls M die Vereinigung von endlich vielen Rechtecken R_1, \dots, R_n ist, deren Inneres paarweise disjunkt ist, ist natürlich $F(M) = F(R_1) + \dots + F(R_n)$.

Diese Beobachtung legt nun folgendes Vorgehen nahe: Für $k \in \mathbb{N}$ zerlegen wir die Ebene durch achsenparallele Geraden $x = n2^{-k}$ und $y = n2^{-k}$ mit $n \in \mathbb{Z}$ in Quadrate der Kantenlänge 2^{-k} und Fläche 2^{-2k} . Dann bezeichnen wir mit $s_k(M)$ die Summe der Flächen aller Quadrate, die ganz in M liegen, und mit $S_k(M)$ die Summe der Flächen

aller Quadrate, die einen nicht leeren Durchschnitt mit M haben (vgl. Abb. IX.3.1). Offensichtlich gilt

$$\begin{aligned} s_k(M) &\leq S_k(M), \\ s_k(M) &\leq s_{k+1}(M), \\ S_{k+1}(M) &\leq S_k(M). \end{aligned}$$

Daher konvergieren die Folgen $(s_k(M))_{k \in \mathbb{N}}$ und $(S_k(M))_{k \in \mathbb{N}}$ gegen Zahlen $F_i(M)$ bzw. $F_a(M)$:

$$\begin{aligned} F_i(M) &= \lim_{k \rightarrow \infty} s_k(M), \\ F_a(M) &= \lim_{k \rightarrow \infty} S_k(M). \end{aligned}$$

Es ist immer $F_i(M) \leq F_a(M)$.

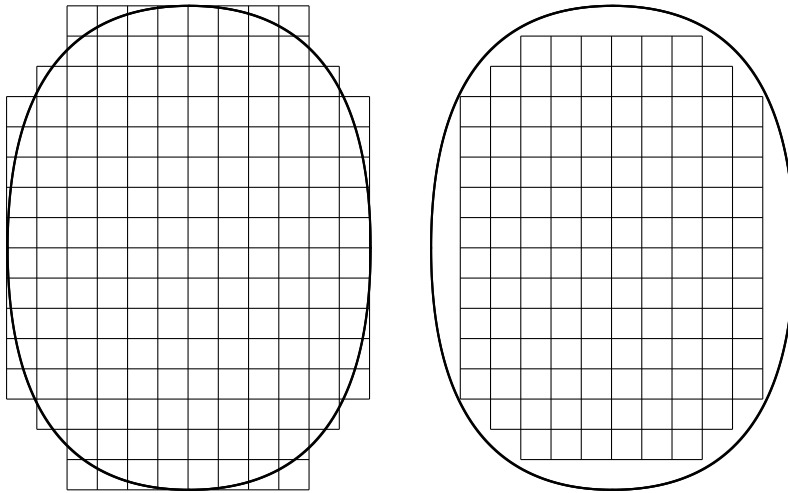


ABBILDUNG IX.3.1. $S_k(M)$ ist die Summe der Flächen der Quadrate im linken Bild, $s_k(M)$ ist die Summe der Flächen der Quadrate im rechten Bild

Wir nennen M RIEMANN-MESSBAR, wenn $F_i(M) = F_a(M)$ ist. In diesem Fall heißt der gemeinsame Wert der FLÄCHENINHALT von M und wird mit $F(M)$ bezeichnet.

Nicht jede Menge ist Riemann-messbar:

BEISPIEL IX.3.1. Betrachte $M = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1, x, y \in \mathbb{Q}\}$. Dann gilt für alle $k \in \mathbb{N}$ $s_k(M) = 0$, $S_k(M) = 1$ und damit $F_i(M) = 0$, $F_a(M) = 1$.

Es gilt aber:

Jede beschränkte Menge, deren Rand aus endlich vielen regulären Kurvenstücken besteht, ist Riemann-messbar.

Eine messbare Menge N mit $F(N) = 0$ heißt NULLMENGE.

BEISPIEL IX.3.2. Jede einpunktige Menge ist eine Nullmenge. Jedes reguläre Kurvenstück ist eine Nullmenge.

Ist M messbar und N eine Nullmenge, so sind $M \cup N$ und $M \setminus N$ ebenfalls messbar und es gilt

$$\begin{aligned} F(M \cup N) &= F(M), \\ F(M \setminus N) &= F(M). \end{aligned}$$

BEISPIEL IX.3.3. $M = \{(x, y) : x^2 + y^2 < 1\}$, $N = \{(0, 0)\} \cup \{(x, y) : x^2 + y^2 = 1\}$, $M \cup N = \{(x, y) : x^2 + y^2 \leq 1\}$, $M \setminus N = \{(x, y) : 0 < x^2 + y^2 < 1\}$ und $F(M) = F(M \cup N) = F(M \setminus N) = \pi$.

IX.3.2. Das Doppelintegral. Im Folgenden betrachten wir nur Mengen $M \subset \mathbb{R}^2$ mit folgender Eigenschaft: Es gibt eine Nullmenge N , so dass gilt:

- $M \cup N$ ist beschränkt und abgeschlossen.
- Je zwei Punkte im Innern von $M \cup N$ können durch ein reguläres Kurvenstück verbunden werden, das ganz im Innern von $M \cup N$ verläuft.
- Der Rand von $M \cup N$ besteht aus endlich vielen regulären Kurvenstücken.

Mengen mit dieser Eigenschaft sind gemäß vorigem Abschnitt Riemann-messbar.

Wir wollen einer auf M stetigen Funktion $f : M \rightarrow \mathbb{R}$ ein Integral zuordnen. Insbesondere soll das Integral der konstanten Funktion 1 der Flächeninhalt von M sein. Dazu zerlegen wir die Ebene wieder in achsenparallele Quadrate der Kantenlänge 2^{-k} mit $k \in \mathbb{N}$. Die Quadrate, die ganz in $M \cup N$ liegen, nummerieren wir von 1 bis n_k . In jedem dieser Quadrate Q_i wählen wir einen Punkt $(x_{i,k}, y_{i,k})$ und setzen

$$Z_k = \sum_{i=1}^{n_k} f(x_{i,k}, y_{i,k}) F(Q_i).$$

Man kann zeigen, dass die Folge $(Z_k)_{k \in \mathbb{N}}$ konvergiert. Diesen Grenzwert nennen wir das **DOPPELINTEGRAL** (oder **GEBIETSINTEGRAL** oder einfach **INTEGRAL**) von f über M und bezeichnen ihn mit

$$\iint_M f(x, y) dF \quad \text{oder} \quad \iint_M f dF.$$

Das Doppelintegral hat folgende

GEOMETRISCHE DEUTUNG:

- $\iint_M dF$ ist der Flächeninhalt $F(M)$ von M .

- Ist $f(x, y) \geq 0$ für alle $(x, y) \in M$, so ist $\iint_M f dF$ das Volumen des senkrecht auf der (x, y) -Ebene stehenden Zylinders mit Grundfläche M und Deckfläche $z = f(x, y)$.

Es gelten folgende

RECHENREGELN:

- (1) $\iint_M (\alpha f + \beta g) dF = \alpha \iint_M f dF + \beta \iint_M g dF$
für $\alpha, \beta \in \mathbb{R}$, $f, g : M \rightarrow \mathbb{R}$.
- (2) $f(x, y) \leq g(x, y)$ für alle $(x, y) \in M$
 $\implies \iint_M f dF \leq \iint_M g dF$.
- (3) $\iint_M f dF = \iint_{M_1} f dF + \iint_{M_2} f dF$,
falls M durch ein reguläres Kurvenstück in die zwei Bereiche M_1 und M_2 zerlegt wird.

Zudem gilt der

MITTELWERTSATZ: Es gibt einen Punkt $(x^*, y^*) \in M$ mit

$$\iint_M f dF = f(x^*, y^*) F(M).$$

IX.3.3. Praktische Berechnung des Doppelintegrals. Für die praktische Berechnung des Doppelintegrals spielen folgende zwei Typen von Mengen eine besondere Rolle (vgl. Abb. IX.3.1):

Typ I: Es gibt ein Intervall $[a, b]$ in \mathbb{R} und zwei stetige Funktionen $g, h : [a, b] \rightarrow \mathbb{R}$ mit $g(x) \leq h(x)$ für alle $x \in [a, b]$ und $M = \{(x, y) : a \leq x \leq b, g(x) \leq y \leq h(x)\}$.

Typ II: Es gibt ein Intervall $[c, d]$ in \mathbb{R} und zwei stetige Funktionen $\ell, r : [c, d] \rightarrow \mathbb{R}$ mit $\ell(y) \leq r(y)$ für alle $y \in [c, d]$ und $M = \{(x, y) : c \leq y \leq d, \ell(y) \leq x \leq r(y)\}$.

Dann gilt für jede auf M stetige Funktion f :

M ist vom Typ I

$$\implies \iint_M f(x, y) dF = \int_a^b \left\{ \int_{g(x)}^{h(x)} f(x, y) dy \right\} dx$$

M ist vom Typ II

$$\implies \iint_M f(x, y) dF = \int_c^d \left\{ \int_{\ell(y)}^{r(y)} f(x, y) dx \right\} dy$$

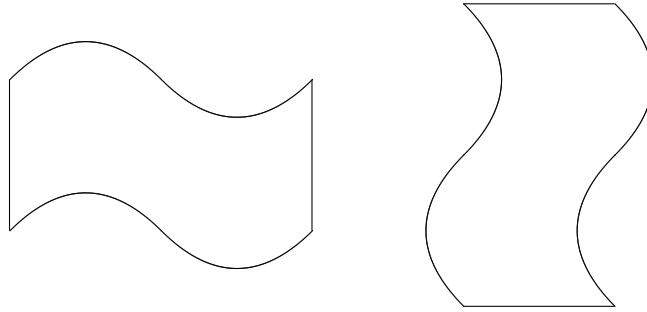


ABBILDUNG IX.3.2. Menge vom Typ I links und Menge vom Typ II rechts

BEISPIEL IX.3.4. Sei M der Viertelkreis im ersten Quadranten mit Mittelpunkt im Ursprung und Radius 1

$$M = \{(x, y); x \geq 0, y \geq 0, x^2 + y^2 \leq 1\}.$$

Die Koordinaten (x_S, y_S) des Schwerpunktes von M sind gegeben durch

$$x_S = \frac{1}{F(M)} \iint_M x dF,$$

$$y_S = \frac{1}{F(M)} \iint_M y dF.$$

M ist eine Menge vom Typ I und II:

$$M = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq \sqrt{1-x^2}\}$$

$$= \{(x, y) : 0 \leq y \leq 1, 0 \leq x \leq \sqrt{1-y^2}\}.$$

Die erste Darstellung ist günstig für die Berechnung von y_S , die zweite für diejenige von x_S :

$$y_S = \frac{1}{\frac{1}{4}\pi} \int_0^1 \left\{ \int_0^{\sqrt{1-x^2}} y dy \right\} dx$$

$$= \frac{4}{\pi} \int_0^1 \frac{1}{2} (1-x^2) dx$$

$$= \frac{2}{\pi} \left(1 - \frac{1}{3}\right)$$

$$= \frac{4}{3\pi},$$

$$\begin{aligned}
 x_S &= \frac{1}{\frac{1}{4}\pi} \int_0^1 \left\{ \int_0^{\sqrt{1-y^2}} x dx \right\} dy \\
 &= \frac{4}{\pi} \int_0^1 \frac{1}{2}(1-y^2) dy \\
 &= \frac{4}{3\pi}.
 \end{aligned}$$

Kompliziertere Mengen zerlegt man durch achsenparallele Schnitte in endlich viele Mengen M_1, \dots, M_n vom Typ I oder II. Dann ist

$$\iint_M f dF = \iint_{M_1} f dF + \dots + \iint_{M_n} f dF$$

und jedes einzelne Integral kann wie in Beispiel [IX.3.4](#) berechnet werden.

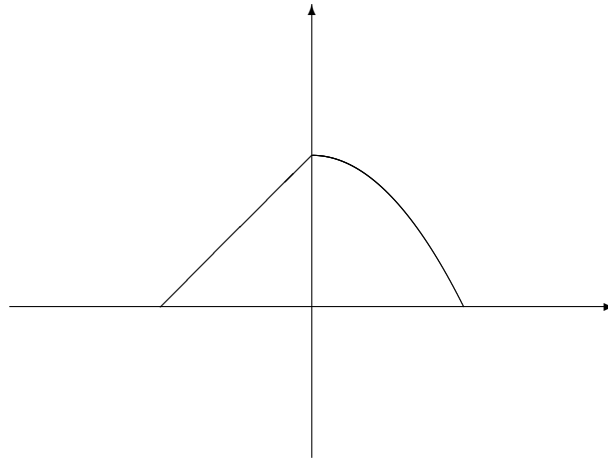


ABBILDUNG IX.3.3. Menge M aus Beispiel [IX.3.5](#)

BEISPIEL IX.3.5. Gesucht ist

$$\iint_M f dF$$

mit

$$f(x, y) = x^2 y$$

und

$$\begin{aligned}
 M &= \{(x, y) : x \geq -1, 0 \leq y \leq 1 + x\} \\
 &\quad \cap \{(x, y) : -1 \leq x \leq 1, y \leq 1 - x^2\}
 \end{aligned}$$

(vgl. Abb. [IX.3.3](#)). Offensichtlich ist $M = M_1 \cup M_2$ mit

$$M_1 = \{(x, y) : -1 \leq x \leq 0, 0 \leq y \leq 1 + x\},$$

$$M_2 = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1 - x^2\}.$$

Damit folgt

$$\begin{aligned}
 \iint_M f dF &= \int_{-1}^0 \left\{ \int_0^{1+x} x^2 y dy \right\} dx + \int_0^1 \left\{ \int_0^{1-x^2} x^2 y dy \right\} dx \\
 &= \int_{-1}^0 \frac{1}{2} x^2 (1+x)^2 dx + \int_0^1 \frac{1}{2} x^2 (1-x^2)^2 dx \\
 &= \frac{1}{2} \int_{-1}^0 (x^2 + 2x^3 + x^4) dx + \frac{1}{2} \int_0^1 (x^2 - 2x^4 + x^6) dx \\
 &= \frac{1}{2} \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{5} \right) + \frac{1}{2} \left(\frac{1}{3} - \frac{2}{5} + \frac{1}{7} \right) \\
 &= \frac{1}{60} + \frac{4}{105} \\
 &= \frac{23}{420}.
 \end{aligned}$$

BEISPIEL IX.3.6. Der MASSENMITTELPUNKT (oder auch SCHWERPUNKT) (x_S, y_S) eines Flächenstückes M mit der Massendichte $\mu(x, y)$ ist gegeben durch

$$\begin{aligned}
 x_S &= \frac{1}{m} \iint_M x \mu(x, y) dF, \\
 y_S &= \frac{1}{m} \iint_M y \mu(x, y) dF, \\
 m &= \iint_M \mu(x, y) dF.
 \end{aligned}$$

Für $M = \{(x, y) : 1 \leq x \leq 2, 1 \leq y \leq x^2\}$ und $\mu(x, y) = x^2 + y^2$ erhalten wir z.B.

$$\begin{aligned}
 m &= \int_1^2 \left\{ \int_1^{x^2} (x^2 + y^2) dy \right\} dx \\
 &= \int_1^2 \left(yx^2 + \frac{1}{3}y^3 \right) \Big|_{y=1}^{y=x^2} dx \\
 &= \int_1^2 \left(x^4 + \frac{1}{3}x^6 - x^2 - \frac{1}{3} \right) dx \\
 &= \left(\frac{1}{5}x^5 + \frac{1}{21}x^7 - \frac{1}{3}x^3 - \frac{1}{3}x \right) \Big|_{x=1}^{x=2} \\
 &= \frac{32}{5} + \frac{128}{21} - \frac{8}{3} - \frac{2}{3} - \frac{1}{5} - \frac{1}{21} + \frac{1}{3} + \frac{1}{3} \\
 &= \frac{31}{5} + \frac{127}{21} - \frac{8}{3} \\
 &= \frac{31}{5} + \frac{71}{21}
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1006}{105} \\
x_S &= \frac{105}{1006} \int_1^2 \left\{ \int_1^{x^2} x(x^2 + y^2) dy \right\} dx \\
&= \frac{105}{1006} \int_1^2 \left(yx^3 + \frac{1}{3}y^3x \right) \Big|_{y=1}^{y=x^2} dx \\
&= \frac{105}{1006} \int_1^2 \left(x^5 + \frac{1}{3}x^7 - x^3 - \frac{1}{3}x \right) dx \\
&= \frac{105}{1006} \left(\frac{1}{6}x^6 + \frac{1}{24}x^8 - \frac{1}{4}x^4 - \frac{1}{6}x^2 \right) \Big|_{x=1}^{x=2} \\
&= \frac{105}{1006} \left(\frac{63}{6} + \frac{255}{24} - \frac{15}{4} - \frac{3}{6} \right) \\
&= \frac{105}{1006} \left(10 + \frac{165}{24} \right) \\
&= \frac{105}{1006} \cdot \frac{80 + 55}{8} \\
&= \frac{14175}{8048} \\
y_S &= \frac{105}{1006} \int_1^2 \left\{ \int_1^{x^2} y(x^2 + y^2) dy \right\} dx \\
&= \frac{105}{1006} \int_1^2 \left(\frac{1}{2}x^2y^2 + \frac{1}{4}y^4 \right) \Big|_{y=1}^{y=x^2} dx \\
&= \frac{105}{1006} \int_1^2 \left(\frac{1}{2}x^6 + \frac{1}{4}x^8 - \frac{1}{2}x^2 - \frac{1}{4} \right) dx \\
&= \frac{105}{1006} \left(\frac{1}{14}x^7 + \frac{1}{36}x^9 - \frac{1}{6}x^3 - \frac{1}{4}x \right) \Big|_{x=1}^{x=2} \\
&= \frac{105}{1006} \left(\frac{127}{14} + \frac{511}{36} - \frac{7}{6} - \frac{1}{4} \right) \\
&= \frac{105}{1006} \left(\frac{127}{14} + \frac{460}{36} \right) \\
&= \frac{105}{1006} \cdot \frac{1143 + 1610}{126} \\
&= \frac{289065}{126756}.
\end{aligned}$$

IX.3.4. Der Satz von Green. Wir betrachten eine beschränkte und abgeschlossene Menge M in \mathbb{R}^2 , deren Rand aus endlich vielen regulären Kurvenstücken besteht. Die Kurvenstücke seien so parametrisiert, dass M immer links zur Durchlaufrichtung liegt (positiver Umlauf). Ist $\mathbf{v} : M \rightarrow \mathbb{R}^2$ ein Vektorfeld, so verstehen wir unter

$$\int_{\partial M} \mathbf{v} \cdot d\mathbf{x}$$

die Summe der Kurvenintegrale über die Kurvenstücke, die den Rand ∂M von M bilden. Ist das Vektorfeld \mathbf{v} stetig differenzierbar, gilt der

SATZ VON GREEN:

$$\int_{\partial M} \mathbf{v} \cdot d\mathbf{x} = \iint_M \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dF.$$

BEISPIEL IX.3.7. Wählen wir $\mathbf{v} = (0, x)^T$ oder $\mathbf{v} = (-y, 0)^T$, ergibt sich jeweils der Flächeninhalt von M :

$$F(M) = \iint_M dF = \int_{\partial M} x dy = - \int_{\partial M} y dx.$$

Ist z.B. M der Bereich, der durch die Zykloide

$$\mathbf{w}_1(t) = a \begin{pmatrix} 2\pi - t + \sin t \\ 1 - \cos t \end{pmatrix}, \quad 0 \leq t \leq 2\pi$$

und die Strecke

$$\mathbf{w}_2(t) = \begin{pmatrix} t \\ 0 \end{pmatrix}, \quad 0 \leq t \leq 2\pi a$$

begrenzt wird, ergibt sich

$$\begin{aligned} F(M) &= - \int_{\partial M} y dx \\ &= - \int_{\mathbf{w}_1} y dx - \underbrace{\int_{\mathbf{w}_2} y dx}_{=0} \\ &= - \int_0^{2\pi} a(1 - \cos t) a(-1 + \cos t) dt \\ &= \int_0^{2\pi} a^2 (1 - \cos t)^2 dt \\ &= a^2 \int_0^{2\pi} \left(1 - \underbrace{2 \cos t}_{f \dots = 0} + \underbrace{\cos^2 t}_{f \dots = \pi} \right) dt \\ &= 3\pi a^2. \end{aligned}$$

BEISPIEL IX.3.8. Wählen wir $\mathbf{v} = (0, \frac{1}{2}x^2)^T$ oder $\mathbf{v} = (-\frac{1}{2}y^2, 0)^T$, erhalten wir für die Komponenten des Schwerpunktes von M :

$$x_S = \frac{1}{F(M)} \iint_M x dF = \frac{1}{2F(M)} \int_{\partial M} x^2 dy,$$

$$y_S = \frac{1}{F(M)} \iint_M y dF = -\frac{1}{2F(M)} \int_{\partial M} y^2 dx.$$

Für die Menge M aus Beispiel IX.3.7 ergibt sich so z.B.

$$\begin{aligned} y_S &= -\frac{1}{6\pi a^2} \int_{\partial M} y^2 dx \\ &= -\frac{1}{6\pi a^2} \int_{\mathbf{w}_1} y^2 dx - \underbrace{\frac{1}{6\pi a^2} \int_{\mathbf{w}_2} y^2 dx}_{=0} \\ &= -\frac{1}{6\pi a^2} \int_0^{2\pi} a^2(1 - \cos t)^2 a(-1 + \cos t) dt \\ &= \frac{a}{6\pi} \int_0^{2\pi} (1 - \cos t)^3 dt \\ &= \frac{a}{6\pi} \int_0^{2\pi} (1 - \underbrace{3\cos t}_{f \dots = 0} + \underbrace{3\cos^2 t}_{f \dots = 3\pi} - \underbrace{\cos^3 t}_{f \dots = 0}) dt \\ &= \frac{a}{6\pi} 5\pi \\ &= \frac{5}{6} a \end{aligned}$$

und wegen der Symmetrie der Integranden

$$\begin{aligned} x_S &= \frac{1}{6\pi a^2} \int_{\partial M} x^2 dy \\ &= \frac{1}{6\pi a^2} \int_{\mathbf{w}_1} x^2 dy + \underbrace{\frac{1}{6\pi a^2} \int_{\mathbf{w}_2} x^2 dy}_{=0} \\ &= \frac{1}{6\pi a^2} \int_0^{2\pi} a^2(2\pi - t + \sin t)^2 a \sin t dt \quad \left| t = s + \pi \right. \\ &= -\frac{a}{6\pi} \int_{-\pi}^{\pi} (\pi - s - \sin s)^2 \sin s ds \\ &= -\frac{a}{6\pi} \int_{-\pi}^{\pi} (\underbrace{\pi^2 \sin s}_{f \dots = 0} + \underbrace{s^2 \sin s}_{f \dots = 0} + \underbrace{\sin^3 s}_{f \dots = 0} \\ &\quad - 2\pi \underbrace{s \sin s}_{f \dots = 2\pi} - 2\pi \underbrace{\sin^2 s}_{f \dots = \pi} + 2 \underbrace{s \sin^2 s}_{f \dots = 0}) ds \\ &= \pi a. \end{aligned}$$

BEISPIEL IX.3.9 (SATZ VON GAUSS IN DER EBENE). Ist $u : M \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, können wir

$$\mathbf{v} = (-u_y, u_x)^T$$

wählen. Dann ist

$$\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} = u_{xx} + u_{yy} = \Delta u$$

und

$$\mathbf{v} \cdot d\mathbf{x} = \text{grad } u \cdot \mathbf{n} ds,$$

wenn \mathbf{n} der nach außen zeigende Einheitsnormalenvektor von ∂M ist. Damit ergibt sich der Satz von Gauß in der Ebene

$$\iint_M \Delta u dF = \int_{\partial M} \text{grad } u \cdot \mathbf{n} ds.$$

IX.4. Integration über Flächen im Raum

IX.4.1. Parameterdarstellungen. Neben der Darstellung von Flächen als Graphen $z = h(x, y)$ oder Niveauflächen $F(x, y, z) = \text{const}$ von Funktionen gibt es die Parameterdarstellung. Sie ist für die Integration von auf Flächen definierten Funktionen besonders vorteilhaft. Ihr liegt die Vorstellung zugrunde, dass die Fläche durch Verformung eines Stückes der Ebene entstanden ist (vgl. Abb. IX.4.1).

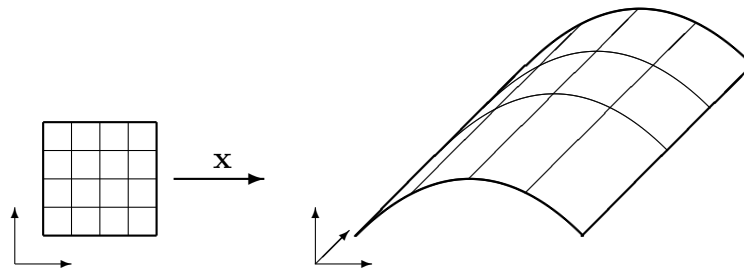


ABBILDUNG IX.4.1. Parameterdarstellung einer Fläche mit Parameterlinien

DEFINITION IX.4.1. Eine PARAMETERDARSTELLUNG eines REGULÄREN FLÄCHENSTÜCKES S im Raum ist gegeben durch eine nicht leere, offene Menge $G \subset \mathbb{R}^2$, eine nicht leere, abgeschlossene Menge $D \subset G$ und die Einschränkung auf D einer stetig differenzierbaren Abbildung

$$\mathbf{x} : G \rightarrow \mathbb{R}^3$$

$$(u, v) \mapsto \mathbf{x}(u, v) = (x(u, v), y(u, v), z(u, v))^T$$

mit folgenden Eigenschaften:

- (1) \mathbf{x} ist injektiv auf D , d.h. $\mathbf{x}(u, v) \neq \mathbf{x}(u', v')$ für alle $(u, v) \neq (u', v')$ in D .
 (2) $\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v) \neq 0$ für alle $(u, v) \in D$.

Die Bedingung (2) besagt, dass die Jacobi-Matrix

$$D\mathbf{x}(u, v) = (\mathbf{x}_u(u, v), \mathbf{x}_v(u, v))$$

für alle $(u, v) \in D$ den Rang 2 hat.

Variiert man nur einen der beiden Parameter u und v und hält den anderen fest, erhält man Kurven auf S , die sog. PARAMETERLINIEN. Die entsprechenden partiellen Ableitungen von \mathbf{x} sind die Tangentenvektoren an die Parameterlinien. Bedingung (2) besagt, dass diese Tangentenvektoren in jedem Punkt von S linear unabhängig sind. Sie spannen daher eine Ebene auf, die sog. TANGENTIALEBENE. Der Vektor $\mathbf{x}_u \times \mathbf{x}_v$ steht senkrecht auf der Tangentialebene. Der normierte Vektor

$$\mathbf{n} = \frac{1}{|\mathbf{x}_u \times \mathbf{x}_v|} \mathbf{x}_u \times \mathbf{x}_v$$

heißt daher der NORMALENVEKTOR zu S in dem entsprechenden Flächenpunkt.

Aus der Identität

$$|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$$

(vgl. Abschnitt I.4.7 (S. 31)) folgt

$ \mathbf{x}_u \times \mathbf{x}_v ^2 = EG - F^2$ <p>mit</p> $E = \mathbf{x}_u \cdot \mathbf{x}_u,$ $G = \mathbf{x}_v \cdot \mathbf{x}_v,$ $F = \mathbf{x}_u \cdot \mathbf{x}_v.$

Die Größen E, F, G heißen die METRISCHEN FUNDAMENTALGRÖSSEN von S . Insbesondere ist F genau dann gleich Null, wenn sich die Parameterlinien in einem rechten Winkel schneiden.

Der RAND ∂S des Flächenstückes S ist das Bild des Randes ∂D von D unter \mathbf{x} :

$$\partial S = \{\mathbf{x}(u, v) : (u, v) \in \partial D\}.$$

Die Oberfläche der in den technischen Anwendungen auftretenden Körper kann in der Regel nicht durch ein einziges reguläres Flächenstück dargestellt werden. Man denke etwa an einen Würfel oder eine Konservendose. Vielmehr ist die Oberfläche die Vereinigung von endlich

vielen regulären Flächenstücken, von denen je zwei an einem oder mehreren Randstücken zusammenstoßen, aber sonst keinen Punkt gemeinsam haben. Wir nennen eine solche Fläche S eine **STÜCKWEISE REGULÄRE FLÄCHE**. Ihr **RAND** besteht aus allen Randstücken, die nur zu einem Flächenstück gehören. Die Fläche heißt **GESCHLOSSEN**, wenn dieser Rand leer ist.

BEISPIEL IX.4.2 (EBENEN). Für $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ mit $\mathbf{a} \times \mathbf{b} \neq 0$ ist

$$\mathbf{x}(u, v) = \mathbf{x}_0 + u\mathbf{a} + v\mathbf{b}$$

eine Parameterdarstellung der Ebene durch den Punkt \mathbf{x}_0 , die von \mathbf{a} und \mathbf{b} aufgespannt wird. Es ist

$$\begin{aligned} \mathbf{x}_u &= \mathbf{a}, & \mathbf{x}_v &= \mathbf{b}, \\ E &= |\mathbf{a}|^2, & G &= |\mathbf{b}|^2, & F &= \mathbf{a} \cdot \mathbf{b}. \end{aligned}$$

BEISPIEL IX.4.3 (GRAPHEN). Für den Graphen $z = h(x, y)$ mit $h : D \rightarrow \mathbb{R}$ erhält man die Parameterdarstellung

$$\mathbf{x}(u, v) = \begin{pmatrix} u \\ v \\ h(u, v) \end{pmatrix}.$$

Die Parameterlinien sind die Schnitte mit den Ebenen $x = \text{const}$ bzw. $y = \text{const}$. Es ist

$$\begin{aligned} \mathbf{x}_u &= \begin{pmatrix} 1 \\ 0 \\ h_x \end{pmatrix}, & \mathbf{x}_v &= \begin{pmatrix} 0 \\ 1 \\ h_y \end{pmatrix}, & \mathbf{x}_u \times \mathbf{x}_v &= \begin{pmatrix} -h_x \\ -h_y \\ 1 \end{pmatrix}, \\ E &= 1 + h_x^2, & G &= 1 + h_y^2, & F &= h_x h_y. \end{aligned}$$

BEISPIEL IX.4.4 (DREHFLÄCHEN). Aus einem regulären Kurvenstück

$$t \mapsto \begin{pmatrix} x(t) \\ 0 \\ z(t) \end{pmatrix}, \quad t_0 \leq t \leq t_1,$$

mit $x(t) > 0$ für alle t entsteht durch Drehung um die z -Achse um den Winkel φ_0 , $0 < \varphi_0 < 2\pi$, ein reguläres Flächenstück. Eine Parameterdarstellung ist

$$\mathbf{x}(t, \varphi) = \begin{pmatrix} x(t) \cos \varphi \\ x(t) \sin \varphi \\ z(t) \end{pmatrix}.$$

Die Parameterlinien $t = \text{const}$ sind die Breitenkreise, die $\varphi = \text{const}$ die Meridiane. Es ist

$$\mathbf{x}_t = \begin{pmatrix} \dot{x}(t) \cos \varphi \\ \dot{x}(t) \sin \varphi \\ \dot{z}(t) \end{pmatrix},$$

$$\begin{aligned}\mathbf{x}_\varphi &= \begin{pmatrix} -x(t) \sin \varphi \\ x(t) \cos \varphi \\ 0 \end{pmatrix}, \\ \mathbf{x}_t \times \mathbf{x}_\varphi &= \begin{pmatrix} -x(t)\dot{z}(t) \cos \varphi \\ -x(t)\dot{z}(t) \sin \varphi \\ x(t)\dot{x}(t) \end{pmatrix} \\ E &= \dot{x}(t)^2 + \dot{z}(t)^2, \\ G &= x(t)^2, \\ F &= 0.\end{aligned}$$

Da stets $F = 0$ ist, bilden die Parameterlinien ein orthogonales Netz. Da bei einer vollen Drehung, die Meridiane $\varphi = 0$ und $\varphi = 2\pi$ aufeinanderfallen, ist die gesamte Drehfläche nur stückweise regulär; man nehme z.B. die beiden zu $0 \leq \varphi \leq \pi$ und $\pi \leq \varphi \leq 2\pi$ gehörenden Flächenstücke.

BEISPIEL IX.4.5 (SPHÄRE). Eine Parameterdarstellung der Sphäre (Kugeloberfläche) mit Mittelpunkt im Ursprung und Radius r ist

$$\mathbf{x}(\varphi, \theta) = \begin{pmatrix} r \cos \varphi \cos \theta \\ r \sin \varphi \cos \theta \\ r \sin \theta \end{pmatrix}, \quad 0 \leq \varphi \leq 2\pi, \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}.$$

Es ist

$$\begin{aligned}\mathbf{x}_\varphi &= \begin{pmatrix} -r \sin \varphi \cos \theta \\ r \cos \varphi \cos \theta \\ 0 \end{pmatrix}, \\ \mathbf{x}_\theta &= \begin{pmatrix} -r \cos \varphi \sin \theta \\ -r \sin \varphi \sin \theta \\ r \cos \theta \end{pmatrix}, \\ \mathbf{x}_\varphi \times \mathbf{x}_\theta &= \begin{pmatrix} r^2 \cos \varphi \cos^2 \theta \\ r^2 \sin \varphi \cos^2 \theta \\ r^2 \sin \theta \cos \theta \end{pmatrix}, \\ E &= r^2 \cos^2 \theta, \\ G &= r^2, \\ F &= 0.\end{aligned}$$

In den Polen $\theta = \pm \frac{\pi}{2}$ ist $\mathbf{x}_\varphi \times \mathbf{x}_\theta = 0$; längs des Halbkreises $\varphi = 0$ ist die Darstellung nicht eindeutig. Die Ausnahmemengen spielen aber als Nullmengen für die Integration in den folgenden Abschnitten keine Rolle.

BEISPIEL IX.4.6 (TORUS). Ein Torus entsteht durch Drehung um die z -Achse eines Kreises in der (x, z) -Ebene mit Radius r und Mittelpunkt $(R, 0)$ mit $0 < r < R$. Eine Parameterdarstellung ist

$$\mathbf{x}(\psi, \varphi) = \begin{pmatrix} (R + r \sin \psi) \cos \varphi \\ (R + r \sin \psi) \sin \varphi \\ r \cos \psi \end{pmatrix}, \quad 0 \leq \psi, \varphi \leq 2\pi.$$

Es ist

$$\mathbf{x}_\psi = \begin{pmatrix} r \cos \psi \cos \varphi \\ r \cos \psi \sin \varphi \\ -r \sin \psi \end{pmatrix},$$

$$\mathbf{x}_\varphi = \begin{pmatrix} -(R + r \sin \psi) \sin \varphi \\ (R + r \sin \psi) \cos \varphi \\ 0 \end{pmatrix},$$

$$\mathbf{x}_\psi \times \mathbf{x}_\varphi = \begin{pmatrix} (R + r \sin \psi)r \sin \psi \cos \varphi \\ (R + r \sin \psi)r \sin \psi \sin \varphi \\ r(R + r \sin \psi) \cos \psi \end{pmatrix}$$

$$E = r^2,$$

$$G = (R + r \sin \psi)^2,$$

$$F = 0.$$

BEISPIEL IX.4.7 (WENDELFLÄCHE). Eine Parameterdarstellung einer Wendelfläche ist

$$\mathbf{x}(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ a\varphi \end{pmatrix}, \quad r_0 \leq r \leq r_1, \quad \varphi_0 \leq \varphi \leq \varphi_1$$

mit $a > 0$. Es ist

$$\mathbf{x}_r = \begin{pmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{pmatrix}, \quad \mathbf{x}_\varphi = \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \\ a \end{pmatrix}, \quad \mathbf{x}_r \times \mathbf{x}_\varphi = \begin{pmatrix} a \sin \varphi \\ -a \cos \varphi \\ r \end{pmatrix}$$

$$E = 1,$$

$$G = r^2 + a^2,$$

$$F = 0.$$

IX.4.2. Der Flächeninhalt. Wir wollen den Flächeninhalt eines regulären Flächenstückes S mit der Parameterdarstellung $\mathbf{x} : D \rightarrow \mathbb{R}^3$ berechnen. Dazu unterteilen wir die (u, v) -Ebene durch Gitterlinien $u = n2^{-k}$, $v = n2^{-k}$ mit $n \in \mathbb{Z}$, $k \in \mathbb{N}$ in Quadrate der Kantenlänge 2^{-k} (vgl. Abb. IX.4.1 (S. 351)). Betrachte ein Quadrat Q , das ganz in D liegt. Der Mittelpunkt von Q habe die Koordinaten (u_0, v_0) . Das Stück $\mathbf{x}(Q)$ von S , das Q entspricht, wird approximiert durch den Ausschnitt

$$\mathbf{x}(u_0, v_0) + s\mathbf{x}_u(u_0, v_0) + t\mathbf{x}_v(u_0, v_0) \quad , \quad -2^{-k-1} \leq s, t \leq 2^{-k-1},$$

der Tangentialebene. Der Flächeninhalt dieses Ebenenstückes ist gemäß Abschnitt I.4.7 (S. 31) gegeben durch $|\mathbf{x}_u(u_0, v_0) \times \mathbf{x}_v(u_0, v_0)|2^{-2k}$, wobei

2^{-2k} der Flächeninhalt von Q ist. Summieren wir über alle Quadrate, die in D liegen, erhalten wir eine Näherung für den Flächeninhalt von S . Wenn die Gitterweite gegen Null strebt, d.h. $k \rightarrow \infty$, konvergieren diese Näherungswerte gegen den Flächeninhalt von S . Damit erhalten wir:

Der FLÄCHENINHALT eines regulären Flächenstückes $\mathbf{x} : D \rightarrow \mathbb{R}^3$ ist

$$O(S) = \iint_D |\mathbf{x}_u \times \mathbf{x}_v| dudv.$$

Mit den metrischen Fundamentalgrößen ergibt sich

$$O(S) = \iint_D \sqrt{EG - F^2} dudv.$$

BEISPIEL IX.4.8. Für die Kugel aus Beispiel IX.4.5 (S. 354) erhalten wir

$$\begin{aligned} O(S) &= \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{r^2 \cos^2(\theta) \cdot r^2 - 0} d\theta d\varphi \\ &= \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cos \theta d\theta d\varphi \\ &= 2\pi r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta d\theta \\ &= 4\pi r^2. \end{aligned}$$

BEISPIEL IX.4.9. Für den Torus aus Beispiel IX.4.6 (S. 354) ergibt sich

$$\begin{aligned} O(S) &= \int_0^{2\pi} \int_0^{2\pi} \sqrt{r^2(R + r \sin \psi)^2 - 0} d\psi d\varphi \\ &= \int_0^{2\pi} \int_0^{2\pi} r(R + r \sin \psi) d\psi d\varphi \\ &= 2\pi r \int_0^{2\pi} (R + r \sin \psi) d\psi \\ &= 2\pi r \cdot 2\pi R \\ &= 4\pi^2 r R. \end{aligned}$$

IX.4.3. Das Oberflächenintegral einer skalaren Funktion.

Wir betrachten ein reguläres Flächenstück S mit der Parameterdarstellung $\mathbf{x} : D \rightarrow \mathbb{R}^3$ und eine stetige Funktion $f : S \rightarrow \mathbb{R}$ auf S . Wir fassen

diese Funktion als eine Belegung auf S auf und wollen ihr dementsprechend ein Integral zuordnen. Dazu können wir wie im vorigen Abschnitt vorgehen. Wir müssen nur den Flächeninhalt des Flächenstückes über dem dortigen Quadrat Q mit dem Wert $f(\mathbf{x}(u_0, v_0))$ gewichten. Die entsprechende Summation und der Grenzübergang $k \rightarrow \infty$ liefern dann:

Das OBERFLÄCHENINTEGRAL der Funktion f über dem Flächenstück S mit der Parameterdarstellung $\mathbf{x} : D \rightarrow \mathbb{R}^3$ ist gegeben durch

$$\iint_S f dO = \iint_D f(\mathbf{x}(u, v)) |\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)| du dv.$$

BEISPIEL IX.4.10. Wir wollen das axiale Drehmoment bzgl. der z -Achse einer Kugelschale mit Radius r , homogener Dichte und Gesamtmasse m berechnen. Dieses ist gegeben durch

$$I = \frac{1}{O(S)} \iint_S m(x^2 + y^2) dO,$$

wobei $O(S)$ die Kugeloberfläche ist. Mit Beispiel IX.4.5 (S. 354) und IX.4.8 (S. 356) ergibt sich

$$\begin{aligned} I &= \frac{m}{4\pi r^2} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \underbrace{(r^2 \cos^2 \varphi \cos^2 \theta + r^2 \sin^2 \varphi \cos^2 \theta)}_{=x^2+y^2} \cdot \underbrace{r^2 \cos \theta}_{=|\mathbf{x}_\varphi \times \mathbf{x}_\theta|} d\theta d\varphi \\ &= \frac{m}{4\pi r^2} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^4 \cos^3 \theta d\theta d\varphi \\ &= \frac{m}{4\pi r^2} 2\pi r^4 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3 \theta d\theta \\ &= \frac{m}{4\pi r^2} 2\pi r^4 \frac{4}{3} \\ &= \frac{2}{3} m r^2. \end{aligned}$$

BEISPIEL IX.4.11. Wir wollen den Schwerpunkt (x_S, y_S, z_S) des Teils der Fläche

$$z = \frac{2}{3}(x^{3/2} + y^{3/2})$$

berechnen, der von den Ebenen $x = 0$, $y = 0$ und $x + y = 1$ geschnitten wird.

Die Parameterdarstellung wird gegeben durch Beispiel IX.4.3 (S. 353) mit

$$D = \{(u, v) : u \geq 0, v \geq 0, u + v \leq 1\}$$

(Dreieck) und

$$h(u, v) = \frac{2}{3}(u^{3/2} + v^{3/2}).$$

Mit Beispiel IX.4.3 ergibt sich

$$\sqrt{EG - F^2} = \sqrt{(1+u)(1+v) - uv} = \sqrt{1+u+v}.$$

Also ist der Flächeninhalt von S

$$\begin{aligned} O(S) &= \int_0^1 \left\{ \int_0^{1-u} \sqrt{1+u+v} \, dv \right\} du \\ &= \int_0^1 \left\{ \frac{2}{3}(1+u+v)^{\frac{3}{2}} \Big|_{v=0}^{v=1-u} \right\} du \\ &= \int_0^1 \frac{2}{3} \left\{ \sqrt{8} - (1+u)^{\frac{3}{2}} \right\} du \\ &= \frac{2}{3} \left\{ \sqrt{8} - \frac{2}{5}(1+u)^{\frac{5}{2}} \Big|_{u=0}^{u=1} \right\} \\ &= \frac{4}{15}(\sqrt{2} + 1). \end{aligned}$$

Für die x -Koordinate des Schwerpunktes ergibt sich damit

$$\begin{aligned} x_S &= \frac{1}{O(S)} \iint_S x dO \\ &= \frac{15}{4(\sqrt{2} + 1)} \int_0^1 \left\{ \int_0^{1-u} u \sqrt{1+u+v} \, dv \right\} du \\ &= \frac{15}{4(\sqrt{2} + 1)} \int_0^1 \frac{2}{3} u \left\{ \sqrt{8} - (1+u)^{\frac{3}{2}} \right\} du \\ &= \frac{5}{2(\sqrt{2} + 1)} \left\{ \underbrace{\sqrt{8} \int_0^1 u du}_{=\frac{1}{2}} - \underbrace{\int_0^1 u(1+u)^{\frac{3}{2}} du}_{=\frac{2}{5}[u(1+u)^{\frac{5}{2}} - \frac{2}{7}(1+u)^{\frac{7}{2}}] \Big|_{u=0}^{u=1}} \right\} \\ &= \frac{1}{14}(26 - 15\sqrt{2}). \end{aligned}$$

Aus Symmetriegründen ist $y_S = x_S$.

Für die z -Koordinate erhalten wir

$$\begin{aligned} z_S &= \frac{1}{O(S)} \iint_S z dO \\ &= \frac{1}{O(S)} \left\{ \iint_S \frac{2}{3} x^{\frac{3}{2}} dO + \iint_S \frac{2}{3} y^{\frac{3}{2}} dO \right\} \\ &= \frac{1}{O(S)} \left\{ \iint_D \frac{2}{3} u^{\frac{3}{2}} \sqrt{1+u+v} \, dudv \right\} \end{aligned}$$

$$+ \left. \iint_D \frac{2}{3} v^{\frac{3}{2}} \sqrt{1+u+v} \, dudv \right\}.$$

Da D invariant ist unter der Vertauschung von u und v , stimmen die beiden Integrale in der letzten Zeile überein. Damit ergibt sich

$$\begin{aligned} z_S &= \frac{4}{3} \frac{1}{O(S)} \iint_D u^{\frac{3}{2}} \sqrt{1+u+v} \, dudv \\ &= \frac{4}{3} \frac{15}{4(\sqrt{2}+1)} \int_0^1 \left\{ \int_0^{1-u} u^{\frac{3}{2}} \sqrt{1+u+v} \, dv \right\} du \\ &= \frac{5}{\sqrt{2}+1} \int_0^1 \frac{2}{3} u^{\frac{3}{2}} \left[\sqrt{8} - (1+u)^{\frac{3}{2}} \right] du \\ &= \frac{5}{\sqrt{2}+1} \frac{2}{3} \left\{ \underbrace{\int_0^1 \sqrt{8} u^{\frac{3}{2}} du}_{=\frac{2}{5}\sqrt{8}} - \int_0^1 \{u(1+u)\}^{\frac{3}{2}} du \right\}. \end{aligned}$$

Für das zweite Integral erhalten wir

$$\begin{aligned} & \int_0^1 \{u(1+u)\}^{\frac{3}{2}} du \\ &= \int_0^1 \left\{ \frac{1}{4} [(2u+1)^2 - 1] \right\}^{\frac{3}{2}} du \quad |t = 2u+1 \\ &= \int_1^3 \frac{1}{8} (t^2 - 1)^{\frac{3}{2}} \frac{1}{2} dt \\ &= \frac{1}{16} \int_1^3 (t^2 - 1)^{\frac{3}{2}} dt \\ &= \frac{1}{64} \left[t(t^2 - 1)^{\frac{3}{2}} - \frac{3}{2} t(t^2 - 1)^{\frac{1}{2}} + \frac{3}{2} \ln(t + \sqrt{t^2 - 1}) \right] \Bigg|_{t=1}^{t=3} \\ &= \frac{1}{64} \left[3 \cdot 8^{\frac{3}{2}} - \frac{9}{2} \cdot 8^{\frac{1}{2}} + \frac{3}{2} \ln(\underbrace{3 + \sqrt{8}}_{=(1+\sqrt{2})^2}) \right] \\ &= \frac{3}{64} [16\sqrt{2} - \sqrt{2} + \ln(1 + \sqrt{2})]. \end{aligned}$$

Damit ergibt sich insgesamt

$$\begin{aligned} z_S &= \frac{5}{\sqrt{2}+1} \cdot \frac{2}{3} \cdot \left\{ \frac{4}{5} \sqrt{2} - \frac{3}{64} [15\sqrt{2} + \ln(1 + \sqrt{2})] \right\} \\ &= \frac{31\sqrt{2} - 15 \ln(1 + \sqrt{2})}{96(1 + \sqrt{2})}. \end{aligned}$$

IX.4.4. Transformationsformel für Gebietsintegrale. Die Parameterdarstellung eines regulären Flächenstückes in der Ebene $z = 0$ hat die Form

$$\mathbf{x}(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \\ 0 \end{pmatrix}.$$

Die Bedingung $|\mathbf{x}_u \times \mathbf{x}_v| \neq 0$ ist dann äquivalent zu

$$|x_u y_v - x_v y_u| = \left| \det \begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix} \right| \neq 0.$$

Da wir die Ebene $z = 0$ mit dem \mathbb{R}^2 identifizieren können (d.h. wir ignorieren die z -Koordinate), nennen wir eine solche Abbildung eine **KOORDINATENTRANSFORMATION** des ebenen Bereiches $D : (u, v) \mapsto (x(u, v), y(u, v))$. Aus dem vorigen Abschnitt folgt dann:

TRANSFORMATIONSFORMEL: Wird die Menge $D \subset \mathbb{R}^2$ unter der injektiven Abbildung $(u, v) \mapsto (x(u, v), y(u, v))$ in die Menge $S \subset \mathbb{R}^2$ abgebildet und gilt $|x_u y_v - x_v y_u| \neq 0$ für alle $(u, v) \in D$, so ist für jede stetige Funktion f auf S

$$\begin{aligned} & \iint_S f(x, y) dx dy \\ &= \iint_D f(x(u, v), y(u, v)) \\ & \quad |x_u(u, v)y_v(u, v) - x_v(u, v)y_u(u, v)| du dv. \end{aligned}$$

BEISPIEL IX.4.12 (AFFINE KOORDINATEN). Für

$$\begin{aligned} x &= x_0 + au + bv, \\ y &= y_0 + cu + dv \end{aligned}$$

mit $ad - bc \neq 0$ erhalten wir

$$\begin{aligned} & \iint_S f(x, y) dx dy \\ &= \iint_D f(x_0 + au + bv, y_0 + cu + dv) |ad - bc| du dv. \end{aligned}$$

So wird z.B. durch die Transformation $x = au, y = bv$ mit $a > 0, b > 0$ die Ellipse

$$E = \left\{ (x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\}$$

auf den Kreis

$$K = \{ (u, v) : u^2 + v^2 \leq 1 \}$$

abgebildet. Daher gilt für die Fläche der Ellipse

$$\begin{aligned} F(E) &= \iint_E dx dy \\ &= \iint_K ab du dv \\ &= abF(K) \\ &= \pi ab. \end{aligned}$$

BEISPIEL IX.4.13 (POLARKOORDINATEN). Für

$$\begin{aligned} x &= r \cos \varphi, \\ y &= r \sin \varphi \end{aligned}$$

erhalten wir

$$\iint_S f(x, y) dx dy = \iint_D f(r \cos \varphi, r \sin \varphi) r dr d\varphi.$$

Für das Trägheitsmoment bzgl. der y -Achse der Kreisscheibe

$$K = \{(x, y) : x^2 + y^2 \leq 1\}$$

erhalten wir so z.B. bei konstanter Dichte ρ mit

$$D = \{(r, \varphi) : 0 \leq r \leq 1, 0 \leq \varphi \leq 2\pi\}$$

den Wert

$$\begin{aligned} M &= \rho \iint_K x^2 dx dy \\ &= \rho \iint_D r^2 \cos^2 \varphi r dr d\varphi \\ &= \rho \int_0^1 \left\{ \int_0^{2\pi} r^3 \cos^2 \varphi d\varphi \right\} dr \\ &= \pi \rho \int_0^1 r^3 dr \\ &= \frac{1}{4} \pi \rho. \end{aligned}$$

BEISPIEL IX.4.14 (ELLIPTISCHE KOORDINATEN). Sie haben die Form

$$\begin{aligned} x &= a \cosh u \cos v, \\ y &= a \sinh u \sin v \end{aligned}$$

mit $a > 0$. Es ist

$$\begin{aligned} x_u y_v - x_v y_u &= a^2 [\sinh^2 u \cos^2 v + \cosh^2 u \sin^2 v] \\ &= a^2 [\cosh^2 u - \cos^2 v]. \end{aligned}$$

IX.4.5. Der Fluss eines Vektorfeldes. Wir betrachten ein reguläres Flächenstück S mit der Parameterdarstellung $\mathbf{x} : D \rightarrow \mathbb{R}^3$ und ein Vektorfeld $\mathbf{v} : S \rightarrow \mathbb{R}^3$ auf S . Wenn wir \mathbf{v} als Geschwindigkeitsfeld einer stationären Flüssigkeitsströmung interpretieren, stellt das Spatprodukt

$$[\mathbf{v}, \mathbf{x}_u du, \mathbf{x}_v dv] = [\mathbf{v}, \mathbf{x}_u, \mathbf{x}_v] dudv$$

das Volumen der Flüssigkeitsmenge dar, die pro Zeiteinheit durch das Oberflächenelement

$$dO = |\mathbf{x}_u \times \mathbf{x}_v| dudv$$

strömt. Da gemäß Abschnitt 1.4.8 (S. 32)

$$\begin{aligned} [\mathbf{v}, \mathbf{x}_u, \mathbf{x}_v] &= \mathbf{v} \cdot (\mathbf{x}_u \times \mathbf{x}_v) \\ &= \mathbf{v} \cdot \mathbf{n} |\mathbf{x}_u \times \mathbf{x}_v| \end{aligned}$$

mit der Flächennormalen

$$\mathbf{n} = \frac{1}{|\mathbf{x}_u \times \mathbf{x}_v|} \mathbf{x}_u \times \mathbf{x}_v$$

aus Abschnitt IX.4.1 ist, ist dieses Flüssigkeitsvolumen gleich $\mathbf{v} \cdot \mathbf{n} dO$.

Daher beschreibt das Integral $\iint_S \mathbf{v} \cdot \mathbf{n} dO$ das Flüssigkeitsvolumen, das pro Zeiteinheit durch das gesamte Flächenstück strömt.

Diese Beobachtung motiviert folgende Definition:

Der FLUSS des Vektorfeldes \mathbf{v} durch das Flächenstück S mit der Parameterdarstellung $\mathbf{x} : D \rightarrow \mathbb{R}^3$ ist gegeben durch

$$\begin{aligned} \iint_S \mathbf{v} \cdot d\mathbf{O} &= \iint_S \mathbf{v} \cdot \mathbf{n} dO \\ &= \iint_D [\mathbf{v}, \mathbf{x}_u, \mathbf{x}_v] dudv. \end{aligned}$$

BEISPIEL IX.4.15. Wir wollen den Fluss des Vektorfeldes

$$\mathbf{v}(x, y, z) = \begin{pmatrix} 2z \\ x + y \\ 0 \end{pmatrix}$$

durch die Kugel

$$S = \{(x, y, z) : x^2 + y^2 + z^2 = r^2\}$$

(von innen nach außen!) berechnen. Die Parameterdarstellung ist gemäß Beispiel IX.4.5 (S. 354)

$$\mathbf{x}(\varphi, \theta) = \begin{pmatrix} r \cos \varphi \cos \theta \\ r \sin \varphi \cos \theta \\ r \sin \theta \end{pmatrix}.$$

Die Normale ist

$$\mathbf{n} = \begin{pmatrix} \cos \varphi \cos \theta \\ \sin \varphi \cos \theta \\ \sin \theta \end{pmatrix}.$$

Sie zeigt in die richtige Richtung, nämlich nach außen. (Sonst müssten wir das Vorzeichen umkehren!) Gemäß Beispiel IX.4.5 ist

$$\begin{aligned} |\mathbf{x}_u \times \mathbf{x}_v| &= \sqrt{EG - F^2} \\ &= r^2 \cos \theta. \end{aligned}$$

Damit erhalten wir für den gesuchten Fluss

$$\begin{aligned} \iint_S \mathbf{v} \cdot \mathbf{n} dO &= \int_0^{2\pi} \left\{ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \begin{pmatrix} 2r \sin \theta \\ r \cos \theta (\cos \varphi + \sin \varphi) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \cos \varphi \cos \theta \\ \sin \varphi \cos \theta \\ \sin \theta \end{pmatrix} r^2 \cos \theta d\theta \right\} d\varphi \\ &= \int_0^{2\pi} \left\{ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2r^3 \underbrace{\sin \theta \cos^2 \theta}_{\int \dots d\theta = 0} \cos \varphi \right. \\ &\quad \left. + r^3 \underbrace{\cos^3 \theta}_{\int \dots d\theta = \frac{4}{3}} \sin \varphi (\cos \varphi + \sin \varphi) d\theta \right\} d\varphi \\ &= \frac{4}{3} r^3 \int_0^{2\pi} \underbrace{\sin \varphi \cos \varphi}_{\int \dots d\varphi = 0} + \underbrace{\sin^2 \varphi}_{\int \dots d\varphi = \pi} d\varphi \\ &= \frac{4}{3} \pi r^3. \end{aligned}$$

IX.4.6. Der Satz von Stokes. Der Satz von Green besitzt eine Verallgemeinerung für räumliche ZWEISEITIGE FLÄCHEN, bei denen man eindeutig von einer unteren und oberen (bzw. inneren und äußeren) Seite reden kann.

Jedes reguläre Flächenstück S ist zweiseitig. Wählt man das Vorzeichen von

$$\mathbf{n} = \pm \frac{1}{|\mathbf{x}_u \times \mathbf{x}_v|} \mathbf{x}_u \times \mathbf{x}_v$$

einheitlich für alle Punkte von S , so wird die Oberseite von S durch die Vorschrift bestimmt:

„ \mathbf{n} weist im Flächenpunkt aufgetragen aus der Oberseite heraus.“

Eine stückweise reguläre Fläche ist dann ZWEISEITIG oder ORIENTIERBAR, wenn sich die einzelnen Flächenstücke so wählen lassen, dass sich der Umlaufsinn über gemeinsame Kanten stetig fortsetzen lässt.

Die Kugel und der Zylinder sind Beispiele für solche orientierbaren Flächen. Das Möbiusband ist ein Beispiel für eine Fläche, die nicht orientierbar ist.

SATZ VON STOKES: Sei S eine orientierbare, stückweise reguläre Fläche mit überschneidungsfreier Randkurve ∂S , die so durchlaufen wird, dass S links liegt und der Umlaufsinn zusammen mit der Normalen von S eine Rechtsschraubung ergibt. Weiter sei $U \subset \mathbb{R}^3$ eine offene Menge, die S enthält, und $\mathbf{v} : U \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Dann ist

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \iint_S (\operatorname{rot} \mathbf{v}) \cdot \mathbf{n} dO.$$

Aus dem Satz von Stokes folgt, dass für je zwei gleich orientierte Flächen S_1, S_2 mit gleicher Randkurve gilt

$$\iint_{S_1} (\operatorname{rot} \mathbf{v}) \cdot \mathbf{n} dO = \iint_{S_2} (\operatorname{rot} \mathbf{v}) \cdot \mathbf{n} dO.$$

Außerdem ergibt sich folgende Interpretation der Rotation:

$(\operatorname{rot} \mathbf{v}(\mathbf{x})) \cdot \mathbf{n}$ gibt im Punkt \mathbf{x} die Zirkulation des Vektorfeldes \mathbf{v} pro Flächeneinheit an, für jedes zu \mathbf{n} senkrechte reguläre Flächenstück durch \mathbf{x} .

