

Mathematische Rechenmethoden

Stefan Weinzierl

9. Februar 2024

Inhaltsverzeichnis

1	Einführung	5
1.1	Literatur	5
2	Schreibweisen	6
3	Zahlen	9
3.1	Die natürlichen Zahlen	9
3.2	Die ganzen Zahlen	11
3.3	Die rationalen Zahlen	12
3.4	Die reellen Zahlen	12
3.5	Die komplexen Zahlen	14
4	Lineare Algebra	17
4.1	Vektorräume	17
4.2	Skalarprodukte	19
4.3	Das Kreuzprodukt	21
4.4	Vektoren in der Physik: Vierervektoren	23
4.5	Lineare Gleichungssysteme	26
4.6	Matrizen	30
4.7	Spuren und Determinanten von quadratischen Matrizen	33
4.8	Berechnung der inversen Matrix	36
4.9	Eigenwerte und Eigenvektoren*	38
4.10	Das Schmidtsche Orthonormierungsverfahren	43
5	Analysis	44
5.1	Folgen	44
5.2	Reihen	45
5.3	Funktionen und Stetigkeit	50
5.3.1	Rationale Funktionen	52
5.3.2	Trigonometrische Funktionen	55
5.3.3	Hyperbolische Funktionen	56
5.4	Differentialrechnung	58
5.4.1	Die Regeln von l'Hospital	60
5.5	Integralrechnung	61
5.5.1	Uneigentliche Integrale	65
5.6	Taylorreihen	67
5.7	Orthogonale Polynome	69
5.7.1	Legendre-Polynome	70
5.7.2	Tschebyscheff-Polynome	71
5.7.3	Laguerre-Polynome	71
5.7.4	Hermite-Polynome	72

6	Differentialgleichungen	73
6.1	Reduktion einer Differentialgleichung höherer Ordnung auf ein System erster Ordnung	76
6.2	Existenz und Eindeutigkeit von Lösungen	77
6.3	Elementare Lösungsmethoden	78
6.4	Lineare Differentialgleichungen	80
6.4.1	Lösung der homogenen Gleichung	81
6.4.2	Lösung der inhomogenen Gleichung	81
6.4.3	Systeme linearer Differentialgleichungen	84
6.4.4	Lineare Differentialgleichungen höherer Ordnung	87
6.5	Lineare Differentialgleichungen mit konstanten Koeffizienten*	88
6.5.1	Systeme von linearen Differentialgleichungen mit konstanten Koeffizienten*	88
6.5.2	Die lineare Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten*	89
6.5.3	Der harmonische Oszillator*	92
6.6	Exakte Differentialgleichungen und integrierende Faktoren*	95
7	Die Eulersche Gamma-Funktion*	98
8	Asymptotisches Verhalten*	101
9	Fehlerrechnung*	103
9.1	Fehlerfortpflanzung*	106
9.2	Fitten von Parametern*	108
10	Differential- und Integralrechnung in mehreren Dimensionen	109
10.1	Topologische Grundbegriffe*	109
10.2	Konvergenz in metrischen Räumen*	112
10.3	Stetigkeit*	113
10.4	Kurven	114
10.5	Funktionen in mehreren Variablen	117
10.6	Vektorfelder	123
10.7	Integralrechnung in mehreren Variablen	132
10.7.1	Die Substitutionsregel als Spezialfall der Transformationsformel	140
10.7.2	Polarkoordinaten im \mathbb{R}^2	140
10.7.3	Zylinderkoordinaten im \mathbb{R}^3	141
10.7.4	Kugelkoordinaten im \mathbb{R}^3	142
10.7.5	Der Maßtensor	143
10.8	Integration auf Mannigfaltigkeiten	144
10.8.1	Definition einer Mannigfaltigkeit*	145
10.8.2	Untermannigfaltigkeiten des \mathbb{R}^{n*}	148
10.8.3	Die Sätze von Gauß und Stokes	151

10.8.4	Differentialformen*	155
10.9	Variationsrechnung	164
11	Partielle Differentialgleichungen	167
11.1	Distributionen	167
11.2	Die Fourier-Transformation	170
11.3	Beispiele partieller Differentialgleichungen	172
11.3.1	Die Potentialgleichung	172
11.3.2	Die Schwingungsgleichung	175
11.3.3	Die Wärmeleitungsgleichung	177

1 Einführung

Dieses Skript basiert zum einen auf einer zweisemestrigen Vorlesung über “Mathematische Rechenmethoden”, die ich im akademischen Jahr 2009/10 an der Universität Mainz gehalten habe, als auch auf einer einsemestrigen Vorlesung mit identischen Titel, die ich im Wintersemester 2013/14 ebenfalls an der Universität Mainz gehalten habe. Klarerweise ist der Stoff einer einsemestrigen Vorlesung gegenüber einer zweisemestrigen Vorlesung zu reduzieren. Themen, die ich nur im Rahmen einer zweisemestrigen Vorlesung behandelt habe, sind mit einem Stern versehen.

Die Behandlung des Satzes von Stokes mittels Differentialformen ist elegant, geht aber auch über das Niveau einer normalen zweisemestrigen Vorlesung hinaus. Dieser Abschnitt wurde im Sommersemester 2010 als Alternativprogramm zu einem zeitgleich stattfindenden sportlichen Großereignis aufgenommen.

1.1 Literatur

- W. Schäfer, K. Georgi, G. Trippler, Mathematik Vorkurs, Teubner 1997
- S. Großmann, Mathematischer Einführungskurs, Teubner 2004
- K. Weltner, Mathematik für Physiker, Springer 2001
- R. Brauner, F. Geiß, Abiturwissen Mathematik, Fischer 2006
- P. van Dongen, Einführungskurs Mathematik und Rechenmethoden, Springer 2015

Skripten:

- S. Groote, Mathematischer Vorkurs,

<https://kodu.ut.ee/~groote/vorkurs>

- K. Hefft, Mathematischer Vorkurs,

<https://www.thphys.uni-heidelberg.de/~hefft/vk1/>

Lehrbücher:

- O. Forster, Analysis 1-3, Vieweg
- G. Fischer, Lineare Algebra, Vieweg
- W. Fischer, I. Lieb, Funktionentheorie, Vieweg
- K.-H. Goldhorn, H.-P. Heinz, Mathematik für Physiker

2 Schreibweisen

$\{a, b, c\}$: Menge der Elemente a , b , und c .

Bemerkung: Die Ordnung spielt keine Rolle: $\{a, b, c\} = \{b, a, c\}$

$a \in A$: a ist ein Element der Menge A .

$A \subset B$: Die Menge A ist eine Teilmenge der Menge B .

Vereinigung: $A \cup B$ enthält alle Elemente sowohl aus A als auch aus B .

$\{a, b, c\} \cup \{c, d, e\} = \{a, b, c, d, e\}$.

Durchschnitt: $A \cap B$ enthält alle Elemente die sowohl in A als auch in B enthalten sind.

$\{a, b, c\} \cap \{c, d, e\} = \{c\}$.

Differenzmenge: $A \setminus B$ enthält alle Elemente, die in A enthalten sind, die aber nicht in B enthalten sind.

$\{a, b, c, d\} \setminus \{b, c\} = \{a, d\}$

$A \times B$: Produktmenge, dies ist die Menge aller geordneten Paare (a, b) , wobei $a \in A$ und $b \in B$ gilt.

$[a, b]$: Intervall, die Grenzen sind im Intervall enthalten: $a \in [a, b]$, $b \in [a, b]$

$]a, b[$: Intervall, die Grenzen sind im Intervall nicht enthalten: $a \notin [a, b]$, $b \notin [a, b]$

Analog: $[a, b[$ und $]a, b]$.

Logisch und: \wedge

Logisch oder: \vee

Negation: \neg

\exists : Es existiert

\forall : Für alle

∞ : Symbol für Unendlich.

Σ : Summenzeichen

$$\sum_{j=1}^n a_j = a_1 + a_2 + a_3 + \dots + a_{n-1} + a_n.$$

\prod : Produktzeichen

$$\prod_{j=1}^n a_j = a_1 \cdot a_2 \cdot a_3 \cdot \dots \cdot a_{n-1} \cdot a_n.$$

$n!$: Fakultät. Bemerkung $0! = 1, 1! = 1$.

Binomialkoeffizient:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

lim: Grenzwert für den Fall, daß sich x dem Wert a annähert.
 $x \rightarrow a$

Ableitung: Sei $f(x)$ eine Funktion von x .

$$f'(x) = \frac{df(x)}{dx} = \frac{d}{dx}f(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{(x+h) - x}.$$

Integral:

$$\int_a^b f(x) dx.$$

In der Mathematik und Physik verwendet man oft griechische Buchstaben:

α	alpha	β	beta	γ	gamma	δ	delta
ϵ oder ε	epsilon	ζ	zeta	η	eta	θ oder ϑ	theta
ι	iota	κ	kappa	λ	lambda	μ	mu
ν	nu	ξ	xi	\omicron	o	π oder ϖ	pi
ρ oder ϱ	rho	σ oder ς	sigma	τ	tau	υ	upsilon
ϕ oder φ	phi	χ	chi	ψ	psi	ω	omega

Griechische Großbuchstaben:

A	Alpha	B	Beta	Γ	Gamma	Δ	Delta
E	Epsilon	Z	Zeta	H	Eta	Θ	Theta
I	Iota	K	Kappa	Λ	Lambda	M	Mu
N	Nu	Ξ	Xi	O	O	Π	Pi
P	Rho	Σ	Sigma	T	Tau	Υ	Upsilon
Φ	Phi	X	Chi	Ψ	Psi	Ω	Omega

Aus dem hebräischen Alphabet verwendet man

\aleph Aleph.

Üblicherweise wird dieser Buchstabe zur Beschreibung der Mächtigkeit der natürlichen Zahlen verwendet.

Aus dem russischen Alphabet verwendet man

□ Sha.

Üblicherweise verwendet man dieses Zeichen zur Notation für das Shuffle-Produkt. Dies ist ein spezielles Produkt, wird aber in dieser Vorlesung nicht vorkommen.

Beispiel: Mit diesen Schreibweisen können wir nun die Riemannsche Definition des Integrals kompakt angeben:

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{j=1}^n f(\xi_j) \Delta x_j, \quad \Delta x_j = x_j - x_{j-1}, \quad x_0 = a, \quad x_n = b, \quad \xi_j \in [x_{j-1}, x_j].$$

Übersetzt bedeutet diese Zeile, daß wir das Intervall $[a, b]$ in n Unterintervalle $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ mit $x_0 = a$ und $x_n = b$ unterteilen. In jedem Unterintervall nehmen wir einen (beliebigen) x -Wert $\xi_j \in [x_{j-1}, x_j]$ und nähern das Integral über dieses Unterintervall durch ein Rechteck an. Die Fläche des Rechtecks ist $f(\xi_j) \Delta x_j$. Wählen wir die Unterteilungen so, daß jedes $\Delta x_j \rightarrow 0$ für $n \rightarrow \infty$, so ergibt sich das Integral als Grenzwert $n \rightarrow \infty$. (Wir werden später das Riemann-Integral detaillierter mit Ober- und Untersummen definieren.)

3 Zahlen

3.1 Die natürlichen Zahlen

\mathbb{N} : Die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, 4, \dots\}$.

Die **Axiome von Peano** für die natürlichen Zahlen:

- (P1) Die Zahl 1 ist eine natürliche Zahl.
- (P2) Falls a eine natürliche Zahl, so ist die nachfolgende Zahl $a+1$ ebenfalls eine natürliche Zahl.
- (P3) Die natürlichen Zahlen sind die minimale Menge, welche die ersten beiden Axiome erfüllt.

\mathbb{N}_0 : Die natürlichen Zahlen mit der Null $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}$.

Der **Induktionsbeweis**: Man ist oft in der Situation eine Aussage der Form

$$f(n) = g(n)$$

für alle $n \in \mathbb{N}$ beweisen zu müssen. Hier bietet sich der Induktionsbeweis an. Der Induktionsbeweis verläuft in zwei Teilen:

1. Induktionsanfang: Im ersten Teil zeigt man zunächst, daß die Behauptung für $n = 1$ richtig ist.
2. Induktionsschritt: Im zweiten Teil nimmt man an, daß die Behauptung für $(n - 1)$ richtig ist und zeigt, daß sie dann auch für n richtig ist.

Man sieht leicht, daß dies die allgemeine Aussage beweist: Für $n = 1$ wird die Aussage im ersten Teil bewiesen. Für $n = 2$ können wir dann verwenden, daß die Aussage für $n = 1$ richtig ist. Somit liegt die Voraussetzung für den zweiten Teil vor und es folgt aufgrund des zweiten Teils die Richtigkeit für $n = 2$. Diese Argumentation läßt sich nun fortsetzen: Da die Aussage für $n = 2$ richtig ist, muß sie aufgrund des zweiten Teils auch für $n = 3$ richtig sein, usw..

Beispiel: Für jede natürliche Zahl n ist die folgende Behauptung zu zeigen:

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

Induktionsanfang: Für $n = 1$ haben wir

$$\text{linke Seite : } \sum_{j=1}^1 j = 1.$$

$$\text{rechte Seite : } \frac{1(1+1)}{2} = 1.$$

Induktionsschritt: Wir dürfen nun annehmen, daß die Behauptung für $n-1$ richtig ist, und müssen zeigen, daß sie dann auch für n gilt. In unserem Fall:

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

Wir diskutieren noch eine Erweiterung, in dem man im Induktionsschritt verwendet, daß die Aussage sowohl für $(n-1)$ als auch für $(n-2)$ richtig ist. In diesem Fall muß im Induktionsanfang die Aussage für $n=1$ und $n=2$ gezeigt werden. Dies sieht man wie folgt: Ist die Aussage für $n=1$ und $n=2$ richtig, so folgt aus dem Induktionsschritt die Richtigkeit für $n=3$.

Beispiel: Die Fibonacci-Zahlen sind definiert durch

$$f_1 = 1, \quad f_2 = 1, \quad f_n = f_{n-1} + f_{n-2}, \quad (n \geq 3).$$

Behauptung: Für $n \in \mathbb{N}$ gilt

$$f_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right].$$

Beweis: Im Induktionsanfang zeigen wir die Behauptung für $n=1$ und $n=2$. Für $n=1$ gilt

$$\frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right) - \left(\frac{1-\sqrt{5}}{2} \right) \right] = \frac{1}{\sqrt{5}} \left[\frac{\sqrt{5}}{2} + \frac{\sqrt{5}}{2} \right] = 1 = f_1.$$

Für $n=2$ gilt

$$\frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^2 - \left(\frac{1-\sqrt{5}}{2} \right)^2 \right] = \frac{1}{\sqrt{5}} \left[\frac{1+2\sqrt{5}+5}{4} - \frac{1-2\sqrt{5}+5}{4} \right] = 1 = f_2.$$

Im Induktionsschritt berechnen wir $f_{n-1} + f_{n-2}$ und verwenden die Induktionsannahme sowohl für f_{n-1} als auch für f_{n-2} :

$$\begin{aligned} f_n &= f_{n-1} + f_{n-2} \\ &= \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n-1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n-1} \right] + \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n-2} - \left(\frac{1-\sqrt{5}}{2} \right)^{n-2} \right] \\ &= \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n-2} \left(\frac{1+\sqrt{5}}{2} + 1 \right) - \left(\frac{1-\sqrt{5}}{2} \right)^{n-2} \left(\frac{1-\sqrt{5}}{2} + 1 \right) \right] \\ &= \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n-2} \left(\frac{1+\sqrt{5}}{2} \right)^2 - \left(\frac{1-\sqrt{5}}{2} \right)^{n-2} \left(\frac{1-\sqrt{5}}{2} \right)^2 \right] \\ &= \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]. \end{aligned}$$

3.2 Die ganzen Zahlen

\mathbb{Z} : Die ganzen Zahlen $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$.

Die ganzen Zahlen bilden bezüglich der Addition eine Gruppe.

Definition einer Gruppe: Wir betrachten eine nicht-leere Menge G mit einer Verknüpfung \circ , d.h. eine Abbildung $\circ : G \times G \rightarrow G$. Man nennt das Paar (G, \circ) eine Gruppe, falls die folgenden Axiome gelten:

- (G1) \circ ist assoziativ: $a \circ (b \circ c) = (a \circ b) \circ c$
- (G2) Es gibt ein links-neutrales Element : $e \circ a = a$ für alle $a \in G$
- (G3) Zu jedem $a \in G$ gibt es ein links-inverses Element $a^{-1} : a^{-1} \circ a = e$

Eine Gruppe (G, \circ) nennt man kommutativ oder Abelsch, falls $a \circ b = b \circ a$.

Bemerkung: Falls (G, \circ) eine Gruppe ist, dann ist das links-neutrale Element identisch mit dem recht-neutralen Element. Ebenso sind links- und rechts-inverses Element identisch.

Bemerkung: In vielen praktischen Fällen ist zunächst eine Menge G und eine Verknüpfung $\circ : G \times G \rightarrow X$ gegeben, wobei X eine Menge ist, die G enthält: $G \subset X$. Um zu zeigen, daß (G, \circ) eine Gruppe bildet, zeigt man zunächst, daß die Verknüpfung abgeschlossen ist, d.h.

$$a \circ b \in G, \quad \forall a, b, \in G.$$

Danach zeigt man dann die einzelnen Gruppenaxiome.

Ein **Ring** ist eine nicht-leere Menge R mit zwei Verknüpfungen, die üblicherweise als $+$ und \cdot geschrieben werden, so daß

- (R1) $(R, +)$ ist eine Abelsche Gruppe.
- (R2) (R, \cdot) ist assoziativ: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- (R3) Es gelten die Distributivgesetze:

$$\begin{aligned} a \cdot (b + c) &= (a \cdot b) + (a \cdot c) \\ (a + b) \cdot c &= (a \cdot c) + (b \cdot c) \end{aligned}$$

Die ganzen Zahlen \mathbb{Z} bilden einen Ring.

3.3 Die rationalen Zahlen

\mathbb{Q} : Die rationalen Zahlen.

$$\mathbb{Q} = \left\{ r \mid r = \frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

Die rationalen Zahlen sind bezüglich der Division abgeschlossen. Sie bilden einen Körper.

Eine nicht-leere Menge K mit zwei Verknüpfungen $+$ und \cdot nennt man **Körper**, falls gilt:

- (K1) $(K, +)$ ist eine Abelsche Gruppe.
- (K2) $(K \setminus \{0\}, \cdot)$ ist eine Abelsche Gruppe.
- (K3) Es gelten die Distributivgesetze:

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

$$(a + b) \cdot c = (a \cdot c) + (b \cdot c)$$

Bruchrechnen:

$$\frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1 \cdot p_2}{q_1 \cdot q_2}, \quad \frac{p_1}{q_1} / \frac{p_2}{q_2} = \frac{p_1 \cdot q_2}{q_1 \cdot p_2}, \quad \frac{c \cdot p_1}{c \cdot q_1} = \frac{p_1}{q_1}, \quad \frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 \cdot q_2 + p_2 \cdot q_1}{q_1 \cdot q_2}.$$

Für Potenzen schreiben wir

$$a^n = \underbrace{a \cdot a \cdot a \cdot \dots \cdot a}_{n \text{ mal}}.$$

Rechnen mit Potenzen:

$$\begin{aligned} a^n \cdot b^n &= (a \cdot b)^n, & \frac{a^n}{b^n} &= \left(\frac{a}{b}\right)^n, \\ a^n \cdot a^m &= a^{n+m}, & \frac{a^n}{a^m} &= a^{n-m}, \\ (a^n)^m &= a^{n \cdot m}. \end{aligned}$$

3.4 Die reellen Zahlen

\mathbb{R} : Die reellen Zahlen.

Die reellen Zahlen bilden einen Körper.

Alle rationalen Zahlen sind in den reellen Zahlen enthalten. Darüberhinaus enthält \mathbb{R} Zahlen, die nicht rational sind. Diese nennt man irrational. So ist zum Beispiel $\sqrt{2}$ eine irrationale Zahl. $\sqrt{2}$ ist Lösung der Gleichung $x^2 = 2$. Zahlen, welche Lösungen einer algebraischen Gleichung

sind, nennt man algebraisch. Darüberhinaus enthält \mathbb{R} irrationale Zahlen, die keine Lösung einer algebraischen Gleichung sind. Solche Zahlen nennt man transzental. Die Kreiszahl π oder die Eulersche Konstante e sind transzental.

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen nennt man **Cauchy-Folge**, falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß

$$|a_n - a_m| < \varepsilon, \quad \forall n, m \geq N.$$

Vollständigkeitsaxiom: Jede Cauchy-Folge konvergiert.

Dies bedeutet, daß es zu jeder Cauchy-Folge eine reelle Zahl $a \in \mathbb{R}$ gibt, mit der Eigenschaft daß zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so daß

$$|a_n - a| < \varepsilon, \quad \forall n \geq N.$$

Die wesentliche Aussage ist, daß der Grenzwert a wieder in der betrachteten Menge (in diesem Fall \mathbb{R}) liegt. Dies ist nicht notwendigerweise der Fall. Als Gegenbeispiel betrachten wir die Menge aller positiven reellen Zahlen $\mathbb{R}_{>0}$ und die Folge $a_n = 1/n$. Dann gilt für jedes Folgenglied $a_n \in \mathbb{R}_{>0}$, aber

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0 \notin \mathbb{R}_{>0}.$$

Für die reellen Zahlen gilt das Vollständigkeitsaxiom, dies ist im wesentlichen die Aussage, daß der Zahlenstrahl keine "Löcher" enthält.

Darüberhinaus sind die reellen Zahlen **angeordnet**: Es sind gewisse Elemente als positiv ausgezeichnet ($x > 0$), so daß die folgenden Axiome erfüllt sind:

- (O1) Es gilt genau eine der Beziehungen $x < 0$, $x = 0$ oder $x > 0$.
- (O2) Aus $x > 0$ und $y > 0$ folgt $x + y > 0$.
- (O3) Aus $x > 0$ und $y > 0$ folgt $x \cdot y > 0$.

Man nennt eine Ordnung **archimedisch**, falls zu jedem $x > 0$ und $y > 0$ ein natürliche Zahl n existiert, so daß

$$n \cdot x > y.$$

Axiomatisch lassen sich die reellen Zahlen als ein Körper, der archimedisch angeordnet ist und in dem jede Cauchy-Folge konvergiert, charakterisieren.

3.5 Die komplexen Zahlen

\mathbb{C} : Die komplexen Zahlen.

Definition: Man definiert die **imaginäre Einheit** i als eine Lösung der Gleichung

$$x^2 = -1.$$

Die komplexen Zahlen \mathbb{C} sind die Menge

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

Wir führen eine Addition und Multiplikation auf \mathbb{C} ein: Sei $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$.

$$\begin{aligned}z_1 + z_2 &= (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2), \\z_1 \cdot z_2 &= (x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2)\end{aligned}$$

Mit dieser Addition und Multiplikation bilden die komplexen Zahlen einen Körper. Dieser Körper ist **algebraisch abgeschlossen**, d.h. die Nullstellen eines jeden Polynoms liegen in dem Körper. Der Körper ist allerdings nicht angeordnet. Das Vollständigkeitsaxiom gilt.

Sei $z = x + iy$ eine komplexe Zahl. Man bezeichnet x als **Realteil** und y als **Imaginärteil**.

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y.$$

Die zu $z = x + iy$ **konjugiert komplexe Zahl** ist

$$z^* = x - iy.$$

Es ist

$$z \cdot z^* = (x + iy) \cdot (x - iy) = x^2 + y^2.$$

Insbesondere ist $z \cdot z^*$ immer eine nicht-negative reelle Zahl. Als **Betrag** der komplexen Zahl bezeichnet man

$$|z| = \sqrt{zz^*} = \sqrt{x^2 + y^2}.$$

Da wir eine komplexe Zahl durch ein Paar (x, y) reeller Zahlen angeben, können wir uns eine komplexe Zahl als einen Punkt in der komplexen Zahlenebene vorstellen (siehe Abbildung 1).

Polardarstellung einer komplexen Zahl:

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

$\varphi \in [0, 2\pi[$ nennt man das **Argument** oder die Phase der komplexen Zahl. Man schreibt

$$\arg(z) = \varphi.$$

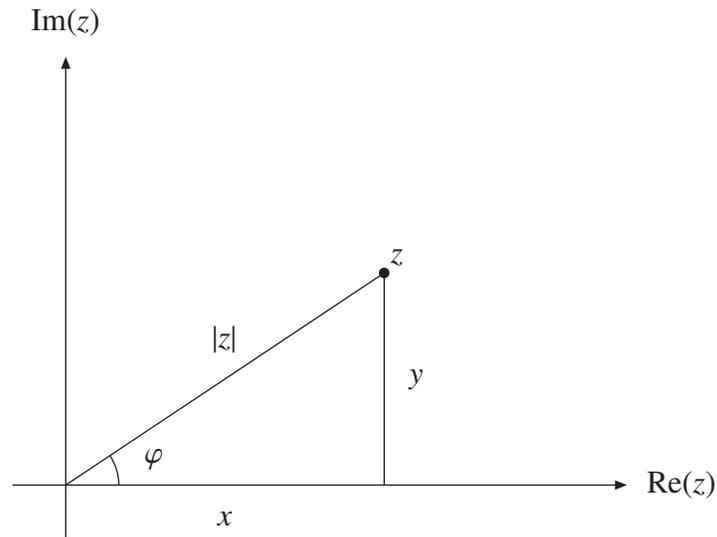


Abbildung 1: Die komplexe Zahlenebene.

Umrechnung: Wir betrachten zunächst den Fall der Umrechnung von der Normalform in die Polarform. Es sei (x, y) gegeben, so daß $z = x + iy$. Wir möchten $|z|$ und φ bestimmen. Der Betrag ist relativ einfach:

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

Zur Bestimmung des Arguments φ betrachten wir eine Fallunterscheidung. Für $x = 0$ setzen wir

$$\varphi = \begin{cases} \frac{\pi}{2}, & \text{für } y > 0, \\ \frac{3\pi}{2}, & \text{für } y < 0. \end{cases}$$

Wir können nun $x \neq 0$ voraussetzen. Dann gilt

$$\tan \varphi = \frac{y}{x}, \quad x \neq 0,$$

Die Auflösung der Gleichung $\tan \varphi = y/x$ nach φ ergibt

$$\begin{aligned} \varphi &= \arctan \frac{y}{x}, & \text{für } x > 0, y \geq 0, \\ \varphi &= \pi + \arctan \frac{y}{x}, & \text{für } x < 0, y \geq 0, \\ \varphi &= \pi + \arctan \frac{y}{x}, & \text{für } x < 0, y < 0, \\ \varphi &= 2\pi + \arctan \frac{y}{x}, & \text{für } x > 0, y < 0. \end{aligned}$$

Für die Umrechnung von der Polarform in die Normalform gilt

$$x = |z| \cos \varphi, \quad y = |z| \sin \varphi.$$

Rechenregeln:

$$\begin{aligned}z_1 - z_2 &= (x_1 + iy_1) - (x_2 + iy_2) = (x_1 - x_2) + i(y_1 - y_2), \\ \frac{z_1}{z_2} &= \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{(x_1 + iy_1) \cdot (x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} = \frac{(x_1x_2 + y_1y_2) + i(-x_1y_2 + y_1x_2)}{x_2^2 + y_2^2}.\end{aligned}$$

In der Polarform sind Multiplikation und Division besonders einfach:

$$\begin{aligned}z_1 \cdot z_2 &= |z_1| |z_2| [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)], \\ \frac{z_1}{z_2} &= \frac{|z_1|}{|z_2|} [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)].\end{aligned}$$

Diese Formeln beweist man mit Hilfe der Additionstheoreme für Sinus und Kosinus.

Aus der Formel für die Multiplikation folgt insbesondere:

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

Diese Gleichung wird auch als **Formel von Moivre** bezeichnet.

Wir werden später komplexwertige Funktionen kennenlernen. Im Vorgriff soll allerdings hier schon die **Formel von Euler** erwähnt werden

$$e^{i\varphi} = \cos \varphi + i \sin \varphi.$$

Diese Formel werden wir später mit Hilfe der Reihendarstellung der Funktionen \exp , \sin und \cos relativ einfach beweisen können.

4 Lineare Algebra

4.1 Vektorräume

Sei K ein Körper und $(V, +)$ eine Abelsche Gruppe. Weiter sei eine zusätzliche Verknüpfung gegeben, die man skalare Multiplikation nennt:

$$\begin{aligned} K \times V &\rightarrow V \\ (k, \vec{v}) &\rightarrow k \cdot \vec{v} \end{aligned}$$

V ist ein K -Vektorraum falls gilt:

- (V1) $(K, +, \cdot)$ ist ein Körper
- (V2) $(V, +)$ ist eine Abelsche Gruppe
- (V3) Es gelten die Distributivgesetze:

$$\begin{aligned} k \cdot (\vec{v}_1 + \vec{v}_2) &= (k \cdot \vec{v}_1) + (k \cdot \vec{v}_2) \\ (k_1 + k_2) \cdot \vec{v} &= (k_1 \cdot \vec{v}) + (k_2 \cdot \vec{v}) \end{aligned}$$

- (V4) Es gilt das Assoziativgesetz:

$$k_1 \cdot (k_2 \cdot \vec{v}) = (k_1 \cdot k_2) \cdot \vec{v}$$

Bemerkung: Bei $(k_1 \cdot k_2)$ ist die Multiplikation im Körper gemeint.

- (V5) Für die Eins gilt:

$$1 \cdot \vec{v} = \vec{v}$$

für alle $\vec{v} \in V$.

Als Grundkörper treten in der Physik fast immer \mathbb{R} oder \mathbb{C} auf. Beispiele für Vektorräume sind der \mathbb{R}^n (mit Grundkörper \mathbb{R}) und der \mathbb{C}^n (mit Grundkörper \mathbb{C}).

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, x_2, \dots, x_n \in \mathbb{R} \right\}.$$

Bei der Summe zweier Vektoren werden die Vektoren komponentenweise addiert:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix}.$$

Der **Nullvektor** $\vec{0}$ ist der Vektor, der in allen Komponenten eine Null enthält. Beispielsweise ist der Nullvektor $\vec{0} \in \mathbb{R}^3$ gegeben durch

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Allgemein gilt für den Nullvektor immer

$$\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}.$$

Bei der skalaren Multiplikation wird jede Komponente mit dem Skalar multipliziert:

$$3 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 12 \\ 15 \\ 18 \end{pmatrix}.$$

Wie bereits erwähnt, schreibt man die Elemente aus dem Vektorraum als Spaltenvektoren, so zum Beispiel:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

Ebenso ist die Schreibweise als Zeilenvektor gebräuchlich:

$$(x, y, z) \in \mathbb{R}^{3*}.$$

V^* bezeichnet den zu V **dualen Vektorraum** (falls V alle Spaltenvektoren enthält, so enthält V^* die Zeilenvektoren). Schreibweise: Man bezeichnet mit \vec{v}^T den zu \vec{v} **transponierten Vektor** (d.h. aus einem Spaltenvektor wird ein Zeilenvektor, und aus einem Zeilenvektor wird ein Spaltenvektor):

$$(x_1, x_2, \dots, x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}.$$

Definition: Vektoren, die in fast allen Komponenten eine Null haben, bis auf eine Komponente, in der sie eine Eins haben, spielen eine wichtige Rolle. Hat so ein Vektor in der i -ten Komponente eine Eins,

$$\vec{e}_i = \left(\underbrace{0, 0, \dots, 0}_{i-1}, 1, 0, \dots, 0 \right)^T,$$

so bezeichnet man diesen Vektor als den ***i*-ten Einheitsvektor**.

Definition: Seien n Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ gegeben. Folgt aus

$$\begin{aligned} a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_n \vec{v}_n &= \vec{0} \\ \Rightarrow a_1 = a_2 = \dots = a_n &= 0, \end{aligned}$$

so nennt man die Vektoren **linear unabhängig**. Anderfalls nennt man sie linear abhängig.

Definition: Sei V ein Vektorraum. Die maximale Anzahl linear unabhängiger Vektoren in V nennt man die **Dimension des Vektorraumes**. Eine Menge linearer unabhängiger Vektoren, die maximal ist, nennt man eine **Basis** von V .

Beispiel: \mathbb{R}^n und \mathbb{C}^n haben die Dimension n . Eine Basis von \mathbb{R}^n und \mathbb{C}^n ist zum Beispiel

$$\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}.$$

Man nennt diese Basis die Standardbasis.

4.2 Skalarprodukte

Wir betrachten zunächst den \mathbb{R}^n . Seien $\vec{x}, \vec{y} \in \mathbb{R}^n$. Die Komponentendarstellung der beiden Vektoren bezüglich der Standardbasis sei

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}.$$

Wir definieren das euklidische Standardskalarprodukt zwischen zwei Vektoren als die Abbildung

$$\begin{aligned} V \times V &\rightarrow \mathbb{R}, \\ \vec{x} \cdot \vec{y} &= x_1 y_1 + x_2 y_2 + \dots + x_n y_n. \end{aligned}$$

Allgemein ist ein euklidische Skalarprodukt eines reellen Vektorraumes eine Abbildung $V \times V \rightarrow \mathbb{R}$, welche die folgenden Bedingungen erfüllt:

- Linear in der ersten Komponente:

$$\begin{aligned} (\vec{x} + \vec{y}) \cdot \vec{z} &= \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}, \\ (\lambda \cdot \vec{x}) \cdot \vec{y} &= \lambda (\vec{x} \cdot \vec{y}). \end{aligned}$$

- Linear in der zweiten Komponente:

$$\begin{aligned} \vec{x} \cdot (\vec{y} + \vec{z}) &= \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}, \\ \vec{x} \cdot (\lambda \cdot \vec{y}) &= \lambda (\vec{x} \cdot \vec{y}). \end{aligned}$$

- Symmetrisch:

$$\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}.$$

- Positiv definit:

$$\vec{x} \cdot \vec{x} > 0, \text{ falls } \vec{x} \neq \vec{0}.$$

Bemerkung: Die Linearität in der zweiten Komponente folgt bereits aus der Linearität in der ersten Komponente und der Symmetrie, so daß man nur Linearität in der ersten Komponente, Symmetrie und positive Definitheit fordern muß. Man bezeichnet eine Abbildung $V \times V \rightarrow \mathbb{R}$ mit den oben genannten Eigenschaften als eine **positiv definite symmetrische Bilinearform**.

Mit diesem Skalarprodukt wird der Vektorraum V zu einem euklidischen Vektorraum. Die Bezeichnung "euklidisch" bezieht sich insbesondere auf Forderung nach positiver Definitheit.

Bemerkung: In der Physik treten auch Skalarprodukte auf, bei denen die Forderung nach positiv Definitheit aufgegeben wird. Ein Beispiel hierfür ist das Skalarprodukt im Minkowskiraum.

Wir betrachten nun \mathbb{C}^n . Seien $\vec{x}, \vec{y} \in \mathbb{C}^n$. In diesem Fall definieren wir das unitäre Standard-skalarprodukt als

$$\begin{aligned} V \times V &\rightarrow \mathbb{C}, \\ \vec{x} \cdot \vec{y} &= x_1^* y_1 + x_2^* y_2 + \dots + x_n^* y_n. \end{aligned}$$

Allgemein ist ein unitäres Skalarprodukt eines komplexen Vektorraumes eine Abbildung $V \times V \rightarrow \mathbb{C}$, welche die folgenden Bedingungen erfüllt:

- Semilinear in der ersten Komponente:

$$\begin{aligned} (\vec{x} + \vec{y}) \cdot \vec{z} &= \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}, \\ (\lambda \cdot \vec{x}) \cdot \vec{y} &= \lambda^* (\vec{x} \cdot \vec{y}). \end{aligned}$$

- Linear in der zweiten Komponente:

$$\begin{aligned} \vec{x} \cdot (\vec{y} + \vec{z}) &= \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}, \\ \vec{x} \cdot (\lambda \cdot \vec{y}) &= \lambda (\vec{x} \cdot \vec{y}). \end{aligned}$$

- Hermitisch:

$$\vec{x} \cdot \vec{y} = (\vec{y} \cdot \vec{x})^*.$$

- Positiv definit:

$$\vec{x} \cdot \vec{x} > 0, \text{ falls } \vec{x} \neq \vec{0}.$$

Bemerkung: Auch hier folgt die Linearität in der zweiten Komponente wieder bereits aus der Semilinearität in der ersten Komponente und der Hermite-Eigenschaft wie man leicht sieht:

$$\begin{aligned}\vec{x} \cdot (\vec{y} + \vec{z}) &= [(\vec{y} + \vec{z}) \cdot \vec{x}]^* = [\vec{y} \cdot \vec{x} + \vec{z} \cdot \vec{x}]^* = (\vec{y} \cdot \vec{x})^* + (\vec{z} \cdot \vec{x})^* = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}, \\ \vec{x} \cdot (\lambda \cdot \vec{y}) &= [(\lambda \cdot \vec{y}) \cdot \vec{x}]^* = [\lambda^* (\vec{y} \cdot \vec{x})]^* = [\lambda^* (\vec{x} \cdot \vec{y})^*]^* = \lambda (\vec{x} \cdot \vec{y}).\end{aligned}$$

Man bezeichnet eine Abbildung $V \times V \rightarrow \mathbb{C}$ mit diesen Eigenschaften als eine **positiv definite Hermitesche Form**.

Mit diesem Skalarprodukt wird der Vektorraum V zu einem unitären Vektorraum.

Definition: Man bezeichnet mit

$$|\vec{x}| = \sqrt{\vec{x} \cdot \vec{x}}$$

die Länge oder den **Betrag von \vec{x}** .

Sei $\vec{x}, \vec{y} \in \mathbb{R}^n$. Der Winkel zwischen den beiden Vektoren ist gegeben durch

$$\vec{x} \cdot \vec{y} = |\vec{x}| |\vec{y}| \cos \varphi,$$

also

$$\varphi = \arccos \frac{\vec{x} \cdot \vec{y}}{|\vec{x}| |\vec{y}|}.$$

Zwei Vektoren stehen **senkrecht** aufeinander ($\varphi = 90^\circ$), falls

$$\vec{x} \cdot \vec{y} = 0.$$

4.3 Das Kreuzprodukt

Sei V der Vektorraum \mathbb{R}^3 oder \mathbb{C}^3 . In einem dreidimensionalen Vektorraum ist zusätzlich das Kreuzprodukt als eine Abbildung

$$\begin{aligned}V \times V &\rightarrow V, \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &= \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}\end{aligned}$$

definiert.

Wichtig: Das Kreuzprodukt gibt es nur in drei Dimensionen !

Das Kreuzprodukt ist anti-symmetrisch:

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}.$$

Der Vektor $\vec{x} \times \vec{y}$ steht senkrecht auf \vec{x} und \vec{y} :

$$\vec{x} \cdot (\vec{x} \times \vec{y}) = 0,$$

$$\vec{y} \cdot (\vec{x} \times \vec{y}) = 0,$$

Für den Betrag von $\vec{x} \times \vec{y}$ gilt:

$$|\vec{x} \times \vec{y}| = |\vec{x}| |\vec{y}| \sin \varphi,$$

wobei φ der Winkel zwischen \vec{x} und \vec{y} ist. Sei

$$\vec{z} = \vec{x} \times \vec{y}.$$

Für die Komponenten von \vec{z} gilt:

$$z_i = \sum_{j=1}^3 \sum_{k=1}^3 \varepsilon_{ijk} x_j y_k$$

Hier wurde der **antisymmetrische Tensor** (oder **Levi-Civita-Tensor**) ε_{ijk} verwendet, der wie folgt definiert ist:

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{für } (i, j, k) \text{ eine gerade Permutation von } (1, 2, 3), \\ -1 & \text{für } (i, j, k) \text{ eine ungerade Permutation von } (1, 2, 3), \\ 0 & \text{sonst.} \end{cases}$$

Eine Permutation $(\sigma_1, \sigma_2, \dots, \sigma_n)$ nennt man gerade, wenn man sie durch eine gerade Anzahl von paarweisen Vertauschungen aus $(1, 2, \dots, n)$ erzeugen kann. Benötigt man eine ungerade Anzahl von Vertauschungen, so nennt man die Permutation ungerade.

Beispiel:

$(3, 2, 1, 5, 4)$ ist eine gerade Permutation (vertausche $1 \leftrightarrow 3$ und $4 \leftrightarrow 5$),

$(1, 5, 3, 4, 2)$ ist eine ungerade Permutation (vertausche $2 \leftrightarrow 5$).

An dieser Stelle definieren wir auch noch das **Kronecker-Delta-Symbol**:

$$\delta_{ij} = \begin{cases} +1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Es sei V der Vektorraum \mathbb{R}^3 oder \mathbb{C}^3 . Wir bezeichnen die Abbildung

$$P : V \rightarrow V, \\ \vec{x} \rightarrow -\vec{x}$$

als **Paritätsoperation** oder **Raumspiegelung**. In Komponenten haben wir

$$P \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \\ -x_3 \end{pmatrix}.$$

Es seien $\vec{x}, \vec{y} \in V$ und

$$\vec{z} = \vec{x} \times \vec{y}.$$

Wir setzen

$$\vec{x}' = P(\vec{x}) = -\vec{x}, \quad \vec{y}' = P(\vec{y}) = -\vec{y},$$

und

$$\vec{z}' = P(\vec{x}) \times P(\vec{y}) = (-\vec{x}) \times (-\vec{y}) = \vec{x} \times \vec{y} = \vec{z}.$$

Es ist daher

$$\vec{z}' = \vec{z} \neq P(\vec{z}) = -\vec{z}.$$

Unter einer Raumspiegelung verhält sich also $\vec{z}' = \vec{x}' \times \vec{y}'$ anders als die Vektoren \vec{x}' und \vec{y}' :

$$\begin{aligned}\vec{x}' &= P(\vec{x}), \\ \vec{y}' &= P(\vec{y}), \\ \vec{z}' &= -P(\vec{z}).\end{aligned}$$

Wir bezeichnen $\vec{z} = \vec{x} \times \vec{y}$ als einen **axialen Vektor** oder **Pseudovektor**.

4.4 Vektoren in der Physik: Vierervektoren

Ein Punkt im Raum wird durch einen Ortsvektor

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

beschrieben. Der zugrundeliegende Vektorraum ist der \mathbb{R}^3 .

In der Physik (insbesondere in der speziellen Relativitätstheorie) fasst man oft die Zeit und die drei Ortskoordinaten zu einem Vierervektor zusammen:

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

Die Komponenten eines Vierervektors notiert man als x^μ , wobei $\mu = 0, 1, 2, 3$, also

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z.$$

Für die Komponenten eines Vierervektors verwendet man üblicherweise griechische Indizes wie zum Beispiel $\mu, \nu, \rho, \sigma, \dots$, für die Komponenten eines Dreiervektors hingegen lateinische Indizes i, j, k, \dots . Es hat sich auch eingebürgert, zu schreiben "es sei x^μ ein Vierervektor" anstelle von der eigentlich korrekten Aussage "es seien x^μ die Komponenten eines Vierervektors".

Für Vierervektoren ist der zugrundeliegende Vektorraum der \mathbb{R}^4 , allerdings nicht mit dem euklidischen Skalarprodukt, sondern dem Minkowski-Skalarprodukt. Seien x und y Vierervektoren, so setzt man

$$x \cdot y = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3.$$

Dieses Skalarprodukt ist nicht positiv definit ! Für

$$x^\mu = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \quad \text{und} \quad y^\mu = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

haben wir

$$\begin{aligned} x \cdot x &= 1^2 - 2^2 - 3^2 - 4^2 = -28, \\ y \cdot y &= 1^2 - 1^2 = 0. \end{aligned}$$

Man bezeichnet Vierervektoren wie folgt:

$$\begin{aligned} x \cdot x > 0 &: \text{zeitartig,} \\ x \cdot x = 0 &: \text{lichtartig,} \\ x \cdot x < 0 &: \text{raumartig.} \end{aligned}$$

Bei Vierervektoren unterscheidet man zwischen hochgestellten und tiefgestellten Indizes. Bisher haben wir nur Vierervektoren mit hochgestellten Indizes betrachtet. Diese werden als **kontravariante Vierervektoren** bezeichnet. Vierervektoren mit tiefgestellten Indizes bezeichnet man als **kovariante Vierervektoren**. Ist ein kontravarianter Vierervektor $x^\mu = (x^0, x^1, x^2, x^3)$ gegeben (auf Spalten- bzw. Zeilenschreibweise müssen wir hier keine Rücksicht nehmen, da x^μ ja eigentlich die einzelnen Komponenten eines Vierervektors bezeichnet), so definieren wir den zugehörigen kovarianten Vierervektor $x_\mu = (x_0, x_1, x_2, x_3)$ durch

$$x_\mu = (x^0, -x^1, -x^2, -x^3).$$

Es gilt also

$$x_0 = x^0, \quad x_1 = -x^1, \quad x_2 = -x^2, \quad x_3 = -x^3.$$

Die räumlichen Komponenten bekommen also beim Übergang von kontravarianten zu kovarianten Vierervektoren ein Minuszeichen, die zeitliche Komponente ändert sich nicht. Etwas formaler

kann man diesen Zusammenhang auch wie folgt ausdrücken: Man definiert zunächst die Größen $g_{\mu\nu}$ und $g^{\mu\nu}$ für $\mu, \nu \in \{0, 1, 2, 3\}$ durch

$$g_{\mu\nu} = g^{\mu\nu} = \begin{cases} 1 & \mu = 0, \nu = 0, \\ -1 & \mu = \nu = j, \quad j \in \{1, 2, 3\} \\ 0 & \mu \neq \nu. \end{cases}$$

Dann gilt, wie man leicht verifiziert

$$x_\mu = \sum_{\nu=0}^3 g_{\mu\nu} x^\nu, \quad x^\mu = \sum_{\nu=0}^3 g^{\mu\nu} x_\nu.$$

Da Summen über Viererindizes oft vorkommen und um Schreibarbeit zu sparen führte Einstein die folgende Summenkonvention ein: Tritt ein Index genau einmal hochgestellt und einmal tiefgestellt auf, so ist eine Summe über diesen Index impliziert. Mit Hilfe der **Einsteinschen Summenkonvention** können wir also schreiben

$$x_\mu = g_{\mu\nu} x^\nu, \quad x^\mu = g^{\mu\nu} x_\nu.$$

In beiden Fällen tritt der Index ν auf der echten Seite einmal hochgestellt und einmal tiefgestellt auf. Man bezeichnet $g_{\mu\nu}$ als die Komponenten des metrischen Tensors, oder kurz als **metrischen Tensor** bzw. noch kürzer als **Metrik**. $g^{\mu\nu}$ wird als **inverse Metrik** bezeichnet. Der Übergang zwischen kontravarianten und kovarianten Indizes verläuft immer mit Hilfe der Metrik bzw. der inversen Metrik.

Kehren wir nochmal zum Minkowski-Skalarprodukt zurück:

$$x \cdot y = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3.$$

Mit Hilfe der kovarianten Vierervektoren und der Einsteinschen Summenkonvention können wir nun auch schreiben

$$x \cdot y = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3 = x_0 y^0 + x_1 y^1 + x_2 y^2 + x_3 y^3 = x_\mu y^\mu.$$

Wir haben bisher nur die Zeit und die Raumkoordinaten zu einem Vierervektor zusammengefasst. Ebenso kann man die Energie und die drei Koordinaten des Impulses zu einem Vierervektor zusammenfassen:

$$p_\mu = \begin{pmatrix} E/c \\ p_x \\ p_y \\ p_z \end{pmatrix}.$$

4.5 Lineare Gleichungssysteme

Definition: Unter einem linearen Gleichungssystem versteht man n Gleichungen mit m Unbekannten x_1, x_2, \dots, x_m der Form

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m &= b_1, \\a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2, \\&\dots \\a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nm}x_m &= b_n.\end{aligned}$$

Die Koeffizienten a_{ij} und b_i sind gegebene reelle oder komplexe Zahlen. Jede Variable kommt nur linear vor und jeder Term auf der linken Seite enthält nur eine Variable, daher der Name "lineares Gleichungssystem". Beispiel:

$$\begin{aligned}3x_1 + 3x_2 + 9x_3 &= 36, \\2x_1 + 3x_2 + 7x_3 &= 29, \\x_2 + 4x_3 &= 14.\end{aligned}$$

Wir betrachten nun einen Algorithmus um ein Gleichungssystem mit n Gleichungen und m Unbekannten systematisch zu vereinfachen und zu lösen. Wir beginnen mit einer trivialen Beobachtung: Offensichtlich können Zeilen vertauscht werden, d.h. das Gleichungssystem

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m &= b_1, \\a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2,\end{aligned}$$

ist äquivalent zu dem Gleichungssystem

$$\begin{aligned}a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2, \\a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m &= b_1.\end{aligned}$$

Desweiteren sei (x_1, x_2, \dots, x_m) ein m -Tupel, welches die Gleichung

$$a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_mx_m = b,$$

erfüllt. Dann erfüllt es auch die Gleichung

$$(ca_1)x_1 + (ca_2)x_2 + (ca_3)x_3 + \dots + (ca_m)x_m = cb,$$

Umgekehrt gilt, daß für $c \neq 0$ jedes m -Tupel, welches die zweite Gleichung erfüllt, auch die erste Gleichung erfüllt. Daraus folgt, daß man die linke und rechte Seite einer Gleichung mit einer konstanten Zahl c ungleich Null multiplizieren darf.

