

Vorlesung Mathematik für Physikstudierende 2

Andreas Knauf*

Sommersemester 2020

Zusammenfassung

Vorlesungsbegleitendes Skript für das zweite Semester. Anregungen und Kritik sind willkommen!

Inhaltsverzeichnis

Zur Notation	iv
Kleines Englisch-Wörterbuch	v
1 Integrationstechniken	1
1.1 Partielle Integration und Substitution	1
1.2 Integration rationaler Funktionen	2
1.2.1 Reelle und komplexe Polynome	2
1.2.2 Rationale Funktionen	5
1.2.3 Integration reeller rationaler Funktionen	7
1.2.4 Weitere berechenbare Integrale	8
2 Matrizen und Vektorraum-Endomorphismen	10
2.1 Die Jordansche Normalform	10
2.1.1 Äquivalenz und Ähnlichkeit	10
2.1.2 Nilpotente Endomorphismen	14
2.1.3 Haupträume	20
2.1.4 Jordan-Zerlegung und Jordansche Normalform	25
2.2 Ein Beispiel	28
2.3 Euklidische und unitäre Vektorräume: Fortsetzung	30

*Department Mathematik, Universität Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen, Raum 02.321. e-mail: knauf@math.fau.de, web: <http://math.fau.de/knauf>

2.3.1	Orthonormalisierung	32
2.3.2	Die adjungierte Abbildung	34
2.3.3	Normale Operatoren	37
2.3.4	Orthogonale und unitäre Operatoren	39
2.3.5	(Anti-) Selbstdjungierte Endomorphismen	44
2.3.6	Anwendung: Quantenmechanik	48
3	Lineare Differentialgleichungen mit konstanten Koeffizienten	49
3.1	Berechnung der Lösung	50
3.2	* Lineare Flüsse in der Ebene	55
3.3	* Ein Beispiel	59
4	Quadratische Formen	62
4.1	Affinitäten und Kongruenzen	62
4.2	Normalformen von Quadriken	65
5	Metrik, Topologie und Stetigkeit	70
5.1	Metrische Räume	70
5.2	Folgen in metrischen Räumen	72
5.3	Topologische Räume	76
5.4	Stetigkeit	82
5.5	Kompaktheit	85
6	Kurven im \mathbb{R}^n	90
6.1	Reguläre und nicht reguläre Kurven	90
6.2	Wechsel der Parametrisierung	94
6.3	Krümmung einer Kurve	95
7	Gewöhnliche Differentialgleichungen	99
7.1	Definitionen und Beispiele	99
7.2	Der Banachscher Fixpunktsatz	103
7.3	Lokale Existenz und Eindeutigkeit der Lösung	107
7.4	Globale Existenz und Eindeutigkeit der Lösung	111
8	Differentialrechnung mehrerer Variablen	115
8.1	Die Landau-Symbole	117
8.2	Definition der totalen Ableitung	119
8.3	Partielle Ableitungen	120
8.4	Der Gradient	124
8.5	Differentiationsregeln	125
8.6	Höhere Ableitungen	128
8.7	Vektoranalysis	131

8.8	Anwendungen der Kettenregel	136
9	Der Satz von Taylor im \mathbb{R}^n	138
9.1	Die Taylor-Formel	138
9.2	Die Taylor-Formel in m Dimensionen	141
9.3	Die Hesse-Matrix einer reellen Funktion	144
9.4	Extremalstellen	145
10	Implizite Funktionen	149
10.1	Das Newton-Verfahren	151
10.2	Das vereinfachte Newton-Verfahren	154
10.3	Konstruktion der impliziten Funktionen	157
11	Extrema mit Nebenbedingungen	163
11.1	1. Methode: Parametrisierung der Nebenbedingungen	164
11.2	2. Methode: Lagrange–Multiplikatoren	165
12 *	Ergänzungen	168
12.1	Physikalische Bedeutung der Fraktale	168
12.2	Physikalische Bedeutung der Topologie	170
	Literatur	173
	Index	174

Da dieses *pdf* [intern](#) und [extern](#) verlinkt ist, kann es sinnvoll sein, es *online* zu benutzen statt es auszudrucken.

Danksagung: Ich danke Frau I. Moch für ihre hervorragende Arbeit beim Schreiben des Manuskriptes. Erlangen, im Februar 2020, A.K.

Vorbemerkung: Mit * gekennzeichnete Kapitel sind nicht Prüfungsstoff. Dieses Skript kann kein Lehrbuch ersetzen. Einige Lehrbücher zur Mathematik für Physikstudierende sind im Literaturverzeichnis erwähnt.

Zur Notation

Teilmengen: Sind A und B Mengen, dann heißt A *Teilmenge* von B (in Zeichen $A \subseteq B$), wenn gilt: $x \in A \Rightarrow x \in B$. Insbesondere gilt $B \subseteq B$. Die *echte Inklusion* $A \subsetneq B$ bedeutet, dass $A \subseteq B$, aber $A \neq B$ gilt. (in der mathematischen Literatur findet man davon abweichend auch das Teilmengenzeichen $A \subset B$.)

Potenzmengen: Ist A eine Menge, dann ist

$$2^A := \{B \mid B \subseteq A\}$$

die *Potenzmenge* von A . Synonym findet man auch die Notationen $\mathfrak{P}(A)$ und $\mathcal{P}(A)$.

Zahlen: Menge $\mathbb{N} = \{1, 2, \dots\}$ der natürlichen Zahlen, $\mathbb{N}_0 = \{0, 1, 2, \dots\}$, Ring $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ der ganzen Zahlen.

Körper $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ der rationalen, reellen bzw. komplexen Zahlen.

Für einen Körper \mathbb{K} bedeutet \mathbb{K}^* die multiplikative Gruppe $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$, und

$$\mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\} = (0, \infty).$$

Intervalle: Für $a, b \in \mathbb{R}$, $a < b$ ist

$$\begin{aligned}(a, b) &:= \{x \in \mathbb{R} \mid x > a, x < b\}, \\(a, b] &:= \{x \in \mathbb{R} \mid x > a, x \leq b\} \text{ etc.}\end{aligned}$$

(Synonym findet man auch die Notation $]a, b[= (a, b)$, $]a, b] = (a, b]$ etc.)

Matrizen: $\text{Mat}(m \times n, \mathbb{K})$ bezeichnet den \mathbb{K} -Vektorraum der $m \times n$ -Matrizen mit Einträgen aus dem Körper \mathbb{K} , und $\text{Mat}(n, \mathbb{K})$ den Ring $\text{Mat}(n \times n, \mathbb{K})$.

Das griechische Alphabet: a) Kleinbuchstaben

α	Alpha	ζ	Zeta	λ	Lambda	π	Pi	ϕ, φ	Phi
β	Beta	η	Eta	μ	My	ρ, ϱ	Rho	χ	Chi
γ	Gamma	θ, ϑ	Theta	ν	Ny	σ, ς	Sigma	ψ	Psi
δ	Delta	ι	Jota	ξ	Xi	τ	Tau	ω	Omega
ϵ, ε	Epsilon	κ	Kappa	\omicron	Omikron	υ	Ypsilon		

b) Großbuchstaben

Γ	Gamma	Θ	Theta	Ξ	Xi	Σ	Sigma	Φ	Phi	Ω	Omega
Δ	Delta	Λ	Lambda	Π	Pi	Υ	Ypsilon	Ψ	Psi		

Kleines Englisch-Wörterbuch

abelian	abelsch	intersection	Durchschnitt
absolute value	Betrag	interval	Intervall
accumulation point	Häufungspunkt	inverse mapping	Umkehrabbildung
area	Fläche	limit	Limes
assertion	Aussage	map	Abbildung
associativity	Assoziativität	metric	Metrik
asymptotic value	Grenzwert	metric space	metrischer Raum
average	Mittelwert	monotonous	monoton
ball	Vollkugel	neighborhood	Umgebung
bound	Schranke	numbers	Zahlen
bounded	beschränkt	- complex	- komplexe
box	Würfel	- integer	- ganze
cardinality	Mächtigkeit	- irrational	- irrationale
cartesian product	kartesisches Produkt	- natural	- natürliche
chain rule	Kettenregel	- rational	- rationale
circle	Kreislinie	- real	- reelle
closed	abgeschlossen	one-to-one	injektiv
complete	vollständig	onto	surjektiv
continuity	Stetigkeit	open	offen
convergent	konvergent	order	Ordnung
convolution	Faltung	partition	Zerlegung
countable	abzählbar	proposition, theorem	Satz
derivative	Ableitung	power series	Potenzreihe
disjoint	disjunkt	power set	Potenzmenge
disk	Kreisscheibe	primes	Primzahlen
distance	Abstand	real part	Realteil
divergent	divergent	relation	Relation
domain	Definitionsbereich	ring	Ring
empty set	leere Menge	root	Wurzel
equivalence class	Äquivalenzklasse	sequence	Folge
field	Körper	set	Menge
fixed point	Fixpunkt	sign	Signum
function	Funktion	stable	stabil
graph	Graph	subsequence	Teilfolge
group	Gruppe	subset	Teilmenge
image	Bild	triangle inequality	Dreiecksungleichung
imaginary part	Imaginärteil	union	Vereinigung
imaginary unit	imaginäre Einheit	unit	Einheit
inequality	Ungleichung	well defined	wohldefiniert

1 Integrationstechniken

1.1 Partielle Integration und Substitution

Mit dem Hauptsatz haben wir ein Mittel in der Hand, die unbestimmten Integrale vieler Funktionen zu berechnen, indem wir umgekehrt schauen, welche Funktionen wir durch Ableitung elementarer Funktionen erhalten.

1.1 Beispiele

- $\int x^b dx = \frac{x^{b+1}}{b+1} + C \quad (b \in (-1, \infty))$,
denn für $F(x) := x^a$ mit $a > 0$ ist $F'(x) = \frac{d}{dx} e^{a \ln x} = a e^{a \ln x} \cdot \frac{1}{x} = a x^{a-1}$.
- $\int \ln |x| dx = x(\ln |x| - 1) + C$,
denn für $f(x) := x \ln |x|$ ist $f'(x) = \ln |x| + 1$. ◇

Im Gegensatz zur Ableitung einer elementaren Funktion¹ ist das unbestimmte Integral einer elementaren Funktion im Allgemeinen nicht elementar². Ein Beispiel: Das unbestimmte Integral der elementaren Funktion $x \mapsto e^{-x^2}$ ist ein Vielfaches der nicht elementaren Fehlerfunktion.

Die Regeln für die Differentiation werden zu Regeln der Integration. So ist es selbstverständlich, dass wir das Integral einer endlichen Summe von Funktionen als die Summe der Integrale der einzelnen Funktionen berechnen können.

Ein wichtiges Beispiel bildet die Produktregel

$$(FG)' = F'G + FG'$$

der Differentiation. Aus ihr leitet sich die *partielle Integration* ab.

1.2 Satz (Partielle Integration)

Für stetig differenzierbare Funktionen $F, G : [a, b] \rightarrow \mathbb{R}$ ist

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

1.3 Beispiel (Partielle Integration) Wir wollen $\int \arcsin(x) dx$ berechnen und setzen $F := \arcsin$. Mit $G(x) := x$ ist $G'(x) = 1$, also

$$\int_0^y \arcsin(x) dx = FG \Big|_0^y - \int_0^y F'(x)G(x) dx = y \arcsin(y) - \int_0^y \frac{x}{\sqrt{1-x^2}} dx.$$

Scheinbar haben wir uns nur ein neues nicht berechenbares Integral eingehandelt; dieses können wir aber mit einer weiteren Integrationsmethode behandeln:

¹ also die Polynome, die Exponentialfunktion, der Logarithmus, die (invers) trigonometrischen Funktionen, und aus diesen durch endlich viele Anwendungen der Grundrechenarten und Kompositionen gewonnene Funktionen.

²Es gibt aber den **Algorithmus von Risch**, mit dem man dies entscheiden kann.

Die Kettenregel $(F \circ G)' = (F' \circ G) \cdot G'$ der Differentiation führt auf die Substitutionsregel der Integration:

1.4 Satz (Substitutionsregel) Für $f \in C(I, \mathbb{R})$ und eine stetig differenzierbare Funktion $G : [a, b] \rightarrow I$ gilt $\int_a^b f(G(x))G'(x) dx = \int_{G(a)}^{G(b)} f(y) dy$.

Beweis: Es sei F eine Stammfunktion von f . Dann ist $f(G(x))G'(x) = (F \circ G)'(x)$, also

$$\int_a^b f(G(x))G'(x) dx = \int_a^b (F \circ G)'(x) dx = F \circ G \Big|_a^b = F \Big|_{G(a)}^{G(b)} = \int_{G(a)}^{G(b)} f(y) dy. \quad \square$$

1.5 Beispiele (Integration durch Substitution)

1. Es ist für alle $c \in \mathbb{R} \setminus \{0\}$, $d \in \mathbb{R}$ und auf $[ca + d, cb + d]$ integrierbaren Funktionen f

$$\int_a^b f(cx + d) dx = \frac{1}{c} \int_{ca+d}^{cb+d} f(y) dy.$$

2. Setze $F(y) := \sqrt{1-y}$ und $G(x) := x^2$, also $F \circ G = \sqrt{1-x^2}$ und $(F \circ G)' = -\frac{x}{\sqrt{1-x^2}}$. Es ergibt sich

$$\int \frac{x}{\sqrt{1-x^2}} dx = -\sqrt{1-x^2} + C. \quad \diamond$$

1.2 Integration rationaler Funktionen

Im Gegensatz zur Differentiation führt die Integration elementarer Funktionen nicht immer auf elementare Funktionen, und es gibt auch keinen Algorithmus für die analytische Integration. Ein solcher existiert aber immerhin für die große Klasse der rationalen Funktionen.

Wir beginnen mit einigen algebraischen Betrachtungen, die die Partialbruchzerlegung einer beliebigen rationalen Funktion in eine Summe einfacherer rationaler Funktionen vorbereiten.

1.2.1 Reelle und komplexe Polynome

1.6 Definition

• Für den Körper $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} heißt eine Funktion $p : \mathbb{K} \rightarrow \mathbb{K}$ der Form

$$p(x) = \sum_{k=0}^n a_k x^k \quad \text{mit} \quad a_k \in \mathbb{K} \quad (1.1)$$

reelles bzw. komplexes Polynom.

- Ist dabei $a_n \neq 0$, dann ist p ein Polynom vom **Grad**

$$\deg(p) := \text{grad}(p) := n.$$

- Das Polynom $0 : \mathbb{K} \rightarrow \mathbb{K}$, $0(x) = 0$ heißt **Nullpolynom** und $\deg(0) := -\infty$.
- Der Ring³ aller Polynome mit Koeffizienten aus \mathbb{K} wird mit $\mathbb{K}[x]$ bezeichnet.

Die Festlegung des Grades des Nullpolynoms erscheint willkürlich, ist aber dadurch gerechtfertigt, dass die folgenden Regeln für den Grad gelten:

Summe und Produkt zweier Polynome $f, g \in \mathbb{K}[x]$ sind Polynome mit

$$\deg(f + g) \leq \max(\deg(f), \deg(g))$$

und

$$\deg(fg) = \deg(f) + \deg(g),$$

wobei wir $k + (-\infty) := -\infty$ setzen⁴.

Die Koeffizienten a_k in (1.1) sind durch p eindeutig bestimmt (denn andernfalls gäbe es eine Darstellung $0(x) = \sum_{k=0}^n a_k x^k$ des Nullpolynoms mit $a_n \neq 0$, was der Identität $0 = \lim_{x \rightarrow \infty} \frac{0(x)}{x^n} = a_n + \lim_{x \rightarrow \infty} \sum_{k=0}^{n-1} a_k x^{k-n}$ widerspräche).

Lassen wir als Argumente x in (1.1) beliebige komplexe Zahlen zu, dann lässt sich ein reelles Polynom als komplexes Polynom auffassen. In diesem Sinn gilt

$$\mathbb{R}[x] \subset \mathbb{C}[x].$$

Reelle Polynome vom Grad $\deg(p) > 0$ brauchen keine Nullstellen zu besitzen, wohl aber komplexe Polynome (**Bsp.:** $p(x) = x^2 + 1 = (x + i)(x - i)$).

³**Definition:** Eine Menge A mit zwei Operationen

$$+ : A \times A \rightarrow A \quad \text{und} \quad \cdot : A \times A \rightarrow A$$

heißt **Ring**, falls gilt:

1. $(A, +)$ ist eine abelsche Gruppe (mit neutralem Element 0 und zu a inversem Element $-a$).
2. Assoziativität: $(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad (a, b, c \in A)$.
3. Distributivität:
 $a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (b + c) \cdot a = b \cdot a + c \cdot a \quad (a, b, c \in A)$.

4. Falls zusätzlich gilt:

$$\exists 1 \in A \quad \forall a \in A : 1 \cdot a = a,$$

heißt A **Ring mit Einselement**.

5. Falls $\forall a, b \in A : a \cdot b = b \cdot a$, heißt A **kommutativer Ring**.

⁴Für $g \neq 0$ gilt $\deg(f \circ g) = \deg(f) \deg(g)$.

Letzteres ist der Inhalt des so genannten *Fundamentalsatzes der Algebra* (Satz 7.22 der *Analysis I*, Beweis: z.B. in HILDEBRANDT [Hi], Band 1).

Der Fundamentalsatz der Algebra ermöglicht es, ein Polynom $p \in \mathbb{C}[x]$ n -ten Grades in ein Produkt von *Linearfaktoren*, d.h. Polynomen vom Grad Eins zu zerlegen. Der Einfachheit halber nehmen wir an, dass p *normiert* ist, d.h. dass der so genannte *Leitkoeffizient* a_n gleich Eins ist.

1.7 Satz *Ist $p \in \mathbb{C}[x]$ ein normiertes Polynom mit $n := \deg(p) \in \mathbb{N}_0$, dann gibt es (bis auf ihre Nummerierung eindeutig bestimmte) Zahlen*

$$x_1, \dots, x_n \in \mathbb{C} \quad \text{mit} \quad p(x) = \prod_{i=1}^n (x - x_i).$$

Beweis:

- Im Fall $n = 0$ gilt $p(x) = 1$. Ist dagegen $n > 0$, dann existiert nach dem Fundamentalsatz der Algebra ein $x_1 \in \mathbb{C}$ mit $p(x_1) = 0$. Wir führen nun Polynomdivision mit Rest und Divisor $p_1(x) = x - x_1$ durch und erhalten eindeutig bestimmte $q, r \in \mathbb{C}[x]$ mit $\deg(r) < \deg(p_1) = 1$ und $p = p_1 q + r$. Da nun $r(x) = p(x_1) - p_1(x_1)q(x_1)$ gilt, ist $r = 0$, also $p = p_1 q$. q ist ein normiertes Polynom mit $\deg(q) = n - 1$. Weiter durch Induktion.
- Ist gleichzeitig $p(x) = \prod_{i=1}^n (x - x'_i)$, dann müssen die *Mengen* $\{x_1, \dots, x_n\}$ und $\{x'_1, \dots, x'_n\}$ der Nullstellen gleich sein, also $x'_1 = x_{\pi(1)}$. Daher muss auch $\prod_{i=2}^n (x - x_i) = \prod_{i=2}^n (x - x'_i)$ gelten, sodass wir induktiv eine Permutation π der Indizes mit $x'_i = x_{\pi(i)}$, $i = 1, \dots, n$ erhalten. \square

Die x_i in der Linearfaktorzerlegung brauchen nicht voneinander verschieden zu sein. Tritt ein Faktor k -mal auf, spricht man von einer *k -fachen Nullstelle*.

Eine ähnliche Zerlegung in Linearfaktoren können wir für reelle Polynome nicht erwarten. Immerhin gilt der *Faktorisierungssatz für reelle Polynome*:

1.8 Satz

1. *Ist $x_1 \in \mathbb{C}$ k -fache Nullstelle des reellen Polynoms $p \in \mathbb{R}[x] \subset \mathbb{C}[x]$, dann ist auch \bar{x}_1 k -fache Nullstelle von p .*
2. *Ist $p \in \mathbb{R}[x]$ normiert, dann gibt es $a_1, \dots, a_l, b_1, \dots, b_l, c_1, \dots, c_m \in \mathbb{R}$ mit $2l + m = \deg(p)$, $a_i^2 - b_i < 0$ und*

$$p(x) = \prod_{i=1}^l (x^2 - 2a_i x + b_i) \cdot \prod_{j=1}^m (x - c_j).$$

Beweis: Für $p \in \mathbb{C}[x]$ der Form $p(x) = \sum_{k=0}^n a_k x^k$ sei $\bar{p} \in \mathbb{C}[x]$ gleich $\bar{p}(x) := \sum_{k=0}^n \bar{a}_k x^k$.

1. Wäre $p \in \mathbb{C}[x]$ ein beliebiges komplexes Polynom mit k -facher Nullstelle x_1 , also $p(x) = (x - x_1)^k q(x)$, dann wäre $\overline{x_1}$ k -fache Nullstelle von \overline{p} , denn $\overline{p}(x) = (x - \overline{x_1})^m \overline{q}(x)$. Nun ist aber $p \in \mathbb{R}[x]$, also $\overline{p} = p$.
2. Nach Satz 1.7 schreiben wir p in der Form

$$p(x) = \prod_{i=1}^m (x - x_i) \cdot \prod_{i=m+1}^n (x - x_i)$$

mit $x_1, \dots, x_m \in \mathbb{R}$ und $x_{m+1}, \dots, x_n \in \mathbb{C} \setminus \mathbb{R}$.

Nach Teil 1. des Satzes ist $l := \frac{n-m}{2} \in \mathbb{N}_0$, und nach Ummummerierung gilt

$$\prod_{i=m+1}^n (x - x_i) = \prod_{i=m+1}^{m+l} (x - x_i)(x - \overline{x_i}).$$

Wir setzen $c_i := x_i$ ($i = 1, \dots, m$), $a_i := \operatorname{Re}(x_{i+m})$ und $b_i := x_{i+m} \overline{x_{i+m}}$ ($i = 1, \dots, l$). \square

Wir können also reelle Polynome im Wesentlichen eindeutig in reelle Linearfaktoren und quadratische Polynome zerlegen.

1.2.2 Rationale Funktionen

Die (reellen wie auch komplexen) Polynome bilden bezüglich punktweiser Addition und Multiplikation einen Ring, jedoch keinen Körper, denn nicht konstante Polynome besitzen kein multiplikativ inverses Element. Gehen wir aber zu den *rationalen Funktionen* über $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} über, d.h. Funktionen, die sich in der Form p/q mit $p, q \in \mathbb{K}[x]$ schreiben lassen, dann existiert für alle $p/q \neq 0$ der Kehrwert q/p . Wir müssen nur darauf achten, dass der Definitionsbereich von p/q die Menge $\{x \in \mathbb{K} \mid q(x) \neq 0\}$ ist.

Nun besitzt eine rationale Funktion keine eindeutige Darstellung p/q , denn wir können ja Zähler und Nenner mit einem beliebigen Polynom $r \neq 0$ erweitern.

Wir nehmen aber an, dass p und q schon teilerfremd sind, was für $\mathbb{K} = \mathbb{C}$ nach Satz 1.7 bedeutet, dass sie keine gemeinsamen Nullstellen besitzen.

Bezeichnet man den Körper der rationalen Funktionen mit $\mathbb{K}(x)$, dann gilt $\mathbb{K}[x] \subset \mathbb{K}(x)$, und wir bezeichnen die Polynome manchmal auch als die *ganzrationalen Funktionen*.

Eine rationale Funktion $p/q \in \mathbb{K}(x)$ können wir eindeutig in der Form

$$\frac{p}{q} = t + \frac{\tilde{p}}{q} \quad \text{mit } t, \tilde{p} \in \mathbb{K}[x] \quad \text{und} \quad \deg(\tilde{p}) < \deg(q) \quad (1.2)$$

schreiben, wobei wir t und \tilde{p} mit dem **Euklidischen Algorithmus** berechnen.

1.9 Beispiel (Polynomdivision) $\frac{x^4}{x^3+2x^2-1} = x - 2 + \frac{4x^2+x-2}{x^3+2x^2-1}$, denn

$$(x^3 + 2x^2 - 1)(x - 2) = x^4 - 4x^2 - x + 2. \quad \diamond$$

Die Polynome $t \in \mathbb{R}[x]$ in (1.2) können wir integrieren. Wir müssen uns also nur noch um die Bestimmung des unbestimmten Integrals rationaler Funktionen p/q kümmern, deren Nenner-Grad $\deg(q)$ größer als der Zähler-Grad $\deg(p)$ ist.

Die Technik der **Partialbruchzerlegung** ermöglicht es, rationale Funktionen aus $\mathbb{C}(x)$ in eine Summe rationaler Funktionen zu zerlegen, deren Nennerpolynom höchstens eine Nullstelle besitzt. Der folgende Satz gilt aber für $\mathbb{K} = \mathbb{C}$ und $\mathbb{K} = \mathbb{R}$.

1.10 Satz (Partialbruchzerlegung)

Es seien $p, q \in \mathbb{K}[x]$, $x_1 \in \mathbb{K}$ und $k \in \mathbb{N}$. Ist $\deg(p) < \deg(q) + k$, dann gibt es eindeutig bestimmte $\tilde{p} \in \mathbb{K}[x]$ und $A_1, \dots, A_k \in \mathbb{K}$ mit

$$\frac{p(x)}{q(x)(x-x_1)^k} = \frac{\tilde{p}(x)}{q(x)} + \sum_{j=1}^k \frac{A_j}{(x-x_1)^j},$$

wobei $\deg(\tilde{p}) < \deg(q)$ gilt.

Beweis: Einfachheit halber nehmen wir an, dass x_1 eine k -fache Nullstelle des Nennerpolynoms ist, d.h. $q(x_1) \neq 0$ gilt. Wir reduzieren nun schrittweise deren Vielfachheit, indem wir zeigen, dass

$$\frac{p(x)}{q(x)(x-x_1)^k} = \frac{p_k(x)}{q(x)(x-x_1)^{k-1}} + \frac{A_k}{(x-x_1)^k} \quad (1.3)$$

mit einzigem Polynom $p_k \in \mathbb{K}[x]$ vom Grad $\deg(p_k) < \deg(q) + k - 1$ und $A_k \in \mathbb{K}$ ist.

Erweitern wir (1.3) mit $(x-x_1)^k$ und setzen wir $x = x_1$ ein, so ergibt sich

$$A_k = \frac{p(x_1)}{q(x_1)} \in \mathbb{K}.$$

Auflösen nach p_k ergibt

$$p_k(x) = \frac{p(x) - A_k q(x)}{x - x_1}.$$

Dies ist aber ein Polynom aus $\mathbb{K}[x]$, denn der Zähler der rechten Seite verschwindet bei Einsetzen von x_1 .

Da nach Voraussetzung $\deg(p) < \deg(q) + k$ gilt, ist $\deg(p - A_k q) < \deg(q) + k$, also $\deg(p_k) < \deg(q) + k - 1$. \square

Ist das Nennerpolynom von der Form

$$q(x) = \prod_{l=1}^m (x - x_l)^{d_l}$$

und sind die Nullstellen x_l der Vielfachheit d_l voneinander verschieden, dann führt mehrfache Anwendung des Satzes zu der Formel

$$\frac{p(x)}{q(x)} = \sum_{l=1}^m \sum_{j=1}^{d_l} \frac{A_j^{(l)}}{(x-x_l)^j}. \quad (1.4)$$

Die Anzahl der zu bestimmenden Koeffizienten $A_j^{(l)}$ ist $\sum_{l=1}^m d_l = \deg(q)$. Nach Multiplikation von (1.4) mit q werden diese durch Koeffizientenvergleich der beiden Polynome berechnet. Das läuft auf die Lösung eines linearen Gleichungssystems hinaus.

1.11 Beispiel (Partialbruchzerlegung)

$\frac{2x^3+4x}{(x+1)^2(x-1)^2} = \frac{A_1^{(1)}}{x+1} + \frac{A_2^{(1)}}{(x+1)^2} + \frac{A_1^{(2)}}{x-1} + \frac{A_2^{(2)}}{(x-1)^2}$ ist äquivalent zu

$$\begin{aligned} 2x^3 + 4x &= A_1^{(1)}(x+1)(x-1)^2 + A_2^{(1)}(x-1)^2 + A_1^{(2)}(x+1)^2(x-1) \\ &\quad + A_2^{(2)}(x+1)^2 \\ &= \left(A_1^{(1)} + A_1^{(2)}\right)x^3 + \left(-A_1^{(1)} + A_2^{(1)} + A_1^{(2)} + A_2^{(2)}\right)x^2 \\ &\quad + \left(-A_1^{(1)} - 2A_2^{(1)} - A_1^{(2)} + 2A_2^{(2)}\right)x \\ &\quad + A_1^{(1)} + A_2^{(1)} - A_1^{(2)} + A_2^{(2)}. \end{aligned}$$

Das lineare Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ -1 & -2 & -1 & 2 \\ 1 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} A_1^{(1)} \\ A_2^{(1)} \\ A_1^{(2)} \\ A_2^{(2)} \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}$$

besitzt die Lösung $A_1^{(1)} = 1$, $A_2^{(1)} = \frac{-3}{2}$, $A_1^{(2)} = 1$, $A_2^{(2)} = \frac{3}{2}$. \diamond

Analog zu Satz 1.10 zeigt man

1.12 Satz *Es seien $p, q \in \mathbb{R}[x]$, $a, b \in \mathbb{R}$ und $k \in \mathbb{N}$. Ist $\deg(p) < \deg(q) + k$, dann gibt es eindeutig bestimmte $\tilde{p} \in \mathbb{R}[x]$ und $A_1, \dots, A_k, B_1, \dots, B_k \in \mathbb{R}$ mit*

$$\frac{p(x)}{q(x)(x^2-2ax+b)^k} = \frac{\tilde{p}(x)}{q(x)} + \sum_{j=1}^k \frac{A_j x + B_j}{(x^2-2ax+b)^j},$$

wobei $\deg(\tilde{p}) < \deg(q)$ gilt.

1.2.3 Integration reeller rationaler Funktionen

Wir können also reelle rationale Funktionen in Summen von Termen der Form $\frac{Ax+B}{(x^2-2ax+b)^j}$ bzw. $\frac{A}{(x-a)^j}$ zerlegen, wobei die Koeffizienten reell sind.

Damit gestattet uns der folgende Satz die Integration rationaler Funktionen:⁵

1.13 Satz *Für $j \in \mathbb{N}$ und $a, b \in \mathbb{R}$ ist*

$$\int \frac{dx}{(x-a)^j} = \begin{cases} \ln|x-a| & , \quad j = 1 \\ \frac{-1}{(j-1)(x-a)^{j-1}} & , \quad j > 1. \end{cases}$$

⁵Wir lassen im Folgenden die Integrationskonstanten weg. Ohnehin ist der Definitionsbereich einer reellen rationalen Funktion die disjunkte Vereinigung endlich vieler offener Intervalle, und für jedes Intervall können wir eine beliebige Konstante zur Stammfunktion addieren.

Ist $a^2 < b$, dann gilt

$$I_j := \int \frac{dx}{(x^2 - 2ax + b)^j} = \begin{cases} \frac{\arctan\left(\frac{x-a}{\sqrt{b-a^2}}\right)}{\sqrt{b-a^2}} & , j = 1 \\ \frac{x-a}{2(b-a^2)(j-1)(x^2-2ax+b)^{j-1}} + \frac{2j-3}{2(b-a^2)(j-1)} I_{j-1} & , j > 1 \end{cases}$$

und

$$\int \frac{x-a}{(x^2-2ax+b)^j} dx = \begin{cases} \frac{1}{2} \ln(x^2 - 2ax + b) & , j = 1 \\ \frac{-1}{2(j-1)(x^2-2ax+b)^{j-1}} & , j > 1. \end{cases}$$

Beweis: Durch Differentiation. □

1.14 Beispiel (Integration einer rationalen Funktionen)

$$\begin{aligned} \int \frac{x+2}{x^2-4x+7} dx &= \int \frac{x-2}{x^2-4x+7} dx + 4 \int \frac{1}{x^2-4x+7} dx \\ &= \frac{1}{2} \ln(x^2 - 4x + 7) + \frac{4}{\sqrt{3}} \arctan\left(\frac{x-2}{\sqrt{3}}\right). \end{aligned} \quad \diamond$$

1.2.4 Weitere berechenbare Integrale

Man kann viele weitere Integrale durch Substitution auf Integrale rationaler Funktionen zurückführen:

1.15 Beispiel (Integration durch Transformation in rationale Funktion)

Mit $u(x) := e^x$, also $x = \ln(u)$ ergibt sich

$$\begin{aligned} \int \frac{e^x+1}{e^x+e^{-x}} dx &= \int \frac{u+1}{u+1/u} \underbrace{\frac{dx}{du}}_{1/u} du = \int \frac{u+1}{u^2+1} du \\ &= \frac{1}{2} \ln(u^2 + 1) + \arctan(u) + c \\ &= \frac{1}{2} \ln(e^{2x} + 1) + \arctan(e^x) + c \end{aligned} \quad \diamond$$

1.16 Satz

Die folgenden unbestimmten Integrale stellen elementare Funktionen dar:

- $\int R(e^{ax}) dx$
- $\int R(\sinh(ax), \cosh(ax)) dx$
- $\int R(\sin(ax), \cos(ax)) dx$
- $\int R(x, \sqrt[k]{ax+b}) dx \quad (k \in \mathbb{N})$
- $\int R(x, \sqrt[k]{\frac{ax+b}{cx+d}}) dx \quad (k \in \mathbb{N})$

$$f) \int R(x, \sqrt{ax^2 + bx + c}) dx$$

Dabei bezeichnet $R(x, y)$ eine **rationale Funktion von x und y** , d.h.

$$R(x, y) = \frac{\sum_{i,j=0}^n a_{ij} x^i y^j}{\sum_{i,j=0}^n b_{ij} x^i y^j}.$$

Beweis: Man muss jeweils nur die Substitution $x = x(u)$ angeben, die auf ein Integral einer rationalen Funktion von u führt:

a) Aus $u(x) := e^{ax}$ (wie in Beispiel 1.15) ergibt sich

$$x = \frac{1}{a} \ln(u) \quad \text{und} \quad \frac{dx}{du} = \frac{1}{au}, \quad \text{also}$$

$$\int R(e^{ax}) dx = \int \frac{R(u)}{au} du.$$

b) lässt sich auf a) zurückführen, denn

$$\sinh(ax) = \frac{1}{2}(e^{ax} - e^{-ax}) \quad \text{und} \quad \cosh(ax) = \frac{1}{2}(e^{ax} + e^{-ax}).$$

c) O.B.d.A. $a = 1$ (lineare Substitution).

Setze $u(x) := \tan\left(\frac{x}{2}\right)$, also $x(u) = 2 \arctan(u)$ und $\frac{dx}{du} = \frac{2}{1+u^2}$. Wegen

$$\cos^2\left(\frac{x}{2}\right) = \frac{1}{1+\tan^2\left(\frac{x}{2}\right)} = \frac{1}{1+u^2}$$

und

$$\sin^2\left(\frac{x}{2}\right) = 1 - \cos^2\left(\frac{x}{2}\right) = \frac{u^2}{1+u^2}$$

ist

$$\cos(x) = \cos^2\left(\frac{x}{2}\right) - \sin^2\left(\frac{x}{2}\right) = \frac{1-u^2}{1+u^2}$$

und

$$\sin(x) = 2 \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right) = \frac{2u}{1+u^2}.$$

$$\Rightarrow \int R(\sin x, \cos x) dx = \int \underbrace{R\left(\frac{2u}{1+u^2}, \frac{1-u^2}{1+u^2}\right)}_{\text{Rationale Funktion von } u} \cdot \frac{2}{1+u^2} du.$$

d) Mit $u(x) := \sqrt[k]{ax+b}$ ist $x = (u^k - b)/a$, also

$$\int R(x, \sqrt[k]{ax+b}) dx = \frac{k}{a} \int R\left(\frac{u^k-b}{a}, u\right) u^{k-1} du$$

e) Setze $u(x) := \sqrt[k]{\frac{ax+b}{cx+d}}$.

f) Durch lineare Substitution

$$t(x) := \frac{2ax+b}{\sqrt{|4ac-b^2|}}$$

Rückführung auf einen der Spezialfälle

- $\int R(t, \underbrace{\sqrt{t^2 + 1}}_{\cosh u}) dt \quad t := \sinh u$
- $\int R(t, \underbrace{\sqrt{t^2 - 1}}_{\sinh u}) dt \quad t := \cosh u$
- $\int R(t, \underbrace{\sqrt{1 - t^2}}_{\cos u}) dt \quad t := \sin u$

Weiter mit b) bzw. c). □

1.17 Beispiel Mit der Substitution $u(x) := \tan \frac{x}{2}$ ist

$$\begin{aligned}
 \int \frac{\cot x}{1 + \cos x} dx &= \int \frac{\cos x}{\sin x \cdot (1 + \cos x)} dx = \int \frac{1 - u^2}{2u \cdot \left(1 + \frac{1 - u^2}{1 + u^2}\right)} \cdot \frac{2}{1 + u^2} du \\
 &= \int \frac{1 - u^2}{2u} du = \frac{1}{2} \ln |u| - \frac{u^2}{4} + c \\
 &= \frac{1}{2} \ln \left| \tan \left(\frac{x}{2} \right) \right| - \frac{1}{4} \tan^2 \left(\frac{x}{2} \right) + c. \quad \diamond
 \end{aligned}$$

In der Praxis benutzt man Computeralgebrasysteme für die Integration rationaler (und vieler anderer) Funktionen.

2 Matrizen und Vektorraum-Endomorphismen

2.1 Die Jordansche Normalform

2.1.1 Äquivalenz und Ähnlichkeit

Wie wir wissen, besteht zwischen Matrizen aus $\text{Mat}(n \times m, K)$ und linearen Abbildungen $\varphi : V \rightarrow W$ m - bzw. n -dimensionaler K -Vektorräume V bzw. W ein enger Zusammenhang.

Ist $B = (b_1, \dots, b_m)$ eine Basis von V und $C = (c_1, \dots, c_n)$ eine Basis von W , dann ist die *darstellende Matrix* $\text{Mat}_C^B(\varphi) \in \text{Mat}(n \times m, K)$ diejenige Matrix, bezüglich derer der Vektor $\sum_{i=1}^m v_i b_i \in V$ unter φ auf den Vektor $\sum_{j=1}^n w_j c_j$ mit

$$w_j = \sum_{i=1}^m (\text{Mat}_C^B(\varphi))_{ji} v_i$$

abgebildet wird.

Wir stellen uns nun die Frage, wie wir durch geeignete Wahl der Basen eine besonders einfache Gestalt der darstellenden Matrix erreichen können. Basiswechsel entspricht der Multiplikation mit einer regulären Matrix. Wir definieren daher:

2.1 Definition Zwei Matrizen $A, B \in \text{Mat}(n \times m, K)$ heißen **äquivalent**, wenn es Matrizen $S \in \text{GL}(n, K)$ und $T \in \text{GL}(m, K)$ gibt mit

$$B = SAT.$$

Offensichtlich ist diese Relation eine Äquivalenzrelation auf $\text{Mat}(n \times m, K)$, denn aus der Gruppeneigenschaft von $\text{GL}(r, K)$ folgen die Eigenschaften von Reflexivität, Symmetrie und Transitivität.

Unsere Aufgabe besteht nun darin, aus jeder Äquivalenzklasse einen eindeutigen Repräsentanten auszuwählen. Dieser *Normalform* soll man die Charakteristika der Äquivalenzklasse möglichst sofort ansehen.

Wir kennen schon ein solches Charakteristikum, den Rang, also die Dimension des Bildes der linearen Abbildung $\varphi : K^m \rightarrow K^n, \varphi(x) := Ax$. Denn wegen der Invertierbarkeit von S und T gilt

$$\text{rg}(SAT) = \text{rg}(A).$$

Da Zeilenrang gleich Spaltenrang ist, gilt $0 \leq \text{rg}(A) \leq \min(m, n)$. Besonders einfache Matrizen vom Rang $r \in \{0, \dots, \min(m, n)\}$ sind sicher die

$$D_r \in \text{Mat}(n \times m, K) \quad \text{mit} \quad (D_r)_{ij} = \delta_{ij} \cdot \begin{cases} 1 & , \quad i \leq r \\ 0 & , \quad i > r. \end{cases}$$

Beispielsweise ist $D_r \in \text{Mat}(4 \times 3, K)$ gleich

$$D_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad D_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

2.2 Satz Ist $r := \text{rg}(A)$, dann ist $A \in \text{Mat}(n \times m, K)$ zu D_r äquivalent, aber nicht zu D_s , falls $s \neq r$.

Beweis:

- Der Defekt der linearen Abbildung $\varphi(x) = Ax$ ist $\text{def}(\varphi) = m - r$. Wir ergänzen eine Basis b_{r+1}, \dots, b_m von $\ker(\varphi) \subset K^m$ zu einer Basis $B = (b_1, \dots, b_m)$ und setzen $c_i := \varphi(b_i)$, $i = 1, \dots, r$. Die $c_i \in K^n$ sind linear unabhängig, denn andernfalls wäre $\text{def}(\varphi) > m - r$. Wir ergänzen sie zu einer Basis $C = (c_1, \dots, c_n)$ des K^n .

Es gilt dann $\text{Mat}_C^B(\varphi) = D_r$.

- Da $\text{rg}(D_s) = s \neq r = \text{rg}(A)$, ist A nicht zu D_s äquivalent. □

Solange wir also keine Zusatzstrukturen wie z.B. Skalarprodukte auf den endlich-dimensionalen Vektorräumen V und W berücksichtigen, hat bezüglich geeigneter Basen die lineare Abbildung $\varphi : V \rightarrow W$ die Form D_r . $\text{rg}(\varphi)$ ist mit anderen Worten die einzige Invariante von φ unter Basistransformationen.

2.3 Korollar

Sind V und W endlich-dimensionale K -Vektorräume und $\varphi, \psi \in L(V, W)$ mit $\text{rg}(\varphi) = \text{rg}(\psi)$, dann existieren $\alpha \in \text{GL}(V)$ und $\beta \in \text{GL}(W)$ mit

$$\psi = \beta \circ \varphi \circ \alpha.$$

Ist nun $\varphi \in L(V)$, also ein Endomorphismus, dann wird man sich im Allgemeinen nicht den Luxus zweier Basen auf V leisten. Die Frage, wie durch geschickte Wahl nur *einer* Basis B von V die darstellende Matrix $\text{Mat}_B(\varphi) \equiv \text{Mat}_B^B(\varphi)$ auf eine möglichst einfache Normalform (die so genannte Jordansche Normalform) gebracht werden kann, ist nun schwieriger als die eben untersuchte Problematik.

Insbesondere können wir nicht mehr erwarten, dass diese Normalform immer eine Diagonalmatrix aus $\text{Mat}(n, K)$ ist, und zwar aus zwei, in den der Definition folgenden Beispielen illustrierten, Gründen.

2.4 Definition

- Sind $\varphi, \psi \in L(V)$ und existiert $\beta \in \text{GL}(V)$ mit $\psi = \beta \circ \varphi \circ \beta^{-1}$, dann heißen φ und ψ **ähnlich** oder **konjugiert**.
- Sind $A, B \in \text{Mat}(n, K)$ und existiert $S \in \text{GL}(n, K)$ mit $B = SAS^{-1}$, dann heißen A und B **ähnlich** oder **konjugiert**.

Ähnlichkeit ist wieder eine Äquivalenzrelation, und der Rang verändert sich unter Konjugationen nicht. Allerdings ist diese Zerlegung in Äquivalenzklassen feiner als die vorher untersuchte, denn das *Spektrum* (hier: die Menge der Eigenwerte) ist unter Konjugationen invariant.

2.5 Beispiele (Ähnlichkeit von Matrizen)

1. $n = 1$. Zwei Körperelemente $A, B \in \text{Mat}(1, K) = K$ sind nur dann ähnlich, wenn sie gleich sind.
2. $n = 2$, $A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Zwar ist $\text{rg}(A) = \text{rg}(B) = 1$, A und B sind also äquivalent, aber nicht ähnlich, denn $\text{spek}(A) = \{0, 1\} \neq \text{spek}(B) = \{0\}$.
 B ist zu überhaupt keiner Diagonalmatrix konjugiert, denn diese hätte das gleiche Spektrum, müsste also die Nullmatrix sein, was wegen $\text{rg}(B) = 1$ nicht sein kann.
3. $A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \in \text{Mat}(2, \mathbb{R})$ ist für $\sin \varphi \neq 0$ zu keiner (reellen) Diagonalmatrix konjugiert, denn das charakteristische Polynom $\chi_A(x) = x^2 - 2x \cos(\varphi) + 1$ besitzt dann keine reellen Nullstellen.

Fasst man allerdings A als Element von $\text{Mat}(2, \mathbb{C})$ auf, dann ist A zur Diagonalmatrix $\begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$ konjugiert, wobei $\lambda = \cos \varphi + i \sin \varphi$ und die konjugierende Matrix z.B. $S = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$, also $S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$. \diamond

Das letzte Beispiel zeigt einmal wieder, dass die komplexen Zahlen sich algebraisch einfacher verhalten als die reellen.

2.6 Definition

1. Zerfällt jedes (normierte) Polynom $p \in K[x]$ n -ten Grades ($n \geq 1$) in Linearfaktoren, d.h.

$$p(x) = (x - \lambda_1) \cdot \dots \cdot (x - \lambda_n),$$

dann heißt der Körper K **algebraisch abgeschlossen**.

2. Ein Körper $\overline{K} \supset K$ heißt **algebraischer Abschluss** des Unterkörpers K , wenn \overline{K} algebraisch abgeschlossen ist, es aber keinen algebraisch abgeschlossenen Körper $K' \supset K$ mit $K' \subsetneq \overline{K}$ gibt.

2.7 Satz Jeder Körper K besitzt einen algebraischen Abschluss \overline{K} , und dieser ist (bis auf Isomorphie) eindeutig.

Beweis: Siehe z.B. VAN DER WAERDEN [vW1]. □

2.8 Beispiel (Algebraischer Abschluss von \mathbb{R}) Der algebraische Abschluss $\overline{\mathbb{R}}$ des Körpers $K := \mathbb{R}$ der reellen Zahlen muss \mathbb{C} enthalten, denn in \mathbb{C} zerfällt das Polynom $x^2 + 1 \in \mathbb{R}[x]$ in die Linearfaktoren $(x + i)(x - i)$.

Im Prinzip könnte es nun komplexe Polynome $p \in \mathbb{C}[x]$ geben, die in \mathbb{C} nicht in Linearfaktoren zerfallen. Der (in der Analysis bewiesene!) *Fundamentalsatz der Algebra* besagt aber, dass \mathbb{C} algebraisch abgeschlossen ist. Also ist $\overline{\mathbb{R}} = \mathbb{C}$. ◇

Wir wollen nun zunächst das Normalformproblem für Endomorphismen bzw. quadratische Matrizen über abgeschlossenen Körpern behandeln und danach auf allgemeine Körper, insbesondere \mathbb{R} , zurückkommen:

2.9 Definition Ein Endomorphismus $\varphi \in L(V)$ eines K -Vektorraumes V heißt

- **nilpotent**, wenn es ein $n \in \mathbb{N}$ mit $\varphi^n = 0$ gibt, und das kleinste solche n heißt dann **Nilpotenzindex**,
- **diagonalisierbar**, wenn V eine Basis besitzt, die aus Eigenvektoren von φ besteht.

Eine Matrix $A \in \text{Mat}(n, K)$ heißt

- **nilpotent**, wenn es ein $n \in \mathbb{N}$ mit $A^n = 0$ gibt, und das kleinste solche n heißt dann **Nilpotenzindex**,
- **diagonalisierbar**, wenn sie konjugiert zu einer Diagonalmatrix ist.

2.10 Beispiele (Nilpotenz und Diagonalisierbarkeit)

1. Für $n \in \mathbb{N}_0$ und $V := \{p \in \mathbb{R}[x] \mid \deg(p) \leq n\}$ ist der Endomorphismus $\varphi(p) := p'$, der einem Polynom p seine Ableitung p' zuordnet, nilpotent mit Nilpotenzindex $n + 1$, denn $\varphi^{n+1} = 0$, aber $\varphi^n(p) = n!$ für $p(x) = x^n$. φ ist aber (außer für $n = 0$) nicht diagonalisierbar, denn nur die konstanten Polynome $x \mapsto p(x) = a \in \mathbb{R} \setminus \{0\}$ sind Eigenvektoren von φ .

2. Über dem gleichen Vektorraum ist der Endomorphismus

$$\varphi \in L(V) \quad , \quad \varphi(p)(x) := x^n p(1/x)$$

nicht nilpotent, aber diagonalisierbar, denn φ besitzt die Basis von Eigenvektoren der Form $x^l \pm x^{n-l}$ mit den Eigenwerten ± 1 (dabei ist $l \in \{0, \dots, [n/2]\}$; für gerade n lässt man den Nullvektor $x^{n/2} - x^{n/2}$ weg).

3. Über dem gleichen Vektorraum ist der Endomorphismus

$$\varphi \in L(V) \quad , \quad \varphi(p)(x) := p(x + 1)$$

weder nilpotent (denn er erhält den Grad von p) noch für $n > 0$ diagonalisierbar (denn die konstanten Polynome sind die einzigen Eigenvektoren von φ).

◇

Die nilpotenten und die diagonalisierbaren Endomorphismen bzw. Matrizen werden die Bausteine sein, aus denen wir die Jordan-Normalform konstruieren.

Als Normalform einer diagonalisierbaren Matrix A wählen wir eine Diagonalmatrix, zu der A konjugiert ist. Diese besitzt die Eigenwerte von A als Diagonalelemente und ist somit bis auf Permutation der Diagonaleinträge eindeutig durch A bestimmt.

2.1.2 Nilpotente Endomorphismen

Um Normalformen für nilpotente Matrizen zu konstruieren, definieren wir für $r \in \mathbb{N}$ die nilpotente Matrix

$$N_r \in \text{Mat}(r, K) \quad , \quad (N_r)_{ik} := \delta_{i,k-1}.$$

Diese besitzt also in der oberen Nebendiagonale lauter Einsen, und alle anderen Einträge sind Null.

Der Nilpotenzindex von N_r ist r denn $(N_r)^l$ besitzt außer Einsen in der l -ten oberen Nebendiagonale nur Nullen, z.B.

$$N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad (N_3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad (N_3)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Durch direkte Summenbildung können wir nilpotente Matrizen der Form

$$\bigoplus_{i=1}^j N_{r_i} \in \text{Mat}(r, K) \quad (2.1)$$

mit $r_1, \dots, r_j \in \mathbb{N}$ bilden, wobei $r := \sum_{i=1}^j r_i$ ist.

2.11 Beispiel $N_3 \oplus N_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $(N_3 \oplus N_2)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. \diamond

Deren Nilpotenzindex ist aber nicht etwa r , sondern nur $\max(r_1, \dots, r_j)$, denn allgemein gilt für Potenzen direkter Summen von quadratischen Matrizen

$$\left(\bigoplus_{i=1}^j A_i \right)^k = \bigoplus_{i=1}^j (A_i)^k.$$

Wir wollen nun zeigen, dass jede nilpotente Matrix zu einer Matrix der Form (2.1) konjugiert ist, sodass wir letztere als Normalformen wählen können. Um die Geometrie des Problems herauszuarbeiten, wechseln wir zu nilpotenten Endomorphismen über.