BEISPIEL IX.4.16. S entstehe durch den Schnitt des Zylinders

$$Z = \{(x, y, z) : x^2 + y^2 = 1\}$$

mit der Ebene

$$E = \{(x, y, z) : x + y + z = 1\};$$

die Normale weise in den ersten Oktanten. Wir wollen

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{x}$$

für das Vektorfeld

$$\mathbf{v}(x, y, z) = \begin{pmatrix} -y^3 \\ x^3 \\ -z^3 \end{pmatrix}$$

berechnen.

Für den direkten Weg benötigen wir eine Parametrisierung von ∂S .

Diese lautet

$$\mathbf{x}(t) = \begin{pmatrix} \cos t \\ \sin t \\ 1 - \cos t - \sin t \end{pmatrix}, 0 \leq t \leq 2\pi.$$

Damit ergibt sich

$$\begin{aligned} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} &= \int_0^{2\pi} \begin{pmatrix} -\sin^3 t \\ \cos^3 t \\ -(1 - \cos t - \sin t)^3 \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \\ \sin t - \cos t \end{pmatrix} dt \\ &= \int_0^{2\pi} \sin^4 t + \cos^4 t + (\cos t - \sin t)(1 - \cos t - \sin t)^3 dt. \end{aligned}$$

Die Berechnung dieses Integrals ist nicht schwierig, aber mühselig. Mit dem Satz von Stokes folgt

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} = \iint_S (\operatorname{rot} \mathbf{v}) \cdot \mathbf{n} dO.$$

Für die Rotation erhalten wir

$$\operatorname{rot} \mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ 3x^2 + 3y^2 \end{pmatrix}.$$

Eine Parameterdarstellung von S ist

$$\mathbf{x}(u, v) = \begin{pmatrix} u \\ v \\ 1 - u - v \end{pmatrix}$$

mit $(u, v) \in K$ und der Einheitskreisscheibe K . Es ist

$$\begin{aligned} \mathbf{x}_u &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \\ \mathbf{x}_v &= \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \\ \mathbf{x}_u \times \mathbf{x}_v &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Die Normale weist also in die richtige Richtung. Damit ergibt sich

$$\begin{aligned} \oint_{\partial S} \mathbf{v} \cdot d\mathbf{x} &= \iint_S (\operatorname{rot} \mathbf{v}) \cdot \mathbf{n} dO \\ &= \iint_K \begin{pmatrix} 0 \\ 0 \\ 3u^2 + 3v^2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} dudv \end{aligned}$$

$$\begin{aligned}
&= \iint_K 3(u^2 + v^2) du dv \quad \left| \begin{array}{l} u = r \cos \varphi, \\ v = r \sin \varphi \end{array} \right. \\
&= \int_0^1 \left\{ \int_0^{2\pi} 3r^2 r d\varphi \right\} dr \\
&= 2\pi \frac{3}{4} \\
&= \frac{3}{2} \pi.
\end{aligned}$$

IX.5. Integration über dreidimensionale Bereiche

IX.5.1. Das Volumen. Wir wollen einer beschränkten Menge $M \subset \mathbb{R}^3$ ein Volumen zuordnen. Falls M ein Quader ist, soll dies das Produkt der Kantenlängen sein. Dazu gehen wir analog zu den ebenen Bereichen in Abschnitt IX.3.1 (S. 341) vor und unterteilen den Raum durch Koordinatenebenen parallel zu den Ebenen $x = n2^{-k}$, $y = n2^{-k}$, $z = n2^{-k}$ mit $n \in \mathbb{Z}$, $k \in \mathbb{N}$ in Würfel mit der Kantenlänge 2^{-k} und dem Volumen 2^{-3k} . Dann bezeichne $s_k(M)$ und $S_k(M)$ die Summe der Volumina aller Würfel, die ganz in M enthalten sind, und diejenige aller Würfel, die mindestens einen Punkt mit M gemeinsam haben. Offensichtlich gilt

$$\begin{aligned}
s_k(M) &\leq S_k(M), \\
s_k(M) &\leq s_{k+1}(M), \\
S_{k+1}(M) &\leq S_k(M)
\end{aligned}$$

für alle $k \in \mathbb{N}$. Daher besitzen die beiden Folgen $(s_k(M))_{k \in \mathbb{N}}$ und $(S_k(M))_{k \in \mathbb{N}}$ jeweils einen Grenzwert $V_i(M)$ bzw. $V_a(M)$. Falls beide übereinstimmen, nennen wir M RIEMANN-MESSBAR und den gemeinsamen Grenzwert das VOLUMEN $V(M)$ von M :

$$V(M) = \lim_{k \rightarrow \infty} s_k(M) = \lim_{k \rightarrow \infty} S_k(M).$$

Besteht eine Menge N nur aus endlich vielen Punkten, regulären Kurvenstücken und regulären Flächenstücken, so ist $V(N) = 0$. Man nennt N daher eine NULLMENGE.

Ist M messbar und N eine Nullmenge, so sind $M \cup N$ und $M \setminus N$ auch messbar und es gilt

$$V(M \cup N) = V(M), \quad V(M \setminus N) = V(M).$$

Für unsere Zwecke können wir uns auf sog. REGULÄRE Bereiche M beschränken, die folgende Eigenschaften haben:

- M ist abgeschlossen und beschränkt.
- Der Rand ∂M von M besteht aus endlich vielen regulären Flächenstücken.
- Das Innere $M \setminus \partial M$ von M ist eine nicht leere, offene Menge.

Man kann zeigen:

Jeder im obigen Sinne reguläre Bereich M ist Riemannmessbar mit positivem Volumen $V(M) > 0$.

IX.5.2. Das Dreifachintegral. Sei M ein regulärer Bereich in \mathbb{R}^3 und $f : M \rightarrow \mathbb{R}$ eine stetige Funktion. Wie im vorigen Abschnitt unterteilen wir den Raum in Würfel der Kantenlänge 2^{-k} . Diejenigen Würfel, die ganz in M liegen, nummerieren wir von 1 bis n_k . Für den i -ten Würfel bezeichnet \mathbf{x}_i den Mittelpunkt. Mit diesen Bezeichnungen können wir die Riemann-Summe

$$Z_k = \sum_{i=1}^{n_k} f(\mathbf{x}_i) 2^{-3k}$$

betrachten. Man kann zeigen, dass die Folge $(Z_k)_{k \in \mathbb{N}}$ konvergiert. Den Grenzwert nennen wir das DREIFACH- oder VOLUMENINTEGRAL von f über M und bezeichnen es mit

$$\iiint_M f dV \quad \text{oder} \quad \iiint_M f dx dy dz.$$

Für das Dreifachintegral gelten die gleichen Rechenregeln wie für Doppelintegrale und Integrale in \mathbb{R} .

IX.5.3. Praktische Berechnung des Dreifachintegrals. Für die praktische Rechnung wichtige Mengen $M \subset \mathbb{R}^3$ lassen sich häufig als „Zylinder“ über einem ebenen Grundbereich D in einer der Koordinatenebenen darstellen. Liegt z.B. D in der (x, y) -Ebene, bedeutet dies, dass es zwei stetige Funktionen g und h auf D gibt, sodass M die Form hat

$$M = \{(x, y, z) : (x, y) \in D, g(x, y) \leq z \leq h(x, y)\}.$$

Ist f eine stetige Funktion auf M , gilt dann für das Dreifachintegral von f über M

$$\iiint_M f dV = \iint_D \left\{ \int_{g(x,y)}^{h(x,y)} f(x, y, z) dz \right\} dx dy.$$

Damit ist die Berechnung des Dreifachintegrals auf diejenige eines Integrals in \mathbb{R} und eines Doppelintegrals zurückgeführt. Ähnliche Formeln gelten, wenn D in der (x, z) - oder (y, z) -Ebene liegt.

BEISPIEL IX.5.1. Wir wollen das Volumen der Menge M berechnen, die von dem Ellipsoid

$$E = \left\{ (x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \right\}$$

und dem Kegel

$$K = \left\{ (x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2} \right\}$$

berandet wird und in dem Bereich $z \geq 0$ liegt. Da sich Ellipsoid und Kegel in der Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{1}{2}$$

schneiden, lässt sich M darstellen als

$$M = \left\{ (x, y, z) : (x, y) \in D, \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}} \leq \frac{z}{c} \leq \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \right\}$$

mit

$$D = \left\{ (x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq \frac{1}{2} \right\}.$$

Damit erhalten wir für das Volumen von M

$$\begin{aligned} V(M) &= \iint_D \left\{ \int_{c\sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}}}^{c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}} dz \right\} dx dy \\ &= c \iint_D \left\{ \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} - \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}} \right\} dx dy. \end{aligned}$$

Zur Berechnung des Doppelintegrals führen wir die Koordinatentransformation

$$\begin{aligned} x &= ar \cos \varphi \\ y &= br \sin \varphi \end{aligned}$$

mit $0 \leq r \leq \frac{1}{\sqrt{2}}$ und $0 \leq \varphi \leq 2\pi$ aus. Die Determinante der Jacobi-Matrix ist

$$\begin{aligned} \det \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} &= \det \begin{pmatrix} a \cos \varphi & -ar \sin \varphi \\ b \sin \varphi & br \cos \varphi \end{pmatrix} \\ &= abr. \end{aligned}$$

Damit erhalten wir

$$\begin{aligned} V(M) &= c \iint_D \left\{ \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} - \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}} \right\} dx dy \\ &= c \int_0^{\frac{1}{\sqrt{2}}} \left\{ \int_0^{2\pi} [\sqrt{1 - r^2} - r] abrd\varphi \right\} dr \\ &= 2\pi abc \int_0^{\frac{1}{\sqrt{2}}} r\sqrt{1 - r^2} - r^2 dr \end{aligned}$$

$$\begin{aligned}
&= 2\pi abc \left[-\frac{1}{3}(1-r^2)^{\frac{3}{2}} - \frac{1}{3}r^3 \right] \Big|_{r=0}^{r=\frac{1}{\sqrt{2}}} \\
&= \frac{2}{3}\pi abc \left[1 - \left(\frac{1}{2}\right)^{\frac{3}{2}} - \left(\frac{1}{2}\right)^{\frac{3}{2}} \right] \\
&= \frac{1}{3}\pi abc(2 - \sqrt{2}).
\end{aligned}$$

IX.5.4. Die Transformationsformel für Volumenintegrale.

Wir betrachten zwei reguläre Bereiche U und M in \mathbb{R}^3 . Unter einer KOORDINATENTRANSFORMATION von U auf M verstehen wir eine stetig differenzierbare Abbildung $\mathbf{x} : U \ni (u, v, w)^T \mapsto \mathbf{x}(u, v, w) \in M$ mit den Eigenschaften:

- \mathbf{x} ist injektiv.
- Die Vektoren $\mathbf{x}_u(u, v, w)$, $\mathbf{x}_v(u, v, w)$, $\mathbf{x}_w(u, v, w)$ sind für alle $(u, v, w)^T \in U$ linear unabhängig.

Ein Quader $u_0 \leq u \leq u_0 + \Delta u$, $v_0 \leq v \leq v_0 + \Delta v$, $w_0 \leq w \leq w_0 + \Delta w$ in U wird durch \mathbf{x} in erster Approximation abgebildet auf ein Spat ausgehend vom Punkt $\mathbf{x}(u_0, v_0, w_0)$ aufgespannt von den Vektoren $\mathbf{x}_u(u_0, v_0, w_0)\Delta u$, $\mathbf{x}_v(u_0, v_0, w_0)\Delta v$ und $\mathbf{x}_w(u_0, v_0, w_0)\Delta w$. Das Volumen dieses Spates ist gemäß Abschnitt I.4.8 (S. 32)

$$\begin{aligned}
\Delta V &= |[\mathbf{x}_u, \mathbf{x}_v, \mathbf{x}_w]| \Delta u \Delta v \Delta w \\
&= |\det(\mathbf{x}_u, \mathbf{x}_v, \mathbf{x}_w)| \Delta u \Delta v \Delta w.
\end{aligned}$$

Die Determinante

$$\det(\mathbf{x}_u, \mathbf{x}_v, \mathbf{x}_w) = \det \begin{pmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{pmatrix}$$

heißt JACOBI-DETERMINANTE oder FUNKTIONALDETERMINANTE der Transformation. Setzt man obige Formel für ΔV in die entsprechenden Riemannschen Summen ein, erhält man:

TRANSFORMATIONSFORMEL FÜR VOLUMENINTEGRALE:
Entsteht der reguläre Bereich $M \subset \mathbb{R}^3$ durch die Koordinatentransformation $\mathbf{x} : (u, v, w)^T \mapsto \mathbf{x}(u, v, w)$ aus dem regulären Bereich U , so gilt für jede stetige Funktion $f : M \rightarrow \mathbb{R}$ auf M die Transformationsformel

$$\begin{aligned}
&\iiint_M f(x, y, z) dx dy dz \\
&= \iiint_U f(\mathbf{x}(u, v, w)) |\det(\mathbf{x}_u, \mathbf{x}_v, \mathbf{x}_w)| du dv dw.
\end{aligned}$$

BEISPIEL IX.5.2 (AFFINE KOORDINATEN). Für die Transformation $\mathbf{x} = A\mathbf{u} + \mathbf{x}_0$ mit $A \in \mathbb{R}^{3 \times 3}$ und $\det A \neq 0$ erhalten wir

$$\iiint_M f(\mathbf{x}) dx dy dz = \iiint_U f(A\mathbf{u} + \mathbf{x}_0) |\det A| du dv dw.$$

Das Ellipsoid

$$M = \left\{ (x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq r^2 \right\}$$

mit $a > 0$, $b > 0$, $c > 0$ ist das Bild der Kugel

$$K = \{(u, v, w) : u^2 + v^2 + w^2 \leq r^2\}$$

unter der Transformation

$$\mathbf{x} : (u, v, w)^T \mapsto (au, bv, cw)^T.$$

Dies entspricht

$$\begin{aligned} \mathbf{x}_0 &= 0, \\ A &= \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}. \end{aligned}$$

Wegen

$$\det A = abc$$

ergibt sich

$$\begin{aligned} V(M) &= \iiint_M dx dy dz \\ &= \iiint_K abc du dv dw \\ &= \frac{4}{3} \pi r^3 abc. \end{aligned}$$

BEISPIEL IX.5.3 (ZYLINDERKOORDINATEN). Dies ist die Transformation

$$\begin{aligned} x &= r \cos \varphi, \\ y &= r \sin \varphi, \\ z &= z. \end{aligned}$$

Die Funktionaldeterminante ist

$$\det \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = r.$$

Also ist

$$\iiint_M f dx dy dz = \iiint_U f(r \cos \varphi, r \sin \varphi, z) r dr d\varphi dz.$$

Der Zylinder

$$M = \{(x, y, z) : x^2 + y^2 \leq 1, 0 \leq z \leq 2\}$$

habe z.B. die Massenverteilung

$$f(x, y, z) = z^2(x^2 + y^2).$$

Dann ist die Gesamtmasse

$$\begin{aligned} M &= \iiint_M f dx dy dz \\ &= \int_0^1 \left\{ \int_0^{2\pi} \left\{ \int_0^2 z^2 r^2 r dz \right\} d\varphi \right\} dr \\ &= \frac{8}{3} 2\pi \int_0^1 r^3 dr \\ &= \frac{4}{3} \pi. \end{aligned}$$

BEISPIEL IX.5.4 (KUGELKOORDINATEN). Dies ist die Transformation

$$\begin{aligned} x &= r \cos \varphi \cos \theta, \\ y &= r \sin \varphi \cos \theta, \\ z &= r \sin \theta. \end{aligned}$$

Die Funktionaldeterminante ist

$$\det \begin{pmatrix} \cos \varphi \cos \theta & -r \sin \varphi \cos \theta & -r \cos \varphi \sin \theta \\ \sin \varphi \cos \theta & r \cos \varphi \cos \theta & -r \sin \varphi \sin \theta \\ \sin \theta & 0 & r \cos \theta \end{pmatrix} = r^2 \cos \theta.$$

Also ist

$$\begin{aligned} &\iiint_M f dx dy dz \\ &= \iiint_U f(r \cos \varphi \cos \theta, r \sin \varphi \cos \theta, r \sin \theta) r^2 \cos \theta dr d\varphi d\theta. \end{aligned}$$

Für die Massenverteilung

$$f(x, y, z) = z^2(x^2 + y^2)$$

in der Einheitskugel ergibt sich so z.B. die Gesamtmasse

$$\begin{aligned} M &= \iiint_M f dx dy dz \\ &= \int_0^1 \left\{ \int_0^{2\pi} \left\{ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \underbrace{r^2 \sin^2 \theta}_{=z^2} \underbrace{r^2 \cos^2 \theta}_{=x^2+y^2} r^2 \cos \theta d\theta \right\} d\varphi \right\} dr \end{aligned}$$

$$\begin{aligned}
&= 2\pi \underbrace{\int_0^1 r^6 dr}_{=\frac{1}{7}} \underbrace{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2 \theta \cos^3 \theta d\theta}_{=\frac{4}{15}} \\
&= \frac{8}{105} \pi.
\end{aligned}$$

IX.5.5. Der Satz von Gauß. Wir betrachten einen regulären Bereich $M \subset \mathbb{R}^3$, dessen Rand ∂M aus endlich vielen regulären Flächenstücken besteht, eine offene Menge O , die M enthält, und ein auf O stetig differenzierbares Vektorfeld $\mathbf{v} : O \rightarrow \mathbb{R}^3$. In jedem Punkt von ∂M bezeichne \mathbf{n} den nach außen weisenden Einheitsnormalenvektor. Dann gilt:

GAUSSSCHER INTEGRALSATZ:

$$\iiint_M \operatorname{div} \mathbf{v} dV = \iint_{\partial M} \mathbf{v} \cdot \mathbf{n} dO.$$

BEISPIEL IX.5.5. Der Fluss des Vektorfeldes

$$\mathbf{v}(\mathbf{x}) = \begin{pmatrix} 2z \\ x + y \\ 0 \end{pmatrix}$$

durch die Oberfläche der Kugel

$$K = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$$

beträgt

$$\begin{aligned}
\iint_{\partial K} \mathbf{v} \cdot \mathbf{n} dO &= \iiint_K \operatorname{div} \mathbf{v} dV \\
&= \iiint_K 1 dV \\
&= \frac{4}{3} \pi.
\end{aligned}$$

BEISPIEL IX.5.6. Der Fluss des Vektorfeldes

$$\mathbf{v}(\mathbf{x}) = \begin{pmatrix} xy^2 \\ x^2y \\ y \end{pmatrix}$$

durch die Oberfläche des Zylinders

$$M = \{(x, y, z) : x^2 + y^2 \leq 1, |z| \leq 1\}$$

beträgt

$$\iint_{\partial M} \mathbf{v} \cdot \mathbf{n} dO = \iiint_M \operatorname{div} \mathbf{v} dV$$

$$\begin{aligned}
&= \iiint_M x^2 + y^2 dx dy dz \quad (\text{Zylinderkoordinaten}) \\
&= \int_0^1 \left\{ \int_0^{2\pi} \left\{ \int_{-1}^1 r^2 r dz \right\} d\varphi \right\} dr \\
&= 4\pi \int_0^1 r^3 dr \\
&= \pi.
\end{aligned}$$

BEISPIEL IX.5.7 (KONTINUITÄTSGLEICHUNG DER STRÖMUNGSMCHANIK). Sei \mathbf{v} das Geschwindigkeitsfeld einer strömenden Flüssigkeit mit der räumlich und zeitlich variablen Dichte ρ . Wir betrachten einen beliebigen räumlichen Bereich M im Strömungsgebiet. Durch die Oberfläche von M tritt pro Zeiteinheit die Masse

$$\iint_{\partial M} \rho \mathbf{v} \cdot \mathbf{n} dO$$

aus. Wegen der Massenerhaltung muss dies gleich der zeitlichen Änderung der Masse sein. Dieses ist

$$-\frac{d}{dt} \iiint_M \rho dV.$$

(Minuszeichen, da der Fluss aus M heraus betrachtet wird!) Damit ergibt sich

$$\begin{aligned}
0 &= \underbrace{\iint_{\partial M} \rho \mathbf{v} \cdot \mathbf{n} dO}_{=\iint_M \operatorname{div}(\rho \mathbf{v}) dV} + \underbrace{\frac{d}{dt} \iiint_M \rho dV}_{=\iint_M \frac{\partial \rho}{\partial t} dV} \\
&= \iiint_M \left\{ \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) \right\} dV.
\end{aligned}$$

Da diese Identität für jeden noch so kleinen Bereich M gilt, kann man aus ihr punktweise die Kontinuitätsgleichung der Strömungsmechanik

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0$$

folgern.

Setzen wir im Gaußschen Integralsatz für \mathbf{v} den Gradienten einer zweimal stetig differenzierbaren Funktion f ein, erhalten wir wegen

$$\operatorname{div}(\operatorname{grad} f) = \Delta f = f_{xx} + f_{yy} + f_{zz}$$

die Identität

$$\iiint_M \Delta f dV = \iint_{\partial M} \operatorname{grad} f \cdot \mathbf{n} dO = \iint_{\partial M} \frac{\partial f}{\partial \mathbf{n}} dO.$$

Ist f eine skalare Funktion und \mathbf{v} ein Vektorfeld, ergibt sich aus der Produktregel

$$\operatorname{div}(f\mathbf{v}) = f \operatorname{div} \mathbf{v} + (\operatorname{grad} f) \cdot \mathbf{v}.$$

Setzen wir $f\mathbf{v}$ an Stelle von \mathbf{v} im Gaußschen Integralsatz ein, erhalten wir hieraus ein mehrdimensionales Analogon zur partiellen Integration:

$$\iiint_M f \operatorname{div} \mathbf{v} dV = \iint_{\partial M} f \mathbf{v} \cdot \mathbf{n} dO - \iiint_M \mathbf{v} \cdot \operatorname{grad} f dV.$$

KAPITEL X

Gewöhnliche Differentialgleichungen II Systeme

X.1. Existenz- und Eindeutigkeitsätze

X.1.1. Grundbegriffe. Wir erweitern die Begriffe aus Abschnitt VI.1.1 (S. 217) auf Systeme von gewöhnlichen Differentialgleichungen, d.h. gDglen bei denen die Lösung eine vektorwertige Funktion $\mathbf{y} : \mathbb{R} \rightarrow \mathbb{R}^k$ ist. Dazu seien $I \subset \mathbb{R}$ ein Intervall, k und n positive natürliche Zahlen und $\mathbf{f} : I \times \mathbb{R}^{nk} \rightarrow \mathbb{R}^k$ eine gegebene Funktion. Ein SYSTEM GEWOEHNLICHER DIFFERENTIALGLEICHUNGEN n -TER ORDNUNG kurz GDGL n -TER ORDNUNG ist ein Ausdruck der Form

$$\mathbf{y}^{(n)} = \mathbf{f}(x, \mathbf{y}, \dots, \mathbf{y}^{(n-1)}),$$

bei dem die gesuchte Lösung \mathbf{y} ihre Werte im \mathbb{R}^k hat. Offensichtlich entsprechen die gDglen aus Kapitel VI dem Spezialfall $k = 1$.

Eine Funktion $\mathbf{y} : J \rightarrow \mathbb{R}^k$ heißt eine LÖSUNG der gDgl, wenn das Intervall J im Intervall I enthalten ist und wenn für alle $x \in J$ gilt

$$\mathbf{y}^{(n)}(x) = \mathbf{f}(x, \mathbf{y}(x), \dots, \mathbf{y}^{(n-1)}(x)).$$

Unter einem ANFANGSWERTPROBLEM kurz AWP verstehen wir eine gDgl n -ter Ordnung wie oben angegeben zusammen mit n Bedingungen der Form

$$\begin{array}{rcl} \mathbf{y}(x_0) & = & \mathbf{y}_0 \\ \vdots & & \vdots \\ \mathbf{y}^{(n-1)}(x_0) & = & \mathbf{y}_{n-1} \end{array}$$

Dabei sind $x_0 \in I$ und $\mathbf{y}_0, \dots, \mathbf{y}_{n-1} \in \mathbb{R}^k$ ein gegebener Anfangspunkt und gegebene Anfangswerte.

Eine Funktion \mathbf{y} heißt eine LÖSUNG des AWP, wenn es positive Zahlen ε_1 und ε_2 gibt, so dass \mathbf{y} auf $(x_0 - \varepsilon_1, x_0 + \varepsilon_2)$ n -mal differenzierbar ist, dort die gDgl löst und die Anfangsbedingungen erfüllt.

BEISPIEL X.1.1. Wir betrachten eine Population bestehend aus einer Beute und einem Räuber. Die Größe der Beute- bzw. Räuberpopulation zur Zeit t sei $x(t)$ bzw. $y(t)$. Die Beute habe eine konstante Vermehrungsrate und eine Sterberate, die proportional zur Größe der Räuberpopulation ist. Die Räuber haben eine konstante Sterberate und eine Wachstumsrate, die proportional zur Größe der Beutepopulation

ist. Dann genügen x, y der gDgl

$$\begin{aligned}\dot{x} &= \alpha x - \beta xy \\ \dot{y} &= -\gamma y + \delta xy\end{aligned}$$

mit positiven Konstanten $\alpha, \beta, \gamma, \delta$. Zur Zeit $t = 0$ erfüllen sie die Anfangsbedingung

$$\begin{aligned}x(0) &= x_0, \\ y(0) &= y_0.\end{aligned}$$

X.1.2. Reduktion der Ordnung. Ein System gewöhnlicher Differentialgleichungen der Ordnung n im \mathbb{R}^k kann stets auf ein System erster Ordnung im \mathbb{R}^{nk} reduziert werden. Dazu definiert man für

$$\mathbf{y}^{(n)} = \mathbf{f}(x, \mathbf{y}, \dots, \mathbf{y}^{(n-1)})$$

die vektorwertige Funktion $\mathbf{z} : I \rightarrow \mathbb{R}^{nk}$ durch

$$\mathbf{z} = \begin{pmatrix} \mathbf{y} \\ \mathbf{y}' \\ \vdots \\ \mathbf{y}^{(n-1)} \end{pmatrix}$$

und die Funktion $\mathbf{F} : I \times \mathbb{R}^{nk} \rightarrow \mathbb{R}^{nk}$ durch

$$\mathbf{F}(x, \mathbf{z}) = \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_{n-1} \\ \mathbf{f}(x, \mathbf{z}) \end{pmatrix}.$$

Dabei bezeichnet \mathbf{z}_j die Komponenten $jk + 1, \dots, (j + 1)k$ von \mathbf{z} . Dann ist $\mathbf{y} : J \rightarrow \mathbb{R}^k$ genau dann eine Lösung der gDgl

$$\mathbf{y}^{(n)} = \mathbf{f}(x, \mathbf{y}, \dots, \mathbf{y}^{(n-1)}),$$

wenn $\mathbf{z} : J \rightarrow \mathbb{R}^{nk}$ eine Lösung der gDgl

$$\mathbf{z}' = \mathbf{F}(x, \mathbf{z})$$

ist.

Wegen dieser Reduktion der Ordnung können wir uns bei den theoretischen Ergebnissen der folgenden Abschnitte auf gDglen erster Ordnung beschränken. Für die praktische Rechnung ist es aber nicht immer empfehlenswert eine gDgl n -ter Ordnung auf ein entsprechendes System 1-ter Ordnung zu transformieren.

BEISPIEL X.1.2. Die Schwingungsgleichung

$$\ddot{x} + r\dot{x} + \omega^2 x = f(t)$$

geht mit

$$\mathbf{y} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$$

in das System erster Ordnung

$$\begin{aligned}\dot{\mathbf{y}} &= \begin{pmatrix} y_2 \\ f(t) - ry_2 - \omega^2 y_1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -\omega^2 & -r \end{pmatrix} \mathbf{y} + \begin{pmatrix} 0 \\ f(t) \end{pmatrix}\end{aligned}$$

über.

X.1.3. Lipschitz-stetige Funktionen. Für den Existenz- und Eindeutigkeitssatz des nächsten Abschnittes benötigen wir den Begriff der Lipschitz-Stetigkeit. Dazu bezeichnen $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Menge und $\mathbf{f} : I \times U \rightarrow \mathbb{R}^n$ eine Funktion.

DEFINITION X.1.3. (1) Die Funktion $\mathbf{f} : I \times U \rightarrow \mathbb{R}^n$ heißt GLEICHMÄSSIG LIPSCHITZ-STETIG (bzgl. U) auf $I \times U$, wenn es eine Zahl $L \geq 0$ gibt mit

$$\|\mathbf{f}(x, \mathbf{y}) - \mathbf{f}(x, \mathbf{z})\| \leq L \|\mathbf{y} - \mathbf{z}\| \quad \text{für alle } x \in I, \mathbf{y}, \mathbf{z} \in U.$$

Dabei bezeichnet $\|\cdot\|$ irgendeine Norm auf \mathbb{R}^n . Die kleinst mögliche Zahl $L \geq 0$, für die die obige Abschätzung gilt, heißt die LIPSCHITZ-KONSTANTE von \mathbf{f} .

(2) Die Funktion $\mathbf{f} : I \times U \rightarrow \mathbb{R}^n$ heißt LIPSCHITZ-STETIG (bzgl. U) auf $I \times U$, wenn es zu jedem $\mathbf{y}_0 \in U$ eine offene Menge $V \subset U$ gibt mit $\mathbf{y}_0 \in V$, so dass \mathbf{f} auf $I \times V$ gleichmäßig Lipschitz-stetig ist.

Eine Lipschitz-stetige Funktion ist immer stetig. Die Umkehrung gilt nicht, wie das folgende Beispiel zeigt.

BEISPIEL X.1.4. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(y) = \sqrt{|y|}$$

ist stetig, aber nicht Lipschitz-stetig. Denn andernfalls, müsste es ein $\varepsilon > 0$ und ein $L \geq 0$ geben mit

$$|f(y) - f(0)| \leq L|y - 0| \quad \text{für alle } y \in \mathbb{R} \text{ mit } |y| \leq \varepsilon.$$

Dies ist aber äquivalent dazu, dass der Differenzenquotient

$$\frac{|f(y) - f(0)|}{|y - 0|}$$

in einer Umgebung der Null beschränkt ist. Aber die Folge

$$\frac{|f(\frac{1}{n}) - f(0)|}{|\frac{1}{n} - 0|} = \frac{\sqrt{\frac{1}{n}}}{\frac{1}{n}} = \sqrt{n}$$

ist unbeschränkt.

Die Beobachtung aus obigem Beispiel gilt allgemein: Eine Funktion $f : I \rightarrow \mathbb{R}$ ist Lipschitz-stetig, wenn der Differenzenquotient

$$\frac{|f(x) - f(y)|}{|x - y|}$$

für alle $x, y \in I$, $x \neq y$ beschränkt ist. Dies ist sicher dann der Fall, wenn f auf I differenzierbar ist. Allgemein gilt:

Ist $\mathbf{f} : I \times U \rightarrow \mathbb{R}^n$ bzgl. der zweiten Variablen differenzierbar, so ist \mathbf{f} Lipschitz-stetig.

X.1.4. Der Satz von Picard-Lindelöf. Wie Beispiel VI.1.6 (S. 220) zeigt, gibt es AWP mit mehreren Lösungen. Dies liegt an der fehlenden Lipschitz-Stetigkeit der dortigen Funktion f (vgl. Beispiel X.1.4).

SATZ VON PICARD-LINDELÖF: Seien $I \subset \mathbb{R}$ ein offenes Intervall, $U \subset \mathbb{R}^n$ eine offene Menge und $\mathbf{f} : I \times U \rightarrow \mathbb{R}^n$ Lipschitz-stetig. Dann gibt es zu jedem $x_0 \in I$ und jedem $\mathbf{y}_0 \in U$ ein $\varepsilon > 0$, so dass das AWP

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}(x)) \quad \text{für alle } x \in (x_0 - \varepsilon, x_0 + \varepsilon)$$

$$\mathbf{y}(x_0) = \mathbf{y}_0$$

eine eindeutige Lösung besitzt.

BEWEISIDEE. Offensichtlich ist $\mathbf{y} : I \rightarrow \mathbb{R}^n$ genau dann ein Lösung des AWP, wenn für alle $x \in I$ gilt

$$\begin{aligned} \mathbf{y}(x) - \mathbf{y}_0 &= \mathbf{y}(x) - \mathbf{y}(x_0) \\ &= \int_{x_0}^x \mathbf{y}'(t) dt \\ &= \int_{x_0}^x \mathbf{f}(t, \mathbf{y}(t)) dt. \end{aligned}$$

Durch geeignete Wahl von $\varepsilon > 0$ und einer offenen Menge $V \subset U$ mit $\mathbf{y}_0 \in V$ kann man erreichen, dass \mathbf{f} auf $[x_0 - \varepsilon, x_0 + \varepsilon] \times V$ gleichmäßig Lipschitz-stetig ist mit Lipschitz-Konstante L und dass $L\varepsilon \leq \frac{1}{2}$ gilt. Bezeichne mit X die Menge aller stetigen Funktionen auf $[x_0 - \varepsilon, x_0 + \varepsilon]$ mit Werten in V . X ist ein normierter Vektorraum mit der Norm

$$\|\mathbf{y}\|_X = \max_{|x-x_0| \leq \varepsilon} |\mathbf{y}(x)|,$$

wobei $|\cdot|$ die Euklidische Norm auf \mathbb{R}^n ist. Auf X definiert man eine Abbildung Φ durch

$$\Phi(\mathbf{y})(x) = \mathbf{y}_0 + \int_{x_0}^x \mathbf{f}(t, \mathbf{y}(t)) dt.$$

Dann ist \mathbf{y} genau dann eine Lösung des AWP, wenn \mathbf{y} ein Fixpunkt von Φ ist, d.h.

$$\Phi(\mathbf{y}) = \mathbf{y}.$$

Wegen der Lipschitz-Stetigkeit von \mathbf{f} erhält man für je zwei Funktionen $\mathbf{y}, \mathbf{z} \in X$ die Abschätzung

$$\begin{aligned} \|\Phi(\mathbf{y}) - \Phi(\mathbf{z})\|_X &= \max_{|x-x_0| \leq \varepsilon} \left| \int_{x_0}^x \mathbf{f}(t, \mathbf{y}(t)) - \mathbf{f}(t, \mathbf{z}(t)) dt \right| \\ &\leq \max_{|x-x_0| \leq \varepsilon} \left| \int_{x_0}^x \underbrace{|\mathbf{f}(t, \mathbf{y}(t)) - \mathbf{f}(t, \mathbf{z}(t))|}_{\leq L|\mathbf{y}(t) - \mathbf{z}(t)|} dt \right| \\ &\leq \max_{|x-x_0| \leq \varepsilon} \left| \int_{x_0}^x \underbrace{L|\mathbf{y}(t) - \mathbf{z}(t)|}_{\leq \|\mathbf{y} - \mathbf{z}\|_X} dt \right| \\ &\leq L|x - x_0| \|\mathbf{y} - \mathbf{z}\|_X \\ &\leq L\varepsilon \|\mathbf{y} - \mathbf{z}\|_X \\ &\leq \frac{1}{2} \|\mathbf{y} - \mathbf{z}\|_X. \end{aligned}$$

Dies besagt, dass Φ eine KONTRAKTION ist. Wir bezeichnen nun mit \mathbf{z}_0 die Funktion, die konstant gleich \mathbf{y}_0 ist, und definieren die Folge $(\mathbf{z}_n)_{n \in \mathbb{N}}$ durch

$$\mathbf{z}_{i+1} = \Phi(\mathbf{z}_i) \quad i = 0, 1, \dots$$

Da Φ eine Kontraktion ist, kann man zeigen, dass die Folge $(\mathbf{z}_n)_{n \in \mathbb{N}}$ gegen ein $\mathbf{y} \in X$ konvergiert und dass diese Funktion der eindeutige Fixpunkt von Φ und damit die eindeutige Lösung des AWP ist. \square

Aus der Lipschitz-Stetigkeit folgt auch die stetige Abhängigkeit der Lösung einer AWP von den Anfangswerten:

STETIGE ABHÄNGIGKEIT VON DEN ANFANGSWERTEN:
Die Funktion $\mathbf{f} : I \times U \rightarrow \mathbb{R}^n$ sei gleichmäßig Lipschitz-stetig mit Lipschitz-Konstante L . Dann gilt für jedes $x_0 \in I$ und je zwei Lösungen $\mathbf{y}_1, \mathbf{y}_2$ der gDgl

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$$

die Abschätzung

$$|\mathbf{y}_1(x) - \mathbf{y}_2(x)| \leq |\mathbf{y}_1(x_0) - \mathbf{y}_2(x_0)| e^{L|x-x_0|}$$

für alle $x \in I$.

X.2. Differentialgleichungen 1. Ordnung

X.2.1. Exakte Differentialgleichungen. Wir betrachten Differentialgleichungen der Form

$$a(x, y) + b(x, y)y' = 0.$$

Dabei sind $a, b : G \rightarrow \mathbb{R}$ zwei stetige Funktionen auf einer offenen, sternförmigen Menge $G \subset \mathbb{R}^2$. Diese gDgl heißt EXAKT, wenn das Vektorfeld

$$\mathbf{v} : G \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (a(x, y), b(x, y))^T$$

ein Gradientenfeld ist, d.h. wenn es eine Stammfunktion $U : G \rightarrow \mathbb{R}$ gibt mit

$$a = U_x, \quad b = U_y.$$

Da G sternförmig ist, ist gemäß Abschnitt IX.2.4 (S. 335) die gDgl exakt, wenn gilt

$$a_y = b_x.$$

Falls die gDgl exakt und U eine Stammfunktion ist, ist die gDgl wegen der Kettenregel äquivalent zu

$$0 = \frac{d}{dx}U(x, y(x)) = U_x + U_y y'.$$

Also ist jede Lösung der gDgl eine Niveaulinie

$$\{(x, y) : U(x, y) = c\}$$

der Stammfunktion. Für eine gegebene Anfangsbedingung

$$y(x_0) = y_0$$

ist die Konstante c durch

$$U(x_0, y_0) = c$$

bestimmt. Falls $U_y(x_0, y_0) \neq 0$ ist, kann wegen des Satzes über implizite Funktionen (vgl. Abschnitt VIII.3.5 (S. 301)) die Gleichung

$$U(x, y) = U(x_0, y_0)$$

in einer Umgebung von x_0 nach y aufgelöst werden.

Insgesamt erhalten wir somit:

Lösung der gDgl

$$a(x, y) + b(x, y)y' = 0.$$

(1) Teste auf Exaktheit, d.h. gilt

$$a_y = b_x?$$

(2) Falls die gDgl exakt ist, bestimme eine Stammfunktion U , d.h. eine Funktion U mit

$$a = U_x, \quad b = U_y.$$

(3) Für $c \in \mathbb{R}$ löse die Gleichung

$$U(x, y) = c$$

nach y auf.

(4) Bei vorgegebenen Anfangsbedingungen $y(x_0) = y_0$ ist

$$c = U(x_0, y_0).$$

BEISPIEL X.2.1. Betrachte das AWP

$$\begin{aligned} 2xy + (2y + x^2)y' &= 0 \\ y(0) &= 1. \end{aligned}$$

Es ist

$$\begin{aligned} a(x, y) &= 2xy, \\ b(x, y) &= 2y + x^2 \end{aligned}$$

und

$$a_y = 2x = b_x.$$

Also ist die gDgl exakt. Zur Bestimmung einer Stammfunktion U gehen wir wie in Abschnitt IX.2.5 (S. 338) vor. Unbestimmte Integration von a liefert

$$U(x, y) = x^2y + c(y).$$

Einsetzen in die Bedingung $U_y = b$ ergibt

$$\begin{aligned} x^2 + c'(y) &= 2y + x^2 \\ \implies c'(y) &= 2y \\ \implies c(y) &= y^2 + \gamma. \end{aligned}$$

Also ist

$$U(x, y) = x^2y + y^2$$

eine Stammfunktion. Wegen $U(0, 1) = 1$ ist die Lösung des AWP gegeben durch die Gleichung

$$x^2y(x) + y(x)^2 = 1.$$

Die quadratische Gleichung (für y !)

$$y^2 + x^2y = 1$$

hat die Lösungen

$$y_{1,2} = -\frac{1}{2}x^2 \pm \sqrt{\frac{1}{4}x^4 + 1}.$$

Also sind

$$y_{1,2}(x) = \frac{1}{2}(-x^2 \pm \sqrt{x^4 + 4})$$

Kandidaten für die Lösung des AWP. Wegen der Anfangsbedingung $y(0) = 1$ kommt nur die Lösung

$$y(x) = \frac{1}{2}(\sqrt{x^4 + 4} - x^2)$$

in Frage.

X.2.2. Der integrierende Faktor. Wir betrachten wieder eine gDgl der Form

$$a(x, y) + b(x, y)y' = 0,$$

interessieren uns aber für den Fall, dass die gDgl nicht exakt ist, d.h. $a_y \neq b_x$. Manchmal gibt es eine Funktion $M : G \rightarrow \mathbb{R}$, so dass die gDgl

$$M(x, y)a(x, y) + M(x, y)b(x, y)y' = 0$$

exakt ist. Falls $M(x, y) \neq 0$ ist für alle $(x, y) \in G$, hat diese exakte gDgl die gleiche Lösungsmenge wie die ursprüngliche gDgl, und wir können die Methoden des vorigen Abschnittes anwenden. Eine Funktion M der beschriebenen Art heißt **INTEGRIERENDER FAKTOR**.

Aus der Bedingung, dass die neue gDgl exakt sein soll, ergibt sich eine partielle Differentialgleichung für M :

$$\begin{aligned} (aM)_y &= (bM)_x \\ \iff a_y M + aM_y &= b_x M + bM_x \\ \iff bM_x - aM_y &= (a_y - b_x)M. \end{aligned}$$

Diese partielle Differentialgleichung für M ist im allgemeinen schwerer zu lösen als die ursprüngliche gDgl. Manchmal hat man aber Glück und kann eine Lösung M „raten“. „Bewährte“ Kandidaten hierfür sind von der Form

$$\begin{array}{ll} m(x), & m(y), \\ m(x+y), & m(x-y), \\ m(x^2+y^2), & m(x^2-y^2), \\ m(x \cdot y) & \end{array}$$

mit einer Funktion $m : \mathbb{R} \rightarrow \mathbb{R}$.

BEISPIEL X.2.2. Betrachte die gDgl

$$x^2y^3 + y + (x^3y^2 - x)y' = 0.$$

Es ist

$$\begin{aligned} a(x, y) &= x^2y^3 + y, \\ b(x, y) &= x^3y^2 - x. \end{aligned}$$

Wegen

$$\begin{aligned} a_y &= 3x^2y^2 + 1, \\ b_x &= 3x^2y^2 - 1 \end{aligned}$$

ist die gDgl nicht exakt. Auf der Suche nach einem integrierenden Faktor machen wir den Ansatz $M(x, y) = m(xy)$. Die partielle Differentialgleichung für M hat dann die Form

$$\begin{aligned} 2m &= (a_y - b_x)M \\ &= bM_x - aM_y \\ &= (x^3y^2 - x)ym' - (x^2y^3 + y)xm' \\ &= -2xym'. \end{aligned}$$

Mit $u = xy$ geht dies über in die gDgl

$$-\frac{1}{u}m = m'$$

mit der Lösung (Trennung der Variablen!)

$$m(u) = \frac{1}{u}.$$

Also ist

$$M(x, y) = \frac{1}{xy}$$

ein integrierender Faktor. Die neue gDgl lautet

$$xy^2 + \frac{1}{x} + (x^2y - \frac{1}{y})y' = 0.$$

Für die Stammfunktion U erhalten wir

$$\begin{aligned} U(x, y) &= \int (xy^2 + \frac{1}{x})dx + c(y) \\ &= \frac{1}{2}x^2y^2 + \ln(x) + c(y). \end{aligned}$$

Einsetzen in die Gleichung

$$U_y = x^2y - \frac{1}{y}$$

ergibt

$$\begin{aligned} x^2y + c'(y) &= U_y \\ &= x^2y - \frac{1}{y} \\ \implies c'(y) &= -\frac{1}{y} \\ \implies c(y) &= -\ln(y) + c. \end{aligned}$$

Also ist

$$U(x, y) = \frac{1}{2}x^2y^2 + \ln\left(\frac{x}{y}\right) + c,$$

und die allgemeine Lösung y unserer gDgl ist implizit gegeben durch die Gleichung

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

BEISPIEL X.2.3. Wir betrachten das Räuber-Beute Modell aus Beispiel X.1.1 (S. 375)

$$\begin{aligned} \dot{x} &= \alpha x - \beta xy && \text{Beute} \\ \dot{y} &= -\gamma y + \delta xy && \text{Räuber.} \end{aligned}$$

Wir multiplizieren die erste Gleichung mit $y(-\gamma + \delta x)$ und die zweite Gleichung mit $x(\alpha - \beta y)$ und subtrahieren die Ergebnisse. Das liefert

$$(-\gamma + \delta x)y\dot{x} - (\alpha - \beta y)x\dot{y} = 0.$$

Das Vektorfeld

$$\mathbf{v}(x, y) = \begin{pmatrix} y(-\gamma + \delta x) \\ -x(\alpha - \beta y) \end{pmatrix}$$

ist wegen

$$\mathbf{v}_{1,y} = -\gamma + \delta x \neq -\alpha + \beta y = \mathbf{v}_{2,x}$$

kein Gradientenfeld. Für das Vektorfeld

$$\mathbf{w}(x, y) = \frac{1}{xy}\mathbf{v}(x, y)$$

erhalten wir dagegen

$$\begin{aligned} \mathbf{w}_{1,y} &= \frac{\partial}{\partial y}\left(-\frac{\gamma}{x} + \delta\right) \\ &= 0 \\ &= \frac{\partial}{\partial x}\left(-\frac{\alpha}{y} + \beta\right) \\ &= \mathbf{w}_{2,x}. \end{aligned}$$

Also ist \mathbf{w} ein Gradientenfeld. Eine Stammfunktion ist

$$\begin{aligned} U(x, y) &= \int \left(\delta - \frac{\gamma}{x}\right) dx + c(y) \\ &= \delta x - \gamma \ln x + c(y). \end{aligned}$$

Einsetzen in

$$U_y = -\frac{\alpha}{y} + \beta$$

ergibt

$$\begin{aligned} c'(y) &= -\frac{\alpha}{y} + \beta \\ \implies c(y) &= -\alpha \ln y + \beta y. \end{aligned}$$

Also sind die Lösungen $(x(t), y(t))$ des Räuber-Beute Modells Niveaulinien der Form

$$U(x, y) = \delta x - \gamma \ln x - \alpha \ln y + \beta y = c.$$

Da die Exponentialfunktion bijektiv ist, können wir statt U auch die Funktion $F = \exp \circ (-U)$ betrachten und erhalten die Niveaulinien

$$\frac{x^\gamma}{e^{\delta x}} \cdot \frac{y^\alpha}{e^{\beta y}} = c.$$

Dies sind eiförmige Kurven im ersten Quadranten. Ist

$$\begin{aligned} x(0) &= \frac{\gamma}{\delta} \\ y(0) &= \frac{\alpha}{\beta}, \end{aligned}$$

so ist die Lösung $(x(t), y(t))$ konstant. Andernfalls liegt sie auf der Niveaulinie mit passendem c .

Man kann zeigen, dass die Lösung periodisch ist, d.h. es gibt ein $T > 0$ mit

$$\begin{aligned} x(t+T) &= x(t) \\ y(t+T) &= y(t) \end{aligned}$$

für alle $t \in \mathbb{R}$.

Die mittlere Populationsgröße über eine Periode T ist

$$\begin{aligned} \bar{x} &= \frac{1}{T} \int_0^T x(t) dt && \text{Beute} \\ \bar{y} &= \frac{1}{T} \int_0^T y(t) dt && \text{Räuber.} \end{aligned}$$

Man kann diese Größen bestimmen, ohne die Periode T und die Lösung $(x(t), y(t))$ explizit zu kennen. Dazu dividiert man die Beutegleichung durch x und integriert von 0 bis T . Dies liefert

$$\begin{aligned} 0 &= \ln x(T) - \ln x(0) && \left| \text{wegen } x(T) = x(0) \right. \\ &= \int_0^T \frac{\dot{x}(t)}{x(t)} dt \\ &= \int_0^T \alpha - \beta y(t) dt \\ &= \alpha T - \beta \int_0^T y(t) dt \\ &= \alpha T - \beta T \bar{y}. \end{aligned}$$

Also ist

$$\bar{y} = \frac{\alpha}{\beta}.$$

Verfährt man analog mit der Räubergleichung, erhält man

$$\bar{x} = \frac{\gamma}{\delta}.$$

Die mittlere Populationsgröße der Beute (des Räubers) hängt also nur vom Verhältnis der Wachstums- zur Sterberate des Räubers (der Beute) ab.

Dies hat interessante Konsequenzen. Wollen wir z.B. die Beute durch ein Gift, das auch für den Räuber schädlich ist, reduzieren, müssen wir α durch $\alpha - \varepsilon$ und γ durch $\gamma + \varepsilon$ ersetzen. Für die mittlere Populationsgröße erhalten wir dann

$$\bar{x} = \frac{\gamma + \varepsilon}{\delta}$$

$$\bar{y} = \frac{\alpha - \varepsilon}{\delta}.$$

Unsere Maßnahme ist also kontraproduktiv: die mittlere Größe der Beutepopulation nimmt sogar zu.

X.3. Systeme linearer Differentialgleichungen

X.3.1. Grundlegende Eigenschaften. Wir betrachten Systeme gDglen der Form

$$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x).$$

Dabei sind $A : I \rightarrow \mathbb{R}^{n \times n}$ eine matrixwertige und $\mathbf{f} : I \rightarrow \mathbb{R}^n$ eine vektorwertige Funktion auf einem Intervall $I \subset \mathbb{R}$. Obige gDgl heißt HOMOGEN, falls $\mathbf{f} = 0$ ist, andernfalls heißt sie INHOMOGEN.

Da die Zuordnung $\mathbf{y} \mapsto \mathbf{y}' - A(x)\mathbf{y}$ linear ist, gelten die gleichen Beobachtungen wie in Abschnitt VI.3.1 (S. 233):

- Sind \mathbf{y}_1 und \mathbf{y}_2 zwei Lösungen der inhomogenen gDgl, so ist $\mathbf{y} = \mathbf{y}_1 - \mathbf{y}_2$ eine Lösung der homogenen gDgl.
- Ist \mathbf{y}_p eine partikuläre Lösung der inhomogenen gDgl und \mathbf{y}_h die allgemeine Lösung der homogenen gDgl, so ist $\mathbf{y}_p + \mathbf{y}_h$ die allgemeine Lösung der inhomogenen gDgl.
- Sind \mathbf{y}_1 und \mathbf{y}_2 zwei Lösungen der homogenen gDgl und α_1, α_2 reelle Zahlen, so ist $\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2$ auch eine Lösung der homogenen gDgl.

Wegen dieser Beobachtungen gehen wir bei der Lösung obiger gDgl wie in Abschnitt VI.3:

Wir bestimmen separat die allgemeine Lösung der homogenen gDgl und eine partikuläre Lösung der inhomogenen gDgl und setzen daraus die allgemeine Lösung der inhomogenen gDgl zusammen.

X.3.2. Fundamentalsysteme. Wir betrachten die homogene lineare gDgl

$$\mathbf{y}' = A(x)\mathbf{y}$$

mit $A : I \rightarrow \mathbb{R}^{n \times n}$. Wie im vorigen Abschnitt bemerkt, ist die Lösungsmenge dieser gDgl ein Vektorraum. Wir behaupten, dass dieser Vektorraum die Dimension n hat. Zum Nachweis dieser Behauptung gehen wir wie folgt vor.

Die Abbildung $F : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$F(x, \mathbf{z}) = A(x)\mathbf{z}$$

ist linear und damit Lipschitz-stetig. Daher besitzt das AWP

$$\begin{aligned} \mathbf{y}' &= A(x)\mathbf{y} \\ \mathbf{y}(x_0) &= \mathbf{y}_0 \end{aligned}$$

zu jedem $x_0 \in I$ und jedem $\mathbf{y}_0 \in \mathbb{R}^n$ eine eindeutige Lösung. Wir halten im Folgenden x_0 fest. Zu den Anfangswerten $\mathbf{y}_0 = \mathbf{e}_1, \dots, \mathbf{y}_0 = \mathbf{e}_n$ erhalten wir dann jeweils eine eindeutige Lösung $\mathbf{y}_1, \dots, \mathbf{y}_n$, d.h.

$$\begin{aligned} \mathbf{y}'_i &= A(x)\mathbf{y}_i \\ \mathbf{y}_i(x) &= \mathbf{e}_i \end{aligned}$$

für $i = 1, \dots, n$. Dabei sind $\mathbf{e}_1, \dots, \mathbf{e}_n$ die kanonischen Einheitsvektoren des \mathbb{R}^n . Offensichtlich sind die Vektoren $\mathbf{y}_1(x_0), \dots, \mathbf{y}_n(x_0)$ linear unabhängig.

Wir nehmen nun an, dass sie für ein $x_1 \neq x_0$ linear abhängig seien. Dann gibt es Zahlen $\alpha_1, \dots, \alpha_n$ mit

$$\alpha_1^2 + \dots + \alpha_n^2 \neq 0$$

und

$$\alpha_1 \mathbf{y}_1(x_1) + \dots + \alpha_n \mathbf{y}_n(x_1) = \mathbf{0}.$$

Setze

$$\mathbf{u} = \alpha_1 \mathbf{y}_1 + \dots + \alpha_n \mathbf{y}_n.$$

Dann löst \mathbf{u} das AWP

$$\begin{aligned} \mathbf{u}' &= A(x)\mathbf{u} \\ \mathbf{u}(x_1) &= \mathbf{0}. \end{aligned}$$

Da dieses AWP die eindeutige Lösung $\mathbf{0}$ hat, folgt $\mathbf{u}(x) = \mathbf{0}$ für alle $x \in I$. Insbesondere folgt

$$\begin{aligned} \mathbf{u}(x_0) = \mathbf{0} &\implies \mathbf{0} = \alpha_1 \mathbf{y}_1(x_0) + \dots + \alpha_n \mathbf{y}_n(x_0) \\ &= \alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n. \end{aligned}$$

Also ist $\alpha_1 = \dots = \alpha_n = 0$. Dies ist ein Widerspruch. Daher sind die Vektoren $\mathbf{y}_1(x), \dots, \mathbf{y}_n(x)$ für alle $x \in I$ linear unabhängig.

Hieraus folgt, dass der Lösungsraum mindestens die Dimension n hat.

Um zu zeigen, dass die Dimension gleich n ist, müssen wir jede beliebige Lösung \mathbf{y} der gDgl als Linearkombination von $\mathbf{y}_1, \dots, \mathbf{y}_n$ darstellen können. Sei also \mathbf{y} irgendeine Lösung der gDgl. Da $\mathbf{e}_1, \dots, \mathbf{e}_n$ eine Basis des \mathbb{R}^n ist, gibt es dann Zahlen β_1, \dots, β_n mit

$$\begin{aligned}\mathbf{y}(x_0) &= \beta_1 \mathbf{e}_1 + \dots + \beta_n \mathbf{e}_n \\ &= \beta_1 \mathbf{y}_1(x_0) + \dots + \beta_n \mathbf{y}_n(x_0).\end{aligned}$$

Setze

$$\mathbf{v} = \beta_1 \mathbf{y}_1 + \dots + \beta_n \mathbf{y}_n$$

und

$$\mathbf{w} = \mathbf{v} - \mathbf{y}.$$

Dann folgt

$$\begin{aligned}\mathbf{w}' &= A\mathbf{w} \\ \mathbf{w}(x_0) &= 0\end{aligned}$$

Da dieses AWP die eindeutige Lösung $\mathbf{w} = 0$ hat, folgt

$$\mathbf{y} = \beta_1 \mathbf{y}_1 + \dots + \beta_n \mathbf{y}_n.$$

Das war zu zeigen.

Wir fassen zusammen:

Die Lösungsmenge der gDgl

$$\mathbf{y}' = A(x)\mathbf{y}$$

ist ein n -dimensionaler Vektorraum. Eine Basis wird von den Lösungen $\mathbf{y}_1, \dots, \mathbf{y}_n$ der AWPe

$$\mathbf{y}'_i = A(x)\mathbf{y}_i$$

$$\mathbf{y}_i(x_0) = \mathbf{e}_i$$

mit $i = 1, \dots, n$ gebildet.

Jede Basis des Lösungsraumes nennt man ein FUNDAMENTALSISTEM für die homogene gDgl. Obige Überlegungen zeigen, wie man ein Fundamentalsystem konstruieren kann.