Die dritte elementare Umformung ist die folgende: Man darf eine Zeile durch die Summe dieser Zeile mit einer anderen Zeile ersetzen, d.h. die Gleichungssysteme

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m = b_1,$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m = b_2,$$

und

$$\begin{aligned} (a_{11} + a_{21})x_1 + (a_{12} + a_{22})x_2 + (a_{13} + a_{23})x_3 + \dots + (a_{1m} + a_{2m})x_m &= b_1 + b_2, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2, \end{aligned}$$

haben die gleichen Lösungen.

Mit Hilfe dieser drei elementaren Umformungen läßt sich nun ein Algorithmus zur systematischen Vereinfachung von linearen Gleichungssystemen angeben:

1. Setze $i = 1$ (Zeilenindex), $j = 1$ (Spaltenindex).
2. Falls $a_{ij} = 0$ suche $k > i$, so daß $a_{kj} \neq 0$ und vertausche Zeilen i und k .
3. Falls ein solches k aus Schritt 2 nicht gefunden werden kann, setze $j \rightarrow j + 1$ und gehe zurück zu Schritt 2.
4. Falls man in Schritt 3 den Wert $j = m + 1$ erreicht, beende den Algorithmus.
5. Multipliziere Zeile i mit $1/a_{ij}$.
6. Für alle Zeilen $k \neq i$ addiere zur Zeile k das $-a_{kj}$ -fache der i -ten Zeile.
7. Setze $i \rightarrow i + 1$ und $j \rightarrow j + 1$ und gehe zurück zu Schritt 2.
8. Falls man in Schritt 7 den Wert $i = n + 1$ oder den Wert $j = m + 1$ erreicht, beende den Algorithmus.

Man nennt diesen Algorithmus den Gauß'schen Eliminationsalgorithmus. In der Praxis schreibt man das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2, \\ &\dots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nm}x_m &= b_n. \end{aligned}$$

wie folgt auf:

$$\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} & b_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} & b_n \end{array}$$

Dies ist ausreichend, da alle Umformungen nur auf die Koeffizienten a_{ij} und b_i wirken.

Wir betrachten das obige Beispiel:

$$\begin{aligned}3x_1 + 3x_2 + 9x_3 &= 36, \\2x_1 + 3x_2 + 7x_3 &= 29, \\x_2 + 4x_3 &= 14.\end{aligned}$$

Aufgeschrieben ergibt dies:

$$\begin{array}{ccc|c}3 & 3 & 9 & 36 \\2 & 3 & 7 & 29 \\0 & 1 & 4 & 14\end{array}$$

Wir beginnen nun mit den Umformungen:

$$\begin{array}{ccc|c}1 & 1 & 3 & 12 \\2 & 3 & 7 & 29 \\0 & 1 & 4 & 14\end{array}$$

$$\begin{array}{ccc|c}1 & 1 & 3 & 12 \\0 & 1 & 1 & 5 \\0 & 1 & 4 & 14\end{array}$$

$$\begin{array}{ccc|c}1 & 0 & 2 & 7 \\0 & 1 & 1 & 5 \\0 & 0 & 3 & 9\end{array}$$

$$\begin{array}{ccc|c}1 & 0 & 2 & 7 \\0 & 1 & 1 & 5 \\0 & 0 & 1 & 3\end{array}$$

$$\begin{array}{ccc|c}1 & 0 & 0 & 1 \\0 & 1 & 0 & 2 \\0 & 0 & 1 & 3\end{array}$$

Das lineare Gleichungssystem ist somit äquivalent zu dem Gleichungssystem

$$\begin{aligned}x_1 &= 1, \\x_2 &= 2, \\x_3 &= 3.\end{aligned}$$

Bemerkung: Durch Umbenennung der Variablen x_1, \dots, x_m (dies ist gleichbedeutend mit Spaltenvertauschungen) lässt sich durch den Gauß'schen Eliminationsalgorithmus die folgende Form erreichen:

$$\begin{array}{ccccccc|c}
1 & 0 & \dots & 0 & a_{1(r+1)} & \dots & a_{1m} & b_1 \\
0 & 1 & \dots & 0 & a_{2(r+1)} & \dots & a_{2m} & b_2 \\
\dots & \dots \\
0 & 0 & \dots & 1 & a_{r(r+1)} & \dots & a_{rm} & b_r \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & b_{r+1} \\
\dots & \dots \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & b_n
\end{array}$$

Man bezeichnet r als den Rang (engl. "rank"). Wir betrachten nun, unter welchen Bedingungen es keine Lösung, eine eindeutige Lösung oder mehrere Lösungen gibt:

- Ist eine der Zahlen b_{r+1}, \dots, b_n ungleich Null, so hat das lineare Gleichungssystem keine Lösung.
- Ist $r = m$ und $b_{r+1} = \dots = b_n = 0$, so gibt es eine eindeutige Lösung.
- Ist $r < m$ und $b_{r+1} = \dots = b_n = 0$, so gibt es mehrere Lösungen.

Bemerkung 1: Im ersten Fall ist notwendigerweise $r < n$. Im zweiten und dritten Fall ist auch der Spezialfall $r = n$ enthalten. Für $r = n$ ist $\{b_{r+1}, \dots, b_n\} = \emptyset$ und der zweite Fall reduziert sich auf $r = n = m$, während sich der dritte Fall auf $r = n$ und $r < m$ reduziert.

Bemerkung 2: Ist $r < n$ und $b_{r+1} = \dots = b_n = 0$, so reduzieren sich die Zeilen $(r + 1)$ bis n auf die triviale Gleichung

$$0 = 0.$$

Diese Zeilen enthalten keine zusätzliche Information und können auch weggelassen werden.

Anwendungen: Lineare Unabhängigkeit von Vektoren. Zur Erinnerung: m Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ nennt man linear unabhängig, falls die Gleichung

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_m \vec{v}_m = \vec{0}$$

nur die Lösung $(\lambda_1, \lambda_2, \dots, \lambda_m) = (0, 0, \dots, 0)$ hat. Andernfalls nennt man sie linear abhängig. Ist der zugrundeliegende Vektorraum n -dimensional, so ergibt die obige Gleichung ausgeschrieben in Komponenten n lineare Gleichungen mit m Unbekannten $\lambda_1, \dots, \lambda_m$. Man kann nun mit Hilfe des Gauß'schen Eliminationsalgorithmus feststellen, ob die Vektoren linear abhängig sind.

Beispiel: Sei

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -1 \\ -4 \\ -7 \end{pmatrix}.$$

Dies führt zu dem linearen Gleichungssystem

$$\lambda_1 + 2\lambda_2 - \lambda_3 = 0,$$

$$\lambda_1 + 3\lambda_2 - 4\lambda_3 = 0,$$

$$\lambda_1 + 4\lambda_2 - 7\lambda_3 = 0.$$

Wir formen dieses Gleichungssystem mit Hilfe des Gauß'schen Eliminationsalgorithmus um:

$$\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 1 & 3 & -4 & 0 \\ 1 & 4 & -7 & 0 \\ \hline 1 & 2 & -1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 2 & -6 & 0 \\ \hline 1 & 0 & 5 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Somit gibt es mehrere Lösungen

$$\lambda_1 = -5t, \quad \lambda_2 = 3t, \quad \lambda_3 = t, \quad t \in \mathbb{R}.$$

Die drei Vektoren sind linear abhängig.

4.6 Matrizen

Definition: Eine rechteckige Anordnung

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

von Elementen a_{ij} aus einem Körper nennt man Matrix. Die Elemente a_{ij} nennt man die Komponenten der Matrix. Eine Matrix mit n Zeilen und m Spalten bezeichnet man als $n \times m$ -Matrix.

Eine Matrix bezeichnet man als **quadratisch**, falls $n = m$.

Eine Matrix bezeichnet man als **Einheitsmatrix**, falls sie quadratisch ist und $a_{ij} = \delta_{ij}$.

Eine Matrix bezeichnet man als **Diagonalmatrix**, falls sie quadratisch ist und $a_{ij} = 0$ für alle $i \neq j$.

Eine Matrix bezeichnet man als **obere Dreiecksmatrix**, falls sie quadratisch ist und $a_{ij} = 0$ für alle $i > j$.

Quadratisch:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Diagonalmatrix:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Einheitsmatrix:

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

obere Dreiecksmatrix:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Seien A und B zwei $n \times m$ -Matrizen. Eine **Addition** von zwei Matrizen mit gleicher Spalten- und Zeilenanzahl ist definiert durch

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{pmatrix}$$

Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1+7 & 2+8 & 3+9 \\ 4+10 & 5+11 & 6+12 \end{pmatrix} \\ = \begin{pmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix}.$$

Eine **Multiplikation mit Skalaren** ist definiert durch

$$\lambda \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \dots & \dots & \dots & \dots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}.$$

Beispiel:

$$3 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 3 \\ 3 \cdot 4 & 3 \cdot 5 & 3 \cdot 6 \end{pmatrix} \\ = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{pmatrix}.$$

Mit dieser Addition und dieser skalaren Multiplikation bilden die $n \times m$ -Matrizen einen Vektorraum. Die Dimension dieses Vektorraumes ist $n \cdot m$. Eine Basis ist gegeben durch die Matrizen

e_{ij} ,

$$e_{ij} = \begin{pmatrix} 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \end{pmatrix},$$

die nur in dem Eintrag in der i -ten Zeile und j -ten Spalte eine Eins haben, ansonsten nur Nullen.

Matrizenmultiplikation: Es sei A eine $n \times k$ -Matrix A und B eine $k \times m$ -Matrix. Es sei betont, daß die Spaltenanzahl von A gleich der Zeilenanzahl von B ist. Unter dieser Voraussetzung definieren wir das Matrixprodukt $A \cdot B$ als eine $n \times m$ -Matrix C wie folgt:

$$A \cdot B = C,$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{k1} & b_{k2} & \dots & b_{km} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nm} \end{pmatrix},$$

wobei

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ik}b_{kj}.$$

Merkregel:

Zeile \times Spalte.

Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 \cdot 7 + 2 \cdot 9 + 3 \cdot 11 & 1 \cdot 8 + 2 \cdot 10 + 3 \cdot 12 \\ 4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11 & 4 \cdot 8 + 5 \cdot 10 + 6 \cdot 12 \end{pmatrix}$$

$$= \begin{pmatrix} 7 + 18 + 33 & 8 + 20 + 36 \\ 28 + 45 + 66 & 32 + 50 + 72 \end{pmatrix}$$

$$= \begin{pmatrix} 58 & 64 \\ 139 & 154 \end{pmatrix}$$

Ein n -dimensionaler Spaltenvektor kann als eine $n \times 1$ -Matrix aufgefasst werden. Ebenso kann ein n -dimensionaler Zeilenvektor als eine $1 \times n$ -Matrix betrachtet werden. Setzt man

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix},$$

so läßt sich das lineare Gleichungssystem

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m &= b_1, \\a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2, \\&\dots \\a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nm}x_m &= b_n.\end{aligned}$$

auch wie folgt schreiben:

$$A \cdot \vec{x} = \vec{b}.$$

4.7 Spuren und Determinanten von quadratischen Matrizen

Wir betrachten im folgenden die quadratischen $n \times n$ -Matrizen und führen die Begriffe **Spur** und **Determinante** ein.

Sei A eine $n \times n$ -Matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Unter der **Spur** (engl. "trace") einer quadratischen Matrix versteht man die Summe der Diagonalelemente:

$$\text{Tr } A = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}.$$

Einige Rechenregeln bezüglich der Spur von Matrizen (A, B seien $n \times n$ -Matrizen, λ ein Skalar):

$$\text{Tr } (A + B) = \text{Tr } A + \text{Tr } B,$$

$$\text{Tr } (\lambda \cdot A) = \lambda \text{Tr } A.$$

Beispiel:

$$\text{Tr} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} = 1 + 6 + 11 + 16 = 34$$

Seien A und B zwei $n \times n$ -Matrizen. In diesem Fall ist das Matrixprodukt

$$A \cdot B$$

wieder eine $n \times n$ -Matrix. Für $n \times n$ -Matrizen ist die Matrizenmultiplikation also abgeschlossen. Das neutrale Element bezüglich der Matrizenmultiplikation ist offensichtlich die Einheitsmatrix

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & & & 0 \\ \dots & & \dots & & \dots \\ 0 & & & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Wir wollen nun untersuchen, unter welchen Bedingungen auch ein inverses Element existiert. Hierzu ist die Definition der Determinante einer quadratischen Matrix hilfreich.

Als **Determinante** einer quadratischen Matrix definiert man

$$\det A = \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_n=1}^n \varepsilon_{i_1 i_2 \dots i_n} a_{1i_1} a_{2i_2} \dots a_{ni_n},$$

wobei $\varepsilon_{i_1 i_2 \dots i_n}$ das total antisymmetrische Symbol in n Dimensionen ist

$$\varepsilon_{i_1 i_2 \dots i_n} = \begin{cases} +1 & \text{für } (i_1 i_2 \dots i_n) \text{ eine gerade Permutation von } (1, 2, \dots, n), \\ -1 & \text{für } (i_1 i_2 \dots i_n) \text{ eine ungerade Permutation von } (1, 2, \dots, n), \\ 0 & \text{sonst.} \end{cases}$$

Für die Determinante existiert auch die folgende Schreibweise

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Sei $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ eine Diagonalmatrix. Dann ist

$$\det D = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n.$$

Ein Verfahren zur Berechnung der Determinante ist der **Laplace'sche Entwicklungssatz**. Zu einer $n \times n$ -Matrix A definieren wir zunächst eine $(n-1) \times (n-1)$ -Matrix A_{ij} , die dadurch entsteht, daß man die i -te Zeile und die j -te Spalte der Matrix A entfernt. Der Laplace'sche Entwicklungssatz für die Entwicklung nach der i -ten Zeile lautet:

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Die Zeile i kann frei gewählt werden, daß Ergebnis ist unabhängig von dieser Wahl. Natürlich kann äquivalent auch nach der j -ten Spalte entwickelt werden:

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Auch hier kann die Zeile j frei gewählt werden, daß Ergebnis ist unabhängig von dieser Wahl. Die Laplace'schen Entwicklungssätze erlauben die rekursive Berechnung einer Determinante. Für eine 1×1 -Matrix haben wir

$$|a_{11}| = a_{11}.$$

Rechenregeln für die Determinante:

$$\det(A \cdot B) = (\det A) \cdot (\det B),$$

$$\det(\lambda \cdot A) = \lambda^n \cdot \det A.$$

Beispiel:

$$\begin{aligned} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 0 & 3 & 0 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{vmatrix} = -3 \cdot \begin{vmatrix} 5 & 7 \\ 9 & 11 \end{vmatrix} \\ &= -3 \cdot (5 \cdot 11 - 7 \cdot 9) = 24. \end{aligned}$$

Explizite Formeln für die Determinanten von 1×1 -Matrizen, 2×2 -Matrizen und 3×3 -Matrizen, sind

$$\begin{aligned} |a_{11}| &= a_{11}, \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}a_{22} - a_{12}a_{21}, \\ \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}. \end{aligned}$$

Wenden wir uns wieder der Existenz eines inversen Elements bezüglich der Matrizenmultiplikation zu. Falls so ein Element existiert bezeichnen wir es mit A^{-1} . Es soll also gelten

$$A \cdot A^{-1} = \mathbf{1}.$$

Nehmen wir auf beiden Seiten die Determinante, so erhalten wir

$$\det A \cdot \det A^{-1} = \det \mathbf{1} = 1,$$

also falls $\det A \neq 0$

$$\det A^{-1} = \frac{1}{\det A}.$$

$\det A \neq 0$ ist eine notwendige Bedingung für die Existenz eines Inversen. Es läßt sich zeigen, daß dies auch hinreichend ist.

Satz: A^{-1} existiert genau dann, wenn $\det A \neq 0$.

Wir betrachten nun die Menge aller quadratischen $n \times n$ -Matrizen mit der Eigenschaft $\det A \neq 0$. Wegen $\det(AB) = \det A \det B$ ist diese Menge abgeschlossen bezüglich der Matrizenmultiplikation. Wie gerade diskutiert wurde, existiert zu jeder Matrix auch ein Inverses. Diese Menge bildet daher bezüglich der Matrizenmultiplikation eine Gruppe, die man als

$$GL(n, \mathbb{R}), \quad \text{bzw.} \quad GL(n, \mathbb{C})$$

bezeichnet.

4.8 Berechnung der inversen Matrix

Sei A eine $n \times n$ -Matrix mit $\det A \neq 0$. Gesucht ist eine $n \times n$ -Matrix X

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix},$$

so daß

$$A \cdot X = \mathbf{1}$$

gilt. Wir multiplizieren die linke Seite aus und betrachten danach die j -te Spalte auf beiden Seiten:

$$\begin{aligned} a_{11}x_{1j} + a_{12}x_{2j} + \dots + a_{1n}x_{nj} &= 0, \\ &\dots = 0 \\ a_{j1}x_{1j} + a_{j2}x_{2j} + \dots + a_{jn}x_{nj} &= 1, \\ &\dots = 0 \\ a_{n1}x_{1j} + a_{n2}x_{2j} + \dots + a_{nn}x_{nj} &= 0. \end{aligned}$$

Diese n Gleichungen bilden ein lineares Gleichungssystem für die Unbekannten $x_{1j}, x_{2j}, \dots, x_{nj}$, welches mit Hilfe des Gaußschen Algorithmus gelöst werden kann. Da dies für jede Spalte j gilt, kann man so alle n^2 Unbekannten x_{ij} bestimmen. Da die Koeffizienten der linken Seite des linearen Gleichungssystems immer gleich sind, verfährt man in der Praxis wie folgt: Man schreibt die Gleichungen wie folgt an

$$\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}$$

und bringt dieses Gleichungssystem mit Hilfe des Gauß'schen Algorithmus auf die Form

$$\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & x_{11} & x_{12} & \dots & x_{1n} \\ 0 & 1 & \dots & 0 & x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots \\ 0 & 0 & \dots & 1 & x_{n1} & x_{n2} & \dots & x_{nn} \end{array}$$

Die inverse Matrix ist dann gegeben durch

$$A^{-1} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix}.$$

Beispiel: Sei A die Matrix

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 3 & 7 \\ 0 & 1 & 4 \end{pmatrix}.$$

Wir bestimmen nun die inverse Matrix:

$$\begin{array}{ccc|ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 2 & 3 & 7 & 0 & 1 & 0 \\ 0 & 1 & 4 & 0 & 0 & 1 \\ \\ 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 1 & 4 & 0 & 0 & 1 \\ \\ 1 & 0 & 2 & 3 & -1 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 3 & 2 & -1 & 1 \\ \\ 1 & 0 & 2 & 3 & -1 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \\ \\ 1 & 0 & 0 & \frac{5}{3} & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 0 & -\frac{8}{3} & \frac{4}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \end{array}$$

Somit ist A^{-1}

$$A^{-1} = \frac{1}{3} \begin{pmatrix} 5 & -1 & -2 \\ -8 & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix}.$$

4.9 Eigenwerte und Eigenvektoren*

Sei V ein n -dimensionaler Vektorraum und A eine $n \times n$ -Matrix. Die Matrix A definiert eine Abbildung

$$\begin{aligned} V &\rightarrow V, \\ \vec{x} &\rightarrow \vec{x}' = A \cdot \vec{x}. \end{aligned}$$

Diese Abbildung ist linear

$$A \cdot (c_1 \vec{x} + c_2 \vec{y}) = c_1 A \cdot \vec{x} + c_2 A \cdot \vec{y}.$$

Lineare Abbildungen zwischen zwei Vektorräumen werden auch als Homomorphismen bezeichnet. Da in unserem Fall die Abbildung von V nach V geht, spricht man auch von einem Epimorphismus.

Wir interessieren uns nun für Vektoren $\vec{x} \in V$, die auf ein Vielfaches ihrer selbst abgebildet werden.

$$A \cdot \vec{x} = \lambda \vec{x}.$$

In diesem Fall bezeichnet man λ als Eigenwert und \vec{x} als Eigenvektor.

Beispiel: Falls $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ eine Diagonalmatrix ist, wobei alle λ_i paarweise verschieden sind, so sind die Basisvektoren \vec{e}_j die Eigenvektoren. Der zu \vec{e}_j zugehörige Eigenwert ist λ_j .

Sei nun eine $n \times n$ -Matrix A gegeben. Wir suchen ein Verfahren zur Bestimmung der Eigenwerte und der Eigenvektoren. Hierzu definieren wir das charakteristische Polynom der Abbildung durch

$$\chi_A(\lambda) = \det(A - \lambda \mathbf{1}).$$

Dies ist ein Polynom vom Grad n in λ .

Satz: Die Nullstellen von χ_A sind die Eigenwerte von A .

Beispiel: Wir betrachten wieder

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 3 & 7 \\ 0 & 1 & 4 \end{pmatrix}.$$

Man findet

$$\chi_A(\lambda) = \begin{vmatrix} 1-\lambda & 1 & 3 \\ 2 & 3-\lambda & 7 \\ 0 & 1 & 4-\lambda \end{vmatrix} = -\lambda^3 + 8\lambda^2 - 10\lambda + 3$$

$$= -(\lambda - 1) \left(\lambda - \frac{1}{2}(7 + \sqrt{37}) \right) \left(\lambda - \frac{1}{2}(7 - \sqrt{37}) \right)$$

Somit hat die Matrix A die Eigenwerte

$$1, \quad \frac{1}{2}(7 + \sqrt{37}), \quad \frac{1}{2}(7 - \sqrt{37}).$$

Sind die Eigenwerte bekannt, so bestimmt man im nächsten Schritt die zugehörigen Eigenvektoren. Sei λ ein (bekannter) Eigenwert der Matrix A . Gesucht werden Vektoren \vec{x} , so daß

$$A \cdot \vec{x} = \lambda \vec{x},$$

bzw.

$$(A - \lambda \cdot \mathbf{1}) \vec{x} = \vec{0}.$$

Dies ist wieder ein lineares Gleichungssystem, welches mit Hilfe des Gauß'schen Algorithmus gelöst werden kann.

Beispiel: Wir bestimmen zu der Matrix A aus dem obigen Beispiel die Eigenvektoren zum Eigenwert $\lambda = 1$:

$$A - \lambda \cdot \mathbf{1} = \begin{pmatrix} 0 & 1 & 3 \\ 2 & 2 & 7 \\ 0 & 1 & 3 \end{pmatrix}.$$

Der Gauß'sche Algorithmus liefert:

$$\begin{array}{ccc|c} 0 & 1 & 3 & 0 \\ 2 & 2 & 7 & 0 \\ 0 & 1 & 3 & 0 \\ \hline 1 & 1 & \frac{7}{2} & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Somit sind alle Vektoren der Form

$$\begin{pmatrix} -\frac{1}{2}t \\ -3t \\ t \end{pmatrix}, \quad t \in \mathbb{R}$$

Eigenvektoren zum Eigenwert $\lambda = 1$. Man bezeichnet diesen Untervektorraum als Eigenraum.

Basiswechsel: Sei V ein n -dimensionaler Vektorraum mit einer Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$. Sei

$$\Phi : V \rightarrow V$$

eine Endomorphismus, gegeben in der Koordinatendarstellung der obigen Basis durch eine Matrix A :

$$\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \rightarrow A \cdot \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}.$$

Dieser Endomorphismus bildet also den Basisvektor \vec{v}_i auf

$$\vec{v}_j \rightarrow \sum_{i=1}^n \vec{v}_i a_{ij}$$

ab. (Das Bild von \vec{v}_j ist die j -te Spalte von A .) Sei nun $\{\vec{w}_1, \dots, \vec{w}_n\}$ eine weitere Basis des Vektorraumes. Wir können jeden Basisvektor \vec{w}_j als Linearkombination der \vec{v}_i 's schreiben:

$$\vec{w}_j = \sum_{i=1}^n \vec{v}_i m_{ij}.$$

Dies definiert eine $n \times n$ -Matrix M

$$M = (m_{ij}).$$

Die Koordinatendarstellungen eines beliebigen Vektors \vec{z} bezüglich der beiden Basen

$$\vec{z} = x_1 \vec{v}_1 + \dots + x_n \vec{v}_n = y_1 \vec{w}_1 + \dots + y_n \vec{w}_n$$

transformieren sich wie

$$\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = M \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}.$$

Da umgekehrt auch die Basisvektoren \vec{v}_i durch die Basisvektoren \vec{w}_j ausgedrückt werden können, ist M invertierbar, also

$$M \in GL(n, K), \quad K = \mathbb{R}, \mathbb{C},$$

und

$$\vec{v}_j = \sum_{i=1}^n \vec{w}_i m_{ij}^{-1}.$$

In der Basis $\{\vec{w}_1, \dots, \vec{w}_n\}$ wird der Endomorphismus Φ durch eine Matrix B beschrieben

$$\vec{w}_j \rightarrow \sum_{i=1}^n \vec{w}_i b_{ij}.$$

Wir bestimmen nun den Zusammenhang von A mit B : Es gilt

$$\vec{w}_j = \sum_{k=1}^n \vec{v}_k m_{kj} \rightarrow \sum_{k=1}^n \sum_{l=1}^n \vec{v}_l a_{lk} m_{kj} = \sum_{i=1}^n \sum_{k=1}^n \sum_{l=1}^n \vec{w}_i m_{il}^{-1} a_{lk} m_{kj},$$

und somit ist

$$B = M^{-1}AM.$$

Wir betrachten nun wieder den Endomorphismus Φ , der in der Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ durch die Matrix A gegeben ist. Wir können nun die Frage stellen, ob es eine neue Basis $\{\vec{w}_1, \dots, \vec{w}_n\}$ gibt, so daß der Endomorphismus in dieser neuen Basis durch eine besonders einfache Matrix B gegeben ist. Eine besonders einfache Matrix ist eine Diagonalmatrix.

Definition: Einen Endomorphismus nennt man **diagonalisierbar**, falls es eine Basis $\{\vec{w}_1, \dots, \vec{w}_n\}$ gibt, so daß in dieser neuen Basis die Abbildung durch eine Diagonalmatrix gegeben ist. In diesem Fall ist es unmittelbar einsichtig, daß die \vec{w}_i 's Eigenvektoren sind. Da auch die Umkehrung gilt, hat man folgenden Satz:

Eine Matrix A läßt sich genau dann durch einen Basiswechsel auf Diagonalgestalt bringen

$$D = M^{-1}AM, \quad D \text{ Diagonalmatrix, } M \in GL(n, K),$$

falls es eine Basis von V aus Eigenvektoren von A gibt.

Die Berechnung der Eigenwerte und der Eigenvektoren liefert ein konstruktives Verfahren zur Diagonalisierung einer Matrix. Seien $\vec{w}_1, \dots, \vec{w}_n$ die Eigenvektoren. Dann ergeben diese Vektoren als Spaltenvektoren betrachtet die Matrix M :

$$M = (\vec{w}_1, \dots, \vec{w}_n).$$

Beispiel: Die Matrix

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

hat hat charakteristische Polynom

$$\chi_A(\lambda) = -(\lambda - 1)^2(\lambda + 1).$$

Der Eigenraum zum Eigenwert $\lambda = 1$ wird von den Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

aufgespannt. Der Vektor

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

ist Eigenvektor zum Eigenwert $\lambda = -1$. Somit ist M gegeben durch

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \quad M^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix},$$

und

$$D = M^{-1}AM = \text{diag}(1, 1, -1).$$

Wir wollen noch genauer untersuchen, unter welchen Bedingungen eine Matrix diagonalisiert werden kann. Bei der Bestimmung der Nullstellen des charakteristischen Polynoms kann der Fall auftreten, daß das Polynom nicht vollständig in Linearfaktoren zerfällt. Dies ist zum Beispiel bei der Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \chi_A(\lambda) = \lambda^2 + 1,$$

über dem Körper der reellen Zahlen der Fall, da

$$\lambda^2 = -1$$

keine Lösung in \mathbb{R} hat.

Bemerkung: Die komplexen Zahlen \mathbb{C} sind algebraisch abgeschlossen, d.h. über \mathbb{C} zerfällt jedes Polynom in Linearfaktoren. Im obigen Beispiel findet man

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

Eine zweite Komplikation kann auftreten, falls das charakteristische Polynom Nullstellen mit höherer Multiplizität hat. Ist die Multiplizität einer Nullstelle größer als die Dimension des zugehörigen Eigenraumes, so ist die Matrix nicht diagonalisierbar. Beispiel: Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \chi_A(\lambda) = \lambda^2,$$

hat den doppelten Eigenwert $\lambda = 0$. Der zugehörige Eigenraum ist aber nur ein-dimensional:

$$\begin{pmatrix} t \\ 0 \end{pmatrix}, \quad t \in \mathbb{R}.$$

A ist nicht diagonalisierbar. Wäre A diagonalisierbar, so wäre die zugehörige Diagonalmatrix die Nullmatrix, was zu einem Widerspruch führt.

Zusammenfassend erhalten wir die folgende Aussage:

Satz: Eine Matrix A mit Einträgen aus dem Körper K ist genau dann diagonalisierbar, falls das charakteristische Polynom in Linearfaktoren zerfällt, und die Dimension des Eigenraumes zum Eigenwert λ_i gleich der Multiplizität der Nullstelle λ_i des charakteristischen Polynoms ist.

Es läßt sich zeigen, daß die Dimension der Eigenräume stets größer oder gleich Eins ist. Daher gilt im besonderen: Falls das charakteristische Polynom in Linearfaktoren zerfällt und alle Nullstellen paarweise verschieden sind, so ist die Matrix A diagonalisierbar.

4.10 Das Schmidtsche Orthonormierungsverfahren

Wir betrachten einen euklidischen bzw. unitären Vektorraum V der Dimension n , d.h einen Vektorraum über den reellen Zahlen mit einer positiv definiten symmetrischen Bilinearform (euklidischer Vektorraum) bzw. einen Vektorraum über den komplexen Zahlen mit einer positiv definiten Hermiteschen Form. Gegeben seien k lineare unabhängige Vektoren $\{\vec{v}_1, \dots, \vec{v}_k\}$. Diese Vektoren spannen einen Untervektorraum W auf (und bilden eine Basis von W). Gesucht ist nun eine Orthonormalbasis $\{\vec{n}_1, \dots, \vec{n}_k\}$ von W . Unter einer Orthonormalbasis versteht man

$$|\vec{n}_i| = 1, \quad \vec{n}_i \cdot \vec{n}_j = 0, \quad \text{für } i \neq j.$$

Hierzu geht man wie folgt vor: Man setzt am Anfang

$$\vec{n}_1 = \frac{1}{|\vec{v}_1|} \vec{v}_1.$$

Seien nun $\vec{n}_1, \dots, \vec{n}_{j-1}$ schon konstruiert. Dann erhält man \vec{n}_j wie folgt: Man definiert zunächst

$$\vec{w} = \vec{v}_j - \sum_{i=1}^{j-1} (\vec{v}_j \cdot \vec{n}_i) \vec{n}_i.$$

Durch diese Konstruktion steht \vec{w} senkrecht zu allen bisher konstruierten Vektoren:

$$\vec{w} \cdot \vec{n}_i = 0, \quad \text{für } i \in \{1, 2, \dots, j-1\}.$$

Wir normieren \vec{w} und erhalten \vec{n}_j :

$$\vec{n}_j = \frac{1}{|\vec{w}|} \vec{w}.$$

Eine typische Anwendung für das Schmidt'sche Orthonormierungsverfahren ist die Konstruktion einer Orthonormalbasis für einen Eigenraum der Dimension $D > 1$.

5 Analysis

5.1 Folgen

Unter einer Folge (a_n) reeller Zahlen versteht man eine Abbildung $\mathbb{N} \rightarrow \mathbb{R}$. Jedem $n \in \mathbb{N}$ wird also ein $a_n \in \mathbb{R}$ zugeordnet.

Beispiel:

$$a_n = \frac{1}{n^2}$$

definiert eine Folge. Explizit:

$$a_1 = 1, \quad a_2 = \frac{1}{4}, \quad a_3 = \frac{1}{9}, \quad a_4 = \frac{1}{16}, \quad a_5 = \frac{1}{25}, \quad \dots$$

Eine Folge (a_n) heißt **konvergent** gegen $a \in \mathbb{R}$, falls es zu jedem $\varepsilon > 0$ eine natürliche Zahl N gibt, so daß

$$|a_n - a| < \varepsilon, \quad \forall n \geq N.$$

In diesem Fall schreibt man

$$\lim_{n \rightarrow \infty} a_n = a.$$

In anderen Worten liegen für eine konvergente Folge ab einem bestimmten N alle Folgenglieder im Intervall $]a - \varepsilon, a + \varepsilon[$.

Eine Folge reeller Zahlen nennt man **divergent**, wenn sie gegen keine reelle Zahl konvergiert. Beispiele für divergente Folgen sind

$$a_n = n, \\ b_n = \begin{cases} 1 & n \text{ gerade} \\ -1 & n \text{ ungerade} \end{cases}$$

Eine Folge heißt nach oben (bzw. unten) beschränkt, falls es ein $c \in \mathbb{R}$ gibt, so daß $a_n \leq c$ (bzw. $a_n \geq c$) für alle $n \in \mathbb{N}$. Die Folge heißt beschränkt, wenn sie nach oben und unten beschränkt ist. Jede konvergente Folge beschränkt.

Bemerkung: Die Umkehrung gilt nicht, eine beschränkte Folge ist nicht notwendiger Weise konvergent, siehe obiges Beispiel mit der Folge (b_n) .

Seien (a_n) und (b_n) zwei konvergente Folgen mit den Grenzwerten

$$\lim_{n \rightarrow \infty} a_n = a, \quad \lim_{n \rightarrow \infty} b_n = b.$$

Dann sind auch die Folgen $(a_n + b_n)$, $(a_n - b_n)$, (λa_n) , $(a_n b_n)$ konvergent ($\lambda \in \mathbb{R}$) und es gilt

$$\begin{aligned}\lim_{n \rightarrow \infty} (a_n + b_n) &= a + b, \\ \lim_{n \rightarrow \infty} (a_n - b_n) &= a - b, \\ \lim_{n \rightarrow \infty} (\lambda a_n) &= \lambda a, \\ \lim_{n \rightarrow \infty} (a_n \cdot b_n) &= a \cdot b.\end{aligned}$$

Ist weiter $b \neq 0$, so gibt es ein $N \in \mathbb{N}$, so daß $b_n \neq 0$ für alle $n \geq N$ und wir können die Folge

$$\left(\frac{a_n}{b_n}\right)_{n \geq N}$$

betrachten. Es gilt

$$\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n}\right) = \frac{a}{b}.$$

Seien (a_n) und (b_n) zwei konvergente Folgen mit $a_n \leq b_n$ für alle n . Dann gilt auch

$$\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n.$$

Bemerkung: Aus $a_n < b_n$ folgt nicht $\lim a_n < \lim b_n$, wie das Beispiel $a_n = 0$ und $b_n = 1/n$ zeigt.

Wir hatten bereits bei der axiomatischen Charakterisierung der reellen Zahlen den Begriff einer Cauchy-Folge eingeführt, den wir uns nochmal in Erinnerung rufen: Eine Folge (a_n) reeller Zahlen nennt man **Cauchy-Folge**, falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß

$$|a_n - a_m| < \varepsilon, \quad \forall n, m \geq N.$$

Satz: Ist eine Folge (a_n) reeller Zahlen konvergent, so ist sie auch eine Cauchy-Folge.

Die Umkehrung dieses Satzes postuliert nennt man als Axiom. Das **Vollständigkeitsaxiom** besagt, daß in \mathbb{R} jede Cauchy-Folge konvergent ist. Wir hatten dies bei der axiomatischen Charakterisierung der reellen Zahlen bereits erwähnt.

Somit gilt in \mathbb{R} , daß eine Folge (a_n) reeller Zahlen genau dann konvergent ist, falls sie eine Cauchy-Folge ist. Der Vorteil der Definition einer Cauchy-Folge gegenüber der Definition des Begriffes Konvergenz besteht darin, daß sich erstere nur auf einzelne Folgenglieder bezieht und keinen Bezug auf einen (eventuellen) Grenzwert nimmt.

5.2 Reihen

Sei (a_n) eine Folge reeller Zahlen. Man betrachtet nun die Folge (s_n) der Partialsummen

$$s_n = \sum_{j=1}^n a_j.$$

Als unendliche Reihe bezeichnet man nun die Folge dieser Partialsummen. Man schreibt

$$\sum_{j=1}^{\infty} a_j = \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{j=1}^n a_j.$$

Eine unendliche Reihe heißt konvergent, wenn die Folge der Partialsummen konvergiert.

Eine unendliche Reihe heißt **absolut konvergent**, falls die Reihe

$$\sum_{j=1}^{\infty} |a_j|$$

konvergent ist. Ist eine Reihe absolut konvergent, so ist sie auch konvergent im gewöhnlichen Sinne. Wir interessieren uns nun für Kriterien, die eine Konvergenz einer unendlichen Reihe garantieren.

Konvergenzkriterium von Cauchy: Die Reihe $\sum_{j=1}^{\infty} a_j$ konvergiert genau dann, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so daß

$$\left| \sum_{j=m}^n a_j \right| < \varepsilon, \quad \forall n \geq m \geq N.$$

Satz: Eine notwendige (aber nicht hinreichende) Bedingung für die Konvergenz einer Reihe ist, daß

$$\lim_{n \rightarrow \infty} a_n = 0.$$

Satz: Eine Reihe $\sum_{j=1}^{\infty} a_j$ mit $a_j \geq 0$ konvergiert genau dann, wenn die Folge der Partialsummen beschränkt ist.

Konvergenzkriterium von Leibniz für alternierende Reihen: Sei (a_n) eine monoton fallende Folge nicht-negativer Zahlen mit $\lim a_n = 0$. Dann konvergiert die Reihe

$$\sum_{j=1}^{\infty} (-1)^j a_j.$$

Satz: Eine absolut konvergente Reihe konvergiert auch im gewöhnlichen Sinne. Diese Tatsache hatten wir bereits oben erwähnt. Die Umkehrung gilt im Allgemeinen nicht: So ist die Reihe

$$\sum_{j=1}^{\infty} \frac{(-1)^j}{j} = -\ln 2$$

konvergent, die Reihe

$$\sum_{j=1}^{\infty} \frac{1}{j}$$

aber divergent. Die Konvergenz der ersten Reihe folgt aus dem Leibnizkriterium für alternierende Reihen: Die Reihe ist alternierend und $a_n = 1/n$ ist eine monoton fallende Folge nicht-negativer Zahlen. Die Divergenz der zweiten Reihe zeigt man indem man die Partialsummen betrachtet:

$$\begin{aligned} S_n &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots + \frac{1}{n} \\ &\geq 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \dots + \frac{1}{n} \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{n}. \end{aligned}$$

Dies zeigt, daß die Folge der Partialsummen nicht nach oben beschränkt ist. Hieraus folgt die Divergenz der harmonischen Reihe.

Majorantenkriterium: Sei $\sum_{j=1}^{\infty} c_j$ eine konvergente Reihe mit lauter nicht-negativen Gliedern und (a_n) eine Folge mit $|a_n| \leq c_n$. Dann konvergiert die Reihe

$$\sum_{j=1}^{\infty} a_j$$

absolut. Man nennt $\sum_{j=1}^{\infty} c_j$ eine Majorante von $\sum_{j=1}^{\infty} a_j$.

Quotientenkriterium: Sei $\sum_{j=1}^{\infty} a_j$ eine Reihe mit $a_n \neq 0$ für alle n und x eine reelle Zahl $0 < x < 1$, so daß

$$\left| \frac{a_{n+1}}{a_n} \right| \leq x, \quad \forall n \geq N.$$

Dann konvergiert die Reihe absolut.

Beispiel 1: Wir betrachten die Reihe

$$\sum_{j=1}^{\infty} \frac{x^j}{j!}.$$

Es ist

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| = \frac{|x|}{n+1} < 1 \quad \text{für } n > |x|.$$

Die Reihe ist nach dem Quotientenkriterium konvergent. Diese Reihe definiert die Exponentialfunktion:

$$\sum_{j=1}^{\infty} \frac{x^j}{j!} = \exp(x) - 1, \quad \text{bzw.} \quad \sum_{j=0}^{\infty} \frac{x^j}{j!} = \exp(x).$$

Beispiel 2: Es sei $|x| < 1$. Wir betrachten die Reihe

$$\sum_{j=1}^{\infty} \frac{x^j}{j}.$$

Es ist

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\frac{x^{n+1}}{(n+1)}}{\frac{x^n}{n}} \right| = \frac{n}{n+1} |x| \leq |x| < 1.$$

Die Reihe ist nach dem Quotientenkriterium konvergent und definiert die Funktion $-\ln(1-x)$:

$$\sum_{j=1}^{\infty} \frac{x^j}{j} = -\ln(1-x).$$

Cauchy-Produkt von Reihen: Seien $\sum_{j=1}^{\infty} a_j$ und $\sum_{j=1}^{\infty} b_j$ zwei absolut konvergente Reihen. Für $n \in \mathbb{N}$ setzen wir

$$c_n = \sum_{j=1}^n a_j b_{n-j+1}.$$

Dann ist auch die Reihe $\sum_{j=1}^{\infty} c_j$ absolut konvergent und es gilt

$$\sum_{j=1}^{\infty} c_j = \left(\sum_{j=1}^{\infty} a_j \right) \left(\sum_{j=1}^{\infty} b_j \right).$$

Bemerkung: Die absolute Konvergenz ist wesentlich für die Gültigkeit des Satzes! Im Allgemeinen gilt, daß Umordnungen innerhalb einer Reihe nur erlaubt sind, falls die Reihe absolut konvergiert.

Wir geben noch einige wichtige Reihen an:

$$\begin{aligned} \exp x &= \sum_{j=0}^{\infty} \frac{x^j}{j!}, \\ -\ln(1-x) &= \sum_{j=1}^{\infty} \frac{x^j}{j}, \quad |x| < 1, \end{aligned}$$

$$\begin{aligned}\sin x &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!}, \\ \cos x &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!}, \\ \sinh x &= \sum_{j=0}^{\infty} \frac{x^{2j+1}}{(2j+1)!}, \\ \cosh x &= \sum_{j=0}^{\infty} \frac{x^{2j}}{(2j)!}.\end{aligned}$$

\sinh und \cosh bezeichnet man als Sinus Hyperbolicus bzw. Kosinus Hyperbolicus. Mit Ausnahme der Reihe für $\ln(1-x)$ konvergieren alle Reihen absolut für alle Werte von x . Man sagt die Reihen haben einen unendlichen **Konvergenzradius**. Die Reihe für $\ln(1-x)$ konvergiert absolut für $|x| < 1$. Somit hat diese Reihe den Konvergenzradius 1. Man spricht von einem Konvergenzradius, da die obigen Reihen auch definiert sind, wenn man die reelle Variable x durch eine komplexe Variable z ersetzt.

Wir betrachten noch $\exp(ix)$:

$$\begin{aligned}\exp(ix) &= \sum_{j=0}^{\infty} \frac{i^j x^j}{j!} = \sum_{j=0}^{\infty} i^{2j} \frac{x^{2j}}{(2j)!} + \sum_{j=0}^{\infty} i^{2j+1} \frac{x^{2j+1}}{(2j+1)!} \\ &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!} + i \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!} = \cos x + i \sin x.\end{aligned}$$

Die Reihendarstellung liefert also einen einfachen Beweis der Formel:

$$\exp(ix) = \cos x + i \sin x.$$

Ebenso findet man

$$\exp x = \cosh x + \sinh x.$$

Man beachte daß für die Umordnung der Reihen die absolute Konvergenz notwendig ist. Man kann die trigonometrischen und die hyperbolischen Funktionen auch durch die Exponentialfunktion ausdrücken:

$$\begin{aligned}\cos x &= \frac{1}{2} (e^{ix} + e^{-ix}), & \sin x &= \frac{1}{2i} (e^{ix} - e^{-ix}), \\ \cosh x &= \frac{1}{2} (e^x + e^{-x}), & \sinh x &= \frac{1}{2} (e^x - e^{-x}).\end{aligned}$$

Eine Anwendung sind die Additionstheoreme für Sinus und Kosinus:

$$\sin(\alpha + \beta) = \sin(\alpha) \cos(\beta) + \cos(\alpha) \sin(\beta),$$

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta).$$

Leichter zu merken ist die entsprechende Formel für die Exponentialfunktion:

$$e^{i(\alpha+\beta)} = e^{i\alpha} e^{i\beta}.$$

Man kann die Additionstheoreme für Sinus und Kosinus aus der Formel für die Exponentialfunktion wie folgt herleiten: Es ist

$$\begin{aligned} e^{i(\alpha+\beta)} &= \cos(\alpha + \beta) + i \sin(\alpha + \beta), \\ e^{i\alpha} &= \cos(\alpha) + i \sin(\alpha), \\ e^{i\beta} &= \cos(\beta) + i \sin(\beta) \end{aligned}$$

und

$$\begin{aligned} e^{i\alpha} e^{i\beta} &= [\cos(\alpha) + i \sin(\alpha)] [\cos(\beta) + i \sin(\beta)] \\ &= \cos(\alpha)\cos(\beta) + i \cos(\alpha)\sin(\beta) + i \sin(\alpha)\cos(\beta) + i^2 \sin(\alpha)\sin(\beta) \\ &= [\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)] + i [\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)]. \end{aligned}$$

Somit folgt aus $e^{i(\alpha+\beta)} = e^{i\alpha} e^{i\beta}$

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

Nimmt man nun den Real- bzw. Imaginärteil dieser Gleichung, so erhält man die Additionstheoreme für Kosinus und Sinus.

5.3 Funktionen und Stetigkeit

Seien D und W Teilmengen von \mathbb{R} . Unter einer reellwertigen Funktion auf D versteht man eine Abbildung

$$\begin{aligned} f &: D \rightarrow W, \\ x &\rightarrow y = f(x). \end{aligned}$$

Man nennt D den Definitionsbereich und W den Wertebereich der Funktion. Eine Funktion f ordnet jedem $x \in D$ ein $y \in W$ zu. Gibt es zu jedem $y \in W$ genau ein $x \in D$ mit $y = f(x)$, so ist die Funktion f umkehrbar. In diesem Fall bezeichnet man mit f^{-1} die Umkehrfunktion:

$$\begin{aligned} f^{-1} &: W \rightarrow D, \\ y &\rightarrow x = f^{-1}(y). \end{aligned}$$

Man sagt eine Funktion hat im Punkte a den Grenzwert c , falls es mindestens eine Folge $(x_n) \in D$ mit $\lim_{n \rightarrow \infty} x_n = a$ gibt. Gilt dann für jede Folge $(x_n) \in D$ mit $\lim_{n \rightarrow \infty} x_n = a$, daß

$$\lim_{n \rightarrow \infty} f(x_n) = c,$$

so bezeichnet man c als den Grenzwert der Funktion $f(x)$ im Punkte a . In diesem Fall schreibt man

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

Die obige Bedingung ist äquivalent zu der Forderung, daß es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß

$$|f(x) - c| < \varepsilon, \quad \forall |x - a| < \delta \quad \text{und} \quad x \in D.$$

Bemerkung: Es wird nicht vorausgesetzt, daß $a \in D$ liegt. Diese Definition macht auch Sinn, falls D ein offenes Intervall ist und der Grenzwert an den Intervallgrenzen betrachtet wird.