Schreiben wir

$$\tilde{N}_r \in L(K^r) \quad , \quad \tilde{N}_r(x) := N_r x$$

für die der Matrix N_r zugeordnete lineare Abbildung, dann fällt auf, dass wegen $\tilde{N}_r(e_1) = 0$ und $\tilde{N}_r(e_s) = e_{s-1}$ ($s = 2, \dots, r$) für die kanonische Basis (e_1, \dots, e_r) des K^r dieser Vektorraum die Basis

$$(e_r, \tilde{N}_r(e_r), \tilde{N}_r^2(e_r), \dots, \tilde{N}_r^{r-1}(e_r))$$

besitzt, wir also eine Basis des Vektorraumes durch wiederholte Anwendung des nilpotenten Endomorphismus auf einen einzigen Vektor $w \in K^r$ gewinnen können. Dabei hätten wir statt $w = e_r$ einen beliebigen Vektor $w \in K^r \setminus \text{im}(\tilde{N}_r)$ mit nicht verschwindender r -ter Komponente zum Ausgangspunkt wählen können.

2.12 Definition Ein r -dimensionaler Unterraum $W \subset V$ heißt **zyklisch** bezüglich des Endomorphismus $\varphi \in L(V)$, wenn es einen Vektor $w \in W$ gibt, sodass

- $(w, \varphi(w), \varphi^2(w), \dots, \varphi^{r-1}(w))$ eine Basis von W ist und
- $\varphi^r(w) = 0$ ist.

w heißt dann **erzeugender Vektor**.

2.13 Beispiele (Erzeugende Vektoren)

1. Bezüglich des Ableitungsoperators $\varphi(p) := p'$ auf dem Polynomraum $V := \mathbb{R}[x]$ ist jeder Unterraum $W := \{p \in V \mid \deg(p) \leq n\}$ zyklisch, wobei jeder Vektor w mit $\deg(w) = n$ erzeugend ist.

Wählen wir als erzeugenden Vektor v das Monom $v(x) = x^n$, dann ergibt sich als Basis

$$(v^{(n)}, v^{(n-1)}, \dots, v^{(1)}, v) = \left(n!, \frac{n!}{1!}x, \dots, \frac{n!}{(n-1)!}x^{n-1}, x^n \right).$$

In dieser Basis besitzt $\varphi|_W$ die Form \tilde{N}_n , also die Normalform.

2. Bezüglich des Operators $\bigoplus_{i=1}^j \tilde{N}_{r_i}$ auf dem Summenraum $V := \bigoplus_{i=1}^j W_i$ mit $W_i := K^{r_i}$ ist der Vektor $0 \oplus \dots \oplus 0 \oplus e_{r_i} \oplus 0 \oplus \dots \oplus 0$ (mit e_{r_i} an der i -ten Stelle) erzeugender Vektor von W_i . \diamond

Nehmen wir einmal an, wir wüssten schon, dass jede nilpotente Matrix $A \in \text{Mat}(r, K)$ konjugiert zu einer Normalform (2.1) ist, dann müssten wir nur j zyklische Vektoren finden, um die konjugierende Matrix S zu berechnen. Wir gehen wieder zum geometrischen Problem nilpotenter Endomorphismen $\varphi \in L(V)$ über. Was wir zunächst berechnen können, sind die Kerne der iterierten Abbildungen φ^k , und tatsächlich reicht das aus.

So ist jeder Vektor $w \in W \setminus \ker(\varphi^{n-1})$ in Beispiel 2.10.1. erzeugender Vektor.

Im allgemeinen Fall ist zunächst zu beweisen, dass V direkte Summe von Unterräumen $V = \bigoplus_{i=1}^j W_i$ ist, sodass die W_i bezüglich φ zyklisch sind.

Die vorhandenen Daten sind zunächst die Dimensionen

$$d_p := \text{def}(\varphi^p) \equiv \dim(V_p) \text{ der Kerne } V_p := \ker(\varphi^p) \quad (p = 0, \dots, r+1).$$

Aus diesen lassen sich die Dimensionen der zyklischen Unterräume W_i berechnen.

Ist q der Nilpotenzindex von φ , dann gilt

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_q = V,$$

und $d_{p+1} > d_p$ für $p < q$, denn wären für ein $p < q$ die Dimensionen von V_p und von V_{p+1} gleich, dann wäre auch $V_p = V_{p+1}$, also auch $V_p = \dots = V_q$.

2.14 Lemma

- Es ist $V_{p+1} = \varphi^{-1}(V_p) \quad (p = 0, \dots, q-1)$.

- φ erzeugt injektive Abbildungen

$$\bar{\varphi}_p : \bar{V}_{p+1} \rightarrow \bar{V}_p \quad v + V_p \mapsto \varphi(v) + V_{p-1}$$

der Quotientenräume $\bar{V}_p := V_p/V_{p-1}$.

Mit anderen Worten:

Sind $v, v' \in V_{p+1}$ und $v - v' \notin V_p$, dann ist auch $\varphi(v) - \varphi(v') \notin V_{p-1}$.

Beweis:

- $V_{p+1} \equiv \ker(\varphi^{p+1}) = \ker(\varphi \circ \varphi^p) = \{v \in V \mid \varphi(v) \in V_p\} = \varphi^{-1}(V_p)$.
- Es sei $w \in V_{p+1}$. Wir müssen zeigen: Ist $\varphi(w) \in V_{p-1}$, dann ist $w \in V_p$.

Der erste Teil des Lemmas besagt aber gerade $V_p = \varphi^{-1}(V_{p-1})$. □

Aus $V_{p+1} = \varphi^{-1}(V_p)$ folgt nur $\varphi(V_{p+1}) \subset V_p$ und nicht etwa $\varphi(V_{p+1}) = V_p$. Die Abbildungen $\bar{\varphi}_p$ sind also im Allgemeinen nicht surjektiv, bilden also \bar{V}_{p+1} nur injektiv in einen Unterraum von \bar{V}_p ab.

Setzen wir nun

$$\bar{d}_p := d_p - d_{p-1} = \dim(V_p) - \dim(V_{p-1}) = \dim(\bar{V}_p), \quad (2.2)$$

dann ist

$$k_p := \bar{d}_p - \bar{d}_{p+1} = \dim(\bar{V}_p) - \dim(\bar{V}_{p+1}) \quad (2.3)$$

die Kodimension dieses Bildes $\bar{\varphi}_p(\bar{V}_{p+1})$ in \bar{V}_p .

Diese Dimensionszählerei nützt uns jetzt bei der Konstruktion einer Basis von V , bezüglich derer φ die nilpotente Normalform (2.1) besitzt.

Wir wählen nämlich für $p = 1, \dots, q$ Vektoren

$$e_{p,1}, \dots, e_{p,k_p} \in V_p$$

so, dass sie Repräsentanten einer Basis von $\bar{V}_p/\bar{\varphi}_p(\bar{V}_{p+1})$ sind. Dies ist möglich, denn ihre Anzahl k_p ist gleich der Dimension von $\bar{V}_p/\bar{\varphi}_p(\bar{V}_{p+1})$.

Diese Vektoren $e_{p,j}$ nennen wir *primitiv*, und leiten aus ihnen die Vektoren

$$\varphi^r(e_{p,j}) \quad (r = 0, \dots, p-1; j = 1, \dots, k_p; p = 1, \dots, q)$$

ab. Deren Zahl ist gleich

$$\sum_{p=1}^q p \cdot k_p = \sum_{p=1}^q p \cdot (\bar{d}_p - \bar{d}_{p+1}) = \sum_{p=1}^q \bar{d}_p = \sum_{p=1}^q (d_p - d_{p-1}) = d_q = \dim(V),$$

und tatsächlich bilden sie eine Basis von V .

2.15 Satz 1. Für $p = 1, \dots, q$ repräsentieren die Vektoren

$$\varphi^{p'-p}(e_{p',j}) \quad (p' = p, \dots, q; j = 1, \dots, k_{p'})$$

eine Basis von $\bar{V}_p \equiv V_p/V_{p-1}$.

2. Die Vektoren

$$\varphi^r(e_{p,j}) \quad (r = 0, \dots, p-1; j = 1, \dots, k_p; p = 1, \dots, q)$$

bilden eine Basis von V .

3. Die Unterräume

$$W_{p,j} := \text{span}(e_{p,j}, \dots, \varphi^{p-1}(e_{p,j})) \quad (j = 1, \dots, k_p; p = 1, \dots, q)$$

sind zyklisch bezüglich φ mit erzeugendem Vektor $e_{p,j}$ und besitzen Dimension p .

4. V ist die innere direkte Summe dieser Unterräume

$$V = \bigoplus_{\substack{p=1, \dots, q \\ j=1, \dots, k_p}} W_{p,j}.$$

Beweis: (Siehe auch BRIESKORN, [Bri, p. 15]):

- Für $p = q$ ist $\bar{V}_{p+1} = V/V$ nulldimensional; also repräsentieren $e_{p,1}, \dots, e_{p,k_p}$ eine Basis von $\bar{V}_p/\varphi_p(\bar{V}_{p+1}) = \bar{V}_p = V_p/V_{p-1}$.
- Behauptung 1. sei schon für $p+1 \leq q$ bewiesen, und wir wollen sie für p beweisen. Wir wissen also, dass die Vektoren

$$\varphi^{p'-(p+1)}(e_{p',j}) \quad (p' = p+1, \dots, q; j = 1, \dots, k_{p'})$$

eine Basis von \bar{V}_{p+1} repräsentieren. Wegen der im letzten Lemma bewiesenen Injektivität von $\bar{\varphi}_p$ repräsentieren ihre Bilder unter φ eine Basis von $\bar{\varphi}_p(\bar{V}_{p+1})$. Die primitiven Vektoren $e_{p,1}, \dots, e_{p,k_p} \in V_p$ sind nun gerade so gewählt, dass sie diese zu einer Basis von \bar{V}_p ergänzen.

- Wären diese Vektoren linear abhängig, gäbe es also eine Linearkombination der $\varphi^r(e_{p,j})$, die eine nichttriviale Darstellung des Nullvektors wäre, dann würden wir einen Koeffizienten $\lambda_{p,j}^{(r)} \neq 0$ mit maximalem $p-r =: p'$ finden. Wir könnten also diese Darstellung in der Form

$$\sum_{p-r=p'} \lambda_{p,j}^{(r)} \varphi^r(e_{p,j}) = - \sum_{p-r < p'} \lambda_{p,j}^{(r)} \varphi^r(e_{p,j})$$

schreiben. Die linke Seite repräsentiert einen Vektor aus $\bar{V}_{p'}$, die rechte Seite ist ein Vektor aus $V_{p'-1}$. Da die Vektoren $\varphi^r(e_{p,j})$ der linken Seite eine Basis von $\bar{V}_{p'} = V_{p'}/V_{p'-1}$ repräsentieren, müssen ihre Koeffizienten alle verschwinden. Widerspruch!

- Nach 2. sind die Vektoren $e_{p,j}, \dots, \varphi^{p-1}(e_{p,j})$ linear unabhängig und bilden daher eine Basis von $W_{p,j}$.
- folgt, da für $(p, j) \neq (p', j')$ die Unterräume $W_{p,j}$ und $W_{p',j'}$ von verschiedenen Basisvektoren aufgespannt werden. \square

Da die auf die Unterräume $W_{p,j}$ restringierte Abbildung $\varphi|_{W_{p,j}}$ in der geordneten Basis

$$(\varphi^{p-1}(e_{p,j}), \varphi^{p-2}(e_{p,j}), \dots, \varphi(e_{p,j}), e_{p,j})$$

die Matrixdarstellung N_p besitzt, haben wir nebenbei gezeigt, dass φ zur nilpotenten Jordan-Normalform

$$\bigoplus_{\substack{p=1, \dots, q \\ j=1, \dots, k_p}} \tilde{N}_p$$

konjugiert ist.

2.16 Beispiel (Nilpotente Jordan-Normalform)

Auf dem Vektorraum $\mathbb{R}[x]$ der Polynome betrachten wir den Operator

$$\hat{\varphi}(p)(x) := p'(x) - xp''(x) + p^{(3)}(x),$$

wobei $p^{(n)}$ die n -te Ableitung des Polynoms p bezeichnet. Ist eine Lösung p der Gleichung $\hat{\varphi}(p) = 0$ gesucht (wobei p kein Polynom mehr sein muss), dann spricht man von einer linearen Differentialgleichung. Lineare Differentialgleichungen, allerdings mit konstanten Koeffizienten, werden uns noch in dieser Vorlesung begegnen. $\hat{\varphi}$ selbst ist nicht nilpotent, aber da

$$\hat{\varphi}^2(p)(x) = (x^2 - 1)p^{(4)}(x) - 2xp^{(5)}(x) + p^{(6)}(x)$$

ist, verschwindet schon die dritte Iterierte φ^3 von $\varphi := \hat{\varphi}|_V$ auf $V := \{p \in \mathbb{R}[x] \mid \deg(p) \leq 4\}$. Mit den Basisvektoren $e_i(x) := x^i$ von V ergibt sich

$$V_0 = \{0\}, \quad V_1 = \text{span}(e_0, e_2), \quad V_2 = \text{span}(e_0, e_1, e_2, e_3), \quad V_3 = V.$$

Daher repräsentiert e_4 eine Basis von $\bar{V}_3 = V_3/V_2$, e_1, e_3 eine Basis von $\bar{V}_2 = V_2/V_1$ und e_0, e_2 eine Basis von $\bar{V}_1 = V_1$. Wir wählen also als primitive Vektoren $e_{3,1} := e_4$, $e_{2,1} := e_1$, sodass $\varphi(e_{3,1}) = 24e_1 - 8e_3$, $\varphi^2(e_{3,1}) = 24(e_2 - e_0)$ und $\varphi(e_{2,1}) = e_0$ ist. Somit besitzt φ die Normalform $\tilde{N}_3 \oplus \tilde{N}_2$. \diamond

Wir können durch Ummummern der Basisvektoren immer erreichen, dass in der nilpotenten Normalform

$$\tilde{N}_{r_1} \oplus \tilde{N}_{r_2} \oplus \dots \oplus \tilde{N}_{r_j} \quad r_1 \geq r_2 \geq \dots \geq r_j$$

gilt. Die Summe $r := r_1 + r_2 + \dots + r_j$ dieser Nilpotenzindizes ist gleich der Dimension des Vektorraumes V . Eine solche Zerlegung einer natürlichen Zahl $r \in \mathbb{N}$ in absteigend geordnete Summanden $r_i \in \mathbb{N}$ heißt eine *(Zahl-)Partition* von r , und $p(r)$ bezeichnet die Anzahl voneinander verschiedener Partitionen von r .

2.17 Beispiele (Zahlpartitionen)

1. $p(1) = 1$.
2. Da $2 = 1 + 1$, ist $p(2) = 2$.
3. Da $3 = 2 + 1 = 1 + 1 + 1$, ist $p(3) = 3$.
4. Da $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$, ist $p(4) = 5$. ◇

2.18 Satz Auf einem r -dimensionalen Vektorraum V ist die Zahl der Äquivalenzklassen konjugierter nilpotenter Endomorphismen gleich $p(r)$, und diese besitzen die Normalformen

$$\tilde{N}_{r_1} \oplus \tilde{N}_{r_2} \oplus \dots \oplus \tilde{N}_{r_j} \quad (2.4)$$

mit $r_1 \geq r_2 \geq \dots \geq r_j \geq 1$ und $\sum_{i=1}^j r_i = r$.

Beweis:

- Dass jeder nilpotente Endomorphismus $\varphi \in L(V)$ zu einer dieser Normalformen konjugiert ist, haben wir schon gezeigt.
- Es ist also nur noch zu zeigen, dass zwei verschiedene solche Normalformen (2.4) nicht konjugiert sind. Nun läßt sich aber die Zahl k_p der Jordan-Blöcke \tilde{N}_{r_i} mit $r_i = p$ nach (2.2) und (2.3) aus den Dimensionen d_p der Kerne V_p von (2.4) berechnen. Umgekehrt ist

$$d_p = \sum_{l=1}^p \bar{d}_l \quad \text{und} \quad \bar{d}_l = \sum_{s=l}^q k_s,$$

wir können die Defekte d_p also aus den Multiplizitäten k_s der Jordan-Blöcke der Größe s berechnen. Unter Konjugation bleiben die d_p , also auch die k_s invariant. Die absteigende Anordnung der Blockgrößen r_i impliziert daher die Eindeutigkeit. □

2.1.3 Haupträume

Bis jetzt haben wir das Normalformproblem für zwei Klassen von Endomorphismen $\varphi \in L(V)$ eines r -dimensionalen K -Vektorraumes V gelöst:

- für diagonalisierbare φ ,
- für nilpotente φ .

Nun wollen wir diese Informationen nutzen, um einen beliebigen Endomorphismus φ (eindeutig) als Summe

$$\varphi = \varphi_D + \varphi_N \quad (2.5)$$

eines diagonalisierbaren und eines nilpotenten Endomorphismus darzustellen, die miteinander kommutieren ($\varphi_D \circ \varphi_N = \varphi_N \circ \varphi_D$).

Da nilpotente $\varphi \in L(V)$ die Normalform (2.1) besitzen, ist ihr charakteristisches Polynom gleich

$$\chi_\varphi(x) = x^r.$$

Dagegen besitzen diagonalisierbare $\varphi \in L(V)$ das charakteristische Polynom

$$\chi_\varphi(x) = \prod_{\lambda \in \text{spek}(\varphi)} (x - \lambda)^{m_\lambda}, \quad (2.6)$$

wobei m_λ die Multiplizität des Eigenwertes λ ist, d.h. $m_\lambda := \text{def}(\varphi - \lambda \cdot \text{Id})$.

Wir benutzen daher das charakteristische Polynom dazu, die Zerlegung (2.5) zu finden.

2.19 Definition Es sei $\varphi \in L(V)$ Endomorphismus des K -Vektorraumes V und $\lambda \in K$ Eigenwert von φ .

- Dann heißt $v \in V \setminus \{0\}$ **Hauptvektor** (oder **verallgemeinerter Eigenvektor**) von φ zum Eigenwert λ , wenn ein $q \in \mathbb{N}$ existiert sodass

$$(\varphi - \lambda \cdot \text{Id})^q v = 0.$$

- $\hat{E}(\lambda) := \bigcup_{q \in \mathbb{N}} \ker((\varphi - \lambda \cdot \text{Id})^q)$ heißt **Hauptraum** (oder **verallgemeinerter Eigenraum**) von φ zum Eigenwert λ .

Für diagonalisierbare φ stimmen die Eigenräume $E(\lambda) = \ker(\varphi - \lambda \text{Id})$ mit den Haupträumen $\hat{E}(\lambda)$ überein, für nilpotente φ dagegen ist ganz V Hauptraum (zum Eigenwert 0).

2.20 Definition Es sei $\varphi \in L(V)$. Dann heißt

- ein Unterraum $U \subset V$ φ -invariant, wenn $\varphi(U) \subset U$,
- eine direkte Zerlegung $V = \bigoplus_{i \in I} U_i$ φ -invariant, wenn alle Unterräume U_i φ -invariant sind.

2.21 Satz Es sei $\psi \in L(V)$. Gilt für ein $q \in \mathbb{N}_0$

$$\ker(\psi^q) = \ker(\psi^{q+1}) \quad (2.7)$$

(mit $\psi^0 = \text{Id}$), dann gilt für alle $p \geq q$

$$\ker(\psi^p) = \ker(\psi^{p+1}), \quad (2.8)$$

im und falls $\dim(V) < \infty$

$$V = \ker(\psi^q) \oplus \text{im}(\psi^q).$$

Diese Zerlegung ist dann ψ -invariant.

Beweis:

- Wir zeigen, dass aus (2.8) auch

$$\ker(\psi^{p+1}) = \ker(\psi^{p+2})$$

folgt. Ist $v \in \ker(\psi^{p+2})$, dann ist $w := \psi(v) \in \ker(\psi^{p+1}) = \ker(\psi^p)$, also $\psi^p(w) = \psi^{p+1}(v) = 0$.

- Es sei $D := \ker(\psi^q) \cap \text{im}(\psi^q)$. Dann ist $\psi^q(D) = \{0\}$, und $D = \psi^q(D')$ für $D' := (\psi^q)^{-1}(D)$. Damit ist $D' \subset \ker(\psi^{2q}) = \ker(\psi^q)$, also $D = \{0\}$.
- Da die Summe von Kern und Bild also direkt ist, gilt nach dem Dimensionssatz

$$\begin{aligned} \dim(\ker(\psi^q) \oplus \text{im}(\psi^q)) &= \dim(\ker(\psi^q)) + \dim(\text{im}(\psi^q)) \\ &= \dim(V) < \infty, \end{aligned}$$

sodass die direkte Summe schon gleich dem gesamten Vektorraum V ist.

- Für beliebige $i \in \mathbb{N}$ gilt

$$\begin{aligned} \psi(\ker(\psi^i)) &= \{\psi(v) \mid \psi^{i-1}(\psi(v)) = 0\} \subset \{u \in V \mid \psi^{i-1}(u) = 0\} \\ &= \ker(\psi^{i-1}) \subset \ker(\psi^i) \end{aligned}$$

und

$$\psi(\text{im}(\psi^i)) = \psi(\psi^i(V)) = \psi^i(\psi(V)) \subset \psi^i(V) = \text{im}(\psi^i). \quad \square$$

2.22 Bemerkungen

1. Ist $\dim(V) = \infty$, dann kann $\ker(\psi^q) \oplus \text{im}(\psi^q)$ ein echter Unterraum von V sein. Z.B. ist dies für $V := \mathbb{R}[x]$ und $\varphi \in L(V)$, $\varphi(p)(x) := xp(x)$ der Fall, denn φ ist injektiv, also ist (2.7) für $q = 0$ und damit auch für $q = 1$ erfüllt, aber $\text{im}(\varphi)$ ist ein echter Unterraum von V , da er die Monome vom Grad 0 nicht enthält.
2. Den kleinsten Index $q \in \mathbb{N}_0$ (soweit vorhanden), für den (2.7) erfüllt ist, nennt man den *Fitting-Index* von φ . \diamond

2.23 Korollar Ist $r := \dim(V) < \infty$, dann ist der Fitting-Index $q \leq r$.

Beweis: $0 = \text{def}(\psi^0) < \text{def}(\psi^1) < \dots < \text{def}(\psi^q) \leq r$. \square

Wir wollen nun zunächst zeigen, dass endlich-dimensionale Vektorräume V sich als direkte Summe

$$V = \bigoplus_{\lambda \in \text{spek}(\varphi)} \hat{E}(\lambda)$$

der verallgemeinerten Eigenräume zerlegen lassen.

Dies ist nur möglich, wenn das charakteristische Polynom in Linearfaktoren zerfällt, also die Form (2.6) besitzt. Wir wissen schon, dass dies immer der Fall ist, wenn der Körper K algebraisch abgeschlossen ist, also z.B. für $K = \mathbb{C}$.

2.24 Satz Wenn das charakteristische Polynom $\chi_\varphi \in K[x]$ des Endomorphismus $\varphi \in L(V)$ auf dem endlich-dimensionalen K -Vektorraum V in Linearfaktoren zerfällt, dann ist

$$V = \bigoplus_{\lambda \in \text{spek}(\varphi)} \hat{E}(\lambda), \quad (2.9)$$

und die Zerlegung (2.9) ist φ -invariant (es gilt also $\varphi(\hat{E}(\lambda)) \subset \hat{E}(\lambda)$).

Beweis:

- Für $r := \dim(V)$ und $v \in \hat{E}(\lambda)$ gilt

$$(\varphi - \lambda \text{Id})^r \circ \varphi(v) = \varphi \circ (\varphi - \lambda \text{Id})^r(v) = \varphi(0) = 0,$$

also $\varphi(\hat{E}(\lambda)) \subset \hat{E}(\lambda)$.

- Für $\lambda, \mu \in \text{spek}(\varphi)$ mit $\lambda \neq \mu$ gilt

$$\hat{E}(\lambda) \cap \hat{E}(\mu) = \{0\}.$$

Denn der Unterraum $U := \hat{E}(\lambda) \cap \hat{E}(\mu)$ ist, wie wir gerade gesehen haben, invariant unter φ .

- Nun besitzt für $\dim(U) > 0$ der restringierte Endomorphismus $\hat{\varphi} := \varphi|_U$ den Eigenwert λ , denn $(\hat{\varphi} - \lambda \text{Id})^r U = \{0\}$.
 - Andererseits besitzt $\hat{\varphi}$ keinen weiteren Eigenwert $\sigma \neq \lambda$, denn mit $(\hat{\varphi} - \sigma \text{Id})v = 0$ für $v \in U \setminus \{0\}$ würde $(\hat{\varphi} - \lambda \text{Id})^r v = (\hat{\varphi} - \lambda \text{Id})^{r-1} \cdot (\sigma - \lambda)v = \dots = (\sigma - \lambda)^r v \neq 0$ folgen.
 - Aus Symmetriegründen müsste aber auch μ der einzige Eigenwert von $\hat{\varphi}$ sein, was die Annahme $\dim(U) > 0$ widerlegt.
- Es seien nun $\lambda_1, \dots, \lambda_k \in \text{spek}(\varphi)$ voneinander verschiedene Eigenwerte.

Wir wollen zeigen, dass die Summe

$$\hat{E}(\lambda_1) + \dots + \hat{E}(\lambda_k)$$

der verallgemeinerten Eigenräume direkt ist, d.h. aus $v_{\lambda_1} + \dots + v_{\lambda_k} = 0$ für $v_{\lambda_i} \in \hat{E}(\lambda_i)$ folgt, dass $v_{\lambda_i} = 0$. Dies wurde gerade für $k = 2$ bewiesen.

Der Induktionsschluss von $k - 1$ auf k geht wie folgt: Es ist

$$(\varphi - \lambda_k \text{Id})^r v_{\lambda_k} = 0,$$

also auch

$$\sum_{i=1}^{k-1} (\varphi - \lambda_k \text{Id})^r v_{\lambda_i} = 0. \quad (2.10)$$

Wegen $\varphi(\hat{E}(\lambda_i)) \subset \hat{E}(\lambda_i)$ ist auch $(\varphi - \lambda_k \text{Id})^r \hat{E}(\lambda_i) \subset \hat{E}(\lambda_i)$. Damit müssen für $i = 1, \dots, k - 1$ diese Vektoren in (2.10) die Nullvektoren sein, denn die Summe $\hat{E}(\lambda_1) + \dots + \hat{E}(\lambda_{k-1})$ ist nach Induktionsvoraussetzung direkt. Es ist also

$$v_{\lambda_i} \in \hat{E}(\lambda_i) \cap \hat{E}(\lambda_k) = \{0\} \quad (i = 1, \dots, k - 1)$$

und damit auch $v_k = -(v_1 + \dots + v_{k-1}) = 0$.

- Wir müssen noch zeigen, dass die direkte Summe

$$\bigoplus_{\lambda \in \text{spek}(\varphi)} \hat{E}(\lambda) \subset V$$

gleich V ist. Dies ist für Dimension $r = 1$ der Fall und sei allgemein für Dimensionen $\leq r - 1$ gezeigt.

Nach Satz 2.21 ist für $\lambda \in \text{spek}(\varphi)$ die Zerlegung

$$V = \hat{E}(\lambda) \oplus \text{im}((\varphi - \lambda \text{Id})^r) \quad (2.11)$$

φ -invariant. Daher können wir (falls nicht schon $\hat{E}(\lambda) = V$ gilt) $\hat{\varphi} := \varphi|_I$ auf dem Bildraum $I := \text{im}((\varphi - \lambda \cdot \text{Id})^r)$ betrachten. Da $\dim(I) < \dim(V)$, besitzt I nach Induktionsvoraussetzung eine $\hat{\varphi}$ -invariante Zerlegung

$$I = \bigoplus_{\mu \in \text{spek}(\hat{\varphi})} I_\mu \quad (2.12)$$

in Haupträume I_μ von $\hat{\varphi}$.

Es ist nach dem folgenden Lemma $\text{spek}(\hat{\varphi}) = \text{spek}(\varphi) \setminus \{\lambda\}$.

Daraus folgt aber

$$I_\mu \subset \hat{E}(\mu) \quad (\mu \in \text{spek}(\varphi)), \quad (2.13)$$

denn wegen der $\hat{\varphi}$ -Invarianz von I ist

$$\{0\} = (\hat{\varphi} - \mu \cdot \text{Id})^r(I_\mu) = (\varphi - \mu \cdot \text{Id})^r(I_\mu).$$

Insgesamt folgt aus (2.11), (2.12) und (2.13)

$$V = \hat{E}(\lambda) \oplus I = \hat{E}(\lambda) \oplus \bigoplus_{\mu \in \text{spek}(\hat{\varphi})} I_\mu = \hat{E}(\lambda) \oplus \bigoplus_{\mu \in \text{spek}(\varphi) \setminus \{\lambda\}} \hat{E}(\mu) = \bigoplus_{\mu \in \text{spek}(\varphi)} \hat{E}(\mu). \quad \square$$

2.25 Lemma

Es sei $\varphi \in L(V)$, $\dim(V) < \infty$ und $U \subset V$ ein φ -invarianter Unterraum.

Dann ist für $\hat{\varphi} := \varphi|_U$ das charakteristische Polynom $\chi_{\hat{\varphi}}$ ein Teiler von χ_φ . Zerfällt also χ_φ , dann auch $\chi_{\hat{\varphi}}$.

Beweis: Ist (b_1, \dots, b_k) eine Basis von U und $B = (b_1, \dots, b_k, \dots, b_r)$ eine Basis von V , dann ist die darstellende Matrix von φ von der Blockform

$$\text{Mat}_B(\varphi) = M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad \text{mit} \quad M_{21} = 0$$

und $M_{11} \in \text{Mat}(k, K)$, $M_{22} \in \text{Mat}(r - k, K)$.

Die gleiche Blockstruktur besitzt $M - x \cdot \mathbb{1}_r$ für $x \in K$. Daher gilt (Übungsaufgabe)

$$\chi_M(x) = \chi_{M_{11}}(x) \cdot \chi_{M_{22}}(x) \quad (x \in K).$$

Wegen $\chi_M = \chi_\varphi$ und $\chi_{M_{11}} = \chi_{\hat{\varphi}}$ folgt daraus die Behauptung. \square

2.1.4 Jordan-Zerlegung und Jordansche Normalform

Satz 2.24 ermöglicht es uns, $\varphi \in L(V)$ in der Form (2.5) als Summe $\varphi = \varphi_D + \varphi_N$ eines diagonalisierbaren und eines nilpotenten Endomorphismus zu schreiben.

Es sei dazu f_λ der Fitting-Index von $\varphi - \lambda \text{Id}$ und

$$R_\lambda \in K[x] \quad \text{das Polynom} \quad R_\lambda(x) := \prod_{\mu \in \text{spek}(\varphi) \setminus \{\lambda\}} (x - \mu)^{f_\mu}. \quad (2.14)$$

Setzen wir φ in diese Polynome ein, dann ist

$$R_\lambda(\varphi)(V) \subset \hat{E}(\lambda),$$

denn V besitzt die φ -invariante Zerlegung (2.12), und $(\varphi - \mu \text{Id})^{f_\mu}(\hat{E}(\mu)) = \{0\}$. Andererseits sind die Polynome $R_\lambda \in K[x]$ *teilerfremd*, d.h. es existiert kein Polynom $R \in K[x]$ vom Grad > 0 , das alle R_λ ohne Rest teilt.

Mithilfe des Euklidischen Algorithmus beweist man (Übungsaufgabe) die Existenz von Polynomen $Q_\lambda \in K[x]$ mit

$$\sum_{\lambda \in \text{spek}(\varphi)} Q_\lambda \cdot R_\lambda = 1. \quad (2.15)$$

2.26 Satz (Jordan-Zerlegung) *Das charakteristische Polynom von $\varphi \in L(V)$ zerfalle in Linearfaktoren. Dann bilden die Endomorphismen*

$$P_\lambda := Q_\lambda(\varphi) \cdot R_\lambda(\varphi)$$

(mit den Polynomen R_λ aus (2.14) und Q_λ aus (2.15)) eine Familie von Projektoren auf die verallgemeinerten Eigenräume $\hat{E}(\lambda)$ mit

$$P_\lambda P_\mu = \delta_{\lambda\mu} P_\lambda \quad \text{und} \quad \sum_{\lambda \in \text{spek}(\varphi)} P_\lambda = \text{Id}_V. \quad (2.16)$$

Weiter gilt:

1. Der Endomorphismus

$$\varphi_D := \sum_{\lambda \in \text{spek}(\varphi)} \lambda P_\lambda$$

ist diagonalisierbar, und seine Eigenräume zu den Eigenwerten λ stimmen mit den verallgemeinerten Eigenräumen $\hat{E}(\lambda)$ von φ zu den gleichen Eigenwerten überein.

2. Der Endomorphismus $\varphi_N := \varphi - \varphi_D$ lässt die Zerlegung (2.9) von V in die verallgemeinerten Eigenräume $\hat{E}(\lambda)$ invariant, und $\varphi_N|_{\hat{E}(\lambda)}$ besitzt Nilpotenzindex f_λ .
3. Es gilt $\varphi_D \circ \varphi_N = \varphi_N \circ \varphi_D$, d.h. φ_D und φ_N kommutieren.
4. $\varphi = \varphi_D + \varphi_N$ ist die einzige Zerlegung von φ in einen diagonalisierbaren und einen nilpotenten Endomorphismus, die miteinander kommutieren.

Beweis: • Die zweite Identität aus (2.16) folgt aus (2.15).

• Dass $P_\lambda(V) \subset \hat{E}(\lambda)$ ist, folgt wegen $R_\lambda(\varphi)(V) \subset \hat{E}(\lambda)$ und der φ -Invarianz der Haupträume. Da andererseits

$$\sum_{\lambda \in \text{spek}(\varphi)} P_\lambda = \text{Id}_V$$

ist, ist nur $P_\lambda(V) = \hat{E}(\lambda)$ mit der direkten Zerlegung (2.9) verträglich.

• Wegen $P_\mu \circ P_\lambda = P_\lambda \circ P_\mu$ ist

$$P_\mu P_\lambda(V) = \hat{E}(\lambda) \cap \hat{E}(\mu) = \begin{cases} \hat{E}(\lambda) & , \mu = \lambda \\ \{0\} & , \mu \neq \lambda. \end{cases}$$

Daher ist $P_\lambda = P_\lambda \text{Id}_V = P_\lambda \left(\sum_\mu P_\mu \right) = P_\lambda P_\lambda$, sodass auch die erste Identität aus (2.16) bewiesen ist.

1. Wegen

$$\text{im}(P_\lambda) = \hat{E}(\lambda) \quad \text{und} \quad P_\lambda P_\lambda = P_\lambda$$

ist $\hat{E}(\lambda)$ Eigenraum von P_λ zum Eigenwert Eins. Also ist $\hat{E}(\lambda)$ Eigenraum von φ_D zum Eigenwert λ .

2. φ_N lässt die Zerlegung (2.13) invariant, weil dies auch φ und φ_D tun. Wegen

$$\varphi_N|_{\hat{E}(\lambda)} = (\varphi - \lambda \text{Id})|_{\hat{E}(\lambda)}$$

ist $\varphi_N^{f_\lambda}|_{\hat{E}(\lambda)} = 0$, und der Fitting-Index ist die kleinste solche Potenz.

3. folgt, da φ_D und φ_N Polynome in φ sind.

4. Sei $\varphi = \psi_D + \psi_N$ mit $\psi_D, \psi_N \in L(V)$ diagonalisierbar, bzw. nilpotent und $\psi_D \circ \psi_N = \psi_N \circ \psi_D$. Dann kommutiert ψ_D auch mit φ , also auch mit dem Polynom φ_D von φ .

Auch φ_N und ψ_N kommutieren, denn

$$\begin{aligned} \varphi_N \circ \psi_N &= (\varphi - \varphi_D) \circ (\varphi - \psi_D) = \varphi \circ \varphi - \varphi_D \circ \varphi - \varphi \circ \psi_D + \varphi_D \circ \psi_D \\ &= \varphi \circ \varphi - \varphi \circ \varphi_D - \psi_D \circ \varphi + \psi_D \circ \varphi_D = (\varphi - \psi_D) \circ (\varphi - \psi_D) \\ &= \psi_N \circ \varphi_N. \end{aligned}$$

Nach dem folgenden Satz ist

$$\varphi_D - \psi_D = \psi_N - \varphi_N$$

damit diagonalisierbar und nilpotent, und damit die Nullabbildung. □

2.27 Satz

Für endlich viele Endomorphismen eines endlich-dimensionalen Vektorraumes gilt:

1. Die Summe kommutierender nilpotenter Endomorphismen ist nilpotent.
2. Kommutierende diagonalisierbare Endomorphismen besitzen eine gemeinsame Basis aus Eigenvektoren.

Beweis:

1. Es seien $\varphi, \psi \in L(V)$ nilpotent und $\varphi^n = \psi^n = 0$. Gilt außerdem $\varphi \circ \psi = \psi \circ \varphi$, dann ist

$$(\varphi + \psi)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} \varphi^k \psi^{2n-k} = 0,$$

denn falls $k \leq n$, dann ist $\psi^{2n-k} = 0$, andernfalls $\varphi^k = 0$. Iteration dieses Arguments zeigt die Aussage für endlich viele kommutierende nilpotente Endomorphismen.

2. Es seien $\varphi, \psi \in L(V)$ diagonalisierbar, dann können wir sie mithilfe der Projektoren P_λ bzw. Q_μ auf die Eigenräume von φ bzw. ψ in der Form

$$\varphi = \sum_{\lambda \in \text{spek}(\varphi)} \lambda P_\lambda, \quad \psi = \sum_{\mu \in \text{spek}(\psi)} \mu Q_\mu \quad (2.17)$$

schreiben. Ist außerdem $[\varphi, \psi] = 0$, dann ist, da die Projektoren P_λ bzw. Q_μ nach dem Satz über die Spektraldarstellung⁶ Funktionen von φ bzw. ψ sind, auch

$$[P_\lambda, Q_\mu] = 0 \quad (\lambda \in \text{spek}(\varphi), \mu \in \text{spek}(\psi)).$$

Daher erfüllen die Operatoren

$$R(\lambda, \mu) := P_\lambda \circ Q_\mu$$

nicht nur die Relation

$$\sum_{\lambda, \mu} R(\lambda, \mu) = \text{Id}, \quad (2.19)$$

sondern sind ebenfalls Projektoren mit

$$R(\lambda, \mu) R(\lambda', \mu') = \delta_{\lambda\lambda'} \delta_{\mu\mu'} R(\lambda, \mu). \quad (2.20)$$

Setzen wir $E(\lambda, \mu) := R(\lambda, \mu)(V)$, dann wegen (2.19) und (2.20)

$$V = \bigoplus_{\lambda, \mu} E(\lambda, \mu),$$

und wegen (2.17) sind die Restriktionen von φ bzw. ψ auf den Unterraum $E(\lambda, \mu)$ das λ -fache bzw. μ -fache der Identität.

Iteration dieses Arguments zeigt die Aussage für endlich viele kommutierende diagonalisierbare Endomorphismen. \square

⁶**Satz (Spektraldarstellung):** Ein diagonalisierbarer Endomorphismus $f \in L(V)$ eines endlich-dimensionalen Vektorraumes V besitzt die Form

$$f = \sum_{\lambda \in \text{spek}(f)} \lambda P_\lambda, \quad \text{wobei} \quad P_\lambda := \prod_{\mu \in \text{spek}(f) \setminus \{\lambda\}} \frac{f - \mu \cdot \text{Id}_V}{\lambda - \mu} \quad (2.18)$$

Projektor auf den Eigenraum $E(\lambda)$ zum Eigenwert $\lambda \in \text{spek}(f)$ ist.

Wir benutzen jetzt die Abkürzung

$$J_r(\lambda) := \lambda \mathbb{1}_r + N_r.$$

Beispielsweise ist $J_4(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$, und der nilpotente Anteil $J_r(0) = N_r$ von $J_r(\lambda)$ kommutiert mit dem Diagonalanteil $\lambda \mathbb{1}_r$.

2.28 Definition Eine Matrix aus $\text{Mat}(r, K)$ heißt **Jordan-Matrix** (oder ist in **Jordan-Normalform**), wenn sie die Gestalt

$$\bigoplus_{i=1}^j J_{r_i}(\lambda_i)$$

(mit $\sum_{i=1}^j r_i = r$ und $\lambda_i \in K$) besitzt. Die Matrizen $J_{r_i}(\lambda_i)$ heißen dann die **Jordan-Blöcke** der Matrix.

2.29 Satz Das charakteristische Polynom von $\varphi \in L(V)$ zerfalle in Linearfaktoren. Dann existiert eine Basis B von V (genannt **Jordan-Basis**), sodass die darstellende Matrix $\text{Mat}_B(\varphi)$ eine Jordan-Matrix ist.

Beweis: Wir gehen von der φ -invarianten Zerlegung (2.9) von V in die verallgemeinerten Eigenräume $\hat{E}(\lambda)$ aus, und betrachten die nilpotenten Endomorphismen

$$(\varphi - \lambda \cdot \text{Id})|_{\hat{E}(\lambda)} \quad (\lambda \in \text{spek}(\varphi))$$

dieser Unterräume. Diese besitzen nach Theorem 2.15 Jordan-Basen $B^{(\lambda)}$. schreiben wir diese hintereinander, erhalten wir eine Jordan-Basis B von φ . \square

2.30 Bemerkung (Uneindeutigkeit der Jordan-Normalform)

Im Allgemeinen gibt es keine kanonische Wahl der Anordnung der Jordan-Basen $B^{(\lambda)}$ (eine Ausnahme ist der Fall $K = \mathbb{R}$, wo wir die Eigenwerte $\lambda \in \text{spek}(\varphi)$ und entsprechend die $B^{(\lambda)}$ nach aufsteigender Größe anordnen können).

Wir sprechen daher von *einer*, nicht *der* Jordan-Normalform von φ .

2.2 Ein Beispiel

Wir wollen eine Jordan-Normalform der reellen Matrix

$$M := \begin{pmatrix} 1/4 & 1 & -3/4 & 0 & -3/4 & 1 \\ 1/2 & 4 & 1/2 & 0 & 7/2 & -3 \\ 9/4 & 3 & 13/4 & 0 & 25/4 & -6 \\ 3/4 & 0 & 3/4 & 2 & 7/4 & -1 \\ -1/2 & 0 & -1/2 & 0 & 1/2 & 1 \\ 0 & 2 & 0 & 0 & 2 & 0 \end{pmatrix}$$

bzw. des durch $\varphi(x) := Mx$ gegebenen Endomorphismus von $V = \mathbb{R}^6$ bestimmen.

1. Als ersten Schritt berechnen wir das *charakteristische Polynom*

$$\chi_M(x) = x^6 - 10x^5 + 41x^4 - 88x^3 + 104x^2 - 64x + 16.$$

Dieses besitzt die Nullstelle 2, was man schon daraus ablesen konnte, dass die vierte Spalte von M außer der 2 an der Diagonalstelle nur aus Nullen besteht. Fortgesetzte Division durch $x - 2$ zeigt, dass diese Nullstelle die Ordnung vier besitzt, und dass

$$\chi_M(x) = (x - 2)^4(x - 1)^2$$

ist. Die Eigenwerte sind also 2 und 1, und diese müssen in der Jordan-Normalform mit Vielfachheit 4 bzw. 2 auf der Diagonale auftauchen.

2. Als nächstes bestimmen wir die *verallgemeinerten Eigenräume* $\hat{E}(\lambda)$ dieser Nullstellen λ , zunächst für $\lambda = 2$.

Der vierdimensionale Kern von $(\varphi - 2\text{Id})^3$ mit darstellender Matrix

$$(M - 2 \cdot \mathbb{1})^3 = \begin{pmatrix} -11/4 & 0 & -7/4 & 0 & -19/4 & 4 \\ 1/2 & 0 & 1/2 & 0 & 3/2 & -1 \\ 13/4 & 0 & 9/4 & 0 & 25/4 & -5 \\ 1/2 & 0 & 1/2 & 0 & 3/2 & -1 \\ -1/2 & 0 & -1/2 & 0 & -3/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

enthält den Vektor $e_{3,1}^{(2)} := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ (wobei in der Notation $e_{p,j}^{(\lambda)}$ der obere Index

λ den Eigenwert bezeichnet, und $e_{p,j}^{(\lambda)}$ primitiver Vektor von $\varphi - \lambda\text{Id}$ ist, vergleiche Kapitel 2.1.2). Der verallgemeinerte Eigenraum $\hat{E}(2)$ wird daher von den Vektoren

$$b_3 := e_{3,1}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 := (\varphi - 2\text{Id}) \left(e_{3,1}^{(2)} \right) = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix},$$

$$b_1 := (\varphi - 2\text{Id})^2 \left(e_{3,1}^{(2)} \right) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

und dem zweiten Eigenvektor $b_4 := e_{1,1}^{(2)} := \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ von φ zum Eigenwert 2 aufgespannt.

3. Analog behandeln wir den zweidimensionalen verallgemeinerten Eigenraum $\hat{E}(1)$. Der Kern von $(\varphi - \text{Id})^2$ mit darstellender Matrix

$$(M - \mathbb{1})^2 = \begin{pmatrix} -1/4 & 2 & -1/4 & 0 & 7/4 & -1 \\ 1/2 & 5 & 1/2 & 0 & 11/2 & -5 \\ 7/4 & 6 & 7/4 & 0 & 31/4 & -8 \\ 1 & 1 & 1 & 1 & 3 & -2 \\ -1/2 & 0 & -1/2 & 0 & -1/2 & -1 \\ 0 & 4 & 0 & 0 & 4 & -3 \end{pmatrix}$$

enthält den Vektor $b_6 := e_{2,1}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$, und

$$b_5 := (\varphi - \text{Id}) \left(e_{2,1}^{(1)} \right) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

4. Schreiben wir diese sechs Vektoren b_1, \dots, b_6 nebeneinander, so erhalten wir die Matrix

$$S^{-1} := \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & -2 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 2 & 0 & 1 & 0 & 0 \end{pmatrix}$$

mit Inverser

$$S = \begin{pmatrix} 1/2 & 0 & 1/2 & 1 & 3/2 & -1 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ -1/2 & 0 & -1/2 & 0 & -1/2 & 1 \\ 7/4 & 0 & 3/2 & 0 & 7/4 & -2 \\ -1/2 & 0 & -1/2 & 0 & -3/2 & 1 \end{pmatrix},$$

sodass die *Jordansche Normalform* von M gleich

$$J = SMS^{-1} = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = J_3(2) \oplus J_1(2) \oplus J_2(1)$$

ist. Nebenbei hätten wir die Matrix S^{-1} nicht invertieren müssen, um J zu berechnen.

2.3 Euklidische und unitäre Vektorräume: Fortsetzung

Im Folgenden greifen wir wieder ein Thema der 'Linearen Algebra I' auf: Skalarprodukte über K -Vektorräumen V mit $K = \mathbb{R}$ oder \mathbb{C} , d.h. euklidische bzw. unitäre Vektorräume.

2.31 Definition Eine *Sesquilinearform* Φ auf V über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ heißt

- **hermitesch**, wenn gilt: $\Phi(w, v) = \overline{\Phi(v, w)}$ ($v, w \in V$),
- **antihermitesch**, wenn gilt: $\Phi(w, v) = -\overline{\Phi(v, w)}$ ($v, w \in V$).
- Φ heißt **Skalarprodukt** auf V , wenn Φ hermitesch und **positiv definit** ist, d.h.

$$\Phi(v, v) > 0 \quad (v \in V \setminus \{0\}).$$

Statt $\Phi(v, w)$ schreiben wir dann $\langle v, w \rangle$.

- Ein \mathbb{K} -Vektorraum V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ heißt für $\mathbb{K} = \mathbb{R}$ **euklidischer Vektorraum**, für $\mathbb{K} = \mathbb{C}$ **unitärer Vektorraum**.

Wir führen also zusätzlich zur Vektorraumstruktur ein zweites Strukturelement ein. Skalarprodukte sind *positiv definite* Sesquilinearformen, d.h. insbesondere gilt $\langle v, v \rangle > 0$ für $v \in V \setminus \{0\}$.

Die Beschränkung auf reelle oder komplexe Vektorräume eröffnet uns eine neue Welt, denn sie ermöglicht die Verschmelzung von Linearer Algebra und Analysis, d.h. z.B. die Betrachtung konvergenter Reihen $\sum_{n=1}^{\infty} v_n$ von Vektoren $v_n \in V$. Die Konkretisierung dieser Perspektive bleibt aber der *Funktionalanalysis* vorbehalten, also einer weiterführenden Vorlesung.

Die **Cauchy–Schwarz–Ungleichung**

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad (v, w \in V)$$

ermöglicht die Definition des Winkels $\varphi \in [0, \pi/2]$ zwischen den von $v \neq 0$ und $w \neq 0$ aufgespannten eindimensionalen Unterräumen:

$$\cos(\varphi) = \frac{|\langle v, w \rangle|}{\|v\| \cdot \|w\|}$$

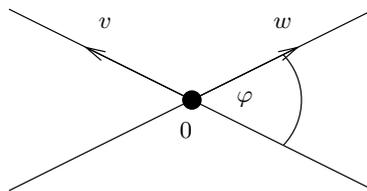


Abbildung 2.1: Winkel zwischen eindimensionalen Unterräumen

Besonders wichtig ist der Fall orthogonaler Vektoren mit $\varphi = \pi/2$ bzw. $\langle v, w \rangle = 0$.

Für ein Skalarprodukt auf einem \mathbb{R} -Vektorraum V ist

$$\langle v, w \rangle = \frac{1}{4} (\langle v + w, v + w \rangle - \langle v - w, v - w \rangle) \quad (v, w \in V),$$

wir können also das Skalarprodukt aus der Norm rekonstruieren.

Analoges gilt wegen der *Polarisationsidentität*

$$\langle v, w \rangle = \frac{1}{4} \left(\langle v + w, v + w \rangle - \langle v - w, v - w \rangle - i \langle v + iw, v + iw \rangle + i \langle v - iw, v - iw \rangle \right) \quad (v, w \in V)$$

auf \mathbb{C} -Vektorräumen V .

Insbesondere können wir die Winkelmessung durch die Längenmessung ersetzen.

2.3.1 Orthonormalisierung

Es sei V ein \mathbb{K} -Vektorraum der Dimension⁷ $1 \leq \dim(V) < \infty$ mit einem Skalarprodukt. Im Fall der arithmetischen Vektorräume $V = \mathbb{C}^n$ und $V = \mathbb{R}^n$ wird oft das *kanonische Skalarprodukt*

$$\langle v, w \rangle = \sum_{i=1}^n \bar{v}_i w_i \quad (v, w \in V) \quad (2.21)$$

verwendet. Im Fall $V = \mathbb{R}^n \subset \mathbb{C}^n$ spielt die Komplexkonjugation des ersten Argumentes keine Rolle. Die Mathematiker benutzen statt (2.21) eine Definition des Skalarproduktes, das im 2. Argument antilinear ist.

2.32 Definition

- Eine nicht leere Teilmenge $M \subset V$ heißt **Orthogonalsystem**, wenn $0 \notin M$ und $\langle v, w \rangle = 0$ ($v \neq w \in M$).
- Sie heißt **Orthonormalsystem**, wenn zusätzlich $\|v\| = 1$ ($v \in M$) gilt.
- Eine Basis B von V heißt **Orthogonalbasis (Orthonormalbasis)**, wenn sie gleichzeitig ein Orthogonalsystem (Orthonormalsystem) ist.