Ist $\mathbf{y}_1(x), \dots, \mathbf{y}_n(x)$ ein Fundamentalsystem für die homogene gDgl, so ist die Determinante

$$W(x) = \det(\mathbf{y}_1(x), \dots, \mathbf{y}_n(x))$$

für alle $x \in I$ von Null verschieden. W heißt die WRONSKI-DETERMINANTE des Fundamentalsystems. Man kann die Wronski-Determinante

bestimmen, ohne das Fundamentalsystem explizit zu kennen. Man kann nämlich zeigen, dass für beliebiges $x_0 \in I$ gilt

$$W(x) = W(x_0) \exp \left(\int_{x_0}^x \text{Spur } A(s) ds \right)$$

mit

$$\text{Spur } A(s) = A_{11}(s) + A_{22}(s) + \dots + A_{nn}(s).$$

X.3.3. Variation der Konstanten. Wir wollen eine partikuläre Lösung der inhomogenen gDgl

$$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$$

bestimmen. Dazu gehen wir ähnlich wie in Abschnitt VI.2.2 (S. 225) vor. Sei $\mathbf{y}_1(x), \dots, \mathbf{y}_n(x)$ ein Fundamentalsystem für die homogene Gleichung. Die Matrix

$$Y(x) = (\mathbf{y}_1(x), \dots, \mathbf{y}_n(x))$$

ist dann für alle $x \in I$ regulär. Wir machen nun den Ansatz

$$\begin{aligned} \mathbf{y}_p(x) &= c_1(x)\mathbf{y}_1(x) + \dots + c_n(x)\mathbf{y}_n(x) \\ &= Y(x)\mathbf{c}(x) \end{aligned}$$

mit unbestimmten Funktionen $c_1, \dots, c_n : I \rightarrow \mathbb{R}$ und

$$\mathbf{c}(x) = (c_1(x), \dots, c_n(x))^T.$$

Setzen wir diesen Ansatz in die inhomogene gDgl ein und beachten, dass $\mathbf{y}_1, \dots, \mathbf{y}_n$ die homogene gDgl lösen, erhalten wir die Bedingung

$$\begin{aligned} 0 &= \mathbf{y}'_p - A(x)\mathbf{y}_p - \mathbf{f}(x) \\ &= \sum_{i=1}^n \{c_i(x) \underbrace{\mathbf{y}'_i(x)}_{=A(x)\mathbf{y}_i(x)} + c'_i(x)\mathbf{y}_i(x)\} \\ &\quad - \sum_{i=1}^n c_i(x)A(x)\mathbf{y}_i(x) - \mathbf{f}(x) \\ &= \sum_{i=1}^n c'_i(x)\mathbf{y}_i(x) - \mathbf{f}(x) \\ &= Y(x)\mathbf{c}'(x) - \mathbf{f}(x). \end{aligned}$$

Da $Y(x)$ regulär ist, folgt

$$\mathbf{c}'(x) = Y(x)^{-1}\mathbf{f}(x).$$

Diese gDgl für \mathbf{c} kann durch komponentenweise Integration gelöst werden. Dies liefert die gesuchte partikuläre Lösung.

BEISPIEL X.3.1. Wir betrachten die gDgl

$$\mathbf{y}' = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \mathbf{y} + \begin{pmatrix} x \\ x^2 \end{pmatrix}.$$

Gemäß Beispiel X.3.2 (S. 392) in Abschnitt X.3.4 (s.u.) ist

$$\mathbf{y}_1(x) = \frac{1}{\sqrt{2}} e^{2x} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\mathbf{y}_2(x) = \frac{1}{\sqrt{2}} e^{4x} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

ein Fundamentalsystem für die homogene gDgl. (Prüft man hier durch Einsetzen nach!) Es ist

$$Y(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{2x} & e^{4x} \\ e^{2x} & -e^{4x} \end{pmatrix}$$

$$\implies Y(x)^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-2x} & e^{-2x} \\ e^{-4x} & -e^{-4x} \end{pmatrix}.$$

Damit lautet die gDgl für \mathbf{c}

$$\mathbf{c}'(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-2x}(x+x^2) \\ e^{-4x}(x-x^2) \end{pmatrix}.$$

In einer Nebenrechnung erhalten wir durch partielle Integration für $a \in \mathbb{R} \setminus \{0\}$ und $k \in \mathbb{N}^*$

$$\int e^{ax} x^k dx = \frac{1}{a} e^{ax} x^k - \frac{k}{a} \int e^{ax} x^{k-1} dx.$$

Damit ergibt sich

$$\mathbf{c}(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-2x}[-\frac{1}{2}x - \frac{1}{4}] + e^{-2x}[-\frac{1}{2}x^2 - \frac{1}{2}x - \frac{1}{4}] \\ e^{-4x}[-\frac{1}{4}x - \frac{1}{16}] - e^{-4x}[-\frac{1}{4}x^2 - \frac{1}{8}x - \frac{1}{32}] \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-2x}[-\frac{1}{2}x^2 - x - \frac{1}{2}] \\ e^{-4x}[\frac{1}{4}x^2 - \frac{1}{8}x - \frac{1}{32}] \end{pmatrix}.$$

Als partikuläre Lösung erhalten wir somit

$$\mathbf{y}_p(x) = Y(x)\mathbf{c}(x)$$

$$= \frac{1}{2} \begin{pmatrix} -\frac{1}{4}x^2 - \frac{9}{8}x - \frac{17}{32} \\ -\frac{3}{4}x^2 - \frac{7}{8}x - \frac{15}{32} \end{pmatrix}$$

$$= -\frac{1}{64} \begin{pmatrix} 8x^2 + 36x + 17 \\ 24x^2 + 28x + 15 \end{pmatrix}.$$

X.3.4. Homogene Systeme mit konstanter Matrix: Der symmetrische Fall. Wir betrachten die gDgl

$$\mathbf{y}' = A\mathbf{y}$$

mit einer symmetrischen Matrix A , d.h. $A^T = A$. Gemäß Abschnitt II.4.8 (S. 91) besitzt A genau n reelle Eigenwerte $\lambda_1, \dots, \lambda_n$, wobei Eigenwerte gemäß ihrer Vielfachheit gezählt werden, mit zugehörigen Eigenvektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$, die normiert und paarweise orthogonal sind, d.h.

$$\mathbf{u}_i^T \mathbf{u}_i = 1$$

für $i = 1, \dots, n$, und

$$\mathbf{u}_i^T \mathbf{u}_j = 0$$

für $i \neq j$. Setze

$$U = (\mathbf{u}_1, \dots, \mathbf{u}_n).$$

U ist orthogonal, d.h.

$$U^T U = U U^T = I.$$

Zur Bestimmung einer Lösung \mathbf{y} der gDgl setzen wir

$$\mathbf{z} = U^T \mathbf{y}.$$

Dann folgt

$$\begin{aligned} \mathbf{z}' &= U^T \mathbf{y}' \\ &= U^T A \mathbf{y} \\ &= U^T A \underbrace{U U^T}_{=I} \mathbf{y} \\ &= U^T A U \mathbf{z} \\ &= D \mathbf{z} \end{aligned}$$

mit

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Die gDgl für \mathbf{z} besitzt offensichtlich das Fundamentalsystem

$$\begin{aligned} \mathbf{z}_1 &= e^{\lambda_1 x} \mathbf{e}_1 \\ \mathbf{z}_2 &= e^{\lambda_2 x} \mathbf{e}_2 \\ &\vdots \\ \mathbf{z}_n &= e^{\lambda_n x} \mathbf{e}_n. \end{aligned}$$

Wegen

$$U \mathbf{e}_i = \mathbf{u}_i \quad \text{für } i = 1, \dots, n$$

ergibt sich hieraus das Fundamentalsystem

$$\begin{aligned} \mathbf{y}_1 &= e^{\lambda_1 x} \mathbf{u}_1 \\ &\vdots \\ \mathbf{y}_n &= e^{\lambda_n x} \mathbf{u}_n \end{aligned}$$

für die ursprüngliche gDgl.

Wir fassen zusammen:

Bestimmung eines Fundamentalsystems für die gDgl

$$\mathbf{y}' = A\mathbf{y}$$

mit einer symmetrischen Matrix A :

- (1) Bestimme alle Eigenwerte $\lambda_1, \dots, \lambda_n$ von A (mit Berücksichtigung der Vielfachheit) und dazu gehörige normierte, paarweise orthogonale Eigenvektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$.
- (2) Das Fundamentalsystem ist

$$\begin{aligned} \mathbf{y}_1 &= e^{\lambda_1 x} \mathbf{u}_1 \\ &\vdots \\ \mathbf{y}_n &= e^{\lambda_n x} \mathbf{u}_n \end{aligned}$$

BEISPIEL X.3.2. Wir betrachten die gDgl

$$\mathbf{y}' = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \mathbf{y}.$$

Das charakteristische Polynom von A lautet

$$\begin{aligned} \det \begin{pmatrix} 3 - \lambda & -1 \\ -1 & 3 - \lambda \end{pmatrix} &= (\lambda - 3)^2 - 1 \\ &= (\lambda - 3 - 1)(\lambda - 3 + 1) \\ &= (\lambda - 4)(\lambda - 2). \end{aligned}$$

Also hat A die Eigenwerte 2 und 4. Die Bestimmungsgleichung für einen Eigenvektor zum Eigenwert 2 lautet

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

Die Bestimmungsgleichung für einen Eigenvektor zum Eigenwert 4 ist

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

Das Fundamentalsystem lautet daher (vgl. Beispiel X.3.1 (S. 390))

$$\mathbf{y}_1(x) = \frac{1}{\sqrt{2}} e^{2x} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\mathbf{y}_2(x) = \frac{1}{2} e^{4x} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

X.3.5. Homogene Systeme mit konstanter Matrix: Der allgemeine Fall. Wir betrachten die gDgl

$$\mathbf{y}' = A\mathbf{y}$$

mit einer nicht symmetrischen Matrix A .

Im Vergleich zum vorigen Abschnitt treten jetzt zwei Probleme auf:

- Die Matrix A hat i.a. komplexe Eigenwerte.
- Die Matrix A ist u.U. nicht diagonalisierbar, d.h. sie besitzt Eigenwerte, deren algebraische Vielfachheit größer ist als ihre geometrische Vielfachheit.

Das erste Problem lösen wir, indem wir wie in Abschnitt VI.3.2 (S. 233) auch komplexe Lösungen betrachten. Bei dem zweiten Problem hilft uns das Konzept der Hauptvektoren aus Abschnitt II.4.10 (S. 93).

Sei dazu $\lambda \in \mathbb{C}$ ein Eigenwert von A und $\mathbf{v} \in \mathbb{C}^n$ ein zugehöriger Hauptvektor der Stufe ℓ , $\ell \geq 1$, d.h.

$$(A - \lambda I)^\ell \mathbf{v} = 0,$$

$$(A - \lambda I)^{\ell-1} \mathbf{v} \neq 0.$$

Wegen

$$\frac{d}{dx} e^{\lambda x} = \lambda e^{\lambda x}$$

liefert Einsetzen von

$$\mathbf{y}(t) = e^{\lambda x} [\mathbf{v} + x(A - \lambda I)\mathbf{v} + \frac{x^2}{2!}(A - \lambda I)^2\mathbf{v} + \dots$$

$$+ \frac{x^{\ell-1}}{(\ell-1)!}(A - \lambda I)^{\ell-1}\mathbf{v}]$$

in die gDgl

$$\mathbf{y}' - A\mathbf{y} = \lambda e^{\lambda x} [\mathbf{v} + x(A - \lambda I)\mathbf{v} + \dots + \frac{x^{\ell-1}}{(\ell-1)!}(A - \lambda I)^{\ell-1}\mathbf{v}]$$

$$+ e^{\lambda x} [(A - \lambda I)\mathbf{v} + \dots + \frac{x^{\ell-2}}{(\ell-2)!}(A - \lambda I)^{\ell-1}\mathbf{v}]$$

$$- e^{\lambda x} [A\mathbf{v} + xA(A - \lambda I)\mathbf{v} + \dots + \frac{x^{\ell-1}}{(\ell-1)!}A(A - \lambda I)^{\ell-1}\mathbf{v}]$$

$$= e^{\lambda x} [(\lambda I - A)\mathbf{v} + x(\lambda I - A)(A - \lambda I)\mathbf{v} +$$

$$\dots + \frac{x^{\ell-1}}{(\ell-1)!}(\lambda I - A)(A - \lambda I)^{\ell-1}\mathbf{v}]$$

$$\begin{aligned}
& + e^{\lambda x}[(A - \lambda I)\mathbf{v} + \dots + \frac{x^{\ell-2}}{(\ell-2)!}(A - \lambda I)^{\ell-1}\mathbf{v}] \\
& = -e^{\lambda x} \frac{x^{\ell-1}}{(\ell-1)!}(A - \lambda I)^{\ell}\mathbf{v} \\
& = 0.
\end{aligned}$$

Also ist \mathbf{y} eine Lösung der gDgl.

Da gemäß Abschnitt II.4.10 (S. 93) eine Basis des \mathbb{C}^n existiert, die aus lauter Hauptvektoren von A besteht, erhalten wir ein Fundamentalsystem *komplexer* Lösungen, wenn wir zu jedem dieser Hauptvektoren und dem entsprechenden Eigenwert obige Konstruktion durchführen.

Das charakteristische Polynom $\det(A - \lambda I)$ hat lauter reelle Koeffizienten. Damit ist mit $\lambda \in \mathbb{C} \setminus \mathbb{R}$ auch $\bar{\lambda}$ ein Eigenwert und beide haben die gleiche algebraische Vielfachheit. Bezeichne mit m die algebraische Vielfachheit des Eigenwertes λ und mit

$$\mathbf{y}_1, \dots, \mathbf{y}_m$$

und

$$\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_m$$

die mit obiger Konstruktion gewonnenen Lösungen der gDgl zu den Eigenwerten λ und $\bar{\lambda}$ und den zugehörigen Hauptvektoren. Dann kann man zeigen, dass

$$\operatorname{Re} \mathbf{y}_1, \operatorname{Im} \mathbf{y}_1, \dots, \operatorname{Re} \mathbf{y}_m, \operatorname{Im} \mathbf{y}_m$$

und

$$\operatorname{Re} \tilde{\mathbf{y}}_1, \operatorname{Im} \tilde{\mathbf{y}}_1, \dots, \operatorname{Re} \tilde{\mathbf{y}}_m, \operatorname{Im} \tilde{\mathbf{y}}_m$$

die gleichen Vektorräume erzeugen. Daher erhalten wir ein Fundamentalsystem *reeller* Lösungen, wenn wir bei jedem Paar $\lambda, \bar{\lambda}$ von komplexen Eigenwerten die Real- und Imaginärteile der komplexen Lösungen zu λ betrachten.

Zusammenfassend erhalten wir folgende Vorschrift:

Bestimmung eines Fundamentalsystems der gDgl

$$\mathbf{y}' = A\mathbf{y}$$

mit einer nicht symmetrischen Matrix A :

- (1) Bestimme alle Eigenwerte von A .
- (2) Bestimme eine Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ des \mathbb{C}^n aus Hauptvektoren von A .
- (3) Für jeden Vektor \mathbf{v}_i und zugehörigen Eigenwert λ setze

$$\begin{aligned}
\mathbf{y}(x) = e^{\lambda x}[\mathbf{v}_i + \dots \\
+ \frac{x^{\ell-1}}{(\ell-1)!}(A - \lambda I)^{\ell-1}\mathbf{v}_i],
\end{aligned}$$

wobei ℓ die Stufe von \mathbf{v}_i ist.

- (4) Für jedes Paar $\lambda, \bar{\lambda}$ komplexer Eigenwerte bestimme $\operatorname{Re} \mathbf{y}$ und $\operatorname{Im} \mathbf{y}$ für alle Lösungen \mathbf{y} zum Eigenwert λ aus Teil (3).

BEISPIEL X.3.3. Wir betrachten die gDgl

$$\mathbf{y}' = \begin{pmatrix} 0 & 1 & -1 \\ -2 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix} \mathbf{y}.$$

Das charakteristische Polynom lautet

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

Die Bestimmungsgleichung für einen Eigenvektor zum Eigenwert 2 ist

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

Die Bestimmungsgleichung für einen Eigenvektor zum Eigenwert 1 lautet

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

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

Insbesondere hat der Eigenwert 1 die algebraische Vielfachheit 2 und die geometrische Vielfachheit 1. Die Bestimmungsgleichung für einen Hauptvektor der Stufe 2 zum Eigenwert 1 lautet

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

Wegen

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

erhalten wir das Fundamentalsystem

$$\begin{aligned} \mathbf{y}_1(x) &= e^{2x} \mathbf{u} \\ &= e^{2x} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \\ \mathbf{y}_2(x) &= e^x \mathbf{v} \\ &= e^x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \\ \mathbf{y}_3(x) &= e^x [\mathbf{w} + x(A - I)\mathbf{w}] \\ &= e^x [\mathbf{w} - x(I - A)\mathbf{w}] \\ &= e^x \begin{pmatrix} 1 - x \\ 1 - x \\ 1 \end{pmatrix}. \end{aligned}$$

BEISPIEL X.3.4. Wir betrachten die gDgl

$$\mathbf{y}' = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ -1 & 1 & 0 \end{pmatrix} \mathbf{y}.$$

Das charakteristische Polynom lautet

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

Die Bestimmungsgleichung für einen Eigenvektor zum Eigenwert -2 ist

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

Für den Eigenwert $1 + i$ ergibt sich die Bestimmungsgleichung

$$\begin{aligned} & \begin{pmatrix} -1 - i & 2 & 0 \\ 0 & -1 - i & 2 \\ -1 & 1 & -1 - i \end{pmatrix} \mathbf{v} = 0 \\ \implies & \begin{pmatrix} -1 & 1 & -1 - i \\ 0 & -1 - i & 2 \\ 0 & 1 - i & 2i \end{pmatrix} \mathbf{v} = 0 \\ \implies & \mathbf{v} = \begin{pmatrix} 4 \\ 2 + 2i \\ 2i \end{pmatrix}. \end{aligned}$$

Damit erhalten wir das Fundamentalsystem

$$\begin{aligned} \mathbf{y}_1(x) &= e^{-2x} \mathbf{u} \\ &= e^{-2x} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned}
\mathbf{y}_2(x) &= \operatorname{Re} [e^{(1+i)x} \mathbf{v}] \\
&= \operatorname{Re} \left[e^x [\cos x + i \sin x] \begin{pmatrix} 4 \\ 2 + 2i \\ 2i \end{pmatrix} \right] \\
&= e^x \begin{pmatrix} 4 \cos x \\ 2 \cos x - 2 \sin x \\ -2 \sin x \end{pmatrix}, \\
\mathbf{y}_3(x) &= \operatorname{Im} [e^{(1+i)x} \mathbf{v}] \\
&= \operatorname{Im} \left[e^x [\cos x + i \sin x] \begin{pmatrix} 4 \\ 2 + 2i \\ 2i \end{pmatrix} \right] \\
&= e^x \begin{pmatrix} 4 \sin x \\ 2 \cos x + 2 \sin x \\ 2 \cos x \end{pmatrix}.
\end{aligned}$$

X.3.6. Lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten. Wir betrachten gDglen der Form

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = f(x).$$

Gemäß Abschnitt X.1.2 (S. 376) können wir die gDgl auch als System 1. Ordnung schreiben. Dazu definieren wir

$$\mathbf{z} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}.$$

\mathbf{z} genügt dann der gDgl

$$\mathbf{z}' = A\mathbf{z} + \mathbf{F}$$

mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix},$$

$$\mathbf{F} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f \end{pmatrix}.$$

Auf diese gDgl können wir die Ergebnisse der vorigen Abschnitte anwenden. Das charakteristische Polynom von A lautet

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

Für eine k -fache Nullstelle λ von p erhalten wir daher für die homogene gDgl n -ter Ordnung die Lösungen

$$\begin{aligned} y_1(x) &= e^{\lambda x} \\ y_2(x) &= x e^{\lambda x} \\ &\vdots \\ y_k(x) &= x^{k-1} e^{\lambda x}. \end{aligned}$$

Diese sind für $\lambda \in \mathbb{R}$ reell.

Falls λ komplex ist, ist $\bar{\lambda}$ auch eine Nullstelle gleicher Vielfachheit, und die reellen Lösungen lauten mit $\lambda = \alpha + i\omega$

$$\begin{aligned} y_1(x) &= e^{\alpha x} \cos(\omega x), & y_2(x) &= e^{\alpha x} \sin(\omega x), \\ y_3(x) &= x e^{\alpha x} \cos(\omega x), & y_4(x) &= x e^{\alpha x} \sin(\omega x), \\ &\vdots & &\vdots \\ y_{2k-1}(x) &= x^{k-1} e^{\alpha x} \cos(\omega x), & y_{2k}(x) &= x^{k-1} e^{\alpha x} \sin(\omega x). \end{aligned}$$

Für die Bestimmung einer partikulären Lösung der inhomogenen gDgl n -ter Ordnung gehen wir wie in Abschnitt X.3.3 (S. 389) vor. Die Matrix Y hat jetzt die Form

$$Y(x) = \begin{pmatrix} y_1(x) & y_2(x) & \dots & y_n(x) \\ y_1'(x) & y_2'(x) & \dots & y_n'(x) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \dots & y_n^{(n-1)}(x) \end{pmatrix},$$

wobei y_1, \dots, y_n ein Fundamentalsystem für die homogene gDgl n -ter Ordnung ist. Die Koeffizienten c_1, \dots, c_n in dem Ansatz

$$y_p(x) = c_1(x)y_1(x) + \dots + c_n(x)y_n(x)$$

sind wegen

$$\mathbf{F}(x) = f(x)\mathbf{e}_n$$

gegeben durch

$$c_i(x) = (-1)^{n-i} \int \frac{\det Y_i(x)}{\det Y(x)} f(x) dx,$$

wobei die Matrix Y_i durch Streichen der i -ten Spalte und der n -ten Zeile der Matrix Y entsteht.

BEISPIEL X.3.5. Betrachte die gDgl

$$y''' - y'' - y' + y = x.$$

Das charakteristische Polynom lautet

$$\begin{aligned} \lambda^3 - \lambda^2 - \lambda + 1 &= (\lambda - 1)(\lambda^2 - 1) \\ &= (\lambda - 1)^2(\lambda + 1). \end{aligned}$$

Also ist

$$\begin{aligned}y_1(x) &= e^x, \\y_2(x) &= xe^x, \\y_3(x) &= e^{-x}\end{aligned}$$

ein Fundamentalsystem für die homogene Gleichung. Für die Matrizen Y und Y_1, Y_2, Y_3 erhalten wir

$$\begin{aligned}Y(x) &= \begin{pmatrix} e^x & xe^x & e^{-x} \\ e^x & (x+1)e^x & -e^{-x} \\ e^x & (x+2)e^x & e^{-x} \end{pmatrix} \\ \implies \det Y(x) &= e^x \det \begin{pmatrix} 1 & x & 1 \\ 1 & x+1 & -1 \\ 1 & x+2 & 1 \end{pmatrix} \\ &= 4e^x \\ Y_1(x) &= \begin{pmatrix} xe^x & e^{-x} \\ (x+1)e^x & -e^{-x} \end{pmatrix} \\ \implies \det Y_1(x) &= -2x - 1 \\ Y_2(x) &= \begin{pmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{pmatrix} \\ \implies \det Y_2(x) &= -2 \\ Y_3(x) &= \begin{pmatrix} e^x & xe^x \\ e^x & (x+1)e^x \end{pmatrix} \\ \implies \det Y_3(x) &= e^{2x}.\end{aligned}$$

Damit ergibt sich die partikuläre Lösung der inhomogenen gDgl

$$\begin{aligned}y_p(x) &= e^x \int \frac{1}{4} e^{-x} (-2x - 1) x dx \\ &\quad - xe^x \int -\frac{2}{4} e^{-x} x dx \\ &\quad + e^{-x} \int \frac{1}{4} e^{-x} e^{2x} x dx \\ &= -\frac{2}{4} (-x^2 - 2x - 2) - \frac{1}{4} (-x - 1) \\ &\quad + \frac{2}{4} x(-x - 1) \\ &\quad + \frac{1}{4} (x - 1) \\ &= x + 1.\end{aligned}$$

Die allgemeine Lösung der inhomogenen gDgl lautet dementsprechend

$$y(x) = x + 1 + ae^x + bxe^x + ce^{-x}$$

mit $a, b, c \in \mathbb{R}$.

X.4. Stabilität

X.4.1. Motivation. Wir betrachten das AWP

$$(X.4.1) \quad \begin{aligned} \mathbf{y}' &= A\mathbf{y} \\ \mathbf{y}(0) &= \mathbf{y}_0 \end{aligned}$$

mit einer symmetrischen $n \times n$ Matrix A .

Gemäß Abschnitt X.1.4 (S. 378) besitzt dieses AWP zu jedem Anfangswert \mathbf{y}_0 eine eindeutige Lösung, die stetig von dem Anfangswert abhängt. Um die Abhängigkeit von dem Anfangswert zu betonen, bezeichnen wir diese Lösung mit $\mathbf{y}(t; \mathbf{y}_0)$.

Gemäß Abschnitt X.1.4 folgt für alle $t > 0$ und zwei beliebige Anfangswerte $\mathbf{u}_0, \mathbf{v}_0$

$$(X.4.2) \quad \|\mathbf{y}(t; \mathbf{u}_0) - \mathbf{y}(t; \mathbf{v}_0)\| \leq e^{Lt} \|\mathbf{u}_0 - \mathbf{v}_0\|$$

mit

$$L = \max\{|\lambda| : \lambda \text{ ist EW von } A\}.$$

Diese Abschätzung besagt, dass die Lösung des AWP (X.4.1) zwar stetig von den Anfangswerten abhängt, dass aber Lösungen zu verschiedenen Anfangswerten exponentiell schnell auseinander laufen können. Die folgenden Beispiele zeigen, dass diese Abschätzung den allgemeinen Fall zwar korrekt widerspiegelt, dass sie aber in Spezialfällen viel zu pessimistisch ist.

BEISPIEL X.4.1. Wir betrachten das AWP (X.4.1) mit

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

Für den Anfangswert $\mathbf{y}_0 = 0$ erhalten wir natürlich die Lösung

$$\mathbf{y}(t; \mathbf{y}_0) = 0 \quad \text{für alle } t > 0.$$

Die Eigenwerte von A sind offensichtlich ± 1 . Damit ergibt sich in Abschätzung (X.4.2) $L = 1$, und wir erhalten für einen beliebigen Anfangswert $\mathbf{u}_0 \neq 0$ die Abschätzung

$$(X.4.3) \quad \|\mathbf{y}(t; \mathbf{u}_0)\| \leq e^t \|\mathbf{u}_0\|.$$

Wir wählen nun den Anfangswert

$$\mathbf{v}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Da

$$A\mathbf{v}_0 = \mathbf{v}_0$$

ist, lautet die Lösung des AWP (X.4.1)

$$\mathbf{y}(t; \mathbf{v}_0) = e^t \mathbf{v}_0 \quad \text{für alle } t > 0.$$

Für diesen Anfangswert ist also die Abschätzung (X.4.3) scharf. Wir betrachten nun den Anfangswert

$$\mathbf{w}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Da

$$A\mathbf{w}_0 = -\mathbf{w}_0$$

ist, lautet die Lösung des AWP (X.4.1)

$$\mathbf{y}(t; \mathbf{w}_0) = e^{-t}\mathbf{w}_0 \quad \text{für alle } t > 0.$$

Inbesondere ist

$$\lim_{t \rightarrow \infty} \mathbf{y}(t; \mathbf{w}_0) = 0.$$

Die Abschätzung (X.4.3) ist also in diesem Fall viel zu pessimistisch.

BEISPIEL X.4.2. Wir betrachten jetzt das AWP (X.4.1) mit

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

Für $\mathbf{y}_0 = 0$ ist wieder $\mathbf{y}(t; \mathbf{y}_0) = 0$ für alle $t > 0$.

Eine leichte Rechnung zeigt, dass

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{ein Eigenvektor zum Eigenwert } -2 \text{ und}$$

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{ein Eigenvektor zum Eigenwert } -4 \text{ ist.}$$

Daher ergibt die Abschätzung (X.4.2) für einen beliebigen Anfangswert $\mathbf{u}_0 \neq 0$

$$(X.4.4) \quad \|\mathbf{y}(t; \mathbf{u}_0)\| \leq e^{4t}\|\mathbf{u}_0\|.$$

Andererseits folgt aus Abschnitt X.3.5 (S. 393)

$$\mathbf{y}(t; \mathbf{u}_0) = \alpha e^{-2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta e^{-4t} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

wobei die Zahlen α, β eindeutig bestimmt sind durch die Bedingung

$$\alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \mathbf{u}_0.$$

Daher gilt für alle Anfangswerte \mathbf{u}_0

$$\lim_{t \rightarrow \infty} \mathbf{y}(t; \mathbf{u}_0) = 0.$$

Die Abschätzung (X.4.4) ist also in jedem Fall viel zu pessimistisch.

Ziel der folgenden Abschnitte ist es, die in den Beispielen X.4.1 und X.4.2 beobachteten Phänomene genauer zu beschreiben.

X.4.2. Stabilitätsbegriffe. Sei $U \subset \mathbb{R}^n$, eine offene, nicht leere Menge und $\mathbf{f} : \mathbb{R} \times U \rightarrow \mathbb{R}^n$ auf $\mathbb{R} \times U$ Lipschitz-stetig bzgl. U . Für beliebiges $t_0 \in \mathbb{R}$ und $\mathbf{y}_0 \in U$ betrachten wir das AWP

$$(X.4.5) \quad \begin{aligned} \mathbf{y}' &= \mathbf{f}(t, \mathbf{y}(t)) \\ \mathbf{y}(t_0) &= \mathbf{y}_0. \end{aligned}$$

Gemäß dem Satz von Picard-Lindelöf aus Abschnitt X.1.4 (S. 378) besitzt das AWP (X.4.5) eine eindeutige Lösung. Wir bezeichnen diese mit $\mathbf{y}(t; t_0, \mathbf{y}_0)$ und setzen voraus, dass sie für alle Zeiten $t \geq t_0$ existiert. Ist speziell $t_0 = 0$, so lassen wir im folgenden das Argument t_0 weg.

DEFINITION X.4.3. (1) Die Lösung $\mathbf{y}(t; t_0, \mathbf{y}_0)$ von (X.4.5) heißt LJAPUNOV-STABIL oder kurz STABIL, wenn es ein $R > 0$ und $C > 0$ gibt, so dass für alle \mathbf{y}_1 mit $\|\mathbf{y}_1 - \mathbf{y}_0\| \leq R$ und alle $t \geq t_0$ gilt

$$\|\mathbf{y}(t; t_0, \mathbf{y}_0) - \mathbf{y}(t; t_0, \mathbf{y}_1)\| \leq C\|\mathbf{y}_0 - \mathbf{y}_1\|.$$

Andernfalls heißt die Lösung $\mathbf{y}(t; t_0, \mathbf{y}_0)$ INSTABIL.

(2) Die Lösung $\mathbf{y}(t; t_0, \mathbf{y}_0)$ heißt ASYMPTOTISCH STABIL, wenn es ein $R > 0$ und eine stetige, nicht negative Funktion $\gamma : [t_0, \infty) \rightarrow \mathbb{R}$ mit

$$\lim_{t \rightarrow \infty} \gamma(t) = 0$$

gibt, so dass für alle \mathbf{y}_1 mit $\|\mathbf{y}_1 - \mathbf{y}_0\| \leq R$ und alle $t \geq t_0$ gilt

$$\|\mathbf{y}(t; t_0, \mathbf{y}_0) - \mathbf{y}(t; t_0, \mathbf{y}_1)\| \leq \gamma(t)\|\mathbf{y}_0 - \mathbf{y}_1\|.$$

BEMERKUNG X.4.4. Ljapunov-Stabilität der Lösung $\mathbf{y}(t; t_0, \mathbf{y}_0)$ bedeutet, dass für alle Anfangswerte \mathbf{y}_1 nahe bei \mathbf{y}_0 die Lösungen $\mathbf{y}(t; t_0, \mathbf{y}_0)$ und $\mathbf{y}(t; t_0, \mathbf{y}_1)$ für $t \rightarrow \infty$ nicht auseinander laufen. Asymptotische Stabilität bedeutet, dass diese Lösungen für $t \rightarrow \infty$ sogar ineinander laufen (vgl. Abb. X.4.1).

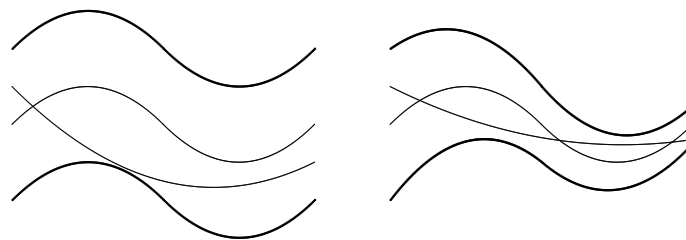


ABBILDUNG X.4.1. Stabile (links) und asymptotisch stabile (rechts) Lösungen einer Differentialgleichung

BEISPIEL X.4.5. Die Nulllösung $\mathbf{y}(t; 0)$ aus Beispiel X.4.1 ist nicht stabil. Die Nulllösung aus Beispiel X.4.2 ist asymptotisch stabil.

X.4.3. Stabilitätskriterien. Wir betrachten zunächst lineare AWP mit konstanten Koeffizienten. Aus den Abschnitten X.3.4 (S. 391) und X.3.5 (S. 393) folgt:

STABILITÄTSKRITERIUM FÜR LINEARE AWP MIT KONSTANTEN KOEFFIZIENTEN: Sei A eine reelle, nicht notwendig symmetrische $n \times n$ Matrix. Für $\mathbf{y}_0 \in \mathbb{R}^n$ bezeichne $\mathbf{y}(t; \mathbf{y}_0)$ die Lösung des AWP

$$(X.4.6) \quad \begin{aligned} \mathbf{y}' &= A\mathbf{y} \\ \mathbf{y}(0) &= \mathbf{y}_0. \end{aligned}$$

(1) Die Realteile der Eigenwerte von A seien alle nicht positiv. Zusätzlich sei für alle Eigenwerte mit verschwindendem Realteil die algebraische gleich der geometrischen Vielfachheit. Dann ist jede Lösung von (X.4.6) Ljapunov-stabil.

(2) Die Realteile der Eigenwerte A seien alle negativ. Dann ist jede Lösung von (X.4.6) asymptotisch stabil.

(3) A habe einen Eigenwert mit positivem Realteil. Dann ist jede Lösung von (X.4.6) instabil.

BEISPIEL X.4.6. Die Matrix aus Beispiel X.4.1 hat die Eigenwerte ± 1 . Daher ist für jeden Anfangswert \mathbf{y}_0 die Lösung $\mathbf{y}(t; \mathbf{y}_0)$ des AWP nicht stabil. Die Matrix aus Beispiel X.4.2 dagegen hat lauter negative Eigenwerte. Daher ist für jeden Anfangswert \mathbf{y}_0 die Lösung $\mathbf{y}(t; \mathbf{y}_0)$ des AWP asymptotisch stabil.

Die folgenden beiden Ergebnisse sagen etwas aus über das Stabilitätsverhalten der Lösungen von (X.4.5), wenn (X.4.5) eine „kleine Störung“ einer linearen gDgl mit konstanten Koeffizienten ist. Diese Ergebnisse werden typischerweise genutzt, um das Stabilitätsverhalten einer Linearisierung einer allgemeinen gDgl um eine Lösung zu untersuchen.

STABILITÄT BELIEBIGER LÖSUNGEN FÜR STÖRUNGEN LINEARER AWP MIT KONSTANTEN KOEFFIZIENTEN: Seien A eine reelle $n \times n$ Matrix, $\mathbf{g} : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine stetige Funktion und $t_0 \in \mathbb{R}$. Es gelte:

- Alle Eigenwerte von A haben negativen Realteil.
- Es gibt eine stetige, nicht negative Funktion $k : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\lim_{t \rightarrow \infty} k(t) = 0$$

und

$$\|\mathbf{g}(t, \mathbf{u}) - \mathbf{g}(t, \mathbf{v})\| \leq k(t)\|\mathbf{u} - \mathbf{v}\|$$

für alle $t \geq t_0$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Dann ist jede Lösung des AWP (X.4.5) mit

$$\mathbf{f}(t, \mathbf{x}) = A\mathbf{x} + \mathbf{g}(t, \mathbf{x})$$

asymptotisch stabil.

STABILITÄT DER NULLLÖSUNG FÜR STÖRUNGEN LINEARER AWP MIT KONSTANTEN KOEFFIZIENTEN: Seien A eine reelle $n \times n$ Matrix, $\mathbf{g} : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine stetige Funktion und $t_0 \in \mathbb{R}$. Es gelte:

- Alle Eigenwerte von A haben negativen Realteil.
- Es gibt eine stetige, nicht negative Funktion $k : \mathbb{R} \rightarrow \mathbb{R}$ mit $k(0) = 0$ und

$$\|\mathbf{g}(t, \mathbf{u})\| \leq k(\|\mathbf{u}\|)\|\mathbf{u}\|$$

für alle $t \geq t_0$, $\mathbf{u} \in \mathbb{R}^n$.

Dann ist die Lösung des AWP (X.4.5) mit

$$\mathbf{f}(t, \mathbf{x}) = A\mathbf{x} + \mathbf{g}(t, \mathbf{x})$$

und $\mathbf{y}_0 = 0$ asymptotisch stabil.

X.4.4. Autonome Systeme. Die Stabilitätskriterien des vorigen Abschnittes lassen sich besonders leicht auf autonome AWP anwenden. Wir betrachten daher in diesem Abschnitt speziell AWP der Form

$$(X.4.7) \quad \begin{aligned} \mathbf{y}' &= \mathbf{F}(\mathbf{y}) \\ \mathbf{y}(0) &= \mathbf{y}_0 \end{aligned}$$

mit stetig differenzierbarer Funktion $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Ist \mathbf{y}_0 eine Nullstelle von \mathbf{F} , d.h.

$$\mathbf{F}(\mathbf{y}_0) = 0,$$

so folgt sofort, dass die Lösung von (X.4.7) konstant ist:

$$\mathbf{y}(t; \mathbf{y}_0) = \mathbf{y}_0 \quad \text{für alle } t \in \mathbb{R}.$$

Ist umgekehrt $\mathbf{y}(t; \mathbf{y}_0)$ eine stationäre Lösung von (X.4.7), d.h., gibt es ein $t_1 \in \mathbb{R}$ mit

$$\mathbf{y}'(t_1; \mathbf{y}_0) = 0,$$

so gilt

$$\mathbf{F}(\mathbf{y}(t_1; \mathbf{y}_0)) = 0,$$

und aus dem Eindeutigkeitssatz folgt

$$\mathbf{y}(t; \mathbf{y}_0) = \mathbf{y}(t_1; \mathbf{y}_0) \quad \text{für alle } t \in \mathbb{R}.$$

Stationäre Lösungen von (X.4.7) sind also genau diejenigen Lösungen, die zu einem Anfangswert \mathbf{y}_0 mit $\mathbf{F}(\mathbf{y}_0) = 0$ gehören. Daher heißen die

Nullstellen von \mathbf{F} auch RUHEPUNKTE des AWP (X.4.7). Der folgende Satz charakterisiert die Stabilität solcher Ruhepunkte. Er folgt aus der Stabilität der Nulllösung von Störungen linearer AWP mit konstanten Koeffizienten.

STABILITÄT VON RUHEPUNKTEN AUTONOMER AWP: Sei \mathbf{y}_0 ein Ruhepunkt von (X.4.7).

(1) Die Jacobi Matrix $D\mathbf{F}(\mathbf{y}_0)$ habe lauter Eigenwerte mit negativem Realteil. Dann ist die stationäre Lösung $\mathbf{y}(t; \mathbf{y}_0) = \mathbf{y}_0$ von (X.4.7) asymptotisch stabil.

(2) Die Jacobi Matrix $D\mathbf{F}(\mathbf{y}_0)$ habe einen Eigenwert mit positivem Realteil. Dann ist die stationäre Lösung $\mathbf{y}(t; \mathbf{y}_0) = \mathbf{y}_0$ von (X.4.7) instabil.

BEISPIEL X.4.7. Betrachte das gedämpfte mathematische Pendel

$$\ddot{x} + 2\alpha\dot{x} + \lambda \sin x = 0$$

mit $\alpha > 0$, $\lambda > 0$. Die zugehörige gDgl 1. Ordnung ist

$$\begin{aligned}\dot{x} &= v \\ \dot{v} &= -2\alpha v - \lambda \sin x.\end{aligned}$$

Diese ist von der Form (X.4.7) mit

$$\mathbf{F}(x, v) = \begin{pmatrix} v \\ -\lambda \sin x - 2\alpha v \end{pmatrix}.$$

Die Nullstellen sind genau die Punkte

$$(x_k, v_k) = (k\pi, 0), \quad k \in \mathbb{N}.$$

Die Jacobi Matrix ist

$$\begin{aligned}D\mathbf{F}(x_k, v_k) &= \begin{pmatrix} 0 & 1 \\ -\lambda \cos x_k & -2\alpha \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -(-1)^k \lambda & -2\alpha \end{pmatrix}\end{aligned}$$

und hat die Eigenwerte $-\alpha \pm \sqrt{\alpha^2 - \lambda(-1)^k}$. Für gerades k sind beide Eigenwerte negativ; für ungerades k ist ein Eigenwert positiv und einer negativ. Mithin sind die Ruhepunkte zu geradem k asymptotisch stabil und die zu ungeradem k instabil.

BEISPIEL X.4.8. Betrachte (X.4.7) mit $n = 2$ und

$$\mathbf{F}(x, y) = \begin{pmatrix} -x^2 - y \\ -x + y^2 \end{pmatrix}.$$

Aus

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

folgt

$$x = y^2$$

und

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

Also sind die Nullstellen

$$(0, 0) \quad \text{und} \quad (1, -1).$$

Für die Jacobi Matrix ergibt sich

$$D\mathbf{F}(x, y) = \begin{pmatrix} -2x & -1 \\ -1 & 2y \end{pmatrix}$$

und somit

$$\begin{aligned} D\mathbf{F}(0, 0) &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \\ &\implies \text{Eigenwerte } \pm 1, \\ D\mathbf{F}(1, -1) &= \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix} \\ &\implies \text{Eigenwerte } -3, -1. \end{aligned}$$

Also ist $(0, 0)$ instabil und $(1, -1)$ asymptotisch stabil.

Das folgende Beispiel zeigt, dass im Fall von rein imaginären Eigenwerten keine Aussage über die Stabilität von Ruhepunkten möglich ist.

BEISPIEL X.4.9. Betrachte (X.4.7) mit $n = 2$ und

$$\mathbf{F}(x, y) = \begin{pmatrix} -y + x^3 \\ x + y^3 \end{pmatrix}.$$

Der einzige Ruhepunkt ist $\mathbf{y}_0 = (0, 0)$, und

$$D\mathbf{F}(\mathbf{y}_0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat die Eigenwerte $\pm i$. Sei $\mathbf{u}(t)$ eine Lösung von (X.4.7) zu einem Anfangswert $\mathbf{u}_0 \neq (0, 0)$. Für

$$r(t)^2 = \|\mathbf{u}(t)\|^2 = u_1(t)^2 + u_2(t)^2$$

folgt dann

$$\begin{aligned} r\dot{r} &= u_1\dot{u}_1 + u_2\dot{u}_2 \\ &= u_1(-u_2 + u_1^3) + u_2(u_1 + u_2^3) \\ &= u_1^4 + u_2^4 \\ &> 0. \end{aligned}$$

Also ist $\dot{r}(t) > 0$ und die Lösung läuft von $\mathbf{y}_0 = (0, 0)$ weg, d.h., \mathbf{y}_0 ist instabil.

Betrachte nun (X.4.7) mit $n = 2$ und

$$\mathbf{F}(x, y) = \begin{pmatrix} -y - x^3 \\ x - y^3 \end{pmatrix}.$$

Wieder ist $\mathbf{y}_0 = (0, 0)$ der einzige Ruhepunkt, und

$$D\mathbf{F}(\mathbf{y}_0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat wieder die Eigenwerte $\pm i$. Für $\mathbf{u}(t)$ und $r(t)$ wie oben folgt jetzt

$$\begin{aligned} r\dot{r} &= u_1\dot{u}_1 + u_2\dot{u}_2 \\ &= u_1(-u_2 - u_1^3) + u_2(u_1 - u_2^3) \\ &= -u_1^4 - u_2^4 \\ &< 0. \end{aligned}$$

Also ist $\dot{r}(t) < 0$ und die Lösung läuft in \mathbf{y}_0 hinein, d.h., \mathbf{y}_0 ist stabil.

KAPITEL XI

Stochastik I Diskrete Modelle

XI.1. Modelle für Zufallsexperimente

XI.1.1. Endliche Wahrscheinlichkeitsräume. Wir betrachten Zufallsexperimente mit endlich vielen möglichen Ausgängen. Diese werden beschrieben durch eine endliche, nicht leere Menge Ω , deren Elemente ω die Versuchsausgänge bezeichnen. Sie heißen ERGEBNISSE oder ELEMENTAREREIGNISSE. Ω heißt ERGEBNISMENGE.

Die Teilmengen von Ω sind die EREIGNISSE, die in dem Modell in Betracht gezogen werden. Genauer:

Wir identifizieren $A \subset \Omega$ mit dem Ereignis, dass ein $\omega \in A$ der beobachtete Versuchsausgang ist.

Dementsprechend bezeichnen $A \cap B$ bzw. $A \cup B$ die Ereignisse, dass A und B bzw. A oder B eintreten. Die leere Menge \emptyset heißt auch das UNMÖGLICHE EREIGNIS; Ω ist das SICHERE EREIGNIS.

Jedem Ereignis ordnen wir eine Wahrscheinlichkeit zu. Die Menge aller möglichen Ereignisse ist die POTENZMENGE $\mathcal{P}(\Omega)$, d.h. die Menge aller Teilmengen von Ω . Eine Abbildung $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$ heißt WAHRSCHEINLICHKEITSVERTEILUNG oder WAHRSCHEINLICHKEITSMASS, wenn sie folgende Eigenschaften hat:

$P(\Omega) = 1$	(NORMIERUNG)
$P(A) \geq 0$	für alle A (POSITIVITÄT)
$P(A \cup B) = P(A) + P(B)$	für alle disjunkten A, B (ADDITIVITÄT)

$P(A)$ heißt die WAHRSCHEINLICHKEIT des Ereignisses A . Das Paar (Ω, P) heißt der dem Zufallsexperiment zugeordnete WAHRSCHEINLICHKEITSRAUM.

BEISPIEL XI.1.1. Wir wollen die Wahrscheinlichkeit bestimmen, dass die Summe der bei zwei Würfeln eines Würfels erhaltenen Augenzahlen mindestens 10 ist. Wir können die Ergebnisse des Zufallsexperimentes „Zweimaliges Werfen eines Würfels“ durch die Paare (i, k)

der beobachteten Augenzahlen beschreiben. Daher ist

$$\Omega = \{(i, k) : 1 \leq i, k \leq 6\}.$$

Ω hat 36 Elemente. Aus Symmetriegründen ist es naheliegend, sie alle als gleich wahrscheinlich zu betrachten. Jedes (i, k) hat also die Wahrscheinlichkeit $\frac{1}{36}$. Die Menge der Ergebnisse, für die die Summe $i + k$ mindestens 10 ist, ist

$$A = \{(6, 6), (6, 5), (5, 6), (6, 4), (4, 6), (5, 5)\}.$$

Da A genau 6 Elemente hat, ist die gesuchte Wahrscheinlichkeit $\frac{6}{36} = \frac{1}{6}$.

Aus den obigen Eigenschaften eines Wahrscheinlichkeitsmaßes folgen leicht weitere Eigenschaften. Für $A, B, A_i \in \mathcal{P}(\Omega)$ gilt

$$\begin{aligned} P(\Omega \setminus A) &= 1 - P(A) \\ P(\emptyset) &= 0 \\ A \subset B &\Rightarrow P(A) \leq P(B) \\ P(A \setminus B) &= P(A) - P(A \cap B) \end{aligned}$$

und

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) \quad \text{falls } A_1, \dots, A_n \text{ paarweise disjunkt} \\ P\left(\bigcup_{i=1}^n A_i\right) &\leq \sum_{i=1}^n P(A_i) \quad \text{für beliebige } A_1, \dots, A_n \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Aus diesen Eigenschaften folgt insbesondere

$$P(A) = \sum_{\omega \in A} P(\{\omega\}).$$

Die Wahrscheinlichkeit von A ist also die Summe der Wahrscheinlichkeiten der Ergebnisse, bei denen A eintritt. P ist also eindeutig bestimmt durch die Werte aller $P(\{\omega\})$ mit $\omega \in \Omega$. Man schreibt auch $P(\omega)$ statt $P(\{\omega\})$. Die Abbildung $\omega \mapsto P(\omega)$ heißt WAHRSCHEINLICHKEITSFUNKTION. Die Wahrscheinlichkeitsverteilung kann also auch durch Angabe der zugehörigen Wahrscheinlichkeitsfunktion beschrieben werden. Offensichtlich gilt für jede Wahrscheinlichkeitsfunktion:

$$\begin{aligned} P(\omega) &\geq 0 \quad \text{für alle } \omega \in \Omega \\ \sum_{\omega \in \Omega} P(\omega) &= 1. \end{aligned}$$

Ein besonders wichtiger Spezialfall einer Wahrscheinlichkeitsfunktion ist die GLEICHVERTEILUNG auf Ω , bei der alle Ergebnisse $\omega \in \Omega$ gleich wahrscheinlich sind. In diesem Fall wird (Ω, P) als LAPLACESCHER WAHRSCHEINLICHKEITSRAUM bezeichnet; das zugehörige Zufallsexperiment heißt LAPLACE-EXPERIMENT. Offensichtlich gilt dann:

$$\begin{aligned} P(\omega) &= \frac{1}{\text{card}(\Omega)} && \text{für alle } \omega \in \Omega, \\ P(A) &= \frac{\text{card}(A)}{\text{card}(\Omega)} && \text{für alle } A \subset \Omega. \end{aligned}$$

XI.1.2. Urnenmodelle. Viele Laplace-Experimente lassen sich durch Urnenmodelle beschreiben. Dabei stellen wir uns vor, dass wir aus einer Urne mit N Kugeln, die von 1 bis N nummeriert sind, sukzessive n Kugeln zufällig ziehen.

Offensichtlich sind zwei Arten des Ziehens zu unterscheiden:

- (1) MIT RÜCKLEGEN: Nach jedem Zug wird die gerade gezogene Kugel wieder in die Urne zurückgelegt; jede Kugel kann mehrmals gezogen werden.
- (2) OHNE RÜCKLEGEN: eine gezogene Kugel wird nicht zurückgelegt; jede Kugel kann nur einmal gezogen werden.

Ein Beispiel für (1) ist das mehrmalige Werfen eines Würfels; ein Beispiel für (2) ist die Ziehung der Lottozahlen.

Man kann das Ergebnis der Folge der Ziehungen dadurch beschreiben, dass man das n -Tupel $(\omega_1, \dots, \omega_n)$ angibt, in dem ω_i die Nummer der im i -ten Zug gezogenen Kugel ist. Offensichtlich kann man hier wiederum zwei Arten von Ziehen unterscheiden:

- (i) MIT REIHENFOLGE: Es kommt auf die Reihenfolge des Erscheinens an; $(1, 2)$ und $(2, 1)$ bezeichnen verschiedene Ergebnisse.
- (ii) OHNE REIHENFOLGE: Es kommt nicht auf die Reihenfolge des Erscheinens an; $(1, 2)$ und $(2, 1)$ bezeichnen gleiche Ergebnisse.

Ein Beispiel für (i) ist das Bestimmen einer Losnummer durch sukzessives Ziehen ihrer Ziffern; ein Beispiel für (ii) ist wieder die Ziehung der Lottozahlen.

Durch diese Unterscheidungen ergeben sich insgesamt vier verschiedene Ergebnismengen. Zu ihrer Beschreibung setzen wir

$$\mathbb{A} = \{1, 2, \dots, N\}.$$

(I) STICHPROBEN IN REIHENFOLGE MIT RÜCKLEGEN (1i):

$$\begin{aligned} \Omega_I &= \mathbb{A}^n, \\ \text{card}(\Omega_I) &= N^n. \end{aligned}$$

(II) STICHPROBEN IN REIHENFOLGE OHNE RÜCKLEGEN (2i):

$$\begin{aligned}\Omega_{II} &= \{(\omega_1, \dots, \omega_n) \in \mathbb{A}^n : \omega_i \neq \omega_j \text{ für } i \neq j\} \\ \text{card}(\Omega_{II}) &= \frac{N!}{(N-n)!} \\ &= N \cdot (N-1) \cdot \dots \cdot (N-n+1).\end{aligned}$$

(III) STICHPROBEN OHNE REIHENFOLGE OHNE RÜCKLEGEN (2ii):

$$\begin{aligned}\Omega_{III} &= \{(\omega_1, \dots, \omega_n) \in \mathbb{A}^n : \omega_1 < \omega_2 < \dots < \omega_n\} \\ \text{card}(\Omega_{III}) &= \binom{N}{n} \\ &= \frac{N!}{(N-n)!n!}.\end{aligned}$$

(IV) STICHPROBEN OHNE REIHENFOLGE MIT RÜCKLEGEN (1ii):

$$\begin{aligned}\Omega_{IV} &= \{(\omega_1, \dots, \omega_n) \in \mathbb{A}^n : \omega_1 \leq \omega_2 \leq \dots \leq \omega_n\} \\ \text{card}(\Omega_{IV}) &= \binom{N+n-1}{n}.\end{aligned}$$

Man kann diese Urnenmodelle auch alternativ wie folgt interpretieren:

Gefragt ist nach der Anzahl der Möglichkeiten, n Murmeln auf N Plätze zu verteilen. Dabei entspricht die Nummer der Ziehung der Nummer der Murmel und die Nummer der Kugel der Nummer des Platzes.

XI.1.3. Anwendungsbeispiele.

BEISPIEL XI.1.2. Es werden vier völlig gleich aussehende Würfel gleichzeitig geworfen. Welches ist die Wahrscheinlichkeit p , dass die vier erscheinenden Augenzahlen verschieden sind?

Da ein gleichzeitiger Wurf von vier Würfeln dem viermaligen Werfen eines Würfels entspricht, handelt es sich um ein Experiment mit Rücklegen. Da die Würfel nicht unterscheidbar sind, könnte man versucht sein, ein Modell ohne Reihenfolge zu betrachten. Dies ist aber *falsch*: Dem Resultat, dass die Augenzahlen 1, 2, 3 und 4 auftreten, entspricht in diesem Modell nur ein Ereignis, da es nicht auf die Reihenfolge ankommt. Tatsächlich entsprechen diesem Resultat jedoch die $4! = 24$ Permutationen der Menge $\{1, 2, 3, 4\}$. Das *richtige* Modell ist also dasjenige mit Berücksichtigung der Reihenfolge. Daher sind alle möglichen Fälle Ω_I mit $N = 6$ und $n = 4$. Die günstigen Fälle (alle Augenzahlen verschieden) sind Ω_{II} mit $N = 6$ und $n = 4$. Daher ist die gesuchte Wahrscheinlichkeit

$$\begin{aligned}p &= \frac{6!}{6^4} \\ &= \frac{3 \cdot 4 \cdot 5 \cdot 6}{6^4}\end{aligned}$$

$$= \frac{5}{18}.$$

BEISPIEL XI.1.3. Wie groß ist die Wahrscheinlichkeit p , dass mindestens 2 von 25 Schülern einer Klasse am gleichen Tag Geburtstag haben?

Als Ereignisraum können wir Ω_I mit $n = 25$ und $N = 365$ wählen. Das Ergebnis $(\omega_1, \dots, \omega_{25})$ bedeutet, dass Schüler Nummer 1 am ω_1 -ten Tag Geburtstag hat, Schüler Nummer 2 am ω_2 -ten Tag usw. Das interessierende Ereignis ist das Komplement des Ereignisses Ω_{II} , das alle Schüler an verschiedenen Tagen Geburtstag haben. Also ist

$$p = 1 - P(\Omega_{II}).$$

Für $P(\Omega_{II})$ erhalten wir

$$\begin{aligned} P(\Omega_{II}) &= \frac{N!}{(N-n)! N^n} \\ &= \frac{N}{N} \cdot \frac{N-1}{N} \cdot \frac{N-2}{N} \cdots \frac{N-n+1}{N} \\ &= 1 \cdot \left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right). \end{aligned}$$

Wenn wir jeden Ausdruck $\left(1 - \frac{k}{N}\right)$ durch $\exp\left(-\frac{k}{N}\right)$ approximieren, erhalten wir

$$\begin{aligned} P(\Omega_{II}) &\approx \exp\left(-\sum_{k=1}^{n-1} \frac{k}{N}\right) \\ &= \exp\left(-\frac{n(n-1)}{2N}\right) \\ &= \exp\left(-\frac{600}{730}\right) \\ &\approx 0.44. \end{aligned}$$

Damit ergibt sich $p \approx 0.56$. Der exakte Wert ohne diese Approximation ist $p = 0.568$.

BEISPIEL XI.1.4. Beim Zahlenlotto „6 aus 49“ werden $n = 6$ Kugeln aus $N = 49$ Kugeln ohne Rücklegen gezogen. Dabei kommt es nicht auf die Reihenfolge an. Also ist die Ereignismenge Ω_{III} mit

$$\begin{aligned} \text{card}(\Omega_{III}) &= \binom{49}{6} \\ &= 13983816. \end{aligned}$$

Daher ist die Wahrscheinlichkeit p_6 für „6 Richtige“

$$\begin{aligned} p_6 &= \frac{1}{\text{card}(\Omega_{III})} \\ &\approx 7.1511 \cdot 10^{-8}. \end{aligned}$$

Wie groß ist die Wahrscheinlichkeit p_4 für „genau 4 Richtige“?

Seien $\tilde{\omega}_1, \dots, \tilde{\omega}_6$ die gezogenen Zahlen. Dann können wir die Menge aller gesuchten Ergebnisse wie folgt erzeugen: Wir ziehen zunächst 4 Kugeln aus $\{\tilde{\omega}_1, \dots, \tilde{\omega}_6\}$ und dann zwei Kugeln aus den 43 Kugeln $\{1, \dots, 49\} \setminus \{\tilde{\omega}_1, \dots, \tilde{\omega}_6\}$. Für das erste Ziehen haben wir

$$\binom{6}{4} = 15$$

Möglichkeiten, für das zweite Ziehen

$$\binom{43}{2} = 21 \cdot 43 = 903$$

Möglichkeiten. Daher ist

$$\begin{aligned} p_4 &= \frac{15 \cdot 903}{13983816} \\ &= \frac{3 \cdot 5 \cdot 43}{7 \cdot 44 \cdot 46 \cdot 47} \\ &\approx 9.68619 \cdot 10^{-4}. \end{aligned}$$

Analog kann man die Wahrscheinlichkeit p_5 für „genau 5 Richtige“ berechnen. Die Wahrscheinlichkeit für „mindestens 4 Richtige“ ist dann

$$p_4 + p_5 + p_6 \approx 9.8714 \cdot 10^{-4}.$$

BEISPIEL XI.1.5. Auf wie viele Arten können sich zwei nicht unterscheidbare Spatzen auf vier Telegraphenleitungen verteilen?

Dies ist Ω_{IV} mit $N = 4$ und $n = 2$. Dementsprechend gibt es

$$\binom{5}{2} = 10$$

Möglichkeiten.