Sei nun $a \in D$. Man bezeichnet eine Funktion als **stetig** im Punkte a , falls

$$\lim_{x \rightarrow a} f(x) = f(a)$$

gilt. Man bezeichnet eine Funktion als in einem Intervall stetig, falls sie in jedem Punkt des Intervalls stetig ist.

Beispiel: Wir betrachten die Heaviside-Funktion, definiert durch

$$\Theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Für diese Funktion gilt $\Theta(0) = 0$, aber

$$\lim_{x \rightarrow 0^+} \Theta(x) = 1.$$

Die Heaviside-Funktion ist im Punkte 0 nicht stetig.

Beispiele von Funktionen, die auf ganz \mathbb{R} stetig sind, sind Polynomfunktionen, $\exp x$, $\sin x$, $\cos x$, $\sinh x$, $\cosh x$.

Satz: Seien $f, g : D \rightarrow \mathbb{R}$ Funktionen, die in a stetig sind und sei $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen

$$\begin{aligned} f + g & : D \rightarrow \mathbb{R}, \\ \lambda \cdot f & : D \rightarrow \mathbb{R}, \\ f \cdot g & : D \rightarrow \mathbb{R} \end{aligned}$$

im Punkte a stetig. Ist ferner $g(a) \neq 0$, so ist auch die Funktion

$$\frac{f}{g} : D' \rightarrow \mathbb{R}$$

in a stetig, wobei $D' = \{x \in D : g(x) \neq 0\}$.

Wir definieren noch den Begriff der **gleichmäßigen Stetigkeit**: Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt in D gleichmäßig stetig, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß

$$|f(x) - f(y)| < \varepsilon \quad \forall |x - y| < \delta.$$

Jede Funktion, die auf D gleichmäßig stetig ist, ist auch in jedem Punkte aus D stetig im herkömmlichen Sinne. Die Umkehrung gilt jedoch nicht. Ist eine Funktion in jedem Punkte $x \in D$ stetig im herkömmlichen Sinne, so genügt es für ein vorgegebenes ε für jeden Punkt ein δ_x zu finden. Dieses δ_x darf mit x variieren. Für die gleichmäßige Stetigkeit wird dagegen gefordert, daß δ von x unabhängig ist.

5.3.1 Rationale Funktionen

Seien $p(x)$ und $q(x)$ Polynomfunktionen. Unter einer rationalen Funktion versteht man eine Funktion

$$R(x) = \frac{p(x)}{q(x)}.$$

Der Definitionsbereich einer rationalen Funktion ist gegeben durch $D = \{x \in \mathbb{R}, q(x) \neq 0\}$. Aufgrund des obigen Satzes ist eine rationale Funktion in ihrem Definitionsbereich stetig.

Rationale Funktionen können in **Partialbrüche** zerlegt werden. Es sei

$$\begin{aligned} p(x) &= p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0, \\ q(x) &= q_m x^m + q_{m-1} x^{m-1} + \dots + q_1 x + q_0. \end{aligned}$$

Wir wollen weiter annehmen, daß das Nennerpolynom in Linearfaktoren zerfällt und die Faktorisierung des Nennerpolynoms bekannt ist:

$$q(x) = c \prod_{j=1}^r (x - x_j)^{\lambda_j}.$$

Hierbei bezeichnet λ_j die Multiziplicität der Nullstelle x_j . Es gilt $\lambda_1 + \dots + \lambda_r = m$. Über den reellen Zahlen zerfällt ein Polynom nicht notwendigerweise in Linearfaktoren, ein Gegenbeispiel ist $x^2 + 1$. Da der Körper der komplexen Zahlen algebraisch abgeschlossen ist, zerfällt über den komplexen Zahlen jedes Polynom in Linearfaktoren. Die Partialbruchzerlegung gilt sowohl für die reellen als auch die komplexen Zahlen. Hat man im reellen Fall irreduzible Polynome vom Grad größer als Eins, geht man zu den komplexen Zahlen über. Beispiel:

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

Ist die Faktorisierung des Nennerpolynoms in Linearfaktoren bekannt, so läßt sich die rationale Funktion schreiben als

$$\begin{aligned} R(x) &= \frac{p(x)}{q(x)} \\ &= P(x) + \sum_{j=1}^r \sum_{k=1}^{\lambda_j} \frac{a_{jk}}{(x-x_j)^k}, \end{aligned}$$

wobei $P(x)$ ein Polynom vom Grad $\deg p(x) - \deg q(x)$ ist und $a_{jk} \in \mathbb{R}$ bzw. $a_{jk} \in \mathbb{C}$.

Berechnung von $P(x)$ und der Konstanten a_{jk} : $P(x)$ bestimmt sich durch Polynomdivision mit Rest. Wir betrachten als Beispiel die rationale Funktion

$$\frac{x^4 + 3x^3 - 12x^2 - 3x + 18}{(x-2)^2(x+2)}$$

Für das Nennerpolynom haben wir

$$(x-2)^2(x+2) = x^3 - 2x^2 - 4x + 8.$$

Polynomdivision mit Rest liefert

$$\begin{array}{r} (x^4 + 3x^3 - 12x^2 - 3x + 18) : (x^3 - 2x^2 - 4x + 8) = x + 5 + (2x^2 + 9x - 22)/(x^3 - 2x^2 - 4x + 8) \\ -(x^4 - 2x^3 - 4x^2 + 8x) \\ \hline 5x^3 - 8x^2 - 11x + 18 \\ -(5x^3 - 10x^2 - 20x + 40) \\ \hline 2x^2 + 9x - 22 \end{array}$$

Somit ist also $P(x) = x + 5$. Für den Rest verwendet man den Ansatz

$$\frac{2x^2 + 9x - 22}{x^3 - 2x^2 - 4x + 8} = \frac{a_{12}}{(x-2)^2} + \frac{a_{11}}{x-2} + \frac{a_{21}}{x+2}.$$

Man bringt die rechte Seite auf den Hauptnenner

$$\frac{a_{12}}{(x-2)^2} + \frac{a_{11}}{x-2} + \frac{a_{21}}{x+2} = \frac{(a_{11} + a_{21})x^2 + (a_{12} - 4a_{21})x + (2a_{12} - 4a_{11} + 4a_{21})}{x^3 - 2x^2 - 4x + 8}$$

Koeffizientenvergleich liefert ein lineares Gleichungssystem:

$$\begin{aligned} a_{11} + a_{21} &= 2, \\ a_{12} - 4a_{21} &= 9, \\ 2a_{12} - 4a_{11} + 4a_{21} &= -22, \end{aligned}$$

dessen Lösung durch

$$a_{12} = 1, \quad a_{11} = 4, \quad a_{21} = -2$$

gegeben ist. Somit erhalten wir das Ergebnis

$$\frac{x^4 + 3x^3 - 12x^2 - 3x + 18}{(x-2)^2(x+2)} = x + 5 + \frac{1}{(x-2)^2} + \frac{4}{x-2} - \frac{2}{x+2}.$$

Bemerkung 1: Die Koeffizienten der Partialbrüche mit der höchsten Potenz einer Nullstelle lassen sich einfacher bestimmen, indem man im Ansatz mit $(x-x_j)^{\lambda_j}$ multipliziert und dann $x = x_j$ setzt. In unserem Beispiel lassen sich so a_{12} und a_{21} bestimmen:

$$\begin{aligned} a_{12} &= \left. \frac{2x^2 + 9x - 22}{(x-2)^2(x+2)}(x-2)^2 \right|_{x=2} = \left. \frac{2x^2 + 9x - 22}{x+2} \right|_{x=2} = \frac{8 + 18 - 22}{4} = 1, \\ a_{21} &= \left. \frac{2x^2 + 9x - 22}{(x-2)^2(x+2)}(x+2) \right|_{x=-2} = \left. \frac{2x^2 + 9x - 22}{(x-2)^2} \right|_{x=-2} = \frac{8 - 18 - 22}{16} = -2. \end{aligned}$$

Bemerkung 2: Wir haben angenommen, daß das Nennerpolynom in Linearfaktoren faktorisiert. Über dem Körper der komplexen Zahlen \mathbb{C} ist dies immer der Fall, da \mathbb{C} algebraisch abgeschlossen ist. Über dem Körper der reellen Zahlen \mathbb{R} gibt es allerdings irreduzible Polynome vom Grad größer als Eins. Ein Beispiel ist

$$x^2 + 1 = (x-i)(x+i).$$

Dieses Polynom läßt sich über \mathbb{R} nicht in Linearfaktoren zerlegen. Tritt dieser Fall auf, so hat man zwei Möglichkeiten: Zum einen kann man von den reellen Zahlen \mathbb{R} zu den komplexen Zahlen \mathbb{C} übergehen. Die zu bestimmenden Koeffizienten a_{jk} sind dann im Allgemeinen komplexe Zahlen.

Die zweite Methode besteht darin, nur mit reellen Zahlen zu arbeiten: Sei

$$q(x) = c \prod_{j=1}^r [q_j(x)]^{\lambda_j}$$

die Zerlegung des Nennerpolynoms in über \mathbb{R} irreduzible Polynome $q_j(x)$. Wir setzen $d_j = \deg q_j(x)$ und setzen voraus, daß der Koeffizient der höchsten Potenz x^{d_j} in $q_j(x)$ gleich Eins ist. Der Ansatz für die Partialbruchzerlegung lautet dann

$$R(x) = P(x) + \sum_{j=1}^r \sum_{k=1}^{\lambda_j} \sum_{l=0}^{d_j-k} \frac{a_{jkl} x^l}{[q_j(x)]^k},$$

wobei die zu bestimmenden Koeffizienten a_{jkl} reell sind.

Beispiel:

$$\frac{x^5 + 1}{x^2 + 1} = x^3 - x + \frac{x+1}{x^2 + 1}.$$

Alternativ erhalten wir über den komplexen Zahlen

$$\frac{x^5 + 1}{x^2 + 1} = x^3 - x + \frac{1(1-i)}{2(x-i)} + \frac{1(1+i)}{2(x+i)}.$$

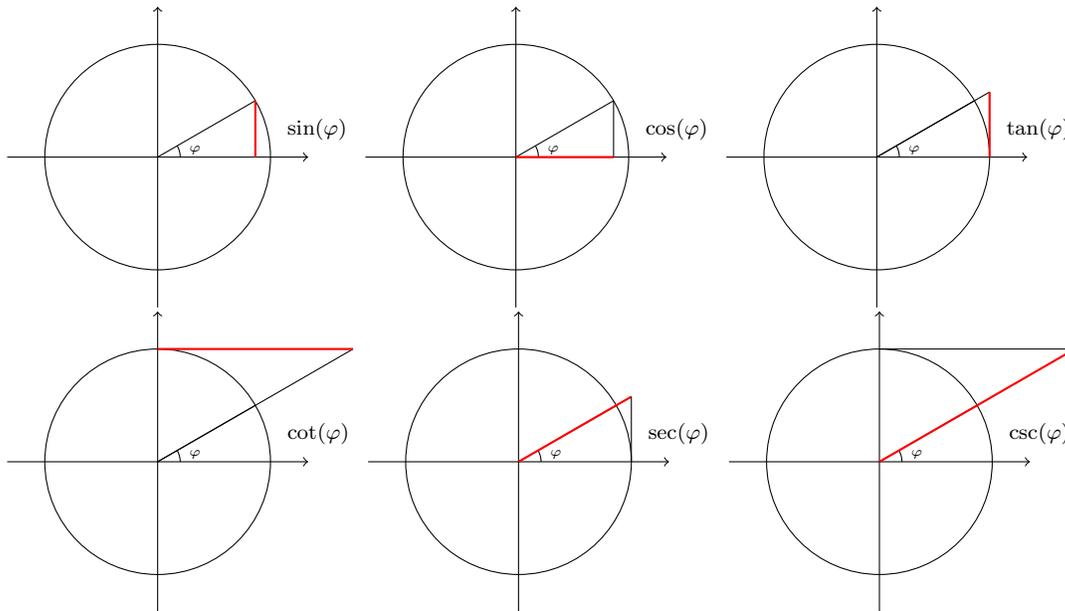


Abbildung 2: Die geometrische Interpretation der Funktionen sin, cos, tan, cot, sec und csc am Einheitskreis.

5.3.2 Trigonometrische Funktionen

Neben den Winkelfunktionen Sinus und Kosinus

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}),$$

gibt es weitere trigonometrische Funktionen:

$$\begin{aligned} \tan x &= \frac{\sin x}{\cos x}, & \text{Tangens} \\ \cot x &= \frac{\cos x}{\sin x}, & \text{Kotangens} \\ \sec x &= \frac{1}{\cos x}, & \text{Sekans} \\ \csc x &= \frac{1}{\sin x}, & \text{Kosekans} \end{aligned}$$

Die geometrische Interpretation dieser Funktionen ist in Abbildung 2 gezeigt. Die Umkehrfunktionen werden mit arcsin, arccos, arctan, etc. bezeichnet:

$$\begin{aligned} \arcsin(x) &= \sin^{-1}(x), & \text{Arkussinus} \\ \arccos(x) &= \cos^{-1}(x), & \text{Arkuskosinus} \\ \arctan(x) &= \tan^{-1}(x), & \text{Arkustangens} \end{aligned}$$

Diese Umkehrfunktionen lassen sich durch den Logarithmus ausdrücken:

$$\begin{aligned}\arcsin(x) &= \frac{1}{i} \ln(ix + \sqrt{1-x^2}), \\ \arccos(x) &= \frac{1}{i} \ln(x + i\sqrt{1-x^2}), \\ \arctan(x) &= \frac{1}{2i} \ln\left(\frac{1+ix}{1-ix}\right).\end{aligned}$$

5.3.3 Hyperbolische Funktionen

Neben den bereits eingeführten hyperbolischen Funktionen

$$\cosh x = \frac{1}{2}(e^x + e^{-x}), \quad \sinh x = \frac{1}{2}(e^x - e^{-x}),$$

definiert man auch

$$\tanh x = \frac{\sinh x}{\cosh x}.$$

Bemerkung: Für \sinh und \cosh gilt

$$\cosh^2 x - \sinh^2 x = 1.$$

Die inversen Funktionen werden als Areafunktionen bezeichnet:

$$\begin{aligned}\operatorname{arsinh}(x) &= \sinh^{-1}(x), \quad \text{Areasinus Hyperbolicus} \\ \operatorname{arcosh}(x) &= \cosh^{-1}(x), \quad \text{Areakosinus Hyperbolicus} \\ \operatorname{artanh}(x) &= \tanh^{-1}(x), \quad \text{Areatangens Hyperbolicus}\end{aligned}$$

Diese Umkehrfunktionen lassen sich ebenfalls durch den Logarithmus ausdrücken:

$$\begin{aligned}\operatorname{arsinh}(x) &= \ln(x + \sqrt{x^2 + 1}), \\ \operatorname{arcosh}(x) &= \ln(x + \sqrt{x^2 - 1}), \\ \operatorname{artanh}(x) &= \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right).\end{aligned}$$

Zur geometrischen Interpretation betrachten wir zunächst nochmal den Einheitskreis $x^2 + y^2 = 1$. Wir betrachten ein Kreissegment mit Winkel 2φ wie in der linken Abbildung in Abbildung 3 gezeigt. Dann ist die Fläche dieses Kreissegments φ . Wir können die trigonometrischen Funktionen daher anstelle als Funktionen des Winkels auch als Funktionen einer Fläche betrachten. Diese letztere Interpretation überträgt sich auf den hyperbolischen Fall

$$x^2 - y^2 = 1.$$

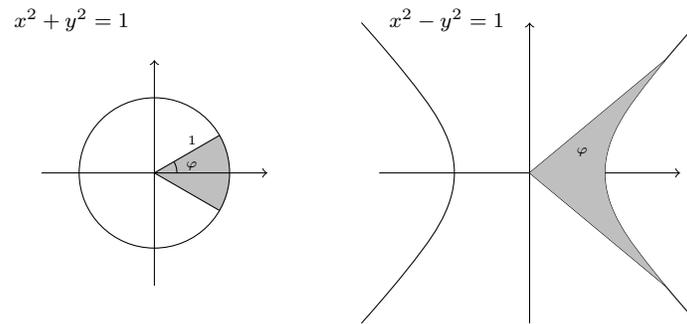


Abbildung 3: Die trigonometrischen Funktionen können sowohl als Funktionen des Winkels als auch als Funktionen einer Fläche betrachtet werden. Letzteres überträgt sich auf den hyperbolischen Fall.

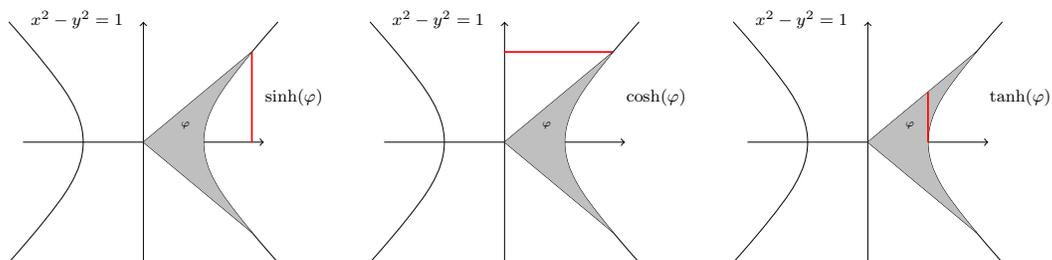


Abbildung 4: Die geometrische Interpretation der Funktionen \sinh , \cosh und \tanh .

Es sei φ die Fläche wie in der rechten Abbildung in Abbildung 3 dargestellt. Die geometrische Interpretation der hyperbolischen Funktionen \sinh , \cosh und \tanh ist in Abbildung 4 gezeigt.

Für den Zusammenhang zwischen den trigonometrischen Funktionen und den hyperbolischen Funktionen gilt

$$\begin{aligned}\sin x &= \frac{1}{i} \sinh(ix), \\ \cos x &= \cosh(ix), \\ \tan x &= \frac{1}{i} \tanh(ix), \\ \arcsin(x) &= \frac{1}{i} \operatorname{arsinh}(ix), \\ \arccos(x) &= \frac{1}{i} \operatorname{arcosh}(x), \\ \arctan(x) &= \frac{1}{i} \operatorname{artanh}(ix).\end{aligned}$$

5.4 Differentialrechnung

Sei $D \subset \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. f nennt man im Punkte $x \in D$ **differenzierbar**, falls es mindestens eine Folge $(\xi_n) \in D \setminus x$ mit $\lim_{n \rightarrow \infty} \xi_n = x$ gibt und für jede solche Folge der Grenzwert

$$f'(x) = \lim_{n \rightarrow \infty} \frac{f(\xi_n) - f(x)}{\xi_n - x} = \lim_{\xi \rightarrow x} \frac{f(\xi) - f(x)}{\xi - x}$$

existiert. Man schreibt auch

$$f'(x) = \frac{df(x)}{dx} = \frac{d}{dx} f(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{(x+h) - x}.$$

Man bezeichnet eine Funktion als in einem Intervall differenzierbar, falls sie in jedem Punkt des Intervalls differenzierbar ist.

Seien $f, g : D \rightarrow \mathbb{R}$ in $x \in D$ differenzierbare Funktionen und $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen $f + g$, $f \cdot g$ und λf in x differenzierbar und es gilt

$$\begin{aligned}(f + g)'(x) &= f'(x) + g'(x), \\ (f \cdot g)'(x) &= f'(x)g(x) + f(x)g'(x), \quad (\text{Produktregel}) \\ (\lambda f)'(x) &= \lambda f'(x).\end{aligned}$$

Dies folgt unmittelbar aus den Rechenregeln für Grenzwerte von Folgen. So ist zum Beispiel

$$\begin{aligned}(f \cdot g)'(x) &= \lim_{h \rightarrow 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} [f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)]\end{aligned}$$

$$\begin{aligned}
&= \lim_{h \rightarrow 0} f(x+h) \frac{g(x+h) - g(x)}{h} + \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} g(x) \\
&= f(x)g'(x) + f'(x)g(x).
\end{aligned}$$

Quotientenregel: Ist weiter $g(x) \neq 0$ für alle $x \in D$, so ist auch die Funktion f/g in x differenzierbar und es gilt

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

Ableitung der Umkehrfunktion: Sei $D \subset \mathbb{R}$ ein abgeschlossenes Intervall, $f : D \rightarrow W$ eine stetige, streng monotone Funktion und $f^{-1} : W \rightarrow D$ die Umkehrfunktion. Ist f im Punkt $x \in D$ differenzierbar und ist $f'(x) \neq 0$, so ist f^{-1} im Punkt $y = f(x)$ differenzierbar und es gilt

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Bemerkung: Die Voraussetzung der strengen Monotonie impliziert nicht $f'(x) \neq 0$, wie das Beispiel $f(x) = x^3$ zeigt: Diese Funktion ist streng monoton, aber $f'(0) = 0$.

Kettenregel: Seien $f : D_1 \rightarrow W_1$ und $g : D_2 \rightarrow W_2$ Funktionen mit $W_1 \subset D_2$. Falls f im Punkte $x \in D_1$ differenzierbar ist und g im Punkte $y = f(x) \in D_2$ differenzierbar ist, so ist die zusammengesetzte Funktion $g \circ f : D_1 \rightarrow W_2$ in x differenzierbar und es gilt

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

Ableitungen einiger Grundfunktionen:

$$\begin{aligned}
f(x) &= x^n, & f'(x) &= nx^{n-1}, \\
f(x) &= e^x, & f'(x) &= e^x.
\end{aligned}$$

Die Ableitung des Logarithmus erhält man mit Hilfe der Regel über die Umkehrfunktion: $f(x) = e^x$, $f^{-1}(y) = \ln y$, $(f^{-1})'(y) = 1/f'(f^{-1}(y)) = 1/\exp(\ln y) = 1/y$, also

$$f(x) = \ln x, \quad f'(x) = \frac{1}{x}.$$

Die Ableitungen von Sinus und Kosinus erhält man aus der Darstellung $\sin(x) = (e^{ix} - e^{-ix})/(2i)$, $\cos(x) = (e^{ix} + e^{-ix})/2$ zu

$$\begin{aligned}
f(x) &= \sin(x), & f'(x) &= \cos(x), \\
f(x) &= \cos(x), & f'(x) &= -\sin(x).
\end{aligned}$$

Die Ableitung aller weiteren trigonometrischen und hyperbolischen Funktionen lassen sich ebenfalls mit den obigen Regeln bestimmen:

$$f(x) = \tan(x), \quad f'(x) = \frac{1}{\cos^2(x)},$$

$$\begin{aligned}
f(x) &= \arcsin(x), & f'(x) &= \frac{1}{\sqrt{1-x^2}}, \\
f(x) &= \arctan(x), & f'(x) &= \frac{1}{1+x^2}, \\
f(x) &= \sinh(x), & f'(x) &= \cosh(x), \\
f(x) &= \cosh(x), & f'(x) &= \sinh(x), \\
f(x) &= \tanh(x), & f'(x) &= \frac{1}{\cosh^2(x)}, \\
f(x) &= \operatorname{arsinh}(x), & f'(x) &= \frac{1}{\sqrt{1+x^2}}, \\
f(x) &= \operatorname{artanh}(x), & f'(x) &= \frac{1}{1-x^2}.
\end{aligned}$$

Höhere Ableitungen: Sei $f : D \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Ist $f' : D \rightarrow \mathbb{R}$ ebenfalls wieder differenzierbar, so bezeichnet man mit

$$f''(x) = \frac{d^2 f(x)}{dx^2} = (f')'(x)$$

die zweite Ableitung. Ist auch $f''(x)$ wieder differenzierbar, so erhält man durch Ableiten die dritte Ableitung $f'''(x)$. Allgemein schreiben wir für die n -te Ableitung

$$f^{(n)}(x) = \frac{d^n f(x)}{dx^n}.$$

Unter der 0-ten Ableitung einer Funktion versteht man die Funktion selbst.

5.4.1 Die Regeln von l'Hospital

Seien $f, g : D \rightarrow \mathbb{R}$ zwei in $x_0 \in D$ stetige Funktionen mit $f(x_0) = g(x_0) = 0$. Weiter seien f und g in einer Umgebung von x_0 differenzierbar. Existiert $\lim_{x \rightarrow x_0} f'(x)/g'(x)$, so gilt:

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Beispiel:

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \lim_{x \rightarrow 0} \frac{\sin x}{2x} = \frac{1}{2} \lim_{x \rightarrow 0} \frac{\cos x}{1} = \frac{1}{2}.$$

Ist $\lim_{x \rightarrow x_0} |f(x)| = \infty$ und $\lim_{x \rightarrow x_0} |g(x)| = \infty$ und existiert $\lim_{x \rightarrow x_0} f'(x)/g'(x)$, so gilt ebenfalls

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Beispiel:

$$\lim_{x \rightarrow 0} \frac{\ln x}{\frac{1}{x}} = \lim_{x \rightarrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \rightarrow 0} (-x) = 0.$$

Bemerkung: Die l'Hospital'schen Regeln gelten auch für $x_0 \rightarrow \pm\infty$.

5.5 Integralrechnung

Definition einer Treppenfunktion: Man nennt $t : [a, b] \rightarrow \mathbb{R}$ **Treppenfunktion**, falls es eine Unterteilung

$$a = x_0 < x_1 < \dots < x_n = b$$

gibt, so daß t auf jedem offenen Intervall $]x_{j-1}, x_j[$ konstant ist. Der Wert auf diesem Intervall sei mit c_j bezeichnet.

Das **Integral einer Treppenfunktion** wird definiert als

$$\int_a^b t(x) dx = \sum_{j=1}^n c_j (x_j - x_{j-1}).$$

Die Menge aller Treppenfunktionen auf dem Intervall $[a, b]$ bilden einen Vektorraum. Wir bezeichnen diesen Vektorraum mit $T[a, b]$. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beliebige beschränkte Funktion und $t \in T[a, b]$. Man schreibt $f \geq t$ falls $f(x) \geq t(x)$ für alle $x \in [a, b]$ gilt.

Wir definieren nun das Ober- und Unterintegral für f :

$$\int_a^{b^*} f(x) dx = \inf \left\{ \int_a^b t(x) dx; t \in T[a, b], t \geq f \right\},$$

$$\int_a^* f(x) dx = \sup \left\{ \int_a^b t(x) dx; t \in T[a, b], t \leq f \right\}.$$

Das **Infimum** einer Menge mit einer Ordnungsrelation $<$ ist die größte untere Schranke. Das Infimum ist nicht notwendigerweise ein Element dieser Menge. Das **Minimum** einer Menge ist dagegen (falls es existiert) das kleinste Element dieser Menge. Es gibt Mengen, die ein Infimum besitzen, aber kein Minimum. Ein Beispiel ist die Menge

$$\left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\}.$$

Das Infimum ist Null, aber die Menge besitzt kein Minimum. Analog definiert man das **Supremum** als die kleinste obere Schranke einer Menge mit einer Ordnungsrelation $<$. Das **Maximum** einer Menge ist (falls es existiert) das größte Element dieser Menge.

Definition des Riemann-Integrals: Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt **Riemann-integrierbar**, falls

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

In diesem Fall setzt man

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

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

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

Seien $f, g : [a, b] \rightarrow \mathbb{R}$ integrierbare Funktionen und $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen $f + g$ und $\lambda \cdot f$ integrierbar und es gilt

$$\begin{aligned} \int_a^b (f + g)(x) dx &= \int_a^b f(x) dx + \int_a^b g(x) dx, \\ \int_a^b (\lambda f)(x) dx &= \lambda \int_a^b f(x) dx. \end{aligned}$$

Desweiteren ist auch die Funktion fg integrierbar, im allgemeinen ist allerdings

$$\int_a^b (f \cdot g)(x) dx \neq \left(\int_a^b f(x) dx \right) \cdot \left(\int_a^b g(x) dx \right).$$

Eine differenzierbare Funktion $F : D \rightarrow \mathbb{R}$ heißt **Stammfunktion** einer Funktion $f : D \rightarrow \mathbb{R}$, falls $F'(x) = f(x)$.

Eine weitere Funktion $G : D \rightarrow \mathbb{R}$ ist genau dann ebenfalls eine Stammfunktion, falls $F - G$ eine Konstante ist.

Man schreibt auch

$$F(x) = \int f(x) dx.$$

Der Ausdruck auf der rechten Seite wird auch als **unbestimmtes Integral** bezeichnet.

Hauptsatz der Differential- und Integralrechnung:

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

Man schreibt auch

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

Stammfunktionen einiger Grundfunktionen:

$$\begin{aligned}
 f(x) &= x^n, & F(x) &= \frac{x^{n+1}}{n+1}, & n &\neq -1, \\
 f(x) &= e^x, & F(x) &= e^x, \\
 f(x) &= \frac{1}{x}, & F(x) &= \ln|x|, \\
 f(x) &= \sin(x), & F(x) &= -\cos(x), \\
 f(x) &= \cos(x), & F(x) &= \sin(x), \\
 f(x) &= \frac{1}{1+x^2}, & F(x) &= \arctan(x).
 \end{aligned}$$

Substitutionsregel: Sei $f : [a, b] \rightarrow W_1$ eine stetig differenzierbare Funktion und $g : D_2 \rightarrow W_2$ eine stetige Funktion mit $W_1 \subset D_2$. Dann gilt

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

Die Substitutionsregel ist die Umkehrung der Kettenregel: Sei $G(x)$ eine Stammfunktion zu $g(x)$. Dann ist

$$\frac{d}{dx} G(f(x)) = g(f(x)) \cdot f'(x)$$

und somit

$$\int_a^b g(f(x)) f'(x) dx = G(f(b)) - G(f(a)).$$

Andererseits ist aber auch

$$\int_{f(a)}^{f(b)} g(x) dx = G(f(b)) - G(f(a)),$$

womit die Substitutionsregel bewiesen ist.

Beispiel: Wir betrachten das Integral

$$I = \int_0^\pi d\theta \sin \theta (5 \cos^2 \theta + 3 \cos \theta + 1).$$

Für die Substitution $u = -\cos \theta$ gilt

$$\frac{du}{d\theta} = \sin \theta$$

und daher ergibt sich mit Hilfe der Substitutionsregel

$$I = \int_{-1}^1 du (5u^2 - 3u + 1) = \left(\frac{5}{3}u^3 - \frac{3}{2}u^2 + u \right) \Big|_{-1}^1 = \frac{16}{3}.$$

Partielle Integration: Seien $f, g : [a, b] \rightarrow \mathbb{R}$ zwei stetig differenzierbare Funktionen. Dann gilt

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

Die partielle Integration ist die Umkehrung der Produktregel. Es sei $H(x) = f(x) \cdot g(x)$ und

$$h(x) = H'(x) = f(x) \cdot g'(x) + f'(x) \cdot g(x).$$

Somit ist

$$\int_a^b h(x) dx = H(x) \Big|_a^b.$$

Setzen wir nun die Ausdrücke für $h(x)$ und $H(x)$ ein, so erhalten wir

$$\int_a^b f(x) \cdot g'(x) dx + \int_a^b f'(x) \cdot g(x) dx = f(x) \cdot g(x) \Big|_a^b.$$

Beispiel: Wir betrachten das Integral

$$I = \int_0^1 dx x e^x.$$

Setzen wir $f(x) = x$ und $g'(x) = \exp(x)$, so läßt sich die partielle Integration anwenden, falls wir eine Stammfunktion zu $g'(x)$ kennen. In diesem Beispiel ist dies besonders einfach, es ist $g(x) = \exp(x)$. Somit erhalten wir

$$I = x e^x \Big|_0^1 - \int_0^1 dx e^x = (x-1) e^x \Big|_0^1 = 1.$$

Integrale über rationale Funktionen: Wir betrachten als Beispiel

$$I = \int_0^1 \frac{x^4 + 3x^3 - 12x^2 - 3x + 18}{(x-2)^2(x+2)} dx.$$

Im ersten Schritt zerlegt man den Integranden mit Hilfe der Partialbruchzerlegung:

$$\frac{x^4 + 3x^3 - 12x^2 - 3x + 18}{(x-2)^2(x+2)} = x + 5 + \frac{1}{(x-2)^2} + \frac{4}{x-2} - \frac{2}{x+2}.$$

Somit ist

$$I = \int_0^1 (x+5) dx + \int_0^1 \frac{dx}{(x-2)^2} + 4 \int_0^1 \frac{dx}{x-2} - 2 \int_0^1 \frac{dx}{x+2}.$$

Wir berechnen nun die einzelnen Integrale:

$$\int_0^1 (x+5) dx = 5x + \frac{1}{2}x^2 \Big|_0^1 = \frac{11}{2},$$

$$\int_0^1 \frac{dx}{(x-2)^2} = -\frac{1}{x-2} \Big|_0^1 = 1 - \frac{1}{2} = \frac{1}{2},$$

$$\int_0^1 \frac{dx}{x-2} = \ln(|x-2|) \Big|_0^1 = \ln 1 - \ln 2 = -\ln 2,$$

$$\int_0^1 \frac{dx}{x+2} = \ln(|x+2|) \Big|_0^1 = \ln 3 - \ln 2,$$

Somit erhalten wir

$$I = \frac{11}{2} + \frac{1}{2} + 4(-\ln 2) - 2(\ln 3 - \ln 2) = 6 - 2 \ln 2 - 2 \ln 3 = 6 - 2 \ln 6.$$

5.5.1 Uneigentliche Integrale

Unter einem uneigentlichen Integral versteht man ein Integral, bei dem eine Integrationsgrenze unendlich ist oder bei dem der Integrand an einer Integrationsgrenze nicht definiert ist. Es kann auch eine Kombination der beiden Fälle auftreten.

Wir betrachten zunächst den Fall, daß eine Integrationsgrenze unendlich ist. Sei $f : [a, \infty[\rightarrow \mathbb{R}$ eine Funktion, die über jedem Intervall $[a, \Lambda]$ mit $a < \Lambda < \infty$ Riemann-integrierbar ist. Falls der Grenzwert

$$\lim_{\Lambda \rightarrow \infty} \int_a^\Lambda f(x) dx$$

existiert, nennt man das Integral von a bis Unendlich konvergent und man setzt

$$\int_a^{\infty} f(x) dx = \lim_{\Lambda \rightarrow \infty} \int_a^{\Lambda} f(x) dx.$$

Analog definiert man das Integral für das Intervall $] - \infty, b]$.

Beispiel:

$$\int_1^{\infty} \frac{dx}{x^2} = \lim_{\Lambda \rightarrow \infty} \int_1^{\Lambda} \frac{dx}{x^2} = - \lim_{\Lambda \rightarrow \infty} \frac{1}{x} \Big|_1^{\Lambda} = 1 - \lim_{\Lambda \rightarrow \infty} \frac{1}{\Lambda} = 1.$$

Wir betrachten nun den Fall, daß der Integrand an einer Intervallgrenze nicht definiert ist. Sei $f :]a, b] \rightarrow \mathbb{R}$ eine Funktion, die über jedem Teilintervall $[a + \varepsilon, b]$ mit $0 < \varepsilon < (b - a)$ Riemann-integrierbar ist. Falls der Grenzwert

$$\lim_{\varepsilon \rightarrow 0^+} \int_{a+\varepsilon}^b f(x) dx$$

existiert, nennt man das Integral über $]a, b]$ konvergent und man setzt

$$\int_a^b f(x) dx = \lim_{\varepsilon \rightarrow 0^+} \int_{a+\varepsilon}^b f(x) dx$$

Analog definiert man das Integral für den Fall in der die Funktion an der oberen Intervallgrenze nicht definiert ist.

Beispiel:

$$\int_0^1 \frac{dx}{\sqrt{x}} = \lim_{\varepsilon \rightarrow 0^+} \int_{\varepsilon}^1 \frac{dx}{\sqrt{x}} = 2 \lim_{\varepsilon \rightarrow 0^+} \sqrt{x} \Big|_{\varepsilon}^1 = 2 - 2 \lim_{\varepsilon \rightarrow 0^+} \sqrt{\varepsilon} = 2.$$

Allgemeiner Fall: Sei $f :]a, b[\rightarrow \mathbb{R}$, $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$, eine Funktion, die über jedem Teilintervall $[\alpha, \beta] \subset]a, b[$ Riemann-integrierbar ist und sei $c \in]a, b[$ beliebig. Falls die beiden uneigentlichen Integrale

$$\int_a^c f(x) dx = \lim_{\alpha \rightarrow a^+} \int_{\alpha}^c f(x) dx, \quad \int_c^b f(x) dx = \lim_{\beta \rightarrow b^-} \int_c^{\beta} f(x) dx$$

existieren, nennt man das Integral über $]a, b[$ konvergent und man setzt

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

5.6 Taylorreihen

Wir haben bereits die Reihendarstellung einiger Funktionen, wie zum Beispiel der Exponentialfunktion, Sinus oder Kosinus kennengelernt. In diesem Abschnitt geht es um die systematische Entwicklung von Funktionen in Potenzreihen.

Taylorische Formel: Sei $I \subset \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal stetig differenzierbare Funktion. Dann gilt für $a \in I$ und $x \in I$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x).$$

Hierbei bezeichnet $f^{(n)}$ die n -te Ableitung. Für das Restglied gilt

$$R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt.$$

Eine alternative Form für das Restglied geht auf Lagrange zurück: Es gibt ein ξ zwischen a und x (d.h. $\xi \in [a, x]$ für $x > a$ bzw. $\xi \in [x, a]$ für $x < a$), so daß

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1}.$$

Bemerkung: Dies ist eine Existenzaussage, ξ ist im allgemeinen schwer zu bestimmen.

In der Praxis verwendet man die ersten n Terme der Taylorentwicklung, um eine Funktion zu approximieren und vernachlässigt das Restglied. Das vernachlässigte Restglied liefert den Fehler dieser Abschätzung.

Sei nun $f : I \rightarrow \mathbb{R}$ eine beliebig oft differenzierbare Funktion und $a \in I$. Wir definieren die Taylorreihe einer Funktion f um den Entwicklungspunkt a :

$$T_f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n.$$

Bemerkungen:

1. Der Konvergenzradius der Taylorreihe ist nicht notwendig > 0 .
2. Falls die Taylorreihe von f konvergiert, konvergiert sie nicht notwendigerweise gegen f .
3. Die Taylorreihe konvergiert genau für diejenigen $x \in I$ gegen $f(x)$, für die das Restglied gegen Null konvergiert.

Wir geben ein Gegenbeispiel zu Punkt 2 an: Wir betrachten die Taylorreihe der Funktion

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & x \neq 0, \\ 0 & x = 0, \end{cases}$$

im Punkte $a = 0$. f ist beliebig oft differenzierbar und es gilt

$$f^{(n)}(0) = 0.$$

Die Taylorreihe von f um den Nullpunkt ist also identisch Null.

Beispiel 1: In der speziellen Relativitätstheorie ist die Gesamtenergie eines Teilchens mit Geschwindigkeit \vec{v} und Ruhemasse m gegeben durch

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Für $|\vec{v}| \ll c$ ist $x = \frac{v^2}{c^2}$ eine kleine Größe und wir können in x entwickeln:

$$\frac{1}{\sqrt{1-x}} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \mathcal{O}(x^4).$$

Somit erhalten wir für die Gesamtenergie des Teilchens

$$E = mc^2 + \frac{1}{2}mv^2 + mc^2\mathcal{O}\left(\frac{v^4}{c^4}\right).$$

Der erste Term gibt die Ruheenergie an, der zweite Term die nicht-relativistische kinetische Energie.

Beispiel 2: Wir betrachten den Fall, daß sich ein nicht-relativistisches Teilchen in einem Potential

$$V(x) = V_0 \left(1 - \cos \frac{x}{x_0}\right)$$

bewegt. Weiter wollen wir annehmen, daß die Auslenkungen klein sind, d.h. es gilt $|x| \ll x_0$. In diesem Fall liefert die Taylorreihe eine gute Näherung:

$$V(x) = V_0 \left[\frac{1}{2} \frac{x^2}{x_0^2} - \frac{1}{24} \frac{x^4}{x_0^4} \right] + \mathcal{O}\left(\frac{x^6}{x_0^6}\right).$$

Berücksichtigen wir nur den führenden Term der Taylorreihe,

$$V(x) = \frac{1}{2}V_0 \frac{x^2}{x_0^2} + \mathcal{O}\left(\frac{x^4}{x_0^4}\right),$$

so reduziert sich das Problem auf den harmonischen Oszillator.

5.7 Orthogonale Polynome

Wir betrachten Polynomfunktionen auf einem Intervall $[a, b]$. Die Menge aller Polynomfunktionen $P : [a, b] \rightarrow \mathbb{R}$ bilden einen (unendlich dimensionalen) \mathbb{R} -Vektorraum. Eine Basis dieses Vektorraumes ist zum Beispiel durch die Monome

$$1, x, x^2, x^3, x^4, \dots$$

gegeben. Ebenso bildet eine Sequenz von Polynomen $P_0(x), P_1(x), \dots$, wobei $P_j(x)$ vom Grad j ist, eine Basis dieses Vektorraumes.

Wir wollen nun ein Skalarprodukt auf diesem Vektorraum definieren. Zunächst führen wir den Begriff einer **Gewichtsfunktion** ein. Man nennt eine Funktion $w : [a, b] \rightarrow \mathbb{R}$ eine Gewichtsfunktion, falls

1. $w(x) \geq 0$ für alle $x \in [a, b]$;
2. $\int_a^b dx w(x) > 0$;
3. $\int_a^b dx w(x)x^j < \infty$ für alle $j = 0, 1, 2, \dots$

Bemerkung: Es ist nicht gefordert, daß die Gewichtsfunktion ein Polynom sein muß.

Seien nun $P_i(x)$ und $P_j(x)$ zwei Polynome. Mit Hilfe einer Gewichtsfunktion definieren wir nun ein Skalarprodukt wie folgt:

$$\langle P_i, P_j \rangle = \int_a^b dx w(x) P_i(x) P_j(x).$$

Bemerkung: Diese Definition hängt von der Wahl der Gewichtsfunktion ab !

Man nennt zwei Polynome orthogonal, falls

$$\langle P_i, P_j \rangle = \int_a^b dx w(x) P_i(x) P_j(x) = 0.$$

Gesucht ist nun zu einer gegebenen Gewichtsfunktion eine Basis des Vektorraumes der Polynome über $[a, b]$, die orthogonal bezüglich des Skalarproduktes ist. In anderen Worten, gesucht ist eine Sequenz von Polynomen $P_0(x), P_1(x), \dots$, wobei $P_j(x)$ vom Grad j ist, so daß

$$\int_a^b dx w(x) P_i(x) P_j(x) = 0 \quad \text{für } i \neq j.$$

Die folgende Tabelle zeigt die Intervalle und die Gewichtsfunktionen der üblichen orthogonalen Polynome:

Intervall	Gewichtsfunktion	Polynome	Bemerkungen
$[-1, 1]$	1	Legendre	
$[-1, 1]$	$1/\sqrt{1-x^2}$	Tschebyscheff	
$[-1, 1]$	$(1-x^2)^{\mu-1/2}$	Gegenbauer	$\mu > -1/2$
$[-1, 1]$	$(1-x)^\alpha(1+x)^\beta$	Jacobi	$\alpha, \beta > -1$
$[0, \infty]$	$x^\alpha e^{-x}$	(verallgemeinerte) Laguerre	$\alpha > -1$
$[-\infty, \infty]$	e^{-x^2}	Hermite	

Die eigentlichen Laguerre-Polynome entsprechen $\alpha = 0$.

Für die Nullstellen von orthogonalen Polynomen gilt: Die Nullstellen von orthogonalen Polynomen sind reell, einfach und befinden sich im Intervall $[a, b]$. Seien $x_1 < x_2 < \dots < x_n$ die Nullstellen des Polynoms $P_n(x)$, dann liegt in jedem Intervall $[a, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n], [x_n, b]$ genau eine Nullstelle des Polynoms $P_{n+1}(x)$.

5.7.1 Legendre-Polynome

Die Legendre-Polynome $P_n(x)$ sind auf dem Intervall $[-1, 1]$ mit der Gewichtsfunktion $w(x) = 1$ definiert. Sie sind wie folgt normiert:

$$\int_{-1}^1 dx P_n(x) P_m(x) = \frac{2}{2n+1} \delta_{nm}.$$

Explizit sind die Polynome gegeben durch

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{[n/2]} (-1)^m \binom{n}{m} \binom{2n-2m}{n} x^{n-2m},$$

wobei $[n/2]$ die größte ganze Zahl kleiner oder gleich $n/2$ bezeichnet. Rekursiv erhält man diese Polynome durch

$$P_0(x) = 1, \quad P_1(x) = x, \quad (n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$\frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} P_n(x) z^n, \quad -1 \leq x \leq 1, \quad |z| < 1.$$

Formel von Rodrigues:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

Die Legendre-Polynome erfüllen die Differentialgleichung

$$(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0.$$

5.7.2 Tschebyscheff-Polynome

Die Tschebyscheff-Polynome der ersten Art $T_n(x)$ sind auf dem Intervall $[-1, 1]$ mit der Gewichtsfunktion $w(x) = 1/\sqrt{1-x^2}$ definiert. Sie sind wie folgt normiert:

$$\int_{-1}^1 dx \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) = \begin{cases} \frac{\pi}{2} \delta_{nm}, & n \neq 0, \\ \pi \delta_{nm}, & n = 0. \end{cases}$$

Explizit sind die Polynome gegeben durch

$$T_n(x) = \frac{n}{2} \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \frac{(n-m-1)!}{m!(n-2m)!} (2x)^{n-2m}.$$

Rekursiv erhält man diese Polynome durch

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$\frac{1-xz}{1-2xz+z^2} = \sum_{n=0}^{\infty} T_n(x) z^n, \quad -1 \leq x \leq 1, \quad |z| < 1.$$

Formel von Rodrigues:

$$T_n(x) = \frac{(-1)^n (1-x^2)^{1/2} \sqrt{\pi}}{2^n \Gamma(n + \frac{1}{2})} \frac{d^n}{dx^n} (1-x^2)^{n-1/2}.$$

Die Tschebyscheff-Polynome erfüllen die Differentialgleichung

$$(1-x^2)T_n''(x) - xT_n'(x) + n^2T_n(x) = 0.$$

5.7.3 Laguerre-Polynome

Die Laguerre-Polynome $L_n(x)$ sind auf dem Intervall $[0, \infty]$ mit der Gewichtsfunktion $w(x) = e^{-x}$ definiert. Sie sind wie folgt normiert:

$$\int_0^{\infty} dx e^{-x} L_n(x) L_m(x) = \delta_{nm}.$$

Explizit sind die Polynome gegeben durch

$$L_n = \sum_{m=0}^n \frac{(-1)^m}{m!} \binom{n}{n-m} x^m$$

Rekursiv erhält man diese Polynome durch

$$L_0(x) = 1, \quad L_1(x) = 1 - x, \quad (n + 1)L_{n+1}(x) = ((2n + 1) - x)L_n(x) - nL_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$(1 - z)^{-1} \exp\left(\frac{xz}{z - 1}\right) = \sum_{n=0}^{\infty} L_n(x)z^n, \quad |z| < 1.$$

Formel von Rodrigues:

$$L_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n} (x^n e^{-x})$$

Die Laguerre-Polynome erfüllen die Differentialgleichung

$$xL_n''(x) + (1 - x)L_n'(x) + nL_n(x) = 0.$$

5.7.4 Hermite-Polynome

Die Hermite-Polynome $H_n(x)$ sind auf dem Intervall $[-\infty, \infty]$ mit der Gewichtsfunktion $w(x) = e^{-x^2}$ definiert. Sie sind wie folgt normiert:

$$\int_{-\infty}^{\infty} dx e^{-x^2} H_n(x) H_m(x) = 2^n \sqrt{\pi} n! \delta_{nm}.$$

Explizit sind die Polynome gegeben durch

$$H_n(x) = n! \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \frac{(2x)^{n-2m}}{m!(n-2m)!}.$$

Rekursiv erhält man diese Polynome durch

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$e^{-t^2+2xt} = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) t^n$$

Formel von Rodrigues:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$$

Die Hermite-Polynome erfüllen die Differentialgleichung

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0.$$

6 Differentialgleichungen

Wir wollen nun Gleichungen betrachten, die eine unbekannte Funktion und deren Ableitung enthalten. Diese Gleichungen nennt man **Differentialgleichungen**. Tritt nur die Ableitung nach einer Variablen auf, spricht man von einer **gewöhnlichen Differentialgleichung**. Hängt dagegen die gesuchte Funktion von mehreren Variablen ab, und treten Ableitungen nach verschiedenen Variablen auf, so spricht man von einer **partiellen Differentialgleichung**. Wir behandeln zunächst gewöhnliche Differentialgleichungen. Beispiele sind

$$(i) \quad f_1'(x) - b = 0,$$

$$(ii) \quad f_2''(x) - a = 0,$$

$$(iii) \quad f_3'(x) + \lambda f_3(x) = 0,$$

$$(iv) \quad f_4''(x) + \omega^2 f_4(x) = 0.$$

Durch Differenzieren überzeugt man sich, daß

$$(i) \quad f_1(x) = bx + c_0,$$

$$(ii) \quad f_2(x) = \frac{1}{2}ax^2 + c_1x + c_0,$$

$$(iii) \quad f_3(x) = c_0e^{-\lambda x},$$

$$(iv) \quad f_4(x) = c_1 \sin(\omega x) + c_2 \cos(\omega x)$$

Lösungen sind. Hierbei sind c_0, c_1, c_2 frei wählbare Konstanten. Wir möchten nun Methoden betrachten, die uns erlauben diese Lösungen zu bestimmen. Die vier obigen Beispiele sind die am häufigsten vorkommenden Differentialgleichungen in der Physik: Beispiel (i) entspricht der Bewegungsgleichung eines Teilchens mit konstanter Geschwindigkeit, Beispiel (ii) entspricht der Bewegungsgleichung eines Teilchens mit konstanter Beschleunigung, Beispiel (iii) beschreibt den radioaktiven Zerfall und Beispiel (iv) den harmonischen Oszillator.