Insbesondere mit Orthonormalbasen $B = (b_1, \dots, b_n)$ lässt sich einfach rechnen, denn es gilt die Basisdarstellung

$$v = \sum_{i=1}^n \lambda_i b_i \quad \text{mit} \quad \lambda_i := \langle b_i, v \rangle \quad (v \in V).$$

Dies lässt sich folgendermaßen geometrisch beweisen:

2.33 Lemma Ist $B = (b_1, \dots, b_n)$ eine Orthonormalbasis, dann ist

$$P_i \in L(V) \quad , \quad P_i(v) := \langle b_i, v \rangle b_i \quad (i = 1, \dots, n)$$

eine Familie von Projektoren, d.h.

$$P_i P_j = \delta_{ij} P_j \quad \text{und} \quad \sum_{i=1}^n P_i = \text{Id}_V.$$

Beweis:

⁷Ist $\dim(V) = \infty$, dann wird der Basisbegriff oft verändert, und man verlangt nicht mehr, dass jeder Vektor *endliche* Linearkombination von Basisvektoren sein muss. Siehe z.B. REED und SIMON, Functional Analysis, Kapitel II.3.

- $P_i P_j v = \langle b_i, \langle b_j, v \rangle b_j \rangle b_i = \langle b_j, v \rangle \langle b_i, b_j \rangle b_i = \langle b_j, v \rangle \delta_{ij} b_i \quad (v \in V)$.
- Weiter sind für $k = 1, \dots, n$ die $Q_k := \sum_{i=1}^k P_i$ Projektoren auf $V_k := \text{span}(b_1, \dots, b_k)$, denn $Q_k^2 = \sum_{i,j=1}^k P_i P_j = \sum_{i,j=1}^k \delta_{ij} P_j = Q_k$, und $Q_k(\sum_{i=1}^n \lambda_i b_i) = \sum_{j=1}^k \sum_{i=1}^n \lambda_i P_j(b_i) = \sum_{i=1}^k \lambda_i b_i$. Für $k = n$ ist damit $Q_k = \text{Id}_V$. \square

2.34 Definition Eine Projektion $P = P^2 \in L(V)$ heißt **Orthogonalprojektion**, wenn $\ker(P) = \text{im}(P)^\perp$.

Wir wissen schon, dass für eine Projektion $V = \ker(P) \oplus \text{im}(P)$ gilt, wir können also jeden Vektor $v \in V$ eindeutig in $v = v_K + v_I$ mit $v_K := v - P(v) \in \ker(P)$ und $v_I := P(v) \in \text{im}(P)$ zerlegen.

Umgekehrt ist $P = P^2$ genau dann eine Orthogonalprojektion, wenn gilt:

$$\langle v - P(v), P(v) \rangle = 0 \quad (v \in V)$$

Weiter ist eine Projektion genau dann Orthogonalprojektion, wenn

$$\langle v, P(w) \rangle = \langle P(v), w \rangle \quad (v, w \in V),$$

denn

$$\begin{aligned} \langle v, P(w) \rangle &= \langle P(v), P(w) \rangle + \langle v - P(v), P(w) \rangle \\ &= \langle P(v), w \rangle + \langle v - P(v), P(w) \rangle + \langle P(v), P(w) - w \rangle. \end{aligned}$$

2.35 Beispiele (Projektionen und Orthogonalprojektionen)

1. Die Projektionen $P_b(v) := \langle b, v \rangle b$ für $b \in V$, $\|b\| = 1$ sind Orthogonalprojektionen mit $\text{im}(P_b) = \text{span}(b)$ und $\ker(P_b) = \text{span}(b)^\perp$.
2. Die Projektion $P(v) := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} v$ auf dem Vektorraum \mathbb{R}^2 mit Skalarprodukt $\langle v, w \rangle = v_1 w_1 + v_2 w_2$ ist keine Orthogonalprojektion, denn $\ker(P) = \text{span}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$ steht nicht senkrecht auf $\text{im}(P) = \text{span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$. \diamond

Wie können wir nun aus einer beliebigen Basis $B = (b_1, \dots, b_n)$ von V eine Orthonormalbasis (ONB) konstruieren? Dies wird im Beweis des folgenden Satzes vorgeführt. Das Verfahren heißt *Gram-Schmidt-Orthonormalisierung*.

2.36 Satz (E. Schmidt) Ist $B = (b_1, \dots, b_n)$ eine Basis von V , dann existiert eine ONB $C = (c_1, \dots, c_n)$ von V mit

$$\text{span}(c_1, \dots, c_k) = \text{span}(b_1, \dots, b_k) \quad (k = 1, \dots, n).$$

Beweis:

- Wir setzen $c_1 := b_1 / \|b_1\|$, sodass $\|c_1\| = 1$.
- Haben wir für $k < n$ schon eine Orthonormalbasis (c_1, \dots, c_k) von $V_k := \text{span}(b_1, \dots, b_k)$ gefunden, dann ist der Projektor $Q_k = \sum_{i=1}^k P_i$ wegen $\langle Q_k(v), w \rangle = \sum_{i=1}^k \langle P_i(v), w \rangle = \sum_{i=1}^k \langle v, P_i(w) \rangle = \langle v, Q_k(w) \rangle$ Orthogonalprojektor auf V_k . Wir setzen

$$c_{k+1} := \frac{(\text{Id} - Q_k)b_{k+1}}{\|(\text{Id} - Q_k)b_{k+1}\|},$$

projizieren also b_{k+1} auf den zu V_k orthogonalen Unterraum, und normieren auf Länge 1. Wegen

$$\langle c_i, (\text{Id} - Q_k)b_{k+1} \rangle = \langle (\text{Id} - Q_k)c_i, b_{k+1} \rangle = \langle 0, b_{k+1} \rangle \quad (i = 1, \dots, k)$$

und $\|c_{k+1}\| = 1$ bildet (c_1, \dots, c_{k+1}) eine ONB von V_{k+1} . □

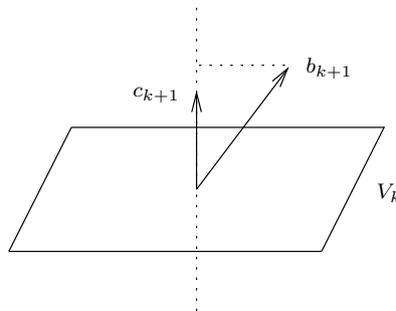


Abbildung 2.2: Gram-Schmidt-Orthonormalisierung

2.3.2 Die adjungierte Abbildung

Es seien V und W euklidische bzw. unitäre Vektorräume.

2.37 Definition

Eine lineare Abbildung $\varphi^* \in L(W, V)$ heißt zu $\varphi \in L(V, W)$ **adjungiert**, wenn

$$\langle \varphi^*(w), v \rangle = \langle w, \varphi(v) \rangle \quad (v \in V, w \in W).$$

Gibt es eine solche adjungierte Abbildung, dann ist sie eindeutig.

2.38 Satz Ist $\dim(V) < \infty$, dann existiert φ^* und besitzt bezüglich einer ONB (b_1, \dots, b_n) die Form

$$\varphi^*(w) = \sum_{i=1}^n \langle \varphi(b_i), w \rangle b_i \quad (w \in W).$$

Beweis: Der Fall $\dim(V) = 0$ ist trivial, und aus der Identität $v = \sum_{j=1}^n \langle b_j, v \rangle b_j$ ergibt sich

$$\begin{aligned} \left\langle \sum_{i=1}^n \langle \varphi(b_i), w \rangle b_i, v \right\rangle &= \sum_{i,j=1}^n \langle w, \varphi(b_i) \rangle \langle b_i, \langle b_j, v \rangle b_j \rangle \\ &= \sum_{i,j=1}^n \langle w, \varphi(b_i) \rangle \langle b_j, v \rangle \delta_{ij} = \sum_{j=1}^n \langle w, \varphi(\langle b_j, v \rangle b_j) \rangle \\ &= \langle w, \varphi(v) \rangle. \quad \square \end{aligned}$$

Wir erinnern uns an die Diskussion des Dualraumes V^* eines K -Vektorraumes V . Dessen Elemente sind die linearen Abbildungen $V \rightarrow K$.

Ist nun $w \in V$ und besitzt V ein Skalarprodukt, dann ist die Abbildung

$$I(w) \in V^* \quad , \quad I(w)(v) := \langle w, v \rangle$$

ein lineares Funktional auf V , und für $w \neq 0$ ist auch $I(w) \neq 0$, denn $I(w)(w) = \langle w, w \rangle > 0$. Wir erhalten also eine Injektion

$$I = I_V : V \rightarrow V^*.$$

Diese ist *konjugiert-linear*, es gilt also

$$I(w_1 + w_2) = I(w_1) + I(w_2) \quad , \quad I(k \cdot w) = \bar{k} \cdot I(w).$$

Damit ist für $K = \mathbb{C}$ die Injektion *nicht* linear, für $K = \mathbb{R}$ wegen $\bar{k} = k$ aber schon.

Ist nun $\dim(V) < \infty$ und $B = (b_1, \dots, b_n)$ eine Orthonormalbasis, dann ist $(I(b_1), \dots, I(b_n))$ gleich der *dualen Basis*, also der Basis (b_1^*, \dots, b_n^*) von V^* mit

$$b_i^*(b_k) = \delta_{ik} \quad (i, k = 1, \dots, n).$$

Damit ist $I_V : V \rightarrow V^*$ surjektiv und auch $I_V^{-1} : V^* \rightarrow V$ konjugiert-linear.

2.39 Satz Sind V und W endlich-dimensionale euklidische oder unitäre Vektorräume, und ist $\varphi \in L(V, W)$, dann gilt für die adjungierte Abbildung

$$\varphi^* = I_V^{-1} \circ \varphi^t \circ I_W. \quad (2.22)$$

Ist B eine ONB von V und C eine ONB von W , dann gilt für die darstellende Matrizen $A := \text{Mat}_C^B(\varphi)$ und $A^* := \text{Mat}_B^C(\varphi)$

$$A^* = \overline{A^t} \quad , \quad \text{d.h.} \quad (A^*)_{i,k} = \overline{(A)_{k,i}} \quad (i = 1, \dots, \dim(V), \\ j = 1, \dots, \dim(W)).$$

Beweis: Nach Definition der *transponierten* Abbildung $\varphi^t : W^* \rightarrow V^*$ gilt

$$\varphi^t(w^*)(v) = w^*(\varphi(v)) \quad (v \in V, w^* \in W^*).$$

Daher ist für alle $v \in V$ und $w \in W$

$$\varphi^t(I_W(w))(v) = I_W(w)(\varphi(v)) = \langle w, \varphi(v) \rangle = \langle \varphi^*(w), v \rangle = I_V(\varphi^*(w))(v)$$

oder kurz

$$\varphi^t \circ I_W = I_V \circ \varphi^*.$$

Wegen der Invertierbarkeit von I_V impliziert dies Gleichung (2.22).

Nach einem Satz über die transponierte Abbildung φ^t ist

$$A^t = \text{Mat}_{B^*}^{C^*}(\varphi^t)$$

die darstellende Matrix von φ^t bezüglich der dualen Basen. Da I_V^{-1} und I_W konjugiert-linear sind und $I_V^{-1}(B^*) = B$, $I_W(C) = C^*$ gilt, ergibt sich daraus $A^* = \overline{A^t}$. \square

Die adjungierte Abbildung hat also den gleichen Rang wie φ .

2.40 Korollar *Sind V und W endlich-dimensionale euklidische oder unitäre Vektorräume, und ist $\varphi \in L(V, W)$, dann gilt für die adjungierte Abbildung*

$$\text{im}(\varphi^*) = \ker(\varphi)^\perp \subset V \quad \text{und} \quad \ker(\varphi^*) = \text{im}(\varphi)^\perp \subset W.$$

Beweis: Dies ergibt sich aus dem entsprechenden Satz für die transponierte Abbildung, denn für einen Unterraum $U \subset V$ ist der zu U in V^* orthogonale Unterraum

$$\{v^* \in V^*, v^*(u) = 0 \quad \forall u \in U\}$$

gleich dem Bild von

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \quad \forall u \in U\}$$

unter der Abbildung $I_V : V \rightarrow V^*$, und analog für W . \square

Damit ist φ^* genau dann surjektiv, wenn φ injektiv ist, und umgekehrt.

2.41 Beispiel (Adjungierte Abbildung)

Die adjungierte Abbildung zur Projektion $\varphi(v) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} v$ im \mathbb{R}^2 ist $\varphi^*(v) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} v$.

φ^* projiziert auf $\ker(\varphi)^\perp = \text{span} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$ entlang der Richtung $\text{im}(\varphi)^\perp = \text{span} \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$. \diamond

Wir haben gesehen, dass eine Projektion $P \in L(V)$ genau dann eine Orthogonalprojektion ist, wenn $P^* = P$ ist.

2.3.3 Normale Operatoren

Diese so genannte Selbstadjungiertheit eines Operators ist Spezialfall der folgenden Eigenschaft.

2.42 Definition Existiert die adjungierte Abbildung φ^* zu $\varphi \in L(V)$ und gilt $\varphi^* \circ \varphi = \varphi \circ \varphi^*$, dann heißt φ **normal**.

2.43 Beispiel (Normalität)

$\varphi(v) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} v$ ist nicht normal, denn $\varphi^*(v) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} v$ und

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}. \quad \diamond$$

2.44 Satz Für normale Endomorphismen $\varphi \in L(V)$ gilt

$$\ker(\varphi) = \ker(\varphi^*) \quad \text{und} \quad \text{im}(\varphi) = \text{im}(\varphi^*).$$

Beweis:

- Allgemein gilt für einen zweiten Endomorphismus ψ $\ker(\psi \circ \varphi) \supset \ker(\varphi)$. Für $\psi = \varphi^*$ ist aber nach Corollar 2.40 $\ker(\varphi^*) = \text{im}(\varphi)^\perp$, sodass $\ker(\varphi^* \circ \varphi) = \ker(\varphi)$. Analog gilt $\ker(\varphi \circ \varphi^*) = \ker(\varphi)$, woraus wegen der Normalität von φ die Beziehung $\ker(\varphi) = \ker(\varphi^*)$ folgt.
- Allgemein gilt $\text{im}(\varphi \circ \psi) \subset \text{im}(\varphi)$, aber für $\psi = \varphi^*$ wegen $\ker(\varphi) = \text{im}(\varphi^*)^\perp$ sogar $\text{im}(\varphi \circ \varphi^*) = \text{im}(\varphi)$. Analog folgt $\text{im}(\varphi^* \circ \varphi) = \text{im}(\varphi^*)$ und wegen der Normalität von φ die zweite Beziehung. \square

2.45 Satz Ein Endomorphismus $\varphi \in L(V)$ eines endlich-dimensionalen unitären Vektorraums V ist genau dann normal, wenn V eine ONB aus Eigenvektoren von φ besitzt.

Beweis:

- Besitzt V eine solche ONB, dann ist φ von der Form $\varphi = \sum_{\lambda \in \text{spek}(\varphi)} \lambda P_\lambda$ mit der Familie von Projektoren P_λ auf die Eigenräume zu den Eigenwerten λ . Da diese orthogonal sind, gilt $P_\lambda^* = P_\lambda$, also

$$\varphi^* \circ \varphi = \varphi \circ \varphi^* = \sum_{\lambda \in \text{spek}(\varphi)} |\lambda|^2 P_\lambda.$$

- Ist φ normal, also $[\varphi, \varphi^*] = 0$, dann gilt auch $[\varphi_D, \varphi_D^*] = [\varphi_N, \varphi_N^*] = 0$ für den diagonalisierbaren bzw. nilpotenten Anteil von $\varphi = \varphi_D + \varphi_N$, denn diese sind Polynome in φ (Satz 2.26). Daraus folgt zunächst $\varphi_N = 0$. Denn wegen Satz 2.44 gilt $\ker(\varphi_N) = \ker(\varphi_N^*)$. Damit ist $\ker(\varphi_N) = \text{im}(\varphi_N^\perp)$, sodass $\ker(\varphi_N^2) = \ker(\varphi_N)$, also $\ker(\varphi_N) = V$ ist.
- Gemäß Satz 2.26 (Jordan-Zerlegung) können wir also $\varphi = \varphi_D$ in der Form

$$\varphi = \sum_{\lambda \in \text{spek}(\varphi)} \lambda P_\lambda$$

Beweis: Dies folgt aus Satz 2.45 zusammen mit der reellen Jordan-Normalform. Denn $\varphi_{\mathbb{C}}$ besitzt ja eine orthonormale Eigenbasis, die mit dem Eigenvektor c zum Eigenwert $\mu_j + i\nu_j$ auch den Eigenvektor \bar{c} zum Eigenwert $\mu_j - i\nu_j$ enthält. Da diese normiert sind und aufeinander senkrecht stehen, gilt für

$$a := \sqrt{2} \operatorname{Re}(c) = \frac{c + \bar{c}}{\sqrt{2}} \quad , \quad b := \sqrt{2} \operatorname{Im}(c) = \frac{c - \bar{c}}{\sqrt{2}i}$$

$$\langle a, a \rangle = \frac{1}{2}(\langle c, c \rangle + \langle \bar{c}, \bar{c} \rangle) = 1 \quad , \quad \langle b, b \rangle = \frac{1}{2}(\langle c, c \rangle + \langle \bar{c}, \bar{c} \rangle) = 1$$

und

$$\langle a, b \rangle = \frac{1}{2i} \langle c + \bar{c}, c - \bar{c} \rangle = 0. \quad \diamond$$

Im Folgenden schauen wir uns einige Beispiele normaler Operatoren genauer an.

2.3.4 Orthogonale und unitäre Operatoren

2.48 Definition

Sind V, W euklidische (bzw. unitäre) Vektorräume, dann heißt $\varphi \in L(V, W)$

- **Isometrie**, wenn $\|\varphi(v)\| = \|v\| \quad (v \in V)$,
- **orthogonal** (bzw. **unitär**), wenn zusätzlich $\operatorname{ran}(\varphi) = W$.

Unter Benutzung der Polarisationsidentität folgern wir für Isometrien die Gleichung

$$\langle \varphi(v_1), \varphi(v_2) \rangle = \langle v_1, v_2 \rangle \quad (v_1, v_2 \in V).$$

Es bleiben also nicht nur Längen, sondern auch Winkel unverändert. Sind V und W endlich-dimensional, dann ist eine Isometrie genau dann orthogonal bzw. unitär, wenn $\dim(V) = \dim(W)$.

Während bijektive lineare Abbildungen $\varphi : V \rightarrow W$ die Vektorraumstrukturen ineinander überführen, werden durch orthogonale bzw. unitäre Transformationen auch die Skalarprodukte ineinander übergeführt.

Dies ist für K -Vektorräume V, W immer möglich, wenn V und W die gleiche endliche Dimension n besitzen, denn dann können wir ja ONBs (b_1, \dots, b_n) von V und (c_1, \dots, c_n) von W konstruieren und aufeinander abbilden:

$$\varphi(b_i) := c_i \quad (i = 1, \dots, n).$$

Lineare Fortsetzung ergibt eine orthogonale bzw. unitäre Transformation $\varphi \in L(V, W)$.

Wir können also, wenn wir diese Abbildungen untersuchen wollen, ohne Verlust an Allgemeinheit orthogonale Endomorphismen des \mathbb{R}^n bzw. unitäre Endomorphismen des \mathbb{C}^n studieren, wobei wir jeweils das kanonische Skalarprodukt (2.21) verwenden.

der Form

$$O = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

schreiben. Denn die erste Spalte ist ein Vektor der Länge 1 im \mathbb{R}^2 , also von der Form $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$, während der zweite Spaltenvektor auf dem ersten senkrecht steht, also von der Form $\pm \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$ ist. Das positive Vorzeichen ergibt sich aus der Forderung $\det(A) = 1$. Die Matrix O ist also schon in der reellen Normalform. Geometrisch ist die lineare Abbildung $v \mapsto Ov$ eine Drehung um den Winkel φ .

• Analog können wir $O \in O(2) \setminus SO(2)$ für ein eindeutiges $\varphi \in [0, 2\pi)$ in der Form

$$O = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$$

schreiben, denn diesmal muss ja $\det(A) = -1$ gelten. Damit ist aber O genau dann in Normalform, wenn $\sin \varphi = 0$ ist, denn O besitzt die Eigenwerte 1 und -1 mit den Eigenvektoren $\begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{pmatrix}$ und $\begin{pmatrix} -\sin(\varphi/2) \\ \cos(\varphi/2) \end{pmatrix}$. Geometrisch ist die lineare Abbildung $v \mapsto Ov$ eine Spiegelung an der durch den ersten Eigenvektor aufgespannten Geraden.

• Sowohl die Drehgruppe $SO(2)$ als auch die Menge $O(2) \setminus SO(2)$ der Achsenpiegelungen sind damit in den Raum $\text{Mat}(2, \mathbb{R}) \cong \mathbb{R}^4$ eingebettete Kreislinien, parametrisiert durch φ .

• Die Gruppe $SO(2)$ ist abelsch, denn

$$\begin{pmatrix} \cos \varphi_1 & -\sin \varphi_1 \\ \sin \varphi_1 & \cos \varphi_1 \end{pmatrix} \begin{pmatrix} \cos \varphi_2 & -\sin \varphi_2 \\ \sin \varphi_2 & \cos \varphi_2 \end{pmatrix} = \begin{pmatrix} \cos(\varphi_1 + \varphi_2) & -\sin(\varphi_1 + \varphi_2) \\ \sin(\varphi_1 + \varphi_2) & \cos(\varphi_1 + \varphi_2) \end{pmatrix}.$$

Dagegen ist $O(2)$ nicht abelsch, denn beispielsweise ist

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \quad \diamond$$

2.51 Beispiel (Die orthogonale Gruppe $O(3)$ des \mathbb{R}^3)

• Ist $O \in SO(3)$, dann ist die reelle Normalform von O von der Gestalt

$$S^{-1}OS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix},$$

wobei die Transformationsmatrix S aus $SO(3)$ gewählt werden kann. Wir können daher den Betrag des *Drehwinkels* φ aus

$$\text{tr}(O) = \text{tr}(S^{-1}OS) = 1 + 2 \cos(\varphi)$$

ablesen.

• Die *Drehachse* von $O \neq \mathbb{1}$ ist der Eigenraum zum Eigenwert 1 und damit eindimensional. Die zu diesem eindimensionalen Unterraum orthogonale Ebene

bleibt unter O als Menge invariant, die Vektoren dieser Ebene werden aber um den Winkel φ gedreht.

- $O \in O(3) \setminus SO(3)$ besitzt die reelle Normalform

$$S^{-1}OS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}.$$

Dies entspricht einer Drehung um den Winkel φ , gefolgt von einer Spiegelung an der Drehebene. Letztere ist für $\varphi \neq \pi$ eindeutig definiert.

- Wir können folgendermaßen eine stetige Abbildung

$$\phi : B \rightarrow SO(3)$$

von der Einheitskugel $B := \{x \in \mathbb{R}^3 \mid \|x\| \leq 1\}$ des \mathbb{R}^3 auf die Drehgruppe definieren: $\phi(0) := \mathbb{1}$, und für $x \in B \setminus \{0\}$ sei $\phi(x)$ die Drehung mit Achse $\text{span}(x) \subset \mathbb{R}^3$ um den Winkel $\pi \cdot |x|$ in positiver Richtung bezüglich x . Es gilt $\phi(x) = \phi(y)$ für $x \neq y$ genau, wenn

$$y = -x \in \partial B := \{x \in \mathbb{R}^3 \mid \|x\| = 1\}.$$

Denn Drehungen um die gleiche Achse mit Winkel π bzw. $-\pi$ sind einander gleich. ∂B ist die Oberfläche der Vollkugel B , und die Punkte x und $-x \in \partial B$ sind Antipoden auf dieser zweidimensionalen Sphäre.

Die Vollkugel B mit identifizierten Antipoden auf ∂B ist damit ein topologisches Modell von $SO(3)$ (siehe auch KNÖRRER [Kn, Kapitel 6]).

- Schon $SO(3)$ ist keine abelsche Gruppe mehr, denn Drehungen um verschiedene Achsen kommutieren nicht. Beispiel: Drehung mit Winkel π um x -Achse und z -Achse

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Die Isometrien des unitären Vektorraumes \mathbb{C}^n bilden die Gruppe $U(n)$. Die Determinante $\det(U)$ einer unitären Matrix U ist eine komplexe Zahl, die wegen $\overline{\det(U)} = \det(U^*) = \det(U^{-1}) = \det(U)^{-1}$ vom Betrag 1 ist. Die Isometrien U von \mathbb{C} beispielsweise bestehen aus der Multiplikation mit der komplexen Zahl $U \in U(1) = \{c \in \mathbb{C} \mid |c| = 1\}$. Wegen des Determinantenproduktsatzes ist

$$\det : U(n) \rightarrow U(1)$$

ein Gruppenhomomorphismus, und damit $SU(n) = \det^{-1}(1)$ eine normale Untergruppe von $U(n)$.

2.52 Beispiel (Die unitäre Gruppe $U(2)$ des \mathbb{C}^2)

• Ist $U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \in U(2)$, dann muss $|u_{11}|^2 + |u_{21}|^2 = 1 = |u_{12}|^2 + |u_{22}|^2$ und $\overline{u_{11}}u_{12} + \overline{u_{21}}u_{22} = 0$ gelten. Dies ist genau dann erfüllt, wenn der erste Spaltenvektor $\begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}$ die Länge 1 besitzt, und der zweite Spaltenvektor von der Form $\begin{pmatrix} u_{12} \\ u_{22} \end{pmatrix} = \lambda \begin{pmatrix} -\overline{u_{21}} \\ \overline{u_{11}} \end{pmatrix}$ für ein $\lambda \in U(1)$ ist. In diesem Fall ist

$$\det(U) = u_{11}u_{22} - u_{21}u_{12} = \lambda(|u_{11}|^2 + |u_{21}|^2) = \lambda.$$

• Ist also $U \in SU(2)$, dann ist

$$U = \begin{pmatrix} v & -\overline{w} \\ w & \overline{v} \end{pmatrix} \quad \text{mit} \quad v, w \in \mathbb{C} \quad \text{und} \quad |v|^2 + |w|^2 = 1.$$

Der erste Spaltenvektor $\begin{pmatrix} v \\ w \end{pmatrix}$ legt also schon die Matrix U fest, und $\begin{pmatrix} v \\ w \end{pmatrix} \in \mathbb{C}^2$ ist ein beliebiger Vektor der Länge 1.

Die Vektoren $\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} v_1 + iv_2 \\ w_1 + iw_2 \end{pmatrix}$ der Länge $v_1^2 + v_2^2 + w_1^2 + w_2^2 = 1$ bilden aber die Einheitskugel $S^3 \subset \mathbb{C}^2 \cong \mathbb{R}^4$. Die dreidimensionale Einheitskugel S^3 bildet damit ein topologisches Modell der Gruppe $SU(2)$. \diamond

2.53 Beispiel (Diskrete Fourier-Transformation) Es sei für ein $N \in \mathbb{N}$

$$V_N := \{f : \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{C}\} \quad , \quad \langle f, g \rangle := \sum_{x \in \mathbb{Z}/N\mathbb{Z}} \overline{f(x)}g(x)$$

der unitäre Vektorraum der komplexen Funktionen auf der abelschen Gruppe $\mathbb{Z}/N\mathbb{Z}$. Diese besitzt N Elemente, z.B. durch $0, 1, \dots, N-1$ repräsentiert, und in Anwendungen ist $f \in V_N$ oft als Diskretisierung einer Funktion $F : [0, T] \rightarrow \mathbb{C}$ definiert: $f(x) := F(x/T)$ ($x = 0, \dots, N-1$).

Die *diskrete Fourier-Transformation* $U_N \in L(V_N)$ ist nun durch

$$(U_N f)(k) := N^{-1/2} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} f(x)e^{-2\pi i k x / N} \quad (k \in \mathbb{Z}/N\mathbb{Z})$$

definiert, sodass die adjungierte Abbildung gleich

$$(U_N^* g)(x) = N^{-1/2} \sum_{k \in \mathbb{Z}/N\mathbb{Z}} g(k)e^{2\pi i k x / N} \quad (x \in \mathbb{Z}/N\mathbb{Z})$$

ist.⁸ Damit ist U_N unitär, denn

$$\begin{aligned} (U_N^* U_N f)(y) &= N^{-1} \sum_{x, k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k(y-x)/N) f(x) \\ &= N^{-1} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} N \delta_{y,x} f(x) = f(y) \end{aligned}$$

⁸Tatsächlich ist U_N wohldefiniert, denn für $k' - k \in N\mathbb{Z}$ und $x' - x \in N\mathbb{Z}$ ist $\exp(2\pi i k' x' / N) = \exp(2\pi i k x / N)$.

folgt (mit $z = y - x$) aus der Relation

$$\sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) = N \delta_{z,0} \quad (z \in \mathbb{Z}/N\mathbb{Z}). \quad (2.23)$$

Die Identität (2.23) für $z = 0$ folgt aus $e^0 = 1$, während für $z \neq 0$ die Anwendung der Bijektion $k \mapsto k + 1$ auf $\mathbb{Z}/N\mathbb{Z}$

$$\begin{aligned} \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) &= \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i (k + 1) z / N) \\ &= e^{2\pi i z / N} \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) \end{aligned}$$

oder

$$(1 - e^{2\pi i z / N}) \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) = 0$$

ergibt. Da der Faktor $1 - e^{2\pi i z / N} \neq 0$ ist, verschwindet die Summe.

Wir haben mit dem Nachweis der Unitarität von U_N auch gezeigt, dass (b_1, \dots, b_N) mit $b_k(x) := \frac{\exp(2\pi i k x / N)}{\sqrt{N}}$ eine ONB von V_N ist. Anders gesagt, entspricht der Fourier-Transformation von $f \in V_N$ die "Entwicklung von f nach den Wellen b_k "

$$f = \sum_{k=0}^{N-1} \lambda_k b_k \quad \text{mit} \quad \lambda_k := \langle b_k, f \rangle = (U_N f)(k). \quad \diamond$$

2.3.5 (Anti-) Selbstadjungierte Endomorphismen

2.54 Definition Ist V ein euklidischer bzw. unitärer Vektorraum, dann heißt ein Endomorphismus $\varphi \in L(V)$ (dessen Adjungierte $\varphi^* \in L(V)$ existiert)

- **selbstadjungiert**, wenn $\varphi^* = \varphi$,
- **antiselbstadjungiert**, wenn $\varphi^* = -\varphi$ gilt.

Die Begriffe (anti-) symmetrisch und (anti-) hermitesch werden ebenfalls benutzt, der erste eher bei euklidischen und der zweite eher bei unitären Vektorräumen. Analoge Definitionen gelten für Matrizen aus $\text{Mat}(n, K)$ über den Körpern $K = \mathbb{R}$ oder \mathbb{C} .

(Anti-) Selbstadjungierte Endomorphismen sind normal, denn

$$\varphi^* \circ \varphi = \pm \varphi \circ \varphi = \varphi \circ \varphi^*.$$

2.55 Satz

Ist V ein unitärer Vektorraum, $\dim(V) < \infty$ und $\varphi \in L(V)$, dann ist

- $\text{spek}(\varphi) \subset \mathbb{R}$, falls φ selbstadjungiert und
- $\text{spek}(\varphi) \subset i \cdot \mathbb{R}$, falls φ antiselbstadjungiert ist.

Beweis: Wegen der Normalität von φ ist φ diagonalisierbar, und

$$\varphi = \sum_{\lambda \in \text{spek}(\varphi)} \lambda P_\lambda,$$

wobei die P_λ Orthogonalprojektionen sind. Damit ist $\varphi^* = \sum_{\lambda \in \text{spek}(\varphi)} \bar{\lambda} P_\lambda$.

Falls $\varphi^* = \varphi$ gilt, ist damit $\bar{\lambda} = \lambda$, also $\text{spek}(\varphi) \subset \mathbb{R}$.

Dagegen ist für $\varphi^* = -\varphi$ der Eigenwert λ imaginär, denn $\bar{\lambda} = -\lambda$. □

2.56 Korollar Ist V ein euklidischer Vektorraum endlicher Dimension, und $\varphi \in L(V)$ selbstadjungiert, dann besitzt V eine ONB aus Eigenvektoren von φ .

Beweis: Auch $\varphi_{\mathbb{C}} \in L(V_{\mathbb{C}})$ ist selbstadjungiert, das charakteristische Polynom von $\varphi_{\mathbb{C}}$ besitzt also nur reelle Nullstellen. Daher zerfällt das charakteristische Polynom von φ in Linearfaktoren, sodass φ diagonalisierbar ist. Wegen der Normalität von φ sind die Eigenräume $E(\lambda)$ zu den Eigenwerten $\lambda \in \text{spek}(\varphi)$ orthogonal. □

Dass Eigenvektoren v und w zu den Eigenwerten $\mu \neq \lambda$ für selbstadjungierte Endomorphismen aufeinander senkrecht stehen, sieht man auch folgendermaßen:

$$0 = \langle \varphi(v), w \rangle - \langle v, \varphi(w) \rangle = (\lambda - \mu) \langle v, w \rangle, \text{ also } \langle v, w \rangle = 0.$$

Jeder Endomorphismus $\varphi \in L(V)$, dessen Adjungierte φ^* existiert, lässt sich eindeutig in die Summe

$$\varphi = \varphi_s + \varphi_a \tag{2.24}$$

eines selbstadjungierten Operators $\varphi_s := \frac{1}{2}(\varphi + \varphi^*)$ und eines antiselbstadjungierten Operators $\varphi_a := \frac{1}{2}(\varphi - \varphi^*)$ zerlegen. Denn ist $\varphi = \psi_s + \psi_a$ eine zweite solche Zerlegung, dann ist der Operator $\pi := \varphi_s - \psi_s$ selbstadjungiert, aber wegen $\pi = \psi_a - \varphi_a$ auch antiselbstadjungiert, sodass $\pi = 0$ folgt.

2.57 Bemerkung (Zerlegung normaler Endomorphismen)

Die Zerlegung (2.24) ist dann besonders nützlich, wenn φ_s und φ_a kommutieren. Dies ist wegen

$$[\varphi, \varphi^*] = [\varphi_s + \varphi_a, \varphi_s - \varphi_a] = 2[\varphi_a, \varphi_s]$$

genau dann der Fall, wenn φ normal ist.

Nach Satz 2.27 können wir damit für unitäre Vektorräume V endlicher Dimension normale φ dadurch diagonalisieren, dass wir φ_s und φ_a diagonalisieren, und eine φ_s und φ_a gemeinsame Basis von Eigenvektoren aufsuchen. ◇

Der Vektorraum $L(K^n)$ der Endomorphismen von \mathbb{R}^n bzw. \mathbb{C}^n besitzt (bezüglich des Standard-Skalarproduktes auf K^n) eine Zerlegung

$$L(K^n) = L_s(K^n) \oplus L_a(K^n) \quad (2.25)$$

in den *reellen* Unterraum $L_s(K^n) := \{\varphi \in L(K^n) \mid \varphi^* = \varphi\}$ selbstadjungierter Endomorphismen und den *reellen* Unterraum $L_a(K^n)$ antiselbstadjungierter Endomorphismen.

2.58 Lemma Die Dimensionen dieser \mathbb{R} -Vektorräume sind

$$\dim_{\mathbb{R}}(L_s(K^n)) = \begin{cases} \frac{1}{2}n(n+1) & , \quad K = \mathbb{R} \\ n^2 & , \quad K = \mathbb{C} \end{cases}$$

und

$$\dim_{\mathbb{R}}(L_a(K^n)) = \begin{cases} \frac{1}{2}n(n-1) & , \quad K = \mathbb{R} \\ n^2 & , \quad K = \mathbb{C} \end{cases} .$$

Beweis: Wir betrachten die darstellenden Matrizen bezüglich der kanonischen Basis des K^n .

- Um eine symmetrische Matrix $A \in \text{Mat}(n, \mathbb{R})$ zu fixieren, müssen wir nur die $\frac{1}{2}n(n+1)$ Einträge oberhalb und einschließlich der Diagonale frei wählen.
- Für eine symmetrische Matrix $A \in \text{Mat}(n, \mathbb{C})$ können wir die $\frac{1}{2}n(n-1)$ komplexen Einträge oberhalb der Diagonale frei wählen, während die Einträge auf der Diagonale reell sein müssen, aber ansonsten frei wählbar sind.
- Die reellen Dimensionen der antisymmetrischen Teilräume ergeben sich aus (2.25):

$$\dim_{\mathbb{R}}(L_a(K^n)) = \dim_{\mathbb{R}}(\text{Mat}_n(K)) - \dim_{\mathbb{R}}(L_s(K^n)). \quad \square$$

Beachtet werden sollte, dass $L_s(\mathbb{C}^n)$ und $L_a(\mathbb{C}^n)$ keine \mathbb{C} -Unterräume sind, denn Multiplikation mit i permutiert diese *reellen* Unterräume.

Ebenso wenig ist im Allgemeinen das Produkt zweier (anti-) selbstadjungierter wieder (anti-) selbstadjungiert.

Im nächsten Kapitel wird die Exponentialabbildung auf Matrizen $A \in \text{Mat}(n, \mathbb{K})$ angewandt. Wir setzen also $\exp(A) := \sum_{n=0}^{\infty} \frac{A^n}{n!}$, und diese matrixwertige Reihe konvergiert.

Wir werden auch sehen, dass für die Vielfachen von A die **Funktionalgleichung** gilt:

$$\exp(t_1 A) \exp(t_2 A) = \exp((t_1 + t_2)A) \quad (t_1, t_2 \in \mathbb{R}).$$

2.59 Satz Für $A \in \text{Mat}(n, K)$, $K = \mathbb{R}$ oder \mathbb{C} , gilt

$$\det(\exp(A)) = \exp(\text{tr}(A)). \quad (2.26)$$

Beweis:

- Ist $A \in \text{Mat}(n, \mathbb{C})$ und ist $J = SAS^{-1}$ Jordan-Normalform, dann ist J eine obere Dreiecksmatrix, sodass

$$\det(\exp(J)) = \prod_{k=1}^n \exp(J_{kk}) = \exp\left(\sum_{k=1}^n J_{kk}\right) = \exp(\text{tr}(J))$$

gilt. Andererseits ist

$$\det(\exp(A)) = \det(S^{-1} \exp(J) S) = \det(\exp(J))$$

und

$$\text{tr}(A) = \text{tr}(S^{-1} J S) = \text{tr}(J),$$

woraus (2.26) folgt.

- Ist $B \in \text{Mat}(n, \mathbb{R})$, dann gilt für die komplexifizierte Matrix $B_{\mathbb{C}}$

$$\det_{\mathbb{R}}(B) = \det_{\mathbb{C}}(B_{\mathbb{C}}) \quad \text{und} \quad \text{tr}_{\mathbb{R}}(B) = \text{tr}_{\mathbb{C}}(B_{\mathbb{C}}),$$

woraus (2.26) auch für $A \in \text{Mat}(n, \mathbb{R})$ folgt. □

2.60 Satz Ist $A \in \text{Mat}(n, K)$ antiselbstadjungiert, dann ist

$$\exp(A) \in \begin{cases} \text{SO}(n) & , \text{ falls } K = \mathbb{R} \\ \text{U}(n) & , \text{ falls } K \in \mathbb{C} \end{cases} .$$

Beweis: Für $O := \exp(A) \in \text{Mat}(n, K)$ gilt $O^* = \exp(A)^* = \exp(A^*) = \exp(-A) = O^{-1}$, denn $\exp(A) \exp(-A) = \exp(0) = \mathbb{1}$. Andererseits folgt für $A = -A^t \in \text{Mat}(n, \mathbb{R})$ wegen $\text{tr} A = -\text{tr}(A^t) = -\text{tr}(A)$ die Identität $\det(\exp(A)) = \exp(\text{tr}(A)) = e^0 = 1$. □

2.61 Beispiel (Antiselbstadjungiert und Drehung) Für $K = \mathbb{R}$ und $n = 2$ ist $A \in \text{Mat}_a(2, \mathbb{R})$ von der Form $A = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ für ein $a \in \mathbb{R}$, und

$$\begin{aligned} \exp(A) &= \sum_{n=0}^{\infty} \frac{A^n}{n!} = \sum_{n=0}^{\infty} \frac{a^n \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^n}{n!} \\ &= \left(\sum_{m=0}^{\infty} \frac{a^{2m} (-1)^m}{(2m)!} \right) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left(\sum_{m=0}^{\infty} \frac{a^{2m+1} (-1)^m}{(2m+1)!} \right) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \cos(a) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin(a) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos(a) & -\sin(a) \\ \sin(a) & \cos(a) \end{pmatrix} . \end{aligned}$$

Es handelt sich also um eine Drehung um den Winkel a . ◇

2.3.6 Anwendung: Quantenmechanik

Zerlegt man das Sonnenlicht mit einem Prisma, dann zeigen sich neben einem kontinuierlichen Spektrum feine Linien. Deren Existenz folgt aus der Tatsache, dass die in einem Atom oder Molekül gebundenen Elektronen nur diskrete Energiewerte annehmen können; man spricht von einer *gequantelten* Energie. Die Vermutung, dass diese reellen Energiewerte Eigenwerte eines selbstadjungierten Operators sind, ist naheliegend und zutreffend. Die Theorie dieser so genannten *Schrödinger-Operatoren* ist die Quantenmechanik.

2.62 Beispiel (Spektrum und Eigenwerte) Für den Fall des *Wasserstoffatoms* ist der unitäre Vektorraum des Elektrons der Raum $L^2(\mathbb{R}^3)$ der quadratintegrierbaren Funktionen⁹ $f: \mathbb{R}^3 \rightarrow \mathbb{C}$ mit dem Skalarprodukt

$$\langle f, g \rangle := \int_{\mathbb{R}^3} \bar{f}(x)g(x)dx.$$

Der Schrödinger-Operator H auf $L^2(\mathbb{R}^3)$ ist von der Form

$$(Hf)(x) := -\frac{1}{2}\Delta f(x) - \frac{f(x)}{|x|},$$

wobei $\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2}$ der so genannte *Laplace-Operator* ist, und die Multiplikation mit dem Potential $-\frac{1}{|x|}$ der Kernladung die elektrostatische Anziehung des Elektrons durch den Kern beschreibt. Man kann nun das Spektrum¹⁰ von H berechnen:

$$\text{spek}(H) = [0, \infty) \cup \{-1/(2n^2) \mid n \in \mathbb{N}\}.$$

In einer geeigneten Energieeinheit entspricht diese Teilmenge von \mathbb{R} genau den Energien, die ein Elektron im Kraftfeld eines Wasserstoffatomkerns (eines Protons) annehmen kann: Für positive Energien ist das Elektron ungebunden und kann beliebige Energien annehmen, während die für das Wasserstoffatom beobachteten Spektralwerte Energiedifferenzen $\frac{1}{2} \left(\frac{1}{m^2} - \frac{1}{n^2} \right)$ entsprechen, die dem Übergang von einem zum anderen gebundenen Zustand entsprechen. \diamond

Ein Hauptziel der Quantenmechanik ist die (meist näherungsweise) Berechnung des Spektrums und der Eigenfunktionen der Schrödinger-Operatoren H . Dies ist auch deshalb wichtig, weil es der Schlüssel zur Berechnung der *Zeitentwicklungsoperatoren*

$$U(t) := \exp(-iHt) \quad (t \in \mathbb{R})$$

⁹Genauer: Äquivalenzklassen solcher Funktionen f_1, f_2 , die sich nur um eine Funktion $f_1 - f_2$ mit $\int_{\mathbb{R}^3} |f_1 - f_2|^2 dx = 0$ unterscheiden.

¹⁰Das **Spektrum** von H ist die Menge $\text{spek}(A) := \{\lambda \in \mathbb{C} \mid A - \lambda \text{Id ist nicht invertierbar}\}$ und enthält alle Eigenwerte von A .

ist, die den quantenmechanischen Anfangszustand φ zur Zeit Null auf den Zustand $U(t)\varphi$ zur Zeit t abbilden. Die Operatoren $U(t)$ sind unitär, wenn H selbstadjungiert ist.

3 Lineare Differentialgleichungen mit konstanten Koeffizienten

Differentialgleichungen (DGLn) modellieren Naturvorgänge. Ein Standardbeispiel ist die Newtonsche Kraftgleichung

$$F(q, \dot{q}) = m\ddot{q}, \quad (3.1)$$

wobei die Kraft F von Ort $q \in \mathbb{R}^d$ und Geschwindigkeit $\dot{q} \in \mathbb{R}^d$ abhängt, und gleich Masse $m > 0$ mal Beschleunigung \ddot{q} ist. Gesucht sind Lösungen, d.h. im idealen Fall glatte Funktionen $q : \mathbb{R} \rightarrow \mathbb{R}^d$, die diese Gleichung erfüllen.

3.1 Beispiel $F(q, \dot{q}) = 0$; $q^{(0)} \in \mathbb{R}^d$, $v^{(0)} \in \mathbb{R}^d$ beliebig. $q(t) := q^{(0)} + v^{(0)}t$ erfüllt die Newtonsche Kraftgleichung (freie Bewegung mit Startpunkt $q^{(0)}$ und Geschwindigkeit $v^{(0)}$ zum Zeitpunkt $t = 0$). \diamond

Mit Methoden der Linearen Algebra behandeln wir *lineare* (homogene, zeitunabhängige) Differentialgleichungen (DGLn), die in der Form

$$\dot{x} = Ax \quad \text{mit} \quad A \in \text{Mat}(n, \mathbb{R})$$

geschrieben werden können. Gesucht ist die *Lösung des Anfangswertproblems* (AWP) für $x_0 \in \mathbb{R}^n$, d.h. eine differenzierbare Funktion

$$x : \mathbb{R} \rightarrow \mathbb{R}^n \quad \text{mit} \quad \frac{d}{dt}x(t) = Ax(t) \quad (t \in \mathbb{R}) \quad \text{und} \quad x(0) = x_0.$$

Die *Ordnung* einer DGL ist die höchste vorkommende Ableitungsstufe. Während die Newtonsche DGL (3.1) von 2. Ordnung ist, betrachten wir also hier DGLn 1. Ordnung.

Die Lösungen sind Kurven im \mathbb{R}^n :

3.2 Definition

- *Stetige Abbildungen* $c = (c_1, \dots, c_n)^t : I \rightarrow \mathbb{R}^n$, deren Definitionsbereich I ein Intervall (mit $I \neq \{a\}$) ist, heißen **Kurven** im \mathbb{R}^n .
- $t \in I$ heißt der **Parameter** von c , ihr Bild $c(I) \subseteq \mathbb{R}^n$ auch die **Spur** von c .

- c heißt **(stetig) differenzierbar**, wenn alle Komponenten c_1, \dots, c_n (stetig) differenzierbar sind.
- Die **Ableitung** oder **Geschwindigkeit** einer stetig differenzierbaren Kurve ist

$$c' : I \rightarrow \mathbb{R}^n, \quad c'(t) := \begin{pmatrix} c'_1(t) \\ \vdots \\ c'_n(t) \end{pmatrix} \quad (t \in I).$$

Oft lassen sich Differentialgleichungen höherer Ordnung auf diesen Fall zurückführen und lösen.

3.3 Beispiel (Freie Bewegung)

Im Fall der freien Bewegung setzen wir $n := 2d$ und schreiben $x = \begin{pmatrix} q \\ v \end{pmatrix}$ mit $q, v \in \mathbb{R}^d$ (Ort bzw. Geschwindigkeit). Dann gilt für $A := \begin{pmatrix} 0 & \mathbb{1} \\ 0 & 0 \end{pmatrix} \in \text{Mat}(n, \mathbb{R})$

$$\dot{x} = Ax,$$

und die (eindeutige) Lösung des Anfangswertproblems mit $x^{(0)} = \begin{pmatrix} q^{(0)} \\ v^{(0)} \end{pmatrix}$ ist

$$x(t) = \begin{pmatrix} q^{(0)} + v^{(0)}t \\ v^{(0)} \end{pmatrix} \quad (t \in \mathbb{R}). \quad \diamond$$

3.1 Berechnung der Lösung

Im Gegensatz zu allgemeinen Differentialgleichungen können wir die linearen DGLn mit algebraischen Mitteln lösen.

Wir definieren dazu die Anwendung der Exponentialfunktion auf eine Matrix $A \in \text{Mat}(n, \mathbb{C})$ durch

$$\exp(A) := \sum_{l=0}^{\infty} \frac{A^l}{l!} := \lim_{k \rightarrow \infty} \sum_{l=0}^k \frac{A^l}{l!}, \quad (3.2)$$

also durch die Einsetzung von A in die Potenzreihe der Exponentialfunktion. Synonym findet man auch die Schreibweise $e^A := \exp(A)$.

In (3.2) ist der Limes-Begriff zu klären. Wir benutzen auf dem \mathbb{R}^n die Euklidische Norm $\|v\| := \sqrt{|v_1|^2 + \dots + |v_n|^2}$ und auf $\text{Mat}(n, \mathbb{C})$ die *Operatornorm*

$$\|A\| := \max_{v \in \mathbb{R}^n \setminus \{0\}} \frac{\|Av\|}{\|v\|} \quad (A \in \text{Mat}(n, \mathbb{C})).$$

$\|A\|$ ist also der maximale¹¹ Streckungsfaktor der linearen Abbildung $v \mapsto Av$.

¹¹Da für $A \in \text{Mat}(n, \mathbb{R})$ $\|A\| = \max_{v \in \mathbb{R}^n, \|v\|=1} \|Av\|$ ist, und die Sphäre $S^{n-1} := \{v \in \mathbb{R}^n \mid \|v\| = 1\}$ kompakt ist, wird das Supremum angenommen. Eine analoge Aussage gilt für $A \in \text{Mat}(n, \mathbb{C})$.

Die Operatornorm auf $\text{Mat}(n, \mathbb{K})$ erfüllt neben den Eigenschaften

$$\|A\| \geq 0, \quad \|A\| = 0 \Leftrightarrow A = 0,$$

$$\|\lambda A\| = |\lambda| \|A\| \quad (\lambda \in \mathbb{K}) \quad \text{und} \quad \|A + B\| \leq \|A\| + \|B\|$$

einer Norm noch die Submultiplikativität

3.4 Satz (Submultiplikativität) $\|AB\| \leq \|A\| \|B\| \quad (A, B \in \text{Mat}(n, \mathbb{K}))$

Beweis: denn

$$\begin{aligned} \|AB\| &= \sup_{v \neq 0} \frac{\|ABv\|}{\|v\|} = \sup_{v: Bv \neq 0} \frac{\|ABv\|}{\|Bv\|} \cdot \frac{\|Bv\|}{\|v\|} \\ &\leq \sup_{w \neq 0} \frac{\|Aw\|}{\|w\|} \cdot \sup_{v \neq 0} \frac{\|Bv\|}{\|v\|} = \|A\| \cdot \|B\|, \end{aligned}$$

außer für $B = 0$, wo sowieso beide Seiten Null sind. □

Nun stellen wir fest, dass die obige Definition von $\exp(A)$ auch sinnvoll ist:

3.5 Satz Die Partialsummen $s_k := \sum_{l=0}^k \frac{A^l}{l!}$ bilden eine Cauchy-Folge.

Beweis: Für $k_1 \geq k_0 \geq \|A\|$ gilt nach der Dreiecksungleichung und Satz 3.4

$$\begin{aligned} \|s_{k_1} - s_{k_0}\| &= \left\| \sum_{l=k_0+1}^{k_1} \frac{A^l}{l!} \right\| \leq \sum_{l=k_0+1}^{k_1} \frac{\|A^l\|}{l!} \leq \sum_{l=k_0+1}^{k_1} \frac{\|A\|^l}{l!} \\ &\leq \frac{\|A\|^{k_0+1}}{(k_0+1)!} \sum_{m=0}^{k_1-k_0-1} \frac{\|A\|^m}{(k_0+1)^m} \leq \frac{\|A\|^{k_0+1}}{(k_0+1)!} \left(1 - \frac{\|A\|}{k_0+1}\right)^{-1}. \end{aligned}$$

Dieser Ausdruck geht für $k_0 \rightarrow \infty$ gegen Null, denn die Fakultät wächst schneller als die (reelle) Exponentialfunktion. □

Nebenbei stellen wir fest, dass wir an keiner Stelle vorausgesetzt haben, dass der lineare Endomorphismus bzw. die Matrix reell ist. Dies ist günstig, denn die Jordan–Normalform von A , und damit von $\exp(At)$, kann auch komplex sein.