XI.1.4. Die hypergeometrische Verteilung. In Beispiel XI.1.4 haben wir einen Spezialfall dieser wichtigen Verteilung kennen gelernt. Für den allgemeinen Fall betrachten wir eine Urne, die S schwarze und W weiße Kugeln enthält, insgesamt $N = S + W$ Kugeln. Es werden $n \leq S + W$ Kugeln ohne Zurücklegen gezogen. Die Wahrscheinlichkeit, dass dabei s schwarze und $w = n - s$ weiße Kugeln gezogen werden, beträgt

$$h(s; n, N, S) = \frac{\binom{S}{s} \binom{W}{w}}{\binom{S+W}{n}} \quad (0 \leq s \leq n).$$

Diese Verteilung heißt HYPERGEOMETRISCHE VERTEILUNG.

BEISPIEL XI.1.6. Wie groß ist die Wahrscheinlichkeit, dass beim Skatspiel ein Spieler genau 3 Asse erhält?

Dies entspricht obiger Situation mit $S = 4$ (Zahl aller Asse), $W =$

28 (Zähler aller anderen Karten), $s = 3$ (Zahl der Asse des Spielers) und $n = 10$ (Zahl der Karten des Spielers). Daher ist die gesuchte Wahrscheinlichkeit

$$\begin{aligned} p &= \frac{\binom{4}{3} \binom{28}{7}}{\binom{32}{10}} \\ &= \frac{4 \cdot 22 \cdot 8 \cdot 9 \cdot 10}{32 \cdot 31 \cdot 30 \cdot 29} \\ &= \frac{66}{899} \\ &\approx 7.34\%. \end{aligned}$$

Die Wahrscheinlichkeit, dass mindestens ein Spieler 3 Asse erhält, ist dreimal so groß, da keine zwei Spieler gleichzeitig drei Asse erhalten können.

XI.1.5. Multinomialkoeffizienten. Die Binomialkoeffizienten $\binom{n}{k}$ beschreiben, auf wie viele Arten man eine Menge von n nummerierten Kugeln so in zwei Gruppen einteilen kann, dass die erste Gruppe k Kugeln enthält. Wie viele Möglichkeiten gibt es, die Zahlen $1, \dots, n$ so in r Gruppen einzuteilen, dass die erste Gruppe k_1 Elemente hat, die zweite k_2 Elemente usw.? Dabei muss man natürlich $k_1 + \dots + k_r = n$ voraussetzen. Man kann zuerst auf $\binom{n}{k_1}$ Arten die erste Gruppe auswählen, dann auf $\binom{n-k_1}{k_2}$ Arten die zweite Gruppe usw. Insgesamt erhält man

$$\begin{aligned} &\binom{n}{k_1} \cdot \binom{n-k_1}{k_2} \cdot \binom{n-k_1-k_2}{k_3} \cdot \dots \cdot \binom{n-k_1-\dots-k_{r-1}}{k_r} \\ &= \frac{n!}{k_1!k_2!\dots k_r!} \end{aligned}$$

Möglichkeiten. Diesen Ausdruck bezeichnet man als MULTINOMIALKOEFFIZIENTEN und kürzt ihn mit $\binom{n}{k_1, k_2, \dots, k_r}$ ab:

$$\binom{n}{k_1, k_2, \dots, k_r} = \frac{n!}{k_1!k_2!\dots k_r!} \quad \text{mit } k_1 + \dots + k_r = n.$$

BEISPIEL XI.1.7. 26 Schulkinder haben einen Fußball, vier Tennisschläger, ein Fußballfeld und einen Tennisplatz. Die Zahl der Einteilungen in zwei Fußballmannschaften mit je 11 Spielern und in zwei Tennisteamen mit je 2 Spielern beträgt

$$\begin{aligned} \binom{26}{11, 11, 2, 2} &= \frac{26!}{11!11!2!2!} \\ &= 2 \cdot 5 \cdot 12 \cdot 13 \cdot 14 \cdot 15 \cdot 17 \cdot 19 \cdot 23 \cdot 26 \\ &\approx 6.327 \cdot 10^{10}. \end{aligned}$$

XI.1.6. Identitäten für Binomialkoeffizienten. Aus der Binomischen Formel (vgl. Abschnitt 1.2.8 (S. 23))

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

ergibt sich durch Einsetzen von $x = 1, y = 1$ bzw. $x = -1, y = 1$ bzw. durch Ableiten nach x und anschließendes Einsetzen von $x = 1, y = 1$:

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} &= 2^n, \\ \sum_{k=0}^n (-1)^k \binom{n}{k} &= 0, \\ \sum_{k=1}^n k \binom{n}{k} &= n2^{n-1}. \end{aligned}$$

In Abschnitt 1.2.8 (S. 23) haben wir bereits die folgenden Identitäten kennen gelernt:

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

Da die Summe der Wahrscheinlichkeiten $h(s; n, N, S)$ aus Abschnitt XI.1.4 (S. 414) gleich 1 ist, ergibt sich schließlich

$$\sum_{s=0}^n \binom{S}{s} \binom{W}{n-s} = \binom{S+W}{n} \quad \text{mit } n \leq S+W.$$

XI.2. Bedingte Wahrscheinlichkeit und Unabhängigkeit

XI.2.1. Definition bedingter Wahrscheinlichkeiten. Häufig steht, bevor das Ergebnis eines Zufallsexperimentes bekannt ist, schon die Information zur Verfügung, dass das Ergebnis zu einer gewissen Teilmenge des Ereignisraumes gehört. Z.B. sieht ein Spieler beim Skat seine eigenen Karten. Interessiert sich Spieler 1 für das Ereignis A , dass Spieler 2 zwei Asse hat, so wird er zunächst seine eigenen Asse zählen. Hat er selbst drei oder vier Asse, so ist für ihn die Wahrscheinlichkeit von A natürlich 0; hat er maximal zwei Asse, so ist die Wahrscheinlichkeit von A für ihn positiv.

Für Laplace-Experimente ist der Ansatz für die Definition bedingter Wahrscheinlichkeiten sehr nahe liegend. Waren ursprünglich alle Ergebnisse $\omega \in \Omega$ gleich wahrscheinlich und erhält man nun die Information, dass $\omega \in B$ liegt, so ordnen wir den Ergebnissen in $\Omega \setminus B$ die bedingte Wahrscheinlichkeit 0 zu und betrachten die Ergebnisse aus B als gleich wahrscheinlich unter der bedingten Wahrscheinlichkeit. Dies bedeutet, dass für $A \subset \Omega$ die bedingte Wahrscheinlichkeit von A bei gegebenem B den Wert

$$P(A|B) = \frac{\text{card}(A \cap B)}{\text{card}(B)}$$

erhält. Aus

$$P(A \cap B) = \frac{\text{card}(A \cap B)}{\text{card}(\Omega)}$$

und

$$P(B) = \frac{\text{card}(B)}{\text{card}(\Omega)}$$

ergibt sich daher in diesem Fall

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Wir definieren nun auch für beliebige Wahrscheinlichkeitsräume (Ω, P) und für beliebige Ereignisse B mit $P(B) > 0$ die **BEDINGTE WAHRSCHEINLICHKEIT** $P(A|B)$ von A bei gegebenem B durch diese Formel.

BEISPIEL XI.2.1. Aus einer Urne, die zwei schwarze und drei weiße Kugeln enthält, werden nacheinander zwei Kugeln ohne Zurücklegen gezogen. Wir interessieren uns für die bedingte Wahrscheinlichkeit des Ereignisses A , dass die zweite gezogene Kugel schwarz ist, bei gegebenem Ereignis B , dass die erste gezogene Kugel weiß ist. Da nach dem ersten Ziehen drei von vier Kugeln schwarz sind, sollte diese Wahrscheinlichkeit $\frac{3}{4}$ betragen. Um die Richtigkeit unserer Definition nachzuprüfen, nummerieren wir die weißen Kugeln mit 1 und 2 und die schwarzen Kugeln mit 3, 4 und 5. Dann haben die interessierenden Ereignisse die Form

$$A \cap B = \{(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5)\},$$

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

Es ergibt sich

$$P(A|B) = \frac{\text{card}(A \cap B)}{\text{card}(B)} = \frac{6}{8} = \frac{3}{4}.$$

XI.2.2. Eigenschaften. In der Praxis wird häufig nicht $P(A|B)$ aus $P(A \cap B)$ und $P(B)$ berechnet, sondern umgekehrt $P(A \cap B)$ aus $P(A|B)$ und $P(B)$. Aus der Definition von $P(A|B)$ ergibt sich nämlich die **PRODUKTFORMEL FÜR WAHRSCHEINLICHKEITEN**

$$P(A \cap B) = P(B)P(A|B).$$

BEISPIEL XI.2.2. In Beispiel [XI.2.1](#) können wir zur Berechnung von $P(A \cap B)$ wie folgt argumentieren: Da zu Beginn 2 von 5 Kugeln weiß sind, ist $P(B) = \frac{2}{5}$. Da nach Eintreten von B 3 von 4 Kugeln schwarz sind, ist $P(A|B) = \frac{3}{4}$. Also muss

$$P(A \cap B) = \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{10}$$

sein.

Die obige Produktformel lässt sich wie folgt verallgemeinern: Sind A_1, \dots, A_k Ereignisse mit $P(A_1 \cap \dots \cap A_k) > 0$, so ist

$$P(A_1 \cap \dots \cap A_k) = P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_k|A_1 \cap \dots \cap A_{k-1}).$$

BEISPIEL XI.2.3. Wir wollen die Wahrscheinlichkeit bestimmen, dass beim Skat jeder der drei Spieler genau ein Ass hat. Bezeichne dazu mit A_i das Ereignis, dass Spieler i genau ein Ass erhält. Aus Symmetriegründen können wir annehmen, dass Spieler 1 die ersten 10 ausgeteilten Karten erhält, Spieler 2 die nächsten 10, dann Spieler 3 zehn, und die letzten 2 Karten in den Stock kommen. Dann ist die gesuchte Wahrscheinlichkeit

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2).$$

Es ist

$$P(A_1) = \frac{\binom{4}{1} \binom{28}{9}}{\binom{32}{10}}.$$

Da nach Austeilen der Karten an Spieler 1 noch 22 Karten incl. 3 Assen zu verteilen sind, ist

$$P(A_2|A_1) = \frac{\binom{3}{1} \binom{19}{9}}{\binom{22}{10}}.$$

Analog ist

$$P(A_3|A_1 \cap A_2) = \frac{\binom{2}{1} \binom{10}{9}}{\binom{12}{10}}.$$

Damit ergibt sich

$$P(A_1 \cap A_2 \cap A_3) = \frac{50}{899} \\ \approx 5.56\%.$$

Folgende Eigenschaften bedingter Wahrscheinlichkeiten sind sehr nützlich:

- (1) Sei $P(B) > 0$. Durch $P_B(A) = P(A|B)$ wird ein Wahrscheinlichkeitsmaß auf Ω definiert. Ist $A \subset \Omega \setminus B$ oder $P(A) = 0$, so ist $P(A|B) = 0$.
- (2) (FORMEL VON DER TOTALEN WAHRSCHEINLICHKEIT) Sind B_1, B_2, \dots paarweise disjunkte Teilmengen von Ω , deren Vereinigung Ω ist, so gilt für jedes Ereignis A

$$P(A) = \sum_k P(B_k)P(A|B_k).$$

- (3) (FORMEL VON BAYES) Ist $P(A) > 0$ und gelten die Voraussetzungen von (2), so gilt für jedes i

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_k P(B_k)P(A|B_k)}.$$

- (4) Ist C die Vereinigung der paarweise disjunkten Ereignisse C_1, C_2, \dots mit $P(C_i) > 0$ für alle i und sind die $P(A|C_i)$ alle gleich, so ist $P(A|C) = P(A|C_1)$.

BEISPIEL XI.2.4. Wir greifen Beispiel XI.2.1 (S. 417) auf. Wie groß ist die Wahrscheinlichkeit für das Ereignis C , dass beide ohne Zurücklegen gezogenen Kugeln die gleiche Farbe haben?

Sei B_1 das Ereignis, dass die erste gezogene Kugel weiß ist, und B_2 das Ereignis, dass sie schwarz ist. Wenn die erste gezogene Kugel weiß ist, ist beim zweiten Zug nur eine von vier Kugeln weiß. Damit ist $P(C|B_1) = \frac{1}{4}$. Analog ergibt sich $P(C|B_2) = \frac{1}{2}$. Wegen $P(B_1) = \frac{2}{5}$ und $P(B_2) = \frac{3}{5}$ ergibt sich mit der Formel von der totalen Wahrscheinlichkeit

$$P(C) = P(B_1)P(C|B_1) + P(B_2)P(C|B_2) \\ = \frac{2}{5} \cdot \frac{1}{4} + \frac{3}{5} \cdot \frac{1}{2} \\ = \frac{4}{10} \\ = \frac{2}{5}.$$

BEISPIEL XI.2.5. Eine Krankheit komme bei ca. 0.5% der Bevölkerung vor. Ein Test zur Erkennung der Krankheit führt bei 99% der Kranken zu einer Reaktion, aber auch bei 2% der Gesunden. Wie groß ist die Wahrscheinlichkeit, dass eine Person, bei der die Reaktion auftritt, tatsächlich krank ist?

Zur Lösung denken wir uns die Bevölkerung von 1 bis N nummeriert. B_1 sei die Menge der Kranken und B_2 diejenige der Gesunden. Also ist $\text{card}(B_1) \approx 0.005N$ und $\text{card}(B_2) \approx 0.995N$. A bezeichne die Menge der Personen, bei denen der Test zur Reaktion führt. Dann ist $\text{card}(A \cap B_1) \approx 0.99 \text{card}(B_1)$ und $\text{card}(A \cap B_2) \approx 0.02 \text{card}(B_2)$. Gesucht ist $P(B_1|A)$.

Wir wenden die Formel von Bayes an und ordnen jeder Person bei zufälliger Auswahl die Wahrscheinlichkeit $\frac{1}{N}$ zu. Dann ist

$$\begin{aligned} P(B_1) &= 0.005, \\ P(B_2) &= 0.995, \\ P(A \cap B_1) &= 0.99 \cdot 0.005, \\ P(A \cap B_2) &= 0.02 \cdot 0.995. \end{aligned}$$

Damit ergibt sich mit der Formel von Bayes

$$\begin{aligned} P(B_1|A) &= \frac{P(A \cap B_1)}{P(A)} \\ &= \frac{0.99 \cdot 0.005}{0.99 \cdot 0.005 + 0.02 \cdot 0.995} \\ &= \frac{495}{2485} \\ &\approx 0.2. \end{aligned}$$

Von allen Personen, bei denen der Test eine Reaktion zeigt, sind also nur 20% tatsächlich erkrankt.

XI.2.3. Unabhängigkeit. Zwei Ereignisse A und B heißen UNABHÄNGIG, wenn $P(A \cap B) = P(A)P(B)$ ist. Ist $P(B) > 0$, ist dies äquivalent zu $P(A) = P(A|B)$.

Unabhängigkeit von A und B drückt aus, dass A und B wahrheitstheoretisch in dem Sinn keinerlei Einfluss aufeinander haben, dass die Information „ B tritt ein“ keinen Einfluss auf die Wahrscheinlichkeit von A hat. Dies muss man deutlich von der realen Beeinflussung unterscheiden.

BEISPIEL XI.2.6. Wir werfen einen Würfel zweimal. Sei A das Ereignis, dass die Summe der beiden Augenzahlen gerade ist, und B das Ereignis, dass die zweite Augenzahl gerade ist. Offensichtlich ist

$$P(A) = P(B) = \frac{1}{2}$$

und

$$P(A \cap B) = \frac{1}{4} = P(A)P(B).$$

Also sind die beiden Ereignisse unabhängig, obwohl B mitbestimmt, ob A eintritt.

Wir müssen den Begriff der Unabhängigkeit auf Familien von Ereignissen übertragen. Dazu sagen wir, dass für die endliche Familie $\{A_i : i \in J\}$ von Ereignissen die **PRODUKTFORMEL** gilt, wenn

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$$

ist. Eine beliebige Familie $\{A_i : i \in J\}$ von Ereignissen heißt **UNABHÄNGIG**, wenn für jede endliche Teilfamilie die Produktformel gilt.

Damit ergeben sich folgende Eigenschaften:

- (1) Jede Teilfamilie einer unabhängigen Familie von Ereignissen ist unabhängig. Eine Familie ist genau dann unabhängig, wenn jede endliche Teilfamilie unabhängig ist.
- (2) Ist $\{A_i : i \in J\}$ eine Familie unabhängiger Ereignisse, k ein nicht zu J gehörender Index und $P(A_k) = 0$ oder $P(A_k) = 1$, so ist auch $\{A_i : i \in J \cup \{k\}\}$ unabhängig.
- (3) Ist $\{A_i : i \in J\}$ unabhängig und B_i für jedes i eines der Ereignisse A_i , $\Omega \setminus A_i$, \emptyset oder Ω , so ist $\{B_i : i \in J\}$ unabhängig.
- (4) Ist $J = \{1, \dots, n\}$ endlich, so ist $\{A_i : i \in J\}$ genau dann unabhängig, wenn für jede Wahl von $B_i \in \{A_i, \Omega \setminus A_i\}$ die Produktformel für B_1, \dots, B_n gilt.

XI.2.4. Produktexperimente. Wir nehmen an, dass wir schon Modelle $(\Omega_1, P_1), \dots, (\Omega_n, P_n)$ für gewisse Zufallsexperimente kennen, und wollen nun ein Modell für das Experiment konstruieren, das in der unabhängigen Hintereinanderausführung aller dieser Telexperimente besteht. Dazu liegt es nahe, als Ereignisraum das kartesische Produkt

$$\Omega_1 \times \dots \times \Omega_n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \Omega_i, 1 \leq i \leq n\}$$

zu wählen und als Wahrscheinlichkeitsfunktion

$$P(\omega) = P_1(\omega_1) \cdot \dots \cdot P_n(\omega_n).$$

Man bezeichnet (Ω, P) als **PRODUKT DER WAHRSCHEINLICHKEITSRÄUME** (Ω_i, P_i) und schreibt

$$\Omega = \Omega_1 \times \dots \times \Omega_n,$$

$$P = P_1 \times \dots \times P_n.$$

Ist A_i ein Ereignis im i -ten Experiment, so bezeichnet das kartesische Produkt $A_1 \times \dots \times A_n$ das Ereignis in Ω , dass für alle i im i -ten Teilexperiment A_i eintritt. Die Wahrscheinlichkeit für dieses Ereignis ist

$$P(A_1 \times \dots \times A_n) = P(A_1) \cdot \dots \cdot P(A_n).$$

XI.2.5. Binomialverteilung. Wir betrachten ein Experiment, das in der unabhängigen n -fachen Wiederholung eines Einzelexperimentes mit nur zwei verschiedenen möglichen Ausgängen besteht. Wir bezeichnen diese beiden Ausgänge mit 0 und 1. Dann ist $\Omega_i = \{0, 1\}$ und $\Omega = \{0, 1\}^n$. Da die Teilexperimente Wiederholungen des gleichen Experimentes sind, sollen in allen Teilexperimenten die gleichen Wahrscheinlichkeiten auftreten. $p = P_i(1)$ soll also nicht von i abhängen. p heißt oft **ERFOLGSWAHRSCHEINLICHKEIT**. Natürlich ist dann $P_i(0) = 1 - p$. Dann ist im Produktmodell $P(\omega) = p^k(1 - p)^{n-k}$, wenn k die Anzahl der Einsen in ω ist. Ein Experiment dieser Form nennt man **BERNOULLI-EXPERIMENT** und P heißt **BERNOULLI-VERTEILUNG**.

Das Ereignis, dass insgesamt k Einsen auftreten wird durch

$$E_k = \{\omega \in \Omega : \sum_{i=1}^n \omega_i = k\}$$

beschrieben. Die Zahl der Elemente von E_k ist gleich der Zahl der Möglichkeiten, die k Zeitpunkte in $\{1, \dots, n\}$ festzulegen, an denen die Einsen auftreten, also $\binom{n}{k}$. Es folgt

$$P(E_k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Die Terme

$$b_{n,p}(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{mit } 0 \leq k \leq n$$

bestimmen die Wahrscheinlichkeitsverteilung. Sie sind nicht negativ und haben die Summe 1. Man nennt sie **BINOMIALVERTEILUNG** mit Parametern n und p oder kurz $b_{n,p}$ -Verteilung.

BEISPIEL XI.2.7. Wie groß ist die Wahrscheinlichkeit, bei 10 Würfeln eines Würfels 3 Sechsen zu erhalten?

Wir können die geworfenen Sechsen als Erfolge auffassen und ihnen den Wert 1 zuordnen; die anderen Ergebnisse sind Misserfolge und erhalten den Wert 0. Dann ist $p = \frac{1}{6}$. Die gesuchte Wahrscheinlichkeit ist

$$\binom{10}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^7 \approx 15.5\%.$$

BEISPIEL XI.2.8. Die Wahrscheinlichkeit für die Geburt eines Jungen sei $p = 0.51$. Aufeinanderfolgende Geburten seien unabhängig. Dann ist die Wahrscheinlichkeit, dass in einer Familie mit 4 Kindern 2 Jungen und 2 Mädchen vorkommen, gleich

$$\binom{4}{2} 0.51^2 0.49^2 \approx 37.47\%.$$

XI.2.6. Multinomialverteilung. Hier hat man wieder n unabhängige, identische Teilversuche, aber jeder Teilversuch hat nun r mögliche Ausgänge. Die Teilerperimente sind beschreibbar durch $P_i(j) = p_j$, $j = 1, \dots, r$, wobei die Wahrscheinlichkeiten p_1, \dots, p_r beliebig vorgegeben sind mit $p_j \geq 0$, $j = 1, \dots, r$, und $p_1 + \dots + p_r = 1$. Die Wahrscheinlichkeit, in den n Teilversuchen k_1 -mal das Ergebnis 1, k_2 -mal das Ergebnis 2 usw. zu erreichen, ist

$$\binom{n}{k_1, \dots, k_r} p_1^{k_1} \cdot \dots \cdot p_r^{k_r}.$$

XI.2.7. Geometrische Verteilung. Die Wahrscheinlichkeit, dass in einem Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p der erste Erfolg genau im k -ten Versuch eintritt, ist $p(1-p)^{k-1}$, $k \geq 1$. Man k als Ergebnis eines Experimentes auffassen, das darin besteht, zu beobachten, wann in einer Folge von Bernoulli-Experimenten der erste Erfolg eintritt. Die zugehörige Ereignismenge ist dann $\Omega = \mathbb{N} \setminus \{0\}$. Die zugehörige Wahrscheinlichkeitsfunktion ist $P(k) = p(1-p)^{k-1}$. Die hierdurch definierte Verteilung heißt GEOMETRISCHE VERTEILUNG. Man beachte, dass im Gegensatz zu allen bisherigen Beispielen der Ereignisraum jetzt nicht mehr endlich ist.

XI.2.8. Negative Binomialverteilung. Dies ist eine Verallgemeinerung der geometrischen Verteilung. Sei $f(k; r, p)$ die Wahrscheinlichkeit, dass bei $n \geq r + k$ Bernoulli-Experimenten mit Erfolgswahrscheinlichkeit p genau k Misserfolge dem r -ten Erfolg vorangehen. Man kann nachrechnen, dass

$$f(k; r, p) = \binom{k+r-1}{k} p^r (1-p)^k$$

ist. Die hierdurch bei festem r auf $\Omega = \mathbb{N}$ definierte Verteilung nennt man NEGATIVE BINOMIALVERTEILUNG oder PASCAL-VERTEILUNG.

XI.3. Zufallsvariable, Erwartungswert, Varianz

XI.3.1. Zufallsvariable. Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum und \mathcal{X} eine beliebige Menge. Eine Abbildung $X : \Omega \rightarrow \mathcal{X}$ heißt (\mathcal{X} -wertige) ZUFALLSVARIABLE.

BEMERKUNG XI.3.1. Da Ω endlich ist, ist die Bildmenge $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ einer Zufallsvariablen X auf Ω stets endlich.

BEISPIEL XI.3.2. Vor einer Wahl werden zufällig ausgewählte Bürger gefragt, welche Partei sie wählen wollen. Hier ist Ω die Menge der Bürger, P die Gleichverteilung auf Ω , \mathcal{X} die Menge der wählbaren Parteien vereinigt mit der „Partei“ der Nichtwähler und $X(\omega)$ die Partei, die Bürger ω angeblich wählen will.

BEISPIEL XI.3.3. Wir interessieren uns für die Summe der Augenzahlen beim Würfeln mit zwei Würfeln. Hier ist $\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$, P die Gleichverteilung auf Ω , $\mathcal{X} = \{2, \dots, 12\}$ und $X((i, j)) = i + j$.

An Zufallsvariablen interessiert uns vor allem ihre Verteilungsfunktion. Diese gibt an, wie wahrscheinlich die einzelnen Ergebnisse von X sind. Genauer versteht man unter der VERTEILUNGSFUNKTION von $X : \Omega \rightarrow \mathcal{X}$ das Wahrscheinlichkeitsmaß P_X auf $\mathcal{X}_X = X(\Omega)$, das definiert ist durch

$$P_X(x) = P(\{\omega \in \Omega : X(\omega) = x\}) \quad \text{mit } x \in \mathcal{X}_X.$$

$P_X(x)$ ist also die Wahrscheinlichkeit, dass X den Wert x annimmt.

Die Verteilungsfunktion einer Zufallsvariablen X kann durch ein STABDIAGRAMM veranschaulicht werden. Dazu erstellt man zunächst eine Liste der möglichen Werte $x_k \in \mathcal{X}$ von X , berechnet für jedes x_k die Wahrscheinlichkeit $P_X(x_k)$ und zeichnet dann senkrecht über den Punkten x_k der x -Achse Striche der Länge $P_X(x_k)$.

BEISPIEL XI.3.4. Die Verteilungsfunktion der Zufallsvariablen aus Beispiel XI.3.3 ist durch Tabelle XI.3.1 bestimmt. Das zugehörige Stabdiagramm ist in Abbildung XI.3.1 wiedergegeben.

TABELLE XI.3.1. Verteilungsfunktion der Zufallsvariablen aus Beispiel XI.3.3

x_k	2	3	4	5	6	7	8	9	10	11	12
$P_X(x_k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

XI.3.2. Gemeinsame Verteilung mehrerer Zufallsvariabler.

Sind auf einem Wahrscheinlichkeitsraum (Ω, P) mehrere Zufallsvariable X_1, \dots, X_n mit evtl. verschiedenen Wertebereichen $\mathcal{X}_1, \dots, \mathcal{X}_n$ definiert, kann man sie zu einer einzigen Zufallsvariablen X mit Wertebereich

$$\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$$

zusammenfassen, indem man

$$X(\omega) = (X_1(\omega), \dots, X_n(\omega))$$

setzt. Die Verteilungsfunktion von X nennt man die GEMEINSAME VERTEILUNGSFUNKTION von X_1, \dots, X_n .

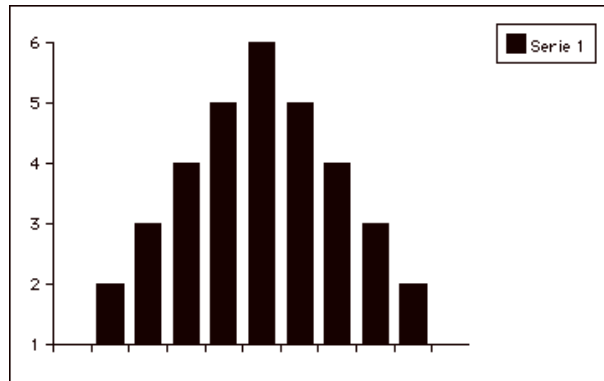


ABBILDUNG XI.3.1. Stabdiagramm der Zufallsvariablen aus Beispiel XI.3.3

BEISPIEL XI.3.5. Wir betrachten die Bernoulli-Verteilung P zu $p \in (0, 1)$ auf $\Omega = \{0, 1\}^n$ (vgl. Abschnitt XI.2.5 (S. 422)). Für $\omega \in \Omega$ sei

$$S(\omega) = \omega_1 + \dots + \omega_n$$

die Zahl der Erfolge. Für ω mit $S(\omega) \geq 1$ setzen wir

$$N(\omega) = \min\{j \geq 1 : \omega_j = 1\}.$$

Dies ist die Wartezeit bis zum ersten Erfolg. Ist $S(\omega) = 0$, setzen wir $N(\omega) = n + 1$. S und N sind zwei Zufallsvariable auf (Ω, P) . Ihre gemeinsame Verteilung wird beschrieben durch die Angabe der Wahrscheinlichkeiten

$$p(k, h) = P(S = k, N = h)$$

mit $0 \leq k \leq n$ und $1 \leq h \leq n + 1$. Wir wollen diese Verteilung bestimmen.

Offenbar ist

$$p(0, n + 1) = (1 - p)^n$$

und

$$p(0, h) = 0$$

für $1 \leq h \leq n$. Ist $S(\omega) = k \geq 1$ und $N(\omega) = h$, so muss gelten $\omega_i = 0$ für $i < h$ und $\omega_h = 1$, und es müssen genau $k - 1$ Einsen unter $\omega_{h+1}, \dots, \omega_n$ vorkommen. Es gibt $\binom{n-h}{k-1}$ solche Elemente, und jedes hat die Wahrscheinlichkeit $p^k(1-p)^{n-k}$. Also ist

$$p(k, h) = \binom{n-h}{k-1} p^k (1-p)^{n-k}$$

für $k \geq 1$. Ist $k - 1 > n - h$, so gibt es kein solches ω , und es ist $\binom{n-h}{k-1} = 0$. Die obige Formel bleibt also richtig.

XI.3.3. Unabhängigkeit. Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum. Eine Familie $X_i, i \in I$, von Zufallsvariablen $X_i : \Omega \rightarrow \mathcal{X}_i$ heißt UNABHÄNGIG, wenn für jede Wahl von $A_i \subset \mathcal{X}_i, i \in I$, die Ereignisse $\{X_i \in A_i\}, i \in I$, unabhängig sind.

Sei nun X_1, \dots, X_n eine unabhängige Familie von Zufallsvariablen. Dann gilt gemäß Abschnitt XI.2.3 (S. 420) speziell für alle $(x_1, \dots, x_n) \in \mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i).$$

Die gemeinsame Verteilung von X_1, \dots, X_n ist also das Produkt der Verteilungen der X_i .

Die Unabhängigkeit von Zufallsvariablen bleibt unter der Komposition von Abbildungen erhalten: Sind $X_i, i \in I$, unabhängige Zufallsvariable und $F_i : \mathcal{X}_i \rightarrow \mathcal{Y}_i$ beliebige Abbildungen mit beliebigen Wertebereichen, so sind die Zufallsvariablen $F_i \circ X_i, i \in I$, auch unabhängig.

XI.3.4. Erwartungswert. Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum und X eine reellwertige Zufallsvariable auf Ω . Dann ist der ERWARTUNGSWERT von X definiert durch

$$EX = E(X) = \sum_{\omega \in \Omega} X(\omega)P(\omega).$$

BEISPIEL XI.3.6. Für die Zufallsvariable „Summe der Augenzahlen beim Werfen von zwei Würfeln“ aus Beispiel XI.3.4 erhalten wir den Erwartungswert

$$\begin{aligned} EX &= \frac{1}{36} \{1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + 4 \cdot 5 + 5 \cdot 6 + 6 \cdot 7 \\ &\quad + 5 \cdot 8 + 4 \cdot 9 + 3 \cdot 10 + 2 \cdot 11 + 1 \cdot 12\} \\ &= \frac{252}{36} \\ &= 7. \end{aligned}$$

Sind X, Y reellwertige Zufallsvariable auf Ω und ist $\lambda \in \mathbb{R}$, so gilt

$$\begin{aligned} E(\lambda X) &= \lambda E(X) \\ E(X + Y) &= E(X) + E(Y) \\ X, Y \text{ unabhängig} &\implies E(XY) = E(X)E(Y). \end{aligned}$$

Die ersten beiden Eigenschaften folgen direkt aus der Definition des Erwartungswertes. Zur Nachweis der dritten Eigenschaft bezeichnen wir mit x_i , $i \in I$, und y_j , $j \in J$, die (endlich vielen!) möglichen Werte von X bzw. Y . Dann folgt mit dem vorigen Abschnitt

$$\begin{aligned}
 E(XY) &= \sum_{\omega \in \Omega} X(\omega)Y(\omega)P(\omega) \\
 &= \sum_{i \in I} \sum_{j \in J} x_i y_j \underbrace{P(X = x_i, Y = y_j)}_{P(X=x_i)P(Y=y_j)} \\
 &= \sum_{i \in I} \sum_{j \in J} x_i y_j P(X = x_i)P(Y = y_j) \\
 &= \left\{ \sum_{i \in I} x_i P(X = x_i) \right\} \left\{ \sum_{j \in J} y_j P(Y = y_j) \right\} \\
 &= E(X)E(Y).
 \end{aligned}$$

BEISPIEL XI.3.7 (ERWARTUNGSWERT DER BINOMIALVERTEILUNG). Für die Zufallsvariable S aus Beispiel XI.3.5 erhalten wir

$$\begin{aligned}
 E(S) &= \sum_{k=0}^n k \underbrace{P(S = k)}_{= \binom{n}{k} p^k (1-p)^{n-k}} \\
 &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n k \underbrace{\binom{n}{k}}_{= n \binom{n-1}{k-1}} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n n \binom{n-1}{k-1} p p^{k-1} (1-p)^{(n-1)-(k-1)} \\
 &= np \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell} \\
 &= np.
 \end{aligned}$$

Das gleiche Ergebnis erhalten wir einfacher mit folgendem „Trick“, der auch in anderen Situationen hilfreich ist: Definiere die Zufallsvariablen S_1, \dots, S_n durch die Vorschrift $S_i(\omega) = 1$, falls das i -te Telexperiment ein Erfolg ist, $S_i(\omega) = 0$ sonst. Dann ist $S = S_1 + \dots + S_n$. Für jedes i ist aber offensichtlich $E(S_i) = p$. Daher ist

$$E(S) = \sum_{i=1}^n E(S_i) = np.$$

BEISPIEL XI.3.8 (ERWARTUNGSWERT DER HYPERGEOMETRISCHEN VERTEILUNG). In einer Urne befinden sich S schwarze und W weiße Kugeln, insgesamt $N = S + W$ Kugeln. Wir ziehen daraus $n \leq N$ Kugeln ohne Zurücklegen (vgl. Abschnitt XI.1.4 (S. 414)). Die Zufallsvariable X sei die Zahl der dabei gezogenen schwarzen Kugeln. Um $E(X)$ zu berechnen, bezeichnen wir mit X_i die Zufallsvariable, die den Wert 1 liefert, wenn im i -ten Zug eine schwarze Kugel gezogen wird, und die sonst den Wert 0 gibt. Dann ist $X = X_1 + \dots + X_n$, und für jedes i gilt $E(X_i) = \frac{S}{N}$. Damit folgt

$$E(X) = n \frac{S}{N}.$$

XI.3.5. Varianz und Kovarianz. Seien X und Y zwei reellwertige Zufallsvariable auf einem endlichen Wahrscheinlichkeitsraum (Ω, P) . Die VARIANZ von X ist definiert durch

$$\text{Var}(X) = E((X - E(X))^2).$$

Die Größe

$$\sigma_X = \sqrt{\text{Var}(X)}$$

heißt die STREUUNG oder STANDARDABWEICHUNG von X .

Die KOVARIANZ von X und Y ist definiert durch

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))).$$

Die Größe

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

heißt der KORRELATIONSKOEFFIZIENT von X und Y . Die Zufallsvariablen X und Y heißen UNKORRELIERT, wenn $\text{Cov}(X, Y) = 0$ ist.

Die Varianz ist ein Maß für die Streuung einer Zufallsvariablen um den Erwartungswert EX , der ein Maß für den Mittelwert ist. Ist $EX = 0$ und haben die Werte x_1, \dots, x_n von X alle die gleiche Wahrscheinlichkeit $\frac{1}{n}$, so ist

$$\sigma_X = \sqrt{\frac{1}{n} \sum x_i^2}.$$

Bis auf den Skalierungsfaktor $\frac{1}{\sqrt{n}}$ ist also σ_X der Euklidische Abstand des Punktes (x_1, \dots, x_n) vom Ursprung.

Eine positive Kovarianz von X und Y bedeutet, dass X die Tendenz hat, dort groß zu sein, wo auch Y groß ist.

Sind X, Y, X_i reellwertige Zufallsvariable und a, b, c, d reelle Zahlen, so gelten folgende Eigenschaften:

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ \text{Var}(aX + b) &= a^2 \text{Var}(X) \\ \text{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\ \text{Cov}(aX + b, cY + d) &= ac \text{Cov}(X, Y) \\ \text{Cov}(X, Y) &= \text{Cov}(Y, X) \\ |\text{Cov}(X, Y)| &\leq \sigma_X \sigma_Y \\ \text{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ X, Y \text{ unabhängig} &\implies \text{Cov}(X, Y) = 0 \\ X_1, \dots, X_n \text{ unabhängig} &\implies \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) \end{aligned}$$

Zum Nachweis dieser Eigenschaften beginnen wir mit der dritten. Da der Erwartungswert einer konstanten Zufallsvariablen diese Konstante ist, erhalten wir

$$\begin{aligned} \text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY - XE(Y) - YE(X) + E(X)E(Y)) \\ &= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Die erste Eigenschaft ist der Spezialfall $X = Y$ der dritten Eigenschaft. Aus der dritten Eigenschaft folgt $\text{Cov}(X, Y) = 0$, wenn eine der Zufallsvariablen X oder Y konstant ist. Aus der Definition der Kovarianz folgt zudem, dass die Abbildung $X, Y \mapsto \text{Cov}(X, Y)$ bilinear ist. Damit folgt

$$\begin{aligned} \text{Cov}(aX + b, cY + d) &= \text{Cov}(aX, cY) + \text{Cov}(aX, d) \\ &\quad + \text{Cov}(b, cY) + \text{Cov}(b, d) \\ &= ac \text{Cov}(X, Y). \end{aligned}$$

Die zweite Eigenschaft ist der Spezialfall $X = Y, a = c$ der soeben bewiesenen vierten Eigenschaft.

Die fünfte Eigenschaft folgt direkt aus der Definition.

Die sechste Eigenschaft ist eine Konsequenz der Cauchy-Schwarzschen

Ungleichung (vgl. Abschnitt [II.6.6](#) (S. 107)).

Zum Nachweis der siebten Eigenschaft setze $\bar{X}_i = X_i - E(X_i)$. Dann folgt

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n X_i\right) &= E\left(\left(\sum_{i=1}^n \bar{X}_i\right)^2\right) \\ &= E\left(\sum_{i=1}^n \bar{X}_i^2 + \sum_{i \neq j} \bar{X}_i \bar{X}_j\right) \\ &= \sum_{i=1}^n E(\bar{X}_i^2) + \sum_{i \neq j} E(\bar{X}_i \bar{X}_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j). \end{aligned}$$

Sind X und Y unabhängig, so auch $X - E(X)$ und $Y - E(Y)$. Damit folgt die achte Eigenschaft:

$$\begin{aligned} \text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(X - E(X))E(Y - E(Y)) \quad (\text{wg. Unabhängigkeit}) \\ &= 0. \end{aligned}$$

Die letzte Eigenschaft schließlich folgt aus der siebten und achten Identität.

BEISPIEL XI.3.9 (VARIANZ DER BINOMIALVERTEILUNG). Wir betrachten die Zufallsvariable S aus [Beispiel XI.3.5](#) (S. 425) und definieren die Zufallsvariablen S_i wie in [Beispiel XI.3.7](#) (S. 427). Da sich S_i nur auf das i -te Telexperiment bezieht und die Telexperimente unabhängig sind, sind die S_i unabhängig. Daher ist

$$\begin{aligned} \text{Var}(S) &= \text{Var}(S_1 + \dots + S_n) \\ &= \text{Var}(S_1) + \dots + \text{Var}(S_n) \\ &= n \text{Var}(S_1), \end{aligned}$$

da jedes S_i die gleiche Varianz hat. Weiter ist

$$\begin{aligned} \text{Var}(S_1) &= E(S_1^2) - E(S_1)^2 \\ &= p - p^2 \\ &= p(1 - p). \end{aligned}$$

Insgesamt folgt

$$\text{Var}(S) = np(1 - p).$$

BEISPIEL XI.3.10 (VARIANZ DER HYPERGEOMETRISCHEN VERTEILUNG). Wir betrachten die Zufallsvariablen X und X_1, \dots, X_n aus Beispiel XI.3.8 (S. 428). Setze $p = \frac{S}{N}$. Alle X_i haben die gleiche Varianz:

$$\begin{aligned}\operatorname{Var}(X_i) &= \operatorname{Var}(X_1) \\ &= E(X_1^2) - E(X_1)^2 \\ &= p - p^2 \\ &= p(1 - p).\end{aligned}$$

Ebenso gilt aus Symmetriegründen für alle $i \neq j$

$$\begin{aligned}\operatorname{Cov}(X_i, X_j) &= \operatorname{Cov}(X_1, X_2) \\ &= E(X_1 X_2) - E(X_1)E(X_2).\end{aligned}$$

Weiter ist

$$\begin{aligned}E(X_1 X_2) &= P(\{X_1 = 1\} \cap \{X_2 = 1\}) \\ &= P(\{X_1 = 1\})P(\{X_2 = 1\} | \{X_1 = 1\}) \\ &= \frac{S}{N} \frac{S-1}{N-1} \\ &= p \frac{S-1}{N-1}.\end{aligned}$$

Damit folgt

$$\begin{aligned}\operatorname{Cov}(X_i, X_j) &= p \frac{S-1}{N-1} - p^2 \\ &= p \left(\frac{S-1}{N-1} - p \right)\end{aligned}$$

für $i \neq j$ und

$$\begin{aligned}\operatorname{Var}(X) &= \sum_{i=1}^n \operatorname{Var}(X_i) + \sum_{i \neq j} \operatorname{Cov}(X_i, X_j) \\ &= np(1-p) + (n^2 - n)p \underbrace{\left(\frac{S-1}{N-1} - p \right)}_{= \frac{S-1}{N-1} - \frac{S}{N} = \frac{S-N}{N(N-1)}} \\ &= np(1-p) + n(n-1)p \underbrace{\frac{S-N}{N(N-1)}}_{= \frac{1}{N-1}(p-1)} \\ &= np(1-p) \left(1 - \frac{n-1}{N-1} \right) \\ &= np(1-p) \frac{N-n}{N-1}.\end{aligned}$$

Tabelle XI.3.2 fasst die Ergebnisse der vorigen Beispiele zusammen.

TABELLE XI.3.2. Erwartungswert und Varianz einiger Verteilungen

Verteilung	Erwartungswert	Varianz
binomial	np	$np(1-p)$
hypergeometrisch $h(s; n, N, S), p = \frac{S}{N}$	np	$np(1-p)\frac{N-n}{N-1}$

XI.3.6. Das schwache Gesetz der großen Zahlen. Seien X_1, \dots, X_n unabhängige, reellwertige Zufallsvariable mit gleichem Erwartungswert und $\text{Var}(X_i) \leq M$ für alle i . Dann gilt für alle $\varepsilon > 0$:

SCHWACHES GESETZ DER GROSSEN ZAHLEN:

$$P\left(\left|\frac{1}{n}(X_1 + \dots + X_n) - E(X_1)\right| \geq \varepsilon\right) \leq \frac{M}{\varepsilon^2 n} \xrightarrow{n \rightarrow \infty} 0.$$

Es besagt, dass bei immer häufigerer, unabhängiger Wiederholung eines Experimentes das Ergebnis im Mittel fast sicher gleich dem Erwartungswert ist.

BEISPIEL XI.3.11. Für die Zufallsvariablen S und S_1, \dots, S_n des Bernoulli-Experimentes aus den Beispielen XI.3.5 (S. 425) und XI.3.9 haben wir für alle i

$$E(S_i) = p,$$

$$\text{Var}(S_i) = p(1-p) \leq \frac{1}{4}.$$

$h_n = \frac{1}{n}S$ ist die relative Häufigkeit, mit der ein Erfolg eintritt. Die Wahrscheinlichkeit, dass diese relative Häufigkeit um mehr als ε von dem Erwartungswert p abweicht, ist

$$P(|h_n - p| \geq \varepsilon) \leq \frac{1}{4\varepsilon^2 n}.$$

BEISPIEL XI.3.12. Wir betrachten ein für uns vorteilhaftes Spiel, bei dem wir auf Dauer fast sicher verlieren.

Die Spielregeln lauten:

- Der Anfangseinsatz beträgt 1 Euro.
- Es wird immer wieder eine Münze geworfen. Fällt Zahl geht die Hälfte des aktuellen Einsatzes verloren, fällt Kopf gewinnen wir $\frac{2}{3}$ des aktuellen Einsatzes hinzu.

Da der Gewinn höher ist als der Verlust, ist dieses Spiel vorteilhaft für uns. Dennoch werden wir auf Dauer fast sicher verlieren.

Um dies einzusehen, bezeichnen wir mit X_{n-1} den Einsatz vor dem n -ten Wurf, $n \geq 1$. Dann ist $X_0 = 1$ Euro und $X_n = \frac{1}{2}X_{n-1}$, wenn im n -ten Wurf Zahl kommt, und $X_n = \frac{5}{3}X_{n-1}$, wenn im n -ten Wurf Kopf fällt. Setze $Y_n = \frac{1}{2}$, falls im n -ten Wurf Zahl kommt, und $Y_n = \frac{5}{3}$,

falls im n -ten Wurf Kopf fällt. Dann ist $X_n = Y_1 \cdot \dots \cdot Y_n$. Die Y_i sind unabhängig. Es ist

$$\begin{aligned} E(Y_i) &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{5}{3} \\ &= \frac{13}{12} \\ &> 1. \end{aligned}$$

Daher ist

$$\begin{aligned} E(X_n) &= \prod_{i=1}^n E(Y_i) \\ &= \left(\frac{13}{12}\right)^n \\ &\xrightarrow{n \rightarrow \infty} \infty. \end{aligned}$$

Setze

$$\mu = E(\log Y_i).$$

Dann ist

$$\begin{aligned} \mu &= \frac{1}{2} \log\left(\frac{1}{2}\right) + \frac{1}{2} \log\left(\frac{5}{3}\right) \\ &< \frac{1}{2} \log\left(\frac{1}{2}\right) + \frac{1}{2} \log(2) \\ &= 0. \end{aligned}$$

Mit $\varepsilon = \frac{1}{2}|\mu| = -\frac{1}{2}\mu$ gilt nach dem schwachen Gesetz der großen Zahlen

$$P\left(\left|\frac{1}{n}(\log(Y_1) + \dots + \log(Y_n)) - \mu\right| \leq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 1.$$

Wegen

$$\log(X_n) = \log(Y_1) + \dots + \log(Y_n)$$

und

$$\left\{\left|\frac{1}{n} \log(X_n) - \mu\right| \leq -\frac{\mu}{2}\right\} \subset \left\{\frac{1}{n} \log(X_n) \leq \frac{\mu}{2}\right\}$$

folgt

$$P\left(\frac{1}{n} \log(X_n) \leq \frac{\mu}{2}\right) \xrightarrow{n \rightarrow \infty} 1.$$

Für große n gilt also mit einer Wahrscheinlichkeit nahe bei 1

$$X_n \leq e^{n\frac{\mu}{2}}.$$

Wegen $\mu < 0$ strebt dieser Ausdruck exponentiell schnell gegen Null. Dieses scheinbar paradoxe Verhalten liegt daran, dass man mit sehr kleiner Wahrscheinlichkeit sehr große Gewinne machen kann.

Wir können uns das Ergebnis heuristisch auch leicht wie folgt klar

machen: Nach $2n$ Spielrunden sollten wir etwa n -mal gewonnen und n -mal verloren haben. Dann beträgt unser Kapital

$$X_{2n} \approx \left(\frac{1}{2}\right)^n \left(\frac{5}{3}\right)^n = \left(\frac{5}{6}\right)^n.$$

Wegen $\frac{5}{6} < 1$ strebt dieser Ausdruck für $n \rightarrow \infty$ aber gegen Null.

XI.4. Grundbegriffe der Schätztheorie

XI.4.1. Motivation. Ein Teich enthält eine unbekannte Zahl N von Fischen, die geschätzt werden soll. Um dies zu erreichen, fangen wir W Fische, markieren sie mit einem weißen Fleck und setzen sie wieder aus. Wir warten eine Weile und fangen nun n Fische. Sei x die Zahl der markierten Fische in diesem Fang. Eine von dieser Zahl x abhängige Schätzung $\hat{N}(x)$ für N erhalten wir mit folgender Plausibilitätsbetrachtung: Wenn x nicht zu klein ist, sollte der Anteil $\frac{x}{n}$ der markierten Fische im zweiten Fang etwa dem Anteil $\frac{W}{N}$ der markierten Fische am Gesamtbestand entsprechen, d.h. $\frac{x}{n} \approx \frac{W}{N}$. Da wir x , n und W kennen, können wir diese Näherungsgleichung nach N auflösen und erhalten die Schätzung

$$N \approx \hat{N}(x) = \left\lceil \frac{Wn}{x} \right\rceil.$$

Dabei bezeichnet $\lceil a \rceil$ die größte ganze Zahl $\leq a$.

Wir können diese Plausibilitätsbetrachtung auf ein etwas sichereres mathematisches Fundament stellen. Dazu beschreiben wir den zweiten Fang durch das Modell des Ziehens ohne Rücklegen von n Kugeln aus einer Urne mit W weißen und $N - W$ schwarzen Kugeln. Die Wahrscheinlichkeit, dass x weiße Kugeln gezogen werden, ist dann

$$P_N(x) = \frac{\binom{W}{x} \binom{N-W}{n-x}}{\binom{N}{n}} \quad \text{mit } 0 \leq x \leq n.$$

Man beachte, dass die Abhängigkeit von N hier eine ganz andere ist als bisher: N ist nicht ein Ereignis, sondern ein unbekannter, zu bestimmender Parameter.

Der MAXIMUM-LIKELIHOOD ANSATZ zur Schätzung von N besagt, dass man bei gegebenem W und x den Wert von N so bestimmt, dass $P_N(x)$ maximal wird. Um N so zu bestimmen, betrachten wir

$$\begin{aligned} \frac{P_N(x)}{P_{N-1}(x)} &= \frac{\binom{W}{x} \binom{N-W}{n-x} \binom{N-1}{n}}{\binom{N}{n} \binom{W}{x} \binom{N-1-W}{n-x}} \\ &= \frac{(N-W)(N-n)}{N(N-W-n+x)}. \end{aligned}$$

Offensichtlich gilt

$$P_N(x) > P_{N-1}(x)$$

$$\begin{aligned} \iff (N - W)(N - n) &> N(N - W - n + x) \\ \iff Wn &> Nx. \end{aligned}$$

Analoge Äquivalenzen gelten auch für die Beziehungen $<$ und $=$. Daher liefert der Maximum-Likelihood Ansatz die gleiche Schätzung wie unser Plausibilitätsargument.

XI.4.2. Der allgemeine Rahmen von Schätzproblemen. Zur Beschreibung eines Schätzproblems mit endlichem Stichprobenraum benötigen wir:

- eine nicht leere, endliche Menge \mathcal{X} , den STICHPROBENRAUM,
- eine Familie $\{P_\vartheta : \vartheta \in \Theta\}$ von Wahrscheinlichkeitsmaßen auf \mathcal{X} und
- eine zu schätzende Funktion $g(\vartheta)$.

\mathcal{X} ist die Menge der möglichen Beobachtungsergebnisse. Durch die unterschiedliche Notation zu den vorigen Abschnitten – \mathcal{X} statt Ω – soll betont werden, dass jedes $x \in \mathcal{X}$ tatsächlich beobachtbar sein muss. Bei wahrscheinlichkeitstheoretischen Modellen treten dagegen auch Stichproben- bzw. Ereignisräume Ω auf, deren Elemente ω nicht beobachtbar sind.

BEISPIEL XI.4.1. Im Rahmen des vorigen Abschnittes ist $\mathcal{X} = \{1, 2, \dots, n\}$, wenn n die Zahl der Fische im zweiten Fang ist. Weiter ist $\vartheta = N$ die unbekannte Zahl der Fische im Teich. $P_\vartheta = P_N$ ist die hypergeometrische Verteilung $h(\cdot; n, N, W)$ und die Funktion g ist gegeben durch $g(\vartheta) = \vartheta = N$.

In obigem Beispiel ist die zu schätzende Funktion $g(\vartheta) = \vartheta$. Ist dagegen z.B. die Varianz einer Binomialverteilung zu schätzen, ist $\vartheta = p$ und $g(p) = np(1 - p)$.

Bezeichne mit \mathcal{Y} den Wertebereich der zu schätzenden Funktion g . Jede Abbildung $T : \mathcal{X} \rightarrow \mathcal{Y}$ heißt ein SCHÄTZER von g . Diese Sprechweise lässt natürlich auch völlig unsinnige Schätzer zu.

Häufig deutet man in der Notation an, dass geschätzt wird, und setzt ein „Dach“ über die zu schätzende Größe. \hat{N} wäre also ein Schätzer für N , \hat{p} ein Schätzer für p und \hat{g} ein Schätzer für g .

XI.4.3. Maximum-Likelihood Schätzer. Sei $x \in \mathcal{X}$ eine feste Beobachtung. Die Funktion

$$L_x : \Theta \rightarrow [0, 1]$$

mit

$$L_x(\vartheta) = P_\vartheta(x)$$

nennt man LIKELIHOOD-FUNKTION. Wenn L_x einen Maximalwert in $\hat{\vartheta} = \hat{\vartheta}(x)$ annimmt, d.h.

$$L_x(\hat{\vartheta}) = \sup\{L_x(\vartheta) : \vartheta \in \Theta\},$$

nennt man $\hat{\vartheta}(x)$ eine MAXIMUM-LIKELIHOOD SCHÄTZUNG von ϑ und $g(\hat{\vartheta})$ eine Maximum-Likelihood Schätzung von $g(\vartheta)$.

Häufig ist Θ ein Intervall in \mathbb{R} , und eine Maximum-Likelihood Schätzung kann wie in Abschnitt IV.2.1 (S. 140) durch Differentiation gefunden werden. Dabei ist es oft einfacher mit der Funktion $\mathcal{L}_x = \ln L_x$ zu arbeiten. Wegen der Monotonie des Logarithmus haben \mathcal{L}_x und L_x ihre Maxima an der gleichen Stelle.

BEISPIEL XI.4.2. In n Bernoulli-Experimenten soll die Erfolgswahrscheinlichkeit p aus der Zahl x der Erfolge geschätzt werden. Es ist

$$\begin{aligned} L_x(p) &= \binom{n}{x} p^x (1-p)^{n-x} \\ \mathcal{L}_x(p) &= \ln L_x(p) \\ &= \ln \binom{n}{x} + x \ln p + (n-x) \ln(1-p) \\ \frac{d}{dp} \mathcal{L}_x(p) &= \frac{x}{p} - \frac{n-x}{1-p}. \end{aligned}$$

Also liegt das eindeutige Extremum bei $\frac{x}{n}$. Wie man leicht nachprüft, ist dies auch ein Maximum. $\frac{x}{n}$ ist die Maximum-Likelihood Schätzung für die Erfolgswahrscheinlichkeit p .

XI.4.4. Erwartungstreue. Ist $T : \mathcal{X} \rightarrow \mathbb{R}$ ein Schätzer, so bezeichnen wir den Erwartungswert von T bzgl. P_ϑ mit $E_\vartheta(T)$:

$$E_\vartheta(T) = \sum_{x \in \mathcal{X}} T(x) P_\vartheta(x).$$

Die Beobachtung des Experimentes, das dem Schätzproblem zugrunde liegt, können wir als eine Zufallsvariable X auffassen, d.h. $X(x) = x$. Mit dieser Notation können wir den Begriff der Erwartungstreue definieren:

Ein Schätzer \hat{g} für $g(\vartheta)$ heißt ERWARTUNGSTREU, wenn für alle $\vartheta \in \Theta$ gilt

$$E_\vartheta(\hat{g}(X)) = g(\vartheta).$$

Speziell heißt $\hat{\vartheta}$ ein ERWARTUNGSTREUER SCHÄTZER von ϑ , wenn für alle $\vartheta \in \Theta$ gilt

$$E_\vartheta(\hat{\vartheta}(X)) = \vartheta.$$

Die Differenz

$$b(\vartheta, \hat{g}) = E_\vartheta(\hat{g}(X)) - g(\vartheta)$$

heißt der BIAS (wörtlich: Vorurteil) des Schätzers \hat{g} . Ein Schätzer ist also genau dann erwartungstreu (engl.: UNBIASED), wenn sein Bias gleich Null ist.

BEISPIEL XI.4.3. Ist X binomial verteilt mit den Parametern n und p , so ist $E(\frac{X}{n}) = p$ (vgl. Abschnitt XI.3.5 (S. 428)). Daher ist die Schätzung $\hat{p}(X) = \frac{X}{n}$ aus Beispiel XI.4.2 erwartungstreu.