Definition: Sei D eine Teilmenge von \mathbb{R}^2 und

$$G : D \rightarrow \mathbb{R}, \\ (x, y) \rightarrow G(x, y)$$

eine stetige Funktion. Dann nennt man

$$y' = G(x, y)$$

eine **Differentialgleichung erster Ordnung**. Unter einer Lösung versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

$$f : I \rightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

- Der Graph von f ist in D enthalten, d.h.

$$\Gamma_f = \{(x, y) \in I \times \mathbb{R} : y = f(x)\} \subset D.$$

- Es gilt

$$f'(x) = G(x, f(x)).$$

Beispiel: $G(x, y) = -\lambda y$ führt auf die Differentialgleichung

$$f'(x) = -\lambda f(x).$$

Bemerkung: Hängt die Funktion G nur von x , aber nicht von y ab, so hat man

$$f'(x) = G(x)$$

und man erhält eine Lösung durch Integration:

$$f(x) = c + \int_{x_0}^x G(x) dx.$$

Wir definieren noch Differentialgleichungen höherer Ordnungen: Hierzu betrachten wir zunächst Systeme von Differentialgleichungen erster Ordnung und zeigen dann, daß sich Differentialgleichungen n -ter Ordnung auf ein System von n Differentialgleichungen erster Ordnung reduzieren lassen.

Sei D eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned} \vec{G} &: D \rightarrow \mathbb{R}^n, \\ (x, \vec{y}) &\rightarrow \vec{G}(x, \vec{y}) \end{aligned}$$

eine stetige Funktion. Dann nennt man

$$\vec{y}' = \vec{G}(x, \vec{y})$$

ein **System von n Differentialgleichungen erster Ordnung**. Unter einer Lösung versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

$$\vec{f} : I \rightarrow \mathbb{R}^n$$

mit folgenden Eigenschaften:

- Der Graph von \vec{f} ist in D enthalten.
- Es gilt

$$\vec{f}'(x) = \vec{G}(x, \vec{f}(x)).$$

\vec{G} and \vec{f} sind Vektoren. Schreiben wir diese in Komponenten

$$\vec{G} = \begin{pmatrix} G_1 \\ \dots \\ G_n \end{pmatrix}, \quad \vec{f} = \begin{pmatrix} f_1 \\ \dots \\ f_n \end{pmatrix},$$

so ergibt sich ein System von n Differentialgleichungen erster Ordnung:

$$\begin{aligned} f_1'(x) &= G_1(x, f_1(x), \dots, f_n(x)), \\ f_2'(x) &= G_2(x, f_1(x), \dots, f_n(x)), \\ &\dots \\ f_n'(x) &= G_n(x, f_1(x), \dots, f_n(x)). \end{aligned}$$

Beispiel: Für $n = 2$ führt die Abbildung

$$\vec{G}(x, y_0, y_1) = \begin{pmatrix} y_1 \\ -\omega^2 y_0 \end{pmatrix}$$

auf das System

$$\begin{aligned} f_0'(x) &= f_1(x), \\ f_1'(x) &= -\omega^2 f_0(x). \end{aligned}$$

Sei D eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned} \tilde{G} &: D \rightarrow \mathbb{R}, \\ (x, \vec{y}) &\rightarrow \tilde{G}(x, \vec{y}) \end{aligned}$$

eine stetige Funktion. Dann nennt man

$$y^{(n)} = \tilde{G}(x, y, y', \dots, y^{(n-1)})$$

eine **Differentialgleichungen n -ter Ordnung**. Unter einer Lösung versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

$$f : I \rightarrow \mathbb{R}$$

mit folgenden Eigenschaften:

- Die Menge

$$\{(x, y_0, y_1, \dots, y_{n-1}) \in I \times \mathbb{R}^n : y_j = f^{(j)}(x), 0 \leq j < n\}$$

ist in D enthalten.

- Es gilt

$$f^{(n)}(x) = \tilde{G}(x, f(x), f'(x), \dots, f^{(n-1)}(x)).$$

Beispiel: $\tilde{G}(x, y_1, y_2) = -\omega^2 y_1$ führt auf die Differentialgleichung zweiter Ordnung

$$f''(x) = -\omega^2 f(x).$$

6.1 Reduktion einer Differentialgleichung höherer Ordnung auf ein System erster Ordnung

Sei nun eine Differentialgleichung n -ter Ordnung gegeben:

$$y^{(n)} = \tilde{G}(x, y, y', \dots, y^{(n-1)})$$

Für

$$\vec{y} = \begin{pmatrix} y_0 \\ \dots \\ y_{n-2} \\ y_{n-1} \end{pmatrix}, \quad \vec{G} = \begin{pmatrix} y_1 \\ \dots \\ y_{n-1} \\ \tilde{G}(x, \vec{y}) \end{pmatrix},$$

betrachten wir das System von n Differentialgleichungen erster Ordnung:

$$\vec{y}' = \vec{G}(x, \vec{y})$$

Ausgeschrieben in Komponenten haben wir

$$\begin{aligned} y_0' &= y_1, \\ y_1' &= y_2, \\ &\dots \\ y_{n-2}' &= y_{n-1}, \\ y_{n-1}' &= \tilde{G}(x, y_0, y_1, \dots, y_{n-1}). \end{aligned}$$

Wir betrachten nun die n -te Ableitung von y_0 :

$$\begin{aligned} y_0^{(n)} &= y_1^{(n-1)} = y_2^{(n-2)} = \dots = y_{n-1}' = \tilde{G}(x, y_0, y_1, \dots, y_{n-1}) \\ &= \tilde{G}(x, y_0, y_0', \dots, y_0^{(n-1)}). \end{aligned}$$

Somit erhält man die Lösung der Differentialgleichung n -ter Ordnung, indem man zunächst das zugehörige System von n Differentialgleichungen erster Ordnung löst. Ist diese Lösung $\vec{f}(x)$, so ist die erste Komponente von \vec{f} Lösung der Differentialgleichung n -ter Ordnung. (Die weiteren Komponenten von \vec{f} sind die Ableitungen der ersten Komponente.)

Beispiel: Zu der Differentialgleichung zweiter Ordnung

$$f''(x) = -\omega^2 f(x).$$

gehört das System von Differentialgleichungen erster Ordnung

$$\begin{aligned} f_0'(x) &= f_1(x), \\ f_1'(x) &= -\omega^2 f_0(x). \end{aligned}$$

6.2 Existenz und Eindeutigkeit von Lösungen

In diesem Abschnitt beschäftigen wir uns mit Existenz und Eindeutigkeit von Lösungen eines Systems von Differentialgleichungen erster Ordnung. Wie wir im obigen Abschnitt gesehen haben, lassen sich Differentialgleichungen höherer Ordnung immer auf ein System erster Ordnung zurückführen.

Definition: Sei D eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned}\vec{G} &: D \rightarrow \mathbb{R}^n, \\ (x, \vec{y}) &\rightarrow \vec{G}(x, \vec{y})\end{aligned}$$

eine Funktion. Man sagt, \vec{G} genügt in D einer **Lipschitz-Bedingung** mit der Lipschitz-Konstanten $L \geq 0$, falls für alle $(x, \vec{y}), (x, \vec{z}) \in D$ gilt

$$|\vec{G}(x, \vec{y}) - \vec{G}(x, \vec{z})| \leq L |\vec{y} - \vec{z}|.$$

Man sagt, \vec{G} genügt in D lokal einer Lipschitz-Bedingung, falls jeder Punkt $(x, \vec{y}) \in D$ eine Umgebung U besitzt, so daß \vec{G} in $D \cap U$ einer Lipschitz-Bedingung mit einer von U abhängigen Konstanten L genügt.

Satz: Sei D offen. Ist die Funktion $\vec{G}(x, \vec{y})$ bezüglich der Variablen $\vec{y} = (y_1, \dots, y_n)$ stetig partiell differenzierbar, so genügt \vec{G} in D lokal einer Lipschitz-Bedingung.

Satz über die **Eindeutigkeit von Lösungen**: Wir setzen voraus, daß die Funktion \vec{G} in D lokal einer Lipschitz-Bedingung genügt. Seien $\vec{f}(x)$ und $\vec{g}(x)$ zwei Lösungen der Differentialgleichung

$$\vec{y}' = \vec{G}(x, \vec{y})$$

über einem Intervall $I \subset \mathbb{R}$. Gilt dann

$$\vec{f}(x_0) = \vec{g}(x_0)$$

für ein $x_0 \in I$, so folgt

$$\vec{f}(x) = \vec{g}(x)$$

für alle $x \in I$.

Satz über die **Existenz von Lösungen** von Picard – Lindelöf: Sei D offen und $\vec{G} : D \rightarrow \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(x_0, \vec{y}_0) \in D$ ein $\varepsilon > 0$ und eine Lösung

$$\vec{f} : [x_0 - \varepsilon, x_0 + \varepsilon] \rightarrow \mathbb{R}^n$$

der Differentialgleichung $\vec{y}' = \vec{G}(x, \vec{y})$ mit der Anfangsbedingung

$$\vec{f}(x_0) = \vec{y}_0.$$

Bemerkung: Überträgt man dies auf eine Differentialgleichung n -ter Ordnung, so wird eine Lösung eindeutig durch n Anfangsbedingungen $f(x_0), f'(x_0), \dots, f^{(n-1)}(x_0)$ bestimmt.

6.3 Elementare Lösungsmethoden

Differentialgleichung mit separierten Variablen: Seien $I, J \subset \mathbb{R}$ offene Intervalle und

$$h : I \rightarrow \mathbb{R}, \quad k : J \rightarrow \mathbb{R},$$

zwei stetige Funktionen. Es gelte ausserdem $k(y) \neq 0$ für alle $y \in J$. Wir betrachten eine Differentialgleichung erster Ordnung

$$y' = G(x, y)$$

Die Funktion G ist auf dem Gebiet $D = I \times J$ definiert und wir nehmen an, daß die Variablen sich trennen lassen:

$$G(x, y) = h(x)k(y).$$

Sei nun $(x_0, y_0) \in I \times J$. Wir setzen

$$H(x) = \int_{x_0}^x h(t) dt, \quad K(y) = \int_{y_0}^y \frac{dt}{k(t)}.$$

Es sei $I' \subset I$ ein Intervall mit $x_0 \in I'$ und $H(I') \subset K(J)$. Dann existiert genau eine Lösung $f : I' \rightarrow \mathbb{R}$ mit

$$f(x_0) = y_0.$$

Diese Lösung erfüllt die Beziehung

$$K(f(x)) = H(x).$$

Beispiel: Wir betrachten die Differentialgleichung

$$y' = 2xe^{-y}$$

und suchen eine Lösung zu der Anfangsbedingung $f(0) = c$. Die Variablen sind klarerweise getrennt. Für dieses Beispiel können wir

$$h(x) = 2x, \quad k(y) = e^{-y}$$

setzen. Wir erhalten

$$H(x) = 2 \int_0^x t dt = x^2,$$
$$K(y) = \int_c^y \frac{dt}{e^{-t}} = e^y - e^c.$$

Somit

$$e^{f(x)} - e^c = x^2.$$

Umgeformt ergibt sich

$$f(x) = \ln(x^2 + e^c).$$

Als zweites Beispiel betrachten wir die Differentialgleichung

$$y' = y^2.$$

Gesucht ist eine Lösung zu der Anfangsbedingung $y(0) = 1$. Wir haben

$$\frac{dy}{y^2} = dx,$$

und somit liefert die Integration

$$-\frac{1}{y} = x + c.$$

Durch Auflösen nach y erhält man

$$y = -\frac{1}{x + c}.$$

Die Anfangsbedingung $y(0) = 1$ liefert $c = -1$, somit lautet die Lösung

$$y = \frac{1}{1 - x}.$$

Diese Lösung hat einen Pol bei $x = 1$. Dies veranschaulicht die Bedeutung des Satzes von Picard-Lindelöf, der eine Lösung lokal um den Punkt $x_0 = 0$ garantiert: In diesem Fall erhält man eine Lösung auf dem Intervall $]-\infty, 1[$ zu der Anfangsbedingung $y(0) = 1$. Die Lösung läßt sich nicht stetig über den Punkt $x = 1$ hinaus fortsetzen.

Wir betrachten einen weiteren Typ von Differentialgleichungen: Sei $J \subset \mathbb{R}$ ein Intervall, $g : J \rightarrow \mathbb{R}$ eine stetige Funktion und

$$D = \left\{ (x, y) \in \mathbb{R}^* \times \mathbb{R}, \frac{y}{x} \in J \right\}.$$

Wir betrachten die Differentialgleichung

$$y' = g\left(\frac{y}{x}\right).$$

Diese Differentialgleichung kann auf eine Differentialgleichung mit getrennten Variablen zurückgeführt werden. Hierzu betrachten wir die Differentialgleichung

$$z' = \frac{1}{x}(g(z) - z).$$

Sei $I \subset \mathbb{R}^*$ ein Intervall und $(x_0, y_0) \in D$ ein Punkt mit $x_0 \in I$. Eine Funktion $f(x)$ ist genau dann Lösung der ersten Differentialgleichung zum Anfangswert $f(x_0) = y_0$, falls die Funktion

$$\tilde{f}(x) = \frac{f(x)}{x}$$

Lösung der zweiten Differentialgleichung mit der Anfangsbedingung $\tilde{f}(x_0) = y_0/x_0$ ist.

Beweis: Setzen wir $z = y/x$, so ist

$$\frac{dz}{dx} = -\frac{y}{x^2} + \frac{y'}{x} = \frac{1}{x} \left(y' - \frac{y}{x} \right) = \frac{1}{x} \left[g\left(\frac{y}{x}\right) - \frac{y}{x} \right] = \frac{1}{x} (g(z) - z).$$

Beispiel:

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

Die Substitution $z = y/x$ führt auf

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

also

$$\frac{dz}{1+z^2} = \frac{dx}{x}.$$

Mit der Anfangsbedingung $z(x_0) = z_0 = y_0/x_0$ findet man

$$\arctan z - \arctan z_0 = \ln \frac{x}{x_0},$$

also

$$z = \tan \left(\ln \frac{x}{x_0} + \arctan z_0 \right).$$

Die Rücksubstitution $y = xz$ liefert

$$y = x \tan \left(\ln \frac{x}{x_0} + \arctan \frac{y_0}{x_0} \right).$$

6.4 Lineare Differentialgleichungen

Wir betrachten nun einen wichtigen Typ von Differentialgleichungen. Ist die Funktion $G(x, y)$ linear in der Variablen y , so spricht man von einer **linearen Differentialgleichung**. Es sei angemerkt, daß nicht gefordert wird, daß $G(x, y)$ auch linear in x ist. Ebenso spricht man bei einem System von n Differentialgleichungen erster Ordnung von einem linearen Differentialgleichungssystem, falls $\vec{G}(x, \vec{y})$ linear in \vec{y} ist.

Wir betrachten zunächst eine lineare Differentialgleichung erster Ordnung. Da $G(x, y)$ linear in y ist, lässt sich diese Funktion zweier Variablen aufgrund der Linearität mit Hilfe zweier Funktionen $a(x)$ und $b(x)$ wie folgt schreiben:

$$G(x, y) = a(x)y + b(x).$$

Sei I ein Intervall und seien $a, b : I \rightarrow \mathbb{R}$ zwei stetige Funktionen. Die lineare Differentialgleichung erster Ordnung lautet

$$y' = a(x)y + b(x).$$

Man spricht von einer **homogenen** linearen Differentialgleichung erster Ordnung, falls die Gleichung von der Form

$$y' = a(x)y$$

ist. Ist im ersten Fall $b(x) \neq 0$, so spricht man auch von einer inhomogenen Gleichung.

6.4.1 Lösung der homogenen Gleichung

Wir betrachten zunächst die homogene Gleichung. Wir werden später sehen, daß die Lösungen der homogenen Gleichung auch eine Rolle bei der Lösung der inhomogenen Gleichung spielen. Die homogene Differentialgleichung ist eine Gleichung mit getrennten Variablen und man erhält als Lösung zum Anfangswert $y(x_0) = c$

$$y(x) = c \exp\left(\int_{x_0}^x a(t) dt\right).$$

Beispiel: Die Differentialgleichung

$$y' = 2xy$$

hat zum Anfangswert $y(0) = 1$ die Lösung

$$y(x) = \exp\left(2 \int_0^x t dt\right) = \exp(x^2).$$

6.4.2 Lösung der inhomogenen Gleichung

Wir betrachten nun eine inhomogene lineare Differentialgleichung erster Ordnung.

$$y' = a(x)y + b(x).$$

Sei $\varphi : I \rightarrow \mathbb{R}$ eine Lösung der zugehörigen homogenen Differentialgleichung, d.h.

$$\varphi'(x) = a(x)\varphi(x).$$

Sei weiter $\varphi(x_0) \neq 0$. Dann ist auch $\varphi(x) \neq 0$ für alle $x \in I$, da die Exponentialfunktion stets positiv ist. Für die gesuchte Lösung der inhomogenen Gleichung machen wir daher den Ansatz

$$f(x) = \varphi(x) g(x).$$

Nun ist

$$f'(x) = \varphi'(x) g(x) + \varphi(x) g'(x) = a(x) \varphi(x) g(x) + \varphi(x) g'(x).$$

Andererseits ist

$$a(x) f(x) + b(x) = a(x) \varphi(x) g(x) + b(x),$$

und somit

$$\varphi(x) g'(x) = b(x).$$

Da nach Voraussetzung $\varphi(x) \neq 0$ ist, erhält man

$$g'(x) = \frac{b(x)}{\varphi(x)},$$

und die Funktion $g(x)$ ergibt sich zu

$$g(x) = c + \int_{x_0}^x \frac{b(t)}{\varphi(t)} dt.$$

Wir fassen zusammen: Die Lösung der inhomogenen Gleichung $y' = a(x)y + b(x)$ mit der Anfangsbedingung $f(x_0) = c$ ergibt sich zu

$$f(x) = \left(c + \int_{x_0}^x \frac{b(t)}{\varphi(t)} dt \right) \varphi(x),$$

wobei

$$\varphi(x) = \exp\left(\int_{x_0}^x a(t) dt\right)$$

eine Lösung der zugehörigen homogenen Differentialgleichung ist. Lösungen der homogenen Gleichung haben die Form $c' \varphi(x)$. Bei der Lösung der inhomogenen Gleichung spricht man von der "Variation der Konstanten", da die Konstante c' durch die Funktion $g(x)$ ersetzt wird.

Beispiel: Wir betrachten die inhomogene Gleichung

$$y' = 2xy + x^3$$

und suchen die Lösung zur Anfangsbedingung $f(0) = 1$. Die Lösungen der homogenen Gleichung haben die Form

$$\varphi(x) = c \exp(x^2).$$

Wir machen nun den Ansatz

$$f(x) = g(x) e^{x^2}.$$

Für $g(x)$ erhält man die Differentialgleichung

$$g'(x) = x^3 e^{-x^2},$$

und somit

$$g(x) = c + \int_0^x t^3 e^{-t^2} dt.$$

Wir berechnen das Integral mit Hilfe der Substitution $s = t^2$ und partieller Integration

$$\begin{aligned} \int_0^x t^3 e^{-t^2} dt &= \frac{1}{2} \int_0^{x^2} s e^{-s} ds = -\frac{1}{2} \int_0^{x^2} s \left(\frac{d}{ds} e^{-s} \right) ds = -\frac{1}{2} s e^{-s} \Big|_0^{x^2} + \frac{1}{2} \int_0^{x^2} e^{-s} ds \\ &= -\frac{1}{2} x^2 e^{-x^2} - \frac{1}{2} e^{-x^2} + \frac{1}{2} = \frac{1}{2} - \frac{1}{2} (1 + x^2) e^{-x^2}. \end{aligned}$$

Man erhält als Lösung der inhomogenen Gleichung

$$f(x) = \left[c + \frac{1}{2} - \frac{1}{2} (1 + x^2) e^{-x^2} \right] e^{x^2} = -\frac{1}{2} (1 + x^2) + \left(c + \frac{1}{2} \right) e^{x^2}.$$

Wir bestimmen noch die Konstante c aus der Anfangsbedingung $f(0) = 1$:

$$\begin{aligned} -\frac{1}{2} + \left(c + \frac{1}{2} \right) &= 1, \\ c &= 1. \end{aligned}$$

Somit lautet die vollständige Lösung

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

6.4.3 Systeme linearer Differentialgleichungen

Wir betrachten ein System von n linearen Differentialgleichungen erster Ordnung. Es ist zweckmässig, nicht nur reellwertige Funktionen, sondern auch komplexwertige Funktionen zu betrachten. Dies ist eine naheliegende Erweiterung. Ein System von n komplexwertigen Differentialgleichungen ist äquivalent zu einem System von $2n$ reellwertigen Differentialgleichungen, da man eine komplexwertige Funktion immer in zwei reellwertige Funktionen aufspalten kann, die den Real- und Imaginärteil beschreiben. Sei K also \mathbb{R} oder \mathbb{C} . Weiter sei $I \subset \mathbb{R}$ ein Intervall. Beachte das I immer eine Teilmenge von \mathbb{R} ist. Gegeben seien n^2 Funktionen

$$a_{ij} : I \rightarrow K, \quad 1 \leq i, j \leq n, \\ x \rightarrow a_{ij}(x),$$

die wir in einer Matrix $A(x)$ zusammenfassen:

$$A(x) = \begin{pmatrix} a_{11}(x) & a_{12}(x) & \dots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \dots & a_{2n}(x) \\ \dots & \dots & \dots & \dots \\ a_{n1}(x) & a_{n2}(x) & \dots & a_{nn}(x) \end{pmatrix}$$

Weiter seien n Funktionen

$$b_i : I \rightarrow K, \quad 1 \leq i \leq n, \\ x \rightarrow b_i(x),$$

gegeben, die wir in einem Vektor $\vec{b}(x)$ zusammenfassen:

$$\vec{b}(x) = \begin{pmatrix} b_1(x) \\ b_2(x) \\ \dots \\ b_n(x) \end{pmatrix}.$$

Dann beschreibt die Gleichung

$$\vec{y}' = A(x)\vec{y} + \vec{b}(x)$$

ein System von n inhomogenen linearen Differentialgleichungen. Die zugehörige homogene Gleichung lautet

$$\vec{y}' = A(x)\vec{y}.$$

Satz: Sei I ein offenes Intervall und seien alle Funktionen a_{ij} und b_i stetig. Dann hat die inhomogene Differentialgleichung auf ganz I genau eine Lösung zu der Anfangsbedingung $\vec{y}(x_0) = \vec{c}$.

Bemerkung: Der Satz von Picard-Lindelöf garantiert nur die Existenz einer Lösung in der Umgebung eines Punktes x_0 . Die wesentliche Aussage dieses Satzes ist, daß eine Lösung auf ganz I

existiert.

Wir betrachten zunächst wieder die homogene Gleichung. Wir bezeichnen mit L die Menge aller Lösungen $\vec{\varphi} : I \rightarrow K^n$ der homogenen Gleichung

$$\vec{\varphi}'(x) = A(x)\vec{\varphi}(x).$$

Satz: Die Menge L aller Lösungen ist ein Vektorraum über K . Der Beweis ist recht einfach: Zunächst ist die Menge L nicht leer, da die Nullfunktion $\vec{0}$ offensichtlich eine (triviale) Lösung der Differentialgleichung ist. Seien nun $\vec{\varphi}, \vec{\psi} \in L$ und $\lambda \in K$. Nun ist

$$\begin{aligned}(\vec{\varphi} + \vec{\psi})' &= \vec{\varphi}' + \vec{\psi}' = A\vec{\varphi} + A\vec{\psi} = A(\vec{\varphi} + \vec{\psi}), \\ (\lambda\vec{\varphi})' &= \lambda\vec{\varphi}' = \lambda A\vec{\varphi} = A(\lambda\vec{\varphi}).\end{aligned}$$

Satz: Der Vektorraum L hat die Dimension n . Seien $\vec{\varphi}_1(x), \dots, \vec{\varphi}_n(x)$ Lösungen der homogenen Gleichung. Dann sind die folgenden drei Aussagen äquivalent:

- $\vec{\varphi}_1(x), \dots, \vec{\varphi}_n(x)$ sind linear unabhängig und bilden somit eine Basis von L .
- Es existiert ein $x_0 \in I$, so daß die Vektoren $\vec{\varphi}_1(x_0), \dots, \vec{\varphi}_n(x_0) \in K^n$ linear unabhängig sind.
- Für alle $x_0 \in I$ sind die Vektoren $\vec{\varphi}_1(x_0), \dots, \vec{\varphi}_n(x_0)$ linear unabhängig.

Die wesentliche Aussage dieses Satzes besteht darin, daß es für einen Satz von Lösungsfunktionen $\vec{\varphi}_1(x), \dots, \vec{\varphi}_n(x)$ genügt, für einen Punkt x_0 zu überprüfen, ob die Vektoren $\vec{\varphi}_1(x_0), \dots, \vec{\varphi}_n(x_0)$ linear unabhängig sind. Damit sind dann auch automatisch die Lösungsfunktionen linear unabhängig. Der Test, ob eine Menge von Lösungsfunktionen linear unabhängig ist, reduziert sich also auf ein Problem der linearen Algebra.

Wir bezeichnen eine Basis des Vektorraumes L als ein **Lösungs-Fundamentalsystem**. Wir können die Lösungen $\vec{\varphi}_j(x)$ in Spaltenform schreiben:

$$\vec{\varphi}_j(x) = \begin{pmatrix} \varphi_{1j}(x) \\ \dots \\ \varphi_{nj}(x) \end{pmatrix}, \quad 1 \leq j \leq n.$$

Dies definiert eine Matrix

$$\Phi(x) = \begin{pmatrix} \varphi_{11}(x) & \dots & \varphi_{1n}(x) \\ \dots & & \dots \\ \varphi_{n1}(x) & \dots & \varphi_{nn}(x) \end{pmatrix}.$$

Mit Hilfe des vorherigen Satzes bilden die Lösungen $\vec{\varphi}_1(x), \dots, \vec{\varphi}_n(x)$ eine Basis, falls für einen Punkt $x_0 \in I$

$$\det \Phi(x_0) \neq 0$$

gilt. Die allgemeine Lösung des homogenen Systems lässt sich schreiben als

$$\vec{\varphi}(x) = c_1 \vec{\varphi}_1(x) + \dots + c_n \vec{\varphi}_n(x),$$

also

$$\vec{\varphi}(x) = \Phi(x)\vec{c}, \quad \vec{c} = \begin{pmatrix} c_1 \\ \dots \\ c_n \end{pmatrix}.$$

Wir betrachten nun die inhomogene Gleichung

$$\vec{y}' = A(x)\vec{y} + \vec{b}(x),$$

und bezeichnen mit \tilde{L} die Menge aller Lösungen $\vec{f}: I \rightarrow K^n$ der inhomogenen Gleichung. Wie zuvor bezeichne L die Menge der Lösungen der zugehörigen homogenen Gleichung.

Satz: Sei $\vec{f}_0 \in \tilde{L}$ eine Lösung der inhomogenen Gleichung. Dann gilt

$$\tilde{L} = \vec{f}_0 + L.$$

Man erhält also die allgemeine Lösung der inhomogenen Gleichung aus der Summe einer speziellen Lösung der inhomogenen Gleichung plus der allgemeinen Lösung der homogenen Gleichung.

Man bestimmt eine spezielle Lösung der inhomogenen Gleichung mit Hilfe der Variation der Konstanten: Beschreibe $\Phi(x)$ das Fundamentalsystem der homogenen Gleichung. Für die Lösung der inhomogenen Gleichung machen wir den Ansatz

$$\vec{f}(x) = \Phi(x)\vec{g}(x).$$

Dies eingesetzt liefert

$$\vec{f}'(x) = \Phi'(x)\vec{g}(x) + \Phi(x)\vec{g}'(x) = A(x)\Phi(x)\vec{g}(x) + \Phi(x)\vec{g}'(x) \stackrel{!}{=} A(x)\Phi(x)\vec{g}(x) + \vec{b}(x),$$

also

$$\Phi(x)\vec{g}'(x) = \vec{b}(x).$$

Da $\det \Phi(x) \neq 0$ für alle $x \in I$ ist, existiert $\Phi^{-1}(x)$. Somit

$$\vec{g}'(x) = \Phi^{-1}(x)\vec{b}(x)$$

und

$$\vec{g}(x) = \vec{c} + \int_{x_0}^x \Phi^{-1}(t)\vec{b}(t) dt.$$

6.4.4 Lineare Differentialgleichungen höherer Ordnung

Wir hatten schon gezeigt, daß sich eine Differentialgleichung n -ter Ordnung immer auf ein System von n Differentialgleichungen erster Ordnung reduzieren läßt. In diesem Abschnitt soll auf speziellen Fall einer linearen Differentialgleichung n -ter Ordnung näher eingegangen werden.

Es sei $I \subset \mathbb{R}$ ein Intervall und $a_k : I \rightarrow K$ ($0 \leq k < n$) und $b : I \rightarrow K$ stetige Funktionen. Wie zuvor betrachten wir entweder reellwertige oder komplexwertige Funktionen. Man nennt

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = b(x)$$

lineare Differentialgleichung n -ter Ordnung. Ist $b(x) = 0$, so spricht man von einer homogenen Gleichung. Das zugehörige System von n Differentialgleichungen erster Ordnung ist gegeben durch

$$\begin{aligned} y_0' &= y_1, \\ y_1' &= y_2, \\ &\dots \\ y_{n-2}' &= y_{n-1}, \\ y_{n-1}' &= -a_{n-1}(x)y_{n-1} - \dots - a_1(x)y_1 - a_0(x)y_0 + b(x), \end{aligned}$$

also

$$\vec{y} = \begin{pmatrix} y_0 \\ y_1 \\ \dots \\ y_{n-2} \\ y_{n-1} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \dots & & & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0(x) & -a_1(x) & -a_2(x) & \dots & -a_{n-1}(x) \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ b(x) \end{pmatrix}.$$

Wir übersetzen nun die Aussagen über Systeme linearer Differentialgleichungen erster Ordnung auf die lineare Differentialgleichung n -ter Ordnung:

Satz: Sei L die Menge aller Lösungen der homogenen linearen Differentialgleichung n -ter Ordnung und \tilde{L} die Menge aller Lösungen der inhomogenen Gleichung. Dann gilt:

1. L ist ein Vektorraum der Dimension n .
2. Für ein beliebiges $f_0(x) \in \tilde{L}$ gilt

$$\tilde{L} = f_0 + L.$$

3. Ein n -Tupel $\varphi_1(x), \dots, \varphi_n(x) \in L$ von Lösungen der homogenen Gleichung bildet genau dann eine Basis von L , falls für ein $x_0 \in I$ (und damit für alle $x \in I$) gilt:

$$W(x_0) = \det \begin{pmatrix} \varphi_1(x_0) & \dots & \varphi_n(x_0) \\ \varphi_1'(x_0) & \dots & \varphi_n'(x_0) \\ \dots & \dots & \dots \\ \varphi_1^{(n-1)}(x_0) & \dots & \varphi_n^{(n-1)}(x_0) \end{pmatrix} \neq 0.$$

Man bezeichnet $W(x)$ als **Wronski-Determinante**.

Beispiel: Die Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$$

besitzt auf dem Intervall $I =]0, \infty[$ die Lösungen

$$\varphi_1(x) = x, \quad \varphi_2(x) = \sqrt{x},$$

wie man leicht durch Einsetzen überprüft. Die Wronski-Determinante ist

$$W(x) = \begin{vmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{vmatrix} = -\frac{1}{2}\sqrt{x}.$$

Es ist $W(x) \neq 0$ für alle $x \in I$, und somit spannen φ_1 und φ_2 den Lösungsraum auf.

6.5 Lineare Differentialgleichungen mit konstanten Koeffizienten*

6.5.1 Systeme von linearen Differentialgleichungen mit konstanten Koeffizienten*

Wir betrachten das lineare Differentialgleichungssystem

$$\vec{y}' = A\vec{y} + \vec{b}(x),$$

wobei A eine $n \times n$ -Matrix ist. "Konstante Koeffizienten" bedeutet, daß A von x unabhängig ist, $\vec{b}(x)$ darf dagegen von x abhängen. Wir suchen ein Fundamentalsystem für die homogene Gleichung

$$\vec{y}' = A\vec{y},$$

eine Lösung für die inhomogene Gleichung kann dann mittels der Technik der Variation der Konstanten erhalten werden.

Satz: Sei \vec{v}_λ ein Eigenvektor der Matrix A zum Eigenwert λ . Dann ist

$$\vec{\varphi}(x) = \vec{v}_\lambda e^{\lambda x}$$

eine Lösung der homogenen Gleichung.

Beweis: Für einen Eigenvektor gilt per Definition $A\vec{v}_\lambda = \lambda\vec{v}_\lambda$. Wir berechnen $\vec{\varphi}'(x)$:

$$\vec{\varphi}'(x) = \vec{v}_\lambda \frac{d}{dx} e^{\lambda x} = \lambda \vec{v}_\lambda e^{\lambda x} = A\vec{v}_\lambda e^{\lambda x} = A\vec{\varphi}(x).$$

Satz: Besitzt die Matrix A eine Basis $\vec{v}_{\lambda_1}, \dots, \vec{v}_{\lambda_n}$ von Eigenvektoren mit den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$, so bilden die Funktionen

$$\vec{\varphi}_k(x) = \vec{v}_{\lambda_k} e^{\lambda_k x}, \quad 1 \leq k \leq n,$$

ein Fundamentalsystem von Lösungen der homogenen Differentialgleichung.

Beweis: Oben wurde bereits gezeigt, daß jede dieser Funktionen eine Lösung der homogenen Gleichung ist. Es bleibt zu zeigen, daß diese n Lösungen den gesamten Lösungsraum aufspannen. Hierzu genügt es zu zeigen, daß die Vektoren $\vec{\varphi}_k(x=0) = \vec{v}_{\lambda_k}$, $k = 1, \dots, n$, linear unabhängig sind. Nach Voraussetzung bilden die $\vec{v}_{\lambda_1}, \dots, \vec{v}_{\lambda_n}$ eine Basis des K^n , sie sind daher linear unabhängig.

Fazit: Die Bestimmung eines Fundamentalsystems reduziert sich für ein System linearer Differentialgleichungen mit konstanten Koeffizienten auf die Bestimmung von Eigenwerten und Eigenvektoren einer Matrix.

Bemerkung: Nicht immer existiert eine Basis aus Eigenvektoren der Matrix A . Eine solche Basis existiert genau dann, wenn A diagonalisierbar ist. Eine Matrix ist aber immer triagonalisierbar, d.h. sie kann durch einen Basiswechsel auf eine obere Dreiecksgestalt gebracht werden. Hat man ein lineares Differentialgleichungssystem mit einer nicht-diagonalisierbaren Matrix, so bringt man das System zunächst auf eine obere Dreiecksform. Man löst dann die letzte Gleichung und setzt diese Lösung dann in die vorletzte Gleichung ein. Dies ergibt eine inhomogene Differentialgleichung (in einer Variablen) die man mit den bekannten Methoden lösen kann. Man iteriert nun diese Prozedur, bis alle Gleichungen gelöst sind.

6.5.2 Die lineare Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten*

In diesem Abschnitt betrachten wir eine lineare Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten. Dieser Fall kommt in der Anwendung oft vor. Für diesen Typ von Differentialgleichungen gibt es eine gute Lösungstheorie über den komplexen Zahlen. Die Bestimmung eines Fundamentalsystems ist wesentlich einfacher im Vergleich zum vorherigen Abschnitt. Daher betrachten wir in diesem Abschnitt nur die komplexen Zahlen.

Es seien $(n + 1)$ komplexe Zahlen $a_i \in \mathbb{C}$, $(0 \leq i \leq n)$ gegeben. Wir betrachten die homogene Differentialgleichung

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0.$$

Zu dieser Differentialgleichung betrachten wir das Polynom

$$P(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0 \in \mathbb{C}[\lambda].$$

Satz: Ist λ_0 eine Nullstelle von $P(\lambda)$, so ist

$$\varphi(x) = e^{\lambda_0 x}$$

eine Lösung der homogenen Differentialgleichung.

Zum Beweis betrachten wir zunächst die j -te Ableitung von $\varphi(x)$:

$$\frac{d^j}{dx^j} \varphi(x) = \lambda_0^j \varphi(x).$$

Eingesetzt in die linke Seite der Differentialgleichung erhält man

$$\begin{aligned} a_n \varphi^{(n)} + a_{n-1} \varphi^{(n-1)} + \dots + a_1 \varphi' + a_0 \varphi &= a_n \lambda_0^n \varphi + a_{n-1} \lambda_0^{n-1} \varphi + \dots + a_1 \lambda_0 \varphi + a_0 \varphi \\ &= (a_n \lambda_0^n + a_{n-1} \lambda_0^{n-1} + \dots + a_1 \lambda_0 + a_0) \varphi = P(\lambda_0) \varphi = 0. \end{aligned}$$

Die Suche nach den Lösungen der homogenen Differentialgleichung reduziert sich daher auf die Bestimmung der Nullstellen des Polynoms $P(\lambda)$.

Satz: Falls die Nullstellen $\lambda_1, \dots, \lambda_n$ des Polynoms $P(\lambda)$ paarweise verschieden sind, so bilden die Funktionen

$$\varphi_i(x) = e^{\lambda_i x}, \quad 1 \leq i \leq n,$$

ein Fundamentalsystem für die Lösungen der homogenen Differentialgleichung.

Beweis: Wir müssen überprüfen, daß die Lösungen linear unabhängig sind. Hierzu betrachten wir die Wronski-Determinante am Punkte $x_0 = 0$:

$$W(0) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_n^2 \\ \dots & \dots & \dots & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{vmatrix}$$

Eine Determinante dieser Form bezeichnet man als Vandermondesche Determinante. Für die Vandermondesche Determinante gilt

$$W(0) = \prod_{i>j} (\lambda_i - \lambda_j) \neq 0,$$

da nach Voraussetzung alle Nullstellen paarweise verschieden sind.

Wir müssen noch den Fall betrachten, daß einige Nullstellen mehrfach auftreten. Das Polynom $P(\lambda)$ habe also die paarweise verschiedenen Nullstellen λ_i mit Vielfachheiten ν_i , wobei $1 \leq i \leq r$ und $\nu_1 + \dots + \nu_r = n$ gilt.

Satz: Die homogene lineare Differentialgleichung

$$P\left(\frac{d}{dx}\right)y = 0$$

besitzt ein Lösungs-Fundamentalsystem aus folgenden Funktionen:

$$\varphi_{ij}(x) = x^j e^{\lambda_i x}, \quad 1 \leq i \leq r, \quad 0 \leq j \leq \nu_i - 1.$$

Beispiel: Gesucht ist ein Lösungs-Fundamentalsystem der Differentialgleichung

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

Das zugehörige Polynom lautet

$$P(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2.$$

Also ist ein Fundamentalsystem gegeben durch

$$\{e^x, xe^x\}$$

und die allgemeine Lösung lautet

$$y = (c_1 x + c_0) e^x.$$

Den obigen Satz sieht man leicht ein, wenn man

$$\left(\frac{d}{dx} - \lambda\right)^n y = 0$$

betrachtet. Für $j \in \mathbb{N}_0$ ist

$$\left(\frac{d}{dx} - \lambda\right) x^j e^{\lambda x} = j x^{j-1} e^{\lambda x}$$

und somit

$$\left(\frac{d}{dx} - \lambda\right)^n x^j e^{\lambda x} = j(j-1)\dots(j-n+1) x^{j-n} e^{\lambda x}$$

Für $j < n$ ist einer der Vorfaktoren Null, und somit ist $x^j e^{\lambda x}$ Lösung der homogenen Differentialgleichung für alle $0 \leq j < n$.

Die inhomogene Gleichung

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = b(x).$$

löst man wie üblich, indem man zuerst die allgemeine Lösung der homogenen Gleichung bestimmt. Man führt dann die Differentialgleichung n -ter Ordnung auf ein System von n Gleichungen erster Ordnung zurückführt und bestimmt hierfür eine spezielle Lösung. Die allgemeine Lösung ist dann gegeben als die Summe einer speziellen Lösung der inhomogenen Gleichung plus die allgemeine Lösung der homogenen Gleichung.

Für den speziellen Fall, daß der inhomogene Term von der Form

$$b(x) = ce^{\kappa x}$$

ist, wobei κ **keine** Nullstelle des zur homogenen Gleichung assoziierten Polynoms $P(\lambda)$ ist, d.h.

$$P(\kappa) \neq 0,$$

gibt es ein vereinfachtes Verfahren: Die Funktion

$$f(x) = \frac{c}{P(\kappa)} e^{\kappa x}$$

ist **eine** spezielle Lösung der inhomogenen Gleichung.

6.5.3 Der harmonische Oszillator*

Gesucht ist eine reellwertige Funktion $x(t)$ die folgende Differentialgleichung erfüllt:

$$\frac{d^2}{dt^2}x + 2\mu \frac{d}{dt}x + \omega_0^2 x = A \cos(\omega_{ext}t)$$

In der Physik ist es üblich, Ableitungen nach der Zeit mit einem Punkt zu kennzeichnen:

$$\dot{x}(t) = \frac{d}{dt}x(t).$$

Somit schreibt man die obige Gleichung auch oft als

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = A \cos(\omega_{ext}t)$$

ω_0 ist die Eigenfrequenz des harmonischen Oszillators, der Term proportional zu μ beschreibt die Dämpfung, die rechte Seite eine äußere treibende Kraft. Wir können $\mu \geq 0$ und $\omega_0 > 0$ annehmen.

Es ist einfacher, zunächst eine komplexwertige Funktion $z(t)$ zu suchen, die die Differentialgleichung

$$\ddot{z} + 2\mu\dot{z} + \omega_0^2 z = Ae^{i\omega_{ext}t}$$

erfüllt, und dann den Realteil zu nehmen:

$$x(t) = \operatorname{Re} z(t).$$

Wir betrachten zunächst die homogene Gleichung:

$$\ddot{z} + 2\mu\dot{z} + \omega_0^2 z = 0.$$

Das zugehörige Polynom lautet

$$P(\lambda) = \lambda^2 + 2\mu\lambda + \omega_0^2.$$

Wir bestimmen die Nullstellen von P . Hierzu betrachten wir die Fälle $\mu^2 < \omega_0^2$, $\mu^2 = \omega_0^2$ und $\mu^2 > \omega_0^2$ getrennt.

1. Fall: $\mu^2 < \omega_0^2$. Dieser Fall beschreibt eine kleine Dämpfung.

$$\lambda = -\mu \pm i\sqrt{\omega_0^2 - \mu^2} = -\mu \pm i\omega,$$

wobei wir

$$\omega = \sqrt{\omega_0^2 - \mu^2}$$

gesetzt haben. Wir erhalten somit das Lösungsfundamentalsystem

$$\varphi_1(t) = e^{-\mu t} e^{i\omega t}, \quad \varphi_2(t) = e^{-\mu t} e^{-i\omega t},$$

und die allgemeine Lösung der homogenen Gleichung lautet in diesem Fall

$$\varphi(t) = c_1 e^{-\mu t} e^{i\omega t} + c_2 e^{-\mu t} e^{-i\omega t}.$$

2. Fall: $\mu^2 = \omega_0^2$. In diesem Fall erhalten wir eine doppelte Nullstelle

$$\lambda = -\mu = -\omega.$$

Das Lösungsfundamentalsystem lautet

$$\varphi_1(t) = e^{-\mu t}, \quad \varphi_2(t) = t e^{-\mu t},$$

und die allgemeine Lösung der homogenen Gleichung lautet in diesem Fall

$$\varphi(t) = c_1 e^{-\mu t} + c_2 t e^{-\mu t}.$$

Den Fall $\mu^2 = \omega_0^2$ bezeichnet man als aperiodischen Grenzfall.

3. Fall: $\mu^2 > \omega_0^2$. In diesem Fall erhalten wir reelle Nullstellen

$$\lambda = -\mu \pm \sqrt{\mu^2 - \omega_0^2}.$$

Beide Nullstellen sind negativ. Wir setzen

$$\mu_1 = \mu + \sqrt{\mu^2 - \omega_0^2}, \quad \mu_2 = \mu - \sqrt{\mu^2 - \omega_0^2}.$$

Das Lösungsfundamentalsystem lautet

$$\varphi_1(t) = e^{-\mu_1 t}, \quad \varphi_2(t) = e^{-\mu_2 t},$$

und die allgemeine Lösung der homogenen Gleichung lautet in diesem Fall

$$\varphi(t) = c_1 e^{-\mu_1 t} + c_2 e^{-\mu_2 t}.$$

Den Fall $\mu^2 > \omega_0^2$ bezeichnet man als Kriechfall.