Um nun zu sehen, dass

$$t \mapsto x(t) := \exp(At)x_0$$

die eindeutige Lösung des Anfangswertproblems

$$\dot{x} = Ax \quad , \quad x(0) = x_0$$

ist, müssen wir in der Lage sein, die matrixwertige Abbildung

$$\mathbb{R} \rightarrow L(\mathbb{K}^n) \quad , \quad t \mapsto \exp(At)$$

nach dem Zeitparameter zu differenzieren.

Später werden wir uns auch fragen, wie die Lösungen von eventuellen Parametern der linearen DGL abhängen. Für diese Art von Fragestellungen wertet man die Exponentialfunktion für mehr als ein Argument aus, man betrachtet also die Abbildung

$$\exp : L(\mathbb{K}^n) \rightarrow L(\mathbb{K}^n) \quad , \quad A \mapsto \exp(A)$$

in ihrer Abhängigkeit von Argument A . Hier hilft das *Weierstraß-Kriterium*:

3.6 Satz (Weierstraß) *Es sei $(V, \|\cdot\|)$ ein Banach-Raum, $X \subseteq V$ und*

$$f_l : X \rightarrow V \quad (l \in \mathbb{N}_0)$$

seien Funktionen mit $\sup_{x \in X} \|f_l(x)\| \leq a_l$ mit $\sum_{l=0}^{\infty} a_l < \infty$. Dann konvergiert die Reihe $\sum_{l=0}^{\infty} f_l$ auf X absolut und gleichmäßig.

Die Beweismethode für reelle Funktionen lässt sich verallgemeinern:

Beweis: Wegen der Vollständigkeit des metrischen Raumes V haben wir punktweise Konvergenz der Partialsummen $s_l := \sum_{m=0}^l f_m$:

$$s(x) := \lim_{l \rightarrow \infty} s_l(x) \quad (x \in X).$$

Für $\varepsilon > 0$ gibt es ein $m \in \mathbb{N}$ mit $\sum_{l=m+1}^n a_l < \varepsilon$ ($n > m$). Daher folgt

$$\|s_n(x) - s_m(x)\| \leq \sum_{l=m+1}^n a_l < \varepsilon, \quad (x \in X),$$

also gleichmäßige Konvergenz. □

Es gilt nun:

3.7 Satz 1. *Die Exponentialreihe (3.2) ist für alle $a > 0$ und A aus*

$$\{A \in \text{Mat}(n, \mathbb{C}) \mid \|A\| \leq a\}$$

gleichmäßig konvergent.

2. Für $B, C \in \text{Mat}(n, \mathbb{C})$ mit $BC = CB$ gilt

$$\exp(B + C) = \exp(B) \exp(C). \quad (3.3)$$

Insbesondere ist

$$\boxed{\exp(At_1) \exp(At_2) = \exp(A(t_1 + t_2))}, \quad (t_1, t_2 \in \mathbb{R}) \quad (3.4)$$

3. Für $A \in L(\mathbb{K}^n)$ ist die Abbildung

$$\mathbb{R} \rightarrow L(\mathbb{K}^n) \quad , \quad t \mapsto \exp(At)$$

stetig differenzierbar, und es gilt

$$\boxed{\frac{d}{dt} \exp(At) = A \exp(At)}. \quad (t \in \mathbb{R}). \quad (3.5)$$

4. $t \mapsto x(t) := \exp(At)x_0$ ist die eindeutige Lösung des Anfangswertproblems $\dot{x} = Ax$, $x(0) = x_0$ die Differenz $z(t) := y(t) - x(t)$ das Anfangswertproblem $\dot{z} = Az$, $z(0) = 0$ löst.

Beweis:

- Setzen wir in unserem Fall $f_l : \underbrace{L(\mathbb{K}^n)}_V \rightarrow L(\mathbb{K}^n)$, $A \mapsto \frac{A^l}{l!}$, so ist zwar $\sup_{A \in V} \|f_l(A)\|$ (außer für $l = 0$) gleich ∞ . Wählen wir aber als Definitionsbereich die Vollkugel $X := \{A \in V \mid \|A\| \leq c\}$ mit Radius $c > 0$, dann ist $a_l := \sup_{A \in X} \|f_l(A)\| = \frac{1}{l!} \sup_{A \in X} \|A^l\| \leq \frac{c^l}{l!}$, und die Reihe $\sum_{l=0}^{\infty} a_l \leq e^c$ konvergiert für jede Wahl von c .
- Durch Einsetzen in die Definition von \exp :

$$\begin{aligned} \exp(B+C) &= \sum_{n=0}^{\infty} \frac{1}{n!} (B+C)^n \stackrel{(!)}{=} \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^n \binom{n}{i} B^i C^{n-i} \\ &= \sum_{n=0}^{\infty} \sum_{i=0}^n \frac{1}{i!(n-i)!} B^i C^{n-i} = \left(\sum_{i=0}^{\infty} \frac{1}{i!} B^i \right) \left(\sum_{j=0}^{\infty} \frac{1}{j!} C^j \right) = \exp(B) \exp(C). \end{aligned}$$

Da Vielfache einer Matrix miteinander kommutieren, folgt (3.4) aus (3.3).

- Als Limes gleichmäßig konvergenter stetiger Funktionen $s_l \upharpoonright_X : X \rightarrow V$ ist die Exponentialabbildung auch stetig, und, da auch die Ableitungen Ds_l auf X gleichmäßig konvergent sind, auch differenzierbar. Damit ist auch die Formel (3.5) richtig, und aus ihr ergibt sich auch die Stetigkeit der Ableitung.
- Die Eindeutigkeitsaussage der Lösung einer DGL wird später viel allgemeiner behandelt. \square

3.8 Bemerkungen (Funktionalgleichung von \exp)

- Als Gegenbeispiel zu (3.3) gilt für die Dreiecksmatrizen $B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ und $C := B^t = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$:

$$\exp(Bt) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad \exp(Ct) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

also

$$\exp(Bt) \exp(Ct) = \begin{pmatrix} 1+t^2 & t \\ t & 1 \end{pmatrix} \quad \text{und} \quad \exp(Ct) \exp(Bt) = \begin{pmatrix} 1 & t \\ t & 1+t^2 \end{pmatrix}.$$

Dagegen folgt aus $(B+C)^n = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & n \text{ gerade} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & n \text{ ungerade} \end{cases}$

$$\exp((B+C)t) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sum_{m=0}^{\infty} \frac{t^{2m}}{(2m)!} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sum_{m=0}^{\infty} \frac{t^{2m+1}}{(2m+1)!} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

- Wegen (3.4) bildet der die DGL $\dot{x} = Ax$ lösende lineare Fluss

$$\Phi_t : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \Phi_t(x) = \exp(At)x \quad (t \in \mathbb{R})$$

eine einparametrische Gruppe. \diamond

3.9 Beispiele (Matrix-Exponential)

1. $A = \text{diag}(\lambda_1, \dots, \lambda_n)$. Dann ist $A^l = \text{diag}(\lambda_1^l, \dots, \lambda_n^l)$, also

$$\exp(At) = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \quad \text{und} \quad x(t) = \begin{pmatrix} e^{\lambda_1 t} x_1^{(0)} \\ \vdots \\ e^{\lambda_n t} x_n^{(0)} \end{pmatrix}.$$

2. $A = N_n$ (nilpotente Blockmatrix). Dann ist $\exp(At) = \sum_{l=0}^n \frac{(At)^l}{l!}$, also z.B. für $n = 4$

$$\exp(N_4 t) = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 \\ 0 & 1 & t & t^2/2 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

3. $A = A_1 \oplus A_2$. Dann ist $A^l = A_1^l \oplus A_2^l$, also $\exp(At) = \exp(A_1 t) \oplus \exp(A_2 t)$. \diamond

Diese Beispiele ermöglichen zusammengenommen die Exponentialfunktion einer Jordan-Matrix

$$J_{r_1}(\lambda_1) \oplus \dots \oplus J_{r_k}(\lambda_k) \in \text{Mat}(r, \mathbb{C})$$

zu bilden, denn $J_r(\lambda) = \lambda \mathbb{1}_r + N_r$, also

$$\begin{aligned} \exp(J_r(\lambda)t) &= \exp(\lambda t \mathbb{1} + N_r t) \\ &= \exp(\lambda t \mathbb{1}) \cdot \exp(N_r t) = e^{\lambda t} \exp(N_r t). \end{aligned}$$

Die Exponentialfunktion einer beliebigen reellen Matrix $A \in \text{Mat}(n, \mathbb{R})$ bilden wir, indem wir sie mittels $S \in \text{GL}(n, \mathbb{C})$ zu einer komplexen Jordan-Matrix

$$J = S^{-1} A S$$

konjugieren. Dann ist

$$\exp(At) = \exp(S J S^{-1} t) = S \exp(J t) S^{-1},$$

denn bei der Bildung der Potenzen

$$(S J S^{-1})^l = S J S^{-1} \cdot \dots \cdot S J S^{-1}$$

fallen die inneren Paare $S^{-1} S = \mathbb{1}_n$ heraus.

3.2 * Lineare Flüsse in der Ebene

Die Summe der algebraischen Vielfachheiten der Eigenwerte λ mit $\operatorname{Re}(\lambda) < 0$ heißt der **Index** von A und wird $\operatorname{Ind}(A)$ geschrieben.

Im Limes von Zeiten $t \rightarrow \infty$ konvergieren die Lösungen mit Anfangswerten im entsprechenden Unterraum gegen Null, weshalb dieser auch der *stabile Unterraum* genannt wird. Wir schauen uns dies im Fall der Dimension 2 genauer an.

Zueinander ähnliche Matrizen führen zu Flüssen, die sich nur durch eine Basis-Transformation unterscheiden.

Wir betrachten für $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \operatorname{Mat}(2, \mathbb{R})$ den Fluss $\Phi_t^{(A)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ der linearen DGL $\dot{x} = Ax$. Die Größen

$$\operatorname{tr}(A) = a_{11} + a_{22} \quad , \quad \det(A) = a_{11}a_{22} - a_{12}a_{21} \quad \text{und} \quad D := \operatorname{tr}(A)^2 - 4\det(A)$$

sind invariant unter Konjugationen $A \mapsto SAS^{-1}$, und die Eigenwerte $\lambda_{1/2} \in \mathbb{C}$ sind gleich

$$\lambda_{1/2} = \frac{1}{2}(\operatorname{tr}(A) \pm \sqrt{D}).$$

Nur wenn die Diskriminante $D = 0$ ist, kann also die komplexe Jordan-Normalform von A aus einem Jordan-Block der Größe 2 bestehen, und nur dann ist die Konjugations-Klasse von A nicht schon durch $\operatorname{tr}(A)$ und $\det(A)$ festgelegt.

Nicht hyperbolisch ist die Matrix A genau dann, wenn mindestens einer der Eigenwerte verschwindenden Realteil hat. Dies ist genau dann der Fall, wenn

1. $\det(A) = 0$, also sogar ein Eigenwert 0 ist oder
2. $\det(A) > 0$, aber $\operatorname{tr}(A) = 0$ gilt, also die Eigenwerte rein imaginär sind.

In der $(\operatorname{tr}, \det) \in \mathbb{R}^2$ -Ebene bilden diese Bedingungen die Abszisse bzw. positive Ordinate und trennen damit drei offene Gebiete ab, siehe Abb. 3.1.

- $\operatorname{Ind}(A) = 0$ gilt für den Quadranten $\det(A) > 0 < \operatorname{tr}(A)$
- $\operatorname{Ind}(A) = 1$ gilt für $\det(A) < 0$. Hier sind beide Eigenwerte reell.
- $\operatorname{Ind}(A) = 2$ entspricht dem Quadranten $\det(A) > 0 > \operatorname{tr}(A)$.

Wie im letzten Abschnitt bewiesen, sind die Flüsse innerhalb jedes dieser drei Gebiete untereinander konjugiert, aber Flüsse für Matrizen mit verschiedenen Indices sind nicht konjugiert.

Der Fall $\operatorname{Ind}(A) = 1$ ist, entsprechend dem Vorzeichen von $\operatorname{tr}(A)$, noch weiter unterteilbar. In der Situation $\operatorname{tr}(A) = 0$ zweier Eigenwerte $\lambda_1 = -\lambda_2 \in \mathbb{R}$ wird das Phasenraumvolumen durch den Fluss erhalten, während er für $\operatorname{tr}(A) < 0$ im Limes großer Zeiten mit exponentieller Rate gegen Null geht. Der Fall $\operatorname{tr}(A) > 0$ ist in Abb. 3.2 dargestellt.

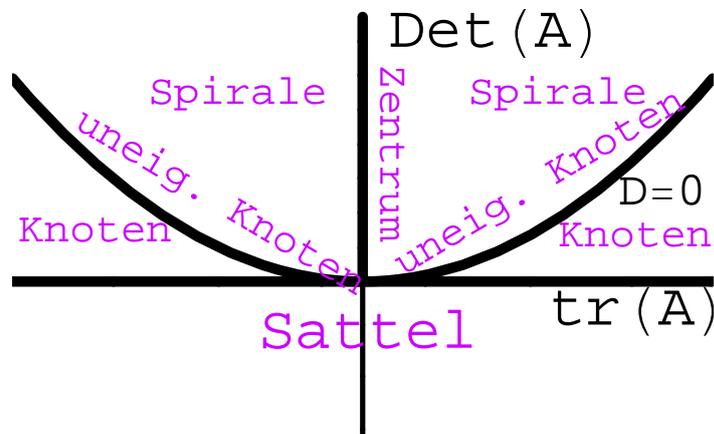


Abbildung 3.1: Verzweigungsdiagramm für Matrizen $A \in \text{Mat}(2, \mathbb{R})$, mit Diskriminante $D = \text{tr}(A)^2 - 4 \det(A)$.

$\text{tr}(A) < 0$: Senke; $\text{tr}(A) = 0$: Volumenerhaltender Fluss; $\text{tr}(A) > 0$: Quelle

Die durch die Gleichung $D = 0$ definierte Parabel trennt die Konjugationsklassen $\text{Ind}(A) = 0$ und $\text{Ind}(A) = 2$ noch weiter auf. Für $D > 0$, also reelle Eigenwerte, erhalten wir sog. *Knoten* als Phasenraumporträts, siehe Abb. 3.3. Diese werden *stabil* genannt, wenn $\dim(E^s) = 2$, also $\text{Ind}(A) = 2$ ist und *instabil* für $\text{Ind}(A) = 0$.

Für $D = 0$ kann die Jordan–Normalform von A ein nichttrivialer Jordan–Block sein. Ein entsprechendes Phasenporträt, *uneigentlicher Knoten* genannt, findet sich ebenfalls in Abb. 3.3.

Ist zusätzlich $\text{tr}(A) = 0$, sind also beide Eigenwerte gleich Null, erhält man einen eindimensionalen Eigenraum von Gleichgewichtslagen, wie in Abb. 3.4 (links). Der Fall $\text{tr}(A) = 0$, $\det(A) > 0$ führt zu imaginären Eigenwerten und periodischen Orbits (Abb. 3.4 rechts), auch *Zentren* genannt.

Endlich ist für $D < 0$ und $\text{tr}(A) < 0$ die Bewegung spiralförmig und stabil (Abb. 3.5), während $D < 0$ und $\text{tr}(A) > 0$ zu sog. *instabilen Spiralen* führt.

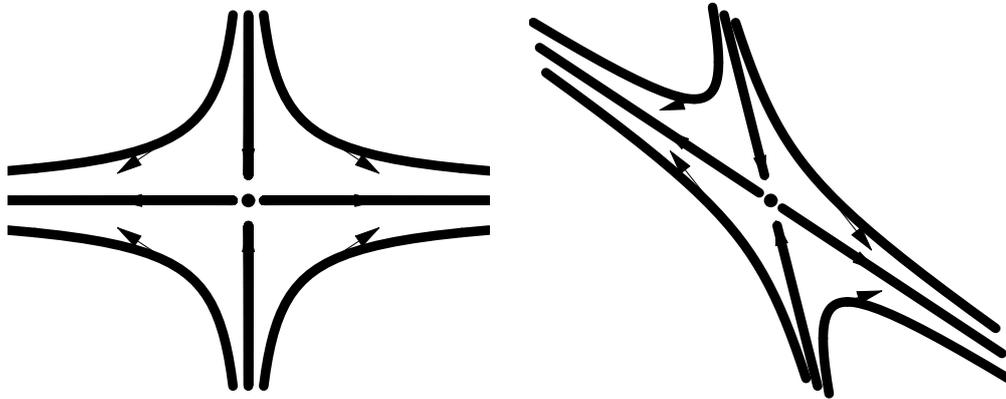


Abbildung 3.2: Phasenporträt von Satteln der DGL $\dot{x} = Ax$. Links: Systemmatrix $A = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$; Rechts: zu A ähnliche Matrix.

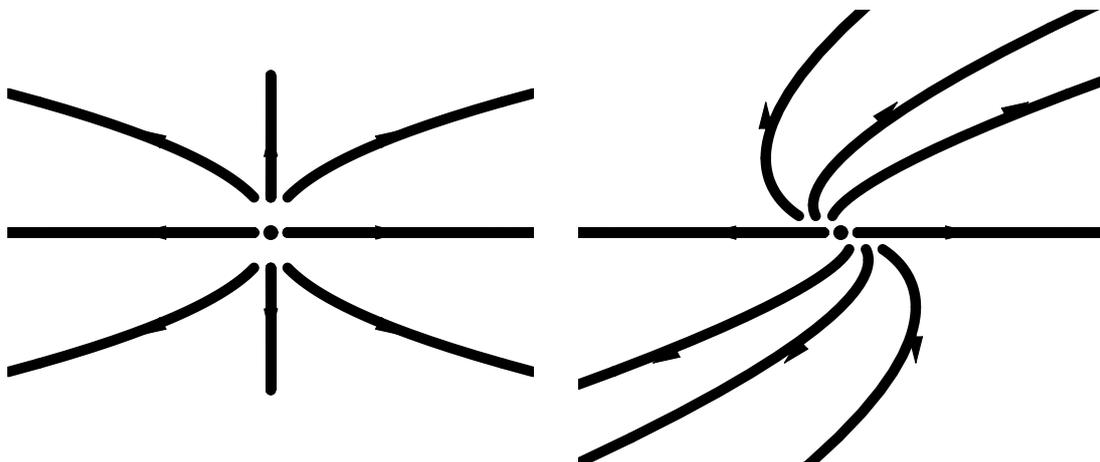


Abbildung 3.3: Phasenporträts von Knoten der DGL $\dot{x} = Ax$. Links: instabiler Knoten, für $A = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}$; Rechts: instabiler uneigentlicher Knoten, für $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

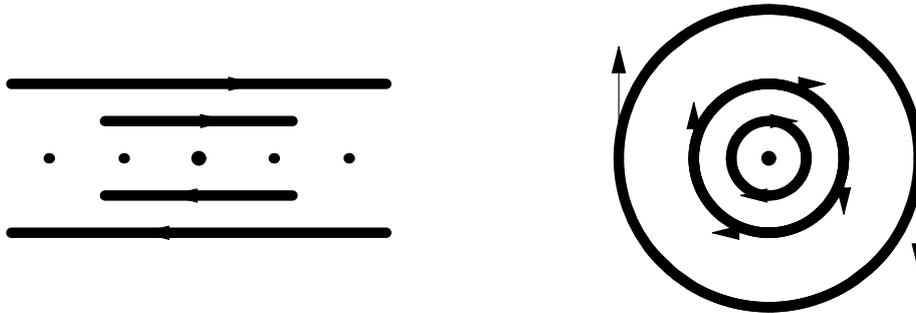


Abbildung 3.4: Phasenporträts von $\dot{x} = Ax$. Fall rein imaginärer Eigenwerte. Links: Nilpotente Matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; Rechts: Zentrum, für antisymmetrische Matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

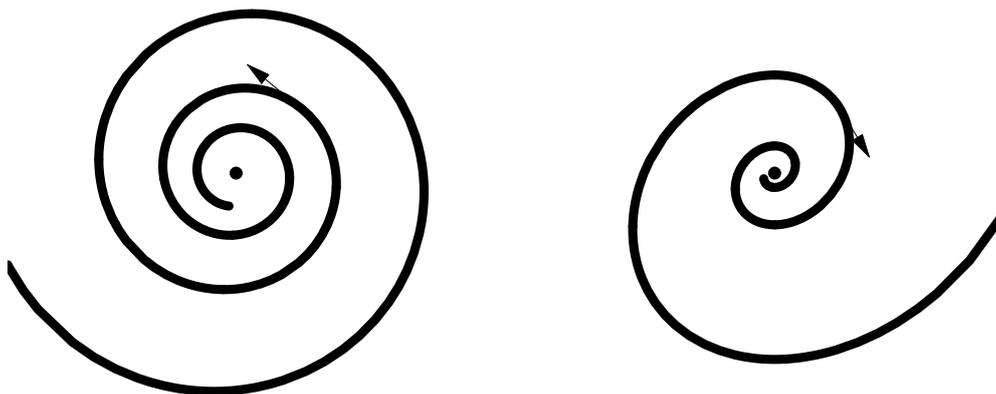


Abbildung 3.5: Phasenporträts von stabilen Spiralen der DGL $\dot{x} = Ax$. Links: $A = \begin{pmatrix} -1/5 & -1 \\ 1 & -1/5 \end{pmatrix}$; rechts: eine nicht zu A ähnliche Matrix

3.3 * Ein Beispiel

Es soll die lineare Differentialgleichung

$$\dot{q}_1 = v_1 \quad , \quad \dot{q}_2 = v_2 \quad , \quad \dot{v}_1 = q_2 - q_1 \quad , \quad \dot{v}_2 = q_1 - q_2$$

gelöst werden. Diese beschreibt die Bewegung zweier Massenpunkte auf der Geraden mit Orten q_i und Geschwindigkeiten v_i , die sich gegenseitig mit einer zu ihrem Abstand proportionalen Kraft anziehen.

Schreiben wir $(x_1, x_2, x_3, x_4) = (v_1, v_2, q_1, q_2)$, dann entspricht diese DGL 2. Ordnung $\dot{x} = Mx$ mit

$$M := \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Wir stellen die Matrix m des Dgl-Systems auf.

Die Koordinaten sind die Geschwindigkeiten und Orte (v_1, v_2, q_1, q_2) der beiden Massenpunkte.

$$m = \{\{0, 0, -1, 1\}, \{0, 0, 1, -1\}, \{1, 0, 0, 0\}, \{0, 1, 0, 0\}\}; \text{MatrixForm}[m]$$

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

Wir berechnen die Eigenwerte und Eigenvektoren von m .

Die erste Klammer von Eigensystem[m] enthält die Nullstellen des charakteristischen Polynoms von m mit ihren Multiplizitäten.

Die zweite Klammer enthält die dazugehörigen Eigenvektoren.

Eigensystem[m]

$$\{\{0, 0, -i\sqrt{2}, i\sqrt{2}\}, \{0, 0, 1, 1\}, \{0, 0, 0, 0\}, \{i\sqrt{2}, -i\sqrt{2}, -1, 1\}, \{-i\sqrt{2}, i\sqrt{2}, -1, 1\}\}$$

Der erste Eigenvektor $\{0,0,1,1\}$ zum Eigenwert 0 entspricht einer Verschiebung der beiden Orte der Massenpunkte um den gleichen Wert.

Da m auf dem verallgemeinerten Eigenraum $\hat{E}(0)$ nicht diagonalisierbar ist, wird kein zweiter Eigenvektor zum Eigenwert 0 gefunden.

Daher berechnen wir die (Eigenwerte und) Eigenvektoren von m^2 :

Eigensystem[$m.m$]

$$\{\{-2, -2, 0, 0\}, \{0, 0, -1, 1\}, \{-1, 1, 0, 0\}, \{0, 0, 1, 1\}, \{1, 1, 0, 0\}\}$$

Dabei ändert sich die Reihenfolge der (quadrierten) Eigenwerte.

Der zweite Eigenvektor $\{1,1,0,0\}$ zum Eigenwert 0 entspricht einer Bewegung beider Massen mit gleicher Anfangsgeschwindigkeit vom Nullpunkt weg.

Die Linearkombination $\{0,0,-1,1\} + \{-1,1,0,0\} = \{-1,1,-1,1\}$ der beiden gefundenen Eigenvektoren zum Eigenwert -2 entspricht Anfangsbedingungen,

bei denen sich die beiden Massenpunkte von ihrem Schwerpunkt bei 0 entfernen,

während sie sich bei der Linearkombination $\{0,0,-1,1\} - \{-1,1,0,0\} = \{1,-1,-1,1\}$ diesem annähern.

Wir wenden die Exponentialfunktion auf die Matrix m an:

em = FullSimplify[MatrixExp[$m t$]]; MatrixForm[em]

$$\begin{pmatrix} \cos\left[\frac{t}{\sqrt{2}}\right]^2 & \sin\left[\frac{t}{\sqrt{2}}\right]^2 & -\frac{\sin[\sqrt{2} t]}{\sqrt{2}} & \frac{\sin[\sqrt{2} t]}{\sqrt{2}} \\ \sin\left[\frac{t}{\sqrt{2}}\right]^2 & \cos\left[\frac{t}{\sqrt{2}}\right]^2 & \frac{\sin[\sqrt{2} t]}{\sqrt{2}} & -\frac{\sin[\sqrt{2} t]}{\sqrt{2}} \\ \frac{1}{4} (2t + \sqrt{2} \sin[\sqrt{2} t]) & \frac{1}{4} (2t - \sqrt{2} \sin[\sqrt{2} t]) & \cos\left[\frac{t}{\sqrt{2}}\right]^2 & \sin\left[\frac{t}{\sqrt{2}}\right]^2 \\ \frac{1}{4} (2t - \sqrt{2} \sin[\sqrt{2} t]) & \frac{1}{4} (2t + \sqrt{2} \sin[\sqrt{2} t]) & \sin\left[\frac{t}{\sqrt{2}}\right]^2 & \cos\left[\frac{t}{\sqrt{2}}\right]^2 \end{pmatrix}$$

Nun wollen wir die allgemeine Lösung des Differentialgleichungssystems bestimmen. Diese hängt von den Konstanten $C[1], C[2], C[3]$ und $C[4]$ ab.

```
FullSimplify[DSolve[{q1''[t] == q2[t] - q1[t], q2''[t] == q1[t] - q2[t]}, {q1[t], q2[t]}, t]]
```

$$\left\{ \left\{ q_1[t] \rightarrow \frac{1}{4} \left(2 (t (C[1] + C[2]) + C[3] + C[4]) + 2 (C[3] - C[4]) \cos[\sqrt{2} t] + \sqrt{2} (C[1] - C[2]) \sin[\sqrt{2} t] \right), \right. \right.$$

$$\left. \left. q_2[t] \rightarrow \frac{1}{4} \left(2 (t (C[1] + C[2]) + C[3] + C[4]) + 2 (-C[3] + C[4]) \cos[\sqrt{2} t] + \sqrt{2} (-C[1] + C[2]) \sin[\sqrt{2} t] \right) \right\} \right\}$$

Jetzt geben wir Anfangsorte und Anfangsgeschwindigkeiten an und lassen das Dgl-System lösen:

```
DSolve[{q1''[t] == q2[t] - q1[t], q2''[t] == q1[t] - q2[t],
q1[0] == 0, q2[0] == 1, q1'[0] == 1/10, q2'[0] == 0}, {q1[t], q2[t]}, t]
```

$$\left\{ \left\{ q_1[t] \rightarrow \frac{1}{80} e^{-i\sqrt{2} t} \left(-20 + i\sqrt{2} + 40 e^{i\sqrt{2} t} - 20 e^{2i\sqrt{2} t} - i\sqrt{2} e^{2i\sqrt{2} t} + 4 e^{i\sqrt{2} t} t \right), \right. \right.$$

$$\left. \left. q_2[t] \rightarrow \frac{1}{80} e^{-i\sqrt{2} t} \left(20 - i\sqrt{2} + 40 e^{i\sqrt{2} t} + 20 e^{2i\sqrt{2} t} + i\sqrt{2} e^{2i\sqrt{2} t} + 4 e^{i\sqrt{2} t} t \right) \right\} \right\}$$

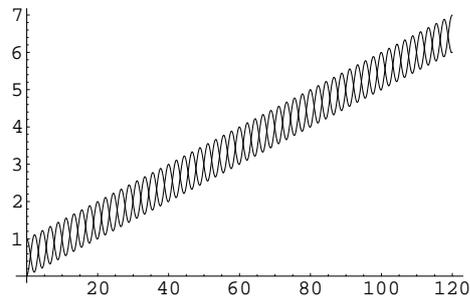
Wir haben vergessen, die Trigonometrischen Additionstheoreme anzuwenden, um die Ausdrücke zu vereinfachen:

```
x = FullSimplify[Evaluate[{q1[t], q2[t]} /. %]]
```

$$\left\{ \left\{ \frac{1}{40} \left(20 + 2 t - 20 \cos[\sqrt{2} t] + \sqrt{2} \sin[\sqrt{2} t] \right), \frac{1}{40} \left(20 + 2 t + 20 \cos[\sqrt{2} t] - \sqrt{2} \sin[\sqrt{2} t] \right) \right\} \right\}$$

q_1 und q_2 werden bis zur Zeit t_{\max} gezeichnet.

```
Plot[Evaluate[x], {t, 0, tmax}, PlotPoints -> 300]
```



- Graphics -

Warum bemühen wir eigentlich die Lineare Algebra, statt die Differentialgleichungen gleich numerisch zu lösen? Weil wir dann auf einen Schlag die Lösung für alle Zeiten finden. Numerik aber verschlechtert sich mit der Zeit. Hier explodiert aus mir unbekanntem Gründen die numerische Lösung zur Zeit $t=109$:

```

tmax = 120;
NDSolve[{Q1''[t] == Q2[t] - Q1[t], Q2''[t] == Q1[t] - Q2[t],
  Q1[0] == 0, Q2[0] == 1, Q1'[0] == 0.1, Q2'[0] == 0}, {Q1, Q2}, {t, 0, tmax}]

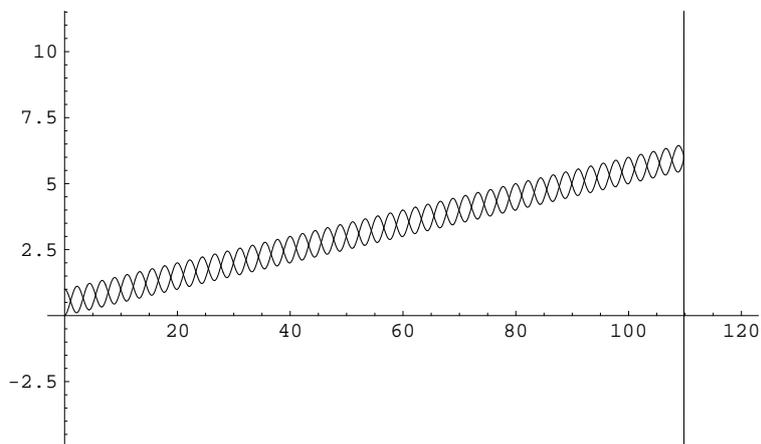
NDSolve::mxst : Maximum number of 1000 steps reached at the point t == 109.68734614586384`.

{{Q1 -> InterpolatingFunction[{{0., 109.687}}, <>],
  Q2 -> InterpolatingFunction[{{0., 109.687}}, <>]}}

Plot[Evaluate[{Q1[t], Q2[t]} /. %], {t, 0, tmax}, PlotPoints -> 300]

InterpolatingFunction::dmval : Input value {109.754} lies
  outside the range of data in the interpolating function. Extrapolation will be used.
InterpolatingFunction::dmval : Input value {110.169} lies
  outside the range of data in the interpolating function. Extrapolation will be used.
InterpolatingFunction::dmval : Input value {109.955} lies
  outside the range of data in the interpolating function. Extrapolation will be used.
General::stop : Further output of InterpolatingFunction::dmval will be suppressed during this calculation.

```



- Graphics -

4 Quadratische Formen

Bis jetzt wurden in dieser Vorlesung hauptsächlich solche Teilmengen $M \subset V$ von \mathbb{K} -Vektorräumen V betrachtet, die selbst Vektorräume sind, oder bei denen zumindest die Menge $\{m_1 - m_2 \mid m_1, m_2 \in M\}$ aller Differenzvektoren einen Unterraum von V bildet (nicht leerer *affiner Unterraum* M). Allgemein treten affine Unterräume als Lösungsmengen linearer Gleichungssysteme auf.

4.1 Affinitäten und Kongruenzen

Jetzt werden wir verallgemeinert Quadriken untersuchen. Beispiele solcher Quadriken für $V = \mathbb{R}^2$ sind neben den Geraden die Ellipsen, Parabeln und Hyperbeln.

Wie bisher betrachten wir endlich-dimensionale Vektorräume über dem Körper $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$, also ohne Beschränkung der Allgemeinheit den \mathbb{K}^n . Den \mathbb{K} -Vektorraum der Polynome in n Unbestimmten nezeichnen wir mit $\mathbb{K}[x_1, \dots, x_n]$.

4.1 Definition Eine Teilmenge $Q \subset \mathbb{K}^n$ heißt **Quadrik**, wenn ein quadratisches Polynom $p \in \mathbb{K}[x_1, \dots, x_n]$ mit

$$Q = \{(x_1, \dots, x_n) \in \mathbb{K}^n \mid p(x_1, \dots, x_n) = 0\}$$

existiert.

Wir können nun mit $x := (x_1, \dots, x_n)^t$ dieses Polynom in der Form

$$p(x) = x^t A x + b^t x + c \quad (x \in \mathbb{K}^n)$$

schreiben, wobei $A \in \text{Mat}(n, k)$, $b \in \mathbb{K}^n$ und $c \in \mathbb{K}$ eindeutig bestimmt sind, wenn wir $A = A^t$ (also die Symmetrie von A fordern (Letzteres können wir immer durch Übergang von A zu $\frac{1}{2}(A + A^t)$ erreichen).

Man sieht dem Polynom p nicht direkt die geometrische Form der Quadrik Q an. Überhaupt müssen wir uns zunächst über diese Formen ins Bild setzen. Dazu ist es naheliegend, eine Koordinatentransformation des \mathbb{K}^n vorzunehmen, die das Polynom p vereinfacht. In Frage kommen zunächst die so genannten Affinitäten:

4.2 Definition Eine Abbildung $f : V \rightarrow W$ von \mathbb{K} -Vektorräumen heißt **affin**, wenn die Abbildung

$$\tilde{f} : V \rightarrow W \quad \tilde{f}(x) := f(x) - f(0)$$

linear ist. Sie heißt **Affinität**, wenn sie zusätzlich bijektiv ist.

Die Affinitäten $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ bilden eine Gruppe, denn jedes solche f besitzt die Form $f(x) = Ax + v$ mit $A \in \text{GL}(n, \mathbb{K})$ und $v \in \mathbb{K}^n$, sodass

$$f^{-1}(y) = A^{-1}y - A^{-1}v \quad \text{und} \quad f \circ g(x) = ABx + (v + Aw)$$

für $g(x) = Bx + w$ gilt.

4.3 Beispiel (Affinitäten) Beispiele für Affinitäten des \mathbb{R}^n sind die Drehungen $f(x) = Ox$ mit $O \in \text{SO}(n)$, die Streckungen $f(x) = cx$ um einen Faktor $c \in \mathbb{R} \setminus \{0\}$ oder die Translationen $f(x) = x + v$ um den Vektor $v \in \mathbb{R}^n$. \diamond

Es ist nun rationell, sowohl die Affinitäten als auch die Quadriken des \mathbb{K}^n im \mathbb{K}^{n+1} darzustellen. Ist nämlich

$$f(x) = Ax + v \quad (x \in \mathbb{K}^n)$$

eine Affinität, dann gilt für die erweiterte Matrix

$$E_f := \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in \text{Mat}(n+1, \mathbb{K})$$

die Eigenschaft eines Gruppenhomomorphismus

$$E_g E_f = \begin{pmatrix} 1 & 0 \\ w & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w+Bv & BA \end{pmatrix} = E_{g \circ f},$$

(wobei $g(x) = Bx + w$ eine weitere Affinität des \mathbb{K}^n sei).

Ordnen wir dem Vektor $x \in \mathbb{K}^n$ den erweiterten Vektor $E_x := \begin{pmatrix} 1 \\ x \end{pmatrix} \in \mathbb{K}^{n+1}$ zu, dann gilt

$$E_f E_x = \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = \begin{pmatrix} 1 \\ Ax+v \end{pmatrix} = E_{f(x)}. \quad (4.1)$$

Ordnen wir andererseits dem quadratischen Polynom $p \in \mathbb{K}[x_1, \dots, x_n]$ der Form

$$p(x) = (x, Ax) + (b, x) + c$$

(mit der Bilinearform $(b, x) = b_1 x_1 + \dots + b_n x_n$) die Matrix

$$S_p := \begin{pmatrix} c & b^t/2 \\ b/2 & (A+A^t)/2 \end{pmatrix} \in \text{Mat}(n+1, \mathbb{K})$$

zu, dann ist S_p symmetrisch ($S_p^t = S_p$), und es gilt

$$p(x) = (E_x, S_p E_x). \quad (4.2)$$

Wir wollen nun durch eine Affinität f des \mathbb{K}^n die p zugeordnete Quadrik $p^{-1}(0) \subset \mathbb{K}^n$ auf eine möglichst einfache Form bringen. D.h. wir substituieren $x = f(y)$ in $p(x)$. Wegen (4.2) und (4.1) bedeutet dies

$$\begin{aligned} p(f(y)) &= (E_{f(y)}, S_p E_{f(y)}) = (E_f E_y, S_p E_f E_y) \\ &= (E_y, S'_p E_y) \quad \text{mit} \quad S'_p := E_f^t S_p E_f. \end{aligned} \quad (4.3)$$

Zur Lösung dieses Normalformproblemles für den Spezialfall $\mathbb{K} = \mathbb{R}$ schauen wir uns zunächst das verwandte Problem der Normalform unter orientierungserhaltenden Kongruenzen an:

4.4 Definition Eine Affinität $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt

- **Kongruenz**, wenn sie Abstände erhält, d.h.

$$\|f(x) - f(y)\| = \|x - y\| \quad (x, y \in \mathbb{R}^n). \quad (4.4)$$

- Sie heißt **orientierungserhaltend**, wenn $\det(A) > 0$ (mit $f(x) = Ax + v$) gilt.

Eine Affinität $f(x) = Ax + v$ ist genau dann eine Kongruenz, wenn $A \in O(n)$ gilt, und genau dann eine orientierungserhaltende Kongruenz, wenn $A \in SO(n)$.

Weiter kann man zeigen, dass jede abstandserhaltende Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (für die also (4.4) gilt) eine Affinität ist.

4.2 Normalformen von Quadriken

Wir wollen jetzt Quadriken — zunächst durch Kongruenzen, danach allgemeiner durch Affinitäten — auf eine möglichst einfache Normalform bringen.

4.5 Beispiel (Ellipsen) Beispielsweise wird dabei eine Ellipse um den Nullpunkt zentriert, und die Hauptachsen werden entlang der Abszisse bzw. Ordinate ausgerichtet. Verschiedene Ellipsen unterscheiden sich danach noch durch die Längen ihrer Hauptachsen, also zwei reelle Parameter.

Läßt man statt Kongruenzen allgemeine Affinitäten zu, dann kann man durch Skalierung der Abszissen- und Ordinatenrichtung alle Ellipsen in einen Kreis vom Radius Eins überführen. \diamond

Das Beispiel verallgemeinert sich für Kongruenzen folgendermaßen:

4.6 Satz Ist $p \in \mathbb{R}[x_1, \dots, x_n]$ ein quadratisches Polynom der Form $p(x) = (x, Ax) + (b, x) + c$ mit $A = A^t$ und $m := \text{rg}(A)$, dann existiert eine orientierungserhaltende Kongruenz f des \mathbb{R}^n mit

$$p \circ f(y) = \begin{cases} \sum_{i=1}^m \lambda_i y_i^2 & , \text{rg}(E_p) = m \\ \sum_{i=1}^m \lambda_i y_i^2 + c & , \text{rg}(E_p) = m + 1 \\ \sum_{i=1}^m \lambda_i y_i^2 + 2\|b\|y_{m+1} & , \text{rg}(E_p) = m + 2. \end{cases}$$

Dabei sind $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ die von Null verschiedenen Eigenwerte der symmetrischen Matrix A .

Beweis: Gemäß (4.3) besitzt das quadratische Polynom $p \circ f$ die Form $p \circ f(y) = (E_y, S'_p E_y)$ mit $S'_p = E_f^t S_p E_f$. Wir setzen f aus vier orientierungserhaltenden Kongruenzen zusammen: $f = f_4 \circ f_3 \circ f_2 \circ f_1$.

1. Dabei ist $f_1(x) := O_1 x$, wobei wir zunächst nur $O_1 \in O(n)$ fordern. Damit ist die Transponierte von $E_{f_1} = \begin{pmatrix} 1 & 0 \\ 0 & O_1 \end{pmatrix}$ gleich $E_{f_1}^t = \begin{pmatrix} 1 & 0 \\ 0 & O_1^{-1} \end{pmatrix}$ und es gibt ein O_1 mit

$$E_{f_1}^t S_p E_{f_1} = \begin{pmatrix} c & \tilde{b}^t \\ \tilde{b} & D \end{pmatrix} \quad \text{und} \quad D = \text{diag}(\lambda_1, \dots, \lambda_m, 0, \dots, 0).$$

Ist $O_1 \notin SO(n)$, also f_1 nicht orientierungserhaltend, dann können wir dies durch Übergang zur Drehmatrix $\text{diag}(-1, 1, \dots, 1)O_1 \in SO(n)$ erzwingen. Ist $\tilde{b} = 0$, dann ist der Satz damit bewiesen.

2. Wir wählen f_2 von der Form $f_2(x) = x + v_2$ mit

$$(v_2)_i := \begin{cases} -\tilde{b}_i/\lambda_i & , \quad i \leq m \\ 0 & , \quad i > m. \end{cases}$$

Dann ist $E_{f_2 \circ f_1}^t S_p E_{f_2 \circ f_1} = \begin{pmatrix} c & \hat{b}^t \\ \hat{b} & D \end{pmatrix}$, wobei $\hat{b}_1 = \dots = \hat{b}_m = 0$ ist.

3. Wegen der Annahme $\tilde{b} \neq 0$ ist auch $\hat{b} \neq 0$. Es gibt aber eine Drehung $O_3 \in \text{SO}(n)$ von der Form $O_3 = \mathbb{1}_m \oplus \tilde{O}_3$ mit $\tilde{O}_3 \in \text{SO}(n-m)$, die $O_3 \hat{b} = \mu \cdot e_{m+1}$ mit $\mu := \|\hat{b}\|$ bewirkt, \hat{b} also in Richtung des $(m+1)$ -ten Einheitsvektors dreht. Damit gilt für $f_3(x) := O_3 x$

$$E_{f_3 \circ f_2 \circ f_1}^t S_p E_{f_3 \circ f_2 \circ f_1} = \begin{pmatrix} c & \mu \cdot e_{m+1}^t \\ \mu \cdot e_{m+1} & D \end{pmatrix},$$

die Diagonalmatrix D bleibt also unverändert.

4. Ist $v_4 := -\frac{c}{2\mu} e_{m+1}$, dann ist für $f_4(x) := x + v_4$ und $f = f_4 \circ f_3 \circ f_2 \circ f_1$

$$S_p' = E_f^t S_p E_f = \begin{pmatrix} 0 & \mu e_{m+1}^t \\ \mu e_{m+1} & D \end{pmatrix},$$

also $p \circ f(y) = \sum_{i=1}^m \lambda_i y_i^2 + 2\|b\|y_{m+1}$. □

4.7 Bemerkung (Hauptachsentransformation) Ist $b = 0$ und $c = 0$, dann ist p die quadratische Form $p(x) = (x, Ax)$, und der Satz besagt, dass es eine Drehung $x = f(y) = Oy$ des \mathbb{R}^n gibt, die p in Diagonalfom $p \circ f(y) = \sum_{i=1}^m \lambda_i y_i^2$ transformiert. $O \in O(n)$ heißt dann auch *Hauptachsentransformation*. ◇

Wir kehren zu unserer ursprünglichen Fragestellung der Normalformen von Quadriken unter *Affinitäten* zurück und zeigen:

4.8 Satz Ist $p \in \mathbb{R}[x_1, \dots, x_n]$ ein quadratisches Polynom der Form $p(x) = (x, Ax) + (b, x) + c$ mit $A = A^t$ und $m := \text{rg}(A)$, dann existiert eine orientierungserhaltende Affinität g des \mathbb{R}^n , für die die Quadrik $(p \circ g)^{-1}(0) \subset \mathbb{R}^n$ durch die Gleichung

$$\begin{cases} z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_m^2 = 0 & , \quad \text{rg}(E_p) = m \\ z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_m^2 = 1 & , \quad \text{rg}(E_p) = m + 1 \\ z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_m^2 = 2z_{m+1} & , \quad \text{rg}(E_p) = m + 2 \end{cases}$$

beschrieben wird. Dabei ist k die Zahl der positiven Eigenwerte von A .

4.9 Bemerkung (Signatur) Das Zahlentripel $(k, m-k, n-m)$ heißt *Signatur* einer quadratischen Form p und ist gleich der Zahl der positiven, negativen bzw. Null-Eigenwerte der symmetrischen Matrix A , jeweils mit ihrer Multiplizität. Der *Trägheitssatz von Sylvester* besagt, dass diese Zahlen gleich den Dimensionen der Unterräume in einer beliebigen Zerlegung

$$\mathbb{R}^n = V_+ \oplus V_- \oplus V_0$$

ist, bei der die Restriktion von p auf V_0 verschwindet, auf V_+ positiv definit und auf V_- negativ definit ist. Der *Rang* der quadratischen Form p ist $m = \text{rg}(A)$.

Beweis:

- Ist $g_1(z) = Mz$ mit $M := \text{diag}(|\lambda_1|^{-1/2}, \dots, |\lambda_m|^{-1/2}, 1, \dots, 1)$, dann ist unter Verwendung der orientierungserhaltenden Kongruenz f aus Satz 4.6

$$p \circ f \circ g_1(z) = \sum_{i=1}^k z_i^2 - \sum_{i=k+1}^m z_i^2, \text{ falls } \text{rg}(E_p) = m.$$

- Ist dagegen $\text{rg}(E_p) = m + 1$, dann setzen wir $g_1(z) = Mz$ mit

$$M := \text{diag} \left(\left| \frac{c}{\lambda_1} \right|^{1/2}, \dots, \left| \frac{c}{\lambda_m} \right|^{-1/2}, 1, \dots, 1 \right),$$

sodass $p \circ f \circ g_1(z) = |c| \left(\sum_{i=1}^k z_i^2 - \sum_{i=k+1}^m z_i^2 + \text{sign}(c) \right)$ ist.

Ist $\text{sign}(c) = +1$, dann permutieren wir noch die z_1, \dots, z_k mit den z_{k+1}, \dots, z_m . Dies ist eine orthogonale Transformation, und wir können wieder die Positivität der Determinante durch Multiplikation einer Koordinate mit -1 erzwingen.

Die Quadrik von $p \circ f \circ g_1$ hängt nicht vom Vorfaktor $|c| > 0$ ab.

- Im letzten Fall $\text{rg}(E_p) = m + 2$ wird durch Übergang zur Variablen $z_{m+1} := \|b\|y_{m+1} - \frac{c}{2}$ reskaliert und die Konstante c gleichzeitig durch Null ersetzt. \square

Durch Verwendung allgemeiner Affinitäten können wir Längen ändern, und beispielsweise jede Ellipse in den Einheitskreis überführen. Daher haben wir für jede Dimension n nur *endlich viele* affine Normalformen.

4.10 Beispiele Normalformen von Quadriken]

1. Für $n = 1$ kommen die Fälle $Q = \{0\}$, $Q = \{-1, 1\}$, $Q = \mathbb{R}$ und $Q = \emptyset$ vor, entsprechend den Kardinalitäten der Lösungsmengen quadratischer Gleichungen.
2. Für $n = 2$ können wir zunächst die Quadriken der eindimensionalen Normalformen mit \mathbb{R} multiplizieren, und erhalten die *Gerade* $Q = \mathbb{R} \times \{0\}$, die *Doppelgerade* $Q = \mathbb{R} \times \{-1, 1\}$, die *Ebene* $Q = \mathbb{R}^2$ und die *leere Menge*.

Die Gleichung $z_1^2 + z_2^2 = 0$ hat als Lösungsmenge den *Nullpunkt*, während für $z_1^2 - z_2^2 = 0$ die beiden Winkelhalbierenden die Quadrik $Q = \text{span} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cup \text{span} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ergeben.

Diese sind die *Asymptoten* der *Hyperbel* $Q = \{(z_1, z_2) \in \mathbb{R}^2 \mid z_1^2 - z_2^2 = 1\}$.

Die Quadrik zur Gleichung $z_1^2 + z_2^2 = 1$ ist natürlich der *Kreis* mit Radius 1 um den Nullpunkt.

Zuletzt beschreibt die Gleichung $z_1^2 = 2z_2$ eine *Parabel*.