BEISPIEL XI.4.4. Häufig erfolgt die Messung einer unbekanntem Größe μ durch Ausführen von n unabhängigen Zufallsexperimenten, die μ als Erwartungswert haben. Mathematisch bedeutet dies, dass wir n unabhängige Zufallsvariable X_1, \dots, X_n und eine unbekanntem Familie P_ϑ von Wahrscheinlichkeiten haben, so dass $E_\vartheta(X_i) = \mu$ ist für alle i und ϑ . Sei

$$\begin{aligned} g_1(\vartheta) &= E_\vartheta(X_i) \\ &= \mu \end{aligned}$$

und

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

der MITTELWERT. Dann ist für jedes ϑ

$$\begin{aligned} E_\vartheta(\bar{X}) &= \frac{1}{n} \sum_{i=1}^n E_\vartheta(X_i) \\ &= \mu. \end{aligned}$$

Also ist der Mittelwert ein erwartungstreuem Schätzer für g_1 . Haben die X_i alle die gleiche unbekanntem Varianz σ^2 , so bestimmt die Familie P_ϑ auch diese Größe. Setze in diesem Fall

$$\begin{aligned} g_2(\vartheta) &= \sigma^2 \\ &= \text{Var}(X_i). \end{aligned}$$

Wir wählen

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

als Schätzer für σ^2 . (Beachte, dass durch $n-1$ geteilt wird!) Da die X_i unabhängig sind, erhalten wir für jedes ϑ

$$\begin{aligned} E_\vartheta((\bar{X} - \mu)^2) &= \text{Var}_\vartheta(\bar{X}) \\ &= \frac{1}{n^2} \text{Var}_\vartheta(X_1 + \dots + X_n) \\ &= \frac{1}{n^2} (\text{Var}_\vartheta(X_1) + \dots + \text{Var}_\vartheta(X_n)) \\ &= \frac{\sigma^2}{n} \end{aligned}$$

und

$$\begin{aligned}
 E_{\vartheta}((X_i - \bar{X})^2) &= E_{\vartheta}(((X_i - \mu) - (\bar{X} - \mu))^2) \\
 &= E_{\vartheta}((X_i - \mu)^2) - 2E_{\vartheta}((X_i - \mu)(\bar{X} - \mu)) \\
 &\quad + E_{\vartheta}((\bar{X} - \mu)^2) \\
 &= \sigma^2 - \frac{2}{n} \sum_{j=1}^n \underbrace{E_{\vartheta}((X_i - \mu)(X_j - \mu))}_{=0 \text{ für } j \neq i} + \frac{\sigma^2}{n} \\
 &= \sigma^2 - \frac{2}{n} E_{\vartheta}((X_i - \mu)^2) + \frac{\sigma^2}{n} \\
 &= \sigma^2 - \frac{\sigma^2}{n} \\
 &= \sigma^2 \left(1 - \frac{1}{n}\right).
 \end{aligned}$$

Hieraus folgt

$$\begin{aligned}
 E_{\vartheta}(s^2) &= \frac{1}{n-1} \sum_{i=1}^n E_{\vartheta}((X_i - \bar{X})^2) \\
 &= \frac{1}{n-1} n \sigma^2 \left(1 - \frac{1}{n}\right) \\
 &= \sigma^2.
 \end{aligned}$$

Also ist s^2 ein erwartungstreuer Schätzer für die Unbekannte Varianz σ^2 .

XI.4.5. Der mittlere quadratische Fehler. Die wohl wichtigste Forderung an einen Schätzer T von $g(\vartheta)$ ist wohl diejenige, dass der Fehler $T(X) - g(\vartheta)$ der Schätzwerte „klein“ sein sollte. Um diese Größe quantitativ zu messen, betrachtet man den MITTLEREN QUADRATISCHEN FEHLER

$$R(\vartheta, T) = E_{\vartheta}((T(X) - g(\vartheta))^2).$$

Man kann den mittleren quadratischen Fehler auch durch die Varianz und den Bias ausdrücken. Berücksichtigt man, dass $E_{\vartheta}(T) - g(\vartheta)$ eine Konstante ist, erhält man nämlich

$$\begin{aligned}
 &E_{\vartheta} \left((T(X) - g(\vartheta))^2 \right) \\
 &= E_{\vartheta} \left(\left((T(X) - E_{\vartheta}(T)) - (g(\vartheta) - E_{\vartheta}(T)) \right)^2 \right)
 \end{aligned}$$

$$\begin{aligned}
 &= E_{\vartheta} \left(\underbrace{\left(T(X) - E_{\vartheta}(T) \right)^2}_{=\text{Var}_{\vartheta}(T)} \right) \\
 &\quad - 2 \underbrace{E_{\vartheta} \left(\left(T(X) - E_{\vartheta}(T) \right) \left(g(\vartheta) - E_{\vartheta}(T) \right) \right)}_{=(g(\vartheta) - E_{\vartheta}(T)) E_{\vartheta}(T(X) - E_{\vartheta}(T)) = 0}} \\
 &\quad + E_{\vartheta} \left(\underbrace{\left(g(\vartheta) - E_{\vartheta}(T) \right)^2}_{=(g(\vartheta) - E_{\vartheta}(T))^2} \right) \\
 &= \text{Var}_{\vartheta}(T) + b(\vartheta, T)^2.
 \end{aligned}$$

Also ist der mittlere quadratische Fehler

$$R(\vartheta, T) = \text{Var}_{\vartheta}(T) + b(\vartheta, T)^2.$$

BEISPIEL XI.4.5. Für den Schätzer $T(X) = \frac{X}{n}$ aus Beispiel XI.4.3 (S. 437) für die binomial verteilte Zufallsvariable X gilt gemäß Beispiel XI.4.3 und Abschnitt XI.3.5 (S. 428)

$$\begin{aligned}
 b(\vartheta, T) &= 0, \\
 \text{Var}_{\vartheta}(T) &= \frac{1}{n^2} \text{Var}_{\vartheta}(X) \\
 &= \frac{1}{n} \vartheta(1 - \vartheta).
 \end{aligned}$$

Also ist der mittlere quadratische Fehler

$$R(\vartheta, T) = \frac{1}{n} \vartheta(1 - \vartheta).$$

Für alle $\vartheta \in \Theta = (0, 1)$ gilt

$$R(\vartheta, T) \leq \frac{1}{4n}.$$

BEISPIEL XI.4.6. Wir greifen Beispiel XI.4.4 (S. 437) auf. Der Bias des Mittelwertes \bar{X} ist Null; seine Varianz ist gemäß Beispiel XI.4.4 gleich $\frac{\sigma^2}{n}$, wenn alle X_i die gleiche Varianz σ^2 haben. Also ist der mittlere quadratische Fehler

$$R(\vartheta, \bar{X}) = \frac{\sigma^2}{n}.$$

XI.5. Approximationen der Binomialverteilung

Für große Werte von n ist die exakte Berechnung der Wahrscheinlichkeit

$$b_{n,p}(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

äußerst mühselig und anfällig für Rundungsfehler. Dies gilt vermehrt für Summen dieser Ausdrücke. Wir wollen daher in diesem Abschnitt Approximationen für $b_{n,p}(k)$ bestimmen, die leichter und mit geringeren Rundungsfehlern berechnet werden können.

Zur Vereinfachung der Sprache führen wir die Notation

$$a_n \sim b_n \iff \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$$

ein. $a_n \sim b_n$ bedeutet, dass es eine Nullfolge $(\theta_n)_{n \in \mathbb{N}}$ gibt mit

$$b_n(1 - \theta_n) \leq a_n \leq b_n(1 + \theta_n)$$

für alle bis auf endlich viele n .

XI.5.1. Approximation von $n!$ und $b_{n,p}(k)$. Unser wichtigstes Hilfsmittel ist die STIRLINGSISCHE FORMEL

$$n! \sim \sqrt{2\pi n} n^n e^{-n}.$$

Für die eingangs erwähnte Folge $(\theta_n)_{n \in \mathbb{N}}$ kann man in diesem Fall zeigen:

$$\frac{1}{12n+1} \leq \theta_n \leq \frac{1}{12n}$$

für alle n .

Tabelle XI.5.1 gibt einen Eindruck von der Genauigkeit der Stirlingschen Formel.

TABELLE XI.5.1. Genauigkeit der Stirlingschen Formel

n	$n!$	$\sqrt{2\pi n} n^n e^{-n}$	$\frac{ n! - \sqrt{2\pi n} n^n e^{-n} }{n!}$
2	2	1.919	4.05%
5	120	118.019	1.65%
10	3628800	3598690	0.83%

Mit der Stirlingschen Formel folgt

$$\begin{aligned} b_{n,p}(k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &\sim \frac{\sqrt{2\pi n} n^n e^{-n}}{\sqrt{2\pi k} k^k e^{-k} \sqrt{2\pi(n-k)} (n-k)^{(n-k)} e^{-(n-k)}} p^k (1-p)^{n-k} \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{n(1-p)}{n-k}\right)^{n-k}. \end{aligned}$$

Wir betrachten nun eine Folge $(k_n)_{n \in \mathbb{N}}$ mit $k_n \sim np$. Dann ist $n - k_n \sim n(1 - p)$ und

$$\begin{aligned} \sqrt{\frac{n}{k_n(n - k_n)}} &\sim \frac{1}{\sqrt{np(1 - p)}} \\ &= \frac{1}{\sigma_n}, \end{aligned}$$

wobei

$$\sigma_n = \sqrt{np(1 - p)}$$

die Standardabweichung der $b_{n,p}$ -Verteilung ist (vgl. Abschnitt [XI.3.5](#) (S. 428)). Wir wollen nun das Verhalten von

$$\chi(n, k_n) = \left(\frac{np}{k_n}\right)^{k_n} \left(\frac{n(1 - p)}{n - k_n}\right)^{n - k_n}$$

studieren. Dazu bilden wir den Logarithmus und setzen

$$t_n = \frac{k_n}{n}.$$

Dann gilt $t_n \rightarrow p$ für $n \rightarrow \infty$ und

$$\begin{aligned} -\ln \chi(n, k_n) &= -k_n \ln\left(\frac{np}{k_n}\right) - (n - k_n) \ln\left(\frac{n(1 - p)}{n - k_n}\right) \\ &= n \left\{ \frac{k_n}{n} \ln\left(\frac{t_n}{p}\right) + \left(1 - \frac{k_n}{n}\right) \ln\left(\frac{1 - \frac{k_n}{n}}{1 - p}\right) \right\} \\ &= n \left\{ t_n \ln\left(\frac{t_n}{p}\right) + (1 - t_n) \ln\left(\frac{1 - t_n}{1 - p}\right) \right\}. \end{aligned}$$

Definiere die Funktion $g : (0, 1) \rightarrow \mathbb{R}$ durch

$$g(t) = t \ln\left(\frac{t}{p}\right) + (1 - t) \ln\left(\frac{1 - t}{1 - p}\right).$$

Dann ist

$$\begin{aligned} g(p) &= 0, \\ g'(p) &= 0, \\ g''(p) &= \frac{1}{p} + \frac{1}{1 - p} \\ &= \frac{1}{p(1 - p)}. \end{aligned}$$

Nach der Taylor Formel ist daher

$$g(t) = \frac{1}{2p(1 - p)}(t - p)^2 + r(t - p) \quad \text{mit } |r(t - p)| \leq c|t - p|^3$$

in einer Umgebung von p .

Wir nehmen nun an, dass nicht nur $t_n \rightarrow p$ gilt, sondern dass sogar $n(t_n - p)^3 \rightarrow 0$ gilt. Dann folgt $n|r(t_n - p)| \rightarrow 0$ und daher

$$\left| -\ln \chi(n, k_n) - \frac{n(t_n - p)^2}{2p(1-p)} \right| \xrightarrow{n \rightarrow \infty} 0.$$

Setzen wir

$$x(n, k_n) = \frac{k_n - np}{\sigma_n},$$

so ist

$$\frac{n(t_n - p)^2}{2p(1-p)} = \frac{x(n, k_n)^2}{2}.$$

Damit erhalten wir

$$\chi(n, k_n) \sim \exp\left(-\frac{x(n, k_n)^2}{2}\right).$$

Insgesamt haben wir somit unter obigen Voraussetzungen die Beziehung

$$b_{n,p}(k_n) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_n} \exp\left(-\frac{x(n, k_n)^2}{2}\right)$$

gezeigt.

Die Funktion

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

heißt DICHTER DER STANDARD-NORMALVERTEILUNG. Damit lässt sich unsere Approximation der $b_{n,p}$ -Verteilung schreiben als

$$b_{n,p}(k) \sim \frac{1}{\sigma_n} \varphi(x(n, k))$$

mit

$$\sigma_n = \sqrt{np(1-p)},$$

$$x(n, k) = \frac{k - pn}{\sigma_n}.$$

Mit Hilfe von Ergebnissen der Funktionentheorie folgt

$$\int_{-\infty}^{\infty} \varphi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1.$$

XI.5.2. Der Satz von Moivre-Laplace. Wir wollen Summen von Wahrscheinlichkeiten $b_{n,p}(k)$ für große n näherungsweise berechnen. Dazu definieren wir

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Die Funktion Φ heißt die VERTEILUNGSFUNKTION DER STANDARD-NORMALVERTEILUNG. Sie hat folgende Eigenschaften

$$\begin{aligned} \lim_{x \rightarrow -\infty} \Phi(x) &= 0 \\ \lim_{x \rightarrow +\infty} \Phi(x) &= 1 \\ \Phi(0) &= \frac{1}{2} \\ \Phi(-x) &= 1 - \Phi(x) \quad \text{für } x \geq 0 \\ \Phi(x) &\leq \Phi(y) \quad \text{für } x \leq y \end{aligned}$$

Sie ist ausführlich tabelliert (z.B. in U. Krengel: Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg 2000, S. 248).

Sei nun S_n eine $b_{n,p}$ -verteilte Zufallsvariable. Setze

$$\begin{aligned} S_n^* &= \frac{S_n - np}{\sigma_n} \\ &= \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}}. \end{aligned}$$

Dann ist $E(S_n^*) = 0$ und $\text{Var}(S_n^*) = 1$. S_n^* heißt daher die STANDARDISIERTE oder NORMIERTE Form von S_n . Nimmt S_n den Wert k an, so hat S_n^* den Wert $x(n, k)$ mit $x(n, k)$ aus dem vorigen Abschnitt.

Der SATZ VON MOIVRE-LAPLACE besagt, dass Φ eine gute Approximation für die Verteilung von S_n^* ist:

$$\lim_{n \rightarrow \infty} P(a \leq S_n^* \leq b) = \Phi(b) - \Phi(a).$$

BEISPIEL XI.5.1. Wie groß ist näherungsweise die Wahrscheinlichkeit, bei 600 Würfeln mit einem Würfel mindestens 90 und höchstens 100 Sechsen zu erhalten?

Es ist

$$n = 600$$

und

$$p = \frac{1}{6},$$

also

$$np = 100$$

und

$$\begin{aligned}\sigma_n &= \sqrt{600 \cdot \frac{1}{6} \cdot \frac{5}{6}} \\ &= \sqrt{\frac{250}{3}} \\ &\approx 9.13\end{aligned}$$

Damit folgt

$$\begin{aligned}P(90 \leq S_n \leq 100) &= P\left(\frac{90 - 100}{\sigma_n} \leq S_n^* \leq \frac{100 - 100}{\sigma_n}\right) \\ &\approx \Phi(0) - \Phi\left(-\frac{10}{9.13}\right) \\ &= 0.5 - \Phi(-1.095) \\ &= 0.5 - (1 - \Phi(1.095)) \\ &\approx 0.36.\end{aligned}$$

BEISPIEL XI.5.2. Wir wollen den Prozentsatz der Wähler der Partei A schätzen. Werden n Wähler befragt und sind darunter S_n Wähler der Partei A , so sei $\frac{S_n}{n}$ der Schätzer für die Wahrscheinlichkeit p , dass ein zufällig ausgewählter Wähler die Partei A wählt. Wie groß muss n sein, damit die Wahrscheinlichkeit eines Irrtums um 1% nicht größer ist als 0.05?

Es muss also gelten

$$P(-0.01 \leq \frac{S_n}{n} - p \leq 0.01) \approx 0.95.$$

Mit

$$\sigma_n = \sqrt{np(1-p)}$$

erhalten wir

$$\begin{aligned}0.95 &\approx P\left(-\frac{0.01n}{\sigma_n} \leq S_n^* \leq \frac{0.01n}{\sigma_n}\right) \\ &\approx \Phi\left(\frac{0.01n}{\sigma_n}\right) - \Phi\left(-\frac{0.01n}{\sigma_n}\right) \\ &= 2\Phi\left(\frac{0.01n}{\sigma_n}\right) - 1\end{aligned}$$

also

$$\Phi\left(\frac{0.01n}{\sigma_n}\right) \approx 0.975.$$

Da Φ streng monoton ist, existiert die Umkehrfunktion Φ^{-1} . Es muss also gelten

$$\frac{0.01n}{\sigma_n} \approx \Phi^{-1}(0.975).$$

Aus einer Tabelle von Φ entnehmen wir

$$\Phi^{-1}(0.975) \approx 1.96.$$

Also muss gelten

$$\begin{aligned} 1.96 &\approx \frac{0.01}{\sigma_n} \\ &= \frac{0.01\sqrt{n}}{\sqrt{p(1-p)}} \\ \implies \sqrt{n} &\approx 196\sqrt{p(1-p)} \\ \implies n &\approx 196^2 p(1-p). \end{aligned}$$

Wir kennen aber p nicht! Aber wir wissen, dass für alle möglichen p , nämlich $0 \leq p \leq 1$, gilt

$$p(1-p) \leq \frac{1}{4}.$$

Damit erhalten wir die Schätzung

$$\begin{aligned} n &\approx \frac{196^2}{4} \\ &= 98^2 \\ &= 9604. \end{aligned}$$

Wir müssen also ca. 9600 Wähler befragen.

Hätten wir z.B. die Zusatzinformation $p \leq 0.1$ zur Verfügung, kämen wir wegen

$$\max_{0 \leq p \leq 0.1} p(1-p) = 0.1 \cdot 0.9 = 0.09$$

mit ca. 3450 Befragungen aus.

XI.5.3. Die Poisson-Approximation. In diesem Abschnitt geben wir eine andere Approximation der $b_{n,p}$ -Verteilung an, die für kleine Werte von p besser ist als die Normalverteilung aus Abschnitt [XI.5.1](#).

Eine Zufallsvariable X heißt POISSON-VERTEILT mit Parameter $\lambda \geq 0$, wenn gilt

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{mit } k = 0, 1, \dots$$

Ist $(p_n)_{n \in \mathbb{N}}$ eine Folge in $[0, 1]$ mit $np_n \rightarrow \lambda$, so gilt für jedes k die POISSON-APPROXIMATION

$$b_{n,p_n}(k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} \\ \xrightarrow{n \rightarrow \infty} e^{-\lambda} \frac{\lambda^k}{k!}.$$

Um dies einzusehen, setze $\lambda_n = np_n$. Dann gilt für festes k

$$\begin{aligned} b_{n,p_n}(k) &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\ &= \frac{1}{k!} \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{n^k} \lambda_n^k \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\ &= \frac{1}{k!} \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-k+1}{n} \lambda_n^k \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k}. \end{aligned}$$

Für $n \rightarrow \infty$ streben die Quotienten $\frac{n}{n}$, $\frac{n-1}{n}$, \dots , $\frac{n-k+1}{n}$ alle gegen 1; nach Voraussetzung gilt

$$\lambda_n^k \rightarrow \lambda^k$$

und

$$\left(1 - \frac{\lambda_n}{n}\right)^{-k} \rightarrow 1.$$

Wegen der Stetigkeit der Exponentialfunktion gilt

$$\left(1 - \frac{\lambda_n}{n}\right)^n \rightarrow e^{-\lambda}.$$

Damit folgt die Poisson-Approximation.

TABELLE XI.5.2. Poisson-Approximation für die Wahrscheinlichkeit, dass unter 91 Personen genau k am heutigen Tage Geburtstag haben

k	$P(X = k) \approx \frac{0.25^k}{k!} e^{-0.25}$
1	0.1947
2	0.0243
5	$6.337 \cdot 10^{-6}$

BEISPIEL XI.5.3. In einem Hörsaal befinden sich 91 Studierende. Die Wahrscheinlichkeit p , heute Geburtstag zu haben, ist $p = \frac{1}{365}$. Die Zahl derer, die heute Geburtstag haben, ist praktisch Poissonverteilt mit $\lambda = \frac{91}{365} \approx 0.25$. Tabelle XI.5.2 gibt die mit der Poisson-Approximation genäherte Wahrscheinlichkeit $P(X = k)$ an, dass heute genau k Personen aus dem Hörsaal Geburtstag haben.

BEISPIEL XI.5.4. Bei der Produktion von Blitzlichtlampen ist mit der Wahrscheinlichkeit $p = 0.015$ eine Lampe schon bei der Produktion defekt. Wie groß muss man n wählen, damit in einem Karton mit n unabhängig gewählten Lampen mit Wahrscheinlichkeit ≥ 0.8 mindestens 100 funktionstüchtige Exemplare sind?

Die Zahl n ist minimal zu wählen unter der Bedingung

$$0.8 \leq P(\text{„höchstens } n - 100 \text{ fehlerhafte Lampen“}) \\ = \sum_{k=0}^{n-100} b_{n,p}(k).$$

Approximieren wir die rechte Seite mit der Poisson-Verteilung zu $\lambda_n = np$, erhalten wir die Bedingung

$$0.8 \leq r_n = e^{-\lambda_n} \sum_{k=0}^{n-100} \frac{\lambda_n^k}{k!}.$$

Mit Hilfe eines Taschenrechners prüfen wir nach, dass $n = 102$ die kleinste Zahl ist, die diese Bedingung erfüllt. Wir müssen also 102 Lampen in den Karton legen.

Würden wir die Wahrscheinlichkeit auf 0.9 erhöhen, müssten wir 103 Lampen in den Karton legen.

XI.6. Tests

XI.6.1. Motivation. Eine Person behauptet, dass sie bei Milchkaffee schmecken kann, ob zuerst der Kaffee oder die Milch in die Tasse gegeben wurde. Wie können wir die Stichhaltigkeit dieser Behauptung nachprüfen?

Ein erster Ansatz geht auf Fisher (1935) zurück. Er schlägt folgende „Versuchs-anordnung“ vor:

In vier Tassen wird zuerst Kaffee und dann Milch gegossen. Diese Tassen nennen wir Tassen vom Typ 1. In weitere vier Tassen wird zuerst Milch und dann Kaffee gegossen. Diese Tassen nennen wir vom Typ 2. Die Kaffee- und Milchmenge ist dabei natürlich bei allen Tassen die gleiche, und es wird gut umgerührt. Nun werden der Person die acht Tassen in beliebiger Reihenfolge (8! Möglichkeiten) vorgesetzt, und sie soll die vier Tassen vom Typ 1

identifizieren. Falls ihr das gelingt, wollen wir ihre Behauptung als wahr akzeptieren.

Die Wahrscheinlichkeit, dass die Person rein zufällig richtig rät, ist aufgrund der hypergeometrischen Verteilung aus Abschnitt [XI.1.4](#) (S. 414) gleich

$$\binom{8}{4}^{-1} = \frac{1}{70}.$$

Wir geben der Person also nur mit dieser Wahrscheinlichkeit Recht, wenn ihre Behauptung nicht stimmt.

Schwieriger wird die Angelegenheit, wenn die Person behauptet, nicht unfehlbar zu sein, aber mit großer Sicherheit die Unterscheidung treffen zu können. Wenn wir ihr die Behauptung in diesem Fall schon dann glauben, wenn sie mindestens drei der vier Tassen vom Typ 1 richtig identifiziert, ist die Wahrscheinlichkeit für einen zufälligen Treffer schon

$$\begin{aligned} \frac{\binom{4}{4}\binom{4}{0} + \binom{4}{3}\binom{4}{1}}{\binom{8}{4}} &= \frac{1 + 4 \cdot 4}{70} \\ &= \frac{17}{70} \\ &\approx 0.24. \end{aligned}$$

Wie kann man in diesem Fall die Gewissheit, der Person auf die Schliche zu kommen, erhöhen, ohne sie dabei zu benachteiligen?

Neyman (1950) schlägt dazu die folgende verbesserte „Versuchsanordnung“ vor:

Der Person wird n -mal die Aufgabe gestellt, zwei Tassen, von denen eine vom Typ 1 und eine vom Typ 2 ist, korrekt zu klassifizieren. Die beiden Tassen werden ihr jeweils in einer zufälligen, durch Münzwurf bestimmten Reihenfolge vorgesetzt. Damit die Person unabhängig von früheren Entscheidungen urteilen kann, wird jeder Teilversuch an einem anderen Tag ausgeführt.

X sei die Zahl der Tage, an denen die Person die Tassen richtig klassifiziert.

Ein mathematisches Modell für diese Versuchsanordnung ist die Annahme, X sei binomial verteilt mit Parametern n und p . Die Hypothese „die Person schwindelt und trifft ihre Entscheidung rein zufällig“ entspricht dem Fall $p = \frac{1}{2}$. Diese Hypothese nennt man die „Nullhypothese“. Die Alternative ist, dass die Person tatsächlich die behauptete Fähigkeit hat, und entspricht dem Fall $p > \frac{1}{2}$. Es wird nun eine Zahl t festgelegt, so dass $P(X \geq t) \leq \alpha$ ist. Dabei ist α eine vorgegebene, kleine, positive Schranke, z.B. $\alpha = 0.05$. Dies ist die „Sicherheitsmarge“,

mit der wir ausschließen wollen, dass wir der Person ungerechtfertigterweise auf den Leim gehen. Falls die Person nun mindestens t -mal die Tassen richtig klassifiziert, wollen wir ihr glauben.

Der Unterschied des Neymanschen Ansatzes zu dem von Fisher liegt darin, dass er präzise Aussagen über die Wahrscheinlichkeit erlaubt, die Nullhypothese zu akzeptieren, wenn die Alternative zutrifft.

XI.6.2. Grundbegriffe der Testtheorie. Von einem Testproblem spricht man, wenn eine zufällige Größe X mit einer unbekanntem Verteilung P_ϑ beobachtet wird, und man aufgrund des beobachteten Wertes x von X entscheiden soll, ob P_ϑ einer bestimmten Menge von Verteilungen angehört oder nicht.

Zur genaueren mathematischen Beschreibung sei \mathcal{X} die Menge der möglichen Werte der Zufallsvariablen X und $\{P_\vartheta : \vartheta \in \Theta\}$ die Menge der in Betracht gezogenen Verteilungen von X . Unter diesen möglichen Verteilungen ist eine echte Teilmenge $\{P_\vartheta : \vartheta \in H\}$ durch zusätzliche Bedingungen ausgezeichnet. Ein TEST ist eine Entscheidungsregel, die für jeden möglichen Wert von X festlegt, ob man sich für die HYPOTHESE „ $\vartheta \in H$ “ oder für die ALTERNATIVE „ $\vartheta \in \Theta \setminus H$ “ entscheiden soll. Man nennt auch kurz H die Hypothese und $K = \Theta \setminus H$ die Alternative. Die Entscheidung für die Hypothese nennt man ANNAHME DER HYPOTHESE, und die Entscheidung für die Alternative nennt man VERWERFEN DER HYPOTHESE.

Ein Test ist also beschrieben durch Angabe der Menge R derjenigen Werte x von X , für die die Hypothese verworfen werden soll. R heißt VERWERFUNGSBEREICH oder KRITISCHER BEREICH des Tests.

Innerhalb des gewählten Modells sind zwei Arten von Fehlern möglich:

- Ist $\vartheta \in H$ und wird die Hypothese verworfen, spricht man von einem FEHLER ERSTER ART.
- Ist $\vartheta \in K$ und wird die Hypothese angenommen, spricht man von einem FEHLER ZWEITER ART.

Praktisch beschreibt man den Verwerfungsbereich R durch eine Funktion $T(x)$, die so gewählt wird, dass große Werte gegen die Hypothese sprechen. T heißt TESTSTATISTIK. Man wählt dann einen KRITISCHEN WERT t und verwirft die Hypothese, wenn $T(x) \geq t$ ist. Es ist also

$$R = \{x : T(x) \geq t\}.$$

Bisher ist das Testproblem so formuliert, dass H und K symmetrische Rollen spielen. Dies ist aber in der Praxis in der Regel nicht der Fall. Vielmehr wird man die Hypothese so wählen, dass sie der etablierten Hypothese oder der bisherigen Erfahrung entspricht. Bei unserem einführenden Beispiel bedeutet dies, dass man als Hypothese

„die Person schwindelt“ nehmen wird. Bei der Einführung eines neuen Medikamentes würde man z.B. die Hypothese „das Medikament ist nicht wirksamer als die bekannten Mittel“ wählen.

Man zieht nur Verwerfungsbereiche in Betracht, für die die Wahrscheinlichkeit eines Fehlers erster Art durch eine vorgegebene kleine Zahl $\alpha > 0$ begrenzt ist. Quantitative Aussagen erhält man durch die GÜTE-FUNKTION

$$\beta(\vartheta) = P_{\vartheta}(X \in R).$$

Sie ordnet jedem ϑ die Verwerfungswahrscheinlichkeit unter P_{ϑ} zu. Der Test hat das Niveau α , wenn für alle $\vartheta \in H$ die Ungleichung $\beta(\vartheta) \leq \alpha$ gilt. Die Wahrscheinlichkeit eines Fehlers erster Art ist dann also maximal α . Gebräuchliche Werte für α sind 0.05, 0.02 oder 0.01.

XI.6.3. Zurück zur Motivation. In unserem motivierenden Beispiel des ersten Abschnittes ist das gesuchte Modell beschrieben durch

$$\begin{aligned} \mathcal{X} &= \{0, 1, \dots, n\}, \\ \Theta &= \left[\frac{1}{2}, 1\right], \\ \vartheta &= p, \\ P_{\vartheta}(X = x) &= b_{n,p}(x) \\ &= \binom{n}{x} p^x (1-p)^{n-x}. \end{aligned}$$

Die Hypothese ist

$$H = \left\{\frac{1}{2}\right\}$$

und die Alternative ist

$$K = \left(\frac{1}{2}, 1\right].$$

Der Verwerfungsbereich ist von der Form

$$R = \{x : x \geq t\}.$$

Sei

$$\beta(p|t, n) = P_p(X \geq t)$$

die Gütefunktion dieses Tests. Als Niveau wollen wir

$$\alpha = 0.05$$

wählen. Wir müssen dann für gegebenes n den Parameter t so wählen, dass

$$\beta\left(\frac{1}{2}|t, n\right) \leq 0.05$$

ist.

Für $n = 5$ kommt nur $t = 5$ in Frage. Denn für $t = 4$ ist

$$\begin{aligned}\beta\left(\frac{1}{2}|4, 5\right) &= \left(\frac{1}{2}\right)^5 + \binom{5}{1}\left(\frac{1}{2}\right)^5 \\ &= \frac{6}{32} \\ &\approx 0.187 \\ &> 0.05.\end{aligned}$$

Die Person müsste also in jedem der 5 Versuche die Tassen richtig klassifizieren, damit wir ihr glauben.

Für $n = 5$ ist

$$\beta(p|5, 5) = p^5.$$

Es ist

$$\beta(0.6|5, 5) \approx 0.08$$

und

$$\beta(0.9|5, 5) \approx 0.59.$$

Hätte also die Person eine Erfolgswahrscheinlichkeit von 0.6 pro Klassifizierung, würde sie doch nur mit Wahrscheinlichkeit 0.08 ihre Fähigkeit beweisen können. Selbst bei einer Erfolgswahrscheinlichkeit von 0.9, würde ihre Behauptung nur mit Wahrscheinlichkeit 0.59 akzeptiert. Die Gütefunktion zeigt also, ob der Test überhaupt in der Lage ist, eine Abweichung von der Hypothese anzuzeigen.

Ist z.B. $p = 0.6$, ist erst bei $n = 42$ zu klassifizierenden Tassenpaaren die Wahrscheinlichkeit wenigstens $\frac{1}{3}$, dass die Behauptung der Person akzeptiert wird. Der kleinste Wert von t mit $\beta\left(\frac{1}{2}|t, 42\right) \leq 0.05$ ist $t = 27$. Die Person müsste also mindestens 27 von 42 Tassenpaaren richtig klassifizieren.

KAPITEL XII

Stochastik II Allgemeine Modelle

XII.1. Wahrscheinlichkeitsmaße mit Dichten

XII.1.1. Ergebnismengen. Im vorigen Kapitel haben wir stets endliche Ergebnismengen Ω betrachtet. Die Bedingung $\text{card}(\Omega) < \infty$ lassen wir jetzt fallen und betrachten allgemeine Ergebnismengen Ω . Die wichtigsten Beispiele für dieses Kapitel sind $\Omega = \mathbb{R}$ und $\Omega = \mathbb{R}^n$.

XII.1.2. σ -Algebren. Bisher waren die Wahrscheinlichkeitsmaße stets auf der ganzen Potenzmenge $\mathcal{P}(\Omega)$ definiert, d.h. jede Teilmenge von Ω war ein mögliches Ereignis. Man kann zeigen, dass dies für unendliche Mengen wie $\Omega = \mathbb{N}$ oder $\Omega = \mathbb{R}$ zu Widersprüchen führt. Daher können wir nur noch Teilmengen \mathcal{A} von $\mathcal{P}(\Omega)$ als mögliche Ereignismengen betrachten. Diese müssen aber gewisse Eigenschaften haben.

Der richtige Begriff ist hier derjenige der σ -Algebra. Eine Menge $\mathcal{A} \subset \mathcal{P}(\Omega)$ von Teilmengen von Ω heißt σ -ALGEBRA (über Ω), wenn folgende drei Eigenschaften erfüllt sind:

$$\begin{array}{l} \Omega \in \mathcal{A} \\ A \in \mathcal{A} \implies \Omega \setminus A \in \mathcal{A} \\ A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \end{array}$$

Sei nun $\mathcal{B} \subset \mathcal{P}(\Omega)$. Dann kann man zeigen, dass es genau eine σ -Algebra \mathcal{A} mit den folgenden beiden Eigenschaften gibt:

- $\mathcal{B} \subset \mathcal{A}$.
- Für jede σ -Algebra $\tilde{\mathcal{A}}$ mit $\mathcal{B} \subset \tilde{\mathcal{A}}$ gilt $\mathcal{A} \subset \tilde{\mathcal{A}}$.

\mathcal{A} ist gewissermaßen die kleinste σ -Algebra, die das Mengensystem \mathcal{B} enthält. \mathcal{A} heißt die von \mathcal{B} ERZEUGTE σ -ALGEBRA.

BEISPIEL XII.1.1. Sei $\Omega = \mathbb{R}$. Betrachte folgende Mengensysteme

$$\begin{array}{l} \mathcal{B}_1 = \{(-\infty, a] : a \in \mathbb{R}\}, \\ \mathcal{B}_2 = \{(-\infty, a) : a \in \mathbb{R}\}, \\ \mathcal{B}_3 = \{(a, b) : a, b \in \mathbb{R}, a < b\}, \\ \mathcal{B}_4 = \{[a, b] : a, b \in \mathbb{R}, a \leq b\}. \end{array}$$

Alle vier Mengensysteme erzeugen die gleiche σ -Algebra \mathcal{B} . Diese heißt die σ -Algebra der BORELMENGEN von \mathbb{R} . Alle Intervalle, egal ob offen oder nicht, egal ob beschränkt oder nicht, sind in \mathcal{B} enthalten.

BEISPIEL XII.1.2. Sei $\Omega = \mathbb{R}^n$. Wir definieren die Relation \leq auf \mathbb{R}^n durch

$$x \leq y \iff x_i \leq y_i \text{ für alle } 1 \leq i \leq n.$$

Mit dieser Relation können wir Intervalle verallgemeinern:

$$[a, b] = \{x \in \mathbb{R}^n : a \leq x \leq b\}.$$

Für $n = 2$ ist z.B. $[a, b]$ das Rechteck, dessen linke untere Ecke die Koordinaten (a_1, a_2) hat, und dessen rechte obere Ecke die Koordinaten (b_1, b_2) hat. Das Mengensystem $\{[a, b] : a, b \in \mathbb{R}^n\}$ aller dieser Intervalle erzeugt eine σ -Algebra \mathcal{B} . Diese heißt wieder σ -Algebra der BORELMENGEN von \mathbb{R}^n . Zu ihr gehören z.B. alle offenen und alle abgeschlossenen Mengen (vgl. Abschnitt VIII.2.1 (S. 283)).

XII.1.3. Wahrscheinlichkeitsmaße. Sei nun Ω eine beliebige Menge und \mathcal{A} irgendeine σ -Algebra über Ω . Eine Abbildung $P : \mathcal{A} \rightarrow [0, 1]$ heißt ein WAHRSCHEINLICHKEITSMASS auf Ω , wenn folgende drei Bedingungen erfüllt sind:

$$P(\Omega) = 1$$

$$P(A) \geq 0 \quad \text{für alle } A \in \mathcal{A}$$

$$A_1, A_2, \dots \in \mathcal{A} \text{ paarweise disjunkt} \implies P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Ein WAHRSCHEINLICHKEITSRAUM ist ein Tripel (Ω, \mathcal{A}, P) mit einer Menge Ω , einer σ -Algebra \mathcal{A} und einem Wahrscheinlichkeitsmaß P . Die Elemente von \mathcal{A} heißen EREIGNISSE.

Im vorigen Kapitel konnten wir auf die Angabe der σ -Algebra \mathcal{A} verzichten, weil diese „per ordre de mufti“ stets die Potenzmenge $\mathcal{P}(\Omega)$ war.

Eine wichtige Konsequenz aus den obigen Eigenschaften eines Wahrscheinlichkeitsmaßes ist:

Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum.

- Ist $B_1 \subset B_2 \subset \dots$ eine wachsende Folge von Ereignissen und B deren Vereinigung, so ist

$$P(B) = \lim_{i \rightarrow \infty} P(B_i).$$

- Ist $C_1 \supset C_2 \supset \dots$ eine fallende Folge von Ereignissen und C deren Durchschnitt, so ist

$$P(C) = \lim_{i \rightarrow \infty} P(C_i).$$

XII.1.4. Wahrscheinlichkeitsmaße mit Dichten. Wir betrachten nun den Fall $\Omega = \mathbb{R}$ und $\mathcal{A} = \mathcal{B}$ die Menge der Borelmengen auf \mathbb{R} . Sei P ein Wahrscheinlichkeitsmaß auf \mathbb{R} . Dann existiert für jedes $x \in \mathbb{R}$ die Zahl

$$F(x) = P((-\infty, x]).$$

Die Funktion $F : \mathbb{R} \rightarrow [0, 1]$ ist monoton wachsend mit

$$\begin{aligned} \lim_{x \rightarrow -\infty} F(x) &= 0, \\ \lim_{x \rightarrow +\infty} F(x) &= 1. \end{aligned}$$

Außerdem existiert für jedes $x \in \mathbb{R}$ der rechtsseitige Grenzwert $F(x+0)$ (vgl. Abschnitt III.3.4 (S. 125)) und stimmt mit $F(x)$ überein. Die Funktion F heißt die zu P gehörende VERTEILUNGSFUNKTION.

Ist umgekehrt $F : \mathbb{R} \rightarrow [0, 1]$ eine Funktion mit obigen Eigenschaften, so wird durch

$$P([a, b]) = F(b) - F(a)$$

ein Wahrscheinlichkeitsmaß auf \mathbb{R} definiert. Die gegebene Funktion F ist dann die zugehörige Verteilungsfunktion. In diesem Sinne besteht eine eindeutige Beziehung zwischen Wahrscheinlichkeitsmaßen auf \mathbb{R} und Verteilungsfunktionen.

Eine Funktion $f : \mathbb{R} \rightarrow [0, \infty)$ mit

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

heißt DICHTE. Sie definiert durch

$$F(x) = \int_{-\infty}^x f(t) dt$$

eine Verteilungsfunktion und damit ein Wahrscheinlichkeitsmaß P auf \mathbb{R} . Die Funktion f heißt dann die zu P gehörende Dichte.

Nicht jedes Wahrscheinlichkeitsmaß auf \mathbb{R} besitzt eine Dichte. Aber Wahrscheinlichkeitsmaße mit Dichte sind für die Anwendungen besonders wichtig.

XII.1.5. Gleichverteilung auf einem Intervall. Seien $a, b \in \mathbb{R}$ mit $a < b$. Die Funktion $f : \mathbb{R} \rightarrow [0, \infty)$ mit

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{falls } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

ist eine Dichte. Sie heißt Dichte der GLEICHVERTEILUNG. Das zugehörige Wahrscheinlichkeitsmaß ordnet jedem Intervall I seinen relativen Anteil an $[a, b]$ zu.

XII.1.6. Exponentialverteilung. Für jedes $\lambda > 0$ wird durch

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & \text{falls } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

eine Dichte definiert. Sie heißt Dichte der EXPONENTIALVERTEILUNG.

XII.1.7. Normalverteilung. Die Funktion

$$\varphi_{\mu, \sigma^2}(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

heißt Dichte der NORMALVERTEILUNG mit Mittelwert μ und Varianz σ^2 . Ist

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

die Funktion aus Abschnitt XI.5.1 (S. 440), so ist

$$\varphi_{\mu, \sigma^2}(x) = \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right).$$

Mit

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

und der Substitutionsregel (vgl. Abschnitt V.2.3 (S. 178)) folgt daher

$$\int_a^b \varphi_{\mu, \sigma^2}(x) dx = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

XII.1.8. Produktdichten. Werden n Telexperimente durch Dichten f_1, \dots, f_n beschrieben, verwendet man

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

als Dichte für die Verteilung auf \mathbb{R}^n , die die unabhängige Hintereinanderausführung der Telexperimente beschreibt.

XII.2. Zufallsvariable und ihre Momente

XII.2.1. Messbare Funktionen. Seien Ω und Ω' zwei beliebige Mengen und \mathcal{A} und \mathcal{A}' zwei beliebige σ -Algebren auf Ω bzw. Ω' . Eine Abbildung $f : \Omega \rightarrow \Omega'$ heißt MESSBAR, wenn für jede Menge $A' \in \mathcal{A}'$ das Urbild

$$f^{-1}(A') = \{x \in \Omega : f(x) \in A'\}$$

unter f in \mathcal{A} enthalten ist.

Man beachte, dass der Begriff der Messbarkeit von den σ -Algebren \mathcal{A} und \mathcal{A}' abhängt.

BEISPIEL XII.2.1. Sei Ω eine endliche Menge und $\mathcal{A} = \mathcal{P}(\Omega)$. Dann ist für jede Menge Ω' und jede σ -Algebra \mathcal{A}' auf Ω' jede Abbildung $f : \Omega \rightarrow \Omega'$ messbar. Aus diesem Grunde haben wir den Begriff der Messbarkeit in Kapitel XI nicht benötigt.

BEISPIEL XII.2.2. Sei $\Omega = \mathbb{R}^m$, $\Omega' = \mathbb{R}^n$ und \mathcal{A} und \mathcal{A}' die σ -Algebren der Borelmengen. Dann ist jede stetige Abbildung $f : \Omega \rightarrow \Omega'$ messbar. Ist speziell $m = 1$, so ist jede stückweise stetige Funktion (vgl. Abschnitt V.1.1 (S. 169)) messbar.

Sind $f : \Omega \rightarrow \Omega'$ und $g : \Omega' \rightarrow \Omega''$ messbar, dann ist auch die Komposition $g \circ f : \Omega \rightarrow \Omega''$ messbar. Sind $f_1, f_2 : \Omega \rightarrow \mathbb{R}$ messbar und $\alpha_1, \alpha_2 \in \mathbb{R}$, so sind auch $\alpha_1 f_1 + \alpha_2 f_2$, $f_1 \cdot f_2$, $\min\{f_1, f_2\}$ und $\max\{f_1, f_2\}$ messbar.

XII.2.2. Zufallsvariable. Seien (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, Ω' eine beliebige Menge und \mathcal{A}' eine σ -Algebra auf Ω' . Eine messbare Funktion $X : \Omega \rightarrow \Omega'$ heißt (Ω' -wertige) ZUFALLSVARIABLE. Jede Zufallsvariable $X : \Omega \rightarrow \Omega'$ definiert durch

$$P_X(A') = P(X \in A')$$

für $A' \in \mathcal{A}'$ ein Wahrscheinlichkeitsmaß auf Ω' . Ist speziell $\Omega' = \mathbb{R}$, so lässt sich dieses Wahrscheinlichkeitsmaß durch seine Verteilungsfunktion (vgl. Abschnitt XII.1.4 (S. 455))

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x)$$

beschreiben. F_X heißt die VERTEILUNGSFUNKTION der reellwertigen Zufallsvariablen X . Besitzt F_X eine Dichte f , so heißt diese die DICHTe der Zufallsvariablen.

Seien nun $X : \Omega \rightarrow \mathbb{R}$ eine reellwertige Zufallsvariable und $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ eine streng monoton wachsende, differenzierbare Funktion. Dann ist $Y = \varphi \circ X$ auch eine reellwertige Zufallsvariable. X habe die Dichte f . Dann folgt aus der Substitutionsregel (vgl. Abschnitt V.2.3 (S. 178)) für jedes $y \in \mathbb{R}$

$$F_Y(y) = P(\varphi \circ X \leq y)$$

$$\begin{aligned}
&= P(X \leq \varphi^{-1}(y)) \quad \text{wegen der Monotonie} \\
&= \int_{-\infty}^{\varphi^{-1}(y)} f(x) dx \quad |x = \varphi^{-1}(t), dx = \frac{1}{\varphi'(\varphi^{-1}(t))} dt \\
&= \int_{-\infty}^y f(\varphi^{-1}(t)) \frac{1}{\varphi'(\varphi^{-1}(t))} dt.
\end{aligned}$$

Also hat $\varphi \circ X$ die Dichte

$$g(y) = \frac{f(\varphi^{-1}(y))}{\varphi'(\varphi^{-1}(y))}.$$

Mit den gleichen Argumenten folgt, dass für eine streng monoton fallende, differenzierbare Funktion $\psi : \mathbb{R} \rightarrow \mathbb{R}$ die Zufallsvariable $\psi \circ X$ die Dichte

$$g(y) = \frac{f(\psi^{-1}(y))}{|\psi'(\psi^{-1}(y))|}$$

hat. Speziell folgt:

- Hat X die Dichte f , so hat $X + a$ die Dichte

$$g(y) = f(y - a).$$

- Ist $c \neq 0$, hat cX die Dichte

$$g(y) = \frac{1}{|c|} f\left(\frac{y}{c}\right)$$

XII.2.3. Unabhängigkeit. Seien $\Omega_i, i \in I$, beliebige Mengen und \mathcal{A}_i σ -Algebren auf den Ω_i , sowie (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $X_i : \Omega \rightarrow \Omega_i$ Zufallsvariable. Dann heißen die X_i UNABHÄNGIG, wenn für alle $A_i \in \mathcal{A}_i$ die Ereignisse $\{X_i \in A_i\}$ unabhängig sind.

Diese Definition ist eine direkte Verallgemeinerung derjenigen aus Abschnitt XI.3.3 (S. 426). Für die praktische Rechnung ist sie aber wenig geeignet. Für den Spezialfall reellwertiger Zufallsvariablen mit Dichten hilft folgendes Ergebnis:

Die reellwertigen Zufallsvariablen X_1, \dots, X_n mit Dichten f_1, \dots, f_n sind genau dann unabhängig, wenn $X = (X_1, \dots, X_n)$ die Produktdichte

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

besitzt.

Seien X_1 und X_2 zwei reellwertige Zufallsvariable mit Dichten f_1 und f_2 . Dann folgt aus Obigem, dass $X_1 + X_2$ die Dichte

$$(f_1 * f_2)(u) = \int_{-\infty}^{\infty} f_1(u - v) f_2(v) dv$$

besitzt. Der Ausdruck $f_1 * f_2$ heißt FALTUNG der Dichten f_1 und f_2 .

BEISPIEL XII.2.3. Betrachte zwei unabhängige, reellwertige Zufallsvariable X_1 und X_2 , die $N(\mu_i, \sigma_i^2)$ -verteilt sind, $i = 1, 2$. Wir behaupten, dass ihre Summe $X_1 + X_2$ dann $N(\mu, \sigma^2)$ -verteilt ist mit

$$\mu = \mu_1 + \mu_2$$

und

$$\sigma^2 = \sigma_1^2 + \sigma_2^2.$$

Wir müssen also für alle $u \in \mathbb{R}$ die Gleichung

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{(u - \mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{(u - v - \mu_1)^2}{2\sigma_1^2}\right) \exp\left(-\frac{(v - \mu_2)^2}{2\sigma_2^2}\right) dv \end{aligned}$$

beweisen. Diese ist äquivalent zu

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left[\frac{(u - v - \mu_1)^2}{\sigma_1^2} + \frac{(v - \mu_2)^2}{\sigma_2^2} \right. \right. \\ & \qquad \qquad \qquad \left. \left. - \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \right] \right) dv \end{aligned}$$

$$= 1.$$

Wir betrachten den Term in eckigen Klammern und erhalten mit der Abkürzung

$$\begin{aligned} x &= u - \mu_1 - \mu_2 \\ y &= v - \mu_2 \end{aligned}$$

die Beziehung

$$\begin{aligned} & \frac{(u - v - \mu_1)^2}{\sigma_1^2} + \frac{(v - \mu_2)^2}{\sigma_2^2} - \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \\ &= \frac{(x - y)^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} - \frac{x^2}{\sigma_1^2 + \sigma_2^2} \\ &= \frac{x^2}{\sigma_1^2} - \frac{2xy}{\sigma_1^2} + \frac{y^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} - \frac{x^2}{\sigma_1^2 + \sigma_2^2} \\ &= \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x^2 - \frac{1}{\sigma_1^2} 2xy + \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 \sigma_2^2} y^2 \\ &= \left(\frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} x - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} y \right)^2. \end{aligned}$$

Daher liefert die Variablentransformation

$$t = \frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} (u - \mu_1 - \mu_2) - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} (v - \mu_2)$$

$$dt = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} dv$$

für alle $a < b$ und alle $u \in \mathbb{R}$

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2} \int_a^b \exp\left(-\frac{1}{2} \left[\frac{(u-v-\mu_1)^2}{\sigma_1^2} + \frac{(v-\mu_2)^2}{\sigma_2^2} - \frac{(u-\mu_1-\mu_2)^2}{\sigma_1^2 + \sigma_2^2} \right]\right) dv \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}(u-\mu_1-\mu_2) - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2}(b-\mu_2)}^{\frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}(u-\mu_1-\mu_2) - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1 \sigma_2}(a-\mu_2)} \exp\left(-\frac{1}{2}t^2\right) dt \\ &\xrightarrow[\substack{a \rightarrow -\infty \\ b \rightarrow \infty}]{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt \\ &= 1. \end{aligned}$$

Dies beweist die Behauptung.

XII.2.4. Erwartungswert. Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $X : \Omega \rightarrow \mathbb{R}$ eine reellwertige Zufallsvariable.

Falls X nur endlich viele Werte x_1, \dots, x_n annimmt, können wir X wie in Abschnitt XI.3.4 (S. 426) einen Erwartungswert durch

$$E(X) = \sum_{i=1}^n x_i P(X = x_i)$$

zuordnen.

Wir wollen diese Einschränkung gerne fallen lassen. Dazu nehmen wir zunächst an, dass X beschränkt ist, d.h. es gibt ein $R > 0$ mit $|X(\omega)| \leq R$ für alle $\omega \in \Omega$. Für $k \in \mathbb{Z}$ und $n \in \mathbb{N}^*$ definieren wir dann

$$\begin{aligned} A_{k,n} &= \left\{ \omega \in \Omega : \frac{k}{n} \leq X(\omega) < \frac{k+1}{n} \right\}, \\ \chi_{A_{k,n}}(\omega) &= \begin{cases} 1 & \text{falls } \omega \in A_{k,n}, \\ 0 & \text{sonst,} \end{cases} \\ X_n &= \sum_{k \in \mathbb{Z}} \frac{k}{n} \chi_{A_{k,n}}. \end{aligned}$$

Da X beschränkt ist, nimmt X_n nur endlich viele Werte an und es ist

$$E(X_n) = \sum_{k \in \mathbb{Z}} \frac{k}{n} P(A_{k,n}).$$

Man kann zeigen, dass die Folge $(E(X_n))_{n \in \mathbb{N}}$ konvergiert. Den Grenzwert nennen wir Erwartungswert von X und bezeichnen ihn mit

$$E(X) = \lim_{n \rightarrow \infty} E(X_n).$$

Falls X die stückweise stetige Dichte f besitzt, erhalten wir für jedes n

$$\begin{aligned} E(X_n) &= \sum_{k \in \mathbb{Z}} \frac{k}{n} P(A_{k,n}) \\ &= \sum_{k \in \mathbb{Z}} \frac{k}{n} \int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x) dx. \end{aligned}$$

Man kann zeigen, dass dieser Ausdruck für $n \rightarrow \infty$ gegen

$$\int_{-\infty}^{\infty} x f(x) dx$$

konvergiert. Daher erhalten wir in diesem Fall

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Wenn X nicht beschränkt ist, benutzen wir diese Identität als Definition des Erwartungswertes:

Die reellwertige Zufallsvariable X auf Ω besitze die stückweise stetige Dichte f . Dann ist der ERWARTUNGSWERT von X definiert durch

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

sofern das uneigentliche Integral existiert.

BEISPIEL XII.2.4. X sei $N(\mu, \sigma^2)$ -verteilt. Für $a < 0 < b$ folgt

$$\begin{aligned} & \int_a^b x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \left| t = \frac{x-\mu}{\sigma}, dt = \frac{1}{\sigma} dx \right. \\ &= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} (\sigma t + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \underbrace{t e^{-\frac{1}{2}t^2}}_{=-\frac{d}{dt} e^{-\frac{1}{2}t^2}} dt + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\ &= \frac{\sigma}{\sqrt{2\pi}} \left[e^{-\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{b-\mu}{\sigma}\right)^2} \right] + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\ & \xrightarrow[a \rightarrow -\infty]{b \rightarrow \infty} \mu. \end{aligned}$$

Also hat X den Erwartungswert μ .

BEISPIEL XII.2.5. X sei exponentiell verteilt mit Parameter λ . Für $a > 0$ gilt

$$\begin{aligned} \int_0^a x \underbrace{\lambda e^{-\lambda x}}_{=-\frac{d}{dx} e^{-\lambda x}} dx &= -x e^{-\lambda x} \Big|_{x=0}^{x=a} + \int_0^a e^{-\lambda x} dx \\ &= -a e^{-\lambda a} - \frac{1}{\lambda} e^{-\lambda a} + \frac{1}{\lambda} \\ &\xrightarrow{a \rightarrow \infty} \frac{1}{\lambda}. \end{aligned}$$

Da $X(x) = 0$ ist für $x < 0$, folgt

$$E(X) = \frac{1}{\lambda}.$$

Offensichtlich gilt für den Erwartungswert

$$E(aX + bY) = aE(X) + bE(Y),$$

wobei X, Y reellwertige Zufallsvariablen und a, b reelle Zahlen sind.

XII.2.5. Varianz. Sei X eine reellwertige Zufallsvariable auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit einem endlichem Erwartungswert $E(X)$. Die VARIANZ von X ist definiert durch

$$\text{Var}(X) = E((X - E(X))^2),$$

sofern dieser Erwartungswert existiert. Die STANDARDABWEICHUNG ist dann definiert durch

$$\sigma = \sqrt{\text{Var}(X)}.$$

Wie in Abschnitt XI.3.5 (S. 428) gilt

$$\text{Var}(X) = E(X^2) - E(X)^2.$$

Hat daher X die stückweise stetige Dichte f , ergibt sich

$$\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - \left[\int_{-\infty}^{\infty} x f(x) dx \right]^2.$$

BEISPIEL XII.2.6. Sei X $N(\mu, \sigma^2)$ -verteilt. Für $a < 0 < b$ erhalten wir wie in Beispiel XII.2.4 (S. 461)

$$\int_a^b x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \left| t = \frac{x-\mu}{\sigma}, dt = \frac{1}{\sigma} dx \right.$$

$$\begin{aligned}
&= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} (\sigma t + \mu)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} t^2 e^{-\frac{1}{2}t^2} dt \\
&\quad + \frac{2\sigma\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} t e^{-\frac{1}{2}t^2} dt \\
&\quad + \frac{\mu^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\
&= \frac{\sigma^2}{\sqrt{2\pi}} t(-e^{-\frac{1}{2}t^2}) \Big|_{t=\frac{a-\mu}{\sigma}}^{t=\frac{b-\mu}{\sigma}} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\
&\quad + \frac{2\sigma\mu}{\sqrt{2\pi}} (-e^{-\frac{1}{2}t^2}) \Big|_{t=\frac{a-\mu}{\sigma}}^{t=\frac{b-\mu}{\sigma}} \\
&\quad + \frac{\mu^2}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt \\
&\xrightarrow[\substack{a \rightarrow -\infty \\ b \rightarrow \infty}]{} \sigma^2 + \mu^2.
\end{aligned}$$

Da $E(X) = \mu$ ist, folgt

$$\text{Var}(X) = \sigma^2.$$

BEISPIEL XII.2.7. X sei exponentiell verteilt mit Parameter λ . Für $a > 0$ erhalten wir

$$\begin{aligned}
\int_0^a x^2 \lambda e^{-\lambda x} dx &= x^2 (-e^{-\lambda x}) \Big|_{x=0}^{x=a} + 2 \int_0^a x e^{-\lambda x} dx \\
&= -a^2 e^{-\lambda a} + \frac{2}{\lambda} x (-e^{-\lambda x}) \Big|_{x=0}^{x=a} + \frac{2}{\lambda} \int_0^a e^{-\lambda x} dx \\
&= -a^2 e^{-\lambda a} - \frac{2}{\lambda} a e^{-\lambda a} - \frac{2}{\lambda^2} e^{-\lambda a} + \frac{2}{\lambda^2} \\
&\xrightarrow{a \rightarrow \infty} \frac{2}{\lambda^2}.
\end{aligned}$$

Wegen $E(X) = \frac{1}{\lambda}$ ergibt sich

$$\text{Var}(X) = \frac{1}{\lambda^2}.$$

Tabelle [XII.2.1](#) fasst die Ergebnisse der vorigen Beispiele zusammen.