Wir betrachten nun den Fall einer schwachen Dämpfung mit einer äußeren treibenden Kraft, d.h. für $\mu^2 < \omega_0^2$ die inhomogene Differentialgleichung

$$\ddot{z} + 2\mu\dot{z} + \omega_0^2 z = A e^{i\omega_{ext} t}.$$

Wir können annehmen, daß ω_{ext} reell und positiv ist. Wie bereits oben gezeigt, ist das Fundamentalsystem der homogenen Gleichung gegeben durch

$$\varphi_1(t) = e^{-\mu t} e^{i\omega t}, \quad \varphi_2(t) = e^{-\mu t} e^{-i\omega t}, \quad \omega = \sqrt{\omega_0^2 - \mu^2}.$$

Für $\mu > 0$ haben wir immer

$$i\omega_{ext} \neq -\mu \pm i\omega,$$

somit ist $(i\omega_{ext})$ keine Nullstelle von $P(\lambda)$ und eine spezielle Lösung der inhomogenen Differentialgleichung ist

$$f_0(t) = \frac{A}{P(i\omega_{ext})} e^{i\omega_{ext} t} = \frac{A}{\omega_0^2 - \omega_{ext}^2 + 2i\mu\omega_{ext}} e^{i\omega_{ext} t} = A \frac{\omega_0^2 - \omega_{ext}^2 - 2i\mu\omega_{ext}}{(\omega_0^2 - \omega_{ext}^2)^2 + 4\mu^2\omega_{ext}^2} e^{i\omega_{ext} t}$$

Die allgemeine Lösung der inhomogenen Gleichung ist somit

$$z(t) = A \frac{\omega_0^2 - \omega_{ext}^2 - 2i\mu\omega_{ext}}{(\omega_0^2 - \omega_{ext}^2)^2 + 4\mu^2\omega_{ext}^2} e^{i\omega_{ext} t} + c_1 e^{-\mu t} e^{i\omega t} + c_2 e^{-\mu t} e^{-i\omega t}.$$

Es bleibt noch der Spezialfall $\mu = 0$ und $\omega = \omega_0 = \omega_{ext}$ zu diskutieren. Man bezeichnet diesen Fall als Resonanzfall. Dies entspricht der Differentialgleichung

$$\ddot{z} + \omega^2 z = A e^{i\omega t}.$$

Das Fundamentalsystem der homogenen Gleichung ist

$$\varphi_1(t) = e^{i\omega t}, \quad \varphi_2(t) = e^{-i\omega t}.$$

Wir schreiben die Differentialgleichung zweiter Ordnung um auf ein System zweier Gleichungen erster Ordnung:

$$\begin{pmatrix} \dot{z}_0 \\ \dot{z}_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} + \begin{pmatrix} 0 \\ A e^{i\omega t} \end{pmatrix}.$$

Das Lösungsfundamentalsystem für dieses System lautet

$$\Phi(t) = \begin{pmatrix} e^{i\omega t} & e^{-i\omega t} \\ i\omega e^{i\omega t} & -i\omega e^{-i\omega t} \end{pmatrix}$$

Das Inverse dieser Matrix lautet

$$\Phi^{-1}(t) = \frac{i}{2\omega} \begin{pmatrix} -i\omega e^{-i\omega t} & -e^{-i\omega t} \\ -i\omega e^{i\omega t} & e^{i\omega t} \end{pmatrix}$$

Wir berechnen nun

$$\begin{aligned} \int_0^t \Phi^{-1}(\tilde{t}) \vec{b}(\tilde{t}) d\tilde{t} &= \frac{i}{2\omega} \int_0^t d\tilde{t} \begin{pmatrix} -i\omega e^{-i\omega \tilde{t}} & -e^{-i\omega \tilde{t}} \\ -i\omega e^{i\omega \tilde{t}} & e^{i\omega \tilde{t}} \end{pmatrix} \begin{pmatrix} 0 \\ Ae^{i\omega \tilde{t}} \end{pmatrix} \\ &= \frac{iA}{2\omega} \int_0^t d\tilde{t} \begin{pmatrix} -1 \\ e^{2i\omega \tilde{t}} \end{pmatrix} = A \begin{pmatrix} -\frac{i}{2\omega} t \\ -\frac{1}{4\omega^2} (1 - e^{2i\omega t}) \end{pmatrix} \end{aligned}$$

Damit ergibt sich eine spezielle Lösung zu

$$\Phi(t) \int_0^t \Phi^{-1}(\tilde{t}) \vec{b}(\tilde{t}) d\tilde{t} = A \begin{pmatrix} -\frac{i}{2\omega} t e^{i\omega t} - \frac{1}{4\omega^2} e^{-i\omega t} + \frac{1}{4\omega^2} e^{i\omega t} \\ \frac{1}{2} t e^{i\omega t} + \frac{i}{4\omega} e^{-i\omega t} - \frac{i}{4\omega} e^{i\omega t} \end{pmatrix}$$

Wir benötigen die erste Komponente dieses Vektors, außerdem können wir Terme, die nur Linearkombinationen der homogenen Lösungen sind, zur Bestimmung der speziellen Lösung weglassen. Wir erhalten also

$$f(t) = -\frac{i}{2\omega} A t e^{i\omega t}$$

als eine spezielle Lösung der inhomogenen Differentialgleichung. Die allgemeine Lösung im Resonanzfall ist somit

$$z(t) = -\frac{i}{2\omega} A t e^{i\omega t} + c_1 e^{i\omega t} + c_2 e^{-i\omega t}.$$

Die Amplitude des ersten Terms wächst linear mit t an.

6.6 Exakte Differentialgleichungen und integrierende Faktoren*

Eine gewöhnliche Differentialgleichung der Form

$$a(x, y) + b(x, y)y' = 0$$

nennt man **exakt**, falls es eine stetig differenzierbare Funktion $F(x, y)$ gibt, so daß

$$a(x, y) = \frac{\partial F(x, y)}{\partial x},$$

$$b(x, y) = \frac{\partial F(x, y)}{\partial y}$$

ist. Ist die Funktion $F(x, y)$ bekannt, so erhält man eine Lösung der Differentialgleichung durch Auflösen der impliziten Gleichung

$$F(x, y) = c$$

nach y , wobei c eine Integrationskonstante ist. Es läßt sich zeigen, daß eine Funktion $F(x, y)$ genau dann existiert, falls

$$\frac{\partial}{\partial y}a(x, y) = \frac{\partial}{\partial x}b(x, y)$$

ist.

Beispiel: Wir betrachten die Differentialgleichung

$$3x^2 \tan y + x^3 (1 + \tan^2 y) y' = 0.$$

Diese Differentialgleichung ist exakt, die Funktion $F(x, y)$ lautet

$$F(x, y) = x^3 \tan y,$$

wie man leicht durch Nachrechnen überprüft:

$$\begin{aligned} \frac{\partial}{\partial x}F(x, y) &= 3x^2 \tan y, \\ \frac{\partial}{\partial y}F(x, y) &= x^3 (1 + \tan^2 y). \end{aligned}$$

Somit müssen wir die Gleichung

$$x^3 \tan y = c$$

lösen. Dies ergibt

$$y = \arctan\left(\frac{c}{x^3}\right).$$

Gilt für eine Differentialgleichung

$$a(x, y) + b(x, y)y' = 0$$

$$\frac{\partial}{\partial y}a(x, y) \neq \frac{\partial}{\partial x}b(x, y),$$

so ist sie nicht exakt. Findet man aber eine Funktion $c(x, y) \neq 0$, so daß

$$\frac{\partial}{\partial y} (a(x, y)c(x, y)) = \frac{\partial}{\partial x} (b(x, y)c(x, y)),$$

gilt, so ist die Differentialgleichung

$$a(x, y)c(x, y) + b(x, y)c(x, y)y' = 0$$

exakt. Da nach Voraussetzung $c(x, y) \neq 0$ für alle x und y ist, so ist eine Lösung dieser Differentialgleichung auch eine Lösung der ursprünglichen Gleichung und umgekehrt. Man bezeichnet die Funktion $c(x, y)$ als integrierenden Faktor.

Beispiel: Wir betrachten für $x > 0$ die Differentialgleichung

$$3 \tan y + x(1 + \tan^2 y)y' = 0.$$

Diese Differentialgleichung ist nicht exakt, da

$$\frac{\partial}{\partial y} (3 \tan y) = 3(1 + \tan^2 y) \neq (1 + \tan^2 y) = \frac{\partial}{\partial x} x(1 + \tan^2 y).$$

Allerdings ist die Funktion

$$c(x, y) = x^2$$

ein integrierender Faktor (und ungleich Null für $x > 0$). Multiplikation mit $c(x, y) = x^2$ ergibt die exakte Differentialgleichung

$$3x^2 \tan y + x^3(1 + \tan^2 y)y' = 0,$$

deren Lösung wir schon oben betrachtet haben.

7 Die Eulersche Gamma-Funktion*

In diesem Abschnitt behandeln wir die Eulersche Gamma-Funktion.

Definition: Sei $x > 0$. Man definiert die Eulersche Gamma-Funktion durch das uneigentliche Integral

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt.$$

Diese Funktion hat die Eigenschaft, daß

$$\begin{aligned}\Gamma(n+1) &= n! \quad \text{für alle } n \in \mathbb{N}, \\ \Gamma(x+1) &= x\Gamma(x) \quad \text{für alle } x > 0.\end{aligned}$$

Die zweite Behauptung beweisen wir mit partieller Integration:

$$\begin{aligned}\Gamma(x+1) &= \lim_{a \rightarrow 0, b \rightarrow \infty} \int_a^b t^x e^{-t} dt = \lim_{a \rightarrow 0, b \rightarrow \infty} -t^x e^{-t} \Big|_a^b + \lim_{a \rightarrow 0, b \rightarrow \infty} x \int_a^b t^{x-1} e^{-t} dt = x \int_0^{\infty} t^{x-1} e^{-t} dt \\ &= x\Gamma(x).\end{aligned}$$

Um $\Gamma(n+1) = n!$ zu zeigen, genügt es nun, da die Funktionalgleichung bereits bewiesen wurde, $\Gamma(1) = 1$ zu zeigen.

$$\Gamma(1) = \lim_{a \rightarrow 0, b \rightarrow \infty} \int_a^b e^{-t} dt = - \lim_{a \rightarrow 0, b \rightarrow \infty} e^{-t} \Big|_a^b = 1.$$

Die Eulersche Gamma-Funktion läßt sich auch axiomatisch definieren. Hierzu betrachten wir zunächst den Begriff "logarithmisch konvex". Sei $I \subset \mathbb{R}$ ein Intervall und bezeichne \mathbb{R}_+^* alle Zahlen $x \in \mathbb{R}$ mit $x > 0$. Wir betrachten eine Funktion $F : I \rightarrow \mathbb{R}_+^*$. Die Funktion F nennt man **logarithmisch konvex**, falls die Funktion $\ln F : I \rightarrow \mathbb{R}$ konvex ist. Übersetzt bedeutet dies, daß für alle $x, y \in I$ und $0 < \lambda < 1$ gilt:

$$F(\lambda x + (1-\lambda)y) \leq F(x)^\lambda F(y)^{1-\lambda}.$$

Satz von H. Bohr: Sei $F : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ eine Funktion mit den folgenden Eigenschaften:

1. $F(1) = 1$,
2. $F(x+1) = xF(x)$ für alle $x \in \mathbb{R}_+^*$,
3. F ist logarithmisch konvex.

Dann gilt $F(x) = \Gamma(x)$ für alle $x \in \mathbb{R}_+^*$.

Es läßt sich zeigen, daß sich die Eulersche Gamma-Funktion auf \mathbb{C} fortsetzen läßt. Als Funktion auf \mathbb{C} hat $\Gamma(z)$ Pole entlang der negativen reellen Achse bei $z = 0, -1, -2, -3, \dots$

Oft benötigt man die Taylorentwicklung der Gamma-Funktion um einen ganzzahligen Wert. Diese Entwicklung erhält man aus der Funktionalgleichung $\Gamma(x+1) = x\Gamma(x)$ indem man die Entwicklung um $x = n, n \in \mathbb{N}$ auf eine Einwicklung um $n = 1$ zurückführt. Die Entwicklung um $x = 1$ gewinnt man aus der Formel

$$\Gamma(1+x) = \exp\left(-\gamma_E x + \sum_{n=2}^{\infty} \frac{(-1)^n}{n} \zeta_n x^n\right).$$

Hierin ist γ_E die Eulersche Konstante definiert durch

$$\gamma_E = \lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \frac{1}{j} - \ln n \right) = 0.5772156649\dots$$

und ζ_n der Wert der Riemannschen Zetafunktion an der Stelle $s = n$.

$$\zeta_n = \sum_{j=1}^{\infty} \frac{1}{j^n}.$$

Bemerkung: Die Riemannsche Zetafunktion ist für $s > 1$ definiert durch

$$\zeta(s) = \sum_{j=1}^{\infty} \frac{1}{j^s}.$$

Diese Funktion läßt sich auf \mathbb{C} fortsetzen. Die Lage der Nullstellen dieser Funktion ist ein ungelöstes Problem der Mathematik.

Wir betrachten noch die Eulersche Beta-Funktion. Sie ist definiert für $x > 0$ und $y > 0$ durch

$$B(x, y) = \int_0^{\infty} \frac{t^{x-1}}{(1+t)^{x+y}} dt.$$

Durch Variablensubstitution läßt sich dieses Integral umschreiben auf

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt.$$

Die Eulersche Beta-Funktion läßt sich durch Eulersche Gamma-Funktionen ausdrücken:

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Wir betrachten noch eine Anwendung der Eulerschen Gamma-Funktion. Wir betrachten zwei Folgen (a_n) und (b_n) deren Folgenglieder alle ungleich Null sind. Wir bezeichnen diese Folgen als **asymptotisch gleich**, falls

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$$

gilt. In diesem Fall schreiben wir

$$a_n \sim b_n.$$

Bemerkung: Es wird nicht vorausgesetzt, daß die Folgen konvergieren.

Satz (Stirling): Die Fakultät hat das asymptotische Verhalten

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

8 Asymptotisches Verhalten*

Im letzten Abschnitt hatten wir bereits asymptotisch gleiche Folgen betrachtet. In diesem Abschnitt wollen wir dies noch etwas vertiefen. Wir betrachten das Verhalten zweier Funktionen $f(x)$ und $g(x)$ in der Umgebung eines Punktes $x = x_0$. Wir sind im wesentlichen an dem Fall interessiert, in dem beide Funktionen an diesem Punkte divergieren. Wir betrachten daher den Grenzwert $x \rightarrow x_0$. Der Fall $x \rightarrow \pm\infty$ ist ein Spezialfall mit $x_0 = \pm\infty$. Wir sagen, die beiden Funktionen haben an diesem Punkte das gleiche asymptotische Verhalten, falls

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 1$$

gilt. Wir schreiben in diesem Fall

$$f(x) \sim g(x) \quad \text{für } x \rightarrow x_0.$$

Bemerkung: Sind beide Funktionen an der Stelle $x = x_0$ endlich und haben sie den gleichen Wert $f(x_0) = g(x_0)$, so haben sie natürlich auch trivialerweise das gleiche asymptotische Verhalten an dieser Stelle.

Wir führen noch die Notation mit einem großen O und einem kleinen o ein. Falls

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0$$

gilt, so schreibt man

$$f(x) = o(g(x)) \quad \text{für } x \rightarrow x_0.$$

Gibt es eine Konstante K , so daß

$$\lim_{x \rightarrow x_0} \frac{|f(x)|}{|g(x)|} \leq K$$

gilt, so schreibt man

$$f(x) = O(g(x)) \quad \text{für } x \rightarrow x_0.$$

Bemerkung: Die Relationen \sim und o implizieren die Relation O .

Die formale Reihe

$$\sum_{j=0}^{\infty} f_j(x)$$

nennt man eine **asymptotische Entwicklung** der Funktion $f(x)$ um den Punkt x_0 falls für jedes n für $x \rightarrow x_0$

$$f(x) - \sum_{j=0}^n f_j(x) = o(f_n(x))$$

gilt. Dies bedeutet

$$\lim_{x \rightarrow x_0} \frac{f(x) - \sum_{j=0}^n f_j(x)}{f_n(x)} = 0, \quad \forall n.$$

9 Fehlerrechnung*

Wir betrachten den Fall, daß eine Person A eine bestimmte Größe experimentell mißt. Sie führt diese Messung öfters durch. Aufgrund der experimentellen Meßgenauigkeit ergeben sich leicht unterschiedliche Werte. Zum Beispiel:

Meßreihe 1:

Messung	1	2	3	4	5	6	7	8	9	10	11	12	13
Ergebnis	2.6	2.3	2.5	2.3	2.6	2.4	2.2	2.3	2.4	2.5	2.6	2.8	2.7

Wir definieren den Mittelwert einer Meßreihe mit n Meßpunkten als

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j.$$

Für die obige Meßreihe ergibt sich

$$\bar{x} = 2.48$$

Wir betrachten nun weiter den Fall, daß eine Person B die gleiche Größe experimentell bestimmt. Person B verwendet allerdings eine schlechtere Meßapparatur und führt weniger Messungen durch. Person B erhält die folgenden Meßwerte:

Meßreihe 2:

Messung	1	2	3	4
Ergebnis	0.3	5.2	3.1	1.4

Der Mittelwert ergibt sich zu

$$\bar{x} = 2.48$$

Im zweiten Fall streuen die einzelnen Messungen wesentlich stärker als im ersten Fall. Es ist daher offensichtlich, daß das Ergebnis von Person A vertrauenswürdiger als das Ergebnis von Person B ist. Wir wollen nun diese Aussage quantitativ machen und suchen ein Maß für die Streuung der Meßpunkte.

Wir beginnen mit einigen Definitionen:

Ω : Ergebnismenge eines Zufallsexperiments,
Zufallsfunktion : Funktion $X : \Omega \rightarrow \mathbb{R}$,

Wahrscheinlichkeitsfunktion einer Zufallsgröße X :

$$W : x \rightarrow P(\omega | X(\omega) = x).$$

Erwartungswert einer Zufallsgröße: Nimmt die Zufallsgröße X die Werte x_1, x_2, \dots, x_n an, so bezeichnet man mit

$$\mu(X) = \sum_{j=1}^n x_j W(x_j)$$

den Erwartungswert von X .

Satz: Entsprechen die einzelnen Messungen einzelnen unabhängigen Realisierungen eines Zufallsexperiments, so ist der Mittelwert \bar{x} eine Schätzung für $\mu(X)$.

Die Varianz einer Zufallsgröße ist definiert durch

$$\text{Var}(X) = \sum_{j=1}^n (x_j - \mu)^2 W(x_j).$$

Die Standardabweichung einer Zufallsgröße ist definiert durch

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

Kennen wir den Erwartungswert μ einer Zufallsgröße und machen n Messungen x_j , so ist

$$\frac{1}{n} \sum_{j=1}^n (x_j - \mu)^2$$

eine Schätzfunktion für die Varianz.

Im allgemeinen ist μ aber nicht bekannt und man verwendet \bar{x} als Schätzung für μ . In diesem Fall ist

$$S^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$$

eine Schätzfunktion für die Varianz der Zufallsgröße X .

Sätze über die Varianz: Sei $c \in \mathbb{R}$ und seien X_1, X_2, \dots, X_n unabhängige Zufallsgrößen. Dann gilt:

$$\begin{aligned} \text{Var}(cX) &= c^2 \text{Var}(X), \\ \text{Var}(X_1 + X_2 + \dots + X_n) &= \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n). \end{aligned}$$

Insbesondere gilt

$$\text{Var}\left(\frac{1}{n} \underbrace{(X + X + \dots + X)}_{n \text{ mal}}\right) = \frac{1}{n^2} (\text{Var}(X) + \text{Var}(X) + \dots + \text{Var}(X)) = \frac{1}{n} \text{Var}(X).$$

Varianz des Mittelwertes: Es interessiert in erster Linie nicht die Varianz der einzelnen Messungen $\text{Var}(X)$, sondern die Varianz des Mittelwertes $\text{Var}(\bar{X})$. Bei n Messungen gilt:

$$\text{Var}(\bar{X}) = \frac{1}{n} \text{Var}(X).$$

Somit erhält man als Schätzung für die Varianz des Mittelwertes

$$S_{\bar{X}}^2 = \frac{1}{n} S^2 = \frac{1}{n(n-1)} \sum_{j=1}^n (x_j - \bar{x})^2.$$

Für die Standardabweichung erhält man

$$\sigma_{\bar{X}} = \sqrt{\frac{1}{n(n-1)} \sum_{j=1}^n (x_j - \bar{x})^2}.$$

Somit findet man für die beiden oben aufgeführten Meßreihen:

$$\text{Meßreihe 1 : } \sigma_{\bar{X}} = 0.05,$$

$$\text{Meßreihe 2 : } \sigma_{\bar{X}} = 1.07.$$

Es ist üblich mit dem Mittelwert auch immer die Standardabweichung anzugeben, also

$$\text{Meßreihe 1 : } x = 2.48 \pm 0.05,$$

$$\text{Meßreihe 2 : } x = 2.48 \pm 1.07.$$

Zur Interpretation der Standardabweichung betrachten wir zunächst kontinuierliche Zufallsgrößen. Die Wahrscheinlichkeitsdichtefunktion $p(x)$ für eine kontinuierliche Zufallsgröße beschreibt

$$P(a < X \leq b) = \int_a^b p(x) dx.$$

Definition: Man nennt eine kontinuierliche Zufallsgröße **normalverteilt**, falls sie die Wahrscheinlichkeitsdichtefunktion

$$p(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right)$$

besitzt. Der Erwartungswert dieser normalverteilten Zufallsgröße ist μ , die Standardabweichung ist σ . Für eine normalverteilte Zufallsgröße gilt:

$$P(\mu - \sigma < X \leq \mu + \sigma) \approx 68.27\%,$$

$$P(\mu - 2\sigma < X \leq \mu + 2\sigma) \approx 95.45\%,$$

$$P(\mu - 3\sigma < X \leq \mu + 3\sigma) \approx 99.73\%.$$

9.1 Fehlerfortpflanzung*

Problemstellung: Gesucht wird eine Größe $f = f(x, y)$ die von zwei weiteren Größen x und y abhängt. Die Funktion f wird als bekannt vorausgesetzt, die Größen x und y werden durch eine Messung mit Fehlern $x \pm \Delta x$ und $y \pm \Delta y$ bestimmt. Gesucht ist nun der Fehler für die Größe f .

Für die Größe f beginnen wir mit der Taylorentwicklung:

$$f(x + \Delta x, y + \Delta y) = f(x, y) + \frac{\partial f(x, y)}{\partial x} \Delta x + \frac{\partial f(x, y)}{\partial y} \Delta y + \dots$$

Wir nehmen an, daß wir n Messungen für die Größen x und y haben, die einzelnen Meßwerte seien mit x_j und y_j bezeichnet. Somit haben wir auch n Ergebnisse für f . Für die Abweichung eines Einzelergebnisses vom Mittelwert gilt für kleine Abweichungen

$$f_j - \bar{f} = \frac{\partial f}{\partial x} \cdot (x_j - \bar{x}) + \frac{\partial f}{\partial y} \cdot (y_j - \bar{y}) + \dots$$

Somit gilt für die Varianz:

$$\begin{aligned} \sigma_f^2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n (f_j - \bar{f})^2 \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \left[(x_j - \bar{x})^2 \left(\frac{\partial f}{\partial x} \right)^2 + (y_j - \bar{y})^2 \left(\frac{\partial f}{\partial y} \right)^2 + 2(x_j - \bar{x})(y_j - \bar{y}) \left(\frac{\partial f}{\partial x} \right) \left(\frac{\partial f}{\partial y} \right) \right] \end{aligned}$$

Wir definieren die **Kovarianz** als

$$\text{Cov}(x, y) = \sigma_{xy} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$$

Somit haben wir

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y} \right)^2 \sigma_y^2 + 2 \left(\frac{\partial f}{\partial x} \right) \left(\frac{\partial f}{\partial y} \right) \sigma_{xy}.$$

Falls x und y unkorreliert sind, gilt $\sigma_{xy} = 0$ und somit

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y} \right)^2 \sigma_y^2,$$

bzw.

$$\sigma_f = \sqrt{\left(\frac{\partial f}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y} \right)^2 \sigma_y^2}.$$

Wir betrachten nun einige Beispiele für diese Formel:

1. $f = x + y$. In diesem Fall haben wir

$$\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2},$$

man sagt, die (absoluten) Fehler addieren sich quadratisch. Für $x = 15 \pm 3$ und $y = 17 \pm 4$ ergibt sich also $f = 32 \pm 5$.

2. $f = x \cdot y$. In diesem Fall findet man

$$\sigma_f = \sqrt{y^2 \sigma_x^2 + x^2 \sigma_y^2},$$

oder anders geschrieben

$$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}.$$

Bei einem Produkt addieren sich die relativen Fehler quadratisch.

3. $f = x - y$. Hier findet man wie bei einer Summe

$$\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2}.$$

4. $f = \frac{x}{y}$. In diesem Fall findet man

$$\sigma_f = \sqrt{\frac{1}{y^2} \sigma_x^2 + \frac{x^2}{y^4} \sigma_y^2}.$$

Schreibt man dies mit Hilfe der relativen Fehler erhält man wie beim Produkt

$$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}.$$

5. Zum Abschluss betrachten wir noch $f = x^a y^b$. Man erhält

$$\sigma_f = \sqrt{(ax^{a-1}y^b)^2 \sigma_x^2 + (bx^a y^{b-1})^2 \sigma_y^2}$$

Auch hier empfiehlt es sich wieder, die Formel in relativen Fehler zu schreiben:

$$\frac{\sigma_f}{f} = \sqrt{a^2 \left(\frac{\sigma_x}{x}\right)^2 + b^2 \left(\frac{\sigma_y}{y}\right)^2}.$$

Wir hatten zuvor den Fall betrachtet, daß eine Größe durch zwei Meßreihen experimentell bestimmt wird:

$$\text{Meßreihe 1 : } x = 2.48 \pm 0.05,$$

$$\text{Meßreihe 2 : } x = 2.48 \pm 1.07.$$

Es stellt sich nun die Frage, wie man diese Ergebnisse miteinander kombiniert. Etwas allgemeiner seien für eine Größe x n Messungen x_j mit Fehlern σ_j gegeben. Dann setzt man

$$x = \frac{\frac{1}{\sigma_1^2}x_1 + \frac{1}{\sigma_2^2}x_2 + \dots + \frac{1}{\sigma_n^2}x_n}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \dots + \frac{1}{\sigma_n^2}}, \quad \sigma = \frac{1}{\sqrt{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \dots + \frac{1}{\sigma_n^2}}}.$$

Für das Beispiel hat man

$$x_1 = 2.48, \quad \sigma_1 = 0.05,$$

$$x_2 = 2.48, \quad \sigma_2 = 1.07.$$

Man findet somit

$$x = 2.48, \quad \sigma = 0.04995.$$

Die zweite Meßreihe liefert keinen wesentlichen Beitrag zur Verbesserung des Fehlers.

9.2 Fitten von Parametern*

Wir betrachten die Situation, in dem in einem Experiment eine Meßkurve bestimmt wird, d.h. zu n bekannten Werten x_1, \dots, x_n werden die Meßpunkte $y_1 \pm \sigma_1, \dots, y_n \pm \sigma_n$ mit den Fehlern $\sigma_1, \dots, \sigma_n$ bestimmt. Wir nehmen an, daß diese n Messungen unabhängig sind. Man sucht nun eine Funktion $y = f(x)$, die diese Meßreihe möglichst gut beschreibt. Da diese Funktion nicht bekannt ist, verwendet man einen Ansatz für diese Funktion mit freien Parametern $\alpha_1, \dots, \alpha_r$:

$$y = f(x, \alpha).$$

Nun bildet man die Größe χ^2 wie folgt:

$$\chi^2(\alpha) = \sum_{j=1}^n \frac{(y_j - f(x_j, \alpha))^2}{\sigma_j^2}.$$

Offensichtlich nimmt χ^2 Werte zwischen 0 und Unendlich an. Die Werte von α , die die Größe χ^2 minimieren, bezeichnet man als den besten Fit.

10 Differential- und Integralrechnung in mehreren Dimensionen

Im letzten Semester haben wir uns eingehend mit der Differential- und Integralrechnung von Funktionen einer Variablen beschäftigt. Hierbei war der Definitionsbereich und Wertebereich jeweils eine Teilmenge von \mathbb{R} . Wir wollen dies nun verallgemeinern, indem wir \mathbb{R} durch \mathbb{R}^n ersetzen.

Wir unterscheiden drei Situationen:

- Ist der Definitionsbereich weiterhin eine Teilmenge von \mathbb{R} , aber der Wertebereich eine Teilmenge von \mathbb{R}^n , so spricht man von einer **Kurve**.
- Ist hingegen der Definitionsbereich eine Teilmenge von \mathbb{R}^n , der Wertebereich aber weiterhin eine Teilmenge von \mathbb{R} , so spricht man von einer **Funktion in mehreren Variablen**.
- Im allgemeinsten Fall ist der Definitionsbereich eine Teilmenge von \mathbb{R}^n und der Wertebereich eine Teilmenge von \mathbb{R}^m . In diesem Fall spricht man von einem **Vektorfeld**.

10.1 Topologische Grundbegriffe*

Um Begriffe wie Stetigkeit und Differenzierbarkeit auf den allgemeinen Fall übertragen zu können, benötigen wir einige Grundbegriffe.

Wir beginnen mit dem Begriff einer **Metrik**: Sei X eine Menge. Eine Metrik auf X ist eine Abbildung

$$d : X \times X \rightarrow \mathbb{R}, \\ (x, y) \rightarrow d(x, y)$$

mit den folgenden Eigenschaften:

$$d(x, y) = 0 \Leftrightarrow x = y, \\ d(x, y) = d(y, x), \\ d(x, y) \leq d(x, z) + d(z, y), \quad \text{Dreiecksungleichung}$$

Man spricht von einem **metrischen Raum**, falls die Menge X eine Metrik besitzt. Die Größe $d(x, y)$ bezeichnet als den Abstand der Punkte x und y .

Bemerkung: Aus den oben aufgeführten Eigenschaften folgt, daß

$$d(x, y) \geq 0$$

für alle $x, y \in X$ gilt. Dies zeigt man wie folgt:

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y).$$

Wir interessieren uns insbesondere für den Fall, daß die Menge X ein Vektorraum ist. Sei V nun also ein reeller Vektorraum. Wir definieren nun den Begriff einer **Norm** auf dem Vektorraum V . Wir werden sehen, daß dieser Begriff eng mit dem Begriff der Metrik verknüpft ist. Unter einer Norm auf V versteht man eine Abbildung

$$\begin{aligned} \|\cdot\| : V &\rightarrow \mathbb{R}, \\ \vec{x} &\rightarrow \|\vec{x}\|, \end{aligned}$$

mit den folgenden Eigenschaften:

$$\begin{aligned} \|\vec{x}\| = 0 &\Leftrightarrow \vec{x} = \vec{0}, \\ \|\lambda\vec{x}\| &= |\lambda| \cdot \|\vec{x}\|, \quad \lambda \in \mathbb{R}, \\ \|\vec{x} + \vec{y}\| &\leq \|\vec{x}\| + \|\vec{y}\|. \end{aligned}$$

Satz: Sei V ein reeller Vektorraum mit einer Norm. Dann wird durch

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$$

eine Metrik auf V definiert. Ein reeller Vektorraum mit einer Norm ist also automatisch ein metrischer Vektorraum.

Beispiele: Eine Norm auf dem \mathbb{R}^n ist gegeben durch

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Man nennt diese Norm die euklidische Norm (oder auch Zwei-Norm). Weitere Beispiele für Normen auf dem \mathbb{R}^n sind die Maximum-Norm oder allgemeiner die p -Normen. Die Maximum-Norm ist definiert durch

$$\|\vec{x}\|_\infty = \max(|x_1|, |x_2|, \dots, |x_n|).$$

Sei $p \geq 1$. Die p -Norm ist definiert durch

$$\|\vec{x}\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{\frac{1}{p}}.$$

Die euklidische Norm ergibt sich für $p = 2$. Man kann weiter zeigen, daß man die Maximum-Norm aus den p -Normen durch den Grenzfall $p \rightarrow \infty$ erhält.

Wir werden oft von einer **offener Menge** sprechen. Dieser Begriff wird durch eine Topologie auf einem Raum definiert. Ein **topologischer Raum** ist eine Menge M zusammen mit einer Familie \mathcal{T} von Untermengen von M , so daß die folgenden Eigenschaften erfüllt sind:

1. $\emptyset \in \mathcal{T}, M \in \mathcal{T}$
2. $U_1, U_2 \in \mathcal{T} \Rightarrow U_1 \cap U_2 \in \mathcal{T}$

3. Für jede Indexmenge A gilt $U_\alpha \in \mathcal{T}; \alpha \in A \Rightarrow \bigcup_{\alpha \in A} U_\alpha \in \mathcal{T}$

\mathcal{T} bezeichnet man als Topologie auf M , die Mengen $U \in \mathcal{T}$ nennt man offene Mengen. Man bezeichnet eine Teilmenge \mathcal{U} als abgeschlossen, falls das Komplement $M \setminus \mathcal{U}$ offen ist. Die Eigenschaft 2 impliziert daß der Durchschnitt endliche vieler offener Mengen wieder offen ist, die Eigenschaft 3 bedeutet, daß die Vereinigung beliebig vieler offener Mengen (auch unendlich vieler) wieder offen ist.

Einen topologischen Raum bezeichnet man als **Hausdorff**-Raum, falls es zu jedem Paar verschiedener Punkte $p_1, p_2 \in M$ offene Mengen $U_1, U_2 \in \mathcal{T}$ gibt, so daß die folgenden Bedingungen erfüllt sind:

$$p_1 \in U_1, \quad p_2 \in U_2, \quad U_1 \cap U_2 = \emptyset.$$

Wir wollen den Begriff einer Topologie für einen metrischen Raum etwas genauer betrachten. Sei X ein metrischer Raum mit der Metrik d . Wir definieren die offene Kugel mit Mittelpunkt x_0 und Radius r durch

$$B(x_0, r) = \{x \in X, d(x, x_0) < r\}$$

Wir bezeichnen eine Teilmenge $U \subset X$ als **Umgebung** des Punktes $x_0 \in X$, falls ein $\varepsilon > 0$ existiert, so daß

$$B(x_0, \varepsilon) \subset U.$$

Wir bezeichnen eine Teilmenge $U \subset X$ als **offen**, falls für alle $x \in U$ ein $\varepsilon > 0$ existiert, so daß

$$B(x, \varepsilon) \subset U.$$

Die so definierten offenen Mengen definieren eine Topologie und wir erhalten den folgenden Satz: Ein metrischer Raum ist ein topologischer Raum. Somit haben wir die Implikationen:

$$\text{normierter Raum} \Rightarrow \text{metrischer Raum} \Rightarrow \text{topologischer Raum}.$$

Wir haben darüberhinaus die zusätzliche Eigenschaft, daß ein metrischer Raum ein Hausdorff-Raum ist. Dies beweist man wie folgt: Seien $x, y \in X$ zwei Punkte mit $x \neq y$. Wir setzen $\varepsilon = d(x, y)/2$ und definieren

$$U = B(x, \varepsilon), \quad V = B(y, \varepsilon).$$

Diese beiden Umgebungen sind disjunkt. Angenommen, dies wäre nicht der Fall. Dann gäbe es ein $z \in U \cap V$ und es gälte

$$2\varepsilon = d(x, y) \leq d(x, z) + d(z, y) < \varepsilon + \varepsilon,$$

also $2\varepsilon < 2\varepsilon$, was zum Widerspruch führt.

Wir definieren noch den Begriff eines Randpunktes: Sei X ein metrischer Raum und Y eine Teilmenge von X . Ein Punkt $x \in X$ heißt **Randpunkt** von Y , wenn in jeder Umgebung von x sowohl ein Punkt von Y als auch ein Punkt $X \setminus Y$ liegt. Die Menge aller Randpunkte von Y wird mit ∂Y bezeichnet.

Wir haben die folgenden Eigenschaften:

- Die Menge $Y \setminus \partial Y$ ist offen.
- Die Menge $Y \cup \partial Y$ ist abgeschlossen.
- Der Rand ∂Y ist abgeschlossen.

Sei X ein metrischer Raum und $A \subset X$ eine Teilmenge. Man nennt A **beschränkt**, falls

$$\sup \{d(x, y) : x, y \in A\} < \infty.$$

Sei A eine Teilmenge des \mathbb{R}^n . Wir bezeichnen A als **kompakt**, falls A abgeschlossen und beschränkt ist.

10.2 Konvergenz in metrischen Räumen*

In diesem Abschnitt betrachten wir metrische Räume. Wir verallgemeinern nun die Konvergenzkriterien, die wir von Folgen reeller Zahlen kennen, auf Folgen von Punkten in metrischen Räumen. Im wesentlichen werden wir dabei den Betrag durch den Abstand (welcher durch die Metrik definiert ist) ersetzen.

Sei X ein metrischer Raum und (x_n) eine Folge von Punkten aus X . Die Folge (x_n) nennt man konvergent gegen den Punkt $x \in X$, in Symbolen

$$\lim_{n \rightarrow \infty} x_n = x,$$

falls zu jeder Umgebung U von x ein $N \in \mathbb{N}$ existiert, so daß

$$x_k \in U \quad \forall k \geq N.$$

Alternativ läßt sich dies wie folgt formulieren: Zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so daß

$$d(x_k, x) < \varepsilon \quad \forall k \geq N.$$

Wir bezeichnen eine Folge als **Cauchy-Folge**, falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so daß

$$d(x_k, x_m) < \varepsilon \quad \forall k, m \geq N.$$

Es ist leicht zu zeigen, daß jede konvergente Folge auch eine Cauchy-Folge ist. Interessanter ist allerdings die Frage, ob jede Cauchy-Folge auch eine konvergente Folge. Räume in der diese

Umkehrung gilt, bezeichnen wir als **vollständige Räume**. Einen Raum bezeichnet man daher als vollständig, falls in ihm jede Cauchy-Folge konvergiert.

Für den \mathbb{R}^n läßt sich zeigen, daß in ihm jede Cauchy-Folge konvergiert, der \mathbb{R}^n ist also vollständig.

Allgemein bezeichnet man einen vollständigen und normierten Vektorraum als **Banachraum**.

10.3 Stetigkeit*

Seien X und Y metrische Räume und

$$f : X \rightarrow Y$$

eine Abbildung. f heißt **stetig** im Punkte $x_0 \in X$, falls

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

d.h. wenn für jede Folge (x_j) von Punkten aus X mit $\lim x_j = x_0$ gilt

$$\lim_{j \rightarrow \infty} f(x_j) = f(x_0).$$

Alternativ kann man die Stetigkeit auch über das ε - δ -Kriterium definieren: f ist genau dann im Punkte x_0 stetig, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß

$$\tilde{d}(f(x), f(x_0)) < \varepsilon \quad \forall x \in X \text{ mit } d(x, x_0) < \delta.$$

Hierbei bezeichnet d die Metrik auf X und \tilde{d} die Metrik auf Y .

Die Abbildung f nennt man stetig auf X , falls f in jedem Punkt $x \in X$ stetig ist.

Wir behandeln noch die gleichmäßige Konvergenz von Funktionenfolgen: Seien X und Y metrische Räume, sowie

$$f_n : X \rightarrow Y, \quad n \in \mathbb{N}$$

und

$$f : X \rightarrow Y$$

Abbildungen. Man sagt, die Folge (f_n) **konvergiert gleichmäßig** gegen f , falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so daß

$$d(f_k(x), f(x)) < \varepsilon \quad \forall x \in X \text{ und } \forall k \geq N.$$

Der Punkt hierbei ist, daß N von x unabhängig ist.

Satz: Seien X und Y metrische Räume und $f_n : X \rightarrow Y$ eine Folge stetiger Funktionen, die gleichmäßig gegen die Funktion $f : X \rightarrow Y$ konvergiert. Dann ist auch f stetig.

Bemerkung: Es genügt nicht, nur die punktweise Konvergenz zu fordern.

10.4 Kurven

Sei $I \subset \mathbb{R}$ ein Intervall. Unter einer Kurve im \mathbb{R}^n versteht man eine stetige Abbildung

$$\vec{f} : I \rightarrow \mathbb{R}^n.$$

Wir bezeichnen mit f_k die Funktion

$$f_k : I \rightarrow \mathbb{R},$$

die die k -te Komponente von f beschreibt. Es ist also

$$\vec{f}(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \\ \dots \\ f_n(t) \end{pmatrix}.$$

Eine Kurve $\vec{f}(t)$ ist differenzierbar, falls alle Funktionen $f_k(t)$ für $1 \leq k \leq n$ differenzierbar sind.

Beispiel 1: Sei $\vec{x}_0 \in \mathbb{R}^n$ und $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$. Die Abbildung

$$\vec{f} : t \rightarrow \vec{v}t + \vec{x}_0$$

beschreibt eine Gerade im \mathbb{R}^n durch den Punkt \vec{x}_0 und mit dem Richtungsvektor \vec{v} .

Beispiel 2: Sei $r > 0$ und $c \neq 0$. Die Abbildung

$$\vec{f} : \mathbb{R} \rightarrow \mathbb{R}^3, \\ \vec{f}(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ ct \end{pmatrix}$$

beschreibt eine Schraubenlinie im \mathbb{R}^3 .

Bemerkung: In der Physik betrachtet man oft die Variable $t \in \mathbb{R}$ als Zeit und $\vec{f}(t) \in \mathbb{R}^n$ als Ort eines Teilchens. Die Kurve $\vec{f}(t)$ beschreibt dann die Bahn eines Teilchens.

Sei $\vec{f} : I \rightarrow \mathbb{R}^n$ eine differenzierbare Kurve. Man bezeichnet die Größe

$$\vec{f}'(t) = \begin{pmatrix} f'_1(t) \\ f'_2(t) \\ \dots \\ f'_n(t) \end{pmatrix} \in \mathbb{R}^n$$

als Tangentialvektor an die Kurve zum Parameterwert t .

Bemerkung: Betrachtet man t als die Zeit und $\vec{f}(t)$ als den Ort eines Teilchens, so entspricht $\vec{f}'(t)$ der Geschwindigkeit.

Wir betrachten nun die Bogenlänge einer Kurve. Hierzu sei $[a, b] \subset \mathbb{R}$ ein Intervall und $\vec{f} : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve. Unterteilt man das Intervall in n Teilintervalle

$$a = t_0 < t_1 < \dots < t_n = b$$

und verbindet man die Punkte $\vec{f}(t_{j-1})$ und $\vec{f}(t_j)$ durch eine Gerade, so erhält man einen Polygonzug. Die Länge dieses Polygonzugs ist

$$\sum_{j=1}^n \|\vec{f}(t_j) - \vec{f}(t_{j-1})\|.$$

Die Länge der Kurve wird nun definiert als der Grenzwert dieser Polygonzüge bei immer feineren Unterteilungen. Man erhält somit für die Länge der Kurve zwischen a und b

$$L = \int_a^b \|\vec{f}'(t)\| dt.$$

Beispiel: Sei $r > 0$. Wir betrachten die Kurve

$$\begin{aligned} \vec{f} &: [0, 2\pi] \rightarrow \mathbb{R}^2, \\ \vec{f}(t) &= \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix}. \end{aligned}$$

Diese Kurve beschreibt eine Kreislinie. Es ist

$$\vec{f}'(t) = \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix}$$

und

$$\|\vec{f}'(t)\| = \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} = r.$$

Somit erhält man für die Kurvenlänge

$$L = \int_0^{2\pi} \|\vec{f}'(t)\| dt = \int_0^{2\pi} r dt = 2\pi r,$$

was natürlich mit dem Umfang des Kreises übereinstimmt.

Sei $\vec{f} : I \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Man nennt \vec{f} **regulär**, falls

$$\vec{f}'(t) \neq \vec{0}$$

für alle $t \in I$ ist. Gilt dagegen $\vec{f}'(t_0) = \vec{0}$, so bezeichnet man t_0 als **singulären** Punkt. Ein Beispiel hierzu ist gegeben durch

$$\begin{aligned} \vec{f} &: \mathbb{R} \rightarrow \mathbb{R}^2, \\ \vec{f}(t) &= \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}. \end{aligned}$$

Diese Kurve hat einen singulären Punkt für $t_0 = 0$.

Wir betrachten nun zwei reguläre Kurven

$$\vec{f}: I_1 \rightarrow \mathbb{R}^n, \quad \vec{g}: I_2 \rightarrow \mathbb{R}^n,$$

die sich in einem Punkte schneiden, es gilt also

$$\vec{f}(t_1) = \vec{g}(t_2)$$

für ein $t_1 \in I_1$ und ein $t_2 \in I_2$. Wir definieren den **Schnittwinkel** θ der beiden Kurven als den Winkel zwischen den Tangentialvektoren an dem Schnittpunkt. Wir erhalten also

$$\cos \theta = \frac{\vec{f}'(t_1) \cdot \vec{g}'(t_2)}{\|\vec{f}'(t_1)\| \|\vec{g}'(t_2)\|}, \quad 0 \leq \theta \leq \pi.$$

Zum Schluss betrachten wir noch Parametertransformationen. Sei $\vec{f}: [a, b] \rightarrow \mathbb{R}^n$ eine Kurve und $[c, d] \subset \mathbb{R}$ ein weiteres Intervall. Sei weiter

$$\varphi: [c, d] \rightarrow [a, b]$$

eine stetige bijektive Abbildung. Dann ist die zusammengesetzte Abbildung

$$\begin{aligned} \vec{g} = \vec{f} \circ \varphi &: [c, d] \rightarrow \mathbb{R}^n, \\ \vec{g}(\tau) &= \vec{f}(\varphi(\tau)), \end{aligned}$$

wieder eine Kurve im \mathbb{R}^n . Man sagt, daß die Kurve \vec{g} aus \vec{f} durch die Parametertransformation φ hervorgeht. Die Kurvenpunkte im \mathbb{R}^n von \vec{f} und \vec{g} sind dieselben, sie werden aber unter Umständen verschieden durchlaufen. Es können die folgenden beiden Fälle auftreten:

1. φ ist streng monoton wachsend. In diesem Fall werden die Punkte von \vec{f} und \vec{g} gleich durchlaufen und man nennt φ orientierungstreu. Für die Anfangs- und Endpunkte gilt in diesem Fall

$$\vec{f}(a) = \vec{g}(c), \quad \vec{f}(b) = \vec{g}(d).$$

2. φ ist streng monoton fallend. In diesem Fall werden die Punkte von \vec{g} umgekehrt durchlaufen bezüglich \vec{f} . Man nennt φ in diesem Fall orientierungsumkehrend. Für die Anfangs- und Endpunkte gilt in diesem Fall

$$\vec{f}(a) = \vec{g}(d), \quad \vec{f}(b) = \vec{g}(c).$$

Betrachtet man die Variable t als die Zeit, so entspricht eine orientierungsumkehrenden Parametertransformation einer Zeitumkehr.

Für den Tangentialvektor der Kurve $\vec{g}(\tau)$ folgt mit Hilfe der Kettenregel

$$\vec{g}'(\tau) = \vec{f}'(\varphi(\tau)) \cdot \varphi'(\tau).$$

Insbesondere unterscheiden sich die Tangentialvektoren $\vec{g}'(\tau)$ und $\vec{f}'(t)$, wobei $t = \varphi(\tau)$ ist, nur um den skalaren Faktor $\varphi'(\tau)$.

Für die Bogenlängen gilt:

$$\int_a^b \|\vec{f}'(t)\| dt = \int_c^d \|\vec{g}'(\tau)\| d\tau.$$

Dies läßt sich leicht mit Hilfe der Substitutionsregel zeigen. Die beiden Kurven haben also die gleiche Bogenlänge, d.h. die Bogenlänge ist von der Parametrisierung unabhängig.

Wir betrachten ein Beispiel: Die Kurve

$$\begin{aligned} \vec{f} &: [-r, r] \rightarrow \mathbb{R}^2, \\ \vec{f}(t) &= \begin{pmatrix} t \\ \sqrt{r^2 - t^2} \end{pmatrix} \end{aligned}$$

beschreibt einen Halbkreis. Sei nun

$$\begin{aligned} \varphi &: [0, \pi] \rightarrow [-r, r], \\ \varphi(\tau) &= r \cos \tau \end{aligned}$$

eine Parametertransformation. Diese Transformation bildet das Intervall $[0, \pi]$ bijektiv auf das Intervall $[-r, r]$ ab. Wir erhalten

$$\begin{aligned} \vec{g} &: [0, \pi] \rightarrow \mathbb{R}^2, \\ \vec{g}(\tau) &= \begin{pmatrix} r \cos \tau \\ r \sin \tau \end{pmatrix}. \end{aligned}$$

Wir sehen, daß auch diese Kurve einen Halbkreis beschreibt.

Bemerkung: Die Parametertransformation φ in diesem Beispiel ist orientierungsumkehrend.

10.5 Funktionen in mehreren Variablen

Sei U eine Teilmenge des \mathbb{R}^n . Wir betrachten nun Funktionen

$$\begin{aligned} f &: U \rightarrow \mathbb{R}, \\ (x_1, \dots, x_n) &\rightarrow f(x_1, \dots, x_n). \end{aligned}$$

Die Funktion f ist **partiell differenzierbar in der i -ten Koordinate**, falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{h}$$

existiert. Man schreibt

$$\frac{\partial}{\partial x_i} f(x_1, \dots, x_i, \dots, x_n) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{h}.$$

Diese Formel zeigt auch, wie man die i -te partielle Ableitung berechnet: Man hält alle anderen Variablen $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ fest und nimmt die gewöhnliche Ableitung nach der Variablen x_i .

Wir nennen eine Funktion **partiell differenzierbar**, falls sie in allen Variablen partiell differenzierbar ist. Ebenso nennen wir eine Funktion **stetig partiell differenzierbar**, falls sie partiell differenzierbar ist und alle Ableitungen stetig sind.