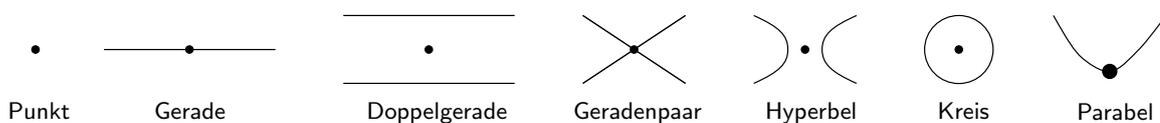
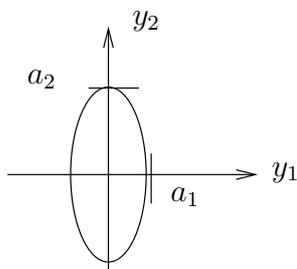


Abbildung 4.1: Affine Normalformen der Quadriken für $n = 2$ (außer $Q = \mathbb{R}^2$, $Q = \emptyset$) mit eingezeichnetem Ursprung des Koordinatensystems

Die Äquivalenzklasseneinteilung bezüglich orientierungserhaltender Kongruenzen ist feiner als die bezüglich Affinitäten. Beispielsweise erhalten wir statt des Kreises die Kongruenznormalformen $\frac{y_1^2}{a_1^2} + \frac{y_2^2}{a_2^2} = 1$ mit $0 < a_1 \leq a_2$ der *Ellipsen*, deren Halbachsen mit den Längen a_i in Richtung der y_i -Achsen gedreht sind:



Wir sehen, dass wir sowohl im affinen als auch im Kongruenzfall Normalformen von gewissen Quadriken der Dimension $n + 1$ erhalten, indem wir im \mathbb{R}^{n+1} (statt im \mathbb{R}^n) die Nullstellenmenge der quadratischen Polynome von Normalformen der Dimension n betrachten. Man spricht dann von *Zylindern*:

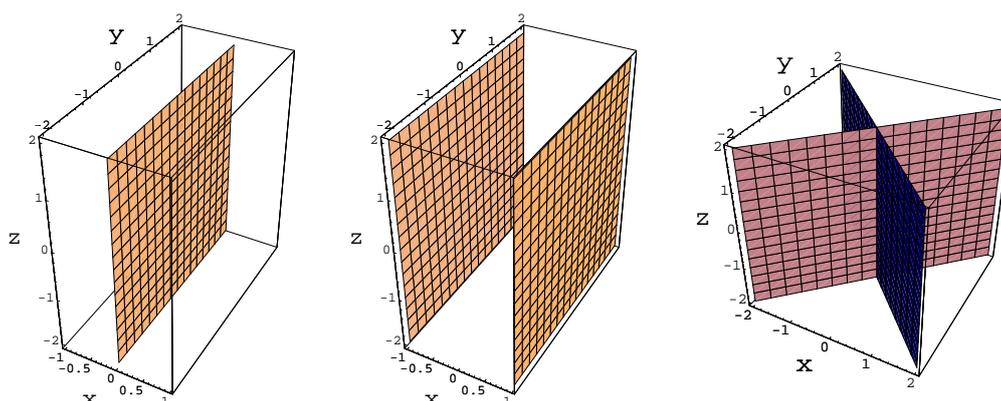


Abbildung 4.2: Affine Normalformen von Quadriken für $n = 3$: x (Ebene), $x^2 - 1$ (Doppel-Ebene), $x^2 - y^2$ (Ebenenpaar)

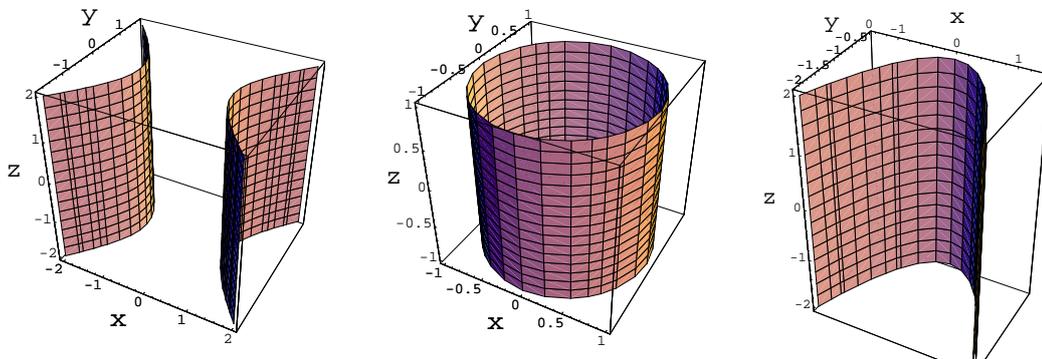


Abbildung 4.3: Affine Normalformen von Quadriken für $n = 3$: $x^2 - y^2 = 1$ (Zylinder über Hyperbel), $x^2 + y^2 = 1$ (Zylinder über Kreis), $x^2 = 2y$ (Zylinder über Parabel)

Daneben finden wir in jeder Dimension jedoch auch Quadriken, die nicht Zylinder sind:

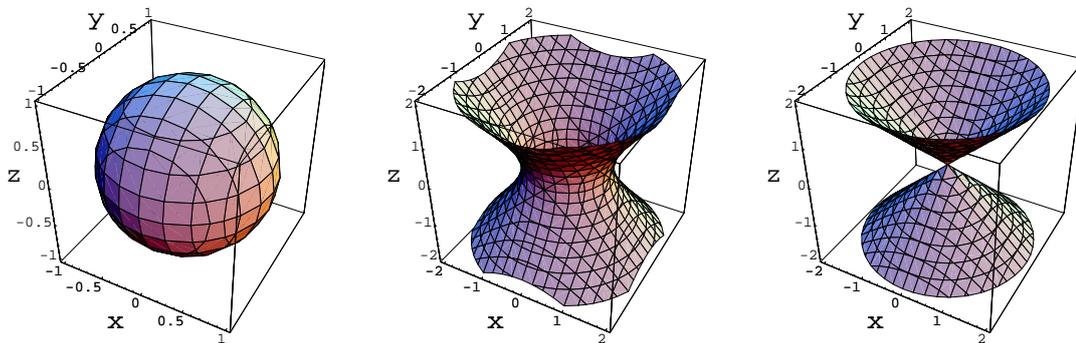


Abbildung 4.4: Affine Normalformen von Quadriken für $n = 3$: $x^2 + y^2 + z^2 = 1$ (Sphäre), $x^2 + y^2 - z^2 = 1$ (einschaliges Hyperboloid), $x^2 + y^2 = z^2$ (Kegel)

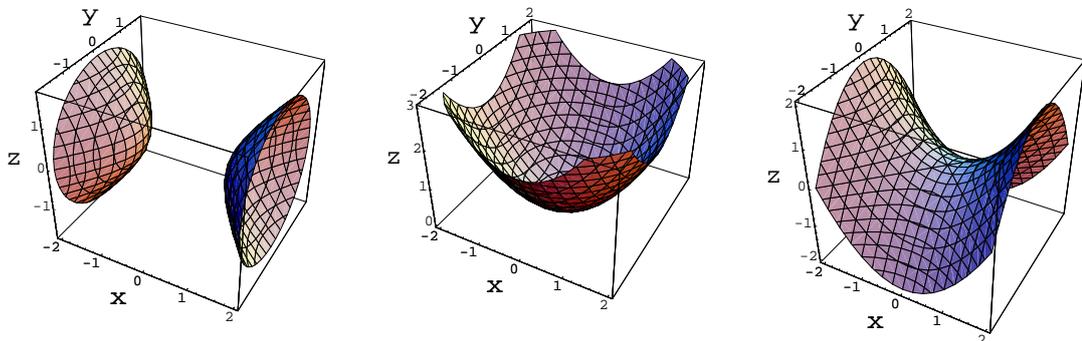


Abbildung 4.5: Affine Normalformen von Quadriken für $n = 3$: $x^2 - y^2 - z^2 = 1$ (zweischaliges Hyperboloid), $x^2 + y^2 - 2z$ (Paraboloid), $x^2 - y^2 - 2z$ (Sattel)

5 Metrik, Topologie und Stetigkeit

Die *Topologie* ist das Teilgebiet der Mathematik, das sich mit Begriffen wie Offenheit, Abgeschlossenheit und Kompaktheit von Mengen, Konvergenz von Folgen und Stetigkeit von Funktionen befasst.

Diese Begriffe kann man z.B. definieren, wenn man Abstände von Punkten der Menge messen kann, d.h. eine Metrik vorliegt.

5.1 Metrische Räume

Der Begriff der *Metrik* stellt eine Abstraktion des Abstands begriffes dar:

5.1 Definition Eine Menge M mit einer Funktion $d : M \times M \rightarrow \mathbb{R}$ heißt **metrischer Raum mit Metrik** d , falls für alle $x, y, z \in M$ gilt:

- $d(x, y) \geq 0$, $d(x, y) = 0 \iff x = y$ (Positivität)
- $d(y, x) = d(x, y)$ (Symmetrie)
- $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung).

Man nennt $d(x, y)$ den **Abstand** von x und y .

5.2 Beispiele (Metrische Räume)

1. Die Menge $B := \{0, 1\}$ bezeichnet ein *Bit*. Wir betrachten Folgen von $n \in \mathbb{N}$ Bits, d.h. Elemente von B^n . Ihr sog. *Hamming-Abstand* ist durch die Metrik $d : B^n \times B^n \rightarrow \{0, 1, \dots, n\}$,

$$d((b_1, \dots, b_n), (c_1, \dots, c_n)) := |\{i \in \{1, \dots, n\} \mid b_i \neq c_i\}|$$

gegeben¹², also durch die Zahl der Stellen, an denen sich die beiden Bitfolgen unterscheiden. Diese Metrik wird in der Informationstheorie benutzt.

2. Für uns besonders wichtig ist die *euklidische* Metrik auf \mathbb{C}

$$d : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{R}, d(x, y) := |x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \quad (5.1)$$

die den Abstand der Punkte x und y in der komplexen Zahlenebene misst. Dass d tatsächlich die Dreiecksungleichung erfüllt, folgt aus

$$d(x, z) = |x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| = d(x, y) + d(y, z).$$

3. Ist (M, d) ein metrischer Raum und N eine Teilmenge von M , dann ist (N, d_N) mit der "auf N restringierten" Metrik¹³

$$d_N : N \times N \rightarrow \mathbb{R}, d_N(x, y) := d(x, y)$$

ebenfalls ein metrischer Raum.

Betrachten wir etwa $M := \mathbb{C}$ mit der Metrik (5.1), dann ist

$$S^1 := \{c \in \mathbb{C} \mid |c| = 1\}$$

geometrisch die *Kreislinie*¹⁴ vom Radius 1 um den Nullpunkt. Der Abstand $d_{S^1}(x, y)$ zwischen zwei Kreispunkten ist dann gleich der Länge der Sekante mit den Endpunkten x und y .

Eine andere sinnvolle Metrik auf S^1 ist durch die (minimale) Winkeldifferenz von x und y gegeben.

4. Auf dem Vektorraum \mathbb{R}^n wird die Länge eines Vektors durch die *euklidische Norm*

$$\|v\| := \sqrt{\sum_{i=1}^n v_i^2} \quad \left(v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n \right)$$

von v definiert. Mit $d(x, y) := \|y - x\|$ ergibt sich daraus eine Metrik auf dem \mathbb{R}^n , genannt die *euklidische Metrik*. Denn es gilt die Dreiecksungleichung $\|v + w\| \leq \|v\| + \|w\|$. Speziell für $n = 1$ ist $d(x, y) = |x - y|$.

¹²mit Schreibweise als Zeilenvektoren

¹³Genauer gesagt wird die Abbildung $d : M \times M \rightarrow \mathbb{R}$ auf $N \times N$ restringiert.

¹⁴Man spricht nicht einfach vom *Kreis*, weil man eine Verwechslung mit der (abgeschlossenen) *Kreisscheibe* $\{c \in \mathbb{C} \mid |c| \leq 1\}$ ausschließen will. Der Index 1 bezieht sich auf die Dimension der Kreislinie. Analog bezeichnet die *d-Sphäre* $S^d := \{x \in \mathbb{R}^{d+1} \mid \sum_{i=1}^{d+1} x_i^2 = 1\}$ die Oberfläche der *Einheitsvollkugel* $\{x \in \mathbb{R}^{d+1} \mid \sum_{i=1}^{d+1} x_i^2 \leq 1\}$ im \mathbb{R}^{d+1} .

5. Auf dem Vektorraum \mathbb{R}^n auch durch die *Maximumsnorm*

$$\|v\|_\infty := \max(|v_1|, \dots, |v_n|) \quad \left(v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n \right)$$

eine Metrik definiert:

$$d_\infty(x, y) := \|x - y\|_\infty, \text{ und es gilt}$$

$$d_\infty(x, y) \leq d(x, y) \leq \sqrt{n}d_\infty(x, y) \quad (x, y \in \mathbb{R}^n),$$

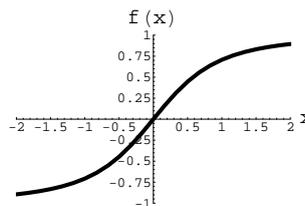
siehe nebenstehende Abbildung für $n = 2$
und $x = 0$.

Geometrisch ist die "Einheitskugel"

$\{v \in \mathbb{R}^n \mid \|v\|_\infty \leq 1\}$ bezüglich der Maximumsnorm ein achsenparalleler um den Nullpunkt zentrierter n -dimensionaler Würfel der Kantenlänge 2.

6. Mit der Bijektion

$$f: \overline{\mathbb{R}} \rightarrow [-1, 1], f(x) := \begin{cases} 1 & , x = \infty \\ \frac{x}{\sqrt{x^2+1}} & , x \in \mathbb{R} \\ -1 & , x = -\infty \end{cases}$$



wird eine Metrik d auf $\overline{\mathbb{R}}$ durch

$$d(x, y) := |f(x) - f(y)| \text{ eingeführt. } \diamond$$

5.3 Definition

Es sei (M, d) ein metrischer Raum, $x \in M$ und $\varepsilon > 0$. Dann heißt

$$U_\varepsilon(x) := \{y \in M \mid d(y, x) < \varepsilon\}$$

(offene) ε -Umgebung von x in M .

Immer gilt $x \in U_\varepsilon(x)$, und im metrischen Raum \mathbb{R} mit euklidischer Metrik ist für $a < b$ das offene Intervall $(a, b) = U_\varepsilon(x)$ mit dem Mittelpunkt $x := \frac{1}{2}(a + b)$ und $\varepsilon := \frac{b-a}{2}$. Im \mathbb{R}^n mit euklidischer Metrik ist $U_\varepsilon(x)$ geometrisch eine n -dimensionale um x zentrierte Vollkugel vom Radius ε . Im Gegensatz zur abgeschlossenen Kugel $\{y \in \mathbb{R}^n \mid d(x, y) \leq \varepsilon\}$ wird sie *offene Kugel* genannt.

5.2 Folgen in metrischen Räumen

5.4 Definition

1. Jede Abbildung $a: \mathbb{N} \rightarrow M$ in eine Menge wird auch **Folge** in M genannt und in der Form $(a_n)_{n \in \mathbb{N}}$ oder (a_1, a_2, a_3, \dots) geschrieben. Für $M = \mathbb{C}$ wird die Folge **komplexwertig**, für $M = \mathbb{R}^n$ **vektorwertig** genannt.

2. Eine Folge $(a_n)_{n \in \mathbb{N}}$ in einem metrischen Raum (M, d) heißt **Cauchy-Folge**, wenn für alle $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ existiert mit

$$d(a_m, a_n) < \varepsilon \quad (m, n \geq N).$$

3. • Eine Folge $(a_n)_{n \in \mathbb{N}}$ in M **konvergiert gegen** $c \in M$, wenn für alle $\varepsilon > 0$ ein $N = N(\varepsilon)$ existiert mit

$$a_n \in U_\varepsilon(c) \quad (n \geq N).$$

- In diesem Fall nennt man c **Grenzwert** oder **Limes** von a und schreibt

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

- Falls es ein $c \in M$ gibt, gegen das die Folge $a : \mathbb{N} \rightarrow M$ konvergiert, heißt sie **konvergent**, sonst **divergent**.

Diese Definitionen erweitern die für die reellen Folgen $a : \mathbb{N} \rightarrow \mathbb{R}$ in der *Analysis I* eingeführten Begriffe. In diesem Fall ist (mit der Metrik $d(x, y) = |x - y|$) die Bedingung $|a_n - c| < \varepsilon$ in (2.7) nämlich gleichbedeutend mit $a_n \in U_\varepsilon(c)$.

5.5 Beispiele (Cauchy-Folgen)

1. Eine reelle Folge ist genau dann konvergent gegen $y \in \mathbb{R}$, wenn sie bezüglich der Metrik d aus Beispiel 5.2.6 gegen y konvergiert. Uneigentliche Konvergenz gegen $\pm\infty$ entspricht aber einfach der Konvergenz gegen $\pm\infty$ bezüglich d .
2. Unter Benutzung der Metriken aus Beispiel 5.2.2 und 5.2.4 können wir jetzt auch Cauchy-Folgen in \mathbb{C} und im \mathbb{R}^n behandeln. \diamond

5.6 Satz Alle konvergenten Folgen sind Cauchy-Folgen.

Beweis: Sei $a = \lim_{n \rightarrow \infty} x_n$. Dann ist $d(x_m, x_n) \leq d(x_m, a) + d(x_n, a)$. \square

Aber nicht in jedem metrischem Raum konvergieren alle Cauchy-Folgen. Daher machen die Mathematiker ihren Wunsch zu einer Definition:

5.7 Definition

- Ein metrischer Raum heißt **vollständig**, wenn in ihm jede Cauchy-Folge konvergiert.
- Ein vollständiger normierter Vektorraum heißt **Banach-Raum**.

Uns ist schon die Vollständigkeit von \mathbb{R} mit der Standardnorm bekannt.

5.8 Satz Die metrischen Räume \mathbb{C} und \mathbb{R}^n ($n \in \mathbb{N}$) mit den euklidischen Normen sind vollständig, also Banachräume.

Beweis:

- Als metrische Räume stimmen (\mathbb{C}, d) und (\mathbb{R}^2, d) mit den Metriken aus Beispiel 5.2.2 und 5.2.4 überein. Daher ist nur der Fall (\mathbb{R}^d, d) zu untersuchen.

- Ist nun $(x_m)_{m \in \mathbb{N}}$ mit $x_m = (x_m^{(1)}, \dots, x_m^{(n)})^T \in \mathbb{R}^n$ eine Cauchy-Folge, existiert also für $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\|x_m - x_\ell\| < \varepsilon \quad (m, \ell \geq N),$$

dann sind die reellen Folgen $(x_m^{(k)})_{m \in \mathbb{N}}$ der k -ten Komponenten ebenfalls Cauchy-Folgen, und es gilt

$$|x_m^{(k)} - x_\ell^{(k)}| < \varepsilon \quad (m, \ell \geq N, k = 1, \dots, n)$$

(dies folgt aus der Ungleichung $|x_m^{(k)} - x_\ell^{(k)}| \leq \|x_m - x_\ell\|$). Letztere Cauchy-Folgen konvergieren aber in \mathbb{R} .

Die Konvergenz der Cauchy-Folge $(x_m)_{m \in \mathbb{N}}$ folgt damit aus dem nächsten Satz. \square

5.9 Satz Eine Folge $(x_m)_{m \in \mathbb{N}}$ mit $x_m = (x_m^{(1)}, \dots, x_m^{(n)})^T \in \mathbb{R}^n$ konvergiert genau dann, wenn die reellen Folgen $(x_m^{(k)})_{m \in \mathbb{N}}$ konvergieren. Es ist dann

$$\left(\lim_{m \rightarrow \infty} x_m \right)^{(k)} = \lim_{m \rightarrow \infty} x_m^{(k)} \quad (k = 1, \dots, n).$$

Beweis:

- Falls $y = (y^{(1)}, \dots, y^{(n)}) := \lim_{m \rightarrow \infty} x_m \in \mathbb{R}^n$ existiert, gibt es für $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\|x_m - y\| < \varepsilon \quad (m \geq N).$$

Dann ist aber auch $|x_m^{(k)} - y^{(k)}| \leq \|x_m - y\| < \varepsilon$.

- Konvergieren dagegen die Koordinatenfolgen, d.h. existieren $y^{(k)} \in \mathbb{R}$ mit

$$y^{(k)} = \lim_{m \rightarrow \infty} x_m^{(k)} \quad (k = 1, \dots, n),$$

dann gibt es auch für jedes $\varepsilon > 0$ ein N , sodass

$$|x_m^{(k)} - y^{(k)}| < \frac{\varepsilon}{\sqrt{n}} \quad (m \geq N, k = 1, \dots, n).$$

Dann ist aber mit $y := (y^{(1)}, \dots, y^{(n)})^T \in \mathbb{R}^n$ auch

$$\|x_m - y\| = \sqrt{\sum_{k=1}^n |x_m^{(k)} - y^{(k)}|^2} < \varepsilon \quad (m \geq N),$$

was $y = \lim_{m \rightarrow \infty} x_m$ impliziert. \square

Wir können also Konvergenz in \mathbb{C} oder im \mathbb{R}^n koordinatenweise überprüfen.

In Anwendungen auf zusammengesetzte Folgen ermöglicht der folgende Satz eine vereinfachte Berechnung von Limiten.

5.10 Satz

1. Für vektorwertige Cauchy-Folgen $a, b : \mathbb{N} \rightarrow \mathbb{R}^n$ gilt

$$\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n.$$

2. Für komplexwertige Cauchy-Folgen $a, b : \mathbb{N} \rightarrow \mathbb{C}$ gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} (a_n + b_n) &= \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n, \\ \lim_{n \rightarrow \infty} (a_n b_n) &= \left(\lim_{n \rightarrow \infty} a_n \right) \left(\lim_{n \rightarrow \infty} b_n \right) \end{aligned}$$

und, falls $b_n \neq 0$ ($n \in \mathbb{N}$) und $\lim_{n \rightarrow \infty} b_n \neq 0$, auch

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}.$$

Beweis: Dies folgt unter Verwendung von Satz 5.9 aus den analogen Aussagen für reelle Folgen. \square

5.11 Definition $y \in M$ heißt **Häufungspunkt** der Folge $(a_n)_{n \in \mathbb{N}}$ im metrischen Raum (M, d) , wenn für alle $\varepsilon > 0$ gilt

$$|\{n \in \mathbb{N} \mid a_n \in U_\varepsilon(y)\}| = \infty.$$

5.12 Beispiele (Häufungsmengen)

1. Ist $y = \lim_{n \rightarrow \infty} a_n$, dann ist y Häufungspunkt der Folge. Es kann dann keine weiteren Häufungspunkte der Folge geben, da diese für jedes $\varepsilon > 0$ schließlich (d.h. für alle $n \geq N(\varepsilon)$) in $U_\varepsilon(y)$ bleibt.
2. Die reelle Folge mit Folgenglied $a_n := (-1)^n$ besitzt genau die Häufungspunkte -1 und 1 .
3. Wir wissen, dass $|\mathbb{Q}| = |\mathbb{N}|$ ist, also eine Bijektion $a : \mathbb{N} \rightarrow \mathbb{Q}$ existiert. Die Menge der Häufungspunkte der reellen Folge a ist ganz \mathbb{R} .
4. Die komplexe geometrische Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n := c^n$ und $c \in \mathbb{C}$ besitzt für $|c| < 1$ den Limes 0 und für $|c| > 1$ keinen Häufungspunkt. Für $c \in S^1$ befinden sich auch alle Häufungspunkte auf der Kreislinie S^1 . \diamond

Durch Auswahl von Folgengliedern kann man aus einer Folge $(a_n)_{n \in \mathbb{N}}$ neue Folgen konstruieren.

5.13 Definition Es sei $a : \mathbb{N} \rightarrow M$ eine Folge und $k : \mathbb{N} \rightarrow \mathbb{N}$ streng monoton wachsend. Dann wird $a \circ k : \mathbb{N} \rightarrow M$ **Teilfolge von a** genannt.

5.14 Satz Es sei $a : \mathbb{N} \rightarrow M$ eine Folge im metrischen Raum (M, d) .

1. Ist $y = \lim_{n \rightarrow \infty} a_n$, dann besitzt auch jede Teilfolge von a den Limes y .
2. Ist y Häufungspunkt der Folge a , dann gibt es eine Teilfolge von a , die gegen y konvergiert.
3. Gibt es eine Teilfolge von a , die gegen $y \in M$ konvergiert, dann ist y Häufungspunkt von a .

Beweis:

1. Zu zeigen ist: Für alle $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit

$$a_{k(n)} \in U_\varepsilon(y) \quad (n \geq N). \quad (5.2)$$

Nach Annahme existiert ein N mit $a_n \in U_\varepsilon(y)$ ($n \geq N$). Da die Indizierung $k : \mathbb{N} \rightarrow \mathbb{N}$ streng monoton ist, gilt $k(n) \geq n$, also (5.2).

2. Ist y Häufungspunkt der Folge a_n , dann gibt es eine streng monoton wachsende Funktion $k : \mathbb{N} \rightarrow \mathbb{N}$ mit $b_n \in U_{1/n}(y)$ für $b_n := a_{k(n)}$. Denn setzen wir $k(1) := \min\{m \in \mathbb{N} \mid a_m \in U_1(y)\}$ und induktiv für $\ell \geq 2$

$$k(\ell) := \min\{m > k(\ell - 1) \mid a_m \in U_{1/\ell}(y)\},$$

dann sind die betrachteten Mengen nie leer, die Funktion k ist streng monoton und $\lim_{n \rightarrow \infty} b_n = y$.

3. Da die Abbildung $k : \mathbb{N} \rightarrow \mathbb{N}$ streng monoton wachsend ist, gilt $|k(\mathbb{N})| = |\mathbb{N}|$. Wenn – wie angenommen – für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, sodass

$$|a_{k(n)} - y| < \varepsilon \quad (n \geq N),$$

dann ist auch die Indexmenge $\{m \in k(\mathbb{N}) \mid m \geq k(N)\}$ unendlich, und für diese Indices m ist $a_m \in U_\varepsilon(y)$. \square

5.3 Topologische Räume

Wir erinnern uns an die Bedeutung der offenen bzw. abgeschlossenen Intervalle $I \subset \mathbb{R}$ in der *Analysis I*.

- Z.B. können wir für Funktionen $f : I \rightarrow \mathbb{R}$ den Differentialquotient definieren, denn jeder Punkt $x \in I$ ist Häufungspunkt der Menge I . Endliche Schnitte offener Intervalle sind offen, beliebige Schnitte i.A. nicht.

- Dagegen bieten *abgeschlossene* Intervalle $I \subset \mathbb{R}$ den Vorteil, dass jede konvergente Folge $a : \mathbb{N} \rightarrow I$ ihren Grenzwert in I annimmt. Beliebige Schnitte abgeschlossener Intervalle sind abgeschlossen.

Diese Begriffe werden jetzt verallgemeinert (siehe auch mein [Topologieskript](#)).

5.15 Definition Ein **topologischer Raum** ist ein Paar (M, \mathcal{O}) , bestehend aus einer Menge M und einer Familie \mathcal{O} von Teilmengen von M (genannt "offene Mengen"), sodass gilt:

1. Beliebige Vereinigungen offener Mengen sind offen.
2. Der Durchschnitt von je zwei offenen Mengen ist offen.
3. \emptyset und M sind offen.

Man nennt \mathcal{O} auch die **Topologie** von (M, \mathcal{O}) .

Die für uns wichtigsten Beispiele topologischer Räume sind die metrischen Räume:

5.16 Definition Für einen metrischen Raum (M, d) heißt

$$\mathcal{O} := \mathcal{O}(d) := \{V \subseteq M \mid \forall x \in V \exists \varepsilon > 0 : U_\varepsilon(x) \subseteq V\} \quad (5.3)$$

(mit $U_\varepsilon(x) := \{y \in M \mid d(y, x) < \varepsilon\}$) die (metrische) **Topologie von** (M, d) .

5.17 Satz $(M, \mathcal{O}(d))$ ist ein topologischer Raum, und die $U_\varepsilon(x)$ sind offen.

Beweis: • $(M, \mathcal{O}(d))$ ist ein topologischer Raum:

1. Ist $x \in \bigcup_{i \in I} V_i$ mit $V_i \in \mathcal{O}$, dann gibt es ein $j \in I$ mit $x \in V_j$, und ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq V_j \subseteq \bigcup_{i \in I} V_i$.
2. Sei $x \in V_1 \cap V_2$ mit $V_1, V_2 \in \mathcal{O}$, dann gibt es $\varepsilon_1, \varepsilon_2 > 0$ mit $U_{\varepsilon_j}(x) \subseteq V_j$, $j = 1, 2$. Daher ist für $\varepsilon := \min(\varepsilon_1, \varepsilon_2) > 0$ $U_\varepsilon(x) \subseteq V_1 \cap V_2$.
3. $\emptyset \in \mathcal{O}$, denn dann ist die Bedingung in (5.3) leer. Ebenso gilt $M \in \mathcal{O}$, denn dann ist die Bedingung in (5.3) für alle $\varepsilon > 0$ erfüllt.

• Die $U_\varepsilon(x)$ sind offen: Für $y \in U_\varepsilon(x)$ ist $\varepsilon' := d(x, y) < \varepsilon$. Wählen wir $\delta := \varepsilon - \varepsilon' > 0$, dann ist wegen der Dreiecksungleichung $U_\delta(y) \subseteq U_\varepsilon(x)$. \square

Oft erzeugen verschiedene Metriken die gleiche Topologie. Insbesondere gilt dies für Metriken auf Vektorräumen, die von äquivalenten Normen abstammen.

Auf jeder Menge M lassen sich Topologien finden:

5.18 Beispiele (In/diskrete Topologie)

1. Die *diskrete Topologie* $\mathcal{O} := 2^M$ aller Teilmengen.

2. Die *indiskrete Topologie* $\mathcal{O} := \{M, \emptyset\}$. ◇

Beispielsweise erzeugt die in der Informatik beliebte Hamming-Metrik auf $M := B^n$, $B := \{0, 1\}$

$$d : B^n \times B^n \rightarrow \mathbb{N}_0, d((b_1, \dots, b_n), (c_1, \dots, c_n)) := \left| \left\{ i \in \{1, \dots, n\} \mid b_i \neq c_i \right\} \right|$$

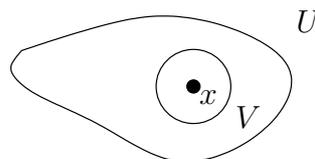
gemäß Def. 5.16 die diskrete Topologie.

Die indiskrete Topologie taucht fast nur in Gegenbeispielen auf.

Wir erweitern unseren topologischen Sprachschatz, indem wir die uns vom Raum \mathbb{R} bekannten Begriffsbildungen verallgemeinern:

5.19 Definition • $A \subseteq M$ heißt **abgeschlossen**, wenn $M \setminus A \in \mathcal{O}$.

• $U \subseteq M$ heißt **Umgebung** von $x \in M$, wenn es eine offene Menge V mit $x \in V \subseteq U$ gibt.



• Für $A \subseteq M$ und $x \in M$ heißt x **innerer** bzw. **äußerer** bzw. **Randpunkt** von A , je nachdem, ob A oder $M \setminus A$ oder keines von beiden Umgebung von x ist.

• $\overset{\circ}{A} := \{x \in M \mid x \text{ ist innerer Punkt von } A\}$ heißt das **Innere** von A .

• $\bar{A} := \{x \in M \mid x \text{ nicht äußerer Punkt von } A\}$ heißt **abgeschlossene Hülle** von A .

• $\partial A := \{x \in M \mid x \text{ Randpunkt von } A\}$ heißt **Rand** von A .

• $x \in M$ heißt **Häufungspunkt** der Teilmenge $A \subseteq M$, wenn für keine Umgebung U von x die Menge $U \cap (A \setminus \{x\})$ leer ist.

5.20 Beispiel

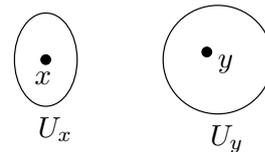
Für $A := (0, 1] \subset \mathbb{R}$ ist $\overset{\circ}{A} = (0, 1)$, $\bar{A} = [0, 1]$ und $\partial A = \{0, 1\}$. \mathbb{Z} besitzt keine Häufungspunkte. ◇

5.21 Bemerkungen (Inneres, Abschluss und Rand)

1. $\overset{\circ}{A} \subseteq A \subseteq \bar{A} = \overset{\circ}{A} \dot{\cup} \partial A$.

2. Eine Menge ist genau dann offen, wenn sie mit ihrem Inneren übereinstimmt.
3. Für alle $A \subseteq M$ ist $\overset{\circ}{A}$ offen und \bar{A} sowie ∂A abgeschlossen.
4. Die Menge der Häufungspunkte von A ist eine Teilmenge von \bar{A} . ◇

5.22 Definition Ein topologischer Raum heißt **Hausdorff-Raum**, wenn je zwei verschiedene Punkte x, y disjunkte Umgebungen U_x, U_y besitzen.



5.23 Bemerkungen 1. Metrische Räume sind hausdorffsch, denn für $x \neq y$ ist $\varepsilon := d(x, y)/2 > 0$ und $U_\varepsilon(x) \cap U_\varepsilon(y) = \emptyset$.

Daher werden die meisten uns begegnenden Räume die Hausdorff-Eigenschaft besitzen.

2. Ein mehr als einpunktiger Raum mit indiskreter Topologie ist nicht hausdorffsch. ◇

5.24 Definition In einem topologischen Raum (X, \mathcal{O}) heißt

- $x \in M$ **Häufungspunkt** einer Folge $a : \mathbb{N} \rightarrow M$, wenn für jede Umgebung U von x und jedes $N \in \mathbb{N}$ ein Folgenglied $a_n \equiv a(n) \in U$ mit $n \geq N$ existiert.
- $x \in M$ heißt **Grenzwert** oder **Limes** einer Folge $a : \mathbb{N} \rightarrow M$, wenn für jede Umgebung U von x ein $N \in \mathbb{N}$ mit $a_n \in U$ für alle $n \geq N$ existiert.

5.25 Bemerkungen (Häufungspunkte)

1. Man muss also zwischen den Häufungspunkt einer Folge und den Häufungspunkten der Menge $a(\mathbb{N}) \subseteq M$ unterscheiden. So besitzt die konstante Folge mit Wert $a_n = x$ den Häufungspunkt x , nicht aber die Menge $\{x\}$.
2. In einem Hausdorff-Raum besitzt jede Folge höchstens einen Grenzwert.

Oft beschreibt man eine Topologie \mathcal{O} , indem man eine so genannte Basis von \mathcal{O} angibt.

5.26 Definition Sei (M, \mathcal{O}) ein topologischer Raum. Eine Teilmenge $B \subseteq \mathcal{O}$ offener Mengen heißt **Basis der Topologie**, wenn jedes $V \in \mathcal{O}$ Vereinigung von Mengen aus B ist.

5.27 Beispiel (Basen von Topologien)

1. Die Menge $B := \{(a, b) \mid a < b\}$ der offenen beschränkten Intervalle bildet eine Basis der (metrischen) Topologie von \mathbb{R} .
2. Allgemeiner bilden nach Definition 5.16 die ε -Umgebungen eine Basis $B := \{U_\varepsilon(x) \mid x \in M, \varepsilon > 0\}$ der Topologie $\mathcal{O}(d)$ des metrischen Raums (M, d) . \diamond

Aus topologischen Räumen kann man auf verschiedene Weise neue topologische Räume konstruieren, z.B. durch Produktbildung bzw. Betrachtung von Teilmengen.

5.28 Definition

$(X^{(i)}, \mathcal{O}^{(i)})$, $i = 1, \dots, n$ seien topologische Räume. Dann ist auf ihrem kartesischen Produkt $X := X^{(1)} \times \dots \times X^{(n)}$ die **Produkttopologie** \mathcal{O} gegeben durch

$$\mathcal{O} := \left\{ W \subseteq X \mid \forall (x^{(1)}, \dots, x^{(n)}) \in W \exists U^{(i)} \in \mathcal{O}^{(i)} \text{ mit } x^{(i)} \in U^{(i)} \right. \\ \left. \text{und } U^{(1)} \times \dots \times U^{(n)} \subseteq W \right\}.$$

5.29 Satz Die Produkttopologie ist eine Topologie auf X .

Beweis: Hausaufgabe. □

Die in X offenen Mengen W müssen also "offene Kästchen" um jeden ihrer Punkte enthalten.

5.30 Bemerkungen (Produkttopologie)

1. Die Bildung der Produkttopologie ist eine assoziative Operation, es genügt also, je zwei Faktoren zu betrachten. Produkte von Basen bilden eine Basis der Produkttopologie.

Beispielsweise bilden die offenen Quader

$$(a_1, b_1) \times \dots \times (a_n, b_n) \subset \mathbb{R}^n$$

eine Basis der Produkttopologie des \mathbb{R}^n .

2. Für metrische Räume (X, d_X) und (Y, d_Y) sind auf dem kartesischen Produkt $Z := X \times Y$

$$d_i : Z \times Z \rightarrow \mathbb{R} \quad \text{für } z_j = (x_j, y_j) \\ d_1(z_1, z_2) := d_X(x_1, x_2) + d_Y(y_1, y_2), \\ d_2(z_1, z_2) := \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}$$

und

$$d_\infty(z_1, z_2) := \max(d_X(x_1, x_2), d_Y(y_1, y_2))$$

Metriken. Sie erzeugen aber alle die gleiche Topologie $\mathcal{O}(d_i)$, die Produkttopologie. \diamond

5.31 Definition Sei (X, \mathcal{O}) ein topologischer Raum und $Y \subseteq X$. Dann heißt (Y, \mathcal{O}_Y) mit $\mathcal{O}_Y := \{U \cap Y \mid U \in \mathcal{O}\}$ **Teilraum von (X, \mathcal{O}) und \mathcal{O}_Y die induzierte Topologie oder Spurtopologie.**

Im Allgemeinen enthält \mathcal{O}_Y also mehr Mengen als die schon in X offenen Mengen $U \subseteq Y, U \in \mathcal{O}$.

5.32 Beispiel $X := \mathbb{R}$ mit der von den offenen Intervallen erzeugten Topologie \mathcal{O} , $Y := (0, 1]$. Dann ist $(\frac{1}{2}, 1] \subseteq Y$ offen und entsprechend $(0, \frac{1}{2}] \subseteq Y$ abgeschlossen in Y (!). \diamond

Analog definiert man auf Teilmengen $Y \subseteq X$ metrischer Räume (X, d) eine $\mathcal{O}(d)_Y = \mathcal{O}(d_Y)$ Metrik d_Y durch Restriktion, und $\mathcal{O}(d)_Y = \mathcal{O}(d_Y)$.

Oft sind solche Teilräume *Niveaumengen* von Funktionen $F : \mathbb{R}^n \rightarrow \mathbb{R}$.

5.33 Beispiele (Niveaumengen)

1. Die n -Sphäre $S^n := \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$.

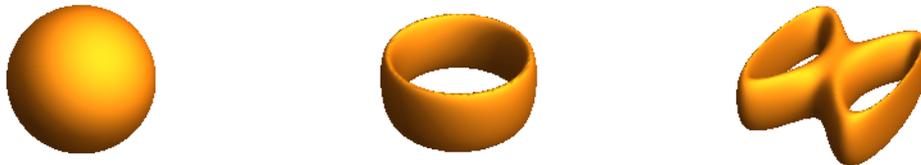


Abbildung 5.1: 2-Sphäre S^2 (links) , 2-Torus \mathbb{T}^2 (Mitte) und Fläche mit 2 Henkeln (rechts)

2. Auch der 2-Torus $\mathbb{T}^2 \subseteq \mathbb{R}^3$, eine Fläche mit einem Henkel, lässt sich als Niveaumenge darstellen:

$$\mathbb{T}^2 = \{x \in \mathbb{R}^3 \mid F(x) = 1\} \text{ mit } F(x) = (\sqrt{x_1^2 + x_2^2} - 2)^2 + x_3^2.$$

3. Die Fläche mit 2 Henkeln in Abbildung 5.1 (rechts) ist die Niveaumenge

$$\{x \in \mathbb{R}^3 \mid F(x) = 1/4\} \text{ von } F(x) = (4x_1^2(1 - x_1^2) - x_2^2)^2 + x_3^2. \quad \diamond$$

Etwas vornehmer heißt die Henkelzahl der Flächen ihr *Geschlecht*. Sie erlaubt diese topologischen Räume voneinander zu unterscheiden. Die *algebraische Topologie* befaßt sich mit solchen den topologischen Räumen zugeordneten so genannten *topologischen Invarianten*.

5.4 Stetigkeit

Eine reelle Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ wurde *stetig bei* $x \in \mathbb{R}$ genannt, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ mit

$$f(U_\delta(x)) \subseteq U_\varepsilon(f(x))$$

existiert, und *stetig*, wenn dies für alle $x \in \mathbb{R}$ gilt. Ganz analog wurde die $\varepsilon - \delta$ -Stetigkeit einer Abbildung zwischen metrischen Räumen definiert. Diesen Stetigkeitsbegriff verallgemeinert man folgendermaßen:

5.34 Definition

- Eine Abbildung $f : X \rightarrow Y$ zwischen topologischen Räumen heißt **stetig bei** $x \in X$, wenn es zu jeder Umgebung V von $f(x)$ eine Umgebung U von x mit $f(U) \subseteq V$ gibt.
- f heißt **stetig**, wenn sie bei jedem Punkt $x \in X$ stetig ist.
- Eine bijektive stetige Abbildung $f : X \rightarrow Y$ heißt **Homöomorphismus**, wenn auch $f^{-1} : Y \rightarrow X$ stetig ist.
- X und Y heißen **homöomorph**, wenn ein Homöomorphismus $f : X \rightarrow Y$ existiert.

5.35 Bemerkung Analog zum Englischen (*homeomorphism*) benutzt man auch das Wort Homeomorphismus := Homöomorphismus \neq Homomorphismus! \diamond

5.36 Satz Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen ist genau dann stetig im Sinn von Definition 5.34, wenn sie $\varepsilon - \delta$ -stetig ist.

Beweis: Hausaufgabe. □

5.37 Satz Es sei $f : X \rightarrow Y$ eine Abbildung zwischen topologischen Räumen. Dann ist f genau dann stetig, wenn die Urbilder $f^{-1}(V)$ offener Mengen immer offen sind.

Beweis:

- Es sei $f : X \rightarrow Y$ stetig und $V \subseteq Y$ offen. Nach Definition existiert für jedes $x \in U := f^{-1}(V)$ eine in U liegende offene Umgebung U_x von x . Deren Vereinigung $\cup_{x \in U} U_x$ ist offen und gleich U .

- Falls aus der Offenheit von $\tilde{V} \subseteq Y$ die Offenheit von $f^{-1}(\tilde{V}) \subseteq X$ folgt, gibt es für alle $x \in X$ und Umgebungen V von $f(x)$ eine offene Umgebung $\tilde{V} \subseteq V$ von $f(x)$, und für $\tilde{U} := f^{-1}(\tilde{V})$ gilt: \tilde{U} ist eine offene Umgebung von x mit $f(\tilde{U}) \subseteq V$. \square

Damit bildet ein Homöomorphismus $f : X \rightarrow Y$ durch

$$\mathcal{O}_X \rightarrow \mathcal{O}_Y \quad , \quad U \mapsto f(U)$$

die Topologie \mathcal{O}_X bijektiv auf \mathcal{O}_Y ab. Homöomorphismen sind daher die Isomorphismen der Topologie!

Es gibt noch einen zweiten Stetigkeitsbegriff:

5.38 Definition

- Eine Abbildung $f : X \rightarrow Y$ zwischen topologischen Räumen heißt **folgenstetig bei** $x \in X$, wenn für jede gegen x konvergente Folge $a : \mathbb{N} \rightarrow X$ die Bildfolge $f \circ a : \mathbb{N} \rightarrow Y$ gegen $f(x)$ konvergiert.
- f heißt **folgenstetig**, wenn sie bei jedem Punkt $x \in X$ folgenstetig ist.

Soweit wir nur Topologien betrachten, die von Metriken kommen, gilt für uns nach Satz 9.4 der *Analysis I*

"stetig = folgenstetig".

Allgemein gilt aber auf *topologischen* Räumen nur "stetig \Rightarrow folgenstetig".

5.39 Beispiele (Homöomorphe topologische Räume)

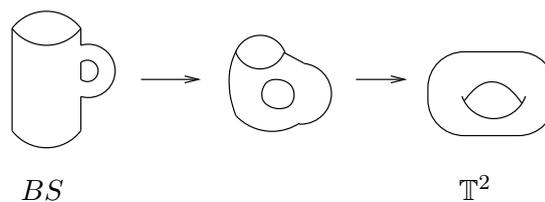
1. Für $\lambda = (\lambda_1, \dots, \lambda_{n+1})$ mit Hauptachsen $\lambda_i > 0$ sei

$$E_\lambda := \{x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} (x_i/\lambda_i)^2 = 1\}$$

das Ellipsoid mit den Hauptachsen der Länge λ_i .

Dann ist $f : E_\lambda \rightarrow S^n$, $x \mapsto \frac{x}{\|x\|}$ ein Homöomorphismus.

2. Ähnlich ist die Oberfläche eines *Bierseidels* zum 2-Torus \mathbb{T}^2 aus Beispiel 5.33 homöomorph:



3. Es lässt sich dagegen zeigen, dass S^2 und \mathbb{T}^2 nicht homöomorph sind. Anschaulich liegt das daran, dass der Torus \mathbb{T}^2 einen Henkel hat, die Sphäre S^2 aber nicht. \diamond

5.40 Satz Sind $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ stetig, so auch $g \circ f : X \rightarrow Z$.

Beweis: Nach Satz 5.37 müssen wir nachprüfen, dass für alle offenen $W \subseteq Z$ auch $U := (g \circ f)^{-1}(W) \subseteq X$ offen ist. Wegen der Stetigkeit von g ist $V := g^{-1}(W) \subseteq Y$ offen, wegen der Stetigkeit von f auch $U = f^{-1}(V)$. \square

5.41 Bemerkung (Topologische Vektorräume)

In der *Funktionalanalysis*, also dem Gebiet der Mathematik, in dem Lineare Algebra und Analysis sich kombinieren, betrachtet man so genannte *topologische Vektorräume*, d.h. Vektorräume mit einer Topologie, die mit Addition und Skalarmultiplikation verträglich ist.

Während nun endliche dimensionale Vektorräume über $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} die Produkttopologie des \mathbb{K}^n tragen, kommen für unendlich dimensionale Vektorräume V oft verschiedene Topologien in Betracht. Oft sind diese Topologien metrische Topologien $\mathcal{O}(d)$ der von einer Norm $\|\cdot\| : V \rightarrow \mathbb{R}$ erzeugten Metrik d . In dieser Topologie ist die Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ immer stetig. \diamond

5.42 Satz Eine lineare Abbildung $f : X \rightarrow Y$ zwischen normierten Vektorräumen $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ ist genau dann stetig, wenn sie *Lipschitz-stetig* ist, d.h. für ein $L > 0$ gilt

$$\|f(x)\|_Y \leq L \cdot \|x\|_X \quad (x \in X). \quad (5.4)$$

Beweis:

- Lipschitz-stetige Abbildungen $f : X \rightarrow Y$ zwischen *metrischen Räumen* sind stetig. Sei nämlich $L > 0$ Lipschitzkonstante, $x \in X$ und V Umgebung von $f(x)$. Dann gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(f(x)) \subseteq V$. Für $\delta := \varepsilon/L$ ist $f(U_\delta(x)) \subseteq U_\varepsilon(f(x))$.
- Wegen $\|f(x_1) - f(x_2)\|_Y = \|f(x)\|_Y$ mit $x := x_1 - x_2$ reicht es aus, die L -Stetigkeit bei der Null zu überprüfen. Ist f stetig, dann gibt es zu $\varepsilon := 1$ ein $\delta > 0$ mit

$$\|f(x)\|_Y \leq 1 \quad (\|x\|_X \leq \delta).$$

Setze $L := 1/\delta$. Dann folgt aus der Linearität von f für $x \neq 0$ und $\hat{x} := \delta x / \|x\|_X$

$$\|f(x)\|_Y = \left\| f(L\|x\|_X \hat{x}) \right\|_Y = \left\| L\|x\|_X f(\hat{x}) \right\|_Y = L\|x\|_X \|f(\hat{x})\|_Y \leq L\|x\|_X,$$

während für $x = 0$ ebenfalls (5.4) folgt. \square

Ist die lineare Abbildung $f : X \rightarrow Y$ stetig, dann definiert man analog zur Matrixnorm die *Operatornorm* von f als

$$\|f\| := \sup_{\|x\|_X=1} \|f(x)\|_Y. \quad (5.5)$$

Diese läßt sich wieder grob als maximaler Streckungsfaktor eines Vektors unter der linearen Abbildung interpretieren, wobei das Supremum in (5.5) nicht angenommen werden muss.

Wie im nächsten Abschnitt gezeigt wird, sind lineare Abbildungen $f : X \rightarrow Y$ immer stetig, falls $\dim(X) < \infty$.

5.43 Beispiel (Unstetige lineare Abbildung)

Es sei

$$X := C^1([0, 1], \mathbb{R}) := \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ stetig differenzierbar}\},$$

und

$$Y := C([0, 1], \mathbb{R}),$$

beide versehen mit der Supremumsnorm. Dann ist die lineare Abbildung

$$D : X \rightarrow Y, \quad f \mapsto f'$$

nach Satz 5.42 *nicht stetig*, denn zwar besitzt $f_n \in X$, $f_n(x) := x^n$ die Norm $\|f_n\| = |f_n(1)| = 1$, aber wegen $Df_n = nf_{n-1}$ ist $\|Df_n\| = n$ ($n \in \mathbb{N}$). \diamond

5.5 Kompaktheit

Im Buch ¹⁵ von KLAUS JÄNICH beginnt das entsprechende Kapitel mit dem Ausruf: "Ah, Kompaktheit! Eine wundervolle Eigenschaft!"

In Kapitel 9.3 des *Analysis I*-Skriptes wurde $A \subseteq \mathbb{R}^n$ kompakt genannt, wenn A beschränkt und abgeschlossen ist.

Der Prototyp eines kompakten Raumes ist daher das abgeschlossene Intervall $[a, b]$, und wir wissen schon, dass darin jede Folge $(x_n)_{n \in \mathbb{N}}$ einen Häufungspunkt $x \in [a, b]$ hat. Das ist eine 'wundervolle Eigenschaft', denn eine geeignete Teilfolge konvergiert dann gegen x .

Dagegen ist \mathbb{R} nicht kompakt, und z.B. die Folge $(x_n := n)_{n \in \mathbb{N}}$ besitzt keinen Häufungspunkt.

In topologischen Räumen definieren wir Kompaktheit ganz anders und zeigen anschließend, dass im Fall des \mathbb{R}^n die Definitionen übereinstimmen.

5.44 Definition

- Eine Teilmenge X eines topologischen Raumes (M, \mathcal{O}) heißt **kompakt**, wenn jede **offene Überdeckung**, d.h. jede Familie $(V_i)_{i \in I}$ mit $V_i \in \mathcal{O}$ und $X \subseteq \bigcup_{i \in I} V_i$, eine **endliche Teilüberdeckung**

$$X \subseteq V_{j_1} \cup \dots \cup V_{j_n} \quad \text{mit} \quad j_1, \dots, j_n \in I$$

besitzt.

¹⁵Klaus Jänich: *Topologie*. Springer, 1999

- X heißt **folgenkompakt**, wenn jede Folge $a : \mathbb{N} \rightarrow X$ eine gegen ein $x \in X$ konvergente Teilfolge besitzt.

5.45 Beispiel $X := (0, 1] \subset \mathbb{R}$ ist nicht kompakt, denn mit $V_i := (\frac{1}{i}, 2)$ ($i \in \mathbb{N}$) ist auch $U_i := V_i \cap X = (\frac{1}{i}, 1]$ in X offen. Die Familie $(U_i)_{i \in \mathbb{N}}$ ist eine offene Überdeckung von X . Gäbe es aber eine endliche Teilüberdeckung $X = U_{j_1} \cup \dots \cup U_{j_n}$, dann wäre mit $j := \max(j_1, \dots, j_n)$ der Punkt $\frac{1}{j} \in X$. Er ist aber in keinem U_{j_i} . Widerspruch! \diamond

Natürlich *existieren* für jede Teilmenge X eines topologischen Raumes M immer endliche offene Überdeckungen, z.B. die einelementige bestehend aus M .