Wie in Abschnitt [XI.3.5](#) (S. 428) gilt die Rechenregel

$$\text{Var}(aX + b) = a^2 \text{Var}(X),$$

TABELLE XII.2.1. Erwartungswert und Varianz der Normal- und Exponentialverteilung

Verteilung	Erwartungswert	Varianz
$N(\mu, \sigma^2)$	μ	σ^2
Exponential λ	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

wobei X eine reellwertige Zufallsvariable und a, b reelle Zahlen sind. Ebenso gilt

$$X_1, \dots, X_n \text{ unabhängig}$$

$$\implies \text{Var}(X_1 + \dots + X_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n).$$

XII.3. Schätzverfahren

XII.3.1. Maximum-Likelihood Schätzung. Es werde eine Zufallsvariable X mit Werten in \mathbb{R}^n beobachtet. Die Verteilung von X hänge von einem unbekanntem Parameter $\vartheta \in \Theta$ ab. Wir nennen sie P_ϑ . P_ϑ habe für jedes $\vartheta \in \Theta$ eine Dichte $f(\cdot; \vartheta)$. Dann ist für jedes $x \in \mathbb{R}^n$ $P_\vartheta(x) = 0$. Wir können also nicht, wie im diskreten Fall, aus der Betrachtung von $P_\vartheta(x)$ Schätzer ableiten. Statt dessen definieren wir nun die Likelihood-Funktion durch die Dichte

$$L_x(\vartheta) = f(x; \vartheta).$$

Wie in Abschnitt XI.4.3 (S. 435) setzen wir

$$\mathcal{L}_x(\vartheta) = \ln L_x(\vartheta).$$

BEISPIEL XII.3.1. Die Zufallsvariablen X_1, \dots, X_n seien unabhängig und jeweils $N(\mu, \sigma^2)$ -verteilt. Dann ist $\vartheta = (\mu, \sigma)$. Die Dichte von X_i ist

$$f_i(x; \vartheta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

Daher hat $X = (X_1, \dots, X_n)$ an der Stelle $x = (x_1, \dots, x_n)$ die Produktdichte

$$\begin{aligned} f(x; \vartheta) &= \prod_{i=1}^n f_i(x; \vartheta) \\ &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right). \end{aligned}$$

Der Maximum-Likelihood Schätzer $\hat{\vartheta} = \hat{\vartheta}(x)$ ist wieder der Parameterwert, der $L_x(\vartheta)$ bzw. $\mathcal{L}_x(\vartheta)$ maximiert. Aus obiger Darstellung der Dichte erhalten wir

$$\mathcal{L}_x(\vartheta) = -n \ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

und somit

$$\begin{aligned} \frac{\partial}{\partial \mu} \mathcal{L}_x(\vartheta) &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \\ &= -\frac{n\mu}{\sigma^2} + \frac{1}{\sigma^2} \sum_{i=1}^n x_i, \\ \frac{\partial}{\partial \sigma} \mathcal{L}_x(\vartheta) &= -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2. \end{aligned}$$

Wir müssen nun drei Fälle unterscheiden:

1. μ IST UNBEKANNT, ABER $\sigma^2 = \sigma_0^2$ IST BEKANNT: Dann ist $\Theta = \{(\mu, \sigma_0) : \mu \in \mathbb{R}\}$, und wir müssen eine Nullstelle $\hat{\mu}$ von $\frac{\partial}{\partial \mu} \mathcal{L}_x(\mu, \sigma_0)$ finden. Aus der Formel für $\frac{\partial}{\partial \mu} \mathcal{L}_x$ folgt sofort

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Indem wir die zweite Ableitung bilden, sehen wir, dass $\hat{\mu}$ wirklich ein Maximum ist.

2. $\mu = \mu_0$ IST BEKANNT, ABER σ^2 IST UNBEKANNT: Jetzt ist $\Theta = \{(\mu_0, \sigma) : \sigma > 0\}$, und wir müssen eine Nullstelle $\hat{\sigma}$ von $\frac{\partial}{\partial \sigma} \mathcal{L}_x(\mu_0, \sigma)$ finden. Aus der Formel für $\frac{\partial}{\partial \sigma} \mathcal{L}_x$ ergibt sich

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2.$$

Wieder folgt durch Bilden der zweiten Ableitung, dass $\hat{\sigma}^2$ tatsächlich ein Maximum ist.

3. μ UND σ^2 SIND BEIDE UNBEKANNT: Die Gleichungen

$$\begin{aligned} \frac{\partial}{\partial \mu} \mathcal{L}_x(\mu, \sigma) &= -\frac{1}{\sigma^2} \left[n\mu - \sum_{i=1}^n x_i \right] \\ &= 0, \\ \frac{\partial}{\partial \sigma} \mathcal{L}_x(\mu, \sigma) &= \frac{1}{\sigma^3} \left[-n\sigma^2 + \sum_{i=1}^n (x_i - \mu)^2 \right] \end{aligned}$$

$$= 0$$

für einen kritischen Punkt liefern die Lösungen

$$\begin{aligned} \hat{\mu} &= \bar{x}, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \end{aligned}$$

Durch Bilden der Hesse-Matrix sehen wir, dass diese Werte wirklich das Maximum liefern.

Der Schätzer $\hat{\mu}$ und der Schätzer $\hat{\sigma}^2$ bei bekanntem Erwartungswert μ_0 sind erwartungstreu. Der Schätzer $\hat{\sigma}^2$ im Fall 3 dagegen ist nicht erwartungstreu. Analog zu Abschnitt XI.4.4 (S. 436) liefert in diesem Fall

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

einen erwartungstreuen Schätzer für die Varianz.

XII.3.2. Die Methode der kleinsten Quadrate. Oft stellt sich das Problem, eine Gerade, eine Parabel oder eine andere „einfache“ Funktion einer gegebenen Menge von Messwerten anzupassen. Zum Beispiel kann eine Größe y in Abhängigkeit von einer Größe x gemessen worden sein und es liegen nun n Messergebnisse $(x_1, y_1), \dots, (x_n, y_n)$ vor. Wenn die Messergebnisse relativ gut auf einer Geraden liegen, können wir einen linearen Zusammenhang der beobachteten Größen vermuten, der durch Messfehler z_i gestört ist. Dann wäre

$$y_i = \alpha + \beta x_i + z_i, \quad i = 1, \dots, n,$$

mit zu bestimmenden Parametern α und β .

Allgemeiner nehmen wir an, dass $\vartheta_1, \dots, \vartheta_p$ unbekannte Parameter sind und dass für bekannte Funktionen φ_i der gemessene Wert bei der i -ten Messung von der Form ist

$$y_i = \varphi_i(\vartheta_1, \dots, \vartheta_p) + z_i$$

mit einem Messfehler z_i . Im Beispiel der Geraden ist

$$\begin{aligned} p &= 2, \\ \vartheta_1 &= \alpha, \\ \vartheta_2 &= \beta, \\ \varphi_i(\vartheta_1, \vartheta_2) &= \vartheta_1 + \vartheta_2 x_i. \end{aligned}$$

Die METHODE DER KLEINSTEN QUADRATE besteht darin, die Parameter $\vartheta_1, \dots, \vartheta_p$ so zu bestimmen, dass die Größe

$$Q = \sum_{i=1}^n [y_i - \varphi_i(\vartheta_1, \dots, \vartheta_p)]^2$$

minimal wird.

Dieser Ansatz kann ad hoc ohne Statistik formuliert werden und wird häufig auch so angewandt. Wir wollen nun zeigen, dass dieser Ansatz statistisch fundiert ist. Dazu nehmen wir an, dass die Messfehler z_i Realisierungen von unabhängigen $N(0, \sigma^2)$ -verteilten Zufallsvariablen Z_1, \dots, Z_n sind. Dann sind die y_i Realisierungen von unabhängigen Zufallsvariablen Y_1, \dots, Y_n , die $N(\varphi_i(\vartheta_1, \dots, \vartheta_p), \sigma^2)$ -verteilt sind. Daher hat $Y = (Y_1, \dots, Y_n)$ die Dichte

$$\left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - \varphi_i(\vartheta_1, \dots, \vartheta_p)]^2 \right).$$

Diese Dichte ist genau dann maximal, wenn obige Größe Q minimal ist. Die Methode der kleinsten Quadrate ist also gerade der Maximum-Likelihood Schätzer für die Parameter $\vartheta_1, \dots, \vartheta_p$.

BEISPIEL XII.3.2 (REGRESSIONSGERADE). Wir wollen eine Gerade $y = \alpha + \beta x$ mit der Methode der kleinsten Quadrate an Messdaten $(x_1, y_1), \dots, (x_n, y_n)$ anpassen. Dann ist

$$Q = Q(\alpha, \beta) = \sum_{i=1}^n (y_i - \beta x_i - \alpha)^2.$$

Die Gleichungen für einen kritischen Punkt $\hat{\alpha}, \hat{\beta}$ lauten

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} Q(\hat{\alpha}, \hat{\beta}) = -2 \sum_{i=1}^n (y_i - \hat{\beta} x_i - \hat{\alpha}) \\ 0 &= \frac{\partial}{\partial \beta} Q(\hat{\alpha}, \hat{\beta}) = -2 \sum_{i=1}^n (y_i - \hat{\beta} x_i - \hat{\alpha}) x_i. \end{aligned}$$

Zur Abkürzung setzen wir

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i, & \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i, \\ \overline{xx} &= \frac{1}{n} \sum_{i=1}^n x_i^2, & \overline{xy} &= \frac{1}{n} \sum_{i=1}^n x_i y_i. \end{aligned}$$

Nach Multiplikation mit $-\frac{1}{2n}$ nehmen dann obige Bestimmungsgleichungen die Form an

$$\begin{aligned} 0 &= \bar{y} - \widehat{\beta}\bar{x} - \widehat{\alpha} \\ 0 &= \overline{xy} - \widehat{\beta}\overline{xx} - \widehat{\alpha}\bar{x}. \end{aligned}$$

Dies liefert die Lösung

$$\begin{aligned} \widehat{\beta} &= \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{xx} - \bar{x}\bar{x}}, \\ \widehat{\alpha} &= \bar{y} - \widehat{\beta}\bar{x}. \end{aligned}$$

Die Gerade $y = \widehat{\alpha} + \widehat{\beta}x$ heißt **REGRESSIONSGERADE**.
Mit den Abkürzungen

$$\begin{aligned} s_{xy} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ s_{xx} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \end{aligned}$$

erhält man wegen

$$\begin{aligned} s_{xy} &= \frac{n}{n-1} \overline{xy} - \frac{n}{n-1} \bar{x}\bar{y}, \\ s_{xx} &= \frac{n}{n-1} \overline{xx} - \frac{n}{n-1} \bar{x}\bar{x} \end{aligned}$$

die alternative Darstellung

$$\widehat{\beta} = \frac{s_{xy}}{s_{xx}}.$$

XII.3.3. Median. Wenn in einem schweizer Bergdorf fünfzig Einheimische und fünf zugezogene Millionäre leben, ist es für die Einheimischen wenig befriedigend, wenn man ihnen erklärt, das durchschnittliche Einkommen in diesem Dorf sei hoch. Dieses Beispiel zeigt, dass der Erwartungswert manchmal kein guter Maßstab für das „Zentrum“ einer Verteilung ist. Einige wenige „Ausreißer“ können ihn stark verfälschen. Man betrachtet daher auch andere Maßzahlen. Die bekannteste davon ist der Median.

Ist Z eine reellwertige Zufallsvariable, so heißt *jede* Zahl μ_m mit

$$P(Z \geq \mu_m) \geq \frac{1}{2}$$

und

$$P(Z \leq \mu_m) \geq \frac{1}{2}$$

ein MEDIAN von Z .

Man beachte, dass der Median μ_m nicht notwendig eindeutig bestimmt ist. Eine Mehrdeutigkeit tritt genau dann auf, wenn es ein Intervall $[a, b]$ gibt mit $a < b$ und $P(Z \geq b) = \frac{1}{2}$ und $P(Z \leq a) = \frac{1}{2}$. Dann ist jede Zahl in dem Intervall $[a, b]$ ein Median. Gibt es dagegen eine Zahl c mit $P(Z \geq c+t) = P(Z \leq c-t)$ für alle $t > 0$, so ist c der eindeutige Median von Z . In diesem Fall heißt Z symmetrisch, und es ist $c = E(Z)$.

Für n Messwerte x_1, \dots, x_n können wir den EMPIRISCHEN MEDIAN $\hat{\mu}_m$ wie folgt ermitteln:

Bezeichne mit $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ die der Größe nach umgeordneten x_i . Dann ist

$$\hat{\mu}_m = \begin{cases} x_{(\frac{n+1}{2})} & \text{falls } n \text{ ungerade,} \\ \frac{1}{2}[x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}] & \text{falls } n \text{ gerade.} \end{cases}$$

BEISPIEL XII.3.3. Für $x_1 = 3.1, x_2 = 4.5, x_3 = 4.5, x_4 = 2.8, x_5 = 2.6, x_6 = 9.8$ erhalten wir $x_{(1)} = 2.6, x_{(2)} = 2.8, x_{(3)} = 3.1, x_{(4)} = 4.5, x_{(5)} = 4.5, x_{(6)} = 9.8$ und $\hat{\mu}_m = \frac{1}{2}[x_{(3)} + x_{(4)}] = \frac{1}{2}[3.1 + 4.5] = 3.8$. Der Mittelwert ist $\bar{x} = 4.55$.

XII.4. Tests

XII.4.1. Vorbemerkungen. Wir beobachten eine, im allgemeinen vektorwertige, Zufallsvariable X , deren Verteilung einer Familie $\{P_\vartheta : \vartheta \in \Theta\}$ angehört. Θ ist die disjunkte Vereinigung der Mengen H und K , der Hypothese und der Alternative. Aufgrund des beobachteten Wertes x von X soll entschieden werden, ob der Parameter ϑ der zu X gehörenden Verteilung in der Menge H liegt oder nicht.

Hierzu bilden wir den LIKELIHOOD-QUOTIENTEN

$$q(x) = \frac{\sup\{L_x(\vartheta) : \vartheta \in K\}}{\sup\{L_x(\vartheta) : \vartheta \in H\}}.$$

Offensichtlich ist $0 \leq q(x)$ für alle x . $q(x) \approx 0$ spricht für die Hypothese, $q(x) \gg 0$ spricht für die Alternative.

Ein TEST ist nun eine messbare Abbildung φ des Wertebereiches \mathcal{X} von X in $[0, 1]$. Wird x beobachtet, besagt $\varphi(x) = 1$, dass die Hypothese verworfen werden soll, und $\varphi(x) = 0$, dass sie angenommen werden soll. Im Fall $0 < \varphi(x) < 1$ soll ein zusätzliches Zufallsexperiment mit Wahrscheinlichkeit $\varphi(x)$ zur Verwerfung führen.

Ein Test φ heißt LIKELIHOOD-QUOTIENTEN TEST, wenn es ein $c > 0$ gibt mit den Eigenschaften

- $q(x) > c \implies \varphi(x) = 1$ und
- $q(x) < c \implies \varphi(x) = 0$.

XII.4.2. Der t -Test. Wir beobachten Zufallsvariable X_1, \dots, X_n , die unabhängig und $N(\mu, \sigma^2)$ -verteilt seien mit unbekanntem (μ, σ^2) . Für ein gegebenes μ_0 sei zu testen, ob $\mu = \mu_0$ ist oder nicht. Dann ist

$$\begin{aligned}\vartheta &= (\mu, \sigma) \\ \Theta &= \{(\mu, \sigma) : \mu \in \mathbb{R}, \sigma > 0\} \\ H &= \{(\mu_0, \sigma) : \sigma > 0\} \\ K &= \{(\mu, \sigma) : \mu \neq \mu_0, \sigma > 0\}.\end{aligned}$$

Die Dichte $f(x; \vartheta)$ von $X = (X_1, \dots, X_n)$ ist gemäß Abschnitt [XII.3.1](#) (S. 464)

$$f(x; \vartheta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right).$$

Aus Stetigkeitsgründen ist

$$\sup\{f(x; \vartheta) : \vartheta \in K\} = \sup\{f(x; \vartheta) : \vartheta \in \Theta\}.$$

Damit ergibt sich aus Abschnitt [XII.3.1](#) (S. 464), dass das Supremum an der Stelle

$$\begin{aligned}\hat{\mu} &= \bar{x} \\ &= \frac{1}{n} \sum_{i=1}^n x_i, \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\end{aligned}$$

angenommen wird. Ebenso folgt, dass das Supremum $\sup\{f(x; \vartheta) : \vartheta \in H\}$ an der Stelle

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2$$

angenommen wird. Daher ist der Likelihood-Quotient

$$\begin{aligned}q(x) &= \frac{f(x; \hat{\mu}, \hat{\sigma})}{f(x; \mu_0, \tilde{\sigma})} \\ &= \left(\frac{\tilde{\sigma}}{\hat{\sigma}}\right)^n \exp\left(-\frac{1}{2\hat{\sigma}^2} \underbrace{\sum_{i=1}^n (x_i - \hat{\mu})^2}_{=n\hat{\sigma}^2} + \frac{1}{2\tilde{\sigma}^2} \underbrace{\sum_{i=1}^n (x_i - \mu_0)^2}_{=n\tilde{\sigma}^2}\right) \\ &= \left(\frac{\tilde{\sigma}}{\hat{\sigma}}\right)^n.\end{aligned}$$

Ist φ irgendein Likelihood-Quotienten Test, gibt es daher ein $c > 0$ mit

$$\begin{aligned}\varphi(x) &= 1 && \text{auf } \left\{ \left(\frac{\tilde{\sigma}}{\hat{\sigma}} \right)^n > c \right\} \\ \varphi(x) &= 0 && \text{auf } \left\{ \left(\frac{\tilde{\sigma}}{\hat{\sigma}} \right)^n < c \right\}.\end{aligned}$$

Mit

$$c' = c^{2/n}$$

ist dies äquivalent zu

$$\begin{aligned}\varphi(x) &= 1 && \text{auf } \left\{ \frac{\tilde{\sigma}^2}{\hat{\sigma}^2} > c' \right\} \\ \varphi(x) &= 0 && \text{auf } \left\{ \frac{\tilde{\sigma}^2}{\hat{\sigma}^2} < c' \right\}.\end{aligned}$$

Es ist

$$\begin{aligned}\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} &= \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= 1 + \frac{(\bar{x} - \mu_0)^2}{\hat{\sigma}^2}.\end{aligned}$$

Wir definieren daher

$$T(x) = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s(x)}$$

mit

$$s(x)^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Dann folgt

$$\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} = 1 + \frac{1}{n-1} |T(x)|^2.$$

Also gibt es für jeden Likelihood-Quotienten Test φ ein $t > 0$ mit

$$\varphi(x) = \begin{cases} 1 & \text{falls } |T(x)| > t, \\ 0 & \text{falls } |T(x)| < t. \end{cases}$$

Unser Testproblem ist damit auf die Bestimmung der Verteilung von $T(x)$ zurückgeführt. Diese Verteilung heißt t -VERTEILUNG mit $n - 1$ Freiheitsgraden oder kurz t_{n-1} -VERTEILUNG. Die zugehörige Dichte ist gegeben durch

$$h_{n-1}(x) = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}}.$$

Dabei ist Γ die Eulersche Gammafunktion aus Abschnitt V.4.3 (S. 196). Für $n \rightarrow \infty$ konvergiert $h_{n-1}(x)$ gegen die Dichte $\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ der Standard-Normalverteilung.

Damit lautet der t -TEST ZUM NIVEAU α AUF DIE HYPOTHESE EINER $N(\mu_0, \sigma^2)$ -VERTEILTEN STICHPROBE:

Berechne die Größen

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s(x) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$T(x) = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s(x)}.$$

Bestimme aus einer Tabelle der t_{n-1} -Verteilung das $(1 - \frac{\alpha}{2})$ -QUANTIL $t_{n-1, 1-\frac{\alpha}{2}}$ aus der Bedingung

$$\int_{-\infty}^{t_{n-1, 1-\frac{\alpha}{2}}} h_{n-1}(x) dx = 1 - \frac{\alpha}{2}.$$

Falls $|T(x)| \leq t_{n-1, 1-\frac{\alpha}{2}}$ ist, wird die Hypothese angenommen, sonst wird sie verworfen.

BEISPIEL XII.4.1. An 15 Glasfaserplatten einer bestimmten Stärke wird die Wärmeleitfähigkeit gemessen. Es ergibt sich der Mittelwert $\bar{x} = 17.1$ und der Wert $s^2 = 0.36$. Zu testen ist die Hypothese $\mu = 17$ zu dem Niveau $\alpha = 0.1$. Aus einer Tabelle der t_{14} -Verteilung lesen wir

das 0.95-Quantil $t_{14,0.95} = 1.76$ ab. Aus den gegebenen Daten ergibt sich

$$T(x) = \frac{\sqrt{14}(17.1 - 17)}{\sqrt{0.36}} \approx 0.624.$$

Also kann die Hypothese akzeptiert werden.

XII.4.3. Der χ^2 -Test. Wir betrachten folgendes Testproblem: Es werden n unabhängige, gleichartige Teilerperimente ausgeführt. Jedes hat $r \geq 2$ mögliche Ausgänge, und der i -te Ausgang hat die Wahrscheinlichkeit p_i . Der Parameter $\vartheta = (p_1, \dots, p_r)$ ist unbekannt. Für einen gegebenen Wahrscheinlichkeitsvektor $\pi = (\pi_1, \dots, \pi_r)$ ist zu testen, ob $\vartheta = \pi$ ist.

Die Wahrscheinlichkeit, die Häufigkeiten x_1, \dots, x_r zu beobachten, ist nach der Multinomialverteilung

$$P_{\vartheta}(x) = \binom{n}{x_1, \dots, x_r} p_1^{x_1} \cdot \dots \cdot p_r^{x_r}.$$

Der Vektor $x = (x_1, \dots, x_r)$ muss dabei natürlich der Bedingung $x_1 + \dots + x_r = n$ genügen. Die Likelihood-Funktion ist $L_x(\vartheta) = P_{\vartheta}(x)$. Bei der Ermittlung des Maximums muss die Nebenbedingung $p_1 + \dots + p_r = 1$ berücksichtigt werden. Dies führt auf das Gleichungssystem

$$\frac{\partial}{\partial p_1} \ln L_x(\vartheta) = \frac{\partial}{\partial p_2} \ln L_x(\vartheta) = \dots = \frac{\partial}{\partial p_r} \ln L_x(\vartheta).$$

Als Lösung ergibt sich der Schätzer $\hat{p}_i(x) = \frac{x_i}{n}$. Damit folgt für den Likelihood-Quotienten

$$q(x) = \frac{P_{\hat{\vartheta}}(x)}{P_{\pi}(x)} \approx \left(\frac{\hat{p}_1}{\pi_1} \cdot \dots \cdot \frac{\hat{p}_r}{\pi_r} \right)^{-\frac{1}{2}} \exp \left(\frac{1}{2} \sum_{i=1}^r \frac{(x_i - n\pi_i)^2}{n\pi_i} \right).$$

Wir definieren daher

$$V^2(x) = \sum_{i=1}^r \frac{(x_i - n\pi_i)^2}{n\pi_i}.$$

Gilt die Hypothese $\vartheta = \pi$, so ist nach dem Gesetz der großen Zahlen $\hat{p}_i(x) = \frac{x_i}{n}$ mit Wahrscheinlichkeit nahe 1 gleich π_i . Daher ist in diesem Fall

$$q(x) \approx \exp\left(\frac{1}{2}V^2(x)\right).$$

Man kann zeigen, dass $V^2(x)$ durch die sog. χ^2 -VERTEILUNG mit $r - 1$ Parametern, kurz χ_{n-1}^2 -VERTEILUNG approximiert wird. Diese hat die Dichte

$$g_{r-1}(x) = \begin{cases} \frac{1}{2^{\frac{r-1}{2}} \Gamma(\frac{r-1}{2})} x^{\frac{r-1}{2}-1} e^{-\frac{x}{2}} & \text{falls } x > 0, \\ 0 & \text{sonst.} \end{cases}$$

Damit lautet der χ^2 -TEST ZUM NIVEAU α :

Berechne die Größe

$$V^2(x) = \sum_{i=1}^r \frac{(x_i - n\pi_i)^2}{n\pi_i}.$$

Bestimme aus einer Tabelle der χ_{r-1}^2 -Verteilung das $(1-\alpha)$ -QUANTIL $\chi_{r-1,1-\alpha}^2$ aus der Bedingung

$$\int_0^{\chi_{r-1,1-\alpha}^2} g_{r-1}(x) dx = 1 - \alpha.$$

Falls $V^2(x) < \chi_{r-1,1-\alpha}^2$ ist, wird die Hypothese angenommen, sonst wird sie verworfen.

BEISPIEL XII.4.2. Man vermutet, dass die Blütenfarben rot, rosa und weiß einer Rosenart mit den Wahrscheinlichkeiten $\frac{1}{4}$, $\frac{1}{2}$ und $\frac{1}{4}$ vererbt werden. Unter einer Auswahl von 320 Nachkommen beobachtet man 102 rote, 156 rosa und 62 weiße Blüten. Die Hypothese soll mit dem χ^2 -Test zum Niveau $\alpha = 0.1$ getestet werden. Es ist

$$\pi = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$$

und

$$(\hat{p}_1, \hat{p}_2, \hat{p}_3) = \left(\frac{102}{320}, \frac{156}{320}, \frac{62}{320}\right).$$

Wir erhalten den Wert

$$V^2(x) = \frac{(102 - 80)^2}{80} + \frac{(156 - 160)^2}{160} + \frac{(62 - 80)^2}{80} \\ \approx 10.2.$$

Aus einer Tabelle der χ^2 -Verteilung mit 2 Parametern erhalten wir das 0.9-Quantil

$$\chi_{2,0.9}^2 = 4.61.$$

Daher muss die Hypothese verworfen werden.

BEISPIEL XII.4.3. Es wird behauptet, dass auf einer Pferderennbahn die Startposition einen Einfluss auf die Gewinnwahrscheinlichkeit hat. In 144 Rennen hatten die Sieger die Startposition 1, 2, ..., 8 mit

den Häufigkeiten 29, 19, 18, 25, 17, 10, 15, 11. Wir wollen die Hypothese, dass alle Startpositionen die gleiche Gewinnwahrscheinlichkeit haben, zu dem Niveau $\alpha = 0.05$ testen. Aus einer Tabelle der χ^2 -Verteilung mit 7 Freiheitsgraden lesen wir das 0.95-Quantil

$$\chi_{7,0.95}^2 = 20.28$$

ab. Da $n\pi_i = 16$ ist für alle i , erhalten wir mit den beobachteten Werten

$$\begin{aligned} V^2(x) &= \frac{1}{16} \{(29 - 16)^2 + (19 - 16)^2 + (18 - 16)^2 + (25 - 16)^2 \\ &\quad + (17 - 16)^2 + (10 - 16)^2 + (15 - 16)^2 + (11 - 16)^2\} \\ &\approx 20.375. \end{aligned}$$

Also wird die Hypothese verworfen.

Die χ^2 -Verteilung wird auch für das Testen der Varianz einer Normalverteilung genutzt. Der χ^2 -TEST ZUM NIVEAU α AUF DIE HYPOTHESE EINER $N(\mu, \sigma_0^2)$ -VERTEILTEN STICHPROBE lautet:

Berechne die Größen

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i \\ s(x)^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\ T(x) &= \frac{(n-1)s(x)^2}{\sigma_0^2}. \end{aligned}$$

Bestimme aus einer Tabelle der χ_{n-1}^2 -Verteilung die $(1 - \frac{\alpha}{2})$ - und $\frac{\alpha}{2}$ - QUANTILE $\chi_{n-1, 1-\frac{\alpha}{2}}^2$ und $\chi_{n-1, \frac{\alpha}{2}}^2$ aus den Bedingungen

$$\begin{aligned} \int_{-\infty}^{\chi_{n-1, 1-\frac{\alpha}{2}}^2} g_{n-1}(x) dx &= 1 - \frac{\alpha}{2}, \\ \int_{-\infty}^{\chi_{n-1, \frac{\alpha}{2}}^2} g_{n-1}(x) dx &= \frac{\alpha}{2}. \end{aligned}$$

Falls

$$\chi_{n-1, \frac{\alpha}{2}}^2 \leq T(x) \leq \chi_{n-1, 1-\frac{\alpha}{2}}^2$$

ist, wird die Hypothese angenommen, sonst wird sie verworfen.

BEMERKUNG XII.4.4. Bei dem oben dargestellten Test wird auf $\sigma^2 = \sigma_0^2$ getestet. Falls auf $\sigma^2 < \sigma_0^2$ getestet werden soll, lautet das Testkriterium

$$T(x) \leq \chi_{n-1, 1-\alpha}^2.$$

Falls auf $\sigma^2 > \sigma_0^2$ getestet werden soll, lautet das Testkriterium

$$T(x) \geq \chi_{n-1, \alpha}^2.$$

BEISPIEL XII.4.5. Nach Angaben des Herstellers eines bestimmten PKW-Typs ist der Benzinverbrauch im Stadtverkehr annähernd normalverteilt mit Erwartungswert 9.5 l/100km und Streuung 2.5 l/100km. Zur Überprüfung dieser Angaben testet eine Verbraucherorganisation 25 PKWs und misst einen Durchschnittsverbrauch von 9.9 l/100km mit einer Streuung von 3.5 l/100km. Das Niveau des Tests soll $\alpha = 0.05$ sein.

Für den Erwartungswert ergeben diese Daten

$$\begin{aligned} T(x) &= \frac{\sqrt{25}(9.9 - 9.5)}{3.5} \\ &\approx 0.571 \\ t_{24, 0.975} &= 2.064. \end{aligned}$$

Daher ist die Aussage über den erwarteten Durchschnittsverbrauch zu akzeptieren.

Für die Varianz liefern diese Daten

$$\begin{aligned} T(x) &= \frac{24 \cdot 3.5^2}{2.5^2} \\ &\approx 47.04 \\ \chi_{24, 0.025}^2 &= 12.40 \\ \chi_{24, 0.975}^2 &= 39.36. \end{aligned}$$

Wegen $T(x) > \chi_{24, 0.975}^2$ ist die Aussage über die Streuung der Verbrauchswerte zu verwerfen.

KAPITEL XIII

Fourier-Analysis

XIII.1. Trigonometrische Polynome und Reihen

XIII.1.1. Periodische Funktionen. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ oder $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt PERIODISCH MIT PERIODE T oder kurz T -PERIODISCH, wenn $T > 0$ ist und $f(t + T) = f(t)$ gilt für alle $t \in \mathbb{R}$. Statt 2π -periodisch sagen wir häufig auch kurz PERIODISCH.

BEISPIEL XIII.1.1. Eine konstante Funktion ist T -periodisch für jedes $T > 0$. Die Funktionen $\sin t$, $\cos t$, e^{it} , $\alpha \cos(nt) + \beta \sin(nt)$ sind alle 2π -periodisch.

Es gelten folgende Rechenregel für T -periodische Funktionen:

- Mit T ist auch nT für jedes $n \in \mathbb{N}^*$ eine Periode.
- Sind f, g T -periodisch und $\alpha, \beta \in \mathbb{C}$, so ist auch $\alpha f + \beta g$ T -periodisch.
- Ist f T -periodisch, so gilt

$$\int_a^{a+T} f(t) dt = \int_0^T f(t) dt \quad \text{für jedes } a \in \mathbb{R}.$$

Durch die Substitution $x = \frac{2\pi}{T}t = \omega t$ mit $\omega = \frac{2\pi}{T}$ wird eine T -periodische Funktion f in eine 2π -periodische Funktion transformiert.

Ist g eine gegebene Funktion auf dem Intervall $[0, T]$ kann man sie auf drei Arten zu einer T -periodischen Funktion f auf \mathbb{R} fortsetzen (vgl. Abbildung XIII.1.1):

- DIREKTE FORTSETZUNG (T -PERIODISCH): Setze

$$f(t) = g(t - nT), \text{ falls } nT \leq t < (n + 1)T, n \in \mathbb{Z}.$$

- GERADE FORTSETZUNG ($2T$ -PERIODISCH): Setze

$$\tilde{f}(t) = \begin{cases} g(t) & \text{für } 0 \leq t \leq T \\ g(-t) & \text{für } -T \leq t < 0 \end{cases}$$

und definiere f durch

$$f(t) = \tilde{f}(t - 2nT) \text{ falls } (2n - 1)T \leq t < (2n + 1)T, \\ n \in \mathbb{Z}.$$

- UNGERADE FORTSETZUNG ($2T$ -PERIODISCH): Setze

$$\widehat{f}(t) = \begin{cases} g(t) & \text{für } 0 \leq t \leq T \\ -g(-t) & \text{für } -T \leq t < 0 \end{cases}$$

und definiere f durch

$$f(t) = \widehat{f}(t - 2nT) \text{ falls } (2n - 1)T \leq t < (2n + 1)T, \\ n \in \mathbb{Z}.$$

Ist g stetig, so ist im ersten Fall f genau dann stetig, wenn $g(0) = g(T)$ ist. Im zweiten Fall ist f ohne weitere Zusatzbedingung an g stetig. Im dritten Fall schließlich ist f genau dann stetig, wenn $g(0) = g(T) = 0$ ist.

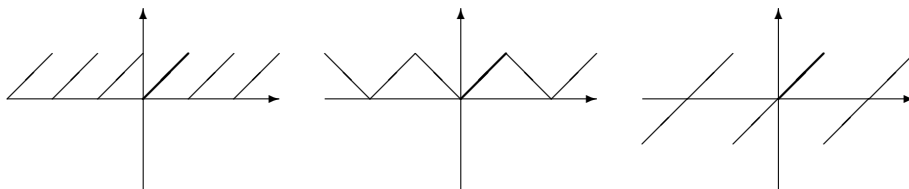


ABBILDUNG XIII.1.1. Periodische Fortsetzungen der Funktion $g(x) = x$ auf $[0, 1]$

XIII.1.2. Trigonometrische Polynome. Eine Funktion der Form

$$f(t) = \sum_{k=-N}^N c_k e^{ik\omega t}$$

mit $\omega \neq 0$, $c_{-N}, \dots, c_N \in \mathbb{C}$ und $|c_{-N}| + |c_N| \neq 0$ nennt man ein TRIGONOMETRISCHES POLYNOM vom GRAD N . Es ist $\frac{2\pi}{\omega}$ -periodisch. Wegen

$$e^{ik\omega t} = \cos(k\omega t) + i \sin(k\omega t)$$

kann jedes trigonometrische Polynom vom Grad N in der Form

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^N [a_k \cos(k\omega t) + b_k \sin(k\omega t)]$$

dargestellt werden. Für die Koeffizienten gelten die Umrechnungsformeln

$$\begin{aligned} c_0 &= \frac{1}{2}a_0, & c_k &= \frac{1}{2}(a_k - ib_k), & c_{-k} &= \frac{1}{2}(a_k + ib_k) \\ a_0 &= 2c_0, & a_k &= c_k + c_{-k}, & b_k &= i(c_k - c_{-k}). \end{aligned}$$

BEISPIEL XIII.1.2. Die Funktion

$$\begin{aligned} f(t) &= \sum_{k=-N}^N e^{ikt} \\ &= 1 + 2 \cos x + \dots + 2 \cos(Nx) \end{aligned}$$

ist ein trigonometrisches Polynom vom Grad N . Aus der geometrischen Summenformel

$$\sum_{k=0}^{n-1} q^k = \begin{cases} \frac{q^n - 1}{q - 1} & \text{falls } q \neq 1 \\ n & \text{falls } q = 1 \end{cases}$$

folgt die Darstellung

$$\sum_{k=-N}^N e^{ikt} = \begin{cases} 2N + 1 & \text{falls } t = 2n\pi, n \in \mathbb{Z}, \\ \frac{\sin((N + \frac{1}{2})t)}{\sin(\frac{1}{2}t)} & \text{sonst.} \end{cases}$$

Sei f ein trigonometrisches Polynom vom Grad N und $T = \frac{2\pi}{\omega}$ seine Periode. Dann gelten folgende Eigenschaften:

- f hat in $[0, T)$ höchstens $2N$ Nullstellen.
- $f(t) = 0$ für alle $t \in \mathbb{R} \iff c_{-N} = \dots = c_N = 0$.
- f ist reellwertig $\iff c_k = \overline{c_{-k}}$ für $k = 0, \dots, N$ (insb. ist $c_0 \in \mathbb{R}$).
- Für $-N \leq k \leq N$ und $0 \leq n \leq N$ ist

$$\begin{aligned} c_k &= \frac{1}{T} \int_0^T f(t) e^{-ik\omega t} dt, \\ a_n &= \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt, \\ b_n &= \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt. \end{aligned}$$

Die vierte Eigenschaft folgt aus der Orthogonalitätsbeziehung (vgl. Beispiel V.2.6 (S. 179))

$$\int_0^T e^{ik\omega t} e^{-i\ell\omega t} dt = \begin{cases} T & \text{falls } k = \ell, \\ 0 & \text{falls } k \neq \ell. \end{cases}$$

XIII.1.3. Trigonometrische Reihen. Einen Ausdruck der Form

$$\sum_{k \in \mathbb{Z}} c_k e^{ik\omega t}$$

nennt man eine TRIGONOMETRISCHE REIHE. Die N -te PARTIALSUMME ist gegeben durch

$$S_N(t) = \sum_{k=-N}^N c_k e^{ik\omega t}.$$

Falls die Folge $(S_N(t))_{N \in \mathbb{N}}$ für jedes $t \in [0, T]$ konvergiert, wird durch ihren Grenzwert eine T -periodische Funktion f definiert. Man sagt dann, f werde durch die trigonometrische Reihe $\sum_{k \in \mathbb{Z}} c_k e^{ik\omega t}$ dargestellt.

BEISPIEL XIII.1.3. Die Partialsummen der trigonometrischen Reihe $\sum_{k \in \mathbb{Z}} e^{ikt}$ sind für kein $t \in \mathbb{R}$ konvergent. Denn gemäß Beispiel XIII.1.2 erhalten wir für $\frac{t}{2\pi} \in \mathbb{Z}$ die Folge $(2N+1)_{N \in \mathbb{N}}$ und für $\frac{t}{2\pi} \notin \mathbb{Z}$ die Folge $(\frac{\sin((N+\frac{1}{2})t)}{\sin(\frac{1}{2}t)})_{N \in \mathbb{N}}$. Beide Folgen sind aber nicht konvergent.

XIII.2. Fourier-Reihen

XIII.2.1. Die Fourier-Reihe einer Funktion. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ sei stückweise stetig und T -periodisch, $T > 0$. Die Zahlen

$$\hat{f}_k = \frac{1}{T} \int_0^T f(t) e^{-ik\omega t} dt, \quad \text{mit } \omega = \frac{2\pi}{T}, k \in \mathbb{Z}$$

heißen die (KOMPLEXEN) FOURIER-KOEFFIZIENTEN von f . Die mit ihnen gebildete trigonometrische Reihe

$$\sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t}$$

heißt die FOURIER-REIHE von f . Wir schreiben

$$f \sim \sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t},$$

um zu betonen, dass die rechts stehende trigonometrische Reihe die Fourier-Reihe der links stehenden Funktion ist. Hierdurch *ist nichts über die Konvergenz der Reihe ausgesagt*.

Wie in Abschnitt XIII.1.2 (S. 478) ist

$$\sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t} = \frac{a_0}{2} + \sum_{n \in \mathbb{N} \setminus \{0\}} [a_n \cos(n\omega t) + b_n \sin(n\omega t)]$$

mit

$$\begin{aligned} a_n &= \hat{f}_n + \hat{f}_{-n} = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt \\ b_n &= i(\hat{f}_n - \hat{f}_{-n}) = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt. \end{aligned}$$

Die Zahlen a_n, b_n heißen die REELLEN FOURIER-KOEFFIZIENTEN von f .

Die Fourier-Koeffizienten einer T -periodischen Schwingung f haben folgende Deutung:

- \widehat{f}_0 : Arithmetischer Mittelwert (Gleichspannungsanteil),
- $2\widehat{f}_1$: Komplexe Amplitude der Grundschwingung,
- $2\widehat{f}_n$: Komplexe Amplituden der n -ten Oberschwingung, $n \geq 2$,
- $\frac{\sum_{n=2}^{\infty} |\widehat{f}_n|^2}{\sum_{n=1}^{\infty} |\widehat{f}_n|^2}$: Oberschwingungsanteil (Klirrfaktor).

BEISPIEL XIII.2.1 (RECHTECKSCHWINGUNG). Definiere $f : [0, 2\pi) \rightarrow \mathbb{R}$ durch

$$f(t) = \begin{cases} 1 & \text{für } 0 \leq t < \pi, \\ -1 & \text{für } \pi \leq t < 2\pi. \end{cases}$$

Dann folgt

$$\begin{aligned} a_n &= \frac{1}{\pi} \left\{ \int_0^{\pi} \cos(nt) dt - \int_{\pi}^{2\pi} \cos(nt) dt \right\} \\ &= 0 \end{aligned}$$

$$\begin{aligned} b_n &= \frac{1}{\pi} \left\{ \int_0^{\pi} \sin(nt) dt - \int_{\pi}^{2\pi} \sin(nt) dt \right\} \\ &= \begin{cases} \frac{4}{n\pi} & \text{falls } n \text{ ungerade} \\ 0 & \text{falls } n \text{ gerade} \end{cases} \end{aligned}$$

Also lautet die Fourier-Reihe

$$\frac{4}{\pi} \sum_{n \in \mathbb{N}} \frac{\sin((2n+1)x)}{2n+1}.$$

XIII.2.2. Rechenregeln. Im Folgenden seien $f, g : \mathbb{R} \rightarrow \mathbb{C}$ stückweise stetige, T -periodische Funktionen mit Fourier-Reihen

$$\begin{aligned} f &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_k e^{ik\omega t}, \\ g &\sim \sum_{k \in \mathbb{Z}} \widehat{g}_k e^{ik\omega t} \end{aligned}$$

und $\omega = \frac{2\pi}{T}$.

Wegen der Linearität des Integrals gilt für $\alpha, \beta \in \mathbb{C}$

$$\begin{aligned} (\widehat{\alpha f + \beta g})_k &= \alpha \widehat{f}_k + \beta \widehat{g}_k \quad \text{für alle } k \in \mathbb{Z} \\ \alpha f + \beta g &\sim \sum_{k \in \mathbb{Z}} (\alpha \widehat{f}_k + \beta \widehat{g}_k) e^{ik\omega t}. \end{aligned}$$

Mit der Linearität der komplexen Konjugation und der Substitution $t \mapsto -t$ folgt

$$\begin{aligned}\overline{f(t)} &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_{-k} e^{ik\omega t} \\ f(-t) &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_{-k} e^{ik\omega t}.\end{aligned}$$

Mit den Substitutionen $t \mapsto ct$, $c > 0$ und $t \mapsto t + a$ ergibt sich

$$\begin{aligned}f(ct) &\sim \sum_{k \in \mathbb{Z}} \widehat{f}_k e^{ik(\omega c)t} \\ f(t+a) &\sim \sum_{k \in \mathbb{Z}} (e^{ik\omega a} \widehat{f}_k) e^{ik\omega t}.\end{aligned}$$

Aus der Definition der Fourier-Koeffizienten folgt

$$e^{i\omega n t} f(t) \sim \sum_{k \in \mathbb{Z}} \widehat{f}_{k-n} e^{ik\omega t}.$$

Mittels partieller Integration ergibt sich

$$f'(t) \sim \sum_{k \in \mathbb{Z}} (ik\omega \widehat{f}_k) e^{ik\omega t}.$$

Das unbestimmte Integral einer T -periodischen Funktion f ist nur dann wieder T -periodisch, wenn der Mittelwert von f verschwindet, d.h. $\widehat{f}_0 = 0$. In diesem Fall gilt:

$$\int_0^t f(s) ds \sim -\frac{1}{T} \int_0^T s f(s) ds + \sum_{k \in \mathbb{Z} \setminus \{0\}} \left(\frac{1}{ik\omega} \widehat{f}_k \right) e^{ik\omega t}.$$

XIII.2.3. Die Bessel-Ungleichung. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ sei stückweise stetig und T -periodisch. Für $N \in \mathbb{N}^*$ bezeichnen wir mit S_N die N -te Partialsumme ihrer Fourier-Reihe

$$S_N(t) = \sum_{k=-N}^N \widehat{f}_k e^{ik\omega t}$$

$$= \frac{a_0}{2} + \sum_{k=1}^N [a_k \cos(k\omega t) + b_k \sin(k\omega t)].$$

Aus der Definition der Fourier-Koeffizienten folgt

$$\begin{aligned} \frac{1}{T} \int_0^T f(t) \overline{S_N(t)} dt &= \sum_{k=-N}^N \widehat{f}_k \underbrace{\frac{1}{T} \int_0^T f(t) e^{-ik\omega t} dt}_{=\widehat{f}_k} \\ &= \sum_{k=-N}^N |\widehat{f}_k|^2. \end{aligned}$$

Aus der Orthogonalitätsbeziehung

$$\frac{1}{T} \int_0^T e^{ik\omega t} e^{-i\ell\omega t} dt = \begin{cases} 1 & \text{falls } k = \ell \\ 0 & \text{falls } k \neq \ell \end{cases}$$

aus Abschnitt XIII.1.2 (S. 478) ergibt sich

$$\begin{aligned} \frac{1}{T} \int_0^T S_N(t) \overline{S_N(t)} dt &= \sum_{k=-N}^N \sum_{\ell=-N}^N \widehat{f}_k \overline{\widehat{f}_\ell} \frac{1}{T} \int_0^T e^{ik\omega t} e^{-i\ell\omega t} dt \\ &= \sum_{k=-N}^N |\widehat{f}_k|^2. \end{aligned}$$

Mit diesen beiden Ergebnissen erhalten wir

$$\begin{aligned} 0 &\leq \frac{1}{T} \int_0^T |f(t) - S_N(t)|^2 dt \\ &= \frac{1}{T} \int_0^T (f(t) - S_N(t)) \overline{(f(t) - S_N(t))} dt \\ &= \frac{1}{T} \int_0^T |f(t)|^2 dt - \frac{1}{T} \int_0^T f(t) \overline{S_N(t)} dt \\ &\quad - \underbrace{\frac{1}{T} \int_0^T \overline{f(t)} S_N(t) dt}_{=\overline{\int_0^T f(t) \overline{S_N(t)} dt}} + \frac{1}{T} \int_0^T S_N(t) \overline{S_N(t)} dt \\ &= \frac{1}{T} \int_0^T |f(t)|^2 dt - \sum_{k=-N}^N |\widehat{f}_k|^2. \end{aligned}$$

Also gilt für alle $N \in \mathbb{N}^*$

$$\sum_{k=-N}^N |\widehat{f}_k|^2 \leq \frac{1}{T} \int_0^T |f(t)|^2 dt.$$

Dies beweist die sog. BESSEL-UNGLEICHUNG

$$\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} [|a_n|^2 + |b_n|^2] = 2 \sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 \leq \frac{2}{T} \int_0^T |f(t)|^2 dt.$$

Da die Koeffizienten einer konvergenten Reihe eine Nullfolge bilden (vgl. Abschnitt VII.1.3 (S. 256)), folgt hieraus das sog. RIEMANN-LEMMA:

$$\begin{aligned} \lim_{k \rightarrow \infty} |\widehat{f}_k| &= 0, & \lim_{k \rightarrow -\infty} |\widehat{f}_k| &= 0, \\ \lim_{n \rightarrow \infty} |a_n| &= 0, & \lim_{n \rightarrow \infty} |b_n| &= 0. \end{aligned}$$

Es gilt sogar folgende Verschärfung:

Die T -periodische Funktion f sei $(m-1)$ -mal differenzierbar auf \mathbb{R} und $f^{(m)}$ sei stückweise stetig auf $[0, T]$. Dann gibt es eine Zahl $M > 0$ mit

$$|\widehat{f}_k| \leq \frac{M}{|k|^{m+1}} \text{ für alle } k \in \mathbb{Z} \setminus \{0\}.$$

XIII.2.4. Konvergenz der Fourier-Reihe. Sei wieder $f : \mathbb{R} \rightarrow \mathbb{C}$ eine stückweise stetige, T -periodische Funktion. Es gilt folgende Verschärfung der Bessel-Ungleichung:

$$\sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 = \frac{1}{T} \int_0^T |f(t)|^2 dt.$$

Dies ist die sog. PARSEVAL-GLEICHUNG.

Aus ihr und dem Beweis der Bessel-Ungleichung folgt, dass die Fourier-Reihe im quadratischen Mittel gegen f konvergiert

$$\lim_{N \rightarrow \infty} \frac{1}{T} \int_0^T |f(t) - S_N(t)|^2 dt = 0.$$

Bezüglich der punktweisen Konvergenz ist die Situation komplizierter (vgl. Beispiele XIII.1.3 (S. 480) und XIII.2.1 (S. 481)). Es gilt:

Die Funktion f sei stückweise differenzierbar auf $[0, T]$. Dann konvergiert die Fourier-Reihe von f für jedes $t \in \mathbb{R}$. Mit $\omega = \frac{2\pi}{T}$ ist

$$\sum_{k=-\infty}^{\infty} \widehat{f}_k e^{ik\omega t} = \frac{1}{2}[f(t+0) + f(t-0)].$$

BEISPIEL XIII.2.2. Betrachte die Funktion $f(t) = t^2$ auf $[-\pi, \pi]$ und setze sie 2π -periodisch fort. Die so auf ganz \mathbb{R} definierte Funktion ist stetig und stückweise differenzierbar. Für die Fourier-Koeffizienten folgt

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \sin(nt) dt \\ &= 0 && \text{wg. Symmetrie} \\ a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \\ &= 2 \frac{\pi^2}{3} \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \cos(nt) dt \\ &= \frac{2}{\pi} \int_0^{\pi} t^2 \cos(nt) dt \\ &= \frac{2}{n\pi} t^2 \sin(nt) \Big|_{t=0}^{t=\pi} - \frac{4}{n\pi} \int_0^{\pi} t \sin(nt) dt \\ &= -\frac{4}{n\pi} \int_0^{\pi} t \sin(nt) dt \\ &= \frac{4}{n^2\pi} t \cos(nt) \Big|_{t=0}^{t=\pi} - \frac{4}{n^2\pi} \int_0^{\pi} \cos(nt) dt \\ &= \frac{4}{n^2} (-1)^{n-1}. \end{aligned}$$

Also gilt für alle $t \in [-\pi, \pi]$

$$t^2 = \frac{\pi^2}{3} - 4 \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\cos(nt)}{n^2}.$$

Wegen

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} (t^2)^2 dt &= \frac{1}{\pi} \int_0^{\pi} t^4 dt \\ &= \frac{1}{5} \pi^4 \end{aligned}$$

folgt aus der Parseval-Gleichung

$$\begin{aligned}
 \sum_{n=1}^{\infty} \frac{1}{n^4} &= \frac{1}{16} \sum_{n=1}^{\infty} |a_n|^2 \\
 &= \frac{1}{16} \left\{ \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} |a_n|^2 - \frac{2}{9} \pi^4 \right\} \\
 &= \frac{2}{16} \sum_{k=-\infty}^{\infty} |\hat{f}_k|^2 - \frac{1}{72} \pi^4 \\
 &= \frac{1}{8} \frac{1}{2\pi} \int_{-\pi}^{\pi} (t^2)^2 dt - \frac{1}{72} \pi^4 \\
 &= \frac{1}{40} \pi^4 - \frac{1}{72} \pi^4 \\
 &= \frac{\pi^4}{90}.
 \end{aligned}$$

XIII.2.5. Anwendung auf gewöhnliche Differentialgleichungen. Wir betrachten das Randwertproblem

$$\begin{aligned}
 a\ddot{x} + b\dot{x} + cx &= f(t) && \text{in } (0, T) \\
 x(0) &= x(T) \\
 \dot{x}(0) &= \dot{x}(T)
 \end{aligned}$$

mit $a, b, c, \in \mathbb{R}$, $a \neq 0$, und einer stückweise stetigen Funktion f . Jede Lösung x dieses Randwertproblems kann T -periodisch festgesetzt werden und liefert eine zweimal differenzierbare T -periodische Funktion auf \mathbb{R} mit stückweise stetiger zweiter Ableitung. Daher gilt für alle t

$$x(t) = \sum_{k=-\infty}^{\infty} \hat{x}_k e^{ik\omega t}$$

mit $\omega = \frac{2\pi}{T}$. Bilden wir die Fourier-Reihen der linken und der rechten Seite der gDgl, erhalten wir aufgrund der Rechenregeln für Fourier-Reihen

$$\begin{aligned}
 f &\sim \sum_{k \in \mathbb{Z}} \hat{f}_k e^{ik\omega t} \\
 a\ddot{x} + b\dot{x} + cx &\sim \sum_{k \in \mathbb{Z}} [a(ik\omega)^2 + b(ik\omega) + c] \hat{x}_k e^{ik\omega t}.
 \end{aligned}$$

Also muss für jede Lösung der gDgl gelten

$$\hat{x}_k = \frac{\hat{f}_k}{a(ik\omega)^2 + b(ik\omega) + c}.$$

Dies bestimmt die Fourier-Reihe der Lösung und damit die Lösung.

XIII.3. Die Fourier-Transformation

XIII.3.1. Definition. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt **FOURIER-TRANSFORMIERBAR**, wenn für jedes $\omega \in \mathbb{R}$ der Grenzwert

$$\int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt = \lim_{R \rightarrow \infty} \int_{-R}^R e^{-i\omega t} f(t) dt$$

existiert. In diesem Fall heißt die Funktion $\mathcal{F}(f) : \mathbb{R} \rightarrow \mathbb{C}$ mit

$$\mathcal{F}(f)(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt$$

die **FOURIER-TRANSFORMIERTE** von f . Statt $\mathcal{F}(f)$ schreibt man gelegentlich auch \hat{f} . Die **INVERSE FOURIER-TRANSFORMATION** einer Funktion $F : \mathbb{R} \rightarrow \mathbb{C}$ ist definiert durch

$$\mathcal{F}^{-1}(F)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} F(\omega) d\omega.$$

Die Fourier-Transformation kann als Grenzübergang $T \rightarrow \infty$ bei Bildung der Fourier-Reihen T -periodischer Funktionen gedeutet werden.

BEISPIEL XIII.3.1. Betrachte den Rechteckimpuls

$$f(t) = \begin{cases} 1 & \text{für } |t| \leq 1, \\ 0 & \text{für } |t| > 1. \end{cases}$$

Für jedes $R > 1$ erhalten wir

$$\begin{aligned} \int_{-R}^R f(t) dt &= \int_{-1}^1 dt \\ &= 2 \end{aligned}$$

und für $\omega \neq 0$

$$\begin{aligned} \int_{-R}^R e^{-i\omega t} f(t) dt &= \int_{-1}^1 e^{-i\omega t} dt \\ &= -\frac{1}{i\omega} e^{-i\omega t} \Big|_{t=-1}^{t=1} \\ &= -\frac{1}{i\omega} (e^{-i\omega} - e^{i\omega}) \\ &= 2 \frac{\sin(\omega)}{\omega}. \end{aligned}$$

Also ist

$$\mathcal{F}(f)(\omega) = \begin{cases} 2 & \text{für } \omega = 0, \\ 2 \frac{\sin(\omega)}{\omega} & \text{für } \omega \neq 0. \end{cases}$$

BEISPIEL XIII.3.2. Sei $a > 0$ und

$$f(t) = \begin{cases} e^{-at} & \text{für } t \geq 0, \\ 0 & \text{für } t < 0. \end{cases}$$

Für $R > 0$ und $\omega \in \mathbb{R}$ erhalten wir

$$\begin{aligned} \int_{-R}^R e^{-i\omega t} f(t) dt &= \int_0^R e^{-(a+i\omega)t} dt \\ &= -\frac{1}{a+i\omega} e^{-(a+i\omega)t} \Big|_{t=0}^{t=R} \\ &= \frac{1}{a+i\omega} (1 - e^{-(a+i\omega)R}) \\ &\xrightarrow{R \rightarrow \infty} \frac{1}{a+i\omega}. \end{aligned}$$

Also ist

$$\mathcal{F}(f)(\omega) = \frac{1}{a+i\omega}, \quad \omega \in \mathbb{R}.$$

BEISPIEL XIII.3.3. Sei wieder $a > 0$ und

$$f(t) = e^{-a|t|}.$$

Dann gilt für $R > 0$

$$\begin{aligned} \int_{-R}^R e^{-i\omega t} f(t) dt &= \int_0^R e^{-(a+i\omega)t} dt + \int_{-R}^0 e^{(a-i\omega)t} dt \\ &= \frac{1}{a+i\omega} (1 - e^{-(a+i\omega)R}) + \frac{1}{a-i\omega} (1 - e^{-(a-i\omega)R}) \\ &\xrightarrow{R \rightarrow \infty} \frac{1}{a+i\omega} + \frac{1}{a-i\omega} \\ &= \frac{2a}{a^2 + \omega^2}. \end{aligned}$$

Also ist

$$\mathcal{F}(f)(\omega) = \frac{2a}{a^2 + \omega^2}, \quad \omega \in \mathbb{R}.$$

XIII.3.2. Rechenregeln. Im Folgenden sind $f, g : \mathbb{R} \rightarrow \mathbb{C}$ stückweise stetige Funktionen mit

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty, \quad \int_{-\infty}^{\infty} |g(t)| dt < \infty.$$

Mit Hilfe der Rechenregeln für Integrale und partieller Integration kann man dann die folgenden Rechenregeln für die Fourier-Transformation beweisen.

$$\mathcal{F}(\alpha f + \beta g) = \alpha \mathcal{F}(f) + \beta \mathcal{F}(g) \quad (\text{Linearität}), \alpha, \beta \in \mathbb{C}$$

$$\mathcal{F}(\bar{f})(\omega) = \overline{\mathcal{F}(f)(-\omega)}$$

$$\mathcal{F}(f(ct))(\omega) = \frac{1}{|c|} \mathcal{F}(f)\left(\frac{\omega}{c}\right) \quad c \neq 0$$

$$\mathcal{F}(f(t-a))(\omega) = e^{-i\omega a} \mathcal{F}(f)(\omega) \quad a \in \mathbb{R}$$

$$\mathcal{F}(e^{i\Omega t} f(t))(\omega) = \mathcal{F}(f)(\omega - \Omega) \quad \Omega \in \mathbb{R}$$

$$\mathcal{F}(f')(\omega) = i\omega \mathcal{F}(f)(\omega)$$

$$\mathcal{F}(tf(t)) = i \frac{d}{d\omega} \mathcal{F}(f).$$

BEISPIEL XIII.3.4. Seien $A \in \mathbb{R}$, $T > 0$ und

$$g(t) = \begin{cases} A & \text{für } |t| \leq \frac{T}{2}, \\ 0 & \text{für } |t| > \frac{T}{2}. \end{cases}$$

Dann ist $g(t) = Af(\frac{2}{T}t)$ mit f aus Beispiel XIII.3.1. Damit folgt aus den Rechenregeln und Beispiel XIII.3.1

$$\mathcal{F}(g)(\omega) = \begin{cases} AT & \text{für } \omega = 0, \\ 2A \frac{\sin(\frac{T}{2}\omega)}{\omega} & \text{für } \omega \neq 0. \end{cases}$$

BEISPIEL XIII.3.5. Sei $a \in \mathbb{R}$, $T > 0$ und

$$g(t) = \begin{cases} 1 & \text{für } |t-a| \leq T, \\ 0 & \text{für } |t-a| > T. \end{cases}$$

Dann ist $g(t) = f(\frac{t-a}{T})$ mit f aus Beispiel XIII.3.1. Damit folgt aus den Rechenregeln und Beispiel XIII.3.1

$$\mathcal{F}(g)(\omega) = \begin{cases} 2T & \text{für } \omega = 0, \\ 2e^{-i\omega a} \frac{\sin(T\omega)}{\omega} & \text{für } \omega \neq 0. \end{cases}$$

BEISPIEL XIII.3.6. Für die Funktion

$$g(t) = \cos(\Omega t) e^{-|a|t}$$

mit $\Omega \in \mathbb{R}$, $a \in \mathbb{R} \setminus \{0\}$ gilt

$$g(t) = \frac{1}{2}(e^{i\Omega t} + e^{-i\Omega t})f(t)$$

mit f aus Beispiel XIII.3.3. Daher ist

$$\mathcal{F}(g)(\omega) = \frac{a}{a^2 + (\omega - \Omega)^2} + \frac{a}{a^2 + (\omega + \Omega)^2}.$$

BEISPIEL XIII.3.7. Für

$$f(t) = e^{-t^2}$$

gilt

$$\begin{aligned} tf(t) &= -\frac{1}{2} \frac{d}{dt}(e^{-t^2}) \\ &= -\frac{1}{2} f'(t). \end{aligned}$$

Damit folgt

$$\begin{aligned} -\frac{i}{2} \omega \mathcal{F}(f)(\omega) &= \mathcal{F}\left(-\frac{1}{2} f'\right)(\omega) \\ &= \mathcal{F}(tf(t))(\omega) \\ &= i \frac{d}{d\omega} \mathcal{F}(f)(\omega). \end{aligned}$$

Also ist $\mathcal{F}(f)$ Lösung der gDgl (in ω !)

$$\frac{dg}{d\omega} = -\frac{\omega}{2} g.$$

Mit der Methode der Trennung der Variablen (vgl. Abschnitt VI.2.1 (S. 221)) folgt

$$\mathcal{F}(f)(\omega) = \mathcal{F}(f)(0) e^{-\frac{\omega^2}{4}}.$$

Weiter ist (vgl. Abschnitt XI.5.1 (S. 440))

$$\mathcal{F}(f)(0) = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}.$$

Daraus ergibt sich insgesamt

$$\mathcal{F}(e^{-t^2})(\omega) = \sqrt{\pi} e^{-\frac{\omega^2}{4}}.$$

XIII.3.3. Existenz- und Eindeutigkeitsätze. Wir nennen eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ absolut integrierbar, wenn sie stückweise stetig ist und das uneigentliche Integral $\int_{-\infty}^{\infty} |f(t)| dt$ existiert.