Beispiel: Wir betrachten die Funktion

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \\ (x_1, x_2, x_3) \rightarrow \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Es ist

$$\frac{\partial}{\partial x_1} f(x_1, x_2, x_3) = \frac{x_1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}.$$

Wir können partielle Ableitungen auch hintereinander ausführen und erhalten höhere Ableitungen:

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(x_1, \dots, x_n) = \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_j} f(x_1, \dots, x_n) \right).$$

Man beachte, daß diese Schreibweise impliziert, daß zunächst die Ableitung nach x_j ausgeführt wird, und das Zwischenergebnis dann nach x_i abgeleitet wird. Wir interessieren uns dafür unter welchen Voraussetzungen das Endergebnis nicht von der Reihenfolge der Ableitungen abhängt:

Satz: Sei f zweimal stetig partiell differenzierbar. Dann gilt für die partiellen Ableitungen

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(x_1, \dots, x_n) = \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} f(x_1, \dots, x_n)$$

Allgemeiner gilt: Ist f k -mal stetig partiell differenzierbar, so vertauschen die k -ten partiellen Ableitungen:

$$\frac{\partial}{\partial x_{i_1}} \dots \frac{\partial}{\partial x_{i_k}} f(x_1, \dots, x_n) = \frac{\partial}{\partial x_{\sigma(i_1)}} \dots \frac{\partial}{\partial x_{\sigma(i_k)}} f(x_1, \dots, x_n),$$

wobei σ eine Permutation von (i_1, \dots, i_k) ist.

Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Wir sagen, daß f in $\vec{x}_0 \in \mathbb{R}^n$ ein **lokales Maximum** hat, falls eine Umgebung $U \subset \mathbb{R}^n$ von \vec{x}_0 existiert, so daß

$$f(\vec{x}_0) \geq f(\vec{x}),$$

für alle $\vec{x} \in U$. Gilt dagegen

$$f(\vec{x}_0) \leq f(\vec{x}),$$

für alle $\vec{x} \in U$, so spricht man von einem **lokalen Minimum**.

Es ist unmittelbar einsichtig, daß eine notwendige Bedingung für das Vorliegen eines lokalen Minimums oder lokalen Maximums das Verschwinden aller partiellen Ableitungen an der Stelle \vec{x}_0 ist:

$$\left. \frac{\partial}{\partial x_i} f(\vec{x}) \right|_{\vec{x}=\vec{x}_0} = 0.$$

Würde eine partielle Ableitung nicht verschwinden, so gibt es in jeder Umgebung von \vec{x}_0 einen Punkt, an dem $f(\vec{x}) < f(\vec{x}_0)$ gilt, sowie einen Punkt an dem $f(\vec{x}) > f(\vec{x}_0)$ gilt. Ist zum Beispiel die i -te partielle Ableitung ungleich Null, so betrachtet man hierzu zwei Punkte, die um einen infinitesimalen positiven bzw. negativen Wert in Richtung des i -ten Einheitsvektors verschoben sind.

Um eine hinreichende Bedingung für das Vorliegen eines lokalen Minimums oder Maximums zu finden betrachten wir die zweiten Ableitungen und definieren die **Hessesche Matrix**:

$$H_{ij}(\vec{x}) = \frac{\partial^2}{\partial x_i \partial x_j} f(\vec{x}), \quad 1 \leq i, j \leq n.$$

Da nach Voraussetzung f zweimal stetig differenzierbar ist, vertauschen die partiellen Ableitungen und die Hessesche Matrix ist offensichtlich symmetrisch:

$$H_{ij}(\vec{x}) = H_{ji}(\vec{x}).$$

Wir bezeichnen eine symmetrische $n \times n$ Matrix A als **positiv definit**, falls für alle $\vec{\xi} \in \mathbb{R}^n \setminus \{\vec{0}\}$ gilt:

$$\vec{\xi}^T A \vec{\xi} > 0.$$

Wir bezeichnen sie als **negativ definit**, falls für alle $\vec{\xi} \in \mathbb{R}^n \setminus \{\vec{0}\}$ gilt:

$$\vec{\xi}^T A \vec{\xi} < 0.$$

Wir bezeichnen die Matrix A als **indefinit**, falls es ein $\vec{\xi} \in \mathbb{R}^n$ und ein $\vec{\eta} \in \mathbb{R}^n$ gibt, so daß

$$\vec{\xi}^T A \vec{\xi} > 0, \quad \vec{\eta}^T A \vec{\eta} < 0.$$

Man findet auch die Begriffe “positiv semi-definit” und “negativ semi-definit”. Eine symmetrische $n \times n$ Matrix A nennt man **positiv semi-definit** bzw. **negativ semi-definit**, falls für alle $\vec{\xi} \in \mathbb{R}^n$ gilt:

$$\vec{\xi}^T A \vec{\xi} \geq 0, \quad \text{bzw.} \quad \vec{\xi}^T A \vec{\xi} \leq 0.$$

Satz: Eine reelle symmetrische $n \times n$ Matrix A besitzt eine Orthonormalbasis $\vec{e}_1, \dots, \vec{e}_n$ von Eigenvektoren und alle Eigenwerte $\lambda_1, \dots, \lambda_n$ sind reell.

$$A \vec{e}_j = \lambda_j \vec{e}_j.$$

Wir betrachten nun den Vektor $\vec{\xi}$ in dieser Basis:

$$\vec{\xi} = \sum_{i=1}^n c_i \vec{e}_i,$$

und erhalten durch Einsetzen:

$$\begin{aligned} \vec{\xi}^T A \vec{\xi} &= \left(\sum_{i=1}^n c_i \vec{e}_i^T \right) A \left(\sum_{j=1}^n c_j \vec{e}_j \right) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j (\vec{e}_i^T A \vec{e}_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n c_i c_j (\vec{e}_i^T \lambda_j \vec{e}_j) = \sum_{j=1}^n \lambda_j c_j^2. \end{aligned}$$

Somit erhalten wir die folgenden Aussagen: Eine reelle symmetrische $n \times n$ Matrix A ist positiv definit, falls alle Eigenwerte positiv sind. A ist negativ definit, falls alle Eigenwerte negativ sind. A ist indefinit, falls mindestens ein positiver und mindestens ein negativer Eigenwert existiert. A ist positiv semi-definit, falls alle Eigenwerte nicht negativ sind. A ist negativ semi-definit, falls alle Eigenwerte nicht positiv sind.

Um zu entscheiden, ob eine symmetrische Matrix positiv definit ist, ist es nicht notwendig die Eigenwerte zu bestimmen. Ein Kriterium, daß die Bestimmung der Eigenwerte vermeidet, wurde von Hurwitz angegeben: Sei

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

eine reelle symmetrische $n \times n$ Matrix. A ist positive definit, falls

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} > 0$$

für alle $k \in \{1, \dots, n\}$ gilt.

Wir kehren zur Betrachtung der lokalen Minima und Maxima einer Funktion zurück. Wir erhalten die folgende Aussage: Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion und $\vec{x}_0 \in \mathbb{R}^n$ ein Punkt, so daß

$$\left. \frac{\partial}{\partial x_j} f(\vec{x}) \right|_{\vec{x}=\vec{x}_0} = 0, \quad \forall 1 \leq j \leq n.$$

Ist die Hessesche Matrix $H_{ij}(\vec{x}_0)$ positiv definit, so besitzt f in \vec{x}_0 ein lokales Minimum. Ist sie negativ definit, so besitzt f in \vec{x}_0 ein lokales Maximum. Ist die Hessesche Matrix indefinit, so sagt man, daß f in \vec{x}_0 einen **Sattelpunkt** besitzt.

Beispiel 1: Sei

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

Im Punkte $\vec{x}_0 = (0, 0)$ verschwinden die partiellen Ableitungen:

$$\left. \frac{\partial f}{\partial x} \right|_{\vec{x}=(0,0)} = 2x|_{\vec{x}=(0,0)} = 0, \quad \left. \frac{\partial f}{\partial y} \right|_{\vec{x}=(0,0)} = 2y|_{\vec{x}=(0,0)} = 0.$$

Die Hessesche Matrix ist gegeben durch

$$H(\vec{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Diese Matrix ist positiv definit und f hat an der Stelle $\vec{x}_0 = (0, 0)$ ein Minimum.

Beispiel 2: Sei nun

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

Im Punkte $\vec{x}_0 = (0, 0)$ verschwinden die partiellen Ableitungen:

$$\left. \frac{\partial f}{\partial x} \right|_{\vec{x}=(0,0)} = 2x|_{\vec{x}=(0,0)} = 0, \quad \left. \frac{\partial f}{\partial y} \right|_{\vec{x}=(0,0)} = -2y|_{\vec{x}=(0,0)} = 0.$$

Die Hessesche Matrix ist gegeben durch

$$H(\vec{x}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

Diese Matrix ist indefinit und f hat an der Stelle $\vec{x}_0 = (0, 0)$ einen Sattelpunkt.

Abbildung 5 zeigt Beispiele für eine Funktion mit einem Minimum und für eine Funktion mit einem Sattelpunkt.

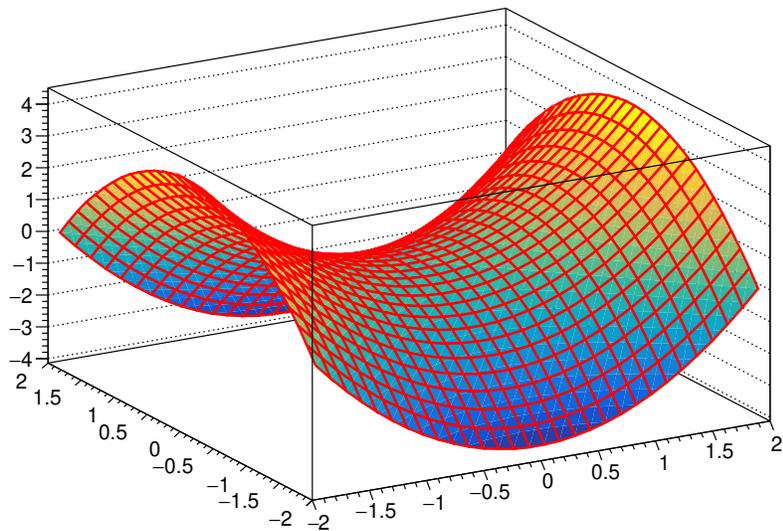
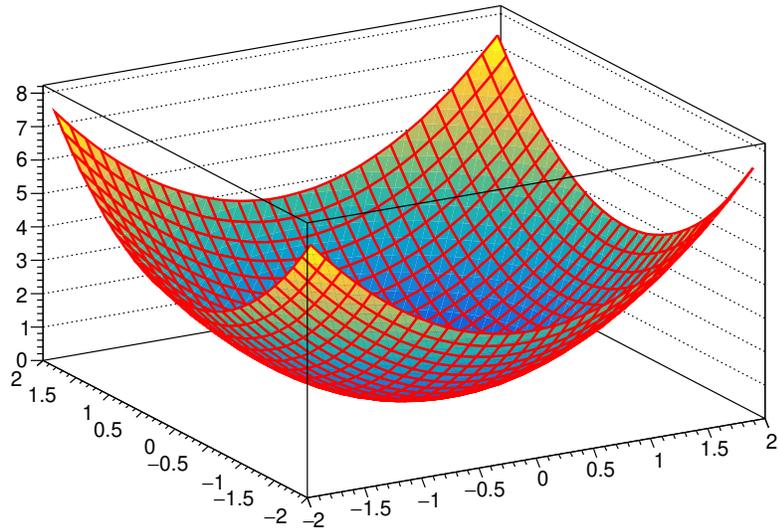


Abbildung 5: Darstellung einer Funktion zweier Variablen mit einem Minimum (oberes Bild) und einem Sattelpunkt (unteres Bild).

10.6 Vektorfelder

Wir betrachten nun den allgemeinen Fall einer Abbildung, in dem der Definitionsbereich U eine offene Teilmenge des \mathbb{R}^n und der Wertebereich W eine Teilmenge des \mathbb{R}^m ist:

$$\begin{aligned} \vec{f} &: U \rightarrow \mathbb{R}^m, \\ (x_1, \dots, x_n) &\rightarrow \vec{f}(x_1, \dots, x_n). \end{aligned}$$

Man bezeichnet \vec{f} als ein Vektorfeld. Jedem Punkt $(x_1, \dots, x_n) \in U$ wird ein Vektor $\vec{f} \in \mathbb{R}^m$ zugeordnet. Schreiben wir \vec{f} in Komponenten

$$\vec{f}(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \dots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$

so haben wir m Abbildungen

$$\begin{aligned} f_j &: U \rightarrow \mathbb{R}, \\ (x_1, \dots, x_n) &\rightarrow f_j(x_1, \dots, x_n). \end{aligned}$$

Wir schreiben im folgenden $\vec{x} = (x_1, \dots, x_n)$.

Wir betrachten drei Beispiele für Vektorfelder:

$$\begin{aligned} \vec{f}_1 &: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ \vec{f}_1(\vec{x}) &= \begin{pmatrix} 1 \\ \sin x_1 \end{pmatrix}, \\ \vec{f}_2 &: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ \vec{f}_2(\vec{x}) &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \\ \vec{f}_3 &: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ \vec{f}_3(\vec{x}) &= \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}. \end{aligned}$$

Diese Vektorfelder sind in Abbildung 6 graphisch dargestellt.

Wir bezeichnen eine Abbildung $\vec{f}: U \rightarrow \mathbb{R}^m$ als im Punkte $\vec{x}_0 \in U$ **total differenzierbar**, falls es eine lineare Abbildung

$$\begin{aligned} A &: \mathbb{R}^n \rightarrow \mathbb{R}^m, \\ \vec{x} &\rightarrow A\vec{x}, \end{aligned}$$

gibt, so daß in einer Umgebung von \vec{x}_0 gilt:

$$\vec{f}(\vec{x}_0 + \vec{\xi}) = \vec{f}(\vec{x}_0) + A\vec{\xi} + o(\|\vec{\xi}\|).$$

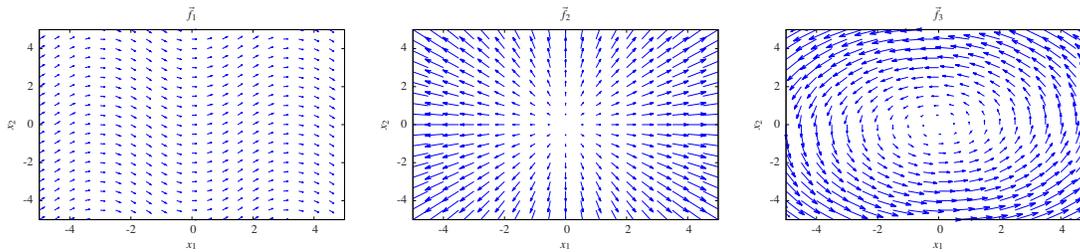


Abbildung 6: Darstellung der drei Vektorfelder aus dem Beispiel.

Die kleine “o”-Schreibweise bedeutet, daß das Restglied durch eine Funktion $\vec{\varphi}(\vec{\xi})$ gegeben ist, für die gilt

$$\lim_{\|\vec{\xi}\| \rightarrow 0} \frac{\vec{\varphi}(\vec{\xi})}{\|\vec{\xi}\|} = \vec{0}.$$

Das Restglied verschwindet also schneller als der lineare Term für $\|\vec{\xi}\| \rightarrow 0$. Die Bedingung an die totale Differenzierbarkeit bedeutet also, daß sich die Abbildung in einer hinreichend kleinen Umgebung von \vec{x}_0 durch eine Konstante $\vec{f}(\vec{x}_0)$ und einen linearen Term $A\vec{\xi}$ beschreiben läßt.

Neben der totalen Differenzierbarkeit haben wir natürlich noch die partiellen Ableitungen der i -ten Komponente f_i nach der j -ten Koordinate:

$$\frac{\partial f_i}{\partial x_j} = \lim_{h \rightarrow 0} \frac{f_i(x_1, \dots, x_j + h, \dots, x_n) - f_i(x_1, \dots, x_j, \dots, x_n)}{h}.$$

Diese partiellen Ableitungen definieren eine $m \times n$ Matrix J_{ij}

$$J_{ij}(\vec{x}) = \frac{\partial f_i}{\partial x_j}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n,$$

die man als **Jacobi-Matrix** oder Funktional-Matrix bezeichnet. Auch die Bezeichnung **Differential** wird verwendet, und man findet die Notation

$$D\vec{f}(\vec{x}) = J(\vec{x}).$$

Für den Zusammenhang zwischen totaler Differenzierbarkeit und partieller Differenzierbarkeit haben wir die folgenden Sätze:

Satz: Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und $\vec{f} : U \rightarrow \mathbb{R}^m$ eine Abbildung, die im Punkte $\vec{x}_0 \in U$ total differenzierbar sei, d.h.

$$\vec{f}(\vec{x}_0 + \vec{\xi}) = \vec{f}(\vec{x}_0) + A\vec{\xi} + o(\|\vec{\xi}\|).$$

Dann ist \vec{f} im Punkte \vec{x}_0 stetig und alle Komponenten $f_j : U \rightarrow \mathbb{R}$ von \vec{f} sind im Punkte \vec{x}_0 partiell differenzierbar und es gilt

$$\frac{\partial f_i}{\partial x_j}(\vec{x}_0) = A_{ij}.$$

Satz: Sei wieder $U \subset \mathbb{R}^n$ eine offene Teilmenge und $\vec{f} : U \rightarrow \mathbb{R}^m$ eine Abbildung. Es sei weiter vorausgesetzt, daß die Abbildung \vec{f} im Punkte $\vec{x}_0 \in U$ stetig partiell differenzierbar ist, d.h. alle partiellen Ableitungen

$$\frac{\partial f_i}{\partial x_j}(\vec{x}_0)$$

existieren und sind stetig. Dann ist \vec{f} in \vec{x}_0 total differenzierbar.

Wir haben also die folgenden Implikationen:

$$\text{stetig partiell differenzierbar} \Rightarrow \text{total differenzierbar} \Rightarrow \text{partiell differenzierbar}.$$

Die Umkehrungen gelten im allgemeinen nicht.

Für die totale Ableitung gilt die Kettenregel: Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen und

$$\vec{g} : U \rightarrow \mathbb{R}^m, \quad \vec{f} : V \rightarrow \mathbb{R}^k$$

Abbildungen mit $\vec{g}(U) \subset V$. Sei weiter vorausgesetzt, daß die Abbildung \vec{g} im Punkte $\vec{x}_0 \in U$ total differenzierbar sei und daß die Abbildung \vec{f} im Punkte $\vec{y}_0 = \vec{g}(\vec{x}_0)$ total differenzierbar sei. Dann ist die zusammengesetzte Abbildung

$$\vec{f} \circ \vec{g} : U \rightarrow \mathbb{R}^k$$

im Punkte \vec{x}_0 total differenzierbar und für ihr Differential gilt

$$D(\vec{f} \circ \vec{g})(\vec{x}_0) = D\vec{f}(\vec{y}_0) \cdot D\vec{g}(\vec{x}_0).$$

Betrachten wir hierzu ein Beispiel:

$$\begin{aligned} \vec{g} &: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ \vec{g}(\vec{x}) &= \begin{pmatrix} x_1^2 + x_2^2 \\ 1 + 2x_1x_2 \end{pmatrix}, \\ \vec{f} &: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ \vec{f}(\vec{y}) &= \begin{pmatrix} y_1 + y_2 \\ y_2^2 \end{pmatrix}. \end{aligned}$$

Die zusammengesetzte Abbildung ist somit

$$\vec{f} \circ \vec{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2,$$

$$(\vec{f} \circ \vec{g})(\vec{x}) = \begin{pmatrix} 1 + x_1^2 + 2x_1x_2 + x_2^2 \\ (1 + 2x_1x_2)^2 \end{pmatrix}.$$

Sei weiter

$$\vec{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{y}_0 = \vec{g}(\vec{0}) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Für die Differentiale findet man

$$D\vec{g}(\vec{x}_0) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}_{x_1=1, x_2=1} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix},$$

$$D\vec{f}(\vec{y}_0) = \begin{pmatrix} 1 & 1 \\ 0 & 2y_2 \end{pmatrix}_{y_1=2, y_2=3} = \begin{pmatrix} 1 & 1 \\ 0 & 6 \end{pmatrix}.$$

Somit ist

$$D\vec{f}(\vec{y}_0) \cdot D\vec{g}(\vec{x}_0) = \begin{pmatrix} 1 & 1 \\ 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 12 & 12 \end{pmatrix}.$$

Andererseits erhält man dieses Ergebnis auch aus der direkten Rechnung:

$$D(\vec{f} \circ \vec{g})(\vec{x}_0) = \begin{pmatrix} 2x_1 + 2x_2 & 2x_1 + 2x_2 \\ 4x_2(1 + 2x_1x_2) & 4x_1(1 + 2x_1x_2) \end{pmatrix}_{x_1=1, x_2=1} = \begin{pmatrix} 4 & 4 \\ 12 & 12 \end{pmatrix}.$$

Wir führen im Zusammenhang mit Vektorfeldern noch einige wichtige Begriffe ein: Sei $U \subset \mathbb{R}^n$ eine offene Menge und $\varphi : U \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion von n Variablen. Die partiellen Ableitungen von φ definieren ein Vektorfeld, welches man als den **Gradienten** von φ bezeichnet:

$$\text{grad } \varphi : U \rightarrow \mathbb{R}^n,$$

$$\text{grad } \varphi(\vec{x}) = \begin{pmatrix} \frac{\partial \varphi(\vec{x})}{\partial x_1} \\ \dots \\ \frac{\partial \varphi(\vec{x})}{\partial x_n} \end{pmatrix}.$$

Der Gradient einer skalaren Funktion ist also ein Vektorfeld, daß in der j -ten Komponente die j -te partielle Ableitung enthält. Führt man den Nabla-Operator $\vec{\nabla}$ ein,

$$\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \dots \\ \frac{\partial}{\partial x_n} \end{pmatrix},$$

so läßt sich dieses Vektorfeld auch als

$$\text{grad } \varphi = \vec{\nabla} \varphi$$

schreiben.

Bemerkung: $\vec{\nabla}$ ist ein Operator, der auf eine Größe, wie zum Beispiel eine Funktion, die abgeleitet werden kann, wirkt. Man sollte diese Größe daher immer mitangeben. Mathematische Beziehungen, in denen die Größe auf die ein Operator wirkt fehlt, machen nur Sinn, wenn sie für alle möglichen Größen des Problems (wie zum Beispiel für alle Testfunktionen) gelten.

Beispiel: Wir betrachten die Funktion

$$\begin{aligned}\varphi &: \mathbb{R}^3 \rightarrow \mathbb{R}, \\ \varphi(\vec{x}) &= x_1^2 + x_2^2 + x_3^2.\end{aligned}$$

Wir erhalten für den Gradienten

$$\text{grad } \varphi(\vec{x}) = \vec{\nabla} \varphi(\vec{x}) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix}.$$

Bemerkung: Wir hatten bereits gesehen, daß eine notwendige Bedingung für das Vorliegen eines lokalen Maximums bzw. eines lokalen Minimums im Punkte \vec{x}_0 das Verschwinden aller partiellen Ableitungen in diesem Punkte ist. Das Verschwinden aller partiellen Ableitungen ist gleichbedeutend mit der Aussage

$$\vec{\nabla} \varphi(\vec{x}_0) = \vec{0},$$

d.h. der Gradient verschwindet.

Sei $U \subset \mathbb{R}^n$ eine offene Menge und $\vec{f}: U \rightarrow \mathbb{R}^n$ ein partiell differenzierbares Vektorfeld. Wir definieren die **Divergenz** dieses Vektorfeldes als eine skalare Funktion der n Variablen

$$\text{div } \vec{f} : U \rightarrow \mathbb{R},$$

die durch

$$\text{div } \vec{f}(\vec{x}) = \sum_{j=1}^n \frac{\partial f_j(\vec{x})}{\partial x_j}$$

gegeben ist. Mit Hilfe des Nabla-Operators schreibt man auch oft

$$\text{div } \vec{f}(\vec{x}) = \vec{\nabla} \cdot \vec{f}(\vec{x}).$$

Beispiel: Wir betrachten das Vektorfeld

$$\begin{aligned}\vec{f} &: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \\ \vec{f}(\vec{x}) &= \begin{pmatrix} x_1^2 + x_2 \\ 3x_2 - x_1 \\ 5x_3 + 7x_2 \end{pmatrix}.\end{aligned}$$

Wir erhalten für die Divergenz

$$\operatorname{div} \vec{f}(\vec{x}) = \vec{\nabla} \cdot \vec{f}(\vec{x}) = 2x_1 + 3 + 5 = 2x_1 + 8.$$

Es ist auch interessant die Divergenz der drei Vektorfelder aus Abbildung 6 zu berechnen. Man findet:

$$\operatorname{div} \vec{f}_1(\vec{x}) = \vec{\nabla} \cdot \vec{f}_1(\vec{x}) = \frac{\partial}{\partial x_1} 1 + \frac{\partial}{\partial x_2} \sin x_1 = 0,$$

$$\operatorname{div} \vec{f}_2(\vec{x}) = \vec{\nabla} \cdot \vec{f}_2(\vec{x}) = \frac{\partial}{\partial x_1} x_1 + \frac{\partial}{\partial x_2} x_2 = 2,$$

$$\operatorname{div} \vec{f}_3(\vec{x}) = \vec{\nabla} \cdot \vec{f}_3(\vec{x}) = \frac{\partial}{\partial x_1} (-x_2) + \frac{\partial}{\partial x_2} x_1 = 0.$$

Von diesen drei Beispielen hat also nur \vec{f}_2 eine nicht-verschwindende Divergenz. Die Divergenz beschreibt die Quellen und Senken eines Vektorfeldes.

Wir betrachten noch die folgende Kombination von Gradient und Divergenz: Sei $U \subset \mathbb{R}^n$ eine offene Menge und $\varphi : U \rightarrow \mathbb{R}$ eine zweimal partiell differenzierbare Funktion von n Variablen. Wir wenden erst den Gradienten auf φ an, und dann die Divergenz auf das resultierende Vektorfeld. Wir erhalten somit wieder eine skalare Funktion:

$$\Delta \varphi : U \rightarrow \mathbb{R},$$

$$\Delta \varphi(\vec{x}) = \operatorname{div} \operatorname{grad} \varphi(\vec{x}) = \sum_{j=1}^n \frac{\partial^2 \varphi(\vec{x})}{\partial x_j^2}.$$

Mit Hilfe des Nabla-Operators können wir wieder schreiben:

$$\Delta \varphi(\vec{x}) = \vec{\nabla} \cdot \vec{\nabla} \varphi(\vec{x}).$$

Wir bezeichnen mit

$$\Delta = \vec{\nabla} \cdot \vec{\nabla} = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$$

den **Laplace-Operator**.

Wir betrachten noch den Spezialfall eines Vektorfeldes in drei Dimensionen:

$$\vec{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3.$$

Wir setzen voraus, daß \vec{A} partiell differenzierbar ist. Hier können wir noch eine weitere Operation einführen, die man als **Rotation** bezeichnet und wie folgt definiert ist:

$$\operatorname{rot} \vec{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3,$$

$$\operatorname{rot} \vec{A}(\vec{x}) = \begin{pmatrix} \frac{\partial A_3(\vec{x})}{\partial x_2} - \frac{\partial A_2(\vec{x})}{\partial x_3} \\ \frac{\partial A_1(\vec{x})}{\partial x_3} - \frac{\partial A_3(\vec{x})}{\partial x_1} \\ \frac{\partial A_2(\vec{x})}{\partial x_1} - \frac{\partial A_1(\vec{x})}{\partial x_2} \end{pmatrix}.$$

Mit Hilfe des Nabla-Operators und des Kreuzproduktes läßt sich dies auch schreiben als

$$\operatorname{rot} \vec{A}(\vec{x}) = \vec{\nabla} \times \vec{A}(\vec{x}).$$

Beispiel: Sei

$$\begin{aligned} \vec{A} &: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \\ \vec{A}(\vec{x}) &= \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}. \end{aligned}$$

Dann ist

$$\operatorname{rot} \vec{A}(\vec{x}) = \vec{\nabla} \times \vec{A}(\vec{x}) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

Kehren wir nocheinmal zu den Vektorfeldern aus Abbildung 6 zurück. Diese Vektorfelder sind Abbildungen von \mathbb{R}^2 nach \mathbb{R}^2 , daher ist die Operation der Rotation nicht unmittelbar darauf anwendbar. Wir können aber trotzdem für ein Vektorfeld $\vec{f} = (f_1, f_2)$ die anti-symmetrische Ableitung

$$\frac{\partial}{\partial x_1} f_2 - \frac{\partial}{\partial x_2} f_1$$

betrachten. Wir finden:

$$\begin{aligned} \frac{\partial}{\partial x_1} f_{12} - \frac{\partial}{\partial x_2} f_{11} &= \frac{\partial}{\partial x_1} \sin x_1 - \frac{\partial}{\partial x_2} 1 = \cos(x_1), \\ \frac{\partial}{\partial x_1} f_{22} - \frac{\partial}{\partial x_2} f_{21} &= \frac{\partial}{\partial x_1} x_2 - \frac{\partial}{\partial x_2} x_1 = 0, \\ \frac{\partial}{\partial x_1} f_{32} - \frac{\partial}{\partial x_2} f_{31} &= \frac{\partial}{\partial x_1} x_1 - \frac{\partial}{\partial x_2} (-x_2) = 2. \end{aligned}$$

Die Rotation beschreibt die Wirbel eines Vektorfeldes.

Rechenregeln: Seien

$$\begin{aligned} \vec{f} &: \mathbb{R}^n \rightarrow \mathbb{R}^n, \\ \vec{g} &: \mathbb{R}^n \rightarrow \mathbb{R}^n, \end{aligned}$$

Vektorfelder und

$$\begin{aligned}\varphi &: \mathbb{R}^n \rightarrow \mathbb{R}, \\ \psi &: \mathbb{R}^n \rightarrow \mathbb{R},\end{aligned}$$

Funktionen. Wir betrachten nun einige Rechenregeln des Nabla-Operators. Wir nehmen an, daß alle Felder und Funktionen zweimal stetig partiell differenzierbar sind. Im Folgenden wollen wir implizit annehmen, daß in Regeln in denen das Vektorprodukt bzw. die Rotation vorkommt, $n = 3$ vorausgesetzt wird. In allen anderen Regeln ist n beliebig.

Rotation eines Gradientenfeldes:

$$\begin{aligned}\operatorname{rot} \operatorname{grad} \varphi &= 0, \\ \vec{\nabla} \times (\vec{\nabla} \varphi) &= 0.\end{aligned}$$

Beweis: Wir betrachten die erste Komponente von $\operatorname{rot} \operatorname{grad} \varphi$:

$$\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_3} \varphi - \frac{\partial}{\partial x_3} \frac{\partial}{\partial x_2} \varphi = 0.$$

Gleiches gilt für die anderen Komponenten. Ein Gradientenfeld ist also rotationsfrei.

Divergenz eines Rotationsfeldes:

$$\begin{aligned}\operatorname{div} \operatorname{rot} \vec{f} &= 0, \\ \vec{\nabla} \cdot (\vec{\nabla} \times \vec{f}) &= 0.\end{aligned}$$

Beweis:

$$\begin{aligned}\vec{\nabla} \cdot (\vec{\nabla} \times \vec{f}) &= \frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_2} f_3 - \frac{\partial}{\partial x_3} f_2 \right) + \frac{\partial}{\partial x_2} \left(\frac{\partial}{\partial x_3} f_1 - \frac{\partial}{\partial x_1} f_3 \right) + \frac{\partial}{\partial x_3} \left(\frac{\partial}{\partial x_1} f_2 - \frac{\partial}{\partial x_2} f_1 \right) \\ &= 0.\end{aligned}$$

Ein Rotationsfeld ist also divergenzfrei.

Produktregeln:

$$\begin{aligned}\operatorname{div} (\varphi \vec{f}) &= (\operatorname{grad} \varphi) \cdot \vec{f} + \varphi \operatorname{div} \vec{f}, \\ \vec{\nabla} (\varphi \vec{f}) &= (\vec{\nabla} \varphi) \cdot \vec{f} + \varphi \vec{\nabla} \vec{f}.\end{aligned}$$

Beweis:

$$\vec{\nabla} (\varphi \vec{f}) = \sum_{j=1}^n \frac{\partial}{\partial x_j} (\varphi f_j) = \sum_{j=1}^n \left(\frac{\partial \varphi}{\partial x_j} \right) f_j + \sum_{j=1}^n \varphi \frac{\partial f_j}{\partial x_j} = (\vec{\nabla} \varphi) \cdot \vec{f} + \varphi \vec{\nabla} \vec{f}.$$

Analog gilt:

$$\begin{aligned}\operatorname{rot}(\varphi \vec{f}) &= (\operatorname{grad} \varphi) \times \vec{f} + \varphi \operatorname{rot} \vec{f}, \\ \vec{\nabla} \times (\varphi \vec{f}) &= (\vec{\nabla} \varphi) \times \vec{f} + \varphi \vec{\nabla} \times \vec{f},\end{aligned}$$

sowie

$$\begin{aligned}\operatorname{div}(\vec{f} \times \vec{g}) &= (\operatorname{rot} \vec{f}) \cdot \vec{g} - \vec{f} \cdot (\operatorname{rot} \vec{g}), \\ \vec{\nabla}(\vec{f} \times \vec{g}) &= (\vec{\nabla} \times \vec{f}) \cdot \vec{g} - \vec{f} \cdot (\vec{\nabla} \times \vec{g}).\end{aligned}$$

Man beachte das für den Laplace-Operator gilt:

$$\Delta(\varphi \psi) = (\Delta \varphi) \psi + 2(\vec{\nabla} \varphi) \cdot (\vec{\nabla} \psi) + \varphi \Delta \psi.$$

Beweis:

$$\begin{aligned}\Delta(\varphi \psi) &= \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2} (\varphi \psi) = \sum_{j=1}^n \frac{\partial}{\partial x_j} \left[\left(\frac{\partial}{\partial x_j} \varphi \right) \psi + \varphi \frac{\partial}{\partial x_j} \psi \right] \\ &= \sum_{j=1}^n \left[\left(\frac{\partial^2}{\partial x_j^2} \varphi \right) \psi + 2 \left(\frac{\partial}{\partial x_j} \varphi \right) \frac{\partial}{\partial x_j} \psi + \varphi \frac{\partial^2}{\partial x_j^2} \psi \right] \\ &= (\Delta \varphi) \psi + 2(\vec{\nabla} \varphi) \cdot (\vec{\nabla} \psi) + \varphi \Delta \psi.\end{aligned}$$

Die zweimalige Anwendung einer Rotation läßt sich Vereinfachen zu

$$\begin{aligned}\operatorname{rot} \operatorname{rot} \vec{f} &= \operatorname{grad}(\operatorname{div} \vec{f}) - \Delta \vec{f}, \\ \vec{\nabla} \times (\vec{\nabla} \times \vec{f}) &= \vec{\nabla}(\vec{\nabla} \cdot \vec{f}) - \Delta \vec{f}.\end{aligned}$$

Beweis: Wir betrachten die erste Komponente von $\operatorname{rot} \operatorname{rot} \vec{f}$:

$$\begin{aligned}\frac{\partial}{\partial x_2} \left[\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right] - \frac{\partial}{\partial x_3} \left[\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right] &= \frac{\partial^2 f_2}{\partial x_1 \partial x_2} + \frac{\partial^2 f_3}{\partial x_1 \partial x_3} - \frac{\partial^2 f_1}{\partial x_2^2} - \frac{\partial^2 f_1}{\partial x_3^2} \\ &= \frac{\partial^2 f_1}{\partial x_1^2} + \frac{\partial^2 f_2}{\partial x_1 \partial x_2} + \frac{\partial^2 f_3}{\partial x_1 \partial x_3} - \frac{\partial^2 f_1}{\partial x_1^2} - \frac{\partial^2 f_1}{\partial x_2^2} - \frac{\partial^2 f_1}{\partial x_3^2} \\ &= \frac{\partial}{\partial x_1} \left[\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3} \right] - \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) f_1 \\ &= \frac{\partial}{\partial x_1} (\vec{\nabla} \cdot \vec{f}) - \Delta f_1.\end{aligned}$$

Der Beweis für die beiden anderen Komponenten verläuft analog.

10.7 Integralrechnung in mehreren Variablen

In diesem Abschnitt möchten wir uns mit der Integration in mehreren Variablen beschäftigen. Sei Q ein achsenparalleler kompakter Quader im \mathbb{R}^n . Q ist gegeben durch

$$Q = I_1 \times I_2 \times \dots \times I_n,$$

wobei jedes $I_j = [a_j, b_j] \subset \mathbb{R}$ ein beschränktes und abgeschlossenes Intervall ist. Auf diesem Quader betrachten wir eine stetige Funktion von n Variablen

$$\begin{aligned} f &: Q \rightarrow \mathbb{R}, \\ (x_1, \dots, x_n) &\rightarrow f(x_1, \dots, x_n). \end{aligned}$$

Halten wir x_2, \dots, x_n fest und integrieren wir über das Intervall I_1 , so erhalten wir eine Funktion von $(n-1)$ Variablen

$$F_1(x_2, \dots, x_n) = \int_{a_1}^{b_1} dx_1 f(x_1, \dots, x_n).$$

Ist f stetig, so ist auch F_1 wieder stetig.

Wir können diesen Prozess nun fortsetzen und definieren F_{12} als eine Funktion von $(n-2)$ Variablen, die wir durch Integration von F_1 über I_2 erhalten:

$$F_{12}(x_3, \dots, x_n) = \int_{a_2}^{b_2} dx_2 F_1(x_2, \dots, x_n) = \int_{a_2}^{b_2} dx_2 \left(\int_{a_1}^{b_1} dx_1 f(x_1, \dots, x_n) \right).$$

Setzen wir dies für alle Variablen fort, so erhalten wir nach der letzten Integration eine reelle Zahl, die wir als das Integral von f über Q bezeichnen:

$$I = \int_{a_n}^{b_n} dx_n \dots \left(\int_{a_2}^{b_2} dx_2 \left(\int_{a_1}^{b_1} dx_1 f(x_1, \dots, x_n) \right) \right).$$

Wir schreiben oft auch

$$I = \int_Q dx_1 \dots dx_n f(x_1, \dots, x_n)$$

oder

$$I = \int_Q d^n x f(x_1, \dots, x_n),$$

bzw.

$$I = \int_Q d^n x f(\vec{x}).$$

Bemerkung: Das Integral I ist unabhängig von der Reihenfolge, in der die Integrationen ausgeführt werden.

Beispiel: Sei $Q = [1, 2] \times [0, 1]$ und

$$f : Q \rightarrow \mathbb{R}, \\ f(x_1, x_2) = x_1^2 + 3x_1x_2 + 5x_2^3.$$

Dann ist

$$\begin{aligned} I &= \int_1^2 dx_1 \int_0^1 dx_2 (x_1^2 + 3x_1x_2 + 5x_2^3) = \int_1^2 dx_1 \left(x_1^2x_2 + \frac{3}{2}x_1x_2^2 + \frac{5}{4}x_2^4 \Big|_{x_2=0}^{x_2=1} \right) \\ &= \int_1^2 dx_1 \left(x_1^2 + \frac{3}{2}x_1 + \frac{5}{4} \right) = \frac{1}{3}x_1^3 + \frac{3}{4}x_1^2 + \frac{5}{4}x_1 \Big|_{x_1=1}^{x_1=2} \\ &= \frac{8}{3} + 3 + \frac{5}{2} - \frac{1}{3} - \frac{3}{4} - \frac{5}{4} = \frac{35}{6} \end{aligned}$$

Wir betrachten nun Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Wir definieren den **Träger** (engl. support) der Funktion f als die abgeschlossene Hülle der Menge aller Punkte, in denen die Funktion von Null verschieden ist:

$$\text{supp } f = \overline{\{\vec{x} \in \mathbb{R}^n : f(\vec{x}) \neq 0\}}.$$

Wir bezeichnen mit $C_c(\mathbb{R}^n)$ die Menge aller stetigen Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit kompakten Träger. Diese Menge ist ein Vektorraum. Aus dieser Definition folgt, daß es zu jedem $f \in C_c(\mathbb{R}^n)$ einen achsenparallelen Quader Q gibt, so daß f außerhalb von Q verschwindet. Für $f \in C_c(\mathbb{R}^n)$ setzen wir nun

$$\int_{\mathbb{R}^n} d^n x f(\vec{x}) = \int_Q d^n x f(\vec{x}).$$

Das Integral hat die folgenden drei Eigenschaften:

1. **Linearität:** Für $f, g \in C_c(\mathbb{R}^n)$ und $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} \int_{\mathbb{R}^n} d^n x [f(\vec{x}) + g(\vec{x})] &= \int_{\mathbb{R}^n} d^n x f(\vec{x}) + \int_{\mathbb{R}^n} d^n x g(\vec{x}), \\ \int_{\mathbb{R}^n} d^n x \lambda f(\vec{x}) &= \lambda \int_{\mathbb{R}^n} d^n x f(\vec{x}). \end{aligned}$$

2. **Monotonie:** Gilt $f(\vec{x}) \leq g(\vec{x})$ für alle $\vec{x} \in \mathbb{R}^n$, so folgt

$$\int_{\mathbb{R}^n} d^n x f(\vec{x}) \leq \int_{\mathbb{R}^n} d^n x g(\vec{x}).$$

3. **Translationsinvarianz:** Für alle $\vec{x}_0 \in \mathbb{R}^n$ gilt

$$\int_{\mathbb{R}^n} d^n x f(\vec{x}) = \int_{\mathbb{R}^n} d^n x f(\vec{x} - \vec{x}_0).$$

Das Integral ordnet also jeder Funktion $f \in C_c(\mathbb{R}^n)$ eine reelle Zahl zu und definiert so eine Abbildung

$$I : C_c(\mathbb{R}^n) \rightarrow \mathbb{R}, \\ f \rightarrow \int_{\mathbb{R}^n} d^n x f(\vec{x}).$$

Eine Abbildung von einem Funktionenraum in die reellen Zahlen nennt man ein **Funktional**. Die obigen drei Eigenschaften des Integrals besagen, daß das Integral ein lineares, monotones und translationsinvariantes Funktional auf dem Funktionenraum $C_c(\mathbb{R}^n)$ ist.

Eine wichtige Aussage der Integralrechnung lautet, daß diese drei Eigenschaften das Integral fast vollständig charakterisieren. Es gilt der folgende Satz: Sei $J : C_c(\mathbb{R}^n) \rightarrow \mathbb{R}$ ein lineares, monotones und translationsinvariantes Funktional. Dann gibt es eine Konstante $c \in \mathbb{R}_+$, so daß

$$J(f) = c \int_{\mathbb{R}^n} d^n x f(\vec{x}).$$

Bemerkung: Ein lineares, monotones und translationsinvariantes Funktional $J : C_c(\mathbb{R}^n) \rightarrow \mathbb{R}$ nennt man auch ein **Haarsches Maß** auf dem \mathbb{R}^n . Der obige Satz sagt aus, daß auf dem \mathbb{R}^n bis auf einen konstanten Faktor genau ein Haarsches Maß gibt.

Wir haben oben das Integral über den \mathbb{R}^n für stetige Funktionen mit kompakten Träger definiert. Nun gibt es darüberhinaus noch weitere Funktionen, für die das Integral sinnvoll definiert werden kann.

Wir erinnern uns an die Definition des Integrals in einer Dimension: Wir hatten zunächst die Menge aller Treppenfunktionen $T[a, b]$ auf einem Intervall $[a, b]$ definiert. Dies sind alle Funktionen t , für die es eine Unterteilung

$$a = x_0 < x_1 < \dots < x_n = b$$

gibt, so daß t auf jedem offenen Intervall $]x_{j-1}, x_j[$ konstant ist. Der Wert auf diesem Intervall sei mit c_j bezeichnet. Das Integral einer Treppenfunktion definiert man als

$$\int_a^b t(x) dx = \sum_{j=1}^n c_j (x_j - x_{j-1}).$$

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beliebige beschränkte Funktion und $t \in T[a, b]$. Man schreibt $f \geq t$ falls $f(x) \geq t(x)$ für alle $x \in [a, b]$ gilt. Für f definiert man das Ober- und Unterintegral wie folgt:

$$\int_a^{b^*} f(x) dx = \inf \left\{ \int_a^b t(x) dx; t \in T[a, b], t \geq f \right\},$$

$$\int_a^* f(x) dx = \sup \left\{ \int_a^b t(x) dx; t \in T[a, b], t \leq f \right\},$$

Ist für eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ das Oberintegral gleich dem Unterintegral, so bezeichnet man die Funktion als Riemann-integrierbar und setzt

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

In einem zweiten Schritt haben wir dann den Integralbegriff auf nicht-kompakte Intervalle erweitert. Dies führte zur Definition der uneigentlichen Integrale:

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{a \rightarrow -\infty} \lim_{b \rightarrow \infty} \int_a^b f(x) dx.$$

Wir wollen nun den Integralbegriff im \mathbb{R}^n auf eine größere Klasse als die stetigen Funktionen mit kompakten Träger erweitern. Dies wird uns auf die Definition des **Lebesgue-Integrals** führen.

Wir betrachten eine Funktionenfolge $f_n \in C_c(\mathbb{R}^n)$ von stetigen Funktionen mit kompakten Träger, die monoton wachsend sein soll, d.h.

$$f_1 \leq f_2 \leq f_3 \leq \dots$$

Bemerkung: Dies ist eine Bedingung, die für jedes \vec{x} gelten muss, d.h.

$$f_1(\vec{x}) \leq f_2(\vec{x}) \leq f_3(\vec{x}) \leq \dots \quad \text{für alle } \vec{x} \in \mathbb{R}^n.$$

Für jedes \vec{x} konvergiert die Folge $f_n(\vec{x})$ entweder gegen eine reelle Zahl oder uneigentlich gegen plus Unendlich. Wir definieren nun eine Abbildung

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\},$$

$$f(\vec{x}) = \lim_{n \rightarrow \infty} f_n(\vec{x}).$$

Die Menge der Funktionen, die sich so als Limiten von Funktionenfolgen aus $C_c(\mathbb{R}^n)$ erhalten lassen, bezeichnen wir als die Menge $\mathcal{H}^\uparrow(\mathbb{R}^n)$. Wir definieren das Integral für $f \in \mathcal{H}^\uparrow(\mathbb{R}^n)$ durch

$$\int_{\mathbb{R}^n} f(\vec{x}) d^n x = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^n} f_n(\vec{x}) d^n x \in \mathbb{R} \cup \{\infty\}.$$

Analog können wir eine monoton fallende Funktionenfolge $f_n \in C_c(\mathbb{R}^n)$ betrachten:

$$f_1 \geq f_2 \geq f_3 \geq \dots$$

Für jedes $\vec{x} \in \mathbb{R}^n$ konvergiert diese Folge entweder gegen eine reelle Zahl oder uneigentlich gegen minus Unendlich. Dies definiert wieder eine Abbildung

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\},$$

$$f(\vec{x}) = \lim_{n \rightarrow \infty} f_n(\vec{x}).$$

Die Menge der Funktionen, die sich so als Limiten von Funktionenfolgen aus $C_c(\mathbb{R}^n)$ erhalten lassen, bezeichnen wir als die Menge $\mathcal{H}^\downarrow(\mathbb{R}^n)$. Wir definieren das Integral für $f \in \mathcal{H}^\downarrow(\mathbb{R}^n)$ durch

$$\int_{\mathbb{R}^n} f(\vec{x}) d^n x = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^n} f_n(\vec{x}) d^n x \in \mathbb{R} \cup \{-\infty\}.$$

Es läßt sich zeigen, daß die Funktionen, die sowohl zu $\mathcal{H}^\uparrow(\mathbb{R}^n)$ als auch zu $\mathcal{H}^\downarrow(\mathbb{R}^n)$ gehören, genau die Funktionen sind, die in $C_c(\mathbb{R}^n)$ liegen:

$$C_c(\mathbb{R}^n) = \mathcal{H}^\uparrow(\mathbb{R}^n) \cap \mathcal{H}^\downarrow(\mathbb{R}^n).$$

Mit Hilfe der Funktionen aus $\mathcal{H}^\uparrow(\mathbb{R}^n)$ und $\mathcal{H}^\downarrow(\mathbb{R}^n)$ definieren wir für $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$ das Ober- und Unterintegral:

$$\int_{\mathbb{R}^n}^* f(\vec{x}) d^n x = \inf \left\{ \int_{\mathbb{R}^n} \varphi(\vec{x}) d^n x; \varphi \in \mathcal{H}^\uparrow(\mathbb{R}^n), \varphi \geq f \right\},$$

$$\int_{\mathbb{R}^n} f(\vec{x}) d^n x = \sup \left\{ \int_{\mathbb{R}^n} \varphi(\vec{x}) d^n x; \varphi \in \mathcal{H}^\downarrow(\mathbb{R}^n), \varphi \leq f \right\}.$$

Man nennt die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$ **Lebesgue-integrierbar**, falls

$$-\infty < \int_{\mathbb{R}^n}^* f(\vec{x}) d^n x = \int_{\mathbb{R}^n} f(\vec{x}) d^n x < \infty$$

gilt. In diesem Fall setzt man

$$\int_{\mathbb{R}^n} f(\vec{x}) d^n x = \int_{\mathbb{R}^n}^* f(\vec{x}) d^n x = \int_{\mathbb{R}^n} f(\vec{x}) d^n x.$$

Den Funktionen aus $\mathcal{H}^\uparrow(\mathbb{R}^n)$ und $\mathcal{H}^\downarrow(\mathbb{R}^n)$ in der Definition des Lebesgue-Integrals entsprechen die Treppenfunktionen in der Definition des Riemann-Integrals.