Es gibt topologische Räume, die kompakt, aber nicht folgenkompakt sind und solche, in denen das Umgekehrte gilt. Aber:

5.46 Satz Ein metrischer Raum (M, d) ist genau dann kompakt, wenn er folgenkompakt ist.

Beweis: In beiden Richtungen führen wir Widerspruchsbeweise:

1. (M, d) sei kompakt. Wir nehmen nun die Existenz einer Folge $a : \mathbb{N} \rightarrow M$ ohne konvergente Teilfolge an.

Dann existiert für jeden Punkt $x \in M$ eine Umgebung V_x , die nur von endlich vielen Folgengliedern getroffen wird. (Warum? Andernfalls existierte ja ein Punkt $x \in M$, dessen Umgebungen $U_{1/n}(x)$ ($n \in \mathbb{N}$) alle von unendlich vielen Folgengliedern getroffen würden. Dann könnten wir eine gegen x konvergente Teilfolge auswählen).

Nach Voraussetzung besitzt aber die offene Überdeckung $(V_x)_{x \in M}$ eine endliche Teilüberdeckung, und damit die Folge nur endlich viele Glieder. Widerspruch!

2. (M, d) sei folgenkompakt, aber nicht kompakt. Es existiert also eine Überdeckung $(V_i)_{i \in I}$ von M ohne endliche Teilüberdeckung.

Für jeden Punkt $x \in M$ und $J(x) := \{i \in I \mid x \in V_i\}$ setzen wir

$$R(x, i) := \min(\sup\{r > 0 \mid U_r(x) \subseteq V_i\}, 1) \quad (i \in J(x)),$$

sodass für diese Radien also $U_{R(x, i)}(x) \subseteq V_i$ gilt.

Als Nächstes wählen wir für jedes $x \in M$ eine Umgebung $V_{j(x)}$ von x mit

$$j(x) \in J(x) \quad \text{so dass} \quad R(x, j(x)) \geq \frac{1}{2} \sup_{i \in J(x)} R(x, i). \quad (5.6)$$

Anschließend konstruieren wir induktiv eine Folge $(x_n)_{n \in \mathbb{N}}$ in M mit

$$x_{n+1} \notin V_{j(x_1)} \cup \dots \cup V_{j(x_n)} \quad (n \in \mathbb{N}) \quad (5.7)$$

(da nach Annahme keine endliche Teilüberdeckung existiert, ist dies möglich).

$(x_n)_{n \in \mathbb{N}}$ besitzt nach Annahme eine konvergente Teilfolge, und wir nehmen o.B.d.A. an, dass $(x_n)_{n \in \mathbb{N}}$ selbst gegen $a \in M$ konvergiert, also $\lim_{n \rightarrow \infty} d(x_n, a) = 0$.

Damit muss wegen (5.7) der Limesradius $\lim_{n \rightarrow \infty} R(x_n, j(x_n)) = 0$ sein.

Andererseits ist $R(a, j(a)) > 0$, es gibt also ein $N \in \mathbb{N}$ mit $d(x_n, a) < \frac{1}{4}R(a, j(a))$ für alle $n \geq N$. Damit ist nicht nur $x_n \in U_{R(a, j(a))}(a)$, sondern sogar die Kugel $U_{\frac{1}{4}R(a, j(a))}(x_n)$ um x_n mit n -unabhängigem Radius in $U_{R(a, j(a))}(a)$ enthalten. Wir haben also für ein $n \in \mathbb{N}$ die Bedingung (5.6) verletzt. Widerspruch! \square

5.47 Korollar *Kompakte metrische Räume sind vollständig.*

Beweis: Denn konvergiert eine Teilfolge einer Cauchy-Folge, so auch die Cauchy-Folge selbst. \square

5.48 Beispiel $\mathbb{Q} \cap [0, 1]$ ist nicht vollständig (da $1/\sqrt{2} \notin \mathbb{Q}$), also nicht kompakt!

Satz 5.46 ermöglicht uns zu zeigen, dass die Kompaktheits-Definition der *Analysis I* Spezialfall unserer jetzigen Definition ist.

5.49 Satz (Heine–Borel) $A \subseteq \mathbb{R}^n$ ist genau dann kompakt im Sinne von Definition 5.44, wenn A beschränkt und abgeschlossen ist.

Beweis:

- Falls A beschränkt und abgeschlossen ist, ist A nach dem Satz von Bolzano–Weierstraß (Satz 7.41 der *Analysis I*) folgenkompakt und damit nach Satz 5.46 kompakt.
- Ist A unbeschränkt, dann existiert eine Folge $(x_n)_{n \in \mathbb{N}}$ in A mit $\lim_{n \rightarrow \infty} \|x_n\| = \infty$. Diese Folge divergiert also, und A ist nicht (folgen)kompakt.
- Gleiches gilt, falls A nicht abgeschlossen ist, also einen nicht in A liegenden Häufungspunkt besitzt. \square

Man könnte nun denken, dass allgemein Abgeschlossenheit und Beschränktheit¹⁶ einer Teilmenge eines metrischen Raumes ihre Kompaktheit impliziert. Das gilt aber noch nicht einmal für Teilmengen normierter Vektorräume:

5.50 Bemerkung (Nicht kompakt, aber abgeschlossen und beschränkt)

Die Einheitskugel $S := \{f \in \ell^2 \mid \|f\|_2 = 1\}$ im *Hilbertschen Folgenraum*

$$\ell^2 := \ell^2(\mathbb{N}) := \{f : \mathbb{N} \rightarrow \mathbb{C} \mid \sum_{n=1}^{\infty} |f(n)|^2 < \infty\}$$

bezüglich der Norm $\|f\|_2 := \sqrt{\sum_{n=1}^{\infty} |f(n)|^2}$ ist beschränkt und abgeschlossen, aber nicht kompakt. Denn die Folge $(e_n)_{n \in \mathbb{N}}$ der (Basis-) Vektoren $e_n(k) := \delta_{n,k}$ ($k \in \mathbb{N}$) liegt in S , besitzt aber wegen $\|e_n - e_m\|_2 = \sqrt{2}$ ($m \neq n$) keine konvergente Teilfolge.

Dieses Beispiel ist in der Quantenmechanik wichtig, weil alle dort benutzten Hilbert-Räume entweder endlichdimensional oder isomorph zu ℓ^2 sind. Die quantenmechanischen Zustände sind dann Punkte in S . \diamond

¹⁶**Definition** Eine Teilmenge $A \subseteq X$ eines metrischen Raums (X, d) heißt **beschränkt**, wenn $A = \emptyset$ oder wenn ihr **Durchmesser** $\text{diam}(A) := \sup_{x, y \in A} d(x, y)$ endlich ist.

5.51 Satz *Kompakte Teilmengen metrischer Räume sind abgeschlossen.*

Beweis: Metrischen Räume sind Hausdorff-Räume, und wir zeigen die Behauptung allgemein für Hausdorff-Räume M . Sei also $X \subseteq M$ kompakt. Zu zeigen ist: $Y := M - X$ ist offen.

Für $y \in Y$ konstruieren wir zum Nachweis eine offene Umgebung $V \subseteq Y$ von y . Für jedes $x \in X$ existieren wegen der Hausdorff-Eigenschaft disjunkte offene Umgebungen $U_x \subseteq M$ von x und $V_x \subseteq M$ von y .

Wegen der Kompaktheit von X gibt es eine endliche Teilüberdeckung

$$U_{x_1} \cup \dots \cup U_{x_n} \supseteq X$$

der Überdeckung $(U_x)_{x \in X}$ von X . Für die offene Umgebung

$$V := V_{x_1} \cap \dots \cap V_{x_n} \subseteq M$$

von y gilt: $U_{x_i} \cap V \subseteq U_{x_i} \cap V_{x_i} = \emptyset$ ($i = 1 \dots, n$). Also ist $V \subseteq M - X = Y$. \square

Gegenbeispiel (Kompakt, aber nicht abgeschlossen):

$M := \{1, 2\}$ mit indiskreter Topologie $\mathcal{O} = \{\emptyset, \{1, 2\}\}$.

$\{1\} \subset \{1, 2\}$ ist zwar als endliche Menge kompakt, aber nicht abgeschlossen, da $\{2\}$ nicht offen ist. \diamond

Die folgenden beiden Sätze können oft zum Nachweis der Kompaktheit benutzt werden.

5.52 Satz

Ist $X \subseteq M$ kompakt und $f : M \rightarrow Y$ stetig, dann ist $f(X) \subseteq Y$ kompakt.

Beweis: Sei $(V_i)_{i \in I}$ eine offene Überdeckung von $f(X)$ und

$$U_i := f^{-1}(V_i) = \{x \in X \mid f(x) \in V_i\}.$$

Dann ist nach Satz 5.37 auch U_i offen, und es gibt eine endliche Teilüberdeckung $X \subseteq U_{j_1} \cup \dots \cup U_{j_m}$. Damit ist $f(X) \subseteq V_{j_1} \cup \dots \cup V_{j_m}$. \square

Ein weiteres Prinzip zur Identifikation kompakter Mengen ist das folgende:

5.53 Satz

Abgeschlossene Teilmengen kompakter topologischer Räume sind kompakt.

Beweis: Es sei $A \subseteq K$ eine abgeschlossene Teilmenge des kompakten Raums K und $(V_i)_{i \in I}$ eine Überdeckung von A durch (in A) offene Teilmengen $V_i \subseteq A$. Nach Definition 5.31 der Teilraumtopologie gibt es daher in K offene Teilmengen $(U_i)_{i \in I}$ mit $V_i = U_i \cap A$. Die U_i bilden zusammen mit der offenen Menge $K \setminus A$ eine offene Überdeckung von K , die wegen Kompaktheit von K eine endliche Teilüberdeckung besitzt:

$$K = (K \setminus A) \cup U_{j_1} \cup \dots \cup U_{j_m}.$$

Damit ist $A = V_{j_1} \cup \dots \cup V_{j_m}$. \square

Eine Abbildung ist genau dann stetig, wenn die Urbilder der offenen Mengen

offen sind, oder äquivalent, wenn die Urbilder der abgeschlossenen Mengen abgeschlossen sind.

So gewinnt man viele Kompakta:

5.54 Beispiel Die n -Sphäre $S^n = \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$ ist *kompakt*, denn

- S^n ist Urbild der abgeschlossenen Menge $\{1\} \subseteq \mathbb{R}$ unter der stetigen Abbildung $x \mapsto \|x\|$, also abgeschlossen, und
- $S^n \subseteq [-1, 1]^{n+1}$. Die Kompaktheit der Quader haben wir aber schon gezeigt. \diamond

5.55 Satz *Stetige Funktionen* $f : X \rightarrow \mathbb{R}$ auf kompakten metrischen Räumen $X \neq \emptyset$ nehmen ihr Minimum und Maximum an, d.h. es gibt $x_{\min}, x_{\max} \in X$ mit

$$\inf_{x \in X} f(x) = f(x_{\min}) \quad , \quad \sup_{x \in X} f(x) = f(x_{\max}).$$

Beweis: Da $f(X) \neq \emptyset$, gibt es eine Folge $(x_n)_{n \in \mathbb{N}}$ mit $\sup_x f(x) = \lim_{n \rightarrow \infty} f(x_n)$. wegen der Äquivalenz von Kompaktheit und Folgenkompaktheit (Satz 5.46) gibt es eine konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$. Wegen der Stetigkeit von f und der Äquivalenz von Stetigkeit und Folgenstetigkeit (Satz 9.4 der *Analysis I*) gilt

$$\lim_{k \rightarrow \infty} f(x_{n_k}) = f(x_{\max}) \quad \text{für} \quad x_{\max} := \lim_{k \rightarrow \infty} x_{n_k}.$$

Analog für das Infimum. \square

Während verschiedene Normen auf einem Vektorraum unterschiedliche geometrische Strukturen erzeugen, können sie einander doch in einem größeren Sinn gleichen:

5.56 Definition Zwei Normen $\|\cdot\|_I$ und $\|\cdot\|_{II}$ auf dem \mathbb{K} -Vektorraum V heißen **äquivalent**, wenn für geeignete $C' \geq C > 0$ gilt:

$$C\|v\|_I \leq \|v\|_{II} \leq C'\|v\|_I \quad (v \in V).$$

Als Anwendung unserer Kompaktheitsätze können wir jetzt den

5.57 Satz *Alle Normen auf einem endlich-dimensionalen \mathbb{K} -Vektorraum ($\mathbb{K} = \mathbb{R}$ oder \mathbb{C}) sind äquivalent*

beweisen:

Beweis: • Ohne Beschränkung der Allgemeinheit betrachten wir den \mathbb{K}^n .

- Es sei $\|\cdot\| : \mathbb{K}^n \rightarrow \mathbb{R}$ eine Norm. Es genügt die Äquivalenz dieser Norm zur euklidischen Norm $\|\cdot\|_2$ nachzuweisen, denn Äquivalenz von Normen ist eine Äquivalenzrelation.
- Wir zeigen zunächst, dass die Abbildung $\|\cdot\| : \mathbb{K}^n \rightarrow \mathbb{R}$ stetig ist. Es sei dazu e_1, \dots, e_n die kanonische Basis des \mathbb{K}^n (mit $\|e_i\|_2 = 1$) und $c_i := \|e_i\|$ die Norm des i -ten Basiselementes.

Dann besitzt ein Vektor $x = \sum_{k=1}^n x_k e_k \in \mathbb{K}^n$ die euklidische Länge $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ und nach der Dreiecksungleichung die Länge

$$\|x\| \leq \sum_{i=1}^n \|x_i e_i\| = \sum_{i=1}^n |x_i| c_i \leq c \sum_{i=1}^n |x_i| \leq cn \|x\|_2$$

mit $c := \max(c_1, \dots, c_n)$. Damit ist die Norm-Abbildung stetig.

- Nun restringieren wir diese auf die Einheitsphäre

$$S_{\mathbb{K}}^{n-1} := \{x \in \mathbb{K}^n \mid \|x\|_2 = 1\}$$

des Euklidischen Vektorraums, betrachten also die Abbildung

$$f : S_{\mathbb{K}}^{n-1} \rightarrow \mathbb{R}, \quad x \mapsto \|x\|.$$

Da $S_{\mathbb{K}}^{n-1}$ nach Beispiel 5.54 kompakt¹⁷ und f stetig ist, nimmt f Minimum $f_{\min} > 0$ und Maximum $f_{\max} < \infty$ an. Daher gilt $f_{\min} \|x\|_2 \leq \|x\| \leq f_{\max} \|x\|_2$. \square

5.58 Satz Sind $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ und $\dim(X) < \infty$ \mathbb{K} -Vektorräume, dann ist jede lineare Abbildung $f : X \rightarrow Y$ stetig.

Beweis: Nach Satz 5.57 können wir annehmen, dass $\|\cdot\|_X$ die euklidische Norm ist. Es sei $e_1, \dots, e_n \in X$ eine Orthonormalbasis und $x = \sum_{i=1}^n c_i e_i \in X$ von der Norm 1, also $\sum_{i=1}^n |c_i|^2 = 1$. Dann ist $\|f(x)\|_Y = \|\sum_{i=1}^n c_i f(e_i)\|_Y \leq \sum_{i=1}^n |c_i| \|f(e_i)\|_Y \leq L := n \max\{\|f(e_i)\|_Y \mid i = 1, \dots, n\}$. L ist damit eine Lipschitzkonstante. \square

6 Kurven im \mathbb{R}^n

Wir beginnen mit der mehrdimensionalen Differentialrechnung, indem wir den einfachsten Fall herausgreifen, den der Kurven im \mathbb{R}^n .

6.1 Reguläre und nicht reguläre Kurven

6.1 Definition

- Stetige Abbildungen $c = (c_1, \dots, c_n)^t : I \rightarrow \mathbb{R}^n$, deren Definitionsbereich I ein Intervall (mit $I \neq \{a\}$) ist, heißen **Kurven im \mathbb{R}^n** .
- $t \in I$ heißt der **Parameter** von c , ihr Bild $c(I) \subseteq \mathbb{R}^n$ auch die **Spur** von c .
- c heißt k -mal (**stetig**) differenzierbar, wenn alle Komponenten c_1, \dots, c_n k -mal (**stetig**) differenzierbar sind.

¹⁷Ist $\mathbb{K} = \mathbb{C}$, dann ist $S_{\mathbb{C}}^{n-1} = S_{\mathbb{R}}^{2n-1}$, wenn man den \mathbb{C}^n reellifiziert d.h. mit dem \mathbb{R}^{2n} identifiziert.

- Die Menge der k -mal stetig differenzierbaren Kurven $c : I \rightarrow \mathbb{R}^n$ wird mit $C^k(I, \mathbb{R}^n)$ bezeichnet und man setzt $C(I, \mathbb{R}^n) := C^0(I, \mathbb{R}^n)$.
- Eine stetig differenzierbare Kurve heißt **regulär**, wenn ihre **Ableitung** oder **Geschwindigkeit**

$$c' : I \rightarrow \mathbb{R}^n \quad , \quad c'(t) := \begin{pmatrix} c'_1(t) \\ \vdots \\ c'_n(t) \end{pmatrix} \quad (t \in I)$$

für alle $t \in I$ ungleich Null ist.

Nach Satz 7.32 der *Analysis 1* ist die Existenz der Ableitung der Kurve bei t

$$c'(t) = \lim_{\delta \rightarrow 0} \frac{c(t + \delta) - c(t)}{\delta} = \lim_{\delta \rightarrow 0} \begin{pmatrix} \frac{c_1(t+\delta) - c_1(t)}{\delta} \\ \vdots \\ \frac{c_n(t+\delta) - c_n(t)}{\delta} \end{pmatrix}$$

gleichbedeutend mit der Existenz der Ableitungen $c'_k(t) = \lim_{\delta \rightarrow 0} \frac{c_k(t+\delta) - c_k(t)}{\delta}$ aller Komponenten c_k von c .

In physikalischem Zusammenhang ist oft die Zeit Parameter der Kurve.

6.2 Beispiele (Kurven)

1. Die Kurve $c : \mathbb{R} \rightarrow \mathbb{R}^2$, $c(t) := x + vt$ für $x, v \in \mathbb{R}^2$ ist eine Gerade durch x mit der Richtung v , siehe Abbildung 6.1. Ihre Spur unterscheidet sich nicht von der Spur der Kurve

$$d : \mathbb{R} \rightarrow \mathbb{R}^2 \quad , \quad d(t) := x + vt^3.$$

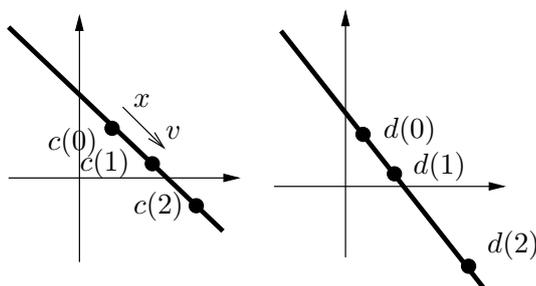
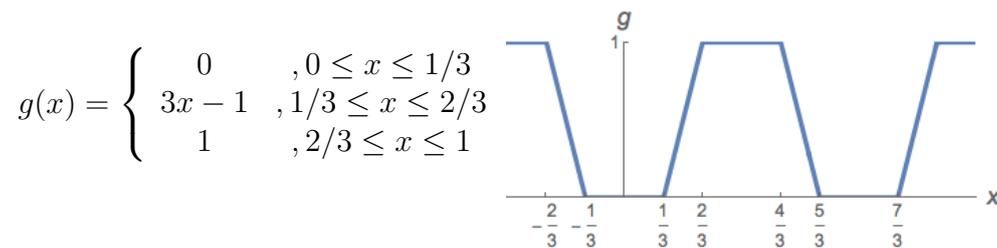


Abbildung 6.1: Zwei Kurven mit gleicher Spur

2. Der Graph einer C^k -Funktion $f : I \rightarrow \mathbb{R}$ ist Bild der C^k -Kurve

$$c : I \rightarrow \mathbb{R}^2, \quad c(t) := \begin{pmatrix} t \\ f(t) \end{pmatrix}.$$

3. Es sei $g : \mathbb{R} \rightarrow [0, 1]$, die gerade 2-periodische Funktion mit



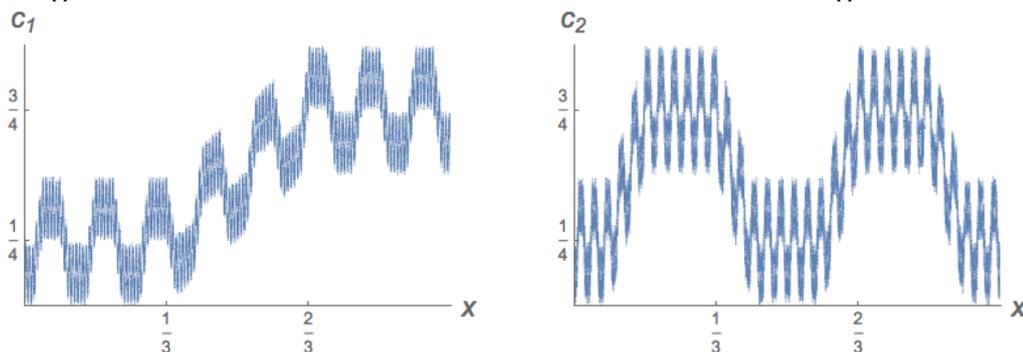
(siehe nebenstehende Abbildung) und $c : [0, 1] \rightarrow \mathbb{R}^2$

$$c(x) := \begin{pmatrix} c_1(x) \\ c_2(x) \end{pmatrix} = \left(\sum_{k=0}^{\infty} \frac{g(4^{2k}x)}{2^{k+1}}, \sum_{k=0}^{\infty} \frac{g(4^{2k+1}x)}{2^{k+1}} \right)^t.$$

Dann sind die Komponenten c_1 und c_2 von c als gleichmäßige Limiten der stetigen Funktionen $c_{i,n} : \mathbb{R} \rightarrow \mathbb{R}$,

$$c_{1,n}(x) := \sum_{k=0}^n \frac{g(4^{2k}x)}{2^{k+1}}, \quad c_{2,n}(x) := \sum_{k=0}^n \frac{g(4^{2k+1}x)}{2^{k+1}}$$

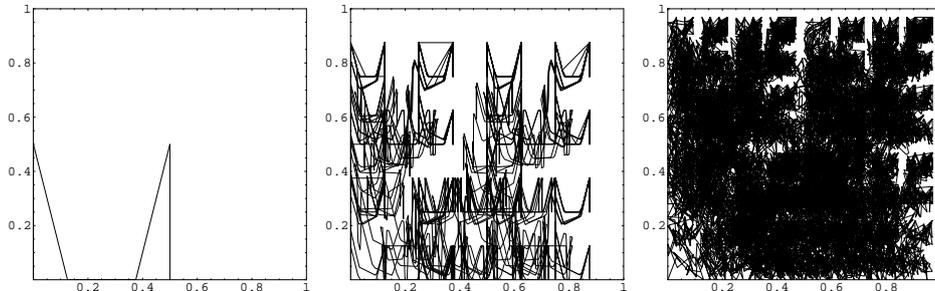
stetiger Funktionenreihen nach dem Weierstraß-Kriterium¹⁸ stetig.



Das Bild (= die Spur) der Kurve c ist aber $c([0, 1]) = [0, 1]^2 \subset \mathbb{R}^2$, denn der Punkt $(y, z) \in [0, 1] \times [0, 1]$ ist Bild von $x = \sum_{k=0}^{\infty} c_k 4^{-k} \in [0, 1]$, falls $y = \sum_{k=0}^{\infty} c_{2k} 2^{-k-1}$, $z = \sum_{k=0}^{\infty} c_{2k+1} 2^{-k-1}$ die Binärdarstellungen von y und z sind, mit $c_k \in \{0, 1\}$.

¹⁸**Satz (Weierstraß):** Der Limes einer gleichmäßig konvergenten Folge stetiger Funktionen ist stetig. Beweis: Siehe z.B.: FORSTER [Fo, §21]

Versuchen Sie, dies selbst zu beweisen (Tipp: für $x = \sum_l c_l 4^{-l}$ ist $g(4^r x) = c_r$).



Die Spuren der Kurven $(c_{1,n}, c_{2,n})$ für $n = 0$ (links), $n = 2$ (Mitte) und $n = 4$ (rechts)

Die Existenz solcher das Quadrat ausfüllender *Peanokurven* war ein Schock für die Mathematiker, da das Bild eine größere Dimension als das Urbild hat. Ein ähnliches Phänomen ist bei der *Koch-* oder *Schneeflockenkurve*, siehe Abb. 6.2 zu beobachten.

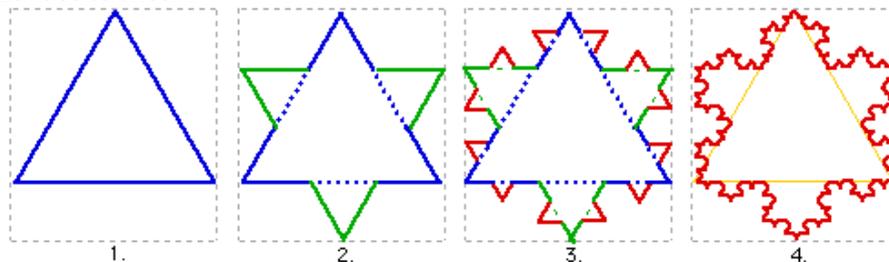


Abbildung 6.2: Konstruktion der Koch-Kurve (rechts)

Versucht man, derartigen Kurven eine Länge zuzuordnen, indem man sie durch Polygonzüge approximiert und den Limes der Länge dieser Polygonzüge bildet, so ergibt sich eine unendliche Länge. c heißt dann *nicht rektifizierbar*. \diamond

Im Gegensatz dazu ist für "vernünftige" Kurven die Länge zwischen zwei Punkten endlich.

6.3 Definition Eine C^1 -Kurve $c : I \rightarrow \mathbb{R}^n$ hat die **Länge**

$$L(c) := \int_I \|c'(t)\| dt = \int_I \sqrt{c_1'(t)^2 + \dots + c_n'(t)^2} dt.$$

Diese Längendefinition entspricht in der physikalischen Analogie der Länge als Zeitintegral über den Geschwindigkeitsbetrag.

Wie Beispiel 6.2.1. zeigt, kann $L(c) = \infty$ vorkommen, aber

6.4 Satz Die Länge $L(c)$ einer C^1 -Kurve $c : [t_0, t_1] \rightarrow \mathbb{R}^n$ ist endlich.

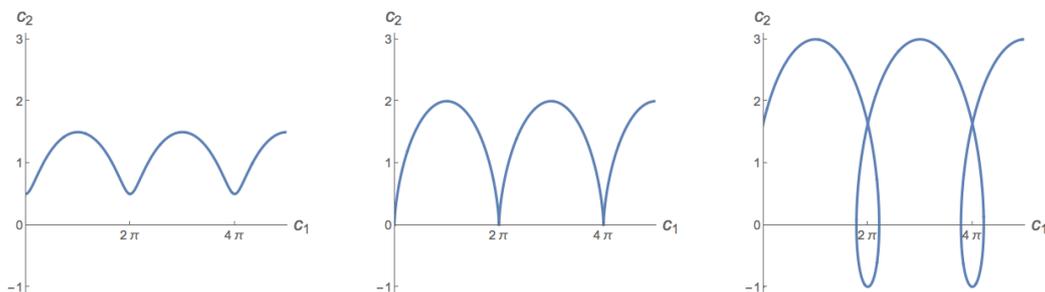


Abbildung 6.3: Zykloiden für Parameter $b = 1$ und $a = 0.5$ (links), $a = 1$ (Mitte) und $a = 2$ (rechts)

Beweis: Die Abbildung $[t_0, t_1] \rightarrow \mathbb{R}, t \mapsto \|c'(t)\|$ ist stetig, besitzt also ein endliches Supremum v auf dem Kompaktum $[t_0, t_1]$. Daher ist $L(c) \leq (t_1 - t_0)v$. \square

Die Spur einer C^1 -Kurve muss nicht glatt aussehen, wenn c nicht regulär ist.

6.5 Beispiel (Regularität von Kurven)

Die *Zykloiden* mit Parametern $a, b > 0$

$$c : \mathbb{R} \rightarrow \mathbb{R}^2, \quad c(t) := \begin{pmatrix} bt - a \sin t \\ b - a \cos t \end{pmatrix}$$

bilden eine Kurvenfamilie, die zwischen Kreislinien ($b = 0$) und Geraden ($a = 0$) interpoliert.

Genau im Fall $a = b$ ist die Kurve nicht regulär, denn dann ist $c'(2\pi n) = 0$ für alle $n \in \mathbb{Z}$. Die Spur von c ist dort nicht glatt, denn $c'(t) = a \begin{pmatrix} 1 - \cos t \\ \sin t \end{pmatrix}$ besitzt den Betrag $\|c'(t)\| = a\sqrt{2(1 - \cos t)}$, sodass $\lim_{\delta \searrow 0} \frac{c'(2\pi n \pm \delta)}{\|c'(2\pi n \pm \delta)\|} = \mp 1$ ist.

Der Fall $a < b$ entspricht der Bahn eines an den Fahrradspeichen angebrachten Reflektors. \diamond

6.2 Wechsel der Parametrisierung

Die Länge einer (injektiven) C^1 -Kurve sollte nur von ihrem Bild, nicht von ihrer Parametrisierung abhängen. Um dies nachzuprüfen, betrachten wir Parametertransformationen:

6.6 Definition Es seien $c : I \rightarrow \mathbb{R}^n$ und $\tilde{c} : \tilde{I} \rightarrow \mathbb{R}^n$ zwei C^1 -Kurven, und $\varphi : \tilde{I} \rightarrow I$ ein **Diffeomorphismus** (d.h. eine bijektive C^1 -Abbildung, deren Umkehrung auch C^1 -Abbildung ist).

\tilde{c} geht aus c durch die Parametertransformation φ hervor, wenn $\tilde{c} = c \circ \varphi$.

Nun muss auf \tilde{I} das Vorzeichen von φ' konstant sein, denn sonst würde die stetige Funktion $\varphi' : \tilde{I} \rightarrow \mathbb{R}$ nach dem Zwischenwertsatz eine Nullstelle $t \in \tilde{I}$ besitzen und die Umkehrfunktion φ^{-1} könnte nicht differenzierbar sein.

Ist $\text{sign}(\varphi') = 1$, dann heißt φ *orientierungserhaltend*, für $\text{sign}(\varphi') = -1$ *orientierungsumkehrend*.

6.7 Satz Für c, \tilde{c} und $\tilde{\varphi}$ wie oben und $\tilde{I} = [t_0, t_1]$ ist $L(c) = L(\tilde{c})$.

Beweis: $L(\tilde{c}) = \int_{t_0}^{t_1} \|\tilde{c}'(t)\| dt = \int_{t_0}^{t_1} \left\| \frac{d}{dt} c(\varphi(t)) \right\| dt$
 $= \int_{t_0}^{t_1} \|c'(\varphi(t)) \cdot \varphi'(t)\| dt = \int_{t_0}^{t_1} \|c'(\varphi(t))\| \left| \frac{d\varphi}{dt}(t) \right| dt = \text{sign}(\varphi') \int_{\varphi(t_0)}^{\varphi(t_1)} \|c'(s)\| ds = L(c). \quad \square$

6.8 Definition Gilt $\|c'(t)\| = 1$ für alle $t \in I$, dann ist die C^1 -Kurve $c : I \rightarrow \mathbb{R}^n$ durch die **Bogenlänge parametrisiert**.

Der Vorteil dieser Parametrisierung liegt in ihrer geometrischen Natürlichkeit, die zu einfacheren Formeln führt.

6.9 Satz Jede reguläre Kurve $c : I \rightarrow \mathbb{R}^n$ kann durch die Bogenlänge parametrisiert werden, d.h. es gibt eine Parametertransformation $\varphi : \tilde{I} \rightarrow I$, sodass die Kurve $c \circ \varphi$ durch die Bogenlänge parametrisiert ist.

Beweis: Es sei für ein beliebig gewähltes $t_0 \in I$

$$\psi(t) := \int_{t_0}^t \|c'(\tau)\| d\tau \quad (t \in I) \quad \text{und} \quad \tilde{I} := \psi(I).$$

Wegen der Stetigkeit des Integranden ist $\psi \in C^1(I, \tilde{I})$, und es gilt $\psi'(t) = \|c'(t)\| > 0$, also für $\varphi := \psi^{-1}$

$$\varphi'(\psi(t)) = \frac{1}{\psi'(t)} = \frac{1}{\|c'(t)\|} \quad (t \in I).$$

Damit gilt für alle $s \in \tilde{I}$ nach der Kettenregel

$$\left\| \frac{d}{ds} (c \circ \varphi)(s) \right\| = \|c'(\varphi(s)) \cdot \varphi'(s)\| = 1,$$

$\tilde{c} := c \circ \varphi : \tilde{I} \rightarrow \mathbb{R}^n$ ist also durch die Bogenlänge parametrisiert. □

6.3 Krümmung einer Kurve

Zwei reguläre Kurven $c : I \rightarrow \mathbb{R}^n$, $\tilde{c} : \tilde{I} \rightarrow \mathbb{R}^n$ können sich an einem Punkt $c(t_0) = \tilde{c}(\tilde{t}_0)$ schneiden. Ihr *Schnittwinkel* ist dann als der Winkel $\varphi \in [0, 2\pi]$ mit

$$\cos \varphi = \frac{\langle c'(t_0), \tilde{c}'(\tilde{t}_0) \rangle}{\|c'(t_0)\| \cdot \|\tilde{c}'(\tilde{t}_0)\|}$$

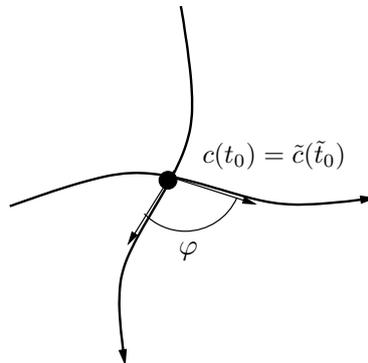


Abbildung 6.4: Winkel zwischen den Geschwindigkeitsvektoren

gegeben, entspricht also dem Winkel zwischen den Geschwindigkeitsvektoren, siehe Abb. 6.4.

Eine reguläre C^2 -Kurve $c : I \rightarrow \mathbb{R}^n$ besitzt die Richtung $\frac{c'(t)}{\|c'(t)\|}$ bei $t \in I$, und diese ändert sich mit t . Die Geschwindigkeit dieser Richtungsänderung muss etwas mit der Beschleunigung $c''(t)$ zu tun haben.

Wir betrachten zunächst Kurven in der Ebene.

6.10 Satz Ist $c : I \rightarrow \mathbb{R}^2$ eine reguläre C^2 -Kurve und $\varphi(t)$ der Winkel zwischen $c'(t)$ und der 1-Richtung, dann ist

$$\varphi'(t) = \frac{c_2''(t)c_1'(t) - c_1''(t)c_2'(t)}{\|c'(t)\|^2} \quad (t \in I).$$

Beweis: Es ist für $c_1'(t) \neq 0$

$$\tan(\varphi(t)) = \frac{c_2'(t)}{c_1'(t)}, \quad \text{also} \quad \frac{d}{dt} \tan \varphi(t) = \frac{c_2''(t)c_1'(t) - c_1''(t)c_2'(t)}{(c_1'(t))^2}$$

und

$$\frac{d}{d\varphi} \tan(\varphi) = 1 + \tan^2(\varphi) = \frac{(c_1')^2 + (c_2')^2}{(c_1')^2}$$

woraus die Formel wegen

$$\varphi'(t) = \frac{\frac{d}{dt} \tan \varphi(t)}{\frac{d}{d\varphi} \tan \varphi(t)}$$

folgt. Für $c_1'(t) = 0$ ist $c_2'(t) \neq 0$ und benutzt man den Cotangens. □

Wir definieren nun die Krümmung als die Richtungsänderung pro Bogenlänge:

6.11 Definition Die Krümmung von c bei $t \in I$ ist

$$k(t) := \frac{\varphi'(t)}{\|c'(t)\|}.$$

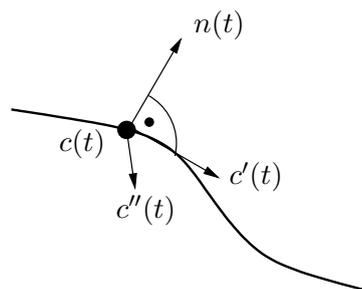
Damit ergibt sich aus Satz 6.10

$$k(t) = \frac{c_2''(t)c_1'(t) - c_1''(t)c_2'(t)}{\|c'(t)\|^3}.$$

$J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ist die Matrix einer Drehung um $\pi/2$. Der *Normalenvektor*

$$n(t) := J \frac{c'(t)}{\|c'(t)\|}$$

steht damit senkrecht auf dem Geschwindigkeitsvektor $c'(t)$ und hat die Länge 1. Damit ist die Krümmung proportional zur Beschleunigung in Normalenrichtung:



$$k(t) = \frac{\langle c''(t), n(t) \rangle}{\|c'(t)\|^2}.$$

6.12 Bemerkungen (Krümmung von Kurven) 1. Bei nach Bogenlänge parametrisierten Kurven ist der Nenner gleich Eins, also $k(t) = \langle c''(t), n(t) \rangle$. Zusätzlich ist die tangentielle Beschleunigungskomponente

$$\langle c''(t), c'(t) \rangle = \frac{1}{2} \frac{d}{dt} \langle c'(t), c'(t) \rangle = \frac{d}{dt} 1 = 0,$$

also die Krümmung betragsmäßig gleich der Beschleunigung:

$$|k(t)| = \|c''(t)\|.$$

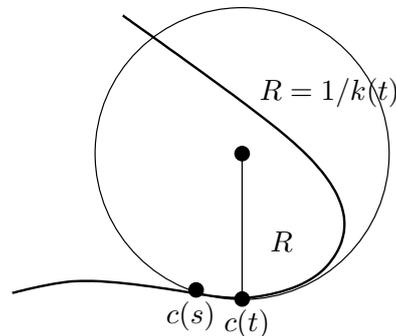
2. Bei Änderung der regulären Parametrisierung von $c : I \rightarrow \mathbb{R}^2$, d.h. für $\tilde{c} := c \circ \psi : \tilde{I} \rightarrow \mathbb{R}^2$ für $\psi : \tilde{I} \rightarrow I$, gilt für die Krümmung von \tilde{c}

$$k_{\tilde{c}}(s) = \text{sign}(\psi') k_c(\psi(s)) \quad (s \in \tilde{I}).$$

Es wird also die Krümmung von c nur mit dem (konstanten) Vorzeichen der Ableitung ψ' der Parametertransformation multipliziert, denn so transformiert sich der Normalenvektor n . Der *Betrag* der Krümmung ändert sich aber nicht. \diamond

Es gibt eine weitere geometrische Interpretation von $k(t)$.

Sei nämlich $c : I \rightarrow \mathbb{R}^2$ irgendeine reguläre C^2 -Kurve, deren Krümmung bei $t \in I$ nicht verschwindet. Dann ist $|k(t)|$ gleich dem inversen Radius des *Schmiegekreises* bei t , also des eindeutig bestimmten Kreises, der die Kurve bei $c(t)$ tangential berührt, und dessen Abstand von $c(s)$ von der Ordnung $\mathcal{O}(|s - t|^3)$ ist.



Die Verallgemeinerung des Krümmungsbegriffes auf Kurven im \mathbb{R}^n ist möglich:

6.13 Definition Eine nach Bogenlänge parametrisierte Kurve $c \in C^2([a, b], \mathbb{R}^n)$ besitzt die (betragsmäßige) **Krümmung** $\|c''(t)\|$.

Nach Bemerkung 6.12.1. stimmt $\|c''(t)\|$ für $n = 2$ mit $|k(t)|$ überein.

Ist eine reguläre C^2 -Kurve c selbst nicht nach Bogenlänge parametrisiert, so definiert man ihre Krümmung durch die einer nach Bogenlänge umparametrisierten Kurve (Die Wahl von Anfangspunkt und Orientierung spielen dabei keine Rolle!).

6.14 Beispiel Für Parameter $R, b > 0$ besitzt die Kurve

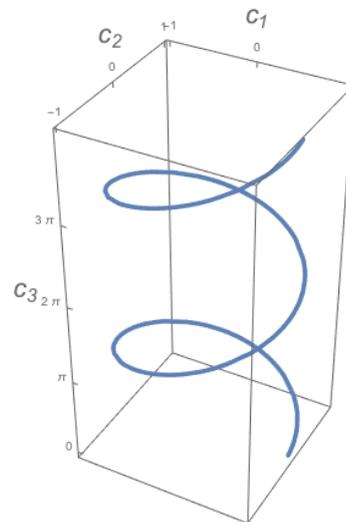
$$c : \mathbb{R} \rightarrow \mathbb{R}^3, \quad c(t) = \begin{pmatrix} R \cos t \\ R \sin t \\ bt \end{pmatrix}$$

die Ableitungen $c'(t) = \begin{pmatrix} -R \sin t \\ R \cos t \\ b \end{pmatrix}$, sodass der Betrag der Geschwindigkeit konstant wird: $\|c'(t)\| = \sqrt{R^2 + b^2}$.

Die Beschleunigung ist $c''(t) = \begin{pmatrix} -R \cos t \\ -R \sin t \\ 0 \end{pmatrix}$, also $\|c''(t)\| = R$ und $\langle c''(t), c'(t) \rangle = 0$. Damit ist die Krümmung

$$k(t) = \frac{\|c''(t)\|}{\|c'(t)\|^2} = \frac{R}{R^2 + b^2} > 0.$$

Die Krümmung der *Helix* (Spirale) c ist also kleiner als die Krümmung $1/R$ des Kreises mit gleichem Radius R . \diamond



Die *Differentialgeometrie* befaßt sich mit der Krümmung von Kurven und analogen Fragen für höherdimensionale Objekte wie Flächen im \mathbb{R}^3 . Ein einführendes Buch ist das von DO CARMO [Ca].

7 Gewöhnliche Differentialgleichungen

7.1 Definitionen und Beispiele

Differentialgleichungen sind so vielfältig wie die Naturvorgänge, die sie beschreiben. Wir beginnen mit (etwas informellen) Definitionen und einer Grobeinteilung:

7.1 Definition

- Eine **Differentialgleichung (DGL)** ist eine Gleichung, in der Ableitungen einer oder mehrerer Funktionen von einer oder mehreren Variablen auftreten. Die gesuchten Unbekannten sind hierbei die Funktionen.
- Hängen die Funktionen von nur **einer** Variablen ab, so heißt die DGL **gewöhnlich**, sonst **partiell**.
- Werden mehrere Funktionen gesucht, so spricht man von einem **Differentialgleichungssystem**, sonst von einer **Einzel-DGL**.

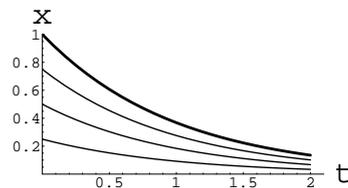
7.2 Beispiel 1. Für $c > 0$ beschreibt die gewöhnliche Einzel-DGL

$$\frac{dx}{dt}(t) = -c \cdot x(t)$$

z.B. radioaktiven Zerfall mit Stoffmenge x als Funktion der Zeit t und Zerfallskonstante c . Ist die Stoffmenge zur Zeit $t = 0$ gleich $x_0 \in \mathbb{R}$, dann ist

$$x(t) = x_0 e^{-ct} \quad (t \in \mathbb{R})$$

die eindeutige Lösung. Wir erhalten also eine einparametrische Schar von Lösungen, die linear vom Anfangswert x_0 abhängt:



2. Die Bahn eines geworfenen Körpers im konstanten Schwerfeld der Erde mit Erdbeschleunigung¹⁹ $g > 0$ wird unter Vernachlässigung der Luftreibung durch das gewöhnliche DGL-System,

$$\frac{d^2 x_1}{dt^2}(t) = 0 \quad , \quad \frac{d^2 x_2}{dt^2}(t) = -g$$

beschrieben. Dabei bezeichnet x_1 die Horizontalkomponente und x_2 die Vertikalkomponente des Ortes als Funktionen der Zeit t .

Für Anfangsort $x_0 = \begin{pmatrix} x_{1,0} \\ x_{2,0} \end{pmatrix}$ und Anfangsgeschwindigkeit $v_0 = \begin{pmatrix} v_{1,0} \\ v_{2,0} \end{pmatrix}$ ist die **Lösung:**

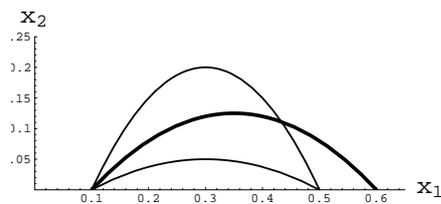
$$x_1(t) = x_{1,0} + v_{1,0}t \quad , \quad x_2(t) = x_{2,0} + v_{2,0}t - \frac{1}{2}gt^2 \quad , \quad (t \in \mathbb{R}).$$

¹⁹In Bodenhöhe ist $g = 9.81m/s^2$

Dies entspricht den Geschwindigkeiten

$$v_1(t) := \frac{d}{dt}x_1(t) = v_{1,0} \quad , \quad v_2(t) := \frac{d}{dt}x_2(t) = v_{2,0} - gt \quad , \quad (t \in \mathbb{R}).$$

Die Zeichnung zeigt verschiedene Wurfbahnen bei gleichem Anfangsort und Betrag der Anfangsgeschwindigkeit, aber unterschiedlicher Richtung der Anfangsgeschwindigkeit. Der Wurf mit Winkel $\pi/4$ führt dabei am weitesten.



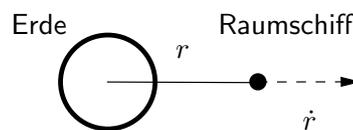
7.3 Definition Die Ordnung des höchsten in der DGL auftretenden Differenti-
alquotienten wird **Ordnung der DGL** genannt.

7.4 Beispiel 1. 7.2.1. ist von erster Ordnung

2. 7.2.2. ist von zweiter Ordnung

3. Die DGL $\frac{d^2 r}{dt^2}(t) = -\frac{M}{r^2(t)}$ ist von zweiter Ordnung. Sie beschreibt die

Bewegung eines sich radial vom Erdmittelpunkt wegbewegenden Raumschiffes. $M > 0$ ist die Erdmasse, r der Abstand des Raumschiffes vom Erdmittelpunkt und $\dot{r} = \frac{d}{dt}r$ die Radialgeschwindigkeit. \diamond



7.5 Definition

Eine gewöhnliches DGL-System für die Funktionen x_1, \dots, x_m heißt **linear**, wenn es die Form

$$\sum_{i=0}^n A^{(i)}(t)x^{(i)}(t) = b(t)$$

hat. Dabei bezeichnet $x^{(i)} := (\frac{d^i}{dt^i}x_1, \dots, \frac{d^i}{dt^i}x_m)^t$ den Vektor der i -ten Ableitungen; $A^{(i)}(t) \in \text{Mat}(m, \mathbb{R})$ und $b(t) \in \mathbb{R}^m$ sind vorgegebene Matrix- bzw. vektorwertige Funktionen.

Andernfalls heißt das DGL-System **nicht linear**.

Eine lineare DGL heißt **homogen**, wenn $b(t) = 0$ für alle t , sonst **inhomogen**.

Die Komponenten b_i von b heißen **Störfunktionen**.

Beispiel 7.4.1. ist linear homogen.
 Beispiel 7.4.2. ist linear inhomogen.
 Beispiel 7.4.3. ist nicht linear.

Ab jetzt werden viele Begriffe nur für Einzel-DGLn eingeführt. Das meiste überträgt sich aber auf DGL-Systeme.

7.6 Definition 1. Eine DGL heißt **implizit**, wenn sie die Form

$$F(t, x, x', \dots, x^{(n)}) = 0 \quad (7.1)$$

hat, **explizit**, wenn sie die Form

$$x^{(n)} = f(t, x, x', \dots, x^{(n-1)}) \quad (7.2)$$

hat.

2. Eine n -mal differenzierbare auf dem offenen Intervall I definierte Funktion $x : I \rightarrow \mathbb{R}$ heißt **explizite Lösung** der DGL (7.1) bzw. (7.2), wenn gilt:

$$F(t, x(t), x'(t), \dots, x^{(n)}(t)) = 0 \quad (t \in I)$$

$$\text{bzw. } x^{(n)}(t) = f(t, x(t), x'(t), \dots, x^{(n-1)}(t)) \quad (t \in I).$$

Beispiele 7.4.1.–3. waren explizite DGLn, für 1. und 2. wurden auch (die) expliziten Lösungen angegeben.

7.7 Beispiel 4. $y \frac{dy}{dx} + x = 0$ ist eine implizite nichtlineare DGL. Die Kreisgleichung $x^2 + y^2 = c > 0$ ist die allgemeine Lösung, aber in *impliziter* Form.

Explizite Lösung: $y(x) = \pm \sqrt{c - x^2}$ für $|x| < \sqrt{c}$, also

$$\frac{dy}{dx} = \mp \frac{x}{\sqrt{c - x^2}} = -\frac{x}{y}.$$

7.8 Definition

- Eine einzelne Lösung (ohne frei wählbare Konstanten) heißt **spezielle** oder **partikuläre Lösung**.
- Eine parameterabhängige Lösung einer DGL n -ter Ordnung heißt **allgemeine Lösung**, wenn sie n frei wählbare Konstanten enthält,
- eine parameterabhängige Lösung heißt **vollständig**, wenn alle speziellen Lösungen durch Wahl geeigneter Parameterwerte aus ihr hervorgehen.

- Eine nicht zu einer parameterabhängigen Lösung gehörende spezielle Lösung heißt **singulär**.

Beispiele 7.2.1. und 2.: Die allgemeinen = vollständigen Lösungen wurden angegeben. Eine partikuläre Lösung von 2. ist z.B. $x_1(t) = 0$, $x_2(t) = -\frac{1}{2}gt^2$.

Beachte: In 2. gab es vier Parameter $x_{1,0}, x_{2,0}, v_{1,0}, v_{2,0}$, denn es waren zwei DGLn zweiter Ordnung, $2 \times 2 = 4$.

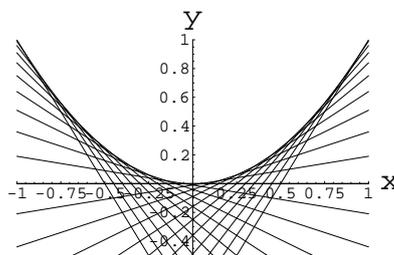
Beispiel 7.7.4.: Hier war $c \geq 0$ Parameter der allgemeinen = vollständigen Lösung.

7.9 Beispiel 5. $(y')^2 - 4xy' + 4y = 0$

ist eine implizite nichtlineare DGL erster Ordnung.

Die *allgemeine* Lösung: $y(x) = 2cx - c^2$, $c \in \mathbb{R}$ ist eine durch c parametrisierte Geradenschar.

Dies ist aber *nicht* die *vollständige* Lösung, denn es existiert noch die *singuläre* Lösung $y(x) = x^2$. Diese ist die Einhüllende der Geradenschar, siehe Abbildung.



- Frage:**
- Wie findet man Lösungen?
 - Woher weiß man, dass man alle gefunden hat?

Diese Frage beschäftigt seit der Zeit Newtons viele Mathematiker (und uns in den nächsten Wochen²⁰).