Es gilt folgendes Kriterium für die Existenz der Fourier-Transformation:

Ist die Funktion f absolut integrierbar, so ist sie Fourier-transformierbar und es gilt die PARSEVAL-GLEICHUNG

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathcal{F}(f)(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |f(t)|^2 dt.$$

Es gilt folgendes Ergebnis für die inverse Fourier-Transformation:

Die Funktion f sei absolut integrierbar und stückweise differenzierbar. Dann gilt für alle $t \in \mathbb{R}$

$$\frac{1}{2}[f(t+0) + f(t-0)] = \frac{1}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^R e^{i\omega t} \mathcal{F}(f)(\omega) d\omega.$$

Ist insbesondere

$$f(t) = \frac{1}{2}[f(t+0) + f(t-0)]$$

für alle $t \in \mathbb{R}$, so gilt

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f.$$

KAPITEL XIV

Partielle Differentialgleichungen

XIV.1. Einführung

XIV.1.1. Beispiele.

BEISPIEL XIV.1.1 (MEMBRAN-, POISSONGLEICHUNG). Sei $\Omega \subset \mathbb{R}^2$ eine offene, beschränkte Menge mit stückweise glattem Rand z.B. ein Kreissegment, das von einer dünnen Membran z.B. dem Trommelfell in seiner Ruhelage eingenommen wird. Auf die Membran wirke eine äußere Kraft f . Diese bewirkt eine Auslenkung $u = u(x) = u(x_1, x_2)$ in vertikaler Richtung. Unter der Annahme, dass die Membran nicht dehnbar und die Auslenkung klein ist, folgt aus dem Prinzip von „Actio = Reactio“, dass die Auslenkung beschrieben wird durch

$$f = -\Delta u = -\frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} \quad \text{in } \Omega.$$

Dies ist eine PARTIELLE DIFFERENTIALGLEICHUNG (kurz PDGL), die sog. MEMBRAN- oder POISSONGLEICHUNG.

Falls die Membran am Rand Γ von Ω eingespannt ist, gilt dort

$$u = 0 \quad \text{auf } \Gamma.$$

Dies ist eine RANDBEDINGUNG, die sog. (HOMOGENE) DIRICHLET-RANDBEDINGUNG.

Falls die Membran am Rand frei gelagert ist, gilt dort

$$\frac{\partial u}{\partial n} = 0 \quad \text{auf } \Gamma.$$

Dabei ist n der nach außen zeigende Einheitsnormalenvektor. Dies ist die sog. (HOMOGENE) NEUMANN-RANDBEDINGUNG.

BEISPIEL XIV.1.2 (PLATTEN-, BIHARMONISCHE GLEICHUNG). Wir ersetzen die Membran aus Beispiel XIV.1.1 durch eine dünne starre Platte und bezeichnen mit u die Auslenkung der Mittelebene. Dann folgt aus den gleichen physikalischen Prinzipien, dass u bestimmt wird durch

$$f = \Delta^2 u = \frac{\partial^4 u}{\partial x_1^4} + 2 \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4} \quad \text{in } \Omega.$$

Dies ist eine pDgl vierter Ordnung, die sog. PLATTEN- oder BIHARMONISCHE GLEICHUNG.

Falls der Rand der Platte fest eingespannt ist, gilt zusätzlich die Randbedingung

$$u = \frac{\partial u}{\partial n} = 0 \quad \text{auf } \Gamma.$$

Ist der Rand dagegen frei gelagert, gilt die Randbedingung

$$u = \Delta u = 0 \quad \text{auf } \Gamma.$$

BEISPIEL XIV.1.3 (GASGLEICHUNG). Wir betrachten die rotationsfreie Strömung eines idealen, kompressiblen Gases. Aus der Rotationsfreiheit folgt für die Geschwindigkeit \mathbf{v} des Gases

$$\mathbf{v} = \nabla u$$

mit einem skalaren Potential u . Aus der Massenerhaltung folgt

$$\operatorname{div}(\rho \mathbf{v}) = 0,$$

wobei $\rho = \rho(\mathbf{v})$ die Dichte des Gases ist. Da das Gas ideal ist, gilt die Zustandsgleichung

$$\rho(\mathbf{v}) = \left[1 - \frac{\gamma - 1}{2} |\mathbf{v}|^2 \right]^{\frac{1}{\gamma - 1}}$$

wobei $\gamma > 1$ der spezifische Wärmekoeffizient ist. Insgesamt erfüllt damit das Potential u die pDgl

$$\operatorname{div} \left[\left(1 - \frac{\gamma - 1}{2} |\nabla u|^2 \right)^{\frac{1}{\gamma - 1}} \nabla u \right] = 0 \quad \text{in } \Omega.$$

Hinzu kommt die Randbedingung

$$u = u_0 \quad \text{auf } \Gamma$$

mit einer gegebenen Funktion u_0 .

BEISPIEL XIV.1.4 (WÄRMELEITUNGSGLEICHUNG). Sei $\Omega \subset \mathbb{R}^3$ eine offene beschränkte Menge mit stückweise glattem Rand, z.B. ein Zylinder. Die Funktion $u(x, t) : \Omega \times [0, \infty] \rightarrow \mathbb{R}$ beschreibe die Temperatur zur Zeit t im Punkt x des Körpers Ω . Wenn dieser einer äußeren

Wärmequelle f ausgesetzt ist, wird der Verlauf der Temperatur durch die pDgl

$$f = \frac{\partial u}{\partial t} - \Delta u = \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} - \frac{\partial^2 u}{\partial x_3^2} \quad \text{in } \Omega \times (0, \infty)$$

beschrieben. Dies ist die sog. WÄRMELEITUNGSGLEICHUNG. Diese Gleichung ist zu ergänzen durch eine Information über die anfängliche Temperaturverteilung, d.h. durch eine ANFANGSBEDINGUNG

$$u(x, 0) = u_0(x) \quad \text{in } \Omega$$

mit einer gegebenen Funktion u_0 .

Wenn der Rand des Körpers künstlich auf einer bestimmten zeitlich nicht notwendig konstanten Temperatur gehalten wird, gilt zusätzlich die Randbedingung

$$u(x, t) = g_D(x, t) \quad \text{auf } \Gamma \times (0, \infty).$$

Ist der Rand des Körpers dagegen isoliert, d.h. findet dort kein Wärmefluss statt, gilt stattdessen die Randbedingung

$$\frac{\partial u}{\partial n}(x, t) = 0 \quad \text{auf } \Gamma \times (0, \infty).$$

Sei nun u eine hinreichend oft differenzierbare Lösung der Wärmeleitungsgleichung zu $f = 0$ und der Randbedingung $u = 0$ auf $\Gamma \times (0, \infty)$. Dann folgt mit dem Integralsatz von Gauß (vgl. Abschnitt IX.5.5 (S. 372))

$$\begin{aligned} 0 &= \int_{\Omega} \left\{ \frac{\partial u}{\partial t} - \Delta u \right\} u dx \\ &= \int_{\Omega} \underbrace{\frac{\partial u}{\partial t} u}_{= \frac{1}{2} \frac{\partial}{\partial t} (u^2)} dx - \int_{\Omega} \underbrace{\Delta u u}_{= \operatorname{div}(u \nabla u) - \nabla u \cdot \nabla u} dx \\ &= \underbrace{\int_{\Omega} \frac{1}{2} \frac{\partial}{\partial t} u^2 dx}_{= \frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} u^2 dx \right)} - \underbrace{\int_{\Omega} \operatorname{div}(u \nabla u) dx}_{= \int_{\Gamma} u \frac{\partial u}{\partial n} ds} + \int_{\Omega} |\nabla u|^2 dx \\ &= \frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} u^2 dx \right) - \int_{\Gamma} \underbrace{u}_{=0} \frac{\partial u}{\partial n} ds + \int_{\Omega} |\nabla u|^2 dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} u^2 dx \right) + \underbrace{\int_{\Omega} |\nabla u|^2 dx}_{\geq 0} \\
&\geq \frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} u^2 dx \right).
\end{aligned}$$

Also ist die „Energie“ $E : (0, \infty) \rightarrow \mathbb{R}$ mit

$$E(t) = \frac{1}{2} \int_{\Omega} u(x, t)^2 dx$$

eine monoton fallende Funktion.

BEISPIEL XIV.1.5 (GRUNDWASSERSTRÖMUNG). Die Funktion $u(x, t)$ beschreibe die räumliche und zeitliche Verteilung einer Flüssigkeit, z.B. Grundwasser in einem porösen Medium $\Omega \subset \mathbb{R}^3$. Dann wird u bestimmt durch die pDgl

$$\frac{\partial u}{\partial t} - \operatorname{div}(D(x, u)\nabla u) + \mathbf{k}(x, u) \cdot \nabla u = f \quad \text{in } \Omega \times (0, \infty).$$

Dabei ist $D(x, z) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}^{3 \times 3}$ die matrixwertige DIFFUSIVITÄT und $\mathbf{k}(x, z) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}^3$ die vektorwertige KONDUKTIVITÄT des Mediums. Die Funktion f beschreibt die Zufuhr (Quellen) bzw. Entnahme (Brunnen) von Flüssigkeit. Die pDgl ist zu ergänzen durch Anfangs- und Randbedingungen ähnlich wie in Beispiel XIV.1.4.

BEISPIEL XIV.1.6 (WELLENGLEICHUNG). Wir betrachten wie in Beispiel XIV.1.1 eine dünne Membran, versetzen sie aber jetzt durch eine zeitlich veränderliche äußere Kraft in Schwingung. Falls die Auslenkung klein ist, tritt keine Dämpfung durch innere Reibung auf. Die Auslenkung $u(x, t)$ am Ort x zur Zeit t wird dann beschrieben durch die pDgl

$$f = \frac{\partial^2 u}{\partial t^2} - \Delta u = \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{x_1^2} - \frac{\partial^2 u}{\partial x_2^2} \quad \text{in } \Omega \times (0, \infty).$$

Dies ist die sog. WELLENGLEICHUNG. Zusätzlich gelten auf $\Gamma \times (0, \infty)$ Randbedingungen wie in Beispiel XIV.1.1, je nachdem ob die Membran eingespannt oder frei gelagert ist. Schließlich muss noch der Anfangszustand des Systems beschrieben werden. Dies geschieht durch die Anfangsbedingungen

$$\begin{aligned}
u(x, 0) &= u_0(x) \quad \text{in } \Omega \\
\frac{\partial u}{\partial t}(x, 0) &= u_1(x) \quad \text{in } \Omega
\end{aligned}$$

mit gegebenen Funktionen u_0 und u_1 .

Betrachte nun speziell den Fall $f = 0$ mit homogener Dirichlet-Randbedingung, d.h. $u(x, t) = 0$ auf $\Gamma \times (0, \infty)$. Sei u eine hinreichend oft differenzierbare Lösung des pDgl. Dann folgt mit dem Gaußschen Integralsatz (vgl. Abschnitt V.5.5 (S. 204))

$$\begin{aligned} 0 &= \int_{\Omega} \left\{ \frac{\partial^2 u}{\partial t^2} - \Delta u \right\} \frac{\partial u}{\partial t} dx \\ &= \int_{\Omega} \underbrace{\frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t}}_{=\frac{1}{2} \frac{d}{dt} \left(\left(\frac{\partial u}{\partial t} \right)^2 \right)} dx - \int_{\Omega} \underbrace{\Delta u \frac{\partial u}{\partial t}}_{=\operatorname{div}(\nabla u \frac{\partial u}{\partial t}) - \nabla u \cdot \nabla \left(\frac{\partial u}{\partial t} \right)} dx \\ &= \underbrace{\int_{\Omega} \frac{1}{2} \frac{\partial}{\partial t} \left(\left(\frac{\partial u}{\partial t} \right)^2 \right) dx}_{=\frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} \left(\frac{\partial u}{\partial t} \right)^2 dx \right)} - \underbrace{\int_{\Omega} \operatorname{div}(\nabla u \frac{\partial u}{\partial t}) dx}_{=\int_{\Gamma} \frac{\partial u}{\partial n} \frac{\partial u}{\partial t} dx} + \underbrace{\int_{\Omega} \nabla u \cdot \nabla \left(\frac{\partial u}{\partial t} \right) dx}_{=\frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} |\nabla u|^2 \right) dx} \\ &= \frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} \left(\frac{\partial u}{\partial t} \right)^2 dx \right) - \underbrace{\int_{\Gamma} \frac{\partial u}{\partial n} \frac{\partial u}{\partial t} dx}_{=0} + \underbrace{\int_{\Omega} \frac{1}{2} \frac{\partial}{\partial t} \left(|\nabla u|^2 \right) dx}_{=\frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} |\nabla u|^2 dx \right)} \\ &= \frac{1}{2} \frac{d}{dt} \left\{ \int_{\Omega} \left(\frac{\partial u}{\partial t} \right)^2 + |\nabla u|^2 dx \right\}. \end{aligned}$$

Also bleibt die „Energie“ $E : (0, \infty) \rightarrow \mathbb{R}$ mit

$$E(t) = \frac{1}{2} \int_{\Omega} \left(\frac{\partial u}{\partial t} \right)^2 + |\nabla u|^2 dx$$

erhalten.

XIV.1.2. Bezeichnungen. Sei $G \subset \mathbb{R}^n$ eine offene Menge und $u : G \rightarrow \mathbb{R}$ eine hinreichend oft differenzierbare Funktion. Für $k \geq 1$ bezeichnen wir zur Abkürzung mit $D^k u$ den „Vektor“ aller partieller Ableitungen der Ordnung k von u :

$$D^k u = \left\{ \frac{\partial^k u}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} : \alpha_1, \dots, \alpha_n \in \mathbb{N}, \alpha_1 + \dots + \alpha_n = k \right\}.$$

Ein DIFFERENTIALOPERATOR m -TER ORDNUNG, $m \geq 1$, hat die Form

$$u \mapsto \mathcal{D}(x, u(x), Du(x), \dots, D^m u(x)) \quad , x \in G.$$

Es heißt QUASILINEAR, wenn er darstellbar ist als

$$\begin{aligned} \mathcal{D}(x, u(x), \dots, D^m u(x)) &= A(x, u(x), \dots, D^m u(x)) \\ &\quad + B(x, u(x), \dots, D^{m-1} u(x)) \end{aligned}$$

mit

$$A(x, u(x), \dots, D^m u(x)) = \sum_{\substack{\alpha_1, \dots, \alpha_n \in \mathbb{N} \\ \alpha_1 + \dots + \alpha_n = m}} a_{\alpha}(x, u(x), \dots, D^{m-1} u(x))$$

$$\frac{\partial^m}{x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u(x).$$

A heißt dann der HAUPTTEIL des Differentialoperators. Der Differentialoperator \mathcal{D} heißt LINEAR, wenn er von der Form ist

$$\mathcal{D}(x, u(x), \dots, D^m u(x)) = \sum_{0 \leq \alpha_1 + \dots + \alpha_n \leq m} a_\alpha(x) \frac{\partial^{\alpha_1 + \dots + \alpha_n}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u(x).$$

Die Funktionen a_α heißen dann die Koeffizienten des Differentialoperators. Der Differentialoperator heißt LINEAR MIT KONSTANTEN Koeffizienten, wenn er linear ist und die Koeffizienten a_α konstante Funktionen sind. Ist \mathcal{D} ein linearer Differentialoperator, so ist die Zuordnung $u \mapsto \mathcal{D}(u)$ eine lineare Abbildung.

Eine PARTIELLE DIFFERENTIALGLEICHUNG m -TER ORDNUNG, kurz PDGL m -TER ORDNUNG, ist eine Gleichung der Form

$$\mathcal{D}(u) = f \quad \text{in } G$$

mit einem Differentialoperator \mathcal{D} m -ter Ordnung und einer gegebenen Funktion $f : \Omega \rightarrow \mathbb{R}$. Eine pDgl heißt quasilinear bzw. linear bzw. linear mit konstanten Koeffizienten, wenn der Differentialoperator quasilinear bzw. linear bzw. linear mit konstanten Koeffizienten ist.

BEISPIEL XIV.1.7. Die pDglen der Beispiele XIV.1.1, XIV.1.3, XIV.1.4, XIV.1.5 und XIV.1.6 sind von zweiter Ordnung; diejenige von Beispiel XIV.1.2 ist von vierter Ordnung. Die pDglen der Beispiele XIV.1.1, XIV.1.2, XIV.1.4 und XIV.1.6 sind linear; diejenigen der Beispiele XIV.1.3 und XIV.1.5 sind quasilinear.

BEISPIEL XIV.1.8. Eine quasilineare pDgl erster Ordnung hat die allgemeine Form

$$a_1(x, u) \frac{\partial u}{\partial x_1} + \dots + a_n(x, u) \frac{\partial u}{\partial x_n} + a_0(x, u) = f \quad \text{in } G$$

mit gegebenen Funktionen $a_0, a_1, \dots, a_n : G \times \mathbb{R} \rightarrow \mathbb{R}$ und $f : G \rightarrow \mathbb{R}$.

XIV.1.3. Lineare partielle Differentialgleichungen zweiter Ordnung. Eine lineare pDgl zweiter Ordnung hat die allgemeine Form

$$\sum_{i=1}^n \sum_{j=1}^n A_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{k=1}^n a_k(x) \frac{\partial u}{\partial x_k} + \alpha(x) u = f \quad \text{in } G.$$

Da es bei den zweiten Ableitungen einer zweimal stetig differenzierbaren Funktion nicht auf die Reihenfolge ankommt, kann man stets voraussetzen, dass $A_{ij}(x) = A_{ji}(x)$ ist für alle $x \in G$. Die Matrix $A(x) = (A_{ij}(x))_{1 \leq i, j \leq n}$ ist also für alle x symmetrisch und besitzt daher lauter reelle Eigenwerte (die von x abhängen!) (vgl. Abschnitt II.4.8 (S. 91)). Je nach Vorzeichen dieser Eigenwerte werden pDglen zweiter Ordnung in drei Typen eingeteilt:

- ELLIPTISCH: die Eigenwerte sind für alle $x \in G$ alle strikt positiv,
- PARABOLISCH: für alle $x \in G$ sind $n - 1$ Eigenwerte strikt positiv und ein Eigenwert gleich Null.
- HYPERBOLISCH: für alle $x \in G$ sind $n - 1$ Eigenwerte strikt positiv und ein Eigenwert strikt negativ.

BEMERKUNG XIV.1.9. Sind die Eigenwerte von $A(x)$ für alle $x \in G$ allesamt strikt negativ, erhält man durch Multiplikation der pDgl mit -1 eine elliptische pDgl, die die gleiche Lösungsmenge hat wie die ursprüngliche pDgl. Analog geht man in den Fällen

- für alle $x \in G$ sind $n - 1$ Eigenwerte strikt negativ und ein Eigenwert gleich Null
- für alle $x \in G$ sind $n - 1$ Eigenwerte strikt negativ und ein Eigenwert strikt positiv

vor und erhält eine parabolische bzw. hyperbolische pDgl mit der gleichen Lösungsmenge wie die ursprüngliche pDgl.

Die Poisson-Gleichung aus Beispiel XIV.1.1 ist elliptisch. Die Wärmeleitungsgleichung aus Beispiel XIV.1.4 ist parabolisch. Die Wellengleichung aus Beispiel XIV.1.6 ist hyperbolisch. Physikalisch beschreiben diese drei Typen verschiedene Phänomene:

- ELLIPTISCH: Minimierung einer Energie,
- PARABOLISCH: Dissipation, d.h. zeitliche Abnahme einer Energie,
- HYPERBOLISCH: Energieerhaltung.

Durch eine geeignete Variablentransformation kann man bei parabolischen und hyperbolischen pDglen stets erreichen, dass der Eigenvektor zum nicht positiven Eigenwert durch $(0, \dots, 0, 1)^T$ gegeben ist. Dementsprechend identifiziert man dann die letzte Komponente der transformierten Variablen mit der Zeit t und schreibt G in der Form $\Omega \times (0, \infty)$ mit $\Omega \subset \mathbb{R}^{n-1}$.

XIV.1.4. Anfangs- und Randbedingungen. Wie in den Beispielen des ersten Abschnittes müssen bei pDglen zusätzliche Anfangs- und Randbedingungen gestellt werden. Für Gleichungen zweiter Ordnung gilt hierfür folgendes Schema:

- ELLIPTISCH: eine Randbedingung auf Γ .
- PARABOLISCH: eine Randbedingung auf $\Gamma \times (0, \infty)$ und eine Anfangsbedingung zur Zeit $t = 0$.
- HYPERBOLISCH: eine Randbedingung auf $\Gamma \times (0, \infty)$ und zwei Anfangsbedingungen zur Zeit $t = 0$.

Die wichtigsten Randbedingungen sind:

- DIRICHLET: $u = g_D$ auf Ω bzw. auf $\Omega \times (0, \infty)$,

- NEUMANN: $\sum_{i=1}^N \sum_{j=1}^N n_i A_{ij} \frac{\partial u}{\partial x_j} = g_N$ auf Ω mit $N = n$ bzw. auf $\Omega \times (0, \infty)$ mit $\Omega \subset \mathbb{R}^{n-1}$ und $N = n - 1$.

Dabei sind g_D und g_N wie f vorgegebene Funktionen. Dirichlet- und Neumann-Bedingungen können auch in der Form kombiniert werden, dass Γ in zwei disjunkte Stücke Γ_D und Γ_N zerfällt, auf denen Dirichlet- bzw. Neumann-Bedingungen gestellt werden. Für Gleichungen zweiter Ordnung dürfen die beiden Bedingungen aber nicht gleichzeitig auf demselben Randstück gefordert werden.

BEISPIEL XIV.1.10. Bei einer ringförmigen Membran, die am inneren Rand eingespannt und am äußeren Rand frei gelagert ist, ist Ω von der Form

$$\Omega = \{x \in \mathbb{R}^2 : R_i^2 < x_1^2 + x_2^2 < R_a^2\}$$

und

$$\Gamma = \Gamma_i \cup \Gamma_a$$

mit

$$\Gamma_{i/a} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = R_{i/a}^2\}$$

Dann gilt auf Γ_i die Dirichlet-Bedingung $u = 0$ und auf Γ_a die Neumann-Bedingung $\frac{\partial u}{\partial n} = 0$.

XIV.2. Die Wärmeleitungsgleichung

XIV.2.1. Vorbemerkungen. Sei $n \geq 1$ und $\Omega \subset \mathbb{R}^n$ eine offene, beschränkte Menge. Falls $n = 1$ ist, soll Ω ein Intervall (a, b) sein. Der Rand Γ besteht dann aus den Punkten a und b . Falls $n \geq 2$ ist, soll der Rand Γ von Ω stückweise glatt sein.

Wir betrachten im Folgenden die Wärmeleitungsgleichung

$$\begin{array}{ll} \frac{\partial u}{\partial t} - \Delta u = f & \text{in } \Omega \times (0, \infty) \\ u = g & \text{auf } \Gamma \times (0, \infty) \\ u(x, 0) = u_0 & \text{in } \Omega \end{array}$$

mit bekannten, hinreichend oft differenzierbaren Funktionen f , g und u_0 . Für alle $x \in \Gamma$ muss zudem die KOMPABILITÄTSBEDINGUNG

$$g(x, 0) = u_0(x)$$

erfüllt sein.

Da der Differentialoperator

$$u \mapsto \frac{\partial u}{\partial t} - \Delta u$$

linear ist, können wir die Lösung der Wärmeleitungsgleichung durch SUPERPOSITION aus Lösungen einfacherer Teilprobleme aufbauen. Genauer machen wir den Ansatz

$$u = v + w + G.$$

Dabei ist $G : \Omega \times [0, \infty) \rightarrow \mathbb{R}$ eine Funktion mit

$$\begin{aligned} G(x, t) &= g(x, t) && \text{für alle } x \in \Gamma, t > 0 \\ G(x, 0) &= 0 && \text{für alle } x \in \Omega; \end{aligned}$$

v soll eine Lösung der Wärmeleitungsgleichung

$$\begin{aligned} \frac{\partial v}{\partial t} - \Delta v &= 0 && \text{in } \Omega \times (0, \infty) \\ v &= 0 && \text{auf } \Gamma \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } \Omega \end{aligned}$$

sein; und w soll eine Lösung der Wärmeleitungsgleichung

$$\begin{aligned} \frac{\partial w}{\partial t} - \Delta w &= F && \text{in } \Omega \times (0, \infty) \\ w &= 0 && \text{auf } \Gamma \times (0, \infty) \\ w(x, 0) &= 0 && \text{in } \Omega \end{aligned}$$

mit

$$F = f - \frac{\partial G}{\partial t} + \Delta G$$

sein. Wie man leicht nachrechnet löst $u = v + w + G$ dann die ursprüngliche pDgl. Die Funktion G heißt eine FORTSETZUNG DER RANDWERTE. Die pDgl für v nennt man eine HOMOGENE WÄRMELEITUNGSGLEICHUNG MIT HOMOGENEN RANDBEDINGUNGEN, diejenige für w eine INHOMOGENE WÄRMELEITUNGSGLEICHUNG MIT HOMOGENEN RAND- UND ANFANGSBEDINGUNGEN.

In den folgenden drei Abschnitten erläutern wir detailliert, wie man im Fall $n = 1$, d.h. eine Raumdimension, die Funktionen G , v und w bestimmt. Da sich die pDgl unter einer Translation $x \mapsto x - a$ nicht ändert, können wir dabei stets annehmen, dass $\Omega = (0, L)$ ist. Im letzten Abschnitt gehen wir dann auf die Modifikationen ein, die im Fall höherer Raumdimension $n \geq 2$ erforderlich sind.

XIV.2.2. Fortsetzung der Randwerte. Diese ist im Fall $n = 1$, $\Gamma = \{0, L\}$ besonders einfach. Wir setzen

$$G(x, t) = \frac{L-x}{L}g(0, t) + \frac{x}{L}g(L, t) \quad 0 \leq x \leq L, t \geq 0.$$

Dann ist

$$F(x, t) = f(x, t) - \frac{L-x}{L} \frac{\partial g}{\partial t}(0, t) - \frac{x}{L} \frac{\partial g}{\partial t}(L, t) \quad 0 \leq x \leq L, t \geq 0.$$

In dem wichtigen Spezialfall, dass die Randbedingung g nicht von t abhängt, gilt daher $F = f$.

XIV.2.3. Lösung der homogenen Gleichung mit homogenen Randbedingungen. Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L). \end{aligned}$$

Zu ihrer Lösung machen wir den SEPARATIONSANSATZ

$$v(x, t) = X(x)T(t).$$

Im Folgenden bezeichnen wir mit $'$ die Ableitung nach der Ortvariablen x und mit $\dot{}$ die Ableitung nach der Zeitvariablen t . Dann muss gelten

$$\begin{aligned} 0 &= \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} \\ &= \dot{T}(t)X(x) - T(t)X''(x) \quad \text{für alle } 0 < x < L, t > 0. \end{aligned}$$

Für alle x, t mit $X(x)T(t) \neq 0$ können wir dann diese Gleichung durch $X(x)T(t)$ dividieren und erhalten nach Umsortieren

$$\frac{X''(x)}{X(x)} = \frac{\dot{T}(t)}{T(t)}.$$

Da die linke Seite dieser Gleichung nur von x , die rechte Seite aber nur von t abhängt, müssen beide Ausdrücke konstant sein. Es muss also ein $\lambda \in \mathbb{R}$ geben mit

$$\frac{X''(x)}{X(x)} = \lambda = \frac{\dot{T}(t)}{T(t)}$$

d.h.

$$\begin{aligned} X''(x) &= \lambda X(x), \\ \dot{T}(t) &= \lambda T(t). \end{aligned}$$

Wir müssen nun die drei Fälle $\lambda > 0$, $\lambda = 0$ und $\lambda < 0$ unterscheiden.

FALL $\lambda > 0$: Gemäß Abschnitt VI.3.2 (S. 233) lautet die allgemeine Lösung der gDgl für X in diesem Fall

$$X(x) = c_1 e^{-\sqrt{\lambda}x} + c_2 e^{\sqrt{\lambda}x}.$$

FALL $\lambda = 0$: In diesem Fall lautet gemäß Abschnitt VI.3.2 (S. 233) die allgemeine Lösung für X

$$X(x) = c_1 + c_2 x.$$

FALL $\lambda < 0$: Gemäß Abschnitt VI.3.2 (S. 233) ergibt sich in diesem Fall für x die allgemeine Lösung

$$X(x) = c_1 \sin(\sqrt{|\lambda|x}) + c_2 \cos(\sqrt{|\lambda|x}).$$

Aus der Randbedingung

$$v(x, t) = 0 \quad \text{für } x \in \{0, L\}, t > 0$$

ergibt sich die Bedingung

$$X(x)T(t) = 0 \quad \text{für } x \in \{0, L\}, t > 0.$$

Diese Gleichung lässt zwei Lösungen zu: $T(t) = 0$ für alle t oder

$$X(0) = X(L) = 0.$$

Die erste liefert offensichtlich die triviale Lösung $v(x, t) = 0$ für alle x, t . Daher betrachten wir sinnvollerweise nur die zweite Bedingung für X . Diese führt in allen drei Fällen auf ein homogenes lineares Gleichungssystem für die Konstanten c_1 und c_2 , das eine vom Nullvektor verschiedene Lösung zulassen soll. Also muss die Determinante der zugehörigen Matrix verschwinden. In den drei Fällen erhalten wir folgende Determinanten:

FALL $\lambda > 0$:

$$\begin{aligned} \det \begin{pmatrix} e^{-\sqrt{\lambda}0} & e^{\sqrt{\lambda}0} \\ e^{-\sqrt{\lambda}L} & e^{\sqrt{\lambda}L} \end{pmatrix} &= \det \begin{pmatrix} 1 & 1 \\ e^{-\sqrt{\lambda}L} & e^{\sqrt{\lambda}L} \end{pmatrix} \\ &= e^{\sqrt{\lambda}L} - e^{-\sqrt{\lambda}L} \\ &= 2 \sinh(\sqrt{\lambda}L). \end{aligned}$$

FALL $\lambda = 0$:

$$\det \begin{pmatrix} 1 & 0 \\ 1 & L \end{pmatrix} = L$$

FALL $\lambda < 0$:

$$\begin{aligned} \det \begin{pmatrix} \sin(\sqrt{|\lambda|}0) & \cos(\sqrt{|\lambda|}0) \\ \sin(\sqrt{|\lambda|}L) & \cos(\sqrt{|\lambda|}L) \end{pmatrix} &= \det \begin{pmatrix} 0 & 1 \\ \sin(\sqrt{|\lambda|}L) & \cos(\sqrt{|\lambda|}L) \end{pmatrix} \\ &= -\sin(\sqrt{|\lambda|}L). \end{aligned}$$

Im ersten und zweiten Fall verschwindet die Determinante für keinen Wert von λ . (Beachte: $\sinh(z)$ hat $z = 0$ als einzige Nullstelle!) Im dritten Fall verschwindet die Determinante genau dann, wenn $\sqrt{|\lambda|}L$ eine Nullstelle des Sinus ist, d.h.

$$\sqrt{|\lambda|}L = k\pi \quad , k \in \mathbb{Z} \setminus \{0\}$$

bzw. äquivalent

$$\lambda = - \left(\frac{k\pi}{L} \right)^2 \quad , k \in \mathbb{Z} \setminus \{0\}.$$

Für gegebenes λ hat die gDgl für T gemäß Abschnitt [VI.2.1](#) (S. [221](#)) die allgemeine Lösung

$$T(t) = ce^{\lambda t}.$$

Insgesamt erhalten wir also mit unserem Ansatz die Funktionen

$$v_k(x, t) = e^{-(\frac{k\pi}{L})^2 t} \sin\left(\frac{k\pi}{L}x\right) \quad , k \in \mathbb{Z} \setminus \{0\}$$

als mögliche Lösungen der pDgl. Da die Funktionen v_k und v_{-k} sich für $k \in \mathbb{N}^*$ nur um ein Vorzeichen unterscheiden lautet die allgemeine Lösung

$$v(x, t) = \sum_{k \in \mathbb{N}^*} c_k e^{-(\frac{k\pi}{L})^2 t} \sin\left(\frac{k\pi}{L}x\right).$$

Jede dieser Funktionen erfüllt konstruktionsgemäß die pDgl und die Randbedingungen. Wir müssen also nur noch die Anfangsbedingung erfüllen. Einsetzen von $t = 0$ ergibt die Bedingung

$$u_0(x) = \sum_{k \in \mathbb{N}^*} c_k \sin\left(\frac{k\pi}{L}x\right).$$

Da nach Voraussetzung

$$u_0(0) = u_0(L) = 0$$

ist (vgl. Abschnitt [XIV.2.1](#) (S. [500](#))) können wir u_0 ungerade zu einer stetigen $2L$ -periodischen Funktion auf \mathbb{R} fortsetzen (vgl. Abschnitt [XIII.1.1](#) (S. [477](#))). Falls u_0 zusätzlich differenzierbar ist, stimmt u_0 mit seiner Fourier-Reihe überein, und die c_k sind die entsprechenden

Fourier-Koeffizienten von u_0 (vgl. Abschnitt XIII.2.1 (S. 480) und XIII.2.4 (S. 484)). Insgesamt erhalten wir damit:

Die Funktion u_0 sei differenzierbar und erfülle

$$u_0(0) = u_0(L) = 0$$

Dann ist die Lösung v der pDgl.

$$\begin{aligned} \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L) \end{aligned}$$

gegeben durch

$$v(x, t) = \sum_{k=1}^{\infty} c_k e^{-(\frac{k\pi}{L})^2 t} \sin\left(\frac{k\pi}{L} x\right)$$

mit

$$c_k = \frac{2}{L} \int_0^L u_0(x) \sin\left(\frac{k\pi}{L} x\right) dx.$$

XIV.2.4. Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen. Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\ w &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ w(x, 0) &= 0 && \text{in } (0, L). \end{aligned}$$

Wegen des vorigen Abschnitts setzen wir die Funktion F für jedes t ungerade bzgl. x $2L$ -periodisch fort und entwickeln sie in eine Fourier-Reihe bzgl. x :

$$F(x, t) \sim \sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin\left(\frac{k\pi}{L} x\right)$$

mit

$$\widehat{F}_k(t) = \frac{2}{L} \int_0^L F(x, t) \sin\left(\frac{k\pi}{L} x\right) dx.$$

Für w machen wir den Ansatz

$$w(x, t) \sim \sum_{k \in \mathbb{N}^*} w_k(t) \sin\left(\frac{k\pi}{L}x\right).$$

Einsetzen in die pDgl liefert

$$\begin{aligned} \sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin\left(\frac{k\pi}{L}x\right) &\sim F(x, t) \\ &= \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} \\ &\sim \sum_{k \in \mathbb{N}^*} [\dot{w}_k(t) + \left(\frac{k\pi}{L}\right)^2 w_k(t)] \sin\left(\frac{k\pi}{L}x\right). \end{aligned}$$

Also ergibt sich für die Fourier-Koeffizienten w_k von w die gDgl

$$\dot{w}_k(t) = -\left(\frac{k\pi}{L}\right)^2 w_k(t) + \widehat{F}_k(t) \quad , k \in \mathbb{N}^*.$$

Wegen der Anfangsbedingung $w(x, 0) = 0$ für $x \in (0, L)$ muss gelten

$$w_k(0) = 0 \quad , k \in \mathbb{N}^*.$$

Damit ergibt sich mit der Methode der Variation der Konstanten (vgl. Abschnitt VI.2.2 (S. 225))

$$w_k(t) = \int_0^t \exp\left(-\left(\frac{k\pi}{L}\right)^2(t-s)\right) \widehat{F}_k(s) ds \quad , k \in \mathbb{N}^*.$$

Falls F stetig ist, kann man zeigen, dass die Fourier-Reihe für w konvergiert. Damit folgt insgesamt:

Die Funktion F sei stetig. Dann ist die Lösung w der pDgl

$$\begin{aligned} \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\ w &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ w(x, 0) &= 0 && \text{in } (0, L). \end{aligned}$$

gegeben durch

$$w(x, t) = \sum_{k=1}^{\infty} w_k(t) \sin\left(\frac{k\pi}{L}x\right)$$

mit

$$w_k(t) = \int_0^t \exp\left(-\left(\frac{k\pi}{L}\right)^2(t-s)\right) \widehat{F}_k(s) ds$$

und

$$\widehat{F}_k(t) = \frac{2}{L} \int_0^L F(x, t) \sin\left(\frac{k\pi}{L}x\right) dx.$$

XIV.2.5. Integraldarstellung. Für $x, \eta \in (0, L)$ und $t \in (0, \infty)$ definieren wir die Funktion $G(x, t, \eta)$ durch

$$G(x, t, \eta) = \frac{2}{L} \sum_{k=1}^{\infty} \exp\left(-\left(\frac{k\pi}{L}\right)^2 t\right) \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{k\pi}{L}\eta\right).$$

G heißt WÄRMELEITUNGSKERN. Mit seiner Hilfe kann man die Ergebnisse der vorigen beiden Abschnitte wie folgt zusammenfassen:

Die Funktion $u_0 : (0, L) \rightarrow \mathbb{R}$ sei differenzierbar und erfülle

$$u_0(0) = u_0(L) = 0.$$

Die Funktion $F : (0, L) \times (0, \infty) \rightarrow \mathbb{R}$ sei stetig. Dann ist die Lösung u der pDgl

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty), \\ u &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ u(x, 0) &= u_0 && \text{in } (0, L) \end{aligned}$$

gegeben durch

$$\begin{aligned} u(x, t) &= \int_0^L G(x, t, \eta) u_0(\eta) d\eta \\ &+ \int_0^L \left\{ \int_0^t G(x, t-s, \eta) F(\eta, s) ds \right\} d\eta. \end{aligned}$$

XIV.2.6. Höhere Raumdimensionen. Im Fall $n \geq 2$ gehen wir ähnlich wie in den Abschnitten [XIV.2.3](#) (S. 502) und [XIV.2.4](#) (S. 505) vor. Für die Funktion v machen wir wieder den Separationsansatz

$$v(x, t) = X(x)T(t).$$

Der einzige Unterschied ist, dass X nun eine Funktion in n Veränderlichen ist. Der Ansatz führt auf die Gleichungen

$$\begin{aligned} \dot{T}(t) &= \lambda T(t), \\ \Delta X &= \lambda X. \end{aligned}$$

Damit erhalten wir nun eine pDgl für X . Man nennt sie das EIGENWERTPROBLEM FÜR DEN LAPLACE-OPERATOR. Wir werden dieses

für spezielle Gebiete $\Omega \subset \mathbb{R}^2$ in den Abschnitten [XIV.4.1](#) (S. 514) und [XIV.4.3](#) (S. 518) lösen. Allgemein kann man zeigen, dass jeder Eigenwert strikt negativ ist, dass insgesamt abzählbar viele Eigenwerte $\lambda_k = -\omega_k^2, k \in \mathbb{N}$, existieren und dass die zugehörigen Eigenfunktionen φ_k paarweise orthogonal sind, d.h.

$$\int_{\Omega} \varphi_k \varphi_\ell = 0$$

für $k \neq \ell$. Die „Fourier-Koeffizienten“

$$\int_{\Omega} u_0 \varphi_k dx \quad \text{und} \quad \int_{\Omega} F(x, t) \varphi_k(x) dx$$

übernehmen dann die Rolle der Fourier-Koeffizienten c_k und $F_k(t)$ in den vorigen Abschnitten. Mit diesen Modifikationen übertragen sich dann die Ergebnisse der vorigen Abschnitte auf den Fall höherer Raumdimensionen.

XIV.3. Die Wellengleichung

XIV.3.1. Vorbemerkung. Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \Delta u &= f && \text{in } \Omega \times (0, \infty) \\ u &= g && \text{auf } \Gamma \times (0, \infty) \\ u(x, 0) &= u_0 && \text{in } \Omega \\ \frac{\partial u}{\partial t}(x, 0) &= u_1 && \text{in } \Omega. \end{aligned}$$

Dabei ist wie im vorigen Abschnitt Ω das Intervall $(0, L)$ mit Rand $\Gamma = \{0, L\}$ oder eine beschränkte, offene Menge in $\mathbb{R}^n, n \geq 2$, mit stückweise glattem Rand. Für alle $x \in \Gamma$ müssen zudem die KOMPABILITÄTSBEDINGUNGEN

$$\begin{aligned} g(x, 0) &= u_0(x) \\ \frac{\partial g}{\partial t}(x, 0) &= u_1(x) \end{aligned}$$

erfüllt sein.

Wir gehen wie im vorigen Paragraphen vor und bestimmen die Lösung u durch die SUPERPOSITION

$$u = v + w + G$$

mit einer Fortsetzung G der Randdaten, der Lösung v der homogenen Gleichung mit homogenen Randbedingungen, d.h. $f = 0$, $g = 0$, und der Lösung w der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen, d.h. $g = 0$, $u_0 = 0$, $u_1 = 0$. Wieder betrachten wir detailliert den Fall einer Raumdimension und erläutern kurz die Modifikationen bei höheren Raumdimensionen. Die Fortsetzung G der Randwerte g wird wie in Abschnitt XIV.2.2 (S. 502) bestimmt. Die Funktion F ist jetzt gegeben durch

$$F = f - \frac{\partial^2 G}{\partial t^2} + \Delta G.$$

XIV.3.2. Lösung der homogenen Gleichung mit homogenen Randbedingungen. Wir betrachten die pDgl

$$\begin{aligned} \frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L) \\ \frac{\partial v}{\partial t}(x, 0) &= u_1 && \text{in } (0, L). \end{aligned}$$

Wie in Abschnitt XIV.2.3 (S. 502) machen wir den Separationsansatz

$$v(x, t) = X(x)T(t).$$

Dieser führt jetzt auf die gDglen

$$\begin{aligned} X''(x) &= \lambda X(x) \\ \ddot{T}(t) &= \lambda T(t) \end{aligned}$$

mit $\lambda \in \mathbb{R}$. Wie in Abschnitt XIV.2.3 (S. 502) folgt aus den Randbedingungen, dass $X(0) = X(L) = 0$ sein muss und dass daher

$$\lambda = - \left(\frac{k\pi}{L} \right)^2, \quad k \in \mathbb{N}^*$$

und

$$X(x) = \sin\left(\frac{k\pi}{L}x\right)$$

gelten muss. Für die Funktion T ergibt sich gemäß Abschnitt VI.3.2 (S. 233) jetzt die allgemeine Lösung

$$T(t) = c_k \cos\left(\frac{k\pi}{L}t\right) + s_k \sin\left(\frac{k\pi}{L}t\right).$$

Daher hat die pDgl die allgemeine Lösung

$$v(x, t) = \sum_{k \in \mathbb{N}^*} \left[c_k \cos\left(\frac{k\pi}{L}t\right) + s_k \sin\left(\frac{k\pi}{L}t\right) \right] \sin\left(\frac{k\pi}{L}x\right).$$

Wir müssen noch die Anfangsbedingungen anpassen. Einsetzen von $t = 0$ liefert die Bedingung

$$\begin{aligned} \sum_{k \in \mathbb{N}^*} c_k \sin\left(\frac{k\pi}{L}x\right) &= v(x, 0) \\ &= u_0(x). \end{aligned}$$

Hieraus können die Koeffizienten c_k wie in Abschnitt XIV.2.3 (S. 502) bestimmt werden. Ableiten nach t und Einsetzen von $t = 0$ liefert die Bedingung

$$\begin{aligned} \sum_{k \in \mathbb{N}^*} s_k \frac{k\pi}{L} \sin\left(\frac{k\pi}{L}x\right) &= \frac{\partial v}{\partial t}(x, 0) \\ &= u_1(x). \end{aligned}$$

Zur Bestimmung der s_k setzen wir wieder u_1 ungerade zu einer $2L$ -periodischen Funktion fort (vgl. Abschnitt XIII.1.1 (S. 477)) und erhalten gemäß Abschnitt XIII.2.1 (S. 480)

$$u_1(x) \sim \sum_{k \in \mathbb{N}^*} \gamma_k \sin\left(\frac{k\pi}{L}x\right)$$

mit

$$\gamma_k = \frac{2}{L} \int_0^L u_1(x) \sin\left(\frac{k\pi}{L}x\right) dx.$$

Ein Vergleich der Fourier-Koeffizienten liefert dann

$$s_k = \frac{L}{k\pi} \gamma_k, \quad k \in \mathbb{N}^*.$$

Insgesamt erhalten wir

Die Funktionen u_0 und u_1 seien differenzierbar und mögen die Bedingungen

$$u_0(0) = u_0(L) = 0$$

$$u_1(0) = u_1(L) = 0$$

erfüllen. Dann ist die Lösung der pDgl

$$\begin{aligned} \frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} &= 0 && \text{in } (0, L) \times (0, \infty) \\ v &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ v(x, 0) &= u_0 && \text{in } (0, L) \end{aligned}$$

$$\frac{\partial v}{\partial t}(x, t) = u_1 \quad \text{in } (0, L)$$

gegeben durch

$$v(x, t) = \sum_{k=1}^{\infty} [c_k \cos(\frac{k\pi}{L}t) + s_k \sin(\frac{k\pi}{L}t)] \sin(\frac{k\pi}{L}x)$$

mit

$$c_k = \frac{2}{L} \int_0^L u_0(x) \sin(\frac{k\pi}{L}x) dx,$$

$$s_k = \frac{L}{k\pi} \frac{2}{L} \int_0^L u_1(x) \sin(\frac{k\pi}{L}x) dx.$$

XIV.3.3. Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen. Wir betrachten jetzt die pDgl

$$\begin{aligned} \frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} &= F \quad \text{in } (0, L) \times (0, \infty) \\ w &= 0 \quad \text{auf } \{0, L\} \times (0, \infty) \\ w(x, 0) &= 0 \quad \text{in } (0, L) \\ \frac{\partial w}{\partial t}(x, 0) &= 0 \quad \text{in } (0, L). \end{aligned}$$

Wie in Abschnitt [XIV.2.4](#) (S. 505) setzen wir die Funktion F für jedes t ungerade bzgl. x $2L$ -periodisch fort und entwickeln sie in eine Fourierreihe bzgl. x :

$$F(x, t) \sim \sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin(\frac{k\pi}{L}x)$$

mit

$$\widehat{F}_k(t) = \frac{2}{L} \int_0^L F(x, t) \sin(\frac{k\pi}{L}x) dx.$$

Für w machen wir wieder den Ansatz

$$w(x, t) \sim \sum_{k \in \mathbb{N}^*} w_k(t) \sin(\frac{k\pi}{L}x).$$

Einsetzen in die pDgl ergibt dann

$$\sum_{k \in \mathbb{N}^*} \widehat{F}_k(t) \sin(\frac{k\pi}{L}x) \sim F$$

$$\begin{aligned}
&= \frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} \\
&\sim \sum_{k \in \mathbb{N}^*} [\ddot{w}_k(t) + (\frac{k\pi}{L})^2 w_k(t)] \sin(\frac{k\pi}{L} x).
\end{aligned}$$

Vergleich der Fourier-Koeffizienten liefert die gDgl

$$\ddot{w}_k(t) + (\frac{k\pi}{L})^2 w_k(t) = \widehat{F}_k(t)$$

für die Koeffizienten w_k . Zusammen mit den Anfangsbedingungen

$$w_k(0) = \dot{w}_k(0) = 0$$

ergibt sich gemäß Abschnitt VI.3.2 (S. 233) und VI.3.3 (S. 236) die Lösung

$$\begin{aligned}
w_k(t) &= -\cos(\frac{k\pi}{L} t) \int_0^t \frac{L}{k\pi} \sin(\frac{k\pi}{L} s) \widehat{F}_k(s) ds \\
&\quad + \sin(\frac{k\pi}{L} t) \int_0^t \frac{L}{k\pi} \cos(\frac{k\pi}{L} s) \widehat{F}_k(s) ds.
\end{aligned}$$

Insgesamt erhalten wir somit:

Die Lösung w der pDgl

$$\begin{aligned}
\frac{\partial^2 w}{\partial t^2} - \frac{\partial^2 w}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\
w &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\
w(x, 0) &= 0 && \text{in } (0, L) \\
\frac{\partial w}{\partial t}(x, 0) &= 0 && \text{in } (0, L)
\end{aligned}$$

ist gegeben durch

$$w(x, t) = \sum_{k \in \mathbb{N}^*}^{\infty} [-c_k(t) \cos(\frac{k\pi}{L} t) + s_k(t) \sin(\frac{k\pi}{L} t)] \sin(\frac{k\pi}{L} x)$$

mit

$$\begin{aligned}
c_k(t) &= \int_0^t \int_0^L \frac{L}{k\pi} \frac{2}{L} F(x, s) \sin(\frac{k\pi}{L} s) \sin(\frac{k\pi}{L} x) dx ds \\
s_k(t) &= \int_0^t \int_0^L \frac{L}{k\pi} \frac{2}{L} F(x, s) \cos(\frac{k\pi}{L} s) \sin(\frac{k\pi}{L} x) dx ds.
\end{aligned}$$

XIV.3.4. Integraldarstellung. Für $x, \eta \in (0, L)$ und $t \in (0, \infty)$ definieren wir die Funktion $H(x, t, \eta)$ durch

$$H(x, t, \eta) = \frac{2}{L} \sum_{k=1}^{\infty} \frac{L}{k\pi} \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{k\pi}{L}t\right) \sin\left(\frac{k\pi}{L}\eta\right).$$

H heißt WELLENGLEICHUNGSKERN. Mit seiner Hilfe kann man die Ergebnisse der vorigen beiden Abschnitte wie folgt zusammenfassen:

Die Funktionen u_0 und u_1 seien differenzierbar und mögen die Bedingungen

$$u_0(0) = u_0(L) = 0$$

$$u_1(0) = u_1(L) = 0$$

erfüllen. Dann ist die Lösung u der pDgl

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} &= F && \text{in } (0, L) \times (0, \infty) \\ u &= 0 && \text{auf } \{0, L\} \times (0, \infty) \\ u(x, 0) &= u_0 && \text{in } (0, L) \\ \frac{\partial u}{\partial t}(x, 0) &= u_1 && \text{in } (0, L) \end{aligned}$$

gegeben durch

$$\begin{aligned} u(x, t) &= \int_0^L \frac{\partial}{\partial t} H(x, t, \eta) u_0(\eta) d\eta \\ &+ \int_0^L H(x, t, \eta) u_1(\eta) d\eta \\ &+ \int_0^L \left\{ \int_0^t H(x, t-s, \eta) F(\eta, s) ds \right\} d\eta. \end{aligned}$$

XIV.3.5. Höhere Raumdimensionen. Das Vorgehen im Fall $n \geq 2$ ist völlig analog zu Abschnitt XIV.2.6 (S. 507). Die Gleichung für T und x ist dann

$$\begin{aligned} \ddot{T}(t) &= \lambda T(t), \\ \Delta X &= \lambda X. \end{aligned}$$

Daher müssen die Fourier-Koeffizienten c_k , s_k und \widehat{F}_k der Abschnitte XIV.3.2 und XIV.3.3 durch die entsprechenden Terme

$$\int_{\Omega} u_0 \varphi_k dx, \quad \int_{\Omega} u_1 \varphi_k dx, \quad \int_{\Omega} F(x, t) \varphi_k(x) dx$$

mit den Eigenfunktionen des Laplace-Operators ersetzt werden.

XIV.3.6. Dämpfung. Wir betrachten nun die pDgl

$$\frac{\partial^2 u}{\partial t^2} + \alpha \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f \quad \text{in } (0, L) \times (0, \infty)$$

mit den bisherigen Rand- und Anfangsbedingungen. Neu ist der Term $\alpha \frac{\partial u}{\partial t}$, der physikalisch eine Dämpfung beschreibt. Die Gleichungen für T in Abschnitt XIV.3.2 und w_k in Abschnitt XIV.3.3 haben jetzt die Form

$$\begin{aligned} \ddot{T}(t) + \alpha \dot{T}(t) &= -\left(\frac{k\pi}{L}\right)^2 T(t) \\ \ddot{w}_k(t) + \alpha \dot{w}_k(t) + \left(\frac{k\pi}{L}\right)^2 w_k(t) &= \widehat{F}_k(t). \end{aligned}$$

Diese können mit den Methoden der Abschnitte VI.3.2 (S. 233) und VI.3.3 (S. 236) gelöst werden. Sofern

$$\alpha^2 < \left(\frac{\pi}{L}\right)^2$$

ist, müssen die Funktion $\sin(\frac{k\pi}{L}t)$ und $\cos(\frac{k\pi}{L}t)$ nun durch die Funktionen

$$e^{-\frac{\alpha}{2}t} \cos\left(\frac{1}{2}\sqrt{\left(\frac{2k\pi}{L}\right)^2 - \alpha^2} t\right)$$

und

$$e^{-\frac{\alpha}{2}t} \sin\left(\frac{1}{2}\sqrt{\left(\frac{2k\pi}{L}\right)^2 - \alpha^2} t\right)$$

ersetzt werden.

XIV.4. Die Poissongleichung

XIV.4.1. Das Eigenwertproblem im Rechteck. Sei

$$\Omega = (0, L) \times (0, H) = \{(x, y) : 0 < x < L, 0 < y < H\}$$

das Rechteck mit Kantenlängen L und H und linker unterer Ecke im Ursprung. Wir betrachten das EIGENWERTPROBLEM:

$$\begin{array}{ll} -\Delta u = \lambda u & \text{in } \Omega \\ u = 0 & \text{auf } \partial\Omega. \end{array}$$

Gesucht sind dabei der EIGENWERT λ und die EIGENFUNKTION $u : R \rightarrow \mathbb{R}$. Dabei interessieren uns natürlich nur nicht triviale Lösungen, d.h. Eigenfunktionen u mit

$$\int_{\Omega} |u(x, y)|^2 dx dy > 0.$$

Vorab überlegen wir uns, dass alle Eigenwerte positiv sind. Ist nämlich

$$-\Delta u = \lambda u,$$

können wir die Gleichung mit u multiplizieren und über Ω integrieren. Mit dem Gaußschen Integralsatz (vgl. Abschnitt IX.5.5 (S. 372)) erhalten wir dann

$$\begin{aligned} \lambda \int_{\Omega} |u|^2 dx &= \int_{\Omega} \underbrace{-\Delta u u}_{= -\operatorname{div}(u \nabla u) + \nabla u \cdot \nabla u} dx \\ &= - \underbrace{\int_{\Omega} \operatorname{div}(u \nabla u) dx}_{= \int_{\partial \Omega} u \frac{\partial u}{\partial n} ds} + \int_{\Omega} |\nabla u|^2 dx \\ &= - \int_{\partial \Omega} \underbrace{u}_{=0} \frac{\partial u}{\partial n} ds + \int_{\Omega} |\nabla u|^2 dx \\ &= \int_{\Omega} |\nabla u|^2 dx. \end{aligned}$$

Offensichtlich ist

$$\int_{\Omega} |\nabla u|^2 dx \geq 0.$$

Wäre

$$\int_{\Omega} |\nabla u|^2 dx = 0,$$

folgte $\nabla u(x) = 0$ für alle $x \in \Omega$. Daher wäre u konstant. Wegen der Randbedingung müsste diese Konstante gleich Null sein. Da wir aber nur an nicht trivialen Lösungen interessiert sind, folgt

$$\int_{\Omega} |\nabla u|^2 dx > 0.$$

Daher folgt aus obiger Gleichung $\lambda > 0$.

Man beachte, dass wir für diese Überlegung die spezielle Gestalt von Ω nicht ausgenutzt haben. Sie gilt für jede offene Menge $\Omega \subset \mathbb{R}^n$, $n \geq 2$, mit stückweise glattem Rand.