Satz: Jede Riemann-integrierbare Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist auch Lebesgue-integrierbar.

Satz: Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Lebesgue-integrierbar, so ist auch die Funktion $|f|$ Lebesgue-integrierbar.

Bemerkung: Der obige Satz, daß jede Riemann-integrierbare Funktion auch Lebesgue-integrierbar ist, bezieht sich nur auf Riemann-Integral im eigentlichen Sinne. Für uneigentliche Integrale ist diese Aussage im allgemeinen nicht korrekt, wie folgendes Gegenbeispiel zeigt:

$$f : \mathbb{R} \rightarrow \mathbb{R},$$

$$f(x) = \frac{(-1)^n}{n}, \quad \text{falls } n \leq x < n+1, \quad n \in \mathbb{Z}.$$

Integriert man diese Funktion von Null bis Unendlich, so gilt für das uneigentliche Riemann-Integral

$$\int_0^\infty dx f(x) = \lim_{n \rightarrow \infty} \int_0^n dx f(x) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \frac{(-1)^{j-1}}{j} = \ln 2.$$

Diese Funktion ist allerdings nicht Lebesgue-integrierbar, denn dann wäre auch die Funktion $|f|$ Lebesgue-integrierbar. Das Integral über $|f|$ entspricht der harmonischen Reihe, welche divergiert.

Die Klasse der Funktionen, die Lebesgue-integrierbar sind, ist größer als die Klasse der (eigentlich) Riemann-integrierbaren Funktionen. Hierzu betrachten wir die Dirichlet-Funktion

$$f : [0, 1] \rightarrow [0, 1],$$

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

Das Riemann-Integral existiert nicht, da alle Obersummen stets 1 und alle Untersummen stets 0 sind. Andererseits läßt sich zeigen, daß die Menge der rationalen Zahlen \mathbb{Q} eine Lebesgue-Nullmenge in der Menge der reellen Zahlen \mathbb{R} ist. Daher existiert das Lebesgue-Integral und hat den Wert Null.

Hierzu müssen wir noch den Begriff einer **Lebesgue-Nullmenge** erläutern: Sei M eine Teilmenge des \mathbb{R}^n . Wir definieren die charakteristische Funktion χ_M durch

$$\chi_M(\vec{x}) = \begin{cases} 1, & x \in M, \\ 0, & x \notin M. \end{cases}$$

Wir bezeichnen eine Menge als integrierbar, falls ihre charakteristische Funktion Lebesgue-integrierbar ist. In diesem Fall definieren wir das Volumen (oder Lebesgue-Maß) der Menge M durch

$$\text{Vol}(M) = \int_{\mathbb{R}^n} d^n x \chi_M(\vec{x}).$$

Wir bezeichnen eine Menge als Nullmenge, falls sie Lebesgue-integrierbar ist und das Lebesgue-Maß Null hat.

Seien zwei Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und $g : \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben. Wir nennen die Funktionen f und g **fast überall gleich**, falls die Menge

$$\{x \in \mathbb{R}^n : f(\vec{x}) \neq g(\vec{x})\}$$

eine Lebesgue-Nullmenge ist.

Satz: Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und $g : \mathbb{R}^n \rightarrow \mathbb{R}$ zwei Funktionen, die fast überall gleich sind. Ist f Lebesgue-integrierbar, so ist auch g Lebesgue-integrierbar und es gilt

$$\int_{\mathbb{R}^n} d^n x f(\vec{x}) = \int_{\mathbb{R}^n} d^n x g(\vec{x}).$$

Fazit: Bei der Berechnung von Integralen kommt es auf Nullmengen nicht an.

Beispiel: Sei $G \subset \mathbb{R}^2$ das Gebiet

$$G = \{\vec{x} \in \mathbb{R}^2 | x_1^2 + x_2^2 \leq 1\}$$

und $f(x_1, x_2)$ die Funktion

$$f : G \rightarrow \mathbb{R} \\ f(x_1, x_2) = 1.$$

Dann ergibt

$$I = \int_G d^2 x \cdot 1 = \int_{-1}^1 dx_1 \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} dx_2 \cdot 1 = 2 \int_{-1}^1 dx_1 \sqrt{1-x_1^2}$$

$$= x_1 \sqrt{1 - x_1^2} + \arcsin(x_1) \Big|_{-1}^1 = \pi$$

die Fläche des Einheitskreises.

Bemerkung: In dieser Rechnung mussten wir eine Stammfunktion zu $\sqrt{1 - x_1^2}$ finden. Wir werden im nächsten Abschnitt sehen, daß wir die Berechnung des Integrals mithilfe einer Koordinatentransformation wesentlich vereinfachen können.

Wir wenden uns nun den Rechen-techniken für Integrale zu. In der Praxis führt man oft eine Koordinatentransformation durch. Wir wollen nun hierzu die Transformationsformel für das Integral betrachten.

Wir kennen bereits für die Integration einer Variable die Substitutionsregel: Sei $\varphi : [a, b] \rightarrow W_1$ eine stetig differenzierbare Funktion und $f : D_2 \rightarrow \mathbb{R}$ eine stetige Funktion mit $W_1 \subset D_2 \subset \mathbb{R}$. Dann gilt

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

Wir verallgemeinern diese Formel nun auf den \mathbb{R}^n . Seien U und V zwei offene Teilmengen des \mathbb{R}^n und

$$\vec{\varphi} : U \rightarrow V$$

eine bijektive stetig differenzierbare Abbildung, die außerdem noch die Bedingung erfüllt, daß auch die Umkehrabbildung

$$(\vec{\varphi})^{-1} : V \rightarrow U$$

stetig differenzierbar ist. Wir bezeichnen mit $D\vec{\varphi}$ die Funktionalmatrix

$$D\vec{\varphi}(\vec{x}) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n}{\partial x_1} & \cdots & \frac{\partial \varphi_n}{\partial x_n} \end{pmatrix}$$

Sei nun

$$f : V \rightarrow \mathbb{R}$$

eine Lebesgue-integrierbare Funktion. Dann gilt

$$\int_U d^n x |\det D\vec{\varphi}(\vec{x})| f(\vec{\varphi}(\vec{x})) = \int_V d^n y f(\vec{y}).$$

Diese Formel bezeichnet man als die **Transformationsformel für Integrale**.

10.7.1 Die Substitutionsregel als Spezialfall der Transformationsformel

Betrachten wir zunächst den Fall $n = 1$. Da φ invertierbar sein soll, nehmen wir an, daß

$$\varphi : [a, b] \rightarrow W_1$$

streng monoton wachsend ist. In diesem Fall ist $W_1 = [\varphi(a), \varphi(b)]$ und

$$D\varphi(x) = \frac{\partial\varphi(x)}{\partial x} = \varphi'(x) > 0.$$

Somit lautet die Transformationsformel

$$\int_a^b dx f(\varphi(x)) \varphi'(x) = \int_{\varphi(a)}^{\varphi(b)} dy f(y).$$

Dies ist die uns schon bekannte Substitutionsregel. Ist φ streng monoton fallend, so ist $W_1 = [\varphi(b), \varphi(a)]$ und $\varphi'(x) < 0$. In diesem Fall hat man

$$\int_a^b dx f(\varphi(x)) \varphi'(x) = - \int_{\varphi(b)}^{\varphi(a)} dy f(y) = \int_{\varphi(a)}^{\varphi(b)} dy f(y).$$

10.7.2 Polarkoordinaten im \mathbb{R}^2

Als nächstes betrachten wir den Fall $n = 2$. Schreiben wir für $\vec{x} \in \mathbb{R}^2$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

so bezeichnet man x_1 und x_2 als die **kartesischen Koordinaten**.

Wir können jeden Punkt im \mathbb{R}^2 auch durch andere Koordinaten beschreiben. Ein Beispiel sind die **Polarkoordinaten** (r, ϕ) . Wir setzen

$$\begin{aligned} x_1 &= r \cos \phi, \\ x_2 &= r \sin \phi, \end{aligned}$$

wobei

$$r \in [0, \infty[, \quad \phi \in [0, 2\pi[.$$

Wir betrachten die offene Menge

$$U = \{(r, \phi) : r > 0, 0 < \phi < 2\pi\} \subset \mathbb{R}^2$$

und die Koordinatentransformation

$$\vec{\varphi} : U \rightarrow \mathbb{R}^2,$$

$$\vec{\varphi}(r, \phi) = \begin{pmatrix} r \cos \phi \\ r \sin \phi \end{pmatrix}.$$

Die Bildmenge von $\vec{\varphi}$ ist der \mathbb{R}^2 bis auf die positive x_1 -Achse. Dies ist der \mathbb{R}^2 bis auf eine Nullmenge. Es ist

$$D\vec{\varphi} = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}.$$

Die Determinante der Jacobi-Matrix ist somit

$$\det D\vec{\varphi} = r(\cos^2 \phi + \sin^2 \phi) = r.$$

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Lebesgue-integrierbare Funktion. Wir erhalten aus der Transformationsformel

$$\int_{\mathbb{R}^2} d^2 x f(x_1, x_2) = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 f(x_1, x_2) = \int_0^{\infty} dr \int_0^{2\pi} d\phi r f(r \cos \phi, r \sin \phi)$$

Beispiel: Sei $f(x_1, x_2)$ die Funktion

$$f = \begin{cases} 1, & x_1^2 + x_2^2 \leq 1, \\ 0, & \text{sonst.} \end{cases}$$

f ist also nur auf dem Gebiet $G = \{\vec{x} \in \mathbb{R}^2 | x_1^2 + x_2^2 \leq 1\}$ von Null verschieden. Auf G ist f konstant und hat dort den Wert 1. Wir erhalten

$$\int_G d^2 x f(x_1, x_2) = \int_0^1 dr \int_0^{2\pi} d\phi r = 2\pi \int_0^1 dr r = \pi.$$

In Polarkoordinaten läßt sich die Fläche des Einheitskreises wesentlich leichter berechnen.

10.7.3 Zylinderkoordinaten im \mathbb{R}^3

Für Probleme, die symmetrisch um eine Achse sind, bietet sich die Verwendung von Zylinderkoordinaten an. Wir nehmen an, daß die Symmetrieachse die x_3 -Achse ist.

Schreiben wir für $\vec{x} \in \mathbb{R}^3$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

so bezeichnet man x_1 , x_2 und x_3 als die kartesischen Koordinaten.

Für die Zylinderkoordinaten schreiben wir

$$x_1 = r \cos \phi,$$

$$\begin{aligned}x_2 &= r \sin \phi, \\x_3 &= z,\end{aligned}$$

wobei

$$r \in [0, \infty[, \quad \phi \in [0, 2\pi[. \quad z \in \mathbb{R}.$$

Wir erhalten aus der Transformationsformel

$$\int_{\mathbb{R}^3} d^3x f(x_1, x_2, x_3) = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{\infty} dx_3 f(x_1, x_2, x_3) = \int_0^{\infty} dr \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} dz r f(r \cos \phi, r \sin \phi, z)$$

10.7.4 Kugelkoordinaten im \mathbb{R}^3

Für Probleme, die kugelsymmetrisch sind, bietet sich die Verwendung von Kugelkoordinaten an. Wir betrachten auch die Transformation auf Kugelkoordinaten für $n = 3$. Wir setzen

$$\begin{aligned}x_1 &= r \sin \theta \cos \phi, \\x_2 &= r \sin \theta \sin \phi, \\x_3 &= r \cos \theta,\end{aligned}$$

$$r \in [0, \infty[, \quad \theta \in [0, \pi], \quad \phi \in [0, 2\pi[.$$

Wir betrachten die offene Menge

$$U = \{(r, \theta, \phi) : r > 0, 0 < \theta < \pi, 0 < \phi < 2\pi\} \subset \mathbb{R}^3$$

und die Koordinatentransformation

$$\begin{aligned}\vec{\varphi} &: U \rightarrow \mathbb{R}^3, \\ \vec{\varphi}(r, \theta, \phi) &= \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \end{pmatrix}.\end{aligned}$$

Die Jacobi-Matrix lautet in diesem Fall

$$D\vec{\varphi}(r, \theta, \phi) = \begin{pmatrix} \sin \theta \cos \phi & r \cos \theta \cos \phi & -r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}.$$

Für die Determinante dieser Matrix findet man

$$\det D\vec{\varphi}(r, \theta, \phi) = r^2 \sin \theta.$$

Sei nun $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ eine Lebesgue-integrierbare Funktion. Aus der Transformationsformel folgt

$$\begin{aligned} \int_{\mathbb{R}^3} d^3x f(x_1, x_2, x_3) &= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{\infty} dx_3 f(x_1, x_2, x_3) \\ &= \int_0^{\infty} dr \int_0^{\pi} d\theta \int_0^{2\pi} d\phi r^2 \sin \theta f(r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta). \end{aligned}$$

Beispiel 1: Wir berechnen das Volumen einer Kugel mit Radius R . Zu diesem Zweck definieren wir

$$f(x_1, x_2, x_3) = \begin{cases} 1, & x_1^2 + x_2^2 + x_3^2 \leq R^2, \\ 0, & \text{sonst.} \end{cases}$$

Dann ist

$$\begin{aligned} \int_{\mathbb{R}^3} d^3x f(x_1, x_2, x_3) &= \int_0^R dr \int_0^{\pi} d\theta \int_0^{2\pi} d\phi r^2 \sin \theta \\ &= 2\pi \int_0^R dr r^2 \int_0^{\pi} d\theta \sin \theta \\ &= 2\pi \cdot \frac{1}{3} R^3 \cdot 2 = \frac{4}{3} \pi R^3. \end{aligned}$$

Beispiel 2: Wir berechnen die Gesamtmasse eines Objektes mit der Dichteverteilung

$$\rho(x_1, x_2, x_3) = \begin{cases} \rho_0 \left(1 - \frac{x_1^2 + x_2^2 + x_3^2}{R^2}\right), & x_1^2 + x_2^2 + x_3^2 \leq R^2, \\ 0, & \text{sonst.} \end{cases}$$

Wir erhalten

$$\begin{aligned} M &= \int_{\mathbb{R}^3} d^3x \rho(x_1, x_2, x_3) = \rho_0 \int_0^R dr \int_0^{\pi} d\theta \int_0^{2\pi} d\phi r^2 \sin \theta \left(1 - \frac{r^2}{R^2}\right) \\ &= 4\pi \rho_0 \int_0^R dr \left(r^2 - \frac{r^4}{R^2}\right) = 4\pi \rho_0 \left(\frac{1}{3} - \frac{1}{5}\right) R^3 \\ &= \frac{8}{15} \pi \rho_0 R^3. \end{aligned}$$

10.7.5 Der Maßtensor

Wir betrachten nochmal die Transformationsformel

$$\int_U d^n x |\det D\vec{\varphi}(\vec{x})| f(\vec{\varphi}(\vec{x})) = \int_V d^n y f(\vec{y}).$$

wobei $\vec{\varphi} : U \rightarrow V$ eine bijektive Abbildung ist und sowohl $\vec{\varphi}$ als auch $\vec{\varphi}^{-1}$ stetig partiell differenzierbar sind. U und V sind offene Teilmengen des \mathbb{R}^n . Die Funktionalmatrix ist eine $n \times n$ -Matrix, dessen Einträge gegeben sind durch

$$(D\vec{\varphi})_{ij} = \frac{\partial \varphi_i}{\partial x_j}.$$

Wir definieren uns nun eine weitere $n \times n$ -Matrix g , deren Einträge durch

$$g_{j_1 j_2} = \sum_{i=1}^n \frac{\partial \varphi_i}{\partial x_{j_1}} \cdot \frac{\partial \varphi_i}{\partial x_{j_2}}$$

gegeben sind. Diese Matrix wird als **Maßtensor** bezeichnet. Offensichtlich gilt in Matrixschreibweise

$$g = (D\vec{\varphi})^T (D\vec{\varphi}).$$

Nehmen wir auf beiden Seiten die Determinante, so findet man

$$\det g = (\det D\vec{\varphi})^2,$$

da die Determinante der transponierten Matrix gleich der Determinante der ursprünglichen Matrix ist. Die rechte Seite ist ein Quadrat und daher nie negativ. Wir können also problemlos die Wurzel nehmen und erhalten

$$\sqrt{\det g} = |\det D\vec{\varphi}|.$$

Man bezeichnet $|g| = \det g$ als **Gramsche Determinante**. Somit läßt sich die Transformationsformel auch schreiben als

$$\int_U d^n x \sqrt{|g|} f(\vec{\varphi}(\vec{x})) = \int_V d^n y f(\vec{y}).$$

Diese Form ist in Hinblick auf Verallgemeinerungen nützlich.

10.8 Integration auf Mannigfaltigkeiten

Die Transformationsformel erlaubt es uns, das Integral über ein n -dimensionales Gebiet V zu reparametrisieren und durch ein Integral über ein (einfacheres) n -dimensionales Gebiet U auszudrücken. Hierbei sind U und V Teilmengen des \mathbb{R}^n .

Wir wollen dies nun verallgemeinern und betrachten nun die Integration über k -dimensionale Gebiete M . Diese Gebiete sollen wieder Teilmengen des \mathbb{R}^n sein, wobei $k \leq n$. Wir sagen, daß M in den \mathbb{R}^n eingebettet ist.

In der Anwendung wichtig sind vor allem Gebiete, die in den dreidimensionalen Raum \mathbb{R}^3 eingebettet sind. Für $k = 2$ spricht man in diesem Fall von einer Fläche und für $k = 1$ spricht man von einer Kurve. Im allgemeinen sind diese eingebetteten Gebiete keine Geraden oder Ebenen, sondern beliebig gekrümmte Objekte. Während der \mathbb{R}^3 ein Vektorraum ist, sind diese eingebetteten Gebiete im allgemeinen keine Vektorräume mehr. Dies führt uns zu der Definition von Mannigfaltigkeiten.

10.8.1 Definition einer Mannigfaltigkeit*

Wir betrachten zunächst zwei topologische Räume M und N sowie eine bijektive Abbildung

$$\varphi : M \rightarrow N$$

zwischen den beiden Räumen. Da f bijektiv sein soll, existiert die Umkehrabbildung

$$\varphi^{-1} : N \rightarrow M.$$

Man bezeichnet φ als **Homeomorphismus**, falls sowohl φ als auch φ^{-1} stetig sind.

Man bezeichnet φ als einen **C^1 -Diffeomorphismus**, falls sowohl φ als auch φ^{-1} stetig partiell differenzierbar sind. Etwas allgemeiner bezeichnet man φ als einen **C^k -Diffeomorphismus**, falls sowohl φ als auch φ^{-1} k -mal stetig partiell differenzierbar sind. Ist sowohl φ als auch φ^{-1} beliebig oft partiell differenzierbar, so spricht man von einem **C^∞ -Diffeomorphismus**. Anstelle von **C^∞ -Diffeomorphismus** verwendet man auch die Sprechweise **glatter Diffeomorphismus**.

Besonders wichtig ist der Fall, in dem einer der beiden Räume, sagen wir N , der \mathbb{R}^n ist. Man kann die Abbildungen φ und φ^{-1} benutzen, um Rechnungen im Raum M auf den Raum \mathbb{R}^n zurückzuführen.

Eine **offene Karte** von M ist ein Paar (U, φ) , wobei U eine offene Untermenge von M und φ ein Homeomorphismus von U auf eine offene Untermenge von \mathbb{R}^n ist.

Eine **differenzierbare Mannigfaltigkeit** der Dimension n ist ein Hausdorff-Raum mit einer Familie offener Karten $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$, so daß

M1:

$$M = \bigcup_{\alpha \in A} U_\alpha.$$

M2: Für jedes Paar $\alpha, \beta \in A$ ist die Abbildung $\varphi_\beta \circ \varphi_\alpha^{-1}$ eine beliebig oft differenzierbare Abbildung von $\varphi_\alpha(U_\alpha \cap U_\beta)$ auf $\varphi_\beta(U_\alpha \cap U_\beta)$.

Eine differenzierbare Mannigfaltigkeit wird auch oft als C^∞ -Mannigfaltigkeit oder glatte Mannigfaltigkeit bezeichnet. Da wir uns nur mit differenzierbaren Mannigfaltigkeiten beschäftigen werden, werden wir oft die Bezeichnung "differenzierbar" weglassen und daher einfach und kurz von Mannigfaltigkeiten sprechen.

Die Familie offener Karten $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ bezeichnet man als **Atlas**.

Für $p \in U_\alpha$ und

$$\varphi_\alpha(p) = (x_1(p), \dots, x_n(p)),$$

bezeichnet man die Menge U_α als die **Koordinatenumgebung** von p . Die Zahlen $x_i(p)$ bezeichnet man als die **lokalen Koordinaten** von p .

Bemerkung: M schaut lokal in jeder Koordinatenumgebung wie der \mathbb{R}^n aus, doch gilt dies nicht global.

Beispiele

a) \mathbb{R}^n : Der Raum \mathbb{R}^n ist eine Mannigfaltigkeit. \mathbb{R}^n kann mit einer einzigen Karte überdeckt werden.

b) S^1 : Die Kreislinie

$$S^1 = \{\vec{x} \in \mathbb{R}^2 \mid \|\vec{x}\|^2 = 1\}$$

ist eine Mannigfaltigkeit. Für einen Atlas benötigt man mindestens zwei Karten. Eine Möglichkeit besteht in der stereographischen Projektion. Sei die Kreislinie durch die Punkte (x_1, x_2) im \mathbb{R}^2 gegeben, die

$$x_1^2 + x_2^2 = 1$$

erfüllen. Wir projizieren vom Punkte $(0, 1)$ (“Nordpol”) auf die x_1 -Achse. Hierzu legen wir eine Gerade durch die Punkte $(0, 1)$ und $p = (x_1, x_2)$ und bestimmen den Schnittpunkt $(z_1, 0)$ mit der x_1 -Achse. Man findet

$$z_1 = \frac{x_1}{1 - x_2}.$$

Dies definiert die Abbildung

$$\begin{aligned} \varphi_1 &: S^1 \setminus \{(0, 1)\} \rightarrow \mathbb{R}, \\ \varphi_1(p) &= z_1. \end{aligned}$$

Die Abbildung φ_1 bildet jeden Punkt der Kreislinie auf die x_1 -Achse ab, mit Ausnahme des Punktes $(0, 1)$. Sie überdeckt also nicht die komplette Mannigfaltigkeit. Wir benötigen also eine zweite Karte, die den Punkt $(0, 1)$ enthält. Hierzu können wir die Projektion vom Punkte $(0, -1)$ (“Südpol”) benutzen. Hier findet man

$$z'_1 = \frac{x_1}{1 + x_2}.$$

Dies definiert nun eine zweite Abbildung

$$\begin{aligned} \varphi_2 &: S^1 \setminus \{(0, -1)\} \rightarrow \mathbb{R}, \\ \varphi_2(p) &= z'_1 \end{aligned}$$

und somit eine zweite Karte. Diese beiden Karten zusammen überdecken die gesamte Kreislinie und bilden somit einen Atlas der Mannigfaltigkeit. Für die erste Karte ergibt sich die Umkehrabbildung φ_1^{-1} zu

$$x_1 = \frac{2z_1}{z_1^2 + 1}, \quad x_2 = \frac{z_1^2 - 1}{z_1^2 + 1}.$$

Für die zweite Karte findet man für φ_2^{-1}

$$x_1 = \frac{2z'_1}{z_1'^2 + 1}, \quad x_2 = -\frac{z_1'^2 - 1}{z_1'^2 + 1}.$$

Wir betrachten noch die Abbildung

$$\begin{aligned} \varphi_2 \circ \varphi_1^{-1} : \mathbb{R} &\rightarrow \mathbb{R} \\ z'_1 &= \frac{x_1}{1 + x_2} = \frac{\frac{2z_1}{z_1^2 + 1}}{1 + \frac{z_1^2 - 1}{z_1^2 + 1}} = \frac{1}{z_1}. \end{aligned}$$

Diese Abbildung ist auf

$$\mathbb{R} \setminus \{0\} = \varphi_1^{-1}(S^1 \setminus \{(0, 1), (0, -1)\})$$

beliebig oft differenzierbar.

c) S^n : Die n -Sphäre, definiert durch

$$S^n = \{\vec{x} \in \mathbb{R}^{n+1} \mid \|\vec{x}\|^2 = 1\}$$

d) $\mathbb{P}^n(\mathbb{R})$: Der projektive Raum definiert als alle Linien durch den Ursprung im \mathbb{R}^{n+1} :

$$(x_0, x_1, \dots, x_n) = \lambda(x'_0, x'_1, \dots, x'_n), \quad \lambda \neq 0.$$

e) Die Menge aller Rotationsmatrizen in zwei Dimensionen:

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix},$$

Die Menge all dieser Matrizen bildet eine Mannigfaltigkeit, die homeomorph zur Kreislinie S^1 ist.

Gegenbeispiele

Zum besseren Verständniss der Definition einer Mannigfaltigkeit seien einige Beispiele angeführt, die **keine** Mannigfaltigkeiten sind:

a) Eine ein-dimensionale Linie vereinigt mit einer zwei-dimensionalen Fläche, wie zum Beispiel durch

$$x_3(x_1^2 + x_2^2) = 0$$

definiert. Die so definierte Punktmenge ist an einigen Punkten homeomorph zu \mathbb{R} , an anderen Punkten homeomorph zu \mathbb{R}^2 . In der Definition einer Mannigfaltigkeit wird aber verlangt, daß die Menge an allen Punkten homeomorph zu \mathbb{R}^n für ein festes n ist.

b) Der Kegel

$$x_1^2 + x_2^2 - x_3^2 = 0.$$

Die Umgebung des Punktes $(0, 0, 0)$ läßt sich nicht homeomorph auf den \mathbb{R}^2 abbilden.

c) Ein einzelnes Kegelsegment

$$x_1^2 + x_2^2 - x_3^2 = 0, \quad x_3 \geq 0.$$

Hier läßt sich zwar eine Umgebung des Punktes $(0, 0, 0)$ stetig auf den \mathbb{R}^2 abbilden, aber nicht differenzierbar.

d) Das Liniensegment

$$[0, 1]$$

Hier gibt es zu den Endpunkten keine offene Umgebungen.

10.8.2 Untermannigfaltigkeiten des \mathbb{R}^{n*}

In der Anwendung sind k -dimensionale Mannigfaltigkeiten von Bedeutung, die in dem \mathbb{R}^n eingebettet sind.

Sei M eine k -dimensionale Mannigfaltigkeit und (V, φ^{-1}) eine Karte, d.h. V ist eine offene Menge von M und φ^{-1} ein Diffeomorphismus

$$\varphi^{-1} : V \rightarrow \mathbb{R}^k.$$

(Wir haben hier gegenüber der Definition einer Karte die Bezeichnungen φ und φ^{-1} vertauscht. Da diese Abbildung invertierbar sind, spielt es keine große Rolle welche man mit φ bzw. φ^{-1} bezeichnet.) Sei $U = \varphi^{-1}(V)$ das Bild von V . U ist eine offene Menge im \mathbb{R}^k . Die Abbildung

$$\varphi : U \rightarrow V$$

geht dann von den lokalen Koordinaten in die Mannigfaltigkeit. Es ist

$$U \subset \mathbb{R}^k, \quad V \subset \mathbb{R}^n.$$

Wir betrachten nun eine Abbildung

$$\vec{\varphi} : \mathbb{R}^k \rightarrow \mathbb{R}^n,$$

mit der Eigenschaft

$$\vec{\varphi}|_U = \varphi.$$

In Komponenten haben wir

$$\vec{\varphi}(\vec{x}) = \begin{pmatrix} \varphi_1(x_1, \dots, x_k) \\ \dots \\ \varphi_n(x_1, \dots, x_k) \end{pmatrix}.$$

Da φ bijektiv ist, gilt für die Funktionalmatrix

$$\text{Rank } D\vec{\varphi}(\vec{x}) = k, \quad \forall \vec{x} \in U.$$

Die Funktionalmatrix ist eine $n \times k$ -Matrix mit Rang k . Wir definieren nun wieder den Maßtensor durch

$$g_{j_1 j_2} = \sum_{i=1}^n \frac{\partial \varphi_i}{\partial x_{j_1}} \cdot \frac{\partial \varphi_i}{\partial x_{j_2}}$$

Der Maßtensor ist eine $k \times k$ -Matrix und es gilt wieder

$$g = (D\vec{\varphi})^T (D\vec{\varphi}).$$

Allerdings handelt es sich hier nun um die Multiplikation einer $k \times n$ -Matrix mit einer $n \times k$ -Matrix. Nehmen wir nun an, daß wir für die Mannigfaltigkeit M nur eine Karte benötigen. Dann definieren wir das Integral einer Funktion $f(y_1, \dots, y_n)$ über die Mannigfaltigkeit M durch

$$\int_M dS(\vec{y}) f(\vec{y}) = \int_U d^k x \sqrt{|g|} f(\vec{\varphi}(\vec{x})).$$

Auf der linken Seite bezeichnet $dS(\vec{y})$ das k -dimensionale Flächenelement (engl. "surface").

Bemerkungen: Ist $k = n$, so reduziert sich diese Definition auf die uns schon bekannte Transformationsformel

$$\int_M d^n y f(\vec{y}) = \int_U d^n x \sqrt{|g|} f(\vec{\varphi}(\vec{x})).$$

Ist hingegen $k = 1$, so stellt die Mannigfaltigkeit M eine Kurve dar, die in den \mathbb{R}^n eingebettet ist. Sei $I \subset M$ ein kompaktes Gebiet dieser Kurve und $[a, b] = \vec{\varphi}^{-1}(I) \subset \mathbb{R}$. Betrachten wir weiter den Spezialfall $f(\vec{y}) = 1$, so erhalten wir

$$\text{vol}(I) = \int_I dS(\vec{y}) = \int_a^b dx \sqrt{|g|} = \int_a^b dx \sqrt{\sum_{i=1}^n \left(\frac{\partial \varphi_i}{\partial x}\right)^2}.$$

Dies ist nichts anderes als die uns schon bekannte Bogenlänge der Kurve.

Als eine Anwendung betrachten wir die Integration auf einer Kugeloberfläche. Sei hierzu

$$M = \{ \vec{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = R^2 \}.$$

Sei U das Gebiet

$$U = [0, \pi] \times [0, 2\pi]$$

und $\vec{\varphi}$ die Abbildung

$$\vec{\varphi} : U \rightarrow \mathbb{R}^3,$$

$$\vec{\varphi}(\theta, \phi) = \begin{pmatrix} R \sin \theta \cos \phi \\ R \sin \theta \sin \phi \\ R \cos \theta \end{pmatrix}.$$

Es ist

$$D\vec{\varphi} = \begin{pmatrix} R \cos \theta \cos \phi & -R \sin \theta \sin \phi \\ R \cos \theta \sin \phi & R \sin \theta \cos \phi \\ -R \sin \theta & 0 \end{pmatrix}$$

und

$$g = (D\vec{\varphi})^T (D\vec{\varphi}) = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{pmatrix}.$$

Somit ergibt sich $\sqrt{|g|}$ zu

$$\sqrt{|g|} = \sqrt{R^4 \sin^2 \theta} = R^2 \sin \theta.$$

Wir wollen nun die Funktion

$$F(\theta, \phi) = \frac{3}{4\pi} \cos^2 \theta$$

über die Kugeloberfläche integrieren:

$$\int_M dS(\vec{y}) F = \int_0^\pi d\theta \int_0^{2\pi} d\phi (R^2 \sin \theta) \left(\frac{3}{4\pi} \cos^2 \theta \right) = \frac{3}{2} R^2 \int_0^\pi \sin \theta \cos^2 \theta = \frac{3}{2} R^2 \left(\frac{2}{3} \right) = R^2.$$

Bemerkung: Auf der Kugeloberfläche bezeichnet man die dort definierten Funktionen

$$Y_{lm}(\theta, \phi) = (-1)^{\frac{m+|m|}{2}} \sqrt{\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos \theta) e^{im\phi}$$

als Kugelflächenfunktionen. Hierbei ist $-l \leq m \leq l$ und die P_l^m stellen die assoziierten Legendre-Funktionen dar:

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x).$$

P_l ist das l -te Legendre-Polynom. Für die Kugelflächenfunktionen mit negativen m gilt

$$Y_{l,-m}(\theta, \phi) = (-1)^m Y_{lm}(\theta, \phi)^*.$$

Die ersten Kugelflächenfunktionen sind explizit gegeben durch

$$\begin{aligned} Y_{00} &= \frac{1}{\sqrt{4\pi}}, \\ Y_{10} &= \sqrt{\frac{3}{4\pi}} \cos \theta, \\ Y_{11} &= -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi}, \\ Y_{20} &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1), \\ Y_{21} &= -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi}, \\ Y_{22} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{2i\phi}. \end{aligned}$$

Auf der Oberfläche der Einheitskugel sind die Kugelflächenfunktionen orthogonal, d.h. es gilt

$$\int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi Y_{lm}(\theta, \phi)^* Y_{l'm'}(\theta, \phi) = \delta_{ll'} \delta_{mm'}.$$

10.8.3 Die Sätze von Gauß und Stokes

In diesem Abschnitt wollen wir zwei Integralsätze näher betrachten. Dies ist zum einen der Satz von Gauß, der es erlaubt das Integral über ein dreidimensionales Raumgebiet der Divergenz eines Vektorfeldes durch ein Flächenintegral über das Vektorfeld selbst zu ersetzen. Als zweites wollen wir dann den Satz von Stokes betrachten, der es erlaubt ein Flächenintegral über die Rotation eines Vektorfeldes durch ein Linienintegral über das Vektorfeld selbst auszudrücken. Diese beiden Sätze sind der Physik der Elektrodynamik sehr wichtig.

Sei $A \subset \mathbb{R}^n$ eine kompakte Teilmenge (der Dimension n) mit glattem Rand. Der Rand von A sei mit ∂A bezeichnet. Da A n -dimensional ist, ist der Rand eine $(n-1)$ -dimensionale Mannigfaltigkeit. Allgemein bezeichnet man $(n-1)$ -dimensionale Untermannigfaltigkeiten des \mathbb{R}^n als **Hyperflächen**. (Diese Bezeichnung leitet sich von dem Fall $n=3$ ab: Im \mathbb{R}^3 sind die zweidimensionalen Untermannigfaltigkeiten Flächen.) Sei

$$\hat{n} : \partial A \rightarrow \mathbb{R}^n$$

das äußere Einheitsnormalenvektorfeld auf dem Rand, d.h. in jedem Punkt \vec{x} von ∂A ist $\hat{n}(\vec{x})$ orthogonal zu jedem Tangentialvektor von ∂A am Punkte \vec{x} . $\hat{n}(\vec{x})$ zeigt vom Rand ∂A nach aussen und ist auf Eins normiert:

$$\hat{n}(\vec{x}) \cdot \hat{n}(\vec{x}) = 1.$$

Betrachten wir nun ein stetig differenzierbares Vektorfeld

$$\vec{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Der **Gaußsche Satz** lautet:

$$\int_A d^n x \operatorname{div} \vec{F}(\vec{x}) = \int_{\partial A} dS(\vec{x}) \hat{n}(\vec{x}) \cdot \vec{F}(\vec{x}).$$

Der Gaußsche Satz besagt, daß das Integral über ein n -dimensionales Gebiet der Divergenz eines Vektorfeldes gleich dem $(n - 1)$ -dimensionalen Integral über dem Rand des Gebietes der Normalkomponente des Vektorfeldes bezüglich des Randes ist.

Bemerkung: Der Satz gilt auch, falls ∂A nur stückweise glatt ist.

Beispiel: Wir betrachten ein Beispiel im \mathbb{R}^3 . Das Vektorfeld sei

$$\begin{aligned} \vec{F} &: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \\ \vec{F}(\vec{x}) &= \frac{\vec{x}}{|\vec{x}|^2}. \end{aligned}$$

Es ist

$$\operatorname{div} \vec{F} = \frac{1}{|\vec{x}|^2}.$$

Sei A die Kugel um den Ursprung mit Radius R . Dann ist

$$\int_A d^n x \operatorname{div} \vec{F}(\vec{x}) = \int_0^R dr \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi r^2 \sin \vartheta \frac{1}{r^2} = 4\pi \int_0^R dr = 4\pi R.$$

Andererseits können das Integral über die Kugeloberfläche betrachten. Es ist

$$dS(\vec{x}) = R^2 \sin \vartheta d\vartheta d\varphi,$$

sowie

$$\vec{F}(\vec{x}) \cdot \hat{n}(\vec{x}) = \frac{1}{R}.$$

Somit erhalten wir

$$\int_{\partial A} dS(\vec{x}) \hat{n}(\vec{x}) \cdot \vec{F}(\vec{x}) = \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi R^2 \sin \vartheta \frac{1}{R} = 4\pi R.$$

Hiermit haben wir den Gaußschen Satz an einem Beispiel verifiziert.

Wir betrachten nun den \mathbb{R}^3 . Sei B eine kompakte Fläche im \mathbb{R}^3 mit glattem Rand. Den Rand bezeichnen wir mit ∂B . Weiter sei

$$\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

ein stetig differenzierbares Vektorfeld.

Der **Satz von Stokes** besagt

$$\int_B dS(\vec{x}) \hat{n}(\vec{x}) \cdot \operatorname{rot} \vec{F}(\vec{x}) = \int_{\partial B} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{F}(\vec{x})$$

Hierbei ist $\hat{t}(\vec{x})$ der Einheits tangentialvektor im Punkte \vec{x} an den Rand ∂B . Diese Orientierungen sind so korreliert, daß die geschlossene Kurve ∂B die in Richtung von $\hat{t}(\vec{x})$ durchlaufen wird und \hat{n} eine Rechtsschraube bilden.

Bemerkung: Der Satz gilt auch, falls ∂B nur stückweise glatt ist.

Auch hier betrachten wir ein Beispiel: Es sei

$$\vec{F} = \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}$$

und B das Einheitsquadrat in der (x_1, x_2) -Ebene (d.h. das Quadrat mit den Ecken $(0, 0, 0)$, $(1, 0, 0)$, $(1, 1, 0)$ und $(0, 1, 0)$). Der Normalenvektor der Fläche B ist immer $\hat{n}(\vec{x}) = (0, 0, 1)^T$ und in diesem Fall unabhängig von \vec{x} . Weiter ist

$$\operatorname{rot} \vec{F}(\vec{x}) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

und somit

$$\int_B dS(\vec{x}) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = 2 \int_0^1 dx_1 \int_0^1 dx_2 = 2.$$

Wir betrachten nun die rechte Seite des Satzes von Stokes. Wir teilen den Weg ∂B in vier Teilabschnitte

$$\partial B = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$$

auf, wobei γ_1 von $(0, 0, 0)$ nach $(1, 0, 0)$ verläuft, γ_2 von $(1, 0, 0)$ nach $(1, 1, 0)$, γ_3 von $(1, 1, 0)$ nach $(0, 1, 0)$ und γ_4 von $(0, 1, 0)$ nach $(0, 0, 0)$. Betrachten wir zunächst γ_1 . Entlang dieses Weges gilt $x_2 = 0 = \text{const}$. Der Einheitstangentenvektor ist $\hat{t}(\vec{x}) = (1, 0, 0)^T$ und unabhängig von \vec{x} . Wir erhalten

$$\int_{\gamma_1} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{F}(\vec{x}) = \int_0^1 dx_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ x_1 \\ 0 \end{pmatrix} = \int_0^1 dx_1 \cdot 0 = 0.$$

Entlang des Weges γ_2 ist $x_1 = 1 = \text{const}$ und $\hat{t}(\vec{x}) = (0, 1, 0)^T$. Somit

$$\int_{\gamma_2} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{F}(\vec{x}) = \int_0^1 dx_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -x_2 \\ 1 \\ 0 \end{pmatrix} = \int_0^1 dx_2 \cdot 1 = 1.$$

Entlang des Weges γ_3 ist $x_2 = 1 = \text{const}$ und $\hat{t}(\vec{x}) = (-1, 0, 0)^T$. Wir setzen $x_1 = 1 - \lambda$, so daß $\lambda = 0$ dem Anfangspunkt und $\lambda = 1$ dem Endpunkt entspricht. Dies ergibt

$$\int_{\gamma_3} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{F}(\vec{x}) = \int_0^1 d\lambda \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 - \lambda \\ 0 \end{pmatrix} = \int_0^1 d\lambda \cdot 1 = 1.$$

Entlang des Weges γ_4 ist $x_1 = 0 = \text{const}$ und $\hat{t}(\vec{x}) = (0, -1, 0)^T$. Wir setzen $x_2 = 1 - \lambda$ und erhalten

$$\int_{\gamma_4} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{F}(\vec{x}) = \int_0^1 d\lambda \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \lambda - 1 \\ 0 \\ 0 \end{pmatrix} = \int_0^1 d\lambda \cdot 0 = 0.$$

Zählen wir alle vier Beiträge zusammen, so erhalten wir

$$\int_{\partial B} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{F}(\vec{x}) = 2,$$

in Übereinstimmung mit der linken Seite des Satzes von Stokes.

Bemerkung: Da die Rotation eines Vektorfeldes involviert ist, gilt der Satz von Stokes in dieser Form nur im \mathbb{R}^3 .

Wir führen noch zwei Sätze von Green an: Seien f und g Funktionen auf dem \mathbb{R}^n und A eine kompakte Teilmenge des \mathbb{R}^n mit glattem Rand. Der **erste Satz von Green** besagt, daß

$$\int_A d^n x \left[f(\vec{x}) \Delta g(\vec{x}) + (\vec{\nabla} f(\vec{x})) \cdot (\vec{\nabla} g(\vec{x})) \right] = \int_{\partial A} dS(\vec{x}) f(\vec{x}) \hat{n}(\vec{x}) \cdot (\vec{\nabla} g(\vec{x})).$$

Dieser Satz ist eine direkte Anwendung des Gauß'schen Satzes wenn man dort das Vektorfeld

$$\vec{F}(\vec{x}) = f(\vec{x}) \vec{\nabla} g(\vec{x})$$

einsetzt und die Produktregel für die Differentiation verwendet,

$$\vec{\nabla} \cdot (f(\vec{x}) \vec{\nabla} g(\vec{x})) = (\vec{\nabla} f(\vec{x})) \cdot (\vec{\nabla} g(\vec{x})) + f(\vec{x}) \Delta g(\vec{x}).$$

Der **zweite Satz von Green** lautet:

$$\int_A d^n x [f(\vec{x}) \Delta g(\vec{x}) - g(\vec{x}) \Delta f(\vec{x})] = \int_{\partial A} dS(\vec{x}) [f(\vec{x}) \hat{n} \cdot \vec{\nabla} g(\vec{x}) - g(\vec{x}) \hat{n} \cdot \vec{\nabla} f(\vec{x})].$$

10.8.4 Differentialformen*

Wir haben im vorherigen Abschnitt die Integralsätze von Gauß und Stokes kennengelernt. Diese Integralsätze sind Spezialfälle eines allgemeineren Satzes, der als der allgemeine Stokessche Integralsatz bekannt ist. Mit Hilfe einer geeigneten Notation läßt sich der allgemeine Stokessche Integralsatz kompakt als

$$\int_A d\omega = \int_{\partial A} \omega$$

darstellen. In dieser Formel stell ω eine Differentialform dar, die wir nun einführen wollen.

Betrachten wir zunächst ein-dimensionale Integrale, welche wir durch einen Grenzwertprozeß definieren können:

$$\int_{\mathbb{R}} dx f(x) = \lim \sum_j f(x_j) \Delta x_j$$

Ebenso zwei-dimensionale Integrale:

$$\int_{\mathbb{R}^2} dx dy g(x, y) = \lim \sum_j \sum_k g(x_j, y_k) \Delta x_j \Delta y_k$$

Anstelle der Funktionen $f(x)$ und $g(x, y)$ werden wir nun neue Objekte

$$f(x) dx, \quad g(x, y) dx \wedge dy$$

einführen, die über ein Gebiet der entsprechenden Dimension integriert werden können, einführen. Der Grund hierfür sind die übersichtlicheren Transformationseigenschaften.

Tangentialvektoren

Sei $I \subset \mathbb{R}$ ein Intervall und $\vec{\gamma} : I \rightarrow M \subset \mathbb{R}^n$ eine differenzierbare Kurve. Man bezeichnet

$$\left. \frac{d}{dt} \vec{\gamma}(t) \right|_{t_0} \in \mathbb{R}^n$$

als Tangentialvektor an M im Punkte $\vec{\gamma}(t_0)$. Die Gesamtheit aller Tangentialvektoren an M im Punkte $p = \vec{\gamma}(t_0)$ wird mit $T_p M$ bezeichnet. $T_p M$ ist ein Vektorraum. Die Dimension des Tangentialraumes ist gleich der Dimension der Mannigfaltigkeit.

Wir bezeichnen mit $T_p^* M$ den dualen Vektorraum von $T_p M$, d.h. die Menge aller Linearformen

$$\phi : T_p M \rightarrow \mathbb{R}.$$

Die Elemente von $\phi \in T_p^* M$ heißen auch Kotangentialvektoren.