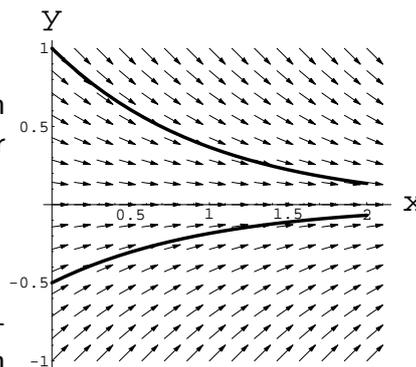
Wir betrachten zunächst die expliziten Einzel-DGLn erster Ordnung

$$y' = f(x, y) \quad (x, y) \in U \subset \mathbb{R}^2, U \text{ offen.}$$

Geometrische Interpretation: Zeichnet man an jedem Punkt $(x, y) \in U$ eine Gerade der Steigung $f(x, y)$, dann ist

$$\text{graph}(\tilde{y}) = \{(x, \tilde{y}(x))\}$$

jeder speziellen Lösung $\tilde{y} \equiv \tilde{y}(x)$ der DGL eine Kurve in U , die überall *tangential* an den lokalen Geraden ist.



Beispiel $y' = -cy$, also $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = -cy$, siehe Abbildung.

²⁰Numerische Methoden zur Lösung von DGLn werden z.B. in [DB] behandelt

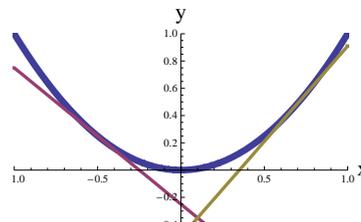
Um also die durch den Punkt (x, y) gehende spezielle Lösung zu finden, bewegt man sich, von (x, y) ausgehend, tangential zum Richtungsfeld.

Vorsicht: Woher wissen wir überhaupt, dass durch jeden Punkt $(x, y) \in U \subset \mathbb{R}^2$ nur *eine* Lösungskurve geht?

7.10 Beispiel (Gegenbeispiele) 1. Implizite DGL, Beispiel 7.7.4.

$$(y')^2 - 4xy' + 4y = 0 \tag{7.3}$$

Hier gehen durch jeden Punkt (x_0, y_0) unterhalb des Graphen der Parabel $y = x^2$ zwei Lösungskurven, d.h. an die Parabel tangentielle Geraden. Deren Steigungen entsprechen den zwei Lösungen der quadratischen Gleichung (7.3) für y' am Punkt (x_0, y_0) .



2. Explizite DGL, f nicht lipschitz-stetig

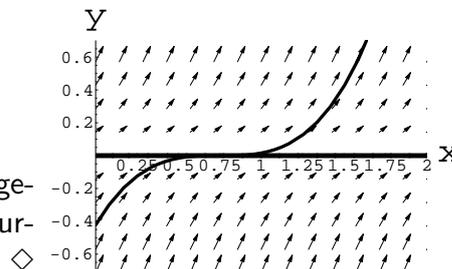
$$y' = f(y) \text{ mit } f(y) := 3\sqrt[3]{y^2}$$

Allgemeine Lösung:

$$y(x) = (x - c)^3, \quad c \in \mathbb{R}$$

Singuläre Lösung: $y(x) = 0$.

Durch jeden Punkt auf der x -Achse gehen damit mindestens zwei Lösungskurven!



Im zweiten Beispiel fällt auf, dass die Funktion f bei 0 nicht lipschitz-stetig, geschweige denn differenzierbar war.

Nach dieser informellen Übersicht über die bei gewöhnlichen DGLn auftretenden Phänomene zeigen wir nun mathematisch rigoros die (lokale) Existenz und Eindeutigkeit der Lösung genügend regulärer expliziter gewöhnlicher DGLn 1. Ordnung. Später werden wir sehen, dass damit auch die gleiche Frage für explizite DGLn höherer Ordnung beantwortet wird.

7.2 Der Banachscher Fixpunktsatz

7.11 Definition

- Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen (X, d_X) und (Y, d_Y) heißt **lipschitz-stetig**, wenn ein $L > 0$ (genannt **Lipschitzkonstante**) existiert mit

$$d_Y(f(x_1), f(x_2)) \leq L d_X(x_1, x_2) \quad (x_1, x_2 \in X).$$

- Eine lipschitz-stetige Abbildung $f : X \rightarrow X$ auf einem vollständigen metrischen Raum (X, d) heißt **kontrahierend**, wenn sie eine Lipschitzkonstante $L < 1$ besitzt. Es sei X eine Menge. Dann heißt $x \in X$ **Fixpunkt** der Abbildung $f : X \rightarrow X$, wenn $f(x) = x$ gilt.

7.12 Satz Lipschitz-stetige Abbildungen sind stetig.

Beweis: Sei $x \in X$ und V Umgebung von $f(x)$. Dann gibt es $\varepsilon > 0$ mit $U_\varepsilon(f(x)) \subseteq V$. Setze $\delta := \varepsilon/L$. Dann ist $f(U_\delta(x)) \subseteq U_\varepsilon(f(x))$. \square

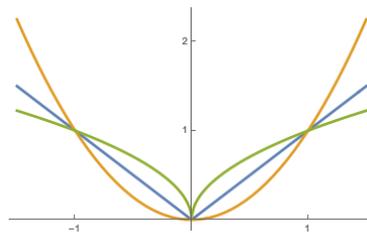
7.13 Bemerkung Die Umkehrung gilt nicht, wie die folgenden stetigen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ zeigen (mit Metrik $d(x_1, x_2) = |x_1 - x_2|$ und $x_1 := 0, x_2 = x$):

- $f(x) := x^2$. Dann ist

$$d(f(x_1), f(x_2)) = x^2 \leq Ld(x_1, x_2) = L|x|$$

nur für $|x| \leq L$.

Da f stetig differenzierbar ist, ist die Funktion zwar wie man sagt *lokal*, d.h. eingeschränkt auf eine geeignete Umgebung eines beliebigen Punktes, aber nicht *global* lipschitz-stetig.



$|x|$, und nicht lokal bzw. nicht global Lipschitz-stetige Funktionen $x \mapsto \sqrt{|x|}$ bzw. $x \mapsto x^2$

- $f(x) := \sqrt{|x|}$. Dann ist $\sqrt{|x|} \leq L|x|$ nur für $|x| \geq L^{-2}$. f ist also bei 0 noch nicht einmal lokal lipschitz-stetig. Andererseits ist f sogar gleichmäßig stetig²¹, also insbesondere stetig. \diamond

7.14 Satz (Banachscher Fixpunktsatz) Eine kontrahierende Abbildung $f : X \rightarrow X$ auf einem vollständigen metrischen Raum (X, d) mit Lipschitzkonstante $\theta < 1$ besitzt genau einen Fixpunkt x^* , und für die m -te Iterierte $x_m := f(x_{m-1})$ von $x_0 \in X$ gilt

$$d(x_m, x^*) \leq d(x_1, x_0) \frac{\theta^m}{1-\theta} \quad (m \in \mathbb{N}). \quad (7.4)$$

²¹**Def.:** Eine Abbildung $f : M \rightarrow N$ zwischen den metrischen Räumen (M, d_M) und (N, d_N) heißt **gleichmäßig stetig**, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt mit $d_N(f(x), f(y)) \leq \epsilon$ für alle $x, y \in M$ mit $d_M(x, y) \leq \delta$. Siehe *Analysis I*, Kapitel 9.3

Beweis: • Für $n > 0$ gilt $d(x_{n+1}, x_n) \leq \theta d(x_n, x_{n-1}) \leq \dots \leq \theta^n d(x_1, x_0)$, also für $n > m$

$$d(x_n, x_m) \leq \sum_{j=m}^{n-1} d(x_{j+1}, x_j) \leq \left(\sum_{j=m}^{n-1} \theta^j \right) d(x_1, x_0) \leq \frac{\theta^m}{1-\theta} d(x_1, x_0),$$

sodass $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge bildet, und wegen Vollständigkeit von (X, d) existiert

$$x^* := \lim_{n \rightarrow \infty} x_n.$$

- Damit ist $d(x^*, x_m) = \lim_{n \rightarrow \infty} (d(x^*, x_n) + d(x_n, x_m)) \leq \frac{\theta^m}{1-\theta} d(x_1, x_0)$.
- Wegen der Stetigkeit von f gilt $f(x^*) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x^*$.
- Gilt für $x \in X$ ebenfalls $f(x) = x$, dann ist $d(x, x^*) = d(f(x), f(x^*)) \leq \theta d(x, x^*)$, also $x = x^*$. Damit ist x^* der einzige Fixpunkt. \square

Wir können hier also mit (7.4) die Zahl der Iterationen, die zum Erreichen einer gewünschten Genauigkeit hinreichend sind, aus $d(x_1, x_0)$ und θ bestimmen.

7.15 Bemerkungen

1. Ist $X \subset \mathbb{R}$ ein abgeschlossenes Intervall und $f : X \rightarrow X$ stetig differenzierbar mit $|f'(x)| \leq \theta < 1$ für alle $x \in X$, dann ist f kontrahierend mit Konstante θ , denn

$$|f(y) - f(x)| = \left| \int_x^y f'(z) dz \right| \leq \left| \int_x^y |f'(z)| dz \right| \leq \left| \int_x^y \theta dz \right| = \theta |y - x|.$$

2. Oft muss man das Problem etwas umformulieren, um die Kontraktionseigenschaft zu bekommen. Beispielsweise kann man den Definitionsbereich verkleinern oder die Gleichung umformen.

So werden wir die eindeutige Existenz von Lösungen gewöhnlicher *Differentialgleichungen* beweisen, indem wir diese als Fixpunkte eines kontrahierenden *Integraloperators* charakterisieren. \diamond

7.16 Beispiel (Berechnung von $\sqrt{2}$)

- Wir suchen also eine Nullstelle von $f(x) := x^2 - 2$, wissen aber schon, dass zwei Nullstellen existieren, also zwei Fixpunkte von $g(x) := x^2 - 2 + x$. Damit können wir den Banachschen Fixpunktsatz nicht direkt auf $g : \mathbb{R} \rightarrow \mathbb{R}$ anwenden.
- Restringieren wir g auf $[1, 2]$, dann ist auf diesem Intervall ist $g' \geq 3$, g also nicht kontrahierend.

Man könnte nun statt f die Funktion cf benutzen, die für $c \in \mathbb{R} \setminus \{0\}$ die gleichen Nullstellen wie f besitzt, aber für $c \in (-\frac{1}{2}, 0)$ eine kontrahierende Abbildung $x \mapsto cf(x) + x$ liefert. Wir gehen aber einen anderen Weg, um die

Flexibilität des Fixpunktverfahrens zu zeigen.

- Nach der vietaschen Wurzelformel ist die Umkehrfunktion von $g \upharpoonright_{[1, \infty)}$ gleich

$$h : [0, \infty) \rightarrow [0, \infty) \quad , \quad h(y) = -\frac{1}{2} + \sqrt{y + 9/4},$$

und $h'(y) = \frac{1}{\sqrt{4y+9}}$, sodass $|h'(y)| \leq \frac{1}{3}$ ist für $y \geq 0$. Damit ist h eine kontra-

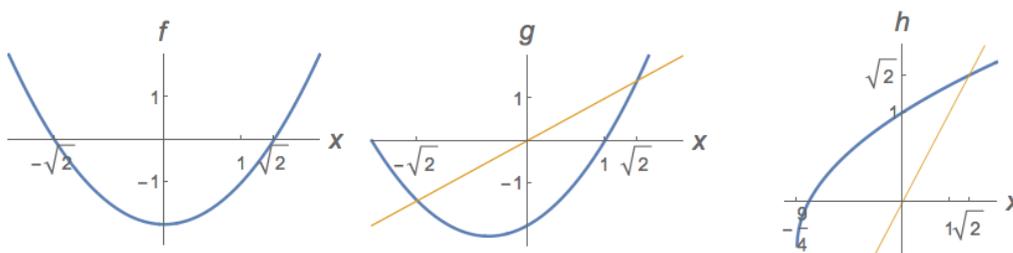


Abbildung 7.1: $\sqrt{2}$ als: Nullstelle von $f(x) = x^2 - 2$ (links), Fixpunkt von $g(x) = f(x) + x$ (Mitte) bzw. Fixpunkt von $h(y) = -\frac{1}{2} + \sqrt{y + 9/4}$ (rechts)

hierende Abbildung auf dem vollständigen metrischen Raum $[0, \infty)$. Außerdem besitzt sie als Umkehrfunktion von $g \upharpoonright_{[1, \infty)}$ den Fixpunkt $\sqrt{2}$.

Mit $y_0 := 0$ ist $y_1 = h(y_0) = 1$, also gemäß (7.4) die Fehlerschranke

$$|y_m - \sqrt{2}| \leq \frac{3}{2} \cdot 3^{-m}.$$

Numerisch ergibt sich gerundet:

m	y_m	Fehler	Fehlerschranke
0	0	1.41421	1.5
1	<u>1</u>	0.41421	0.5
2	<u>1.3027</u>	0.11151	0.16666
3	<u>1.3848</u>	0.02941	0.05555
4	<u>1.4065</u>	0.00771	0.01851
5	<u>1.4122</u>	0.00201	0.00617

Da der Abstand (7.4) vom Fixpunkt exponentiell in der Zahl m der Iterationen schrumpft, wächst die Zahl der (unterstrichenen) richtigen Ziffern ungefähr linear mit m . \diamond

Diese *lineare Konvergenzgeschwindigkeit* ist typisch für kontrahierende Abbildungen und für viele Zwecke ausreichend.

7.3 Lokale Existenz und Eindeutigkeit der Lösung

Wir werden nun sehen, dass bei etwas mehr Regularität von f die DGL lokal eindeutig lösbar ist. Dazu schauen wir uns aber gleich die n -dimensionale Situation an:

7.17 Definition • Ist der erweiterte Phasenraum $U \subseteq \mathbb{R}_t \times \mathbb{R}_x^n$ offen, und das zeitabhängige Vektorfeld $f : U \rightarrow \mathbb{R}^n$ stetig, dann heißt die Gleichung

$$\dot{x} = f(t, x)$$

nichtautonome oder explizit zeitabhängige DGL.

- Ist speziell $U = \mathbb{R}_t \times \tilde{U}$ mit Phasenraum $\tilde{U} \subseteq \mathbb{R}_x^n$ offen und f von der Form $f(t, x) = \tilde{f}(x)$, dann heißt die DGL **autonom** oder **dynamisches System**.
- Eine differenzierbare Funktion $\varphi : I \rightarrow \mathbb{R}_x^n$, $I \subseteq \mathbb{R}_t$ Intervall heißt **Lösung** der DGL, wenn $\text{graph}(\varphi) \subset U$ und

$$\frac{d\varphi}{dt} \Big|_{t=\tau} = f(\tau, \varphi(\tau)) \quad (\tau \in I).$$

- $\varphi : I \rightarrow \mathbb{R}_x^n$ genügt der **Anfangsbedingung** (t_0, x_0) , wenn $t_0 \in I$, $(t_0, x_0) \in U$ und $\varphi(t_0) = x_0$ gilt. φ **löst das Anfangswertproblem (AWP)**, wenn zusätzlich φ eine Lösung der DGL ist.
- Das zeitabhängige Vektorfeld $f : U \rightarrow \mathbb{R}^n$ genügt

- **global einer Lipschitz-Bedingung** mit Konstante L , wenn

$$\|f(t, x_1) - f(t, x_2)\| \leq L\|x_1 - x_2\| \quad ((t, x_i) \in U)$$

- und **(lokal) einer Lipschitz-Bedingung**, wenn jeder Punkt (τ, x) aus U eine Umgebung $V \subseteq U$ besitzt, sodass für eine Konstante $L = L(\tau, x)$

$$\|f(t, x_1) - f(t, x_2)\| \leq L\|x_1 - x_2\| \quad ((t, x_i) \in V).$$

7.18 Bemerkungen (Lipschitz-Bedingung)

1. Man beachte, dass die Lipschitz-Stetigkeit nur bezüglich der x -Variablen gefragt ist.
2. Ist $f \in C^1(U, \mathbb{R}^n)$, dann ist in einem konvexen kompakten Gebiet $K \subset U$ nach Satz 10.6 die lokale Lipschitz-Bedingung erfüllt, mit der Konstante

$$L := \max_{(t,x) \in K} \|D_x f(t, x)\|. \quad \diamond$$

Wir werden in Satz 7.20 die eindeutigen Lösungen des Anfangswertproblems als Fixpunkte einer kontrahierenden Abbildung auf einem Raum stetiger Funktionen finden. Damit wir sicher sein können, dass der Fixpunkt überhaupt existiert, müssen wir gemäß dem Banachschen Fixpunktsatz (Satz 7.14) zunächst kontrollieren, dass der benutzte Funktionenraum ein vollständiger metrischer Raum ist.

7.19 Satz *Es sei $V \subseteq \mathbb{R}^n$ abgeschlossen und M der Raum der Kurven in V , also $M := C([a, b], V)$. Für $f, g \in M$ setzen wir $d(f, g) := \sup_{t \in [a, b]} \|f(t) - g(t)\|$.*

Dann ist (M, d) ein vollständiger metrischer Raum.

Beweis: • Auf $[a, b]$ nimmt die stetige Funktion $t \mapsto \|f(t) - g(t)\|$ ihr Maximum an. Also ist $d(f, g) < \infty$. Andererseits ist $d(f, g) = 0$ genau für $f = g$. Ebenso gilt $d(g, f) = d(f, g)$ und

$$\begin{aligned} d(f, h) &= \sup_{t \in [a, b]} \|(f(t) - g(t)) + (g(t) - h(t))\| \\ &\leq \sup_{t \in [a, b]} (\|f(t) - g(t)\| + \|g(t) - h(t)\|) \\ &\leq \sup_{t \in [a, b]} \|f(t) - g(t)\| + \sup_{t \in [a, b]} \|g(t) - h(t)\| \\ &= d(f, g) + d(g, h). \end{aligned}$$

Damit ist (M, d) ein metrischer Raum.

• Weiter ist für alle $t \in [a, b]$ und eine Cauchy-Folge $(f_m)_{m \in \mathbb{N}}$ von Kurven $f_m \in M$ auch die Folge $(f_m(t))_{m \in \mathbb{N}}$ von Punkten aus V eine Cauchy-Folge, denn es gilt $\|f_m(t) - f_n(t)\| \leq d(f_m, f_n)$. Da \mathbb{R}^n vollständig ist und $V \subseteq \mathbb{R}^n$ abgeschlossen ist, existiert für alle $t \in [a, b]$ der Punkt

$$f(t) := \lim_{m \rightarrow \infty} f_m(t) \in \mathbb{R}^n$$

und ist sogar aus V .

• Ist damit schon der Limes $f \in M$? Das ist noch nicht klar, denn f könnte unstetig sein. Wir verwenden zum Nachweis der Stetigkeit ein sog. $\varepsilon/3$ -Argument:

Für jedes $\varepsilon > 0$ und ein geeignetes $N \in \mathbb{N}$ gilt ja für die Cauchy-Folge $(f_m)_{m \in \mathbb{N}}$ insbesondere

$$d(f_m, f_N) < \varepsilon/3 \quad (m \geq N),$$

also auch für alle $t \in [a, b]$

$$\|f(t) - f_N(t)\| = \lim_{m \rightarrow \infty} \|f_m(t) - f_N(t)\| \leq \varepsilon/3.$$

Da f_N stetig ist, gibt es für jedes $s \in [a, b]$ ein $\delta > 0$ mit

$$\|f_N(t) - f_N(s)\| < \varepsilon/3 \quad (t : |t - s| < \delta).$$

Damit ist nach der Dreiecksungleichung für diese t

$$\begin{aligned} \|f(t) - f(s)\| &\leq \|f(t) - f_N(t)\| + \|f_N(t) - f_N(s)\| + \|f_N(s) - f(s)\| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ &= \varepsilon. \end{aligned}$$

Also ist auch f stetig, d.h. $f \in M$.

- Damit gilt auch $f = \lim_{m \rightarrow \infty} f_m$ im metrischen Raum (M, d) , denn

$$\begin{aligned} d(f, f_N) &= \sup_{t \in [a, b]} \|f(t) - f_N(t)\| = \sup_t \lim_{m \rightarrow \infty} \|f_m(t) - f_N(t)\| \\ &\leq \sup_t \sup_{m \geq N} \|f_m(t) - f_N(t)\| = \sup_{m \geq N} d(f_m, f_N), \end{aligned}$$

und da die f_n eine Cauchy-Folge bilden, geht letzterer Abstand für $N \rightarrow \infty$ gegen Null. Wir haben damit nachgewiesen, dass jede Cauchy-Folge $(f_m)_{m \in \mathbb{N}}$ konvergiert, (M, d) also vollständig ist. \square

7.20 Satz (Picard–Lindelöf) *Das zeitabhängige Vektorfeld $f : U \rightarrow \mathbb{R}^n$ auf der offenen Menge $U \subseteq \mathbb{R}_t \times \mathbb{R}_x^n$ genüge einer Lipschitz-Bedingung auf U . Dann existiert für $(t_0, x_0) \in U$ ein $\varepsilon > 0$, sodass das Anfangswertproblem*

$$\dot{x} = f(t, x) \quad , \quad x(t_0) = x_0$$

eine eindeutige Lösung $\varphi : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$ besitzt.

Beweis: Wir bezeichnen das Zeitintervall mit $I := [t_0 - \varepsilon, t_0 + \varepsilon]$.

- Existiert eine solche Lösung, dann muss sie die Integralgleichung

$$\varphi(t) = x_0 + \int_{t_0}^t f(s, \varphi(s)) ds \quad (t \in I) \quad (7.5)$$

erfüllen, wie man durch Differentiation bzw. Einsetzen von t_0 feststellt. Andererseits ist nach dem Hauptsatz der Differential- und Integralrechnung (Satz 12.15 von [An1]) jede stetige Lösung von (7.5) auch schon differenzierbar und damit eine Lösung des Anfangswertproblems

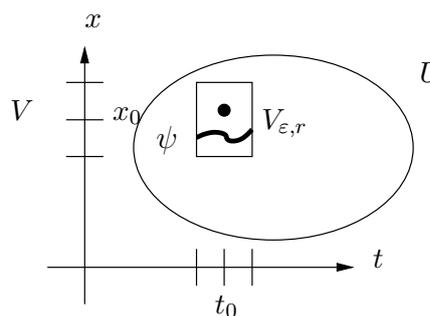
$$\dot{x} = f(t, x) \quad , \quad x(t_0) = x_0.$$

- Es soll nun die Lösung φ als Fixpunkt einer Abbildung A aufgefunden werden. A wird stetige Kurven im Phasenraum in solche abbilden.

Um den Definitionsbereich von A günstig zu wählen, soll der Phasenraumbereich die Vollkugel $V := \overline{U_r(x_0)}$ sein. Wir setzen

$$V_{\varepsilon, r} := I \times V,$$

und wählen r so klein, das $V_{\varepsilon, r} \subseteq U$.



Weiter sei $L > 0$ Lipschitzkonstante von $f|_{V_{\varepsilon, r}}$ und

$$N := \max_{(t, x) \in V_{\varepsilon, r}} \|f(t, x)\| \quad \text{und} \quad \varepsilon := \min \left(r, \frac{r}{N}, \frac{1}{2L} \right). \quad (7.6)$$

Damit ist insbesondere $V_{\varepsilon, r} \subseteq V_{r, r} \subseteq U$.

$$M := C(I, V) \subseteq C(I, \mathbb{R}^n)$$

bezeichne wieder den metrischen Raum der stetigen Funktionen $\psi : I \rightarrow V$ mit Supremumsmetrik

$$d(\psi, \varphi) := \sup_{t \in I} \|\psi(t) - \varphi(t)\|.$$

Nach Satz 7.19 ist (M, d) ein vollständiger metrischer Raum (denn $V \subseteq \mathbb{R}^n$ ist abgeschlossen).

- Wir führen durch

$$(A\psi)(t) := x_0 + \int_{t_0}^t f(s, \psi(s)) ds \quad (t \in I)$$

eine Abbildung $A : M \rightarrow C(I, \mathbb{R}^n)$ ein, von der wir zunächst zeigen wollen, dass ihr Bild in M bleibt. Dazu stellen wir fest, dass gemäß Definition (7.6)

$$\begin{aligned} \left\| \int_{t_0}^t f(s, \psi(s)) ds \right\| &\leq \left| \int_{t_0}^t \|f(s, \psi(s))\| ds \right| \\ &\leq \left| \int_{t_0}^t \max_{(\tilde{t}, x) \in V} \|f(\tilde{t}, x)\| ds \right| \leq |t - t_0|N \leq \varepsilon N \leq r, \end{aligned}$$

sodass $(A\psi)(t) \in V$ für $t \in I$ ist.

- Jetzt müssen wir nur noch beweisen, dass

$$A : M \rightarrow M,$$

die so genannte *Picard-Abbildung*, kontrahierend ist, dass also für ein geeignetes $0 < \theta < 1$ gilt

$$d(A\varphi, A\psi) \leq \theta d(\varphi, \psi) \quad (\varphi, \psi \in M).$$

Tatsächlich ergibt sich aus der Definition (7.6) von ε für $t \in I$

$$\begin{aligned} d(A\varphi, A\psi) &= \sup_{t \in I} \|A\psi(t) - A\varphi(t)\| \\ &= \sup_{t \in I} \left\| \int_{t_0}^t [f(s, \psi(s)) - f(s, \varphi(s))] ds \right\| \\ &\leq \varepsilon \cdot \sup_{s \in I} \|f(s, \psi(s)) - f(s, \varphi(s))\| \\ &\leq \varepsilon L \sup_{s \in I} \|\psi(s) - \varphi(s)\| = \varepsilon L d(\psi, \varphi) \leq \frac{1}{2} d(\psi, \varphi). \end{aligned}$$

A ist damit eine kontrahierende Abbildung auf dem vollständigen metrischen Raum M , besitzt also nach dem Banachschen Fixpunktsatz (Satz 7.14) einen eindeutigen Fixpunkt $\varphi \in M$. Diese Funktion φ erfüllt also die Integralgleichung (7.5) und löst damit das Anfangswertproblem. \square

Die Picard-Iteration, die hier als technisches Beweismittel verwandt wurde, kann auch zur Lösung der DGL verwandt werden:

7.21 Beispiele (Picard-Iteration)

1. Wir approximieren die Lösung des Anfangswertproblems $\dot{x} = x$, $x(0) = x_0$ durch

$$x_0(t) := x_0 \quad (t \in \mathbb{R}) \quad \text{und} \quad x_{i+1}(t) := x_0 + \int_0^t x_i(s) ds,$$

also

$$\begin{aligned} x_1(t) &= x_0(1+t) \\ x_2(t) &= x_0\left(1+t+\frac{t^2}{2}\right) \\ &\vdots \\ x_n(t) &= x_0 \cdot \sum_{i=0}^n \frac{t^i}{i!}. \end{aligned}$$

Da $x(t) = x_0 \sum_{i=0}^{\infty} \frac{t^i}{i!} = x_0 \cdot e^t$, konvergiert für alle $t \in \mathbb{R}$ die n -te iterierte $x_n(t)$ gegen die Lösung $x(t)$, und zwar gleichmäßig auf jedem kompakten Zeitintervall (aber nicht gleichmäßig auf \mathbb{R}).

2. $\dot{x} = 1 + x^2$, $x_0 = 0$

Für $r > 0$ ist $N = \max_{|x| \leq r} \|f(x)\| = 1 + r^2$, und die Lipschitzkonstante $L = \max_{|x| \leq r} \|f'(x)\| = 2r$, also ist gemäß Definition (7.6) $\varepsilon = \min\left(\frac{r}{1+r^2}, \frac{1}{4r}\right)$.

ε wird maximal für $r = \frac{1}{\sqrt{3}}$, d.h. $\varepsilon = \frac{\sqrt{3}}{4}$.

Für Zeiten $|t| < \frac{\sqrt{3}}{4}$ können wir also Konvergenz garantieren.

Picard-Iteration mit $x_0(t) := x_0$ ergibt

$$\begin{aligned} x_0(t) &= 0 \\ x_1(t) &= (Ax_0)(t) = \int_0^t (1 + x_0^2(s)) ds = t \\ x_2(t) &= \int_0^t (1 + s^2) ds = t + t^3/3 \\ x_3(t) &= \int_0^t [1 + (s + s^3/3)^2] ds = t + \frac{t^3}{3} + \frac{2}{15}t^5 + \frac{1}{63}t^7, \\ &\text{etc.} \end{aligned}$$

Die Funktionenfolge konvergiert gegen $\tan(t)$, die Lösung des Anfangswertproblems. Konvergenz haben wir sogar auf dem offenen Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$. \diamond

7.4 Globale Existenz und Eindeutigkeit der Lösung

An Beispiel 7.21.2 sehen wir, dass die Lösung nicht für alle Zeiten existiert, da $f(x) = 1 + x^2$ mit $|x|$ sehr stark anwächst, sodass man in der Zeit $\frac{\pi}{2}$ nach ∞ geführt wird. Dies steht nicht im Widerspruch zur lokalen Lipschitz-Stetigkeit von f .

Dagegen ist in Beispiel 7.21.1 $f(x) = x$ sogar global lipschitz-stetig, und die Lösung existiert für alle Zeiten $t \in \mathbb{R}$. Dies ist ganz allgemein so:

7.22 Satz Ist das (zeitunabhängige) Vektorfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ lipschitz-stetig, dann existiert eine eindeutige Lösung $\varphi : \mathbb{R} \rightarrow \mathbb{R}^n$ des Anfangswertproblems

$$\dot{x} = f(x) \quad , \quad x(0) = x_0.$$

Beweis: Es sei $L \geq 1$ eine Lipschitzkonstante, es gelte also

$$\|f(x_1) - f(x_2)\| \leq L\|x_1 - x_2\| \quad (x_1, x_2 \in \mathbb{R}^n).$$

Wir wählen $r := \|f(x_0)\| + \frac{1}{2L}$, sodass $N = \max_{x \in \overline{U_r(x_0)}} \|f(x)\|$

$$\begin{aligned} N &\leq \|f(x_0)\| + \max_{x \in \overline{U_r(x_0)}} \|f(x) - f(x_0)\| \\ &\leq \|f(x_0)\| + L \max_{x \in \overline{U_r(x_0)}} \|x - x_0\| = r + Lr \end{aligned}$$

ist, und

$$\varepsilon = \min\left(r, \frac{r}{N}, \frac{1}{2L}\right) \geq \min\left(\frac{1}{2L} + \|f(x_0)\|, \frac{1}{1+L}, \frac{1}{2L}\right) = \frac{1}{2L}$$

unabhängig von x_0 wird.

Damit können wir die lokale Lösung $\varphi_0 : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ des Anfangswertproblems $\dot{x} = f(x)$ mit $\varphi_0(0) := x_0$ finden.

$\varphi_0(\varepsilon/2)$ wird der neue Anfangswert für Zeit $t_1 = \frac{\varepsilon}{2}$, und wir erhalten eine Lösung $\varphi_1 : (-\frac{\varepsilon}{2}, \frac{3}{2}\varepsilon) \rightarrow \mathbb{R}^n$, mit

$$\varphi_0(t) = \varphi_1(t) \quad \text{für } t \in (-\varepsilon, \varepsilon) \cap (-\frac{\varepsilon}{2}, \frac{3}{2}\varepsilon) = (-\frac{\varepsilon}{2}, \varepsilon).$$

Durch Zusammensetzung der Lösungen

$$\varphi_k : \left(\varepsilon\left(\frac{k}{2} - 1\right), \varepsilon\left(\frac{k}{2} + 1\right)\right) \rightarrow \mathbb{R}^n \quad \text{mit} \quad \varphi_k\left(\frac{k}{2}\varepsilon\right) := \varphi_{k-1}\left(\frac{k}{2}\varepsilon\right)$$

erhalten wir eine Lösung für alle Zeiten $t \in \mathbb{R}$. □

7.23 Bemerkung (Existenz und Eindeutigkeit für lineare DGLn)

Insbesondere folgt, dass für alle linearen Differentialgleichungen das AWP der Form

$$\dot{x} = Ax \quad , \quad x(0) = x_0$$

für alle Zeiten eindeutig lösbar ist, denn das Vektorfeld $f(x) = Ax$ ist lipschitz-stetig mit Konstante $L = \|A\|$, der Matrixnorm von $A \in \text{Mat}(m, \mathbb{R})$ (wie wir gesehen haben, ist diese Lösung $x(t) = \exp(At)x_0$). ◇

Auch die Differentialgleichungen höherer als erster Ordnung lassen sich mit den beschriebenen Methoden behandeln, indem man aus einer expliziten DGL n -ter Ordnung ein DGL-System von n DGLn erster Ordnung macht.

7.24 Satz Die DGL der Ordnung $n > 1$

$$\frac{d^n x}{dt^n} = F\left(t, x, \frac{dx}{dt}, \dots, \frac{d^{n-1}x}{dt^{n-1}}\right) \quad (7.7)$$

mit $F \in C^1(\mathbb{R}^{n+1})$ ist zum DGL-System

$$\frac{dx}{dt} = f(t, x) \quad \text{mit} \quad f(t, x) := \begin{pmatrix} x_2 \\ \vdots \\ x_n \\ F(t, x_1, \dots, x_n) \end{pmatrix} \quad (7.8)$$

im folgenden Sinn äquivalent:

- Ist $\varphi : I \rightarrow \mathbb{R}$ Lösung von (7.7), dann ist $\psi := \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix} : I \rightarrow \mathbb{R}^n$ Lösung von (7.8).
- Ist umgekehrt $\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix}$ Lösung von (7.8), dann ist ψ_1 Lösung von (7.7).

Beweis: Nach Definition ist φ n -mal differenzierbar, also $\psi : I \rightarrow \mathbb{R}^n$ differenzierbar, und $\psi_k = \frac{d}{dt} \psi_{k-1} \quad (k = 2, \dots, n)$,

$$\frac{d}{dt} \psi_n = \frac{d^n \varphi}{dt^n} = F_n(t, \varphi, \varphi', \dots, \varphi^{(n-1)}) = F_n(t, \psi_1, \dots, \psi_n).$$

Die Argumentation lässt sich umkehren. □

7.25 Beispiel Das himmelsmechanische 2-Körper-Problem beschreibt die z.B. die Bewegung von Erde und Sonne um ihren gemeinsamen Schwerpunkt.

Nach einer Transformation kommt man zum sog. 1-Zentren-Problem der Bewegung einer Punktmasse am Ort $x \in \mathbb{R}^3 \setminus \{0\}$ im Schwerfeld eines im Koordinatenursprung befindlichen Himmelskörpers. Nach dem Gravitationsgesetz ist die Anziehungskraft proportional zu $-\frac{x}{|x|^3}$, also $\ddot{x} = -\gamma \frac{x}{\|x\|^3}$.

Dieses DGL-System zweiter Ordnung lässt sich auf die DGL erster Ordnung

$$\dot{z} = f(z) \quad \text{mit} \quad z = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \text{und} \quad f(z) := \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ -\gamma \frac{x_1}{\|x\|^3} \\ -\gamma \frac{x_2}{\|x\|^3} \\ -\gamma \frac{x_3}{\|x\|^3} \end{pmatrix}$$

umformen. Damit können wir es mit Picard-Iteration oder einem anderen Verfahren für kleine Zeiten lösen. ◇

Wir können auch explizite zeitabhängige DGLn auf autonome zurückführen. Statt der DGL $\dot{x} = f(t, x)$ mit $f : U \rightarrow \mathbb{R}^n$, $U \subseteq \mathbb{R}_t \times \mathbb{R}^n$ offen betrachten wir dazu das autonome DGL-System

$$\dot{y} = g(y) \quad \text{mit} \quad g : U \rightarrow \mathbb{R}^{n+1}, \quad y := \begin{pmatrix} s \\ x \end{pmatrix}, \quad g(y) := \begin{pmatrix} 1 \\ f(y) \end{pmatrix}. \quad (7.9)$$

Wir erhöhen also die Dimension des Phasenraumes um eins, indem wir den Zeitparameter s zum Phasenraumpunkt x hinzufügen. Ausgeschrieben hat (7.9) die Form

$$\frac{d}{dt}s = 1, \quad \frac{d}{dt}x = f(s, x). \quad (7.10)$$

Ist nun $\psi : I \rightarrow U$ eine Lösung von (7.9), dann gilt mit $\psi(t) = \begin{pmatrix} s(t) \\ x(t) \end{pmatrix}$ $s(t) = s(0) + t$, bis auf eine zu wählende additive Konstante ist also die Phasenraumkoordinate s gleich der Zeit t . Die Lösung $\psi = \begin{pmatrix} s \\ x \end{pmatrix}$ des AWP $\dot{y} = g(y)$, $\psi(0) = \begin{pmatrix} t_0 \\ x_0 \end{pmatrix}$ ergibt daher eine Lösung x des AWP $\dot{x} = f(t, x)$, $x(t_0) = x_0$. Man setzt einfach $x(t) := \tilde{x}(t - t_0)$.

Umgekehrt kann man aus einer Lösung des AWP $\dot{x} = f(t, x)$, $x(t_0) = x_0$ durch Ergänzung eine Lösung des AWP von (7.9) konstruieren.

Ist $f : U \rightarrow \mathbb{R}^n$ (in allen Argumenten) lipschitz-stetig, dann auch g . Damit überträgt sich der Existenz- und Eindeigkeitssatz (man beachte aber, dass wir für nicht autonome DGLn keine Lipschitz-Stetigkeit bezüglich t forderten).

8 Differentialrechnung mehrerer Variablen

Im letzten Semester wurde die Differential- und Integralrechnung reeller Funktionen $f : U \rightarrow \mathbb{R}$ (mit $U \subset \mathbb{R}$ offen) entwickelt. Die Differentialrechnung soll nun auf Funktionen $f : U \rightarrow \mathbb{R}^n$ mit $U \subset \mathbb{R}^m$ offen ausgeweitet werden.

Wir haben schon solche Funktionen kennen gelernt:

8.1 Beispiel (Abbildungen $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$)

1. $U \subset \mathbb{R}$ Intervall, $f : U \rightarrow \mathbb{R}^n$ stetig. Dies waren die parametrisierten *Kurven* im Raum.
2. $m = 2$, $n = 1$. Der Graph $\{(x, f(x)) \in \mathbb{R}^2 \times \mathbb{R} \mid x \in U\}$ von f lässt sich als *Fläche über der Ebene* auffassen, siehe Abb. 8.1.

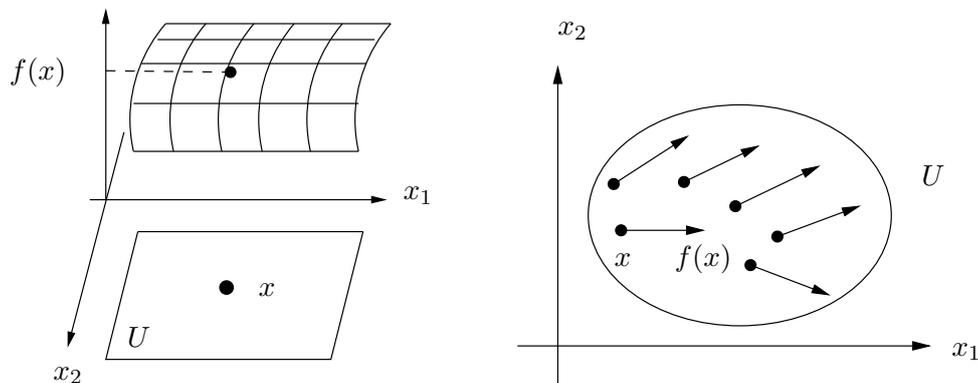


Abbildung 8.1: Links: Graph einer Funktion $f : U \rightarrow \mathbb{R}$ für $U \subset \mathbb{R}^2$. Rechts: Ein Vektorfeld f auf $U \subset \mathbb{R}^2$

3. $m = n$. Hier nennt man f auch *Vektorfeld*, siehe Abb. 8.1 für $m = n = 2$.
4. $m, n \in \mathbb{N}$ beliebig, $n \times m$ -Matrix $A \in \text{Mat}(n \times m, \mathbb{R})$
Lineare Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$, $f(x) := Ax$. ◇

8.2 Bemerkungen 1. Wenn keine Missverständnisse drohen, lässt man oft die Vektorpfeile weg. Diese Konvention werde ich verwenden. In der Literatur findet man stattdessen auch die Schreibweisen \vec{f} und \underline{f} .

2. Allgemein lässt sich bezüglich der kanonischen Koordinaten des \mathbb{R}^n eine Abbildung $f : U \rightarrow \mathbb{R}^n$ als ein n -Tupel

$$f = (f_1, \dots, f_n)^t \quad \text{von Funktionen} \quad f_i : U \rightarrow \mathbb{R} \quad (8.1)$$

schreiben. Bequemlichkeitshalber lasse ich, wo dies nicht zu Missverständnissen führt, das Transpositionszeichen weg, identifiziere also Zeilen- und Spaltenvektoren. \diamond

Wir sind insbesondere an *stetigen* Abbildungen interessiert.

8.3 Satz $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : U \rightarrow \mathbb{R}^n$ ist genau dann stetig, wenn alle Koordinatenfunktionen $f_k : U \rightarrow \mathbb{R}$ stetig sind.

Beweis: Nach Satz 9.4 der *Analysis I* ist f genau dann stetig in $x \in U$, wenn f folgenstetig in $x \in U$ ist, und analog sind die Komponenten f_k von f genau dann stetig in $x \in U$, wenn sie dort folgenstetig sind.

Es sei also $(a_\ell)_{\ell \in \mathbb{N}}$ eine gegen x konvergente Folge von $a_\ell = \begin{pmatrix} a_{\ell,1} \\ \vdots \\ a_{\ell,m} \end{pmatrix} \in \mathbb{R}^m$, und $b_\ell = \begin{pmatrix} b_{\ell,1} \\ \vdots \\ b_{\ell,n} \end{pmatrix} := f(a_\ell) \in \mathbb{R}^n$ das ℓ -te Glied der Bildfolge. Nach Satz 7.32 der *Analysis I* konvergiert $(b_\ell)_{\ell \in \mathbb{N}}$ genau dann gegen $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := f(x)$, wenn die reellen Folgen $(b_{\ell,k})_{\ell \in \mathbb{N}}$ gegen $y_k = f_k(x)$ konvergieren ($k = 1, \dots, n$). Nun ist aber $b_{\ell,k} = f_k(a_\ell)$, letzteres Kriterium also das der Folgenstetigkeit der f_k . \square

Sind also in (8.1) die Koordinatenfunktionen f_1, \dots, f_n stetig, dann auch $f : U \rightarrow \mathbb{R}^n$. Das folgende Beispiel zeigt dagegen, dass eine Funktion $f : U \rightarrow \mathbb{R}$ unstetig sein kann, obgleich alle *partiellen Funktionen* (bei denen man $m - 1$ der Koordinaten einen festen Wert gibt) stetig sind.

8.4 Beispiel (Partielle Funktionen) Für die reelle Funktion zweier Variablen

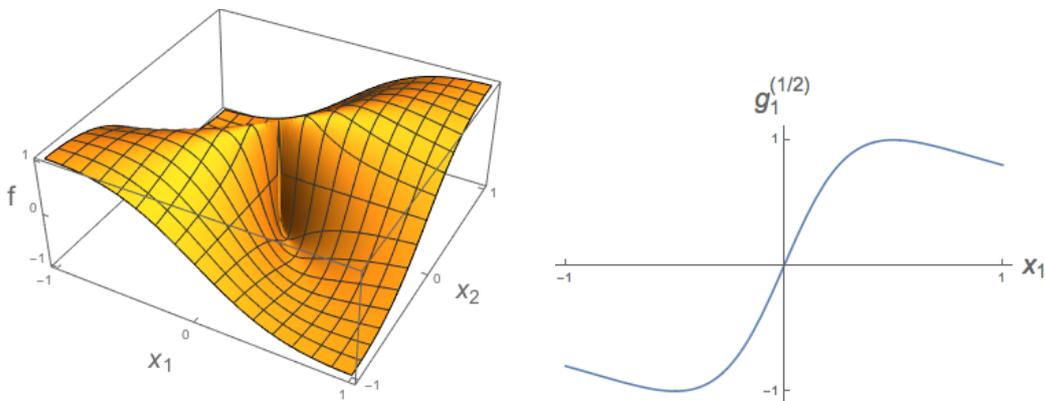
$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad f(x) := \begin{cases} \frac{2x_1x_2}{x_1^2+x_2^2} & , x \neq 0 \\ 0 & , x = 0 \end{cases}$$

und alle $c \in \mathbb{R}$ sind die Funktionen

$$g_1^{(c)}(x) := f(x, c) \quad \text{und} \quad g_2^{(c)}(x) := f(c, x) = g_1^{(c)}(x) \quad (x \in \mathbb{R})$$

einer Variablen stetig. Denn für $c \neq 0$ ist

$$g_l^{(c)}(x) = \frac{2xc}{x^2 + c^2} \quad \text{und} \quad g_l^{(0)}(x) = 0, \quad (l = 1, 2).$$



Dagegen ist $1 = \lim_{t \rightarrow 0} f(t, t) \neq \lim_{t \rightarrow 0} f(t, -t) = -1$, während $f(0) = 0 = \lim_{t \rightarrow 0} f(t, t) = \lim_{t \rightarrow 0} f(t, -t)$ ist. Also ist f unstetig bei $x = 0$. \diamond

8.1 Die Landau-Symbole

In vielen Fällen genügt eine ungefähre Kenntnis einer Funktion. Ist etwa α ein betragsmäßig kleiner Winkel, dann können wir in guter Näherung $\sin(\alpha)$ durch α ersetzen. Um nun ein Maß für Güte der die Näherung zu erhalten, sei allgemein $f : D \rightarrow \mathbb{R}^n$ auf einer Teilmenge $D \subseteq \mathbb{R}^m$ definiert und $a \in \overline{D}$.

8.5 Definition Es sei $g : D \rightarrow \mathbb{R}$.

1. Falls für ein geeignetes $\varepsilon > 0$ ein $C = C(\varepsilon) > 0$ existiert mit

$$\|f(x)\| \leq C|g(x)| \quad (x \in D \text{ mit } \|x - a\| < \varepsilon),$$

heißt f bei a **von der Ordnung groß O von g** , und man schreibt²²

$$f(x) = \mathcal{O}(g(x)) \quad \text{oder} \quad f = \mathcal{O}(g) \quad (x \rightarrow a).$$

2. Falls man dabei sogar $C(\varepsilon) > 0$ so wählen kann, dass $\lim_{\varepsilon \searrow 0} C(\varepsilon) = 0$ gilt, also (unter der Voraussetzung $g(x) \neq 0$ für alle $x \in D$ mit $\|x - a\| < \varepsilon$)

$$\lim_{x \rightarrow a} \frac{\|f(x)\|}{|g(x)|} = 0$$

ist, heißt f bei a **von der Ordnung klein o von g** , und man schreibt

$$f(x) = o(g(x)) \quad \text{oder} \quad f = o(g) \quad (x \rightarrow a).$$

²²Da in der Landau-Notation $f = \mathcal{O}(g)$ der Punkt a nicht auftaucht, darf man nicht vergessen, ihn zu nennen, z.B. in der Form des nachgestellten $(x \rightarrow a)$. Ohne diese Angabe ergibt die Formel $f = \mathcal{O}(g)$ keinen Sinn!

Analog verfährt man für $f : \mathbb{R} \rightarrow \mathbb{R}^n$ und die Punkte $a = \pm\infty$, die man als Punkte der erweiterten Zahlengerade $\overline{\mathbb{R}}$ auffasst, und bei denen man die Umgebungen $[R, \infty)$ bzw. $(-\infty, -R]$ verwendet.

8.6 Beispiele

1. Für $a = 0$ ist $\sin(x) - x = o(x)$ ($x \rightarrow 0$),
denn wir haben in der *Analysis I* bewiesen, dass $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$, also $\lim_{x \rightarrow 0} \frac{\sin(x) - x}{x} = 0$ ist.
2. Genauer gilt sogar

$$\sin(x) - x = \mathcal{O}(x^3) \quad (x \rightarrow 0),$$

denn

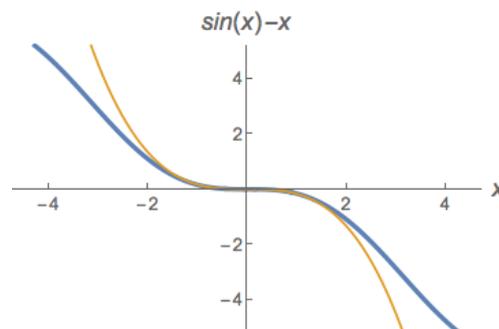
$$|\sin(x) - x| = \left| \sum_{k=1}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \right| = |x^3| \cdot \left| \sum_{k=1}^{\infty} (-1)^k \frac{x^{2(k-1)}}{(2k+1)!} \right|.$$

Letztere Potenzreihe konvergiert auf ganz \mathbb{R} und ist damit stetig, also auf jedem kompakten Intervall $[-\varepsilon, \varepsilon]$ beschränkt.

In der nebenstehenden Abbildung wird der Graph der Funktion

$$x \mapsto \sin(x) - x$$

(fett gezeichnet) mit dem Graph von $x \mapsto \frac{-x^3}{3!}$ verglichen.

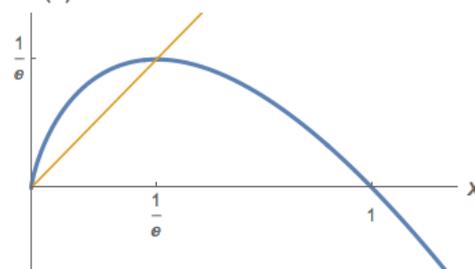


3. Bei $a = \infty$ ist $\sin(x) = \mathcal{O}(1)$, denn der Sinus ist sogar auf ganz \mathbb{R} beschränkt. Dagegen gilt **nicht** $\sin(x) = o(1)$, denn $|\sin((n + \frac{1}{2})\pi)| = 1$ für alle $n \in \mathbb{N}$.
4. Für die (in der Informationstheorie $-x \ln(x)$

wichtige) Funktion

$$f : (0, \infty) \rightarrow \mathbb{R} \quad , \quad f(x) := -x \ln(x)$$

gilt $f(x) = o(1)$ ($x \rightarrow 0$), aber **nicht** $f(x) = \mathcal{O}(x)$ ($x \rightarrow 0$). \diamond



Als Vergleichsfunktionen $g : D \rightarrow \mathbb{R}$ bieten sich die Beträge $g(x) := \|x - a\|^n$ an. Im Eindimensionalen ($D \subseteq \mathbb{R}$) kann man einfach $g(x) := (x - a)^n$ nehmen.

Man kann die Schreibweise $f = \mathcal{O}(g)$ ($x \rightarrow a$) so interpretieren, dass f zur Menge $\mathcal{O}(g)$ der Funktionen gehört, die sich bei a in der beschriebenen Art und Weise verhalten. Dies ist ein \mathbb{R} -Vektorraum. In diesem Sinn können wir aus

$$f_1 = \mathcal{O}(g_1) \quad \text{und} \quad f_2 = \mathcal{O}(g_2) \quad (x \rightarrow a)$$

schließen, dass

$$f_1 + f_2 = \mathcal{O}(g_1) + \mathcal{O}(g_2) = \mathcal{O}(|g_1| + |g_2|) \quad (x \rightarrow a)$$

und

$$f_1 \cdot f_2 = \mathcal{O}(g_1) \cdot \mathcal{O}(g_2) = \mathcal{O}(g_1 \cdot g_2) \quad (x \rightarrow a)$$

ist, und analog für das Klein- o -Symbol. Ist andererseits $g_2 = \mathcal{O}(g_1)$ ($x \rightarrow a$), dann können wir $\mathcal{O}(g_1) + \mathcal{O}(g_2) = \mathcal{O}(g_1)$ ($x \rightarrow a$) schreiben.