Wie in den vorigen Paragraphen machen wir einen Separationsansatz

$$u(x, y) = X(x)Y(y).$$

Im Folgenden bezeichnet $'$ die Ableitung nach x und $\dot{}$ diejenige nach y . Wie in Abschnitt XIV.2.3 (S. 502) führt der Separationsansatz auf die gDglen

$$\begin{aligned} X''(x) &= \mu X(x), \\ \ddot{Y}(y) &= -(\lambda + \mu)Y(y) \end{aligned}$$

mit einer unbekanntenen Konstanten $\mu \in \mathbb{R}$. Wegen der Randbedingung an u muss für die Lösungen X, Y dieser gDglen gelten

$$\begin{aligned} X(0) = X(L) &= 0, \\ Y(0) = Y(H) &= 0. \end{aligned}$$

Wie in Abschnitt XIV.2.3 (S. 502) folgt hieraus, dass

$$\mu = -\left(\frac{k\pi}{L}\right)^2, \quad k \in \mathbb{N}^*$$

und

$$X(x) = \sin\left(\frac{k\pi}{L}x\right)$$

sein muss. Ebenso folgt, dass

$$-(\lambda + \mu) = -\left(\frac{\ell\pi}{H}\right)^2, \quad \ell \in \mathbb{N}^*$$

und

$$Y(y) = \sin\left(\frac{\ell\pi}{H}y\right)$$

sein muss. Aus den Gleichungen für λ und μ ergibt sich insbesondere

$$\lambda = \left(\frac{\ell\pi}{H}\right)^2 + \left(\frac{k\pi}{L}\right)^2, \quad k, \ell \in \mathbb{N}^*.$$

Insgesamt erhalten wir somit:

Die Lösungen λ und u des Eigenwertproblems

$$\begin{aligned} -\Delta u &= \lambda && \text{in } \Omega \\ u &= 0 && \text{auf } \partial\Omega \end{aligned}$$

mit $\Omega = (0, L) \times (0, H)$ sind von der Form

$$\lambda = \left(\frac{\ell\pi}{H}\right)^2 + \left(\frac{k\pi}{L}\right)^2$$

$$u(x, y) = \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right)$$

mit $k, \ell \in \mathbb{N}^*$.

XIV.4.2. Die Poissongleichung im Rechteck. Sei wieder $\Omega = (0, L) \times (0, H)$. Wir betrachten die Poissongleichung mit Dirichlet-Randbedingungen

$$\begin{array}{l} -\Delta u = f \quad \text{in } \Omega \\ u = g \quad \text{auf } \partial\Omega. \end{array}$$

Definiere die Fortsetzung $G : \Omega \rightarrow \mathbb{R}$ der Randwerte g durch

$$\begin{aligned} G(x, y) &= g(x, 0)\left(1 - \frac{y}{H}\right) + g(x, H)\frac{y}{H} \\ &\quad + [g(0, y) - g(0, 0)\left(1 - \frac{y}{H}\right) - g(0, H)\frac{y}{H}]\left(1 - \frac{x}{L}\right) \\ &\quad + [g(L, y) - g(L, 0)\left(1 - \frac{y}{H}\right) + g(L, H)\frac{y}{H}]\frac{x}{L}. \end{aligned}$$

Dann löst

$$v = u - G$$

die Poissongleichung

$$\begin{array}{ll} -\Delta v = F & \text{in } \Omega \\ v = 0 & \text{auf } \partial\Omega \end{array}$$

mit homogenen Dirichlet-Randbedingungen und rechter Seite

$$F = f + \Delta G.$$

Für festes $y \in (0, H)$ können wir die Funktion F bzgl. x ungerade zu einer $2L$ -periodischen Funktion fortsetzen und die Fourier-Reihe bzgl. x dieser Fortsetzung wie in Abschnitt XIII.2.1 (S. 480) berechnen. Die Fourier-Koeffizienten sind dann Funktionen von y . Diese können ungerade zu $2H$ -periodischen Funktionen fortgesetzt und in eine Fourier-Reihe bzgl. y entwickelt werden. Insgesamt erhalten wir eine doppelte Fourier-Reihe

$$F(x, y) \sim \sum_{k \in \mathbb{N}^*} \sum_{\ell \in \mathbb{N}^*} \widehat{F}_{k\ell} \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right)$$

mit

$$\widehat{F}_{k\ell} = \frac{2}{L} \frac{2}{H} \int_0^L \int_0^H F(x, y) \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right) dx dy.$$

Für v machen wir nun den Ansatz

$$v(x, y) \sim \sum_{k \in \mathbb{N}^*} \sum_{\ell \in \mathbb{N}^*} v_{k\ell} \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{\ell\pi}{H}y\right).$$

Einsetzen in die pDgl und Vergleich der Fourier-Koeffizienten ergibt

$$\left[\left(\frac{k\pi}{L}\right)^2 + \left(\frac{\ell\pi}{H}\right)^2 \right] v_{k\ell} = \widehat{F}_{k\ell} \quad k, \ell \in \mathbb{N}^*.$$

Unter geeigneten Stetigkeits- und Differenzierbarkeitsannahmen an f und g kann man zeigen, dass die Fourier-Reihe für v konvergiert. Man erhält dann insgesamt:

Die Lösung u der Poisson-Gleichung

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \\ u &= g && \text{auf } \partial\Omega \end{aligned}$$

im Rechteck $\Omega = (0, L) \times (0, H)$ ist gegeben durch

$$u = v + G$$

mit

$$\begin{aligned} G(x, y) &= g(x, 0)\left(1 - \frac{y}{H}\right) + g(x, H)\frac{y}{H} \\ &+ [g(0, y) - g(0, 0)\left(1 - \frac{y}{H}\right) - g(0, H)\frac{y}{H}]\left(1 - \frac{x}{L}\right) \\ &+ [g(L, y) - g(L, 0)\left(1 - \frac{y}{H}\right) - g(L, H)\frac{y}{H}]\frac{x}{L} \end{aligned}$$

und

$$v(x, y) = \int_0^L \int_0^H G(x, y, \xi, \eta) F(\xi, \eta) d\xi d\eta$$

mit

$$\begin{aligned} G(x, y, \xi, \eta) &= \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{1}{\left(\frac{k\pi}{L}\right)^2 + \left(\frac{\ell\pi}{H}\right)^2} \sin\left(\frac{k\pi}{L}x\right) \\ &\cdot \sin\left(\frac{\ell\pi}{H}y\right) \sin\left(\frac{k\pi}{L}\xi\right) \sin\left(\frac{\ell\pi}{H}\eta\right) \end{aligned}$$

$$F = f + \Delta G.$$

XIV.4.3. Das Eigenwertproblem im Kreis. Sei nun

$$\Omega = B_R(0) = \{(x, y) : x^2 + y^2 < R^2\}$$

der Kreis um den Nullpunkt mit Radius R . Wir betrachten wieder das Eigenwertproblem aus Abschnitt XIV.4.1 (S. 514). Dazu führen wir ebene Polarkoordinaten ein

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad 0 \leq r \leq R, \quad 0 \leq \varphi < 2\pi.$$

Gemäß Beispiel VIII.2.19 (S. 293) lautet die pDgl in den Polarkoordinaten dann

$$\begin{aligned} -\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} &= \lambda u \quad \text{in } (0, R) \times [0, 2\pi) \\ u &= 0 \quad \text{für } r = R. \end{aligned}$$

Wir machen wieder einen Separationsansatz

$$u(r, \varphi) = v(r)\Phi(\varphi)$$

und bezeichnen mit ' die Ableitung bzgl. r und mit $\dot{}$ diejenige bzgl. φ . Dann erhalten wir die beiden gDglen

$$\begin{aligned}\ddot{\Phi} &= \mu\Phi, \\ -r(rv')' - r^2\lambda v &= \mu v\end{aligned}$$

mit einer unbekanntenen Konstanten μ . Aufgrund unseres Ansatzes muss die Funktion Φ 2π -periodisch sein. Daher erhalten wir für μ und Φ die Lösungen

$$\begin{aligned}\mu &= 0, \\ \Phi(\varphi) &= 1\end{aligned}$$

und mit $k \in \mathbb{N}^*$

$$\begin{aligned}\mu &= -k^2, \\ \Phi(\varphi) &= \cos(k\varphi)\end{aligned}$$

und

$$\Phi(\varphi) = \sin(k\varphi).$$

Damit lautet die Bestimmungsgleichung für v und λ

$$r^2v'' + rv' + (r^2\lambda - k^2)v = 0$$

mit $k \in \mathbb{N}$. Da wir uns in Abschnitt [XIV.4.1](#) (S. [514](#)) schon überlegt haben, dass $\lambda > 0$ sein muss, können wir den Ansatz

$$\begin{aligned}v(r) &= w(\rho), \\ \rho &= \sqrt{\lambda}r\end{aligned}$$

machen. Wir erhalten dann für w die gDgl

$$\rho^2 \frac{d^2w}{d\rho^2} + \rho \frac{dw}{d\rho} + (\rho^2 - k^2)w = 0.$$

Dies ist die Besselsche Differentialgleichung aus [Beispiel IX.1.2](#) (S. [326](#)) mit der Lösung

$$\begin{aligned}w(\rho) &= J_k(\rho) \\ &= \frac{1}{\pi} \int_0^\pi \cos(\rho \sin t - kt) dt.\end{aligned}$$

Hieraus ergibt sich für v

$$\begin{aligned}v(r) &= J_k(\sqrt{\lambda}r) \\ &= \frac{1}{\pi} \int_0^\pi \cos(\sqrt{\lambda}r \sin t - kt) dt.\end{aligned}$$

Der Eigenwert λ wird durch die Randbedingung

$$v(R) = 0$$

festgelegt.

Insgesamt erhalten wir:

Die Eigenfunktionen des Laplace-Operators auf dem Kreis $B_R(0)$ sind von der Form

$$J_k(\sqrt{\lambda_k}r) \cos(k\varphi) \quad , k \in \mathbb{N},$$

$$J_\ell(\sqrt{\lambda_\ell}r) \sin(\ell\varphi) \quad , \ell \in \mathbb{N}^*$$

mit

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin t - nt) dt.$$

Die Eigenwerte λ_k , $k \in \mathbb{N}$, sind festgelegt durch die Bedingung

$$J_k(\sqrt{\lambda_k}R) = 0.$$

XIV.4.4. Die Poissongleichung im Kreis. Sei wieder $\Omega = B_R(0)$ der Kreis um Null mit Radius R . Wir betrachten die Poissongleichung mit Dirichlet-Randbedingungen

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \\ u &= g && \text{auf } \partial\Omega. \end{aligned}$$

In Polarkoordinaten lautet diese pDgl

$$-\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} = f \quad \text{für } 0 \leq r < R, 0 \leq \varphi < 2\pi$$

$$u(R, \varphi) = g(\varphi) \quad \text{für } 0 \leq \varphi < 2\pi.$$

Wir setzen nun die Randdaten g fort und definieren dazu

$$G(r, \varphi) = \frac{r}{R} g(\varphi) \quad 0 \leq r \leq R, \quad 0 \leq \varphi < 2\pi.$$

Für

$$v = u - G$$

erhalten wir dann die pDgl

$$-\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 v}{\partial \varphi^2} = F \quad \text{für } 0 \leq r < R, 0 \leq \varphi < 2\pi$$

$$v(R, \varphi) = 0 \quad \text{für } 0 \leq \varphi < 2\pi$$

mit homogenen Dirichlet Randbedingungen und rechter Seite

$$F = f + \Delta G.$$

Wir entwickeln F in eine Fourier-Reihe bzgl. φ

$$F(r, \varphi) \sim \sum_{k \in \mathbb{Z}} \hat{F}_k(r) e^{ik\varphi}$$

mit

$$\widehat{F}_k(r) = \frac{1}{2\pi} \int_0^{2\pi} F(r, \varphi) e^{-ik\varphi} d\varphi$$

und machen für v den Fourier-Ansatz

$$v(r, \varphi) \sim \sum_{k \in \mathbb{Z}} v_k(r) e^{ik\varphi}.$$

Einsetzen in die pDgl und Vergleich der Fourier-Koeffizienten liefert für die v_k die gDglen

$$-\frac{1}{r}(rv'_k)' + \frac{k^2}{r^2}v_k = \widehat{F}_k \quad , k \in \mathbb{Z}.$$

Dies sind lineare gDglen 2. Ordnung mit variablen Koeffizienten. Wir gehen zur Lösung ähnlich vor wie in Abschnitt VI.3 (S. 233) und bestimmen zunächst die allgemeine Lösung der homogenen Gleichung. Sei dazu $k \in \mathbb{Z} \setminus \{0\}$ beliebig, aber fest gewählt. Wir machen den Ansatz

$$w_k(r) = r^\alpha$$

für eine Lösung der homogenen Gleichung und erhalten die Bestimmungsgleichungen

$$\begin{aligned} 0 &= \frac{1}{r}(r\alpha r^{\alpha-1})' + \frac{k^2}{r^2}r^\alpha \\ &= -\alpha^2 r^{\alpha-2} + k^2 r^{\alpha-2} \\ &= (k^2 - \alpha^2)r^{\alpha-2}. \end{aligned}$$

Also lautet die allgemeine Lösung der homogenen Gleichung

$$w_k(r) = a_k r^k + b_k r^{-k}.$$

Für $k = 0$ erhalten wir die allgemeine Lösung

$$w_0(r) = a_0 + b_0 \ln r.$$

Für uns sind aber nur die Lösungen von Interesse, die für $r \rightarrow 0$ beschränkt bleiben. Daher machen wir zur Lösung der inhomogenen Gleichung den Variation-der-Konstanten-Ansatz

$$v_k(r) = a_k(r) r^{|k|} \quad , k \in \mathbb{Z}.$$

Einsetzen in die gDgl ergibt dann mit ein wenig Rechnung für alle $k \in \mathbb{Z}$ die Bestimmungsgleichung

$$-\frac{1}{r^{|k|+1}} [r^{2|k|+1} a'_k]' = \widehat{F}_k$$

für die a_k . Wegen der Randbedingung $a_k(R) = 0$ ergibt sich hieraus die Lösung

$$\begin{aligned} v_0(r) &= a_0(r) \\ &= \int_r^R \frac{1}{s} \int_0^s \tau \widehat{F}_0(\tau) d\tau ds \end{aligned}$$

$$\begin{aligned}
&= -\ln s \int_0^s \tau \widehat{F}_0(\tau) d\tau \Big|_{s=r}^{s=R} \\
&\quad + \int_r^R \ln s - s \widehat{F}_0(s) ds \\
&= -\ln R \int_0^R s \widehat{F}_0(s) ds \\
&\quad + \ln r \int_0^r s \widehat{F}_0(s) ds \\
&\quad + \int_r^R s \ln s \widehat{F}_0(s) ds
\end{aligned}$$

und für $k > 0$

$$\begin{aligned}
v_k(r) &= r^{|k|} a_k(r) \\
&= -r^{|k|} \int_r^R \frac{1}{s^{2|k|+1}} \int_0^s \tau^{|k|+1} \widehat{F}_k(\tau) d\tau ds \\
&= -r^{|k|} \left\{ -\frac{1}{2|k|} s^{-2|k|} \int_0^s \tau^{|k|+1} \widehat{F}_k(\tau) d\tau \Big|_{s=r}^{s=R} \right. \\
&\quad \left. + \int_r^R \frac{1}{2|k|} s^{-2|k|} s^{|k|+1} \widehat{F}_k(s) ds \right\} \\
&= \frac{1}{2|k|} r^{|k|} R^{-2|k|} \int_0^R s^{|k|+1} \widehat{F}_k(s) ds \\
&\quad - \frac{1}{2|k|} r^{-|k|} \int_0^r s^{|k|+1} \widehat{F}_k(s) ds \\
&\quad - \frac{1}{2|k|} r^{|k|} \int_0^R s^{-|k|+1} \widehat{F}_k(s) ds.
\end{aligned}$$

Unter geeigneten Voraussetzungen an f und g kann man wieder zeigen, dass die Fourier-Reihe für v konvergiert. Insgesamt erhalten wir:

Die Lösung u der Poissongleichung mit Dirichlet-Randbedingungen

$$\begin{aligned}
-\Delta u &= f && \text{in } \Omega \\
u &= g && \text{auf } \partial\Omega
\end{aligned}$$

im Kreis um Null mit Radius R ist gegeben durch

$$u = v + G$$

mit

$$G(r, \varphi) = \frac{r}{R} g(\varphi),$$

$$v(r, \varphi) = \sum_{k=-\infty}^{\infty} v_k(r) e^{ik\varphi}$$

und

$$\begin{aligned} v_0(r) &= -\ln R \int_0^R s \widehat{F}_0(s) ds \\ &\quad + \ln r \int_0^r s \widehat{F}_0(s) ds \\ &\quad + \int_r^R s \ln s \widehat{F}_0(s) ds \\ v_k(r) &= \frac{1}{2|k|} \left\{ \left(\frac{r}{R^2} \right)^{|k|} \int_0^R s^{|k|+1} \widehat{F}_k(s) ds \right. \\ &\quad \left. - r^{-|k|} \int_0^r s^{|k|+1} \widehat{F}_k(s) ds \right. \\ &\quad \left. - r^{|k|} \int_r^R s^{-|k|+1} \widehat{F}_k(s) ds \right\}, \quad k \neq 0, \\ \widehat{F}_k(r) &= \frac{1}{2\pi} \int_0^{2\pi} F(r, \varphi) e^{-ik\varphi} d\varphi, \quad k \in \mathbb{Z}, \\ F &= f + \Delta G. \end{aligned}$$

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-Quotienten, 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

VII Potenzreihen

1. Reihen

Definition; Partialsummen; Konvergenz; geometrische Reihe; harmonische Reihe; alternierende harmonische Reihe; absolute und bedingte Konvergenz; alternierende harmonische Reihe ist bedingt konvergent; absolut konvergent \implies konvergent; konvergent \implies Reihenglieder bilden Nullfolge; Leibnizkriterium; Majorantenkriterium; Quotientenkriterium; Summen und Vielfache konvergenter Reihen; Cauchy-Produkt

2. Potenzreihen

Definition; Konvergenzradius; Bestimmung des Konvergenzradius: mittels Quotienten, mittels Wurzeln; Verhalten am Rand des Konvergenzbereiches; Potenzreihen stellen Funktionen dar; gliedweise Differentiation von Potenzreihen; gliedweise Integration von Potenzreihen; Potenzreihen der Exponential-, Logarithmus-, trigonometrischen und inversen trigonometrischen Funktionen; Binomialreihe; Potenzreihen mit beliebigem Entwicklungspunkt; Koeffizientenvergleich

3. Taylorreihen

Taylorpolynom; Taylorformel; Restglied; Extremwerttest; Taylorreihe; Darstellung von Funktionen durch Taylorreihen; Taylorreihen der hyperbolischen und der inversen trigonometrischen Funktionen

4. Anwendungen

Grenzwertbestimmung, Näherungsformeln und Integration mittels Taylorreihen; Taylorreihenansatz zur Lösung gDglen

VIII Differentiation von Funktionen in mehreren Variablen

1. Kurven im \mathbb{R}^n

Parameterdarstellung einer Kurve; Tangentialvektor; Anfangs- und Endpunkt einer Kurve; reguläre Kurven; Bogenlänge; Länge; begleitendes Dreibein; Hauptnormalen- und Binormalenvektor; Krümmung; Torsion; Krümmung und Torsion einer Schraubenfeder

2. Reellwertige Funktionen mehrerer reeller Veränderlicher

Betrag eines Vektors; offener Ball mit Mittelpunkt \mathbf{a} und Radius r ; innere Punkte; Randpunkte; abgeschlossene Mengen; beschränkte Mengen; kompakte Mengen; Beispiele; Niveaumengen; Graphen; Grenzwerte und Stetigkeit von Funktionen mehrerer Veränderlicher; stetige Funktionen nehmen auf kompakten Mengen ihr Minimum und Maximum an; partielle Ableitungen; Gradient; stetige partielle Differenzierbarkeit; Vertauschbarkeit der partiellen Ableitungen; totale Ableitung und lineare Approximation; totale Differenzierbarkeit; stetig total differenzierbar \iff stetig partiell differenzierbar; Näherungsberechnung; Richtungsableitungen; Kettenregel; Transformation auf ebene und räumliche Polarkoordinaten; Laplace-Operator und seine Darstellung in kartesischen Koordinaten und Polarkoordinaten

3. Anwendungen der Differentiation

Gradient gibt Richtung des stärksten Anstiegs; Gradientenverfahren zur Bestimmung von Maxima und Minima; Tangenten; Tangentialebene; Normalendarstellung der Tangentialebene; Taylorformel für Funktionen mehrerer Veränderlicher; Taylorpolynom; Hesse-Matrix; Schmiegequadriken; Satz über implizite Funktionen; Bestimmung der Ableitung impliziter Funktionen; Lösung nichtlinearer Gleichungssysteme; lokale Extrema; stationäre Punkte; lokales Extremum \implies stationärer Punkt; Extremwerttest mittels der Hesse-Matrix; Extrema unter Nebenbedingungen; Lagrange-Multiplikatoren; Bestimmen der Extrema auf kompakten Mengen

4. Vektorwertige Funktionen

Differentiation; Jacobi-Matrix; Newtonverfahren zur Lösung nicht linearer Gleichungssysteme; Kettenregel; Basiswechsel; räumliche Skalaren- und Vektorfelder; starre Drehungen; Zentralfelder; laminare Rohrströmung; Gradient; Divergenz; Rotation; Laplace-Operator; Nabla-Operator; Koordinateninvarianz von Divergenz, Gradient und Rotation; wirbel- und quellfreie Felder

IX Integration von Funktionen in mehreren Variablen

1. Parameterintegrale

eigentliche und uneigentliche Parameterintegrale; Gammafunktion; Besselfunktionen; Fourier-Transformation; Differentiation von eigentlichen Parameterintegralen; Besselsche Differentialgleichung; Differentiation bei variablen Integrationsgrenzen; Differentiation von uneigentlichen Parameterintegralen; Ableitungen der Gammafunktion

2. Kurvenintegrale

Kurvenintegral einer skalaren Funktion; Rechenregeln; Trägheitsmomente; Schwerpunkt; Kurvenintegral eines Vektorfeldes; geschlossene Kurven; Gradientenfeld; Potential; Stammfunktion; Satz von Poincaré; Methoden zur Bestimmung einer Stammfunktion

3. Integration über ebene Bereiche

Flächeninhalt; messbare Mengen; Nullmengen; Doppel- oder Gebietsintegral; geometrische Deutung; Rechenregeln; Mittelwertsatz; praktische Berechnung von Doppelintegralen; Satz von Green; Anwendungen; Satz von Gauß in der Ebene

4. Integration über Flächen im Raum

Parameterdarstellung einer Fläche; Tangentialebene; Normalenvektor; metrische Fundamentalgrößen; Rand einer Fläche; Drehflächen; Sphäre; Torus; Wendelfläche; Flächeninhalt; Oberflächenintegral einer skalaren Funktion; Transformationsformel für Gebietsintegrale; affine Koordinaten; Polarkoordinaten; elliptische Koordinaten; Fluss eines Vektorfeldes; orientierbare Flächen; Beispiele; Möbiusband ist nicht orientierbar; Satz von Stokes; Anwendungen

5. Integration über dreidimensionale Bereiche

Volumen; Nullmengen; reguläre Mengen; Dreifach- oder Volumenintegral; praktische Berechnung von Dreifachintegralen; Jacobi- oder Funktionaldeterminante; Transformationsformel für Volumenintegrale; affine Koordinaten; Zylinderkoordinaten; Kugelkoordinaten; Satz von Gauß; Anwendungen; Kontinuitätsgleichungen der Strömungsmechanik

X Gewöhnliche Differentialgleichungen II: Systeme

1. Existenz- und Eindeutigkeitsätze
Gewöhnliche Differentialgleichung n -ter Ordnung; Anfangswertproblem; Lösung; Reduktion der Ordnung; Schwingungsgleichung; Lipschitz-stetige Funktionen; Zusammenhang mit Differenzierbarkeit; Satz von Picard-Lindlöf; stetige Abhängigkeit der Lösung eines Anfangswertproblems von den Anfangsbedingungen
2. Differentialgleichungen 1. Ordnung
Exakte Differentialgleichungen; Lösung exakter Differentialgleichungen; integrierende Faktoren; Methoden zur Bestimmung integrierender Faktoren; Räuber-Beute-Modelle
3. Systeme linearer Differentialgleichungen
Homogene und inhomogene Gleichungen; partikuläre Lösung; allgemeine Lösungen; Fundamentalsysteme und ihre Bestimmung; Wronski-Determinante; Methode der Variation der Konstanten; homogene Systeme mit konstanter, symmetrischer Matrix; homogene Systeme mit konstanter, allgemeiner Matrix; lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten
4. Stabilität
Motivation der Problemstellung; Ljapunov-Stabilität; asymptotische Stabilität; Instabilität; Stabilität von linearen Systemen mit konstanter Matrix; Stabilität von gestörten linearen Systemen; autonome Systeme und ihre Stabilität

XI Stochastik I: Diskrete Modelle

1. Modelle für Zufallsexperimente
Ergebnisse oder Elementarereignisse; Ergebnismenge; Ereignisse; Potenzmenge; Wahrscheinlichkeitsverteilung oder -maß; Wahrscheinlichkeit; Wahrscheinlichkeitsraum; Wahrscheinlichkeitsfunktion; Laplace-Experimente; Urnenmodelle: Stichproben in Reihenfolge mit Rücklegen, Stichproben in Reihenfolge ohne Rücklegen, Stichproben ohne Reihenfolge ohne Rücklegen, Stichproben ohne Reihenfolge mit Rücklegen; Anwendungsbeispiele; hypergeometrische Verteilung; Multinomialkoeffizienten; Identitäten für Binomialkoeffizienten
2. Bedingte Wahrscheinlichkeit und Unabhängigkeit
Bedingte Wahrscheinlichkeiten; Beispiele; Produktformel; Formel der totalen Wahrscheinlichkeit; Formel von Bayes; Anwendungen; Unabhängigkeit; Produktexperimente; Produkte von Wahrscheinlichkeitsräumen; Bernoulliexperiment und -verteilung; Binomialverteilung; Multinomialverteilung; geometrische Verteilung; negative Binomialverteilung oder Pascalverteilung
3. Zufallsvariable, Erwartungswert, Varianz
Zufallsvariable; Beispiele; Stabdiagramme; gemeinsame Verteilungsfunktion; Unabhängigkeit; Erwartungswert; Erwartungswert der Binomialverteilung; Erwartungswert der hypergeometrischen Verteilung; Varianz; Kovarianz; Streuung; Standardabweichung; Korrelationskoeffizient; unkorrelierte Zufallsvariable; Varianz der Binomialverteilung; Varianz der hypergeometrischen Verteilung; schwaches Gesetz der großen Zahl; Anwendungen

4. Grundbegriffe der Schätztheorie
Motivation; allgemeiner Rahmen; Schätzer; Maximum-Likelihood Schätzer; Erwartungstreue; Mittelwert; mittlerer quadratischer Fehler
5. Approximationen der Binomialverteilung
Stirlingsche Formel; Dichte der Standard-Normalverteilung; Satz von Moivre-Laplace; Verteilungsfunktion der Standard-Normalverteilung; Anwendungen; Poissonapproximation; Poissonverteilung
6. Tests
Motivation; Hypothese; Annehmen und Verwerfen einer Hypothese; Verwerfungsbereich oder kritischer Bereich; Fehler erster und zweiter Art; Teststatistik; kritischer Wert; Gütefunktion

XII Stochastik II: Allgemeine Modelle

1. Wahrscheinlichkeitsmaße mit Dichten
Ergebnismengen; σ -Algebren; Borelmengen; Wahrscheinlichkeitsmaße; Wahrscheinlichkeitsräume; Dichten; Gleichverteilung auf einem Intervall; Exponentialverteilung; Normalverteilung; Produktdichten
2. Zufallsvariable und ihre Momente
Messbare Funktionen; Zufallsvariable; Verteilungsfunktion; Dichte; Unabhängigkeit; Faltung; Faltung von Normalverteilungen; Erwartungswert; Eigenschaften des Erwartungswertes; Varianz; Eigenschaften der Varianz
3. Schätzverfahren
Maximum-Likelihood Schätzung; Methode der kleinsten Quadrate; Regressionsgerade; Median
4. Tests
Likelihood-Quotienten; Likelihood-Quotienten Test; t -Test; χ^2 -Test; Anwendungen

XIII Fourier-Analysis

1. Trigonometrische Polynome und Reihen
Periodische Funktionen; direkte periodische Fortsetzung; gerade periodische Fortsetzung; ungerade periodische Fortsetzung; trigonometrische Polynome; trigonometrische Reihen
2. Fourier-Reihen
Fourier-Koeffizienten; Fourier-Reihe; Rechteckschwingung; Rechenregeln; Bessel-Ungleichung; Riemann-Lemma; Konvergenz der Fourier-Reihe einer Funktion; Parseval-Gleichung; Anwendung auf gewöhnliche Differentialgleichungen
3. Die Fourier-Transformation
Fourier-Transformation; Fourier-Transformierte einer Funktion; Rechteckimpuls; Rechenregeln; Fourier-Transformierte einiger wichtiger Funktionen; Existenz und Eindeutigkeit der Fourier-Transformierten; Parseval-Gleichung

XIV Partielle Differentialgleichungen

1. Einführung
Beispiele: Membrangleichung, Plattengleichung, Gasgleichung, Wärmeleitungsgleichung, Grundwasserströmung, Wellengleichung; Differentialoperatoren; Hauptteil; quasilinear; Koeffizienten; Differentialgleichung; Typen: elliptisch, parabolisch, hyperbolisch; Randbedingungen: Dirichlet, Neumann, gemischt

2. Die Wärmeleitungsgleichung

Kompabilitätsbedingung; Fortsetzung der Randwerte; Lösung der homogenen Gleichung mit homogenen Randbedingungen; Separationsansatz; Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen; Integraldarstellung; Wärmeleitungskern; höhere Raumdimensionen; Zusammenhang mit dem Eigenwertproblem für den Laplace-Operator

3. Die Wellengleichung

Lösung der homogenen Gleichung mit homogenen Randbedingungen; Separationsansatz; Lösung der inhomogenen Gleichung mit homogenen Anfangs- und Randbedingungen; Integraldarstellung; Wellengleichungskern; höhere Raumdimensionen; Dämpfung

4. Die Poissongleichung

Lösung des Eigenwertproblems im Rechteck; Lösung der Poissongleichung im Rechteck; Lösung des Eigenwertproblems im Kreis; Lösung der Poissongleichung im Kreis

Index

- $\chi_A(\lambda)$, 79
- \angle , 30
- $*$, 458
- \cap , 15
- $[\cdot]$, 434
- \circ , 16
- \cup , 15
- \emptyset , 15
- $o(\cdot)$, 289
- ∇ , 323
- $\|\cdot\|$, 377
- ∂ , 352
- \perp , 30
- $[\cdot, \cdot, \cdot]$, 32
- \subset , 15
- \times , 16
- \times , 31
- \rightarrow , 16
- B**, 280
- $B_r(\mathbf{a})$, 283
- $C^0(D, \mathbb{R})$, 289
- $C^k(D, \mathbb{R})$, 289
- $\text{Cov}(X, Y)$, 428
- Δ , 321
- $D\mathbf{f}$, 315
- EX , 426
- $E(X)$, 426, 461
- $F(M)$, 342
- $\mathcal{F}(f)$, 487
- $\mathcal{F}^{-1}(F)$, 487
- Γ_f , 284
- H_f , 299
- $L(\mathbf{x})$, 279
- N**, 280
- Ω , 409
- $P(A|B)$, 417
- Φ , 443
- T**, 279
- $T_n(f; x_0)$, 268, 298
- $V(M)$, 366
- $\text{Var}(X)$, 428
- A^{-1} , 66
- arccos, 158
- arccot, 160
- arcosh, 168
- arcsin, 158
- arctan, 159
- arsinh, 168
- artanh, 168
- A^T , 65
- $|\cdot|$, 20, 29, 45
- $b_{n,p}$, 422
- \mathbb{C} , 45
- \mathcal{P} , 409
- $\mathbb{C}^{m \times n}$, 54
- cos, 25
- cosh, 167
- cot, 129
- det, 72–74
- f' , 133
- dim, 106
- div, 321
- ds , 279
- e , 122
- η_i , 243
- $\eta(t_i, h_i)$, 243
- exp, 122
- \hat{f} , 487
- \hat{f}_k , 480
- f_{x_i} , 286
- $\Gamma(x)$, 196
- grad, 287
- $h(s; n, N, S)$, 414
- $[a, b]$, 19
- $[a, b)$, 19
- Im, 45
- I_n , 65
- inf, 19
- $-\infty$, 19
- ∞ , 19
- \int , 173

\int_a^b , 170
 \iint_M , 343
 \iiint_M , 367
 $(a, b]$, 19
 (a, b) , 19
 Kern, 103
 \lim , 116
 $\lim_{n \rightarrow \infty}$, 125
 $\lim_{x \rightarrow a^-}$, 125
 $\lim_{x \rightarrow a^+}$, 125
 \ln , 131
 \log_a , 167
 \mathbb{N} , 17
 \mathbb{N}^* , 17
 $n!$, 23
 $\binom{n}{k}$, 23
 $\sqrt[n]{z}$, 49
 ω , 409
 ∂D , 283
 $\frac{\partial f}{\partial x_i}$, 286
 $\partial_{\mathbf{v}} f$, 292
 φ , 442
 \mathbb{Q} , 17
 \mathbb{R} , 17
 Rang, 57
 Re, 45
 ρ_{XY} , 428
 $\mathbb{R}^{m \times n}$, 54
 \mathbb{R}^n , 54
 rot, 321
 σ_X , 428
 sin, 25
 sinh, 167
 span, 104
 Spur, 80
 $s(t)$, 279
 sup, 19
 tan, 129
 tanh, 167
 τ , 281
 \mathbb{Z} , 17

a posteriori Fehlerabschätzung, 148
 a priori Fehlerabschätzung, 148
 A-stabil, 253
 Abbildung, 16
 abgeschlossen, 283
 abgeschlossene Menge, 283
 abgeschlossenes Intervall, 19
 ableiten, 133
 Ableitung, 133
 absolut konvergente Reihe, 256
 Absolutglied, 56

Abstand, 42
 Additionstheoreme, 48
 ähnliche Matrizen, 88
 Ähnlichkeits-Differentialgleichung, 229
 äquidistante Knoten, 210
 Affensattel, 300
 affine Koordinaten, 360, 369
 algebraische Vielfachheit, 79
 allgemeine Höhenformel, 166
 allgemeine Lösung, 219
 Alternative, 449
 alternierend, 75
 alternierende harmonische Reihe, 256
 Anfangsbedingung, 495
 Anfangspunkt, 278
 Anfangswertproblem, 219, 375
 Annahme der Hypothese, 449
 Archimedisches Spirale, 205
 Arcuscosinus, 158
 Arcuscotangens, 160
 Arcussinus, 158
 Arcustangens, 159
 Area cosinus hyperbolicus, 168
 Area sinus hyperbolicus, 168
 Argumente, 16
 Assoziativgesetz, 29, 103
 Astroide, 199
 asymptotisch stabil, 403
 autonom, 219
 AWP, 219, 375

 Ball, 283
 barometrische Höhenformel, 167
 Basis, 105
 Basiswechsel, 318
 Bayessche Formel, 419
 bedingt konvergente Reihe, 256
 bedingte Wahrscheinlichkeit, 417
 begleitendes Dreibein, 279
 Bernoulli-Differentialgleichung, 231
 Bernoulli-Experiment, 422
 Bernoulli-Ungleichung, 22
 Bernoulli-Verteilung, 422
 beschränkt, 284
 beschränkte Folge, 115
 beschränkte Menge, 19, 284
 Bessel-Ungleichung, 483
 Besselfunktion, 325
 beste lineare Approximation, 135
 bestimmte Integral, 170
 Betrag, 20, 29, 45
 Bias, 436

- biharmonische Gleichung, 493
- bijektiv, 16
- Bild, 16, 277
- Binomialkoeffizient, 23
- Binomialreihe, 265
- Binomialverteilung, 422
- Binomische Formel, 24
- Binormalenvektor, 280
- Bogenlänge, 201, 279
- Bogenmaß, 25
- Borelmengen, 454

- Cauchy-Produkt, 259
- Cauchy-Schwarzsche Ungleichung, 30, 108
- charakteristisches Polynom, 79, 234
- χ^2 -Test, 474
- χ^2 -Verteilung, 473
- Cosinus, 25
- Cosinus hyperbolicus, 167
- Cotangens, 129
- Cramersche Regel, 77

- Definitionsbereich, 16, 277
- Determinante, 72–74
- Determinantenform, 41
- diagonalisierbare Matrix, 88
- Dichte, 455, 457
- Dichte der
 - Standard-Normalverteilung, 442
- Differentialgleichung, 162
- Differentialoperator m -ter Ordnung, 497
- Differentiation von Potenzreihen, 262
- differenzierbar, 133
- differenzieren, 133
- Diffusivität, 496
- Dimension, 106
- Dirichlet-Randbedingung, 493
- disjunkt, 16
- Distributivgesetz, 30
- Divergenz, 321
- Doppelintegral, 343
- Drehfläche, 353
- Drehmoment, 35
- Drei-Punkte-Gleichung, 41
- Dreiecksungleichung, 20, 29, 108
- Dreifachintegral, 367
- Durchschnitt, 15

- ebene Polarkoordinaten, 293
- Eigenfunktion, 514
- eigentliches Parameterintegral, 325
- Eigenvektor, 78
- Eigenwert, 78, 514
- Eigenwertproblem, 514
- Eigenwertproblem für den
 - Laplace-Operator, 507
- Einheitsmatrix, 65
- Einschrittverfahren, 246
- Elementarereignis, 409
- Elemente, 15
- Elemente einer Matrix, 53
- Ellipse, 25
- elliptische Koordinaten, 361
- elliptische pDgl, 499
- elliptischer Punkt, 300
- empirischen Median, 469
- endlich dimensional, 106
- Endpunkt, 278
- Epizykloide, 198
- Ereignis, 409, 454
- Erfolgswahrscheinlichkeit, 422
- Ergebnis, 409
- Ergebnismenge, 409
- erwartungstreu, 436
- erwartungstreuer Schätzer, 436
- Erwartungswert, 426, 461
- Erwartungswert der
 - Binomialverteilung, 427
- Erwartungswert der
 - hypergeometrischen Verteilung, 428
- erzeugte σ -Algebra, 453
- ESV, 246
- Euklidische Norm, 109
- Euklidisches Skalarprodukt, 107
- Eulersche Gammafunktion, 196, 325
- Eulersche Zahl, 122
- EV, 78
- Evolute, 204
- EW, 78
- exakte Differentialgleichung, 380
- explizites Eulerverfahren, 244
- explizites Runge-Kutta-Verfahren, 248
- Exponentialfunktion, 122
- Exponentialverteilung, 456
- Extremstellen-Test, 305
- Extremum, 140
- Extremwert-Test, 269
- Extremwerttest, 144

- Fakultät, 23
- Faltung, 458
- Fehler erster Art, 449

- Fehler zweiter Art, 449
- Feldlinie, 319
- Fixpunkt, 148
- Fixpunktiteration, 148
- flacher Punkt, 300
- Flächeninhalt, 342, 356
- Flächeninhalt eines Dreieckes, 35
- Fluss, 362
- Folge, 115
- Formel von Bayes, 419
- Formel von der totalen
Wahrscheinlichkeit, 419
- Formel von Sarrus, 73
- Fortsetzung der Randwerte, 501
- Fourier-Koeffizienten, 480
- Fourier-Reihe, 480
- Fourier-transformierbar, 487
- Fourier-Transformierte, 325, 487
- 5-Punkte-Gleichung für den
allgemeinen Kegelschnitt, 78
- Fundamentalsystem, 388
- Funktion, 16
- Funktionaldeterminante, 369
- Funktionalgleichung, 197

- Gammafunktion, 325
- ganze Zahlen, 17
- Gasgleichung, 494
- Gaußsches Eliminationsverfahren,
58, 66
- Gauß-Formeln, 210
- Gaußscher Integralsatz, 372
- gDgl, 218
- gDgl n -ter Ordnung, 375
- Gebietsintegral, 343
- gemeinsame Verteilungsfunktion, 424
- geometrische Reihe, 255
- geometrische Summenformel, 22
- geometrische Verteilung, 423
- geometrische Vielfachheit, 79
- Gerade, 25, 198
- geschlossene Kurve, 334
- geschlossene Fläche, 353
- Geschwindigkeit, 135
- Gewichte einer Quadraturformel, 210
- gewöhnliche Differentialgleichung,
218
- gleich, 15, 16
- gleiche Matrizen, 54
- gleichmäßig Lipschitz-stetig, 377
- gleichmäßig stetig, 125
- Gleichverteilung, 411, 456
- Glieder einer Folge, 115
- Glieder einer Reihe, 255
- globales Maximum, 140, 305
- globales Minimum, 140, 305
- Grad, 126, 478
- Gradient, 287, 321
- Gradientenfeld, 335
- Gradientenverfahren, 295
- Graph, 24, 197, 284, 353
- Gravitationskraft, 320
- Grenzwert, 116, 285
- Grundwasserströmung, 496
- Gütefunktion, 450

- Hängebrücke, 174
- Hagen, G., 320
- halboffenes Intervall, 19
- Halbwertszeit, 163
- harmonische Reihe, 256
- Hauptnormalenvektor, 280
- Hauptsatz der Differential- und
Integralrechnung, 172
- Hauptteil eines Differentialoperators,
498
- Hauptvektor, 94
- Hesse-Matrix, 299
- Hesse-Normalform, 42
- höhere Ableitungen, 139
- homogene Differentialgleichung, 229,
386
- homogene lineare gDgl, 225, 233
- homogene Wärmeleitungsgleichung
mit homogenen
Randbedingungen, 501
- homogenes LGS, 56
- homogenes quadratisches Polynom,
96
- Homogenität, 108
- Householder-Matrix, 90
- Hyperbel, 25
- hyperbolische pDgl, 499
- hyperbolischer Punkt, 300
- hypergeometrische Verteilung, 414
- Hypothese, 449
- Hypozykloide, 198

- Imaginärteil, 45
- implizites Eulerverfahren, 244
- implizites Runge-Kutta-Verfahren,
248
- indefinit, 101
- Induktionsanfang, 20
- Induktionsprinzip, 20
- Induktionsschritt, 21

- Infimum, 19
- inhomogene Differentialgleichung, 386
- inhomogene lineare gDgl, 225, 233
- inhomogene Wärmeleitungsgleichung mit homogenen Rand- und Anfangsbedingungen, 501
- inhomogenes LGS, 56
- injektiv, 16
- Injektivitätskriterium, 154
- Inklusion, 15
- innerer Punkt, 283
- instabil, 403
- Integral, 170
- Integration von Potenzreihen, 263
- integrierbar, 193
- integrierender Faktor, 382
- Inverse, 66
- inverse Fourier-Transformation, 487
- inverse Iteration von Wielandt, 84
- inverse Matrix, 66
- inverse
 - Rayleigh-Quotienten-Iteration, 86
- invertierbare Matrix, 66
- irrationale Zahlen, 17

- Jacobi-Determinante, 369
- Jacobi-Matrix, 315

- kartesische Basis, 33
- kartesisches Koordinatensystem, 24, 33
- Kessel, 151
- Kettenlinie, 174
- Kettenregel, 137, 293, 318
- klassisches Runge-Kutta-Verfahren, 249
- Knoten einer Quadraturformel, 210
- Koeffizient, 56
- Koeffizienten einer Potenzreihe, 260
- Koeffizienten eines
 - Differentialoperators, 498
- Koeffizientenmatrix, 56
- Kommutativgesetz, 28, 30, 102
- Kompabilitätsbedingung, 500, 508
- kompakt, 284
- kompakte Menge, 284
- Komplement, 16
- komplexe Fourier-Koeffizienten, 480
- komplexe Matrix, 53
- komplexe Zahl, 45
- Komposition, 16
- Konduktivität, 496
- konjugiert komplexe Zahl, 46
- Kontinuitätsgleichung der
 - Strömungsmechanik, 373
- Kontraktion, 379
- konvergente Folge, 116
- konvergente Reihe, 255
- Konvergenzkriterium für
 - uneigentliche Integrale, 195
- Konvergenzradius, 260
- konvex, 337
- Koordinatenebene, 33
- Koordinatengleichungen, 38
- Koordinateninvarianz, 322
- Koordinatentransformation, 360, 369
- Korrelationskoeffizient, 428
- Kovarianz, 428
- Kreis, 25, 198
- kritischer Bereich, 449
- kritischer Wert, 449
- Krümmung, 202, 280
- Krümmungskreis, 203
- Krümmungsvektor, 280
- Kugelkoordinaten, 371
- Kurve, 278, 330
- Kurvenintegral, 330, 333
- Kurvenstück, 278

- Länge, 279
- Länge, 200
- Lagrange-Funktion, 309
- Lagrange-Multiplikator, 309
- laminare Rohrströmung, 320
- Laplace-Experiment, 411
- Laplace-Operator, 294, 321
- Laplacescher
 - Wahrscheinlichkeitsraum, 411
- leere Menge, 15
- Leibniz-Kriterium, 257
- Leibnizsche Sektorformel, 207
- LGS, 56
- Likelihood-Funktion, 435
- Likelihood-Quotient, 469
- Likelihood-Quotienten Test, 470
- Limes, 116
- linear abhängig, 104
- linear approximierbar, 289, 315
- linear unabhängig, 104
- lineare Abbildung, 111
- lineare gDgl, 225
- lineare Hülle, 104
- linearer Differentialoperator, 498

- linearer Differentialoperator mit konstanten Koeffizienten, 498
- lineares Gleichungssystem, 56
- Linearität, 107
- Linearkombination, 103
- linksseitiger Grenzwert, 125
- Lipschitz-Konstante, 377
- Lipschitz-stetig, 377
- Ljapunov-stabil, 403
- Lösung einer gDgl, 219
- Logarithmusfunktion, 131
- lokal lösbar, 219
- lokale Lösung, 219
- lokaler Verfahrensfehler, 247
- lokales Extremum, 305
- lokales Maximum, 140, 305
- lokales Minimum, 140, 305
- LR-Zerlegung, 69

- Majorantenkriterium, 257
- Massenmittelpunkt, 332, 347
- Matrix, 53
- Maximum, 140
- Maximum-Likelihood Ansatz, 434
- Maximum-Likelihood Schätzung, 436
- Median, 469
- Membrangleichung, 493
- Menge, 15
- messbar, 366, 457
- Methode der kleinsten Quadrate, 467
- metrische Fundamentalgrößen, 352
- Minimum, 140
- minus unendlich, 19
- Mittelpunktsregel, 210
- Mittelwert, 437
- Mittelwertsatz, 143, 331
- Mittelwertsatz der Integralrechnung, 171
- mittlerer quadratischer Fehler, 438
- Momentengleichung, 39
- monoton fallende Folge, 116
- monoton wachsende Folge, 116
- monotone Folge, 116
- Multinomialkoeffizient, 415

- Nabla-Operator, 323
- nach oben beschränkte Menge, 19
- nach unten beschränkte Menge, 19
- natürliche Basis, 55
- natürliche Zahlen, 17
- negativ definit, 101
- negativ semidefinit, 101
- negative Binomialverteilung, 423

- Neilsche Parabel, 204
- Nennerpolynom, 128
- Neumann-Randbedingung, 493
- Newton, I., 320
- Newton-Cotes-Formeln, 210
- Newtonsches Abkühlungsgesetz, 163
- Newtonverfahren, 150, 316
- nichtorthogonale Zerlegung, 36
- Niveaumenge, 284
- Norm, 108
- Normale, 199
- Normalengleichung, 42
- Normalenvektor, 42, 352
- Normalform einer Quadrik, 97
- Normalverteilung, 456
- Normaxiome, 108
- normiert, 109
- normierte Form, 443
- normierte Vektoren, 89
- notwendige Charakterisierung lokaler Extrema, 305
- n -te Wurzel, 49, 156
- Nullelement, 103
- Nullfolge, 116
- Nullmatrix, 55
- Nullmenge, 342, 366
- Nullpunkt, 24
- Nullvektor, 28, 55, 103

- obere Dreiecksmatrix, 69
- obere Schranke, 19
- Oberflächenintegral, 357
- offen, 283
- offene Menge, 283
- offener Ball, 283
- offenes Intervall, 19
- Ordnung, 247
- Ordnung einer Quadraturformel, 210
- orientierbar, 364
- orthogonal, 30, 109
- orthogonale Matrix, 89
- orthogonale Vektoren, 89
- orthogonale Zerlegung, 30
- Orthogonalitätsbeziehungen, 180
- Orthogonalitätstest, 30
- Orthonormalsystem, 109
- Ortsvektor, 33

- Parabel, 25
- parabolische pDgl, 499
- parabolischer Punkt, 300
- Parallelepiped, 32
- Parallelfäche, 32

- Parameter, 37, 197
- Parameterdarstellung, 37, 41, 197, 351
- Parameterintegral, 325
- Parameterintervall, 197
- Parameterlinie, 352
- Parseval-Gleichung, 484, 490
- Partialbruch, 183
- Partialsomme, 255, 479
- partiell differenzierbar, 287
- partielle Ableitung, 286
- partielle Differentialgleichung m -ter Ordnung, 498
- partielle Differentialgleichung zweiter Ordnung, 493
- partikuläre Lösung, 219
- Pascal-Verteilung, 423
- pDgl, 493
- pDgl m -ter Ordnung, 498
- Periode, 477
- periodisch, 477
- Pivotsuche, 58
- Plattengleichung, 493
- plus unendlich, 19
- Poiseuille, J. L. M., 320
- Poisson-Approximation, 446
- Poisson-verteilt, 445
- Poissongleichung, 493
- Poissonsches Gasgesetz, 165
- Pol, 129
- Polarachse, 205
- Polardarstellung, 205
- Polarkoordinaten, 204, 315, 361
- Polstelle, 129
- Polynom, 79, 126
- positiv definit, 101
- positiv semidefinit, 101
- Positivität, 107, 108
- Potential, 335
- Potenz, 23
- Potenzmenge, 409
- Potenzmethode, 82
- Potenzreihe, 260
- Produkt, 16
- Produkt von
 - Wahrscheinlichkeitsräumen, 421
- Produktformel für
 - Wahrscheinlichkeiten, 418
- Produktregel, 136
- properly posed, 219
- Punkt-Richtungs-Gleichung, 37
- quadratische Form, 96
- quadratische Konvergenz, 152
- quadratisches Polynom, 96
- Quadraturformel, 209
- Quadrik, 97
- Quanti, 472
- quasilinear, 316
- quasilineare Abbildung, 316
- quasilinearer Differentialoperator, 497
- Quelldichte, 321
- Quotientenkriterium, 258
- Quotientenregel, 136
- radioaktiver Zerfall, 163
- räumliche Kugelkoordinaten, 294
- Rand, 283, 352
- Randbedingung, 493
- Randpunkt, 283
- Rang, 57
- rationale Funktion, 128
- rationale Zahlen, 17
- Rayleigh-Quotienten-Iteration, 83
- Realteil, 45
- rechte Seite, 56
- Rechteckschwingung, 481
- rechtsseitiger Grenzwert, 125
- reelle Fourier-Koeffizienten, 481
- reelle Matrix, 53
- reelle Zahlen, 17
- Regeln von de l'Hôpital, 146
- Regressionsgerade, 467, 468
- regulär, 366
- reguläre Matrix, 66
- reguläre Parameterdarstellung, 200
- reguläres Flächenstück, 351
- reguläres Kurvenstück, 278
- Reihe, 255
- rein quadratisch, 96
- rekursive Definition, 22
- Ricatti-Differentialgleichung, 232
- Richtungsableitung, 292
- Richtungsfeld, 220
- Riemann-Lemma, 484
- Riemann-messbar, 342, 366
- Romberg-Verfahren, 214
- Rotation, 321
- Ruhepunkt, 406
- Runge-Kutta-Verfahren, 248
- sachgemäß gestellt, 219
- Satz über die Umkehrfunktion, 154
- Satz über implizite Funktionen, 302
- Satz von Gauß, 350

- Satz von Green, 349
 Satz von Moivre-Laplace, 443
 Satz von Picard-Lindelöf, 378
 Satz von Poincaré, 337
 Satz von Rolle, 143
 Satz von Stokes, 364
 Schätzer, 435
 schiefsymmetrische Matrix, 66
 Schmidtsches
 Orthogonalisierungsverfahren,
 110
 Schmiegeebene, 280
 Schmiegequadratik, 300
 Schursche Normalform, 93
 schwaches Gesetz der großen Zahlen,
 432
 Schwerpunkt, 332, 347
 SDIRK-Verfahren, 250
 Seilkurve, 174
 Sektorfläche, 206
 senkrecht, 30
 Separationsansatz, 502
 sicheres Ereignis, 409
 σ -Algebra, 453
 Simpsonregel, 210
 singuläre Lösung, 219
 singuläre Matrix, 66
 Sinus, 25
 Sinus hyperbolicus, 167
 Skalarenfeld, 319
 skalares Bogenelement, 279
 Skalarprodukt, 30, 107
 Skalarproduktaxiome, 107
 Spaltenvektor, 53
 Spat, 32
 Spatprodukt, 32
 spezielle Lösung, 219
 Sphäre, 354
 Spur, 278
 Stabdiagramm, 424
 stabil, 403
 Stabilität, 252
 Stabilität beliebiger Lösungen für
 Störungen linearer AWP mit
 konstanten Koeffizienten, 404
 Stabilität der Nulllösung für
 Störungen linearer AWP mit
 konstanten Koeffizienten, 405
 Stabilität von Ruhepunkten
 autonomer AWP, 406
 Stabilitätsgebiet, 252
 Stabilitätskriterium für lineare AWP
 mit konstanten Koeffizienten,
 404
 Stammfunktion, 171, 335
 Standardabweichung, 428, 462
 standardisierte Form, 443
 stark diagonal implizite
 Runge-Kutta-Verfahren, 250
 starre Drehung, 319
 stationärer Punkt, 141, 305
 statische Momente, 332
 sternförmig, 337
 stetig, 123, 285
 stetig partiell differenzierbar, 287
 Stichprobenraum, 435
 Stirlingsche Formel, 440
 Streuung, 428
 stückweise reguläre Fläche, 353
 stückweise stetig, 169
 Stufe eines Hauptvektors, 94
 Stufe eines Runge-Kutta-Verfahrens,
 248
 Substitutionsregel, 178, 180
 Superposition, 501, 508
 Supremum, 19
 surjektiv, 16
 Symmetrie, 107
 symmetrische Matrix, 66
 System gewöhnlicher
 Differentialgleichungen n-ter
 Ordnung, 375
 T -periodisch, 477
 t -Test, 472
 t -Verteilung, 472
 Tangens, 129
 Tangens hyperbolicus, 167
 Tangente, 134, 199
 Tangentenvektor, 279
 Tangentialebene, 296, 352
 Tangentialvektor, 278
 Taylor-Formel, 268, 298
 Taylor-Polynom, 268, 298
 Taylor-Reihe, 269
 Teilfolge, 116
 Teilmenge, 15
 Test, 449
 Test für die totale
 Differenzierbarkeit, 290
 Teststatistik, 449
 t_{n-1} -Verteilung, 472
 Torsion, 281
 Torsionsvektor, 281

- Torus, 208, 354
- total differenzierbar, 289, 315
- Trägheitsmomente, 332
- Transformationsformel, 360
- Transformationsformel für
Volumenintegrale, 369
- Transformationsformeln für
Drehungen, 26
- transponierte Matrix, 65
- Trapezregel, 210, 245
- trigonometrische Reihe, 479
- trigonometrisches Polynom, 478
- triviale Linearkombination, 103
- triviale Lösung, 56

- Umkehrabbildung, 16
- unabhängig, 420, 426, 458
- unbestimmtes Integral, 173
- unbiased, 436
- uneigentlich integrierbar, 193
- uneigentliches Integral, 193
- uneigentliches Parameterintegral,
325
- unkorreliert, 428
- unmögliches Ereignis, 409
- untere Dreiecksmatrix, 69
- untere Schranke, 19
- Unterraum, 103
- Untervektorraum, 103
- Urbildraum, 277
- Ursprung, 24, 33

- Varianz, 428, 462
- Varianz der Binomialverteilung, 430
- Varianz der hypergeometrischen
Verteilung, 431
- Variation der Konstanten, 227, 236
- Vektor, 28, 103
- Vektorfeld, 319
- vektorielles Bogenelement, 279
- Vektorprodukt, 31
- Vektorraum, 102
- Vektorraumaxiome, 102
- Vereinigung, 15
- Verfahren von Crank-Nicolson, 245
- Verfahren von Heron, 153
- Verfahrensfunktion, 247
- Vertauschbarkeitskriterium für
partielle Ableitungen, 288
- Verteilungsfunktion, 424, 455, 457
- Verteilungsfunktion der
Standard-Normalverteilung, 443
- Verwerfen der Hypothese, 449

- Verwerfungsbereich, 449
- Vielfachheit, 128
- vollständige Induktion, 20
- vollständige Lösung, 219
- Vollständigkeitsaxiom, 19
- Volumen, 366
- Volumenintegral, 367

- Wärmeleitungsgleichung, 494
- Wärmeleitungskern, 507
- Wahrscheinlichkeit, 409
- Wahrscheinlichkeitsfunktion, 410
- Wahrscheinlichkeitsmaß, 409
- Wahrscheinlichkeitsmaß, 454
- Wahrscheinlichkeitsraum, 409, 454
- Wahrscheinlichkeitsverteilung, 409
- wegunabhängig, 335
- Wellengleichung, 496
- Wellengleichungskern, 513
- Wendelfläche, 355
- Wendepunkt, 144
- Wendepunkttest, 145
- Wertebereich, 16
- Winkel, 29
- Wirbeldichte, 321
- Wronski-Determinante, 237, 388

- x -Achse, 24, 33

- y -Achse, 24, 33

- z -Achse, 33
- Zählerpolynom, 128
- Zahlenfolge, 115
- Zahlengrade, 17
- Zeilenvektor, 53
- zentrales Kraftfeld, 320
- Zentralfeld, 340
- Zentralkraft, 336
- Zufallsvariable, 423, 457
- zusammengesetzte
Mittelpunktsregel, 211
- zusammengesetzte Quadraturformel,
211
- zusammengesetzte Simpsonregel, 211
- zusammengesetzte Trapezregel, 211
- zweiseitig, 364
- zweiseitige Fläche, 363
- Zwischenwertsatz, 124
- Zykloide, 198
- Zylinderkoordinaten, 370