Differentialformen erster Ordnung

Eine Differentialform erster Ordnung ist eine Abbildung

$$\omega : M \rightarrow \bigcup_p T_p^* M$$

mit $\omega(p) \in T_p^* M$. Die Differentialform ω ordnet also jedem Punkt $p \in M$ einen Kotangentialvektor $\omega(p) \in T_p^* M$ zu. Wir bezeichnen den Wert von $\omega(p)$ auf dem Tangentialvektor $v \in T_p M$ mit

$$\langle \omega(p), v \rangle$$

Wir bezeichnen mit $\vec{e}_1, \dots, \vec{e}_n$ die Standardorthonormalbasis des \mathbb{R}^n . Wir können jeden Tangentialraum $T_p M$ als einen Unterraum des \mathbb{R}^n betrachten. (Ist die Mannigfaltigkeit M der \mathbb{R}^n selbst, so ist auch der Tangentialraum an jedem Punkt der \mathbb{R}^n .) Wir führen nun eine Basis des Kotangentialraumes ein: Wir definieren die Differentiale

$$dx_1, \dots, dx_n$$

durch

$$\langle dx_i, \vec{e}_j \rangle = \delta_{ij}.$$

Mit Hilfe dieser Basis des Kotangentialraumes läßt sich jede Differentialform erster Ordnung schreiben als

$$\omega = \sum_{i=1}^n f_i(\vec{x}) dx_i.$$

Eine Differentialform erster Ordnung kann entlang einer Kurve integriert werden. Sei ω eine Differentialform erster Ordnung und $\vec{\gamma} : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve. Wir verwenden die Bezeichnung $\gamma = \vec{\gamma}([a, b])$ für das Bild von $\vec{\gamma}$. Dann wird das Integral von ω über γ definiert als

$$\int_{\gamma} \omega = \int_a^b \langle \omega(\vec{\gamma}(t)), \vec{\gamma}'(t) \rangle dt.$$

Mit Hilfe der Koordinatendarstellung findet man für die rechte Seite

$$\int_a^b \langle \omega(\vec{\gamma}(t)), \vec{\gamma}'(t) \rangle dt = \int_a^b dt \sum_{i=1}^n f_i(\vec{\gamma}(t)) \frac{d\gamma_i(t)}{dt}.$$

Beispiel: Sei

$$\omega = 3dx_1 + (5 + x_1)dx_2 + x_3dx_3$$

und

$$\vec{\gamma} : [0, 1] \rightarrow \mathbb{R}^3, \\ \vec{\gamma}(t) = \begin{pmatrix} t \\ t^2 \\ 1+t \end{pmatrix}.$$

Dann ist

$$\begin{aligned} \int_{\gamma} \omega &= \int_0^1 dt \left[3 \frac{d}{dt} t + (5+t) \frac{d}{dt} t^2 + (1+t) \frac{d}{dt} (1+t) \right] \\ &= \int_0^1 dt [3 + 2t(5+t) + (1+t)] = \int_0^1 dt [4 + 11t + 2t^2] \\ &= 4 + \frac{11}{2} + \frac{2}{3} = \frac{61}{6}. \end{aligned}$$

Differentialformen höherer Ordnung

Wir haben bisher gesehen, daß Differentialformen erster Ordnung über Kurven integriert werden können. Wir möchten nun eine Verallgemeinerung, die eine Integration über höher dimensionale Gebiete erlaubt. Wir beginnen mit der Definition des Dachproduktes von Linearformen: Seien $\omega_1, \dots, \omega_K \in V^*$ Linearformen, also Abbildungen

$$\omega_j : V \rightarrow \mathbb{R}.$$

Dann wird die Abbildung

$$\omega_1 \wedge \dots \wedge \omega_k : V^k \rightarrow \mathbb{R}$$

definiert durch

$$(\omega_1 \wedge \dots \wedge \omega_k)(v_1, \dots, v_k) = \det \begin{pmatrix} \langle \omega_1, v_1 \rangle & \dots & \langle \omega_1, v_k \rangle \\ \dots & \dots & \dots \\ \langle \omega_k, v_1 \rangle & \dots & \langle \omega_k, v_k \rangle \end{pmatrix}$$

Eigenschaften des Dachproduktes:

- Das Dachprodukt ist linear in jedem Argument:

$$\omega_1 \wedge \dots \wedge (a\omega'_i + b\omega''_i) \wedge \dots \wedge \omega_k = a(\omega_1 \wedge \dots \wedge \omega'_i \wedge \dots \wedge \omega_k) + b(\omega_1 \wedge \dots \wedge \omega''_i \wedge \dots \wedge \omega_k)$$

- Das Dachprodukt ist alternierend:

$$\omega_{\sigma(1)} \wedge \dots \wedge \omega_{\sigma(k)} = \text{sign}(\sigma) \cdot \omega_1 \wedge \dots \wedge \omega_k$$

Insbesondere ist

$$\begin{aligned} \omega_1 \wedge \omega_2 &= -\omega_2 \wedge \omega_1, \\ \omega_1 \wedge \omega_1 &= 0. \end{aligned}$$

Wir bezeichnen die Menge aller alternierenden linearen k -Formen auf V mit

$$\wedge^k V^*$$

Definition: Eine **Differentialform der Ordnung** k (oder kurz k -Form) ist eine Abbildung

$$\omega : M \rightarrow \bigcup_p \wedge^k T_p^* M$$

mit $\omega(p) \in \wedge^k T_p^* M$. Für $k = 1$ stimmt diese Definition mit der schon bekannten Definition für Differentialformen erster Ordnung überein. Eine Differentialform nullter Ordnung ist eine reellwertige Funktion.

Koordinatendarstellung von Differentialformen k -ter Ordnung:

$$\omega = \frac{1}{k!} \sum_{i_1, \dots, i_k} f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} = \sum_{i_1 < \dots < i_k} f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

Ableitungen von Differentialformen: Sei

$$\omega = \sum_{i_1 < \dots < i_k} f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

eine k -Form. Dann bezeichnet $d\omega$ die $(k + 1)$ -Form

$$d\omega = \sum_{i_1 < \dots < i_k} df_{i_1 \dots i_k} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k},$$

wobei

$$df_{i_1 \dots i_k} = \sum_{i=1}^n \frac{\partial f_{i_1 \dots i_k}}{\partial x_i} dx_i.$$

Bemerkung: Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Dann bezeichnet man mit df das **totale Differential** von f :

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i.$$

Das totale Differential df einer Funktion f ist eine Einsform.

Rechenregeln: Seien ω und ω' zwei k -Formen und σ eine l -Form. Dann gilt

$$\omega \wedge \sigma = (-1)^{kl} \sigma \wedge \omega.$$

Weiter gilt:

$$\begin{aligned} d(a\omega + b\omega') &= ad\omega + bd\omega', \\ d(\omega \wedge \sigma) &= (d\omega) \wedge \sigma + (-1)^k \omega \wedge (d\sigma), \\ d(d\omega) &= 0. \end{aligned}$$

Zurückziehen von Differentialformen: Sei $U \subset \mathbb{R}^n$ und

$$\omega = \frac{1}{k!} \sum f_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

eine k -Form in U . Weiter sei eine offene Menge $V \subset \mathbb{R}^m$ und eine stetig differenzierbare Abbildung

$$\varphi = (\varphi_1, \dots, \varphi_n) : V \rightarrow U$$

vorgegeben. Dann definiert man eine k -Form $\varphi^*\omega$ in V durch

$$\varphi^*\omega = \frac{1}{k!} \sum (f_{i_1 \dots i_k} \circ \varphi) d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_k}.$$

Hierbei ist $d\varphi_{i_j}$ das totale Differential von φ_{i_j} , also

$$d\varphi_{i_j} = \sum_{l=1}^m \frac{\partial \varphi_{i_j}}{\partial x_l} dx_l.$$

Wir definieren nun die Integration von k -Formen über k -dimensionale Gebiete. Wir beginnen mit dem Fall $k = n$. Sei ω also eine n -Form. In diesem Fall läßt sich ω schreiben als

$$\omega = f(\vec{x}) dx_1 \wedge \dots \wedge dx_n.$$

Sei A eine kompakte Teilmenge des \mathbb{R}^n . Wir setzen

$$\int_A \omega = \int_A d^n x f(\vec{x}).$$

Betrachten wir nun den Fall $k < n$: Sei M eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n und A eine kompakte Teilmenge von M , ebenfalls mit der Dimension k . Sei ferner ω eine k -Form. Wir nehmen an, daß es eine lokale Karte

$$\varphi : \mathbb{R}^k \rightarrow U$$

von M gibt, so daß $A \in U$. Wir möchten nun das Integral von ω über A definieren. Hierzu nutzen wir aus, daß wir die k -Form ω mit Hilfe von φ auf den \mathbb{R}^k zurückziehen können. Auf dem \mathbb{R}^k erhalten wir dann eine k -Form $\varphi^* \omega$, die wir über dem kompakten Gebiet $\varphi^{-1}(A)$ wie oben integrieren können. Wir setzten also

$$\int_A \omega = \int_{\varphi^{-1}(A)} \varphi^* \omega.$$

Für eine Eins-Form reduziert sich diese Definition auf die uns schon bekannte Vorschrift zur Integration einer Eins-Form. In der obigen Definition zur Integration einer k -Form haben wir vorausgesetzt, daß M orientierbar ist und φ die orientierungstreu ist. Diese Begriffe sollen noch kurz erläutert werden: Wir bezeichnen eine Abbildung

$$\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^k$$

als **orientierungstreu**, falls für die Determinante der Funktionalmatrix stets

$$\det D\varphi > 0$$

gilt. Wir erinnern uns, daß eine Mannigfaltigkeit per Definition einen Atlas von Karten besitzt. Seien $\varphi_i : \mathbb{R}^k \rightarrow M$ und $\varphi_j : \mathbb{R}^k \rightarrow M$ zwei Karten. Auf dem Überlapp der Karten gibt es eine Parametertransformation

$$\varphi_i^{-1} \circ \varphi_j : \mathbb{R}^k \rightarrow \mathbb{R}^k.$$

Gibt es einen Atlas, so daß für zwei beliebige überlappende Karten des Atlases die Parametertransformation stets orientierungstreu ist, so nennt man die Mannigfaltigkeit **orientierbar**. Zum besseren Verständnis sei ein Gegenbeispiel angegeben: Das Möbiusband ist nicht orientierbar. Wir bezeichnen eine Mannigfaltigkeit als **positiv orientiert**, falls sie die gleiche Orientierung

wie der \mathbb{R}^k bezüglich der Standardbasis $\vec{e}_1, \dots, \vec{e}_k$ besitzt.

Beispiel für die Integration einer Differentialform: Im \mathbb{R}^3 sei die Differentialform

$$\omega = 3x_3 dx_2 \wedge dx_3 + (x_1^2 + x_2^2) dx_3 \wedge dx_1 + x_1 x_3 dx_1 \wedge dx_2$$

gegeben. Sei M die folgende zweidimensionale Untermannigfaltigkeit

$$M = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = x_1 x_2\}$$

und A die folgende kompakte Teilmenge von M :

$$A = \{(x_1, x_2, x_3) \in M : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}.$$

Wir möchten

$$\int_A \omega$$

berechnen. Wir wählen zunächst eine lokale Karte von M

$$\begin{aligned} \varphi : \mathbb{R}^2 &\rightarrow M, \\ (y_1, y_2) &\rightarrow (y_1, y_2, y_1 y_2) \end{aligned}$$

Die einzelnen Koordinatenabbildungen sind

$$\varphi_1 = y_1, \quad \varphi_2 = y_2, \quad \varphi_3 = y_1 y_2,$$

und daher

$$d\varphi_1 = dy_1, \quad d\varphi_2 = dy_2, \quad d\varphi_3 = y_2 dy_1 + y_1 dy_2.$$

Somit

$$\begin{aligned} \int_A \omega &= \int_{\varphi^{-1}(A)} (\varphi)^* \omega = \\ &= \int_{\varphi^{-1}(A)} 3y_1 y_2 dy_2 \wedge (y_2 dy_1 + y_1 dy_2) + (y_1^2 + y_2^2) (y_2 dy_1 + y_1 dy_2) \wedge dy_1 + y_1 (y_1 y_2) dy_1 \wedge dy_2 \\ &= \int_{\varphi^{-1}(A)} (y_1^2 y_2 - 4y_1 y_2^2 - y_1^3) dy_1 \wedge dy_2 = \int_0^1 dy_1 \int_0^1 dy_2 (y_1^2 y_2 - 4y_1 y_2^2 - y_1^3) = -\frac{3}{4}. \end{aligned}$$

Wir können nun den **allgemeinen Stokesschen Integralsatz** formulieren: Sei M eine orientierbare k -dimensionale Mannigfaltigkeit, A eine kompakte Teilmenge von M mit glattem Rand und ω eine stetig differenzierbare $(k-1)$ -Form auf M . Dann gilt

$$\int_A d\omega = \int_{\partial A} \omega,$$

wobei der Rand ∂A so orientiert ist, daß der äußere Normalenvektor plus die $(k - 1)$ Tangentialvektoren die gleiche Orientierung wie M besitzen, d.h.

$$\det(\hat{n}, \hat{t}_1, \dots, \hat{t}_{k-1}) > 0.$$

Dieser Satz erlaubt es uns, das Integral der Ableitung einer Differentialform über ein kompaktes Gebiet durch das Integral der Differentialform über dem Rand des Gebietes auszudrücken.

Bemerkung: Der Satz gilt auch, falls ∂A nur stückweise glatt ist.

Spezialfälle dieses allgemeinen Satzes sind der Hauptsatz der Integral- und Differentialrechnung, der Satz von Gauß und der klassische Satz von Stokes in drei Dimensionen. Wir wollen auf diese Spezialfälle noch etwas genauer eingehen und beginnen mit dem Fall, in dem sich der allgemeine Satz auf den Satz von Gauß reduziert:

Sei A ein kompaktes n -dimensionales Gebiet und ω eine $(n - 1)$ -Form. Wir können ω wie folgt schreiben:

$$\omega = \sum_{j=1}^n (-1)^{j-1} f_j(\vec{x}) dx_1 \wedge \dots \wedge \widehat{dx}_j \wedge \dots \wedge dx_n,$$

wobei der Hut bedeuten soll, daß das entsprechende Differential ausgelassen wird. Dann ist

$$\begin{aligned} d\omega &= \sum_{j=1}^n \frac{\partial f_j(\vec{x})}{\partial x_j} dx_1 \wedge \dots \wedge dx_j \wedge \dots \wedge dx_n \\ &= \left(\sum_{j=1}^n \frac{\partial f_j(\vec{x})}{\partial x_j} \right) dx_1 \wedge \dots \wedge dx_n. \end{aligned}$$

Setzt man nun

$$\begin{aligned} \vec{f} &: \mathbb{R}^n \rightarrow \mathbb{R}^n, \\ \vec{f}(\vec{x}) &= \begin{pmatrix} f_1(\vec{x}) \\ \dots \\ f_n(\vec{x}) \end{pmatrix}, \end{aligned}$$

so ist

$$\left(\sum_{j=1}^n \frac{\partial f_j(\vec{x})}{\partial x_j} \right) = \operatorname{div} \vec{f}(\vec{x}),$$

und die linke Seite somit

$$\int_A d\omega = \int_A (\operatorname{div} \vec{f}(\vec{x})) dx_1 \wedge \dots \wedge dx_n = \int_A d^n x \operatorname{div} \vec{f}(\vec{x}).$$

Andererseits zeigt man durch eine etwas längere Rechnung

$$\int_{\partial A} \omega = \int_{\partial A} dS(\vec{x}) \vec{f}(\vec{x}) \cdot \hat{n}(\vec{x}).$$

Wir erhalten somit den Satz von Gauß.

Der Fall $n = 1$ und $k = 0$ verdient besondere Betrachtung. In diesem Fall ist die Differentialform ω eine Funktion einer Variablen. Wir schreiben

$$\omega = f(x).$$

Weiter haben wir

$$d\omega = f'(x)dx.$$

Wir betrachten ein kompaktes Intervall $A = [a, b] \subset \mathbb{R}$. In diesem Fall reduziert sich der Satz auf den Hauptsatz der Integral- und Differentialrechnung:

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

Wir betrachten als weiteren Spezialfall den Fall $n = 3$ und $k = 2$. In diesem Fall ist ω eine Eins-Form, die wir als

$$\omega = \sum_{j=1}^3 f_j(\vec{x}) dx_j$$

schreiben können. Wir erhalten nun

$$d\omega = \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial f_j(\vec{x})}{\partial x_i} dx_i \wedge dx_j.$$

Das Dachprodukt ist antisymmetrisch, daher ist $dx_i \wedge dx_i = 0$ und $dx_i \wedge dx_j = -dx_j \wedge dx_i$. Man erhält also

$$\begin{aligned} d\omega &= \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) dx_1 \wedge dx_2 + \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) dx_3 \wedge dx_1 \\ &= (\text{rot } \vec{f}(\vec{x})) \cdot \hat{n}(\vec{x}) dS(\vec{x}). \end{aligned}$$

Hierbei haben wir $\vec{f} = (f_1, f_2, f_3)$ gesetzt. Für die rechte Seite findet man mit Hilfe der Definition des Integrals für eine Eins-Form

$$\int_{\partial B} \omega = \int_{\partial B} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{f}(\vec{x}).$$

Somit erhält man in diesem Fall den ursprünglichen Satz von Stokes:

$$\int_B dS(\vec{x}) \hat{n}(\vec{x}) \cdot \operatorname{rot} \vec{f}(\vec{x}) = \int_{\partial B} ds(\vec{x}) \hat{t}(\vec{x}) \cdot \vec{f}(\vec{x}).$$

Betrachten wir ein Beispiel: Es sei $M = \mathbb{R}^3$ und somit $n = 3$. Weiter sei B das Einheitsquadrat in der (x_1, x_2) -Ebene, dies ist eine zweidimensionale Fläche und daher $k = 2$. Wir betrachten die Eins-Form

$$\omega = -x_2 dx_1 + x_1 dx_2.$$

Es ist

$$d\omega = -dx_2 \wedge dx_1 + dx_1 \wedge dx_2 = 2dx_1 \wedge dx_2.$$

Schreiben wir

$$\omega = f_1 dx_1 + f_2 dx_2 + f_3 dx_3$$

so ist

$$\vec{f} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}$$

und wir erhalten das Beispiel, das wir bereits im Zusammenhang mit der ursprünglichen Version des Satzes von Stokes diskutiert haben.

10.9 Variationsrechnung

In diesem Abschnitt beschäftigen wir uns mit den Eulerschen Differentialgleichungen der Variationsrechnung. Diese spielen in der theoretischen Physik eine wichtige Rolle.

Zunächst betrachten wir einige Hilfssätze. Wir betrachten eine Funktion $f(t, x_1, \dots, x_n)$ von $(n+1)$ Variablen. Diese Funktion integrieren wir über die Variable t von a bis b . Dies liefert eine neue Funktion $g(x_1, \dots, x_n)$, die nun nur noch von den n Variablen x_1 bis x_n abhängt. Die Hilfssätze liefern Aussagen über die Stetigkeit und Differenzierbarkeit der Funktion g .

Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $U \subset \mathbb{R}^n$. Ferner sei

$$f : [a, b] \times U \rightarrow \mathbb{R}$$

eine stetige Funktion. Dann ist die Funktion

$$g : U \rightarrow \mathbb{R}$$

$$g(\vec{x}) = \int_a^b dt f(t, \vec{x})$$

stetig.

Ist darüberhinaus $f(t, x_1, \dots, x_n)$ nach den Variablen x_1, \dots, x_n stetig partiell differenzierbar, dann ist auch $g(x_1, \dots, x_n)$ stetig partiell differenzierbar und es gilt

$$\frac{\partial}{\partial x_i} \int_a^b dt f(t, x_1, \dots, x_n) = \int_a^b dt \frac{\partial}{\partial x_i} f(t, x_1, \dots, x_n).$$

Man darf also “unter dem Integralzeichen differenzieren”.

Wir kommen nun zu den Eulerschen Differentialgleichungen der Variationsrechnung: Sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall und

$$L : I \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \\ (t, \vec{q}, \dot{\vec{q}}) \rightarrow L(t, \vec{q}, \dot{\vec{q}})$$

eine zweimal stetig partiell differenzierbare Funktion von $(2n + 1)$ Variablen. (Hier betrachten wir die Variablen q_i und \dot{q}_i als unabhängig, die Schreibweise wurde nur mit Hinblick auf die Anwendung in der Physik so gewählt.) Zu vorgegebenen Vektoren $\vec{q}_a \in \mathbb{R}^n$ und $\vec{q}_b \in \mathbb{R}^n$ betrachten wir nun die Menge K aller zweimal stetig differenzierbaren Kurven

$$\vec{\varphi} : I \rightarrow \mathbb{R}^n$$

mit der Eigenschaft

$$\vec{\varphi}(a) = \vec{q}_a, \quad \vec{\varphi}(b) = \vec{q}_b.$$

Wir suchen nun diejenige Kurve $\vec{\varphi} \in K$, die das Funktional

$$S[\vec{\varphi}] = \int_a^b dt L(t, \vec{\varphi}(t), \vec{\varphi}'(t))$$

minimiert.

Satz: Notwendige Bedingungen für das Vorliegen eines Minimums sind durch die Eulerschen Differentialgleichungen gegeben:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0, \quad 1 \leq i \leq n.$$

Nehmen wir an $\vec{\varphi}$ minimiert das Funktional S . Dann betrachten wir eine zweimal stetig differenzierbare Kurve $\vec{\lambda}(t)$ mit $\vec{\lambda}(a) = \vec{\lambda}(b) = \vec{0}$. Da $\vec{\varphi}$ das Funktional S minimiert, gilt

$$S[\vec{\varphi}] \leq S[\vec{\varphi} + \varepsilon \vec{\lambda}],$$

mit $\varepsilon \in \mathbb{R}$. Insbesondere gilt

$$\left. \frac{d}{d\varepsilon} S[\vec{\varphi} + \varepsilon \vec{\lambda}] \right|_{\varepsilon=0} = 0.$$

Nun ist aber

$$\begin{aligned} \frac{d}{d\varepsilon} S[\vec{\varphi} + \varepsilon \vec{\lambda}] &= \frac{d}{d\varepsilon} \int_a^b dt L(t, \vec{\varphi}(t) + \varepsilon \vec{\lambda}(t), \vec{\varphi}'(t) + \varepsilon \vec{\lambda}'(t)) \\ &= \int_a^b dt \sum_{i=1}^n \left[\frac{\partial L}{\partial q_i} \lambda_i(t) + \frac{\partial L}{\partial \dot{q}_i} \lambda_i'(t) \right]. \end{aligned}$$

Den zweiten Ausdruck können wir partiell integrieren:

$$\int_a^b dt \frac{\partial L}{\partial \dot{q}_i} \lambda_i'(t) = \left. \frac{\partial L}{\partial \dot{q}_i} \lambda_i(t) \right|_a^b - \int_a^b dt \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \lambda_i(t) = - \int_a^b dt \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \lambda_i(t).$$

Wegen $\lambda_i(a) = \lambda_i(b) = 0$ verschwinden die Randterme. Wir erhalten also

$$\sum_{i=1}^n \int_a^b dt \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \lambda_i(t) = 0.$$

Da dies für beliebige "Variationen" λ_i gilt, folgt die Behauptung

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0.$$

11 Partielle Differentialgleichungen

11.1 Distributionen

Die Verwendung von Distributionen ist oft recht nützlich, um Rechnungen zu vereinfachen. In diesem Abschnitt betrachten wir zunächst die formale Definition einer Distribution und gehen dann näher auf die Diracsche Deltadistribution ein.

Wir beginnen mit der Definition einer Distribution. Hierzu benötigen wir zunächst einmal einen Raum von Funktionen, die wir als **Testfunktionen** bezeichnen möchten. Als Testfunktionen betrachten wir alle beliebig oft differenzierbaren Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit kompakten Träger. Den dazugehörigen Funktionenraum bezeichnen wir als $C_c^\infty(\mathbb{R}^n)$.

Wir sagen eine Folge von Testfunktionen $f_j \in C_c^\infty(\mathbb{R}^n)$ konvergiert gegen eine Testfunktion $f \in C_c^\infty(\mathbb{R}^n)$, falls die folgenden zwei Bedingungen erfüllt sind:

1. Es gibt eine kompakte Menge $K \subset \mathbb{R}^n$, so daß

$$\text{supp}(f) \subset K \quad \text{und} \quad \text{supp}(f_j) \subset K \quad \text{für alle } j.$$

2. Für jeden Multiindex $(i_1, \dots, i_n) \in \mathbb{N}^n$ konvergiert die Folge der Ableitungen

$$\frac{\partial^{i_1+\dots+i_n}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} f_j(x_1, \dots, x_n)$$

gleichmäßig auf K gegen

$$\frac{\partial^{i_1+\dots+i_n}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} f(x_1, \dots, x_n).$$

In diesem Fall schreiben wir

$$f_j \longrightarrow f.$$

Bemerkung: Die Forderung, daß jede Ableitung der Funktionenfolge gleichmäßig gegen die entsprechende Ableitung der Grenzwertfunktion konvergiert, ist eine wesentlich stärkere Bedingung als nur die punktweise oder gleichmäßige Konvergenz.

Wir definieren nun eine Distribution als ein lineares Funktional

$$T : C_c^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}, \\ f \rightarrow T[f],$$

welches die folgende Bedingung erfüllt: Gilt für eine Folge von Testfunktionen $f_j \in C_c^\infty(\mathbb{R}^n)$ und eine Testfunktion $f \in C_c^\infty(\mathbb{R}^n)$ die Relation $f_j \longrightarrow f$, so folgt stets

$$\lim_{j \rightarrow \infty} T[f_j] = T[f].$$

Die Menge der Distributionen im \mathbb{R}^n bilden einen Vektorraum.

Wir betrachten zunächst ein einfaches Beispiel. Sei $g : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetige Funktion. Dann wird durch

$$T_g[f] = \int_{\mathbb{R}^n} d^n x g(\vec{x}) f(\vec{x})$$

eine Distribution definiert. Da auf diese Weise jede stetige Funktion g eine Distribution T_g definiert, unterscheidet man oft nicht zwischen g und T_g . Es soll jedoch betont werden, daß es sich bei g um eine Funktion und bei T_g um ein Funktional auf dem Raum der Testfunktionen handelt.

Als ein weiteres Beispiel betrachten wir die Diracsche Deltadistribution. Sie ist definiert durch

$$T_\delta[f] = f(\vec{0}).$$

In Analogie mit dem ersten Beispiel verwendet man für linke Seite die Schreibweise

$$T_\delta[f] = \int_{\mathbb{R}^n} d^n x \delta(\vec{x}) f(\vec{x}),$$

so daß sich mit Hilfe der Definition die Regel

$$\int_{\mathbb{R}^n} d^n x \delta(\vec{x}) f(\vec{x}) = f(\vec{0})$$

für alle Testfunktionen f ergibt.

Bemerkung 1: Distributionen sind lineare Funktionale auf dem Raum der Testfunktionen. Ein Produkt von Distributionen ist **nicht** definiert. Man sollte sich auch nicht durch die suggestive Schreibweise $\delta(\vec{x})$ dazu verleiten lassen, Distributionen miteinander zu multiplizieren.

Bemerkung 2: $\delta(\vec{x})$ ist keine Funktion, kann aber durch eine Folge von Funktionen dargestellt werden. Hierzu definieren wir den Konvergenzbegriff für Distributionen. Wir sagen, daß eine Folge von Distributionen T_j gegen eine Distribution T konvergiert, falls für alle Testfunktionen $f \in C_c^\infty(\mathbb{R}^n)$ gilt

$$\lim_{j \rightarrow \infty} T_j[f] = T[f],$$

d.h. die Folge der reellen Zahlen $T_j[f]$ konvergiert im herkömmlichen Sinne gegen $T[f]$.

Sei nun $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine integrierbare Funktion mit

$$\int_{\mathbb{R}^n} d^n x f(\vec{x}) = 1.$$

Dann setzt man

$$f_j(\vec{x}) = j^n f(j \cdot \vec{x}).$$

Diese Folge konvergiert dann im Sinne der Distributionen gegen $\delta(\vec{x})$. (Im Sinne einer Folge von Funktionen divergiert diese Folge natürlich für $\vec{x} = \vec{0}$.) Für $f(\vec{x})$ kann man beispielsweise

$$f(\vec{x}) = \frac{1}{(\sqrt{\pi})^n} e^{-|\vec{x}|^2}$$

verwenden. Mit $\varepsilon = 1/j$ ergibt sich dann die Deltadistribution als Grenzwert $\varepsilon \rightarrow 0$ der Folge

$$f_\varepsilon(\vec{x}) = \frac{1}{(\varepsilon \sqrt{\pi})^n} e^{-\frac{|\vec{x}|^2}{\varepsilon^2}}.$$

Völlig analog definiert man $\delta(\vec{x} - \vec{d})$, so daß gilt

$$\int_{\mathbb{R}^n} d^n x \delta(\vec{x} - \vec{d}) f(\vec{x}) = f(\vec{d}).$$

Die Ableitungen einer Distribution werden wie folgt definiert. Sei T_μ eine Distribution. Für die Anwendung auf eine Testfunktion schreiben wir wieder

$$\int_{\mathbb{R}^n} d^n x \mu(\vec{x}) f(\vec{x}).$$

Man definiert die partielle Ableitung nach der i -ten Koordinate durch

$$\int_{\mathbb{R}^n} d^n x \left(\frac{\partial}{\partial x_i} \mu(\vec{x}) \right) f(\vec{x}) = - \int_{\mathbb{R}^n} d^n x \mu(\vec{x}) \left(\frac{\partial}{\partial x_i} f(\vec{x}) \right).$$

Mehrfache Ableitungen werden analog definiert:

$$\int_{\mathbb{R}^n} d^n x \left(\frac{\partial^{i_1+\dots+i_n}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} \mu(\vec{x}) \right) f(\vec{x}) = (-1)^{i_1+\dots+i_n} \int_{\mathbb{R}^n} d^n x \mu(\vec{x}) \left(\frac{\partial^{i_1+\dots+i_n}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} f(\vec{x}) \right).$$

Als eine Anwendung betrachten wir in einer Dimension die Heavysidesche Stufenfunktion. Sie ist definiert durch

$$\Theta(x) = \begin{cases} 0 & \text{für } x < 0, \\ 1 & \text{für } x \geq 0. \end{cases}$$

Diese Funktion ist im Punkte $x = 0$ weder stetig noch differenzierbar. Faßt man dagegen $\Theta(x)$ als Distribution auf, so kann man differenzieren und es gilt

$$\frac{d}{dx} \Theta(x) = \delta(x).$$

Wir stellen also fest, daß im Rahmen der Theorie der Distributionen die Begriffe der Konvergenz und der Differenzierbarkeit wesentlich weiter gefaßt sind.

Zum Abschluß stellen wir noch einige wichtige Eigenschaften und Rechenregeln für die Diracsche Deltadistribution in einer Dimension zusammen:

$$\int_{-\infty}^{\infty} dx \delta(x-a) f(x) = f(a),$$

$$\int_{-\infty}^{\infty} dx \delta'(x-a) f(x) = -f'(a),$$

$$\delta(x-a) f(x) = \delta(x-a) f(a).$$

Hat die Funktion $g(x)$ nur einfache Nullstellen x_j (d.h. $g(x_j) = 0$ aber $g'(x_j) \neq 0$), dann gilt

$$\delta(g(x)) = \sum_j \frac{1}{|g'(x_j)|} \delta(x - x_j).$$

Insbesondere gilt

$$\delta(ax) = \frac{1}{|a|} \delta(x).$$

Weitere Relationen für die Deltadistribution:

$$x\delta(x) = 0,$$

$$x\delta'(x) = -\delta(x).$$

11.2 Die Fourier-Transformation

Fourier-Transformationen spielen in der Anwendung eine große Rolle. Man verwendet sie zum Beispiel zum Lösen partieller Differentialgleichungen.

Wir betrachten eine komplex-wertige Funktion auf dem \mathbb{R}^n , die Lebesgue-integrierbar ist:

$$f : \mathbb{R}^n \rightarrow \mathbb{C},$$

$$\vec{x} \rightarrow f(\vec{x}).$$

Wir definieren die **Fourier-Transformierte** \hat{f} von f durch das Integral

$$\hat{f}(\vec{p}) = \int d^n x f(\vec{x}) e^{i\vec{x} \cdot \vec{p}}.$$

Hierbei bezeichnet

$$\vec{x} \cdot \vec{p} = x_1 p_1 + \dots + x_n p_n$$

das Skalarprodukt im \mathbb{R}^n .

Ist nun auch \hat{f} Lebesgue-integrierbar, so läßt sich in einem ersten Schritt zeigen, daß f mit einer stetigen Funktion, welche im Unendlichen verschwindet, übereinstimmt bis auf Punkte, die eine Lebesgue-Nullmenge bilden. Gehen wir nun davon aus, daß f stetig ist und im Unendlichen verschwindet, so haben wir die Umkehrformel der Fourier-Transformation:

$$f(\vec{x}) = \int \frac{d^n p}{(2\pi)^n} \hat{f}(\vec{p}) e^{-i\vec{x} \cdot \vec{p}}.$$

Bemerkung: Man findet in der Literatur unterschiedliche Konventionen für die Aufteilung der Faktoren (2π) . Hier haben wir die Konvention verwendet, daß in der Definition von \hat{f} kein Faktor auftritt, während in der Formel für die Rücktransformation ein Faktor $1/(2\pi)^n$ auftritt. Dies ist die in der Physik übliche Konvention. In der mathematischen Literatur findet man auch oft eine symmetrische Aufteilung. Innerhalb dieser Konvention haben sowohl die Transformationsformel als auch die Rücktransformationsformel einen Faktor $1/(2\pi)^{n/2}$.

Wir betrachten nun einige Anwendungen: Ist die Funktion f stetig partiell differenzierbar und hat sie einen kompakten Träger, so gilt

$$\frac{\partial}{\partial x_j} f(\vec{x}) = \int \frac{d^n p}{(2\pi)^n} (-ip_j) \hat{f}(\vec{p}) e^{-i\vec{x} \cdot \vec{p}}.$$

Ist f darüberhinaus k -mal stetig partiell differenzierbar, so hat man analog

$$\frac{\partial^k}{\partial x_{j_1} \dots \partial x_{j_k}} f(\vec{x}) = (-i)^k \int \frac{d^n p}{(2\pi)^n} p_{j_1} \dots p_{j_k} \hat{f}(\vec{p}) e^{-i\vec{x} \cdot \vec{p}}.$$

Eine Differentiation nach x_j bringt also in der Fouriertransformierten einen Faktor $(-ip_j)$ herunter.

Wir betrachten noch die **Faltung** zweier Funktionen. Seien f und g zwei Lebesgue-integrierbare Funktionen auf dem \mathbb{R}^n . Dann definiert man die Faltung $f * g$ dieser beiden Funktionen durch

$$(f * g)(\vec{x}) = \int d^n y f(\vec{y}) g(\vec{x} - \vec{y}).$$

Für die Fourier-Transformierte der Faltung gilt

$$\widehat{(f * g)}(\vec{p}) = \hat{f}(\vec{p}) \hat{g}(\vec{p}).$$

Die Fourier-Transformation führt also eine Faltung in ein einfaches Produkt über.

Wir können die Fourier-Transformation auch auf Distributionen erweitern. Man findet für die Fourier-Transformierte der Diracschen Delta-Distribution

$$\hat{\delta}^n(\vec{p}) = \int d^n x \delta^n(\vec{x}) e^{i\vec{x}\cdot\vec{p}} = 1.$$

Die Fourier-Transformierte der Diracschen Delta-Distribution ist also eine Konstante.

11.3 Beispiele partieller Differentialgleichungen

In diesem Abschnitt wollen wir partielle Differentialgleichungen betrachten. Hierbei handelt es sich um Differentialgleichungen, in denen die gesuchte Funktion von mehreren Variablen abhängt und in denen Ableitungen nach mehreren Variablen auftreten. Die Lösung partieller Differentialgleichungen ist im allgemeinen sehr schwierig und wir wollen uns auf einige elementare Beispiele beschränken. Hierbei werden wir die Techniken der Fourier-Transformation und der Distributionen verwenden.

11.3.1 Die Potentialgleichung

Als erstes Beispiel betrachten wir die Potentialgleichung. Sie wird auch als Laplacesche Differentialgleichung oder Poissonsche Differentialgleichung bezeichnet. Gegeben sei im \mathbb{R}^n eine Funktion $\rho(\vec{x})$. Gesucht wird eine Funktion $f(\vec{x})$, welche

$$\Delta f(\vec{x}) = \rho(\vec{x})$$

erfüllt. Wie üblich ist der Laplace-Operator durch

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}$$

gegeben. Die Gleichung

$$\Delta\varphi(\vec{x}) = 0$$

wird als homogene Laplacesche Differentialgleichung bezeichnet. Lösungen der homogenen Laplaceschen Differentialgleichung werden auch als **harmonische Funktionen** bezeichnet.

Unter einer **Fundamental-Lösung** versteht man eine Funktion $g(\vec{x})$, welche die Gleichung

$$\Delta g(\vec{x}) = \delta^n(\vec{x})$$

erfüllt, wobei auf der rechten Seite die Diracsche Delta-Distribution auftritt. Ist eine Fundamental-Lösung bekannt, so erhält man eine spezielle Lösung der inhomogenen Gleichung durch

$$f(\vec{x}) = \int d^n y \rho(\vec{y}) g(\vec{x} - \vec{y}),$$

wie man leicht nachrechnet:

$$\Delta f(\vec{x}) = \int d^n y \rho(\vec{y}) \Delta g(\vec{x} - \vec{y}) = \int d^n y \rho(\vec{y}) \delta^n(\vec{x} - \vec{y}) = \rho(\vec{x}).$$

Bemerkung: Wir können zu dieser einer speziellen Lösung natürlich immer eine Lösung der homogenen Gleichung hinzuaddieren und erhalten wieder eine Lösung der inhomogenen Gleichung.

Wie bestimmt man nun eine Fundamental-Lösung? Hierzu verwendet man die Fourier-Transformation. Wir setzen

$$g(\vec{x}) = \int \frac{d^n p}{(2\pi)^n} \hat{g}(\vec{p}) e^{-i\vec{x}\cdot\vec{p}},$$

$$\delta^n(\vec{x}) = \int \frac{d^n p}{(2\pi)^n} e^{-i\vec{x}\cdot\vec{p}},$$

und bestimmen zunächst $\hat{g}(\vec{p})$. Hierzu setzen wir die Fourier-Darstellung in die zu lösende Differentialgleichung ein:

$$\Delta g(\vec{x}) = \delta^n(\vec{x}),$$

$$\Delta \int \frac{d^n p}{(2\pi)^n} \hat{g}(\vec{p}) e^{-i\vec{x}\cdot\vec{p}} = \int \frac{d^n p}{(2\pi)^n} e^{-i\vec{x}\cdot\vec{p}},$$

$$\int \frac{d^n p}{(2\pi)^n} (-p^2) \hat{g}(\vec{p}) e^{-i\vec{x}\cdot\vec{p}} = \int \frac{d^n p}{(2\pi)^n} e^{-i\vec{x}\cdot\vec{p}}.$$

Da diese Gleichung für alle \vec{x} gelten soll, folgt die Gleichheit der Integranden. Wir erhalten also

$$\hat{g}(\vec{p}) = -\frac{1}{p^2}.$$

Die Fundamental-Lösung $g(\vec{x})$ erhalten wir nun durch die Fourier-Rücktransformation:

$$g(\vec{x}) = - \int \frac{d^n p}{(2\pi)^n} \frac{e^{-i\vec{x}\cdot\vec{p}}}{p^2}$$

Betrachten wir zunächst den Fall $n = 3$: Wir wählen unser Koordinatensystem so, daß \vec{x} entlang der positiven z -Achse liegt: $\vec{x} = (0, 0, r)$. In diesem Fall erhalten wir

$$- \int \frac{d^3 p}{(2\pi)^3} \frac{e^{-i\vec{x}\cdot\vec{p}}}{p^2} = -\frac{1}{(2\pi)^3} \int_0^\infty dp \int_0^\pi d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta e^{-irp \cos \vartheta}$$

$$= -\frac{1}{(2\pi)^2} \int_0^\infty dp \int_{-1}^1 du e^{irpu}, \quad u = -\cos \vartheta$$

$$\begin{aligned}
&= -\frac{1}{(2\pi)^2} \int_0^\infty dp \frac{1}{irp} (e^{irpu} - e^{-irpu}) = -\frac{1}{2\pi^2} \int_0^\infty dp \frac{\sin(rp)}{rp} = -\frac{1}{2\pi^2} \frac{\pi}{2r} \\
&= -\frac{1}{4\pi} \frac{1}{|\vec{x}|}.
\end{aligned}$$

In n Dimensionen findet man durch eine ähnliche Rechnung

$$g(\vec{x}) = \begin{cases} \frac{1}{2\pi} \ln|\vec{x}|, & n = 2, \\ -\frac{1}{(n-2)S_n} \frac{1}{|\vec{x}|^{n-2}}, & n \neq 2, \end{cases}$$

wobei

$$S_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}$$

die Oberfläche der n -dimensionalen Einheitskugel ist.

Wir betrachten noch die Lösungen der homogenen Gleichung. Dies sind die harmonischen Funktionen. Für die harmonischen Funktionen gilt das **Maximumsprinzip**. Dies besagt folgendes: Sei $U \subset \mathbb{R}^n$ ein Gebiet und $\varphi : U \rightarrow \mathbb{R}$ eine harmonische Funktion. Nimmt φ in einem Punkt $\vec{x}_0 \in U$ ihr Maximum an, so ist φ konstant.

Wir können dies auch anders formulieren: Ist $U \subset \mathbb{R}^n$ ein beschränktes Gebiet und $\varphi : \bar{U} \rightarrow \mathbb{R}$ eine stetige Funktion die auf U harmonisch ist. Dann nimmt die Funktion φ ihr Maximum und ihr Minimum auf dem Rand von U an.

Fordert man nun von der Lösung der homogenen Gleichung die Randbedingung

$$\lim_{|\vec{x}| \rightarrow \infty} \varphi(\vec{x}) = 0,$$

so impliziert das Maximumsprinzip

$$\varphi(\vec{x}) = 0$$

auf ganz \mathbb{R}^n .

Betrachten wir nun den Fall $n \geq 3$ und sei $\rho(\vec{x})$ eine zweimal stetig partiell differenzierbare Funktion mit kompakten Träger. Dann hat die inhomogene Gleichung

$$\Delta f(\vec{x}) = \rho(\vec{x})$$

eine eindeutige Lösung zu der Randbedingung

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

11.3.2 Die Schwingungsgleichung

Wir betrachten nun die Schwingungsgleichung. Die homogene Gleichung lautet

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) \varphi(t, \vec{x}) = 0.$$

Diese Gleichung wird auch als Helmholtzsche Differentialgleichung bezeichnet. Die gesuchte Lösung $\varphi(t, \vec{x})$ ist eine Funktion von $(n + 1)$ Variablen (t, x_1, \dots, x_n) . Wie man leicht nachrechnet, stellen die ebenen Wellen

$$\varphi(t, \vec{x}) = e^{-i(\omega(\vec{p})t - \vec{p} \cdot \vec{x})}, \quad \omega(\vec{p}) = c |\vec{p}|,$$

Lösungen der homogenen Gleichung dar. Zur Verifikation betrachten wir

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) e^{-i(\omega(\vec{p})t - \vec{p} \cdot \vec{x})} = \left(-|\vec{p}|^2 + \frac{1}{c^2} \omega^2\right) e^{-i(\omega(\vec{p})t - \vec{p} \cdot \vec{x})} = 0.$$

Somit ist auch jede Überlagerung ebener Wellen eine Lösung:

$$\varphi(t, \vec{x}) = \int \frac{d^n p}{(2\pi)^n} \hat{\varphi}(\vec{p}) e^{-i(\omega(\vec{p})t - \vec{p} \cdot \vec{x})}.$$

Als nächstes wollen wir die Fundamental-Lösungen betrachten, d.h. Lösungen der Gleichung

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) g(t, \vec{x}) = \delta(t) \delta^n(\vec{x}).$$

Hier verfahren wir analog zur Potentialgleichung: Wir betrachten zunächst die Fouriertransformierten

$$\begin{aligned} g(t, \vec{x}) &= \int \frac{d\omega}{2\pi} \int \frac{d^n p}{(2\pi)^n} \hat{g}(\omega, \vec{p}) e^{-i(\omega t - \vec{p} \cdot \vec{x})}, \\ \delta(t) \delta^n(\vec{x}) &= \int \frac{d\omega}{2\pi} \int \frac{d^n p}{(2\pi)^n} e^{-i(\omega t - \vec{p} \cdot \vec{x})}. \end{aligned}$$

Bemerkung 1: Hierbei ist ω nun eine Integrationsvariable und keine Funktion von \vec{p} .

Bemerkung 2: Das Vorzeichen im Exponenten von $\vec{p} \cdot \vec{x}$ ist Konvention und wurde wie bei der Lösung der homogenen Gleichung gewählt.

Mit diesem Ansatz findet man nun für die Fouriertransformierte

$$\left(-p^2 + \frac{\omega^2}{c^2}\right) \hat{g}(\omega, \vec{p}) = 1$$

und somit

$$\hat{g}(\omega, \vec{p}) = \frac{1}{\frac{\omega^2}{c^2} - p^2}.$$

Transformiert man zurück in den Ortsraum, so erhält man

$$g(t, \vec{x}) = \int \frac{d\omega}{2\pi} \int \frac{d^n p}{(2\pi)^n} \frac{e^{-i(\omega t - \vec{p} \cdot \vec{x})}}{\frac{\omega^2}{c^2} - p^2}.$$

Betrachten wir nun wieder den Fall $n = 3$, dies entspricht der physikalischen Situation von drei Raumvariablen und einer Zeitvariablen. Wir haben also

$$g(t, \vec{x}) = \int \frac{d\omega}{2\pi} \int \frac{d^3 p}{(2\pi)^3} \frac{e^{-i(\omega t - \vec{p} \cdot \vec{x})}}{\frac{\omega^2}{c^2} - p^2}.$$

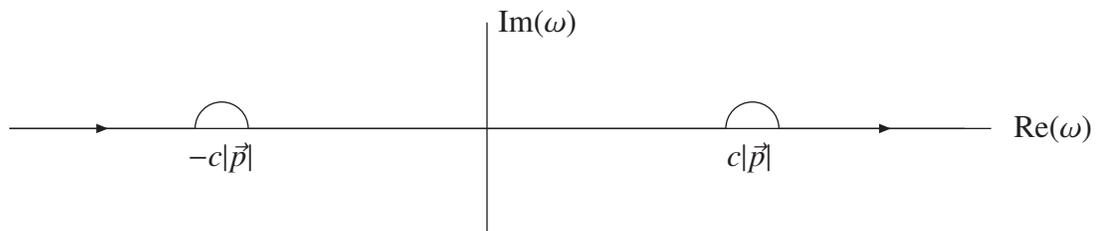
Hier tritt nun zunächst ein Problem auf: Hält man \vec{p} fest und integriert man über ω , so sieht man das der Nenner für

$$\omega = \pm c |\vec{p}|$$

Null wird. Die Lösung dieses Problems kommt aus der Funktionentheorie: Man darf den Integrationsweg für ω , der ursprünglich entlang der reellen Achse in der komplexen ω -Ebene liegt, in die komplexe Ebene verschieben. Auf diese Art können die beiden Pole umgangen werden. Nun hat man aber bei jedem der beiden Pole die Möglichkeit, die Pole entweder “oben herum” oder “unten herum” zu umgehen. Man stellt daher Zusatzforderungen die festlegen, auf welche Art die Pole umgangen werden sollen. Eine mögliche Zusatzforderung wäre zu verlangen, daß

$$g(t, \vec{x}) = 0 \quad \text{für alle } t < 0$$

gilt. Man kann zeigen, daß dies impliziert, daß beide Pole oben herum umgangen werden müssen. Wir haben also den folgenden Integrationsweg:

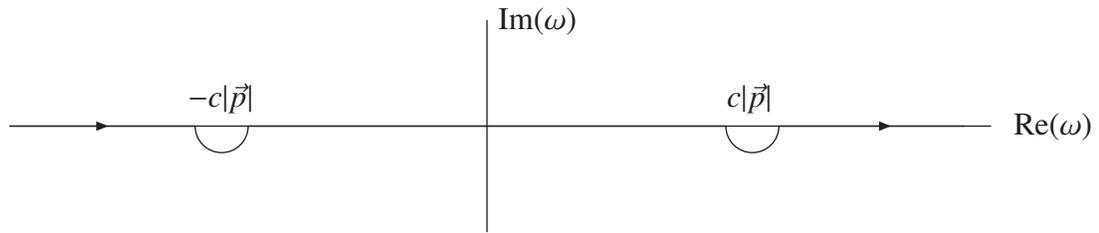


Nach einer etwas längeren Rechnung findet man

$$g(t, \vec{x}) = -\frac{1}{4\pi} \frac{1}{|\vec{x}|} \delta\left(t - \frac{1}{c} |\vec{x}|\right).$$

Offensichtlich ist $g(t, \vec{x}) = 0$ für alle $t < 0$, da in diesem Fall das Argument der Delta-Distribution immer negativ ist. Man bezeichnet diese Fundamental-Lösung als **retardierte Greensche Funktion** und verwendet oft die Schreibweise $g^+(t, \vec{x})$.

Man könnte allerdings auch verlangen, daß $g(t, \vec{x}) = 0$ für alle $t > 0$. In diesem Fall muß man den folgenden Integrationsweg wählen:



d.h. man weicht den Polen nach unten aus. In diesem Fall findet man

$$g(t, \vec{x}) = -\frac{1}{4\pi} \frac{1}{|\vec{x}|} \delta\left(t + \frac{1}{c} |\vec{x}|\right).$$

Man bezeichnet diese Fundamental-Lösung als **avancierte Greensche Funktion** und verwendet die Schreibweise $g^-(t, \vec{x})$ für diese Lösung.

11.3.3 Die Wärmeleitungsgleichung

Die Wärmeleitungsgleichung lautet

$$\left(\Delta - \frac{\partial}{\partial t}\right) f(t, \vec{x}) = 0.$$

Bemerkung: Im Gegensatz zu den obigen Beispielen ist diese Gleichung nicht homogen im Grad der Ableitungen: Die Ableitungen nach den Variablen x_j treten immer zweifach auf, die Ableitung nach der Variablen t dagegen nur einfach.

Wir beschränken uns hier darauf, eine Fundamental-Lösung der Gleichung

$$\left(\Delta - \frac{\partial}{\partial t}\right) g(t, \vec{x}) = \delta^n(\vec{x})\delta(t)$$

anzugeben. Diese lautet

$$g(t, \vec{x}) = \begin{cases} 0, & t \leq 0, \\ -\frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{x^2}{4t}}, & t > 0. \end{cases}$$