Ebenso folgt mit der mengentheoretischen Interpretation des \mathcal{O} -Symbols aus $f_1 - f_2 = \mathcal{O}(g)$

$$f_1 = f_2 + \mathcal{O}(g) \quad (x \rightarrow a).$$

Beispielsweise kann man mit Beispiel 8.6.2 schreiben:

$$\sin(x) = x + \mathcal{O}(x^3) \quad (x \rightarrow 0).$$

Analog schreibt man $f_1 = f_2 + o(g)$ statt $f_1 - f_2 = o(g)$.

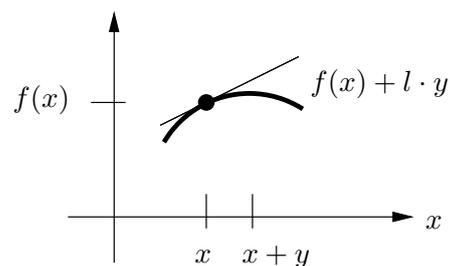
8.2 Definition der totalen Ableitung

Nun wollen wir den Begriff der Ableitung verallgemeinern und lassen uns von der eindimensionalen Situation leiten. Dort war ja für eine bei $x \in U \subset \mathbb{R}$ differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ die Ableitung $f'(x)$ die Zahl $l \in \mathbb{R}$, für die

$$f(x+y) = f(x) + l \cdot y + o(y) \quad (y \rightarrow 0).$$

Wenn l existiert, dann ist die Steigung der Tangentialgeraden (und darum handelt es sich ja) eindeutig bestimmt.

Im \mathbb{R}^m hat man nun mehr Möglichkeiten $x \in U$ zu variieren, y wird ein (betragsmäßig kleiner) Vektor des \mathbb{R}^m sein.



8.7 Definition $f : U \rightarrow \mathbb{R}^n$ mit $U \subset \mathbb{R}^m$ offen heißt bei $x \in U$ **(total) differenzierbar**, wenn es eine Matrix $L \in \text{Mat}(n \times m, \mathbb{R})$ mit

$$f(x + y) = f(x) + Ly + o(\|y\|) \quad (y \rightarrow 0) \quad (8.2)$$

gibt. Dann heißt L **Ableitung von f bei x** .

Wieder gibt es höchstens eine solche Matrix, denn für L_1, L_2 , die die Bedingung erfüllen, muss

$$L_1 y = L_2 y + o(\|y\|)$$

sein. Für $M := L_1 - L_2$ gilt damit $\lim_{y \rightarrow 0} \frac{My}{\|y\|} = 0$, also $M = 0$.

8.8 Bemerkung Insbesondere folgt aus der Differenzierbarkeit bei x die Stetigkeit bei x , denn da $\lim_{y \rightarrow 0} Ly = 0$ und $\lim_{y \rightarrow 0} o(\|y\|) = 0$, ist

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

8.9 Notation

- Man schreibt dann $Df(x)$ statt L und nennt diese (x -abhängige) Matrix *Jacobi-Matrix von f im Punkt x* .
- Ist f für alle $x \in U$ bei $x \in U$ differenzierbar, dann heißt f *(total) differenzierbar*, und $Df : U \rightarrow \text{Mat}(n \times m, \mathbb{R})$ die *Ableitung von f* . \diamond

8.3 Partielle Ableitungen

Wie rechnet man die Jacobi-Matrix aus?

Wir können uns die einzelnen Komponenten von f getrennt vornehmen, also $Df(x)$ zeilenweise berechnen.

8.10 Definition Eine Funktion $g : U \rightarrow \mathbb{R}$ (mit $U \subset \mathbb{R}^m$ offen) besitzt bei $x \in U$ die **partielle Ableitung nach der k -ten Variablen**

$$D_k g(x) = \frac{\partial g}{\partial x_k}(x) \quad (k = 1, \dots, m),$$

wenn für den kanonischen Basisvektor $e_k = (0, \dots, 0, 1, 0, \dots, 0)^t \in \mathbb{R}^m$

$$\frac{\partial g}{\partial x_k}(x) := \lim_{h \rightarrow 0} \frac{g(x + h \cdot e_k) - g(x)}{h}$$

existiert.

8.11 Satz Ist $f : U \rightarrow \mathbb{R}^n$ bei $x \in U$ differenzierbar, dann existieren die partiellen Ableitungen $\frac{\partial f_i}{\partial x_k}(x)$, und

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_m}(x) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \dots & \frac{\partial f_n}{\partial x_m}(x) \end{pmatrix}.$$

Beweis: Für $y = he_k$ mit $h \in \mathbb{R}$ ist

$$f(x + he_k) = f(x) + h \underbrace{Df(x)}_L \cdot e_k + o(h) \quad (k = 1, \dots, m),$$

sodass

$$f_i(x + he_k) = f_i(x) + hL_{ik} + o(h) \quad (i = 1, \dots, n)$$

ist. Damit existiert $\lim_{h \rightarrow 0} \frac{f_i(x + he_k) - f_i(x)}{h} = \frac{\partial f_i}{\partial x_k}(x)$ und ist gleich L_{ik} . \square

8.12 Beispiel (Totalableitung) Für $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = \begin{pmatrix} x_1^3 - 3x_1x_2^2 \\ 3x_1^2x_2 - x_2^3 \end{pmatrix}$ ist $Df(x) = \begin{pmatrix} 3x_1^2 - 3x_2^2 & -6x_1x_2 \\ 6x_1x_2 & 3(x_1^2 - x_2^2) \end{pmatrix}$, und die Ableitung am Punkt $x := \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ ist

$$Df\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 9 & -12 \\ 12 & 9 \end{pmatrix}. \quad \diamond$$

Wir können aber nicht umgekehrt von der Existenz aller partiellen Ableitungen bei x auf die Differenzierbarkeit bei x schließen.

8.13 Beispiel (Unstetige partielle Ableitungen)

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} \frac{x_1x_2}{\|x\|} & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases}.$$

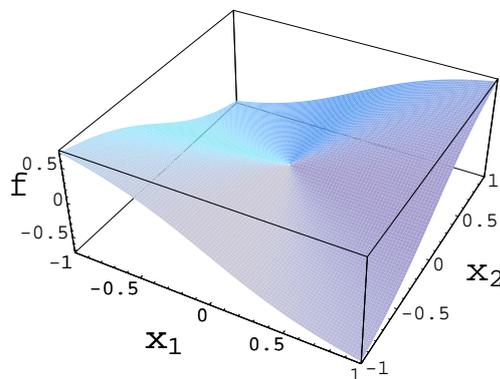
Diese Funktion ist auch stetig am Punkt $0 \in \mathbb{R}^2$, denn $|f(x)| \leq \|x\|$, also $\lim_{x \rightarrow 0} f(x) = 0$. Auch ihre partiellen Ableitungen existieren überall:

$$\frac{\partial f}{\partial x_1}(x) = \begin{cases} \frac{x_2^3}{\|x\|^3} & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases} \quad \text{und analog}$$

für $\frac{\partial f}{\partial x_2}$. Allerdings existiert die totale Ableitung bei Null nicht. Nach Satz 8.11 käme ja nur $Df(0) = (0, 0)$ in Frage, aber für $y := \begin{pmatrix} h \\ h \end{pmatrix}$ ist

$$f(y) - f(0) = f(y) = |h|/\sqrt{2}$$

nicht von der Ordnung $o(\|y\|) = o(h)$. \diamond



Verlangen wir allerdings zusätzlich, dass die partiellen Ableitungen stetig sind, dann gilt die Umkehrung von Satz 8.11:

8.14 Satz Eine auf der offenen Menge $U \subset \mathbb{R}^m$ definierte Funktion $f : U \rightarrow \mathbb{R}^n$ besitze die partiellen Ableitungen $\frac{\partial f_i}{\partial x_k} : U \rightarrow \mathbb{R}$. Falls diese bei $x \in U$ stetig sind, ist f bei x total differenzierbar.

Beweis:

- f ist genau dann total differenzierbar, wenn die Komponenten f_1, \dots, f_n total differenzierbar sind. Dies folgt aus Satz 8.3, wenn man bedenkt, dass das Kriterium (8.2) gleichbedeutend ist mit

$$\lim_{y \rightarrow 0} \frac{f(x+y) - f(x) - Ly}{\|y\|} = 0,$$

und analog $\lim_{y \rightarrow 0} \frac{f_k(x+y) - f_k(x) - l_k y}{\|y\|} = 0$ äquivalent zur totalen Differenzierbarkeit von f_k bei x , mit Ableitung $Df_k(x) = l_k \in \mathbb{R}^m$.

Also nehmen wir $n = 1$ an; f ist dann eine reelle Funktion von m Variablen.

- Wir führen einen Induktionsbeweis in der Zahl m der Variablen. Ist $m = 1$, dann ist das Kriterium

$$f(x+y) = f(x) + \sum_{k=1}^m l_k y_k + o(\|y\|) \quad (y \rightarrow 0) \quad (8.3)$$

gleichbedeutend mit der Existenz der partiellen Ableitung.

- Wir nehmen an, dass die Aussage für Funktionen von $m - 1$ Variablen richtig ist und zeigen ihre Richtigkeit für m Variablen. Dazu schreiben wir $x = (x', x_m) \in \mathbb{R}^m$ mit $x' \in \mathbb{R}^{m-1}$ und analog $y = (y', y_m)$. Die zu beweisende Abschätzung (8.3), aufgelöst

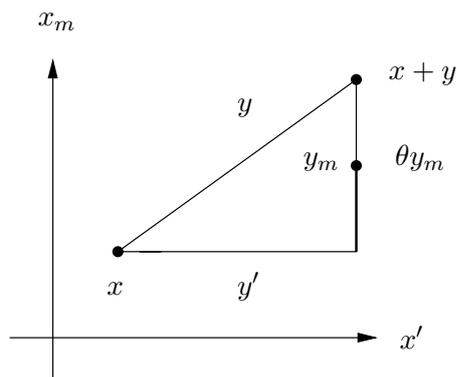


Abbildung 8.2: Aufspaltung von $x + y \in U$

nach dem $o(\|y\|)$ -Term ist damit

$$\begin{aligned}
 A &:= \left| f(x' + y', x_m + y_m) - f(x', x_m) - \sum_{k=1}^m l_k y_k \right| \\
 &\leq \left| f(x' + y', x_m + y_m) - f(x' + y', x_m) - l_m y_m \right| \\
 &\quad + \left| f(x' + y', x_m) - f(x', x_m) - \sum_{k=1}^{m-1} l_k y_k \right| \\
 &= \left| \frac{\partial f}{\partial x_m}(x' + y', x_m + \theta y_m) - l_m \right| \cdot |y_m| \\
 &\quad + \left| f(x' + y', x_m) - f(x', x_m) - \sum_{k=1}^{m-1} l_k y_k \right|
 \end{aligned}$$

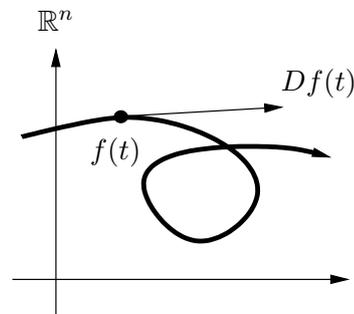
mit $\theta \in (0, 1)$ nach dem Mittelwertsatz der Differentialrechnung. Wir setzen $l_m := \frac{\partial f}{\partial x_m}(x)$. Wegen der Stetigkeit von $\frac{\partial f}{\partial x_m}$ bei x und nach Induktionsvoraussetzung existiert für $\varepsilon > 0$ ein $\delta > 0$ mit

$$\begin{aligned}
 \triangleright &\left| \frac{\partial f}{\partial x_m}(x + y) - l_m \right| < \frac{\varepsilon}{2} \text{ für } \|y\| < \delta \text{ und} \\
 \triangleright &\left| f(x' + y', x_m) - f(x', x_m) - \sum_{k=1}^{m-1} l_k y'_k \right| < \frac{\varepsilon}{2} \|y'\| \text{ für } \|y'\| < \delta.
 \end{aligned}$$

Daher ist $A \leq \frac{\varepsilon}{2} |y_m| + \frac{\varepsilon}{2} \|y'\| \leq \varepsilon \|y\|$ ($\|y\| < \delta$),
sodass $A = o(\|y\|)$ bewiesen ist. \square

Damit haben wir ein handliches Kriterium für die totale Differenzierbarkeit.

- Die Ableitung einer C^1 -Kurve $f : U \rightarrow \mathbb{R}^n$ auf dem offenen Intervall $U \subset \mathbb{R}$ haben wir schon im letzten Kapitel besprochen: Der Tangentialvektor am Punkt $f(t)$ der Kurve ist der Geschwindigkeitsvektor $Df(t) = \begin{pmatrix} f'_1(t) \\ \vdots \\ f'_n(t) \end{pmatrix} \in \mathbb{R}^n$.



- Auch der Fall $n = 1$ ist geometrisch wichtig. Hier haben wir eine Funktion $f : U \rightarrow \mathbb{R}$ mit $U \subset \mathbb{R}^m$, die wir durch ihren Graphen oder auch durch ihre Höhenlinien darstellen können. Der Wert c von f wird hier als Höhe interpretiert und $f^{-1}(c)$ als *Höhenlinie* bezeichnet. ²³

8.15 Beispiel $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x_1^2 - x_2^2$.

Der Graph von f besitzt die Form einer Sattelfläche, siehe Abbildung 8.3. \diamond

²³Die Darstellung einer reellen Funktion durch mehrere Höhenlinien wird auch *Contour-Plot* genannt.

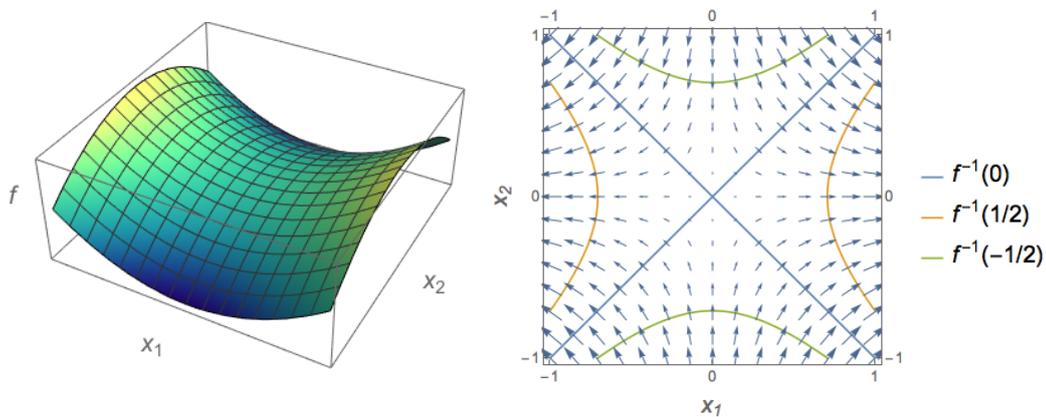


Abbildung 8.3: Graph (links) und Gradientenvektorfeld mit Höhenlinien (rechts) für $f(x) = x_1^2 - x_2^2$

8.4 Der Gradient

In Bsp. 8.15 ist die Ableitung am Punkt x gleich $Df(x) = (2x_1, -2x_2)$.

Interpretieren wir diese durch Transposition als Tangentialvektor an x , so ergibt sich, dass dieser senkrecht auf der Höhenlinie durch x steht und in Richtung des stärksten Anstieges von f weist.

Dies gilt ganz allgemein:

8.16 Definition Der **Gradient** bei $x \in U$ einer total differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^m$ ist der Tangentialvektor

$$\text{grad}f(x) := (Df(x))^t = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_m}(x) \end{pmatrix} \quad \text{an der Stelle } x.$$

8.17 Bemerkungen (Gradient)

1. Ist f total differenzierbar (d.h. für alle $x \in U$ total differenzierbar), dann ist $\text{grad}f : U \rightarrow \mathbb{R}^m$ ein Vektorfeld.
2. Statt $\text{grad}f$ schreibt man auch ∇f (gesprochen: "Nabla f "). ◇

8.18 Beispiel (Gradient) Der Graph von

$$f(x) = \cos(x_1) \cdot \cos(x_2) \quad (x \in \mathbb{R}^2)$$

ähneln einer Eierpappe (siehe Abb. 8.4), und das Gradientenvektorfeld von f ist

$$\nabla f(x) = \begin{pmatrix} -\sin(x_1) \cos(x_2) \\ -\cos(x_1) \sin(x_2) \end{pmatrix}. \quad \diamond$$

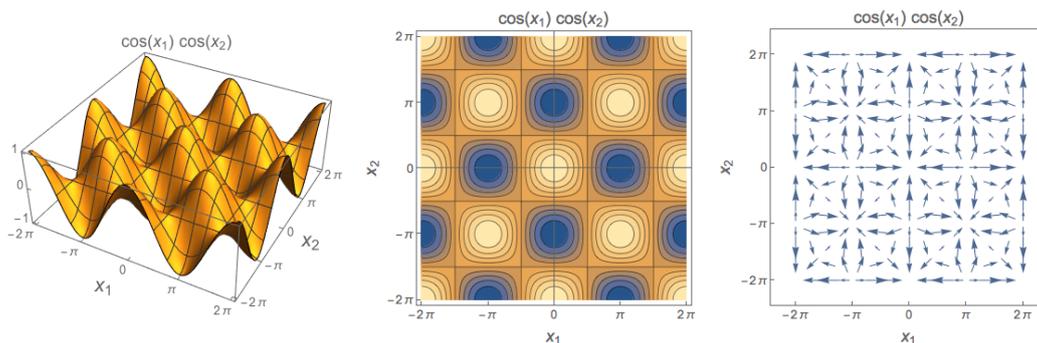


Abbildung 8.4: Graph (links), Contour-Plot (Mitte) und Gradientenvektorfeld (rechts) von $x \mapsto \cos(x_1) \cdot \cos(x_2)$

8.19 Definition Die **Richtungsableitung** der auf der offenen Teilmenge $U \subset \mathbb{R}^m$ total differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ in Richtung $v \in \mathbb{R}^m$ ist durch

$$\partial_v f(x) := Df(x)(v)$$

gegeben.

Damit ist $\partial_v f(x) = \langle \nabla f(x), v \rangle$.

Da die Richtungsableitung der differentielle Anstieg in Richtung v ist, ergibt sich:

8.20 Satz Ist $\nabla f(x) \neq 0$, dann zeigt der Gradient in Richtung des stärksten Anstieges von f .

Beweis: Es sei $\|v\| = 1$. Dann ist $\partial_v f(x) = \|\text{grad} f(x)\| \cos \varphi$, wobei φ der Winkel zwischen $\nabla f(x)$ und v ist. Das Maximum wird für $\varphi = 0$, also $v = \frac{\nabla f(x)}{\|\nabla f(x)\|}$ angenommen. \square

8.5 Differentiationsregeln

Aus der eindimensionalen Differentialrechnung sind uns Differentiationsregeln bekannt:

- Summenregel: $(f + g)' = f' + g'$
- Produktregel: $(f \cdot g)' = f' \cdot g + f \cdot g'$
- Kettenregel: $(g \circ f)' = (g' \circ f) \cdot f'$

- Quotientenregel: $(f/g)' = \frac{f' \cdot g - f \cdot g'}{g^2}$; diese ist eine Folge von Produktregel und Kettenregel.

Wir wollen nun sehen, wie sich diese Regeln in höheren Dimensionen verallgemeinern. Für die Verallgemeinerung der Produktregel überlegen wir uns zunächst, welche Funktionen wir sinnvoll miteinander multiplizieren können. Es sei dabei und im Folgenden $U \subset \mathbb{R}^m$ offen.

1. Funktionen $f, g : U \rightarrow \mathbb{R}^n$ besitzen Werte $f(x), g(x) \in \mathbb{R}^n$. Unter Verwendung des kanonischen Skalarproduktes $\langle v, w \rangle := \sum_{k=1}^n v_k w_k$ von Vektoren $v, w \in \mathbb{R}^n$ bezeichnen wir mit $\langle f, g \rangle$ die reelle Funktion

$$\langle f, g \rangle : U \rightarrow \mathbb{R} \quad , \quad \langle f, g \rangle(x) := \langle f(x), g(x) \rangle .$$

2. Die vektorwertige Funktion $f : U \rightarrow \mathbb{R}^n$ können wir mit einer reellen Funktion $h : U \rightarrow \mathbb{R}$ punktweise multiplizieren und erhalten die vektorwertige Funktion

$$h \cdot f : U \rightarrow \mathbb{R}^n \quad , \quad (h \cdot f)(x) := h(x) \cdot f(x) .$$

8.21 Satz

- Sind $f, g : U \rightarrow \mathbb{R}^n$ und $h : U \rightarrow \mathbb{R}$ bei $x \in U \subset \mathbb{R}^m$ differenzierbar, dann gilt die **Summenregel**

$$D(f + g)(x) = Df(x) + Dg(x) ,$$

und die **Produktregeln**

$$D \langle f, g \rangle(x) = \sum_{k=1}^n (f_k(x) Dg_k(x) + g_k(x) Df_k(x)) \quad (8.4)$$

und

$$D(h \cdot f)(x) = f(x) Dh(x) + h(x) Df(x) .$$

- Ist $f : U \rightarrow \mathbb{R}^n$ bei $x \in U \subset \mathbb{R}^m$ differenzierbar, und $g : V \rightarrow \mathbb{R}^l$ bei $y := f(x) \in V \subset \mathbb{R}^n$ differenzierbar (und $V \supset f(U)$ offen), dann gilt die **Kettenregel**

$$D(g \circ f)(x) = Dg(y) Df(x) .$$

8.22 Bemerkungen (Ableitungs-Regeln)

1. Insbesondere *existieren* die Ableitungen der verknüpften Funktionen!

2. Die Dimensionen der Ableitungs-Matrizen ergeben sich wie folgt:

$$\underbrace{D(h \cdot f)(x)}_{n \times m\text{-Matrix}} = \underbrace{f(x)}_{\substack{\text{Spaltenvektor} \\ \text{Länge } n}} \cdot \underbrace{Dh(x)}_{\substack{\text{Zeilenvektor} \\ \text{Länge } m}} + \underbrace{h(x)}_{\text{Zahl}} \cdot \underbrace{Df(x)}_{n \times m\text{-Matrix}}$$

und

$$\underbrace{D(g \circ f)(x)}_{l \times m\text{-Matrix}} = \underbrace{Dg(y)}_{l \times n\text{-Matrix}} \cdot \underbrace{Df(x)}_{n \times m\text{-Matrix}}.$$

3. Für total differenzierbare Funktionen $f, g : U \rightarrow \mathbb{R}^n$ und $k \in \mathbb{R}$ ergibt sich damit

$$D(f + g) = Df + Dg \quad , \quad D(kf) = kDf,$$

die Ableitung ist also eine lineare Operation!

4. Die erste Produktregel (8.4) soll als Hausaufgabe bewiesen werden. \diamond

Beweis: Wir setzen die Differenzierbarkeit bei x voraus, also

$$f(x + y) - f(x) = Df(x)y + o(\|y\|) \quad \text{und} \quad g(x + y) - g(x) = Dg(x)y + o(\|y\|) \quad (8.5)$$

mit $Df(x), Dg(x) \in \text{Mat}(n \times m, \mathbb{R})$.

• Die *Summenregel* der Differentiation ergibt sich wie in der Analysis reeller Funktionen: Wegen (8.5) ist

$$\begin{aligned} (f + g)(x + y) - (f + g)(x) &= (f(x + y) - f(x)) + (g(x + y) - g(x)) \\ &= (Df(x)y + o(\|y\|)) + (Dg(x)y + o(\|y\|)) \\ &= (Df(x) + Dg(x))y + o(\|y\|). \end{aligned}$$

• Die *Produktregel* folgt aus (8.5) und der dazu analogen definierenden Relation $h(x + y) - h(x) = Dh(x)(y) + o(\|y\|)$ von $Dh(x)$. Mit der Abkürzung $k := hf : U \rightarrow \mathbb{R}^n$ ist nämlich

$$\begin{aligned} k(x + y) - k(x) &= h(x + y)f(x + y) - h(x)f(x) \\ &= f(x + y)(h(x + y) - h(x)) + h(x)(f(x + y) - f(x)) \\ &= (f(x) + \mathcal{O}(\|y\|))(Dh(x)(y) + o(\|y\|)) + h(x)(Df(x)(y) + o(\|y\|)) \\ &= f(x)Dh(x)(y) + h(x)Df(x)(y) + o(\|y\|) \\ &= [f(x)Dh(x) + h(x)Df(x)](y) + o(\|y\|). \end{aligned}$$

• *Kettenregel:* Mit den Abkürzungen $L := Df(x)$, $M := Dg(y)$ und $h := g \circ f : U \rightarrow \mathbb{R}^l$ ergibt sich $g(y + Y) - g(y) = MY + o(\|Y\|)$.

Wir setzen hier $Y := f(x + X) - f(x)$, also $Y = LX + o(\|X\|)$.

Damit ist

$$\begin{aligned} h(x + X) - h(x) &= g(\overbrace{y + f(x + X) - f(x)}^Y) - g(y) \\ &= M(LX + o(\|X\|)) + o(\|f(x + X) - f(x)\|) \\ &= MLX + o(\|X\|), \end{aligned}$$

sodass $Dh(x) = ML$ ist. \square

8.23 Beispiel 1. Ein Wanderer befindet sich im Gebirge, dessen Höhe am Punkt (x, y) gleich $z(x, y)$ sei. Zum Zeitpunkt t sind seine Horizontalkoordinaten $c(t) := (x(t), y(t))$, er befindet sich also in der Höhe $z(x(t), y(t))$, und er bewegt sich mit der Geschwindigkeit

$$\frac{d}{dt} z \circ c(t) = Dz(c(t))Dc(t) = \left\langle \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}, \nabla z(x(t), y(t)) \right\rangle$$

nach oben.

2. Eine Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ hat in Polarkoordinaten (r, φ) (mit $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \tilde{x}(r, \varphi) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$) die Form $G(r, \varphi) := g(r \cos \varphi, r \sin \varphi)$.

Es ist nach der Kettenregel $DG(r, \varphi) = Dg(x)D\tilde{x}(r, \varphi)$ mit $Dg = (g_{x_1}, g_{x_2})$ und $D\tilde{x}(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$, also

$$\begin{aligned} G_r &:= \frac{\partial}{\partial r} G = g_{x_1} \cos \varphi + g_{x_2} \sin \varphi \\ G_\varphi &:= \frac{\partial}{\partial \varphi} G = g_{x_1} (-r \sin \varphi) + g_{x_2} r \cos \varphi. \end{aligned}$$

Damit ergibt sich (durch Auflösen nach den partiellen Ableitungen von g) für $r \neq 0$

$$\nabla g(\tilde{x}(r, \varphi)) = \begin{pmatrix} G_r \cos \varphi - \frac{1}{r} G_\varphi \sin \varphi \\ G_r \sin \varphi + \frac{1}{r} G_\varphi \cos \varphi \end{pmatrix}. \quad \diamond$$

8.6 Höhere Ableitungen

Ist $f : U \rightarrow \mathbb{R}^n$ total differenzierbar, dann ist $g := Df$ eine Abbildung $g : U \rightarrow \mathbb{R}^\ell$ mit $\ell := mn$, denn der \mathbb{R} -Vektorraum $\text{Mat}(n \times m, \mathbb{R})$ hat die Dimension ℓ , ist also zum \mathbb{R}^ℓ isomorph.

Wir können damit die totale Ableitung von g , also die zweite *totale* Ableitung DDf von f untersuchen. Einfacher erscheint es zunächst, zweite und höhere *partielle* Ableitungen zu betrachten.

8.24 Definition *Es sei $f : U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^m$ offen. Existieren für $k \in \mathbb{N}$, $l \leq k$, beliebige $i_1, \dots, i_l \in \{1, \dots, m\}$ und $j = 1, \dots, n$ die partiellen Ableitungen $\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_l}} f_j$ der Ordnung l und sind diese stetig, dann heißt f **k -mal stetig differenzierbar**.*

$$\begin{aligned} C^k(U, \mathbb{R}^n) &:= \{f : U \rightarrow \mathbb{R}^n \mid f \text{ ist } k\text{-mal stetig differenzierbar}\} \\ C(U, \mathbb{R}^n) &:= C^0(U, \mathbb{R}^n) = \{f : U \rightarrow \mathbb{R}^n \mid f \text{ ist stetig}\} \\ C^\infty(U, \mathbb{R}^n) &:= \bigcap_{k \in \mathbb{N}} C^k(U, \mathbb{R}^n) \\ C^k(U) &:= C^k(U, \mathbb{R}). \end{aligned}$$

8.25 Notation Man schreibt auch

$$D_{i_1} \dots D_{i_k} f \quad \text{oder} \quad f_{x_{i_1} \dots x_{i_k}} \quad \text{oder} \quad \frac{\partial^r}{\partial x_{i_1} \dots \partial x_{i_k}} f \quad \text{für} \quad \frac{\partial}{\partial x_{i_1}} \dots \frac{\partial}{\partial x_{i_k}} f.$$

Die erste Schreibweise ist die beste, denn hier wird nicht die Koordinate benannt, nach der partiell differenziert wird, sondern nur ihr Index.

Wenn wir partielle Ableitungen in der Form $f_{x_1 \dots x_k}$ notieren, ist es sinnvoll, die Komponenten einer vektorwertigen Funktion durch einen oben stehenden Index zu bezeichnen (also $f = (f^{(1)}, \dots, f^{(n)})^t$), um Verwechslungen zu vermeiden. \diamond

Die Funktionenräume $C^k(U, \mathbb{R}^n)$ sind \mathbb{R} -Vektorräume, und es gilt

$$C^{k+1}(U, \mathbb{R}^n) \subset C^k(U, \mathbb{R}^n).$$

Einfache Beispiele zeigen, dass nicht immer die Vertauschungsregel

$$\frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(x_1, x_2) = \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(x_1, x_2) \quad (8.6)$$

gilt:

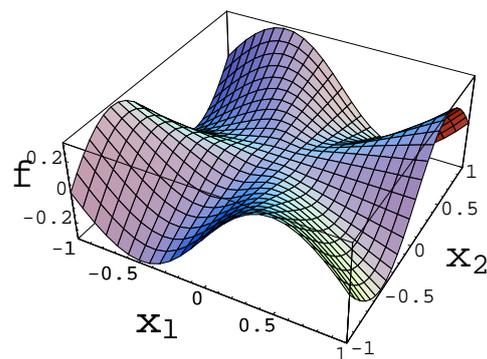
8.26 Beispiel (Vertauschung partieller Ableitungen)

$$\text{Für } f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} \frac{x_1 x_2 (x_1^2 - x_2^2)}{x_1^2 + x_2^2} & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases}$$

existieren die partiellen Ableitungen aus (8.6), und für $x \neq 0$ ist

$$\frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(x) = \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(x) = \frac{(x_1^2 - x_2^2)(x_1^4 + 10x_1^2 x_2^2 + x_2^4)}{(x_1^2 + x_2^2)^3}.$$

Da aber $\frac{\partial}{\partial x_1} f(x) = \frac{x_2(x_1^4 + 4x_1^2 x_2^2 - x_2^4)}{(x_1^2 + x_2^2)^2}$,
 also $\frac{\partial}{\partial x_1} f(0, x_2) = -x_2$, ist
 $\frac{\partial^2}{\partial x_2 \partial x_1} f(0) = -1$, während
 $\frac{\partial^2}{\partial x_1 \partial x_2} f(0) = +1$ gilt. \diamond



Allerdings ist dies für $f \in C^2(U, \mathbb{R}^n)$ der Fall. Allgemeiner gilt:

8.27 Satz Ist $f \in C^k(U, \mathbb{R}^n)$, dann gilt für jede Permutation $\pi \in S_k$:

$$\frac{\partial}{\partial x_{i_{\pi(1)}}} \cdots \frac{\partial}{\partial x_{i_{\pi(k)}}} f = \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_k}} f.$$

Beweis: Jede Permutation $\pi \in S_k$ läßt sich als Produkt von Transpositionen darstellen. Es genügt also die Vertauschung zweier partieller Ableitungen zu betrachten. Höhere Ableitungen erhält man durch Induktion. Wir betrachten den Fall $f \in C^2(U)$, $U \subset \mathbb{R}^2$. In höheren Dimensionen bleiben die anderen Koordinaten unberücksichtigt.

Für kleine $\varepsilon > 0$ ist das Quadrat Q mit den Eckpunkten $x, x + \varepsilon e_1, x + \varepsilon e_2$ und $x + \varepsilon(e_1 + e_2)$ ganz in U enthalten. Wir ordnen Q die Zahl

$$\square f := f(x) - f(x + \varepsilon e_1) - f(x + \varepsilon e_2) + f(x + \varepsilon(e_1 + e_2))$$

zu. In diesem Quadrat liegt ein Punkt y mit

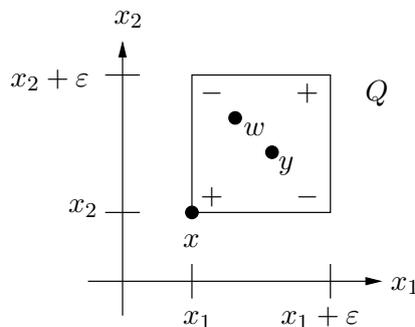


Abbildung 8.5: Vertauschbarkeit der partiellen Ableitungen

$$\square f = \varepsilon^2 \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(y),$$

denn

1. $\square f = g(x_2 + \varepsilon) - g(x_2)$ für $g(z) := f(x_1 + \varepsilon, z) - f(x_1, z)$.
2. Nach dem eindimensionalen Mittelwertsatz gibt es ein $y_2 \in [x_2, x_2 + \varepsilon]$ mit

$$g(x_2 + \varepsilon) - g(x_2) = \varepsilon g'(y_2),$$

also

$$\square f = \varepsilon g'(y_2) = \varepsilon \left(\frac{\partial f}{\partial x_2}(x_1 + \varepsilon, y_2) - \frac{\partial f}{\partial x_2}(x_1, y_2) \right) = \varepsilon (h(x_1 + \varepsilon) - h(x_1))$$

für $h(z) := \frac{\partial f}{\partial x_2}(z, y_2)$.

3. Nochmalige Anwendung des Mittelwertsatzes liefert ein $y_1 \in [x_1, x_1 + \varepsilon]$ mit $h(x_1 + \varepsilon) - h(x_1) = \varepsilon h'(y_1)$, also $\square f = \varepsilon^2 h'(y_1) = \varepsilon^2 \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(y_1, y_2)$.

Aus Symmetriegründen gibt es aber auch ein $w \in Q$ mit $\square f = \varepsilon^2 \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(w)$, also

$$\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(w) = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(y).$$

Für $\varepsilon \searrow 0$ gilt $y \rightarrow x$ und $w \rightarrow x$, also wegen Stetigkeit der partiellen Ableitungen

$$\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(x) = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(x). \quad \square$$

Die von einer $n \times m$ -Matrix A erzeugte lineare Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$, $x \mapsto Ax$ hat die (x -unabhängige) Jacobi-Matrix $Df(x) = A$.

Ist $n = m$, so können wir $\det A$ bilden. Geometrisch entspricht dem Betrag dieser Größe der Faktor, um den das Volumen eines Parallelepipeds P vergrößert wird:

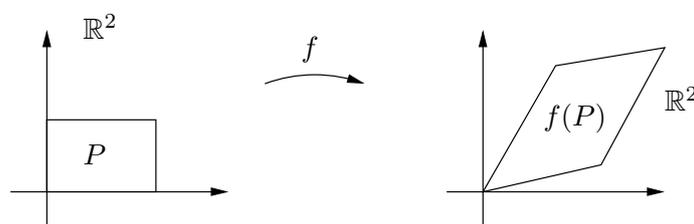


Abbildung 8.6: Geometrische Interpretation der Determinante

Entsprechend können wir die (punktweise definierte) *Funktionaldeterminante* $\det(Df)$ einer Abbildung $f \in C^1(U, \mathbb{R}^m)$ auf einer offenen Teilmenge $U \subset \mathbb{R}^m$ interpretieren.

8.28 Beispiel (Wechsel von Polar- auf Euklidische Koordinaten)

Der Koordinatenwechsel entspricht der Funktion

$$f : \mathbb{R}^+ \times [0, 2\pi] \rightarrow \mathbb{R}^2, \quad f(r, \varphi) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}.$$

Diese besitzt Jacobi-Matrix und -Determinante

$$Df(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}, \quad \det(Df)(r, \varphi) = r.$$

Die geometrische Interpretation des Faktors r wird in Abb. 8.7 dargestellt. \diamond

8.7 Vektoranalysis

In Kapitel 8.4 haben wir mit dem *Gradienten* einen Differentialoperator kennen gelernt, und zwar die lineare Abbildung

$$\text{grad} : C^\infty(U, \mathbb{R}) \rightarrow C^\infty(U, \mathbb{R}^m),$$

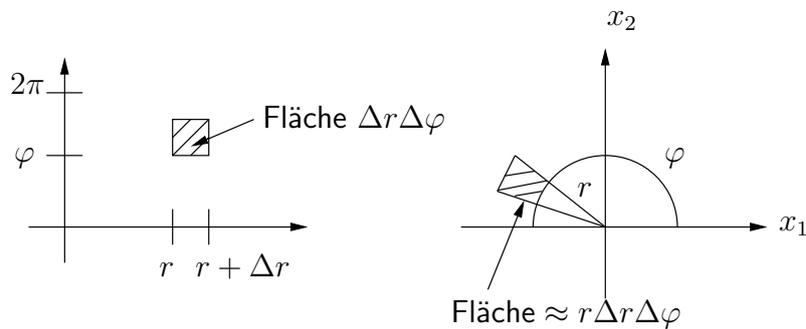


Abbildung 8.7: Funktionaldeterminante für Polarkoordinaten

die einer glatten reellen Funktion $f : U \rightarrow \mathbb{R}$ auf der offenen Menge $U \subseteq \mathbb{R}^m$ das Vektorfeld $\text{grad} f$ auf U zuordnete.

Jetzt werden wir weitere *Differentialoperatoren* untersuchen, also Linearkombinationen partieller Ableitungen. Wichtig sind für uns besonders diejenigen, die sich koordinatenunabhängig definieren lassen und damit einen geometrischen Gehalt besitzen.

Die Divergenz

Wir beginnen mit der Definition der *Divergenz* eines Vektorfeldes $v : U \rightarrow \mathbb{R}^m$ auf der offenen Menge $U \subseteq \mathbb{R}^m$. Wir nehmen dabei an, dass dessen totale Ableitung $U \ni x \mapsto Dv(x) \in \text{Mat}(m, \mathbb{R})$ existiert und stetig ist, und setzen

$$\text{div}(v) := \text{tr}(Dv) = \sum_{i=1}^m \frac{\partial v_i}{\partial x_i},$$

wobei $\text{tr}(M)$ die Spur der quadratischen Matrix M bezeichnet. Damit ist $\text{div}(v)$ eine stetige Funktion auf U . In der Physik schreibt man oft $\nabla \cdot v$ statt $\text{div}(v)$.

Um ein Gefühl für diese Funktion zu bekommen, betrachten wir *lineare* Vektorfelder $v : \mathbb{R}^m \rightarrow \mathbb{R}^m$, d.h. Vektorfelder, die gleichzeitig lineare Abbildungen sind, für die damit eine eindeutig bestimmte Matrix $M \in \text{Mat}(m, \mathbb{R})$ existiert mit $v(x) = Mx$.

Es ist $\text{div}(v)(x) = \sum_{i=1}^m (M)_{ii} = \text{tr}(M)$, die Spur der Matrix, also unabhängig von $x \in \mathbb{R}^m$. Gleichzeitig ist aber, wie aus der *Linearen Algebra* bekannt, $\text{tr}(M)$ gleich der Summe der Eigenwerte der (komplexifizierten) Matrix M , siehe Abb. 8.8.

Die *lineare Differentialgleichung*

$$\dot{x} = v(x) = Mx$$

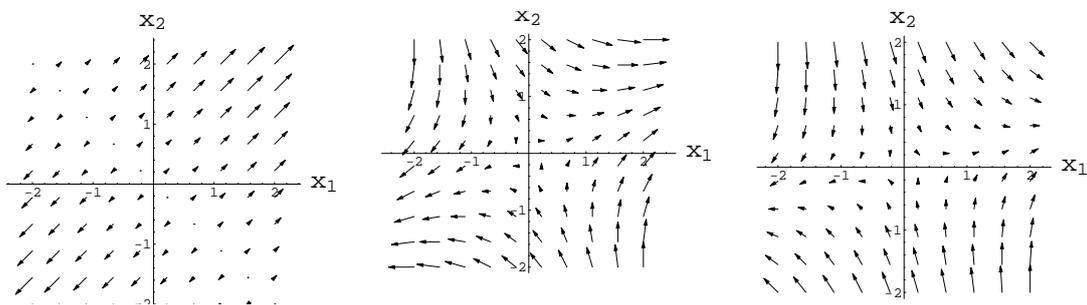


Abbildung 8.8: Lineare Vektorfelder $v(x) = \begin{pmatrix} 1 & 1 \\ 1 & d-1 \end{pmatrix}x$ mit Divergenz d . Links: $d = 2$; Mitte: $d = 0$; Rechts: $d = -2$

besitzt für den Anfangswert $x_0 \in \mathbb{R}^m$ die Lösung

$$x(t) = \exp(Mt)x_0 \quad (t \in \mathbb{R}).$$

Der von den Basisvektoren e_1, \dots, e_m aufgespannte Würfel $[0, 1]^m$ wird unter der Abbildung $\mathbb{R}^m \rightarrow \mathbb{R}^m, x \mapsto \exp(Mt)x$ auf das von den Vektoren $e_1(t) := \exp(Mt)e_1, \dots, e_m(t) := \exp(Mt)e_m$ aufgespannte Parallelotop vom Volumen

$$\det(e_1(t), \dots, e_m(t)) = \det(\exp(Mt))$$

transformiert. Nun gilt aber für eine Matrix $A \in \text{Mat}(m, \mathbb{R})$ immer

$$\det(\exp(A)) = \exp(\text{tr}(A)).$$

Das Volumen des Parallelotops beträgt damit

$$\exp(\text{tr}(M)t) = \exp(\text{div}(v)t).$$

Damit kann die Divergenz als Volumenvergrößerungsfaktor oder *Quellstärke* interpretiert werden.

Der Laplace-Operator

Der Gradient $\text{grad} f$ einer (stetig differenzierbaren) Funktion $f : U \rightarrow \mathbb{R}$ ist ein Vektorfeld auf U . Einem Vektorfeld $v : U \rightarrow \mathbb{R}^m$ können wir umgekehrt die Funktion $\text{div}(v) : U \rightarrow \mathbb{R}$ zuordnen. Zusammensetzung der beiden Operationen ergibt den *Laplace-Operator*

$$\Delta := \text{div grad} : C^2(U) \rightarrow C^0(U).$$

In cartesianen Koordinaten gilt $\text{grad}(f) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{pmatrix}$, also

$$\Delta f = \text{div grad}(f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}.$$

Um ein Gefühl für die Bedeutung des Laplace-Operators zu bekommen, betrachten wir diesmal Funktionen $f: \mathbb{R}^m \rightarrow \mathbb{R}$, die *quadratische Formen* auf \mathbb{R}^m sind, für die also eine Matrix $M \in \text{Mat}(m, \mathbb{R})$ existiert mit

$$f(x) = \frac{1}{2} \langle x, Mx \rangle \quad (x \in \mathbb{R}^m).$$

Der Faktor $\frac{1}{2}$ dient der Vereinfachung der Formeln. Die Matrix M ist durch f nicht eindeutig festgelegt, wohl aber, wenn wir sie als symmetrisch voraussetzen. Dann ist das Vektorfeld

$$\text{grad } f(x) = Mx \quad (x \in \mathbb{R}^m),$$

und $\Delta f(x) = \text{div}(Mx) = \text{tr}(M)$, siehe Abb. 8.9.

Wegen unseres Ansatzes ist Δf konstant. Dies ist ebenso der Fall, wenn wir stattdessen für ein $b \in \mathbb{R}^m$ und $c \in \mathbb{R}$

$$f(x) := \frac{1}{2} \langle x, Mx \rangle + \langle b, x \rangle + c \quad (x \in \mathbb{R}^m)$$

ansetzen.

Betrachten wir speziell die Familie

$$M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}, \text{ dann ist } \Delta f(x) = m_1 + m_2.$$

Wir können m_i als die Krümmung der Kurve $t \mapsto (t, f(e_i t))$ bei $t = 0$ interpretieren. Ist etwa $m_1 > 0$ und $m_2 < 0$, dann besitzt der Graph von f die Form eines Sattels. Ist dessen Krümmung nach oben (in der 1-Richtung) betragsmäßig größer als die Krümmung nach unten (also in der 2-Richtung), dann ist $\Delta f > 0$.

Die Rotation

Wir haben durch Bildung des Gradienten aus einer Funktion $f: U \rightarrow \mathbb{R}$ ein Vektorfeld auf $U \subseteq \mathbb{R}^m$ erzeugt. Können wir jedes Vektorfeld so darstellen?

1. Für $m = 1$ ist das der Fall, denn dann ist $\text{grad}(f) = f'$ und mit dem Hauptsatz der Differential- und Integralrechnung können wir (zumindest wenn $U \subseteq \mathbb{R}$ ein offenes Intervall ist) f bis auf eine additive Konstante aus $\text{grad}(f)$ rekonstruieren.

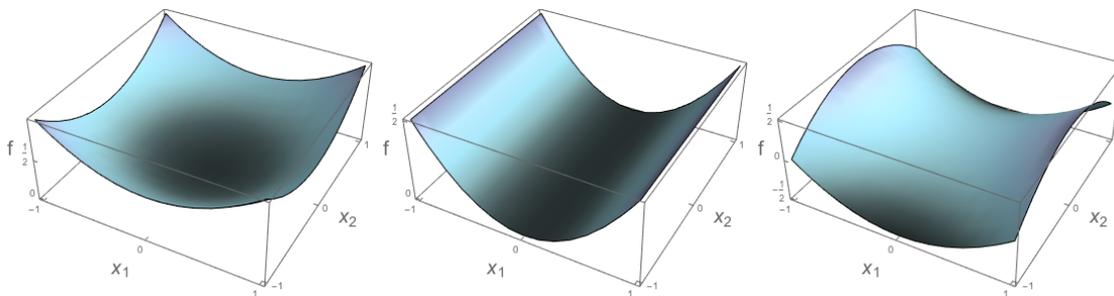


Abbildung 8.9: Graphen quadratischer Funktionen $f(x) = \frac{1}{2} \langle x, \begin{pmatrix} 1 & 0 \\ 0 & \ell \end{pmatrix} x \rangle$ mit $\Delta f = \ell + 1$. Links: $\ell = 1$; Mitte: $\ell = 0$; Rechts: $\ell = -1$

2. Schon für $m = 2$ ist aber nicht jedes Vektorfeld Gradientenvektorfeld. Dies stellen wir mit der *Rotation*, dem Differentialoperator

$$\text{rot} : C^1(U, \mathbb{R}^2) \rightarrow C^0(U, \mathbb{R}) \quad , \quad \text{rot}(v) := \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}$$

fest. Denn

$$\text{rot grad } f = \text{rot} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_2 \partial x_1} = 0,$$

falls f zweimal stetig differenzierbar ist.

Dagegen ist für die Matrix $M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$ die Rotation des linearen Vektorfeldes $v(x) = Mx$ gleich

$$\text{rot } v = m_{21} - m_{12}.$$

Betrachtet man die Tangentialkomponente eines solchen linearen Vektorfeldes entlang einer Kreislinie mit beliebigem Mittelpunkt und Radius, dann ist diese, gemittelt über den Kreis, für $\text{rot}(v) > 0$ im Gegenuhrzeigersinn, für $\text{rot}(v) < 0$ im Uhrzeigersinn orientiert, siehe Abb. 8.10.

3. In $m = 3$ Dimensionen lässt sich ein Differentialoperator definieren, der ebenfalls *Rotation* genannt wird, diesmal aber Vektorfelder auf Vektorfelder abbildet:

$$\text{rot} : C^1(U, \mathbb{R}^3) \rightarrow C^0(U, \mathbb{R}^3) \quad , \quad \text{rot}(v) = \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}.$$

In der Physik schreibt man oft $\nabla \times v$ statt $\text{rot}(v)$.

Wieder gilt $\text{rot grad} = 0$.

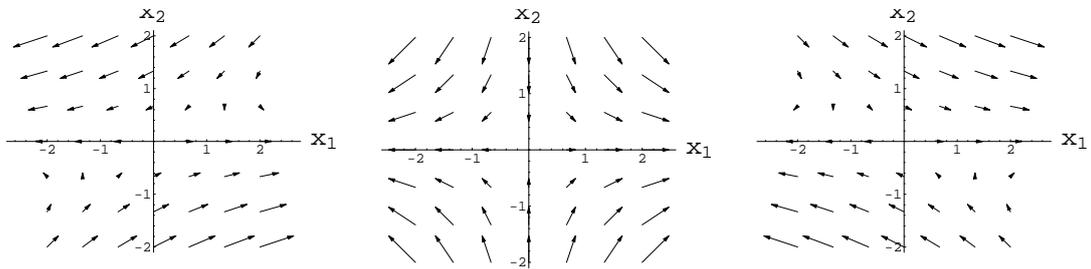


Abbildung 8.10: Divergenzfreie Vektorfelder $v(x) = \begin{pmatrix} 1 & -r \\ 0 & -1 \end{pmatrix}x$ mit $\text{rot}(v) = r$. Links: $r = 2$; Mitte: $r = 0$; Rechts: $r = -2$

8.8 Anwendungen der Kettenregel

Im folgenden Satz betrachten wir ein bestimmtes Integral, bei dem sowohl der Integrand als auch die Integrationsgrenzen von einem Parameter λ abhängen.

8.29 Satz (Differentiation unter dem Integralzeichen) Es sei

$f : \mathbb{R} \times (\lambda_1, \lambda_2) \rightarrow \mathbb{R}$ eine stetige Funktion mit stetiger partieller Ableitung $D_2 f$. $A, B : (\lambda_1, \lambda_2) \rightarrow \mathbb{R}$ seien C^1 -Funktionen. Dann ist

$$F : (\lambda_1, \lambda_2) \rightarrow \mathbb{R} \quad , \quad F(\lambda) := \int_{A(\lambda)}^{B(\lambda)} f(x, \lambda) dx$$

stetig differenzierbar, und

$$F'(\lambda) = f(B(\lambda), \lambda) B'(\lambda) - f(A(\lambda), \lambda) A'(\lambda) + \int_{A(\lambda)}^{B(\lambda)} D_2 f(x, \lambda) dx. \quad (8.7)$$

Beweis: Es sei $\Phi : \mathbb{R} \times \mathbb{R} \times (\lambda_1, \lambda_2) \rightarrow \mathbb{R}$ gegeben durch $\Phi(a, b, \lambda) := \int_a^b f(x, \lambda) dx$.

Wir schreiben $F = \Phi \circ \Psi$ mit

$$\Psi : (\lambda_1, \lambda_2) \rightarrow \mathbb{R} \times \mathbb{R} \times (\lambda_1, \lambda_2) \quad , \quad \Psi(\lambda) := (A(\lambda), B(\lambda), \lambda)$$

und wenden die Kettenregel an.

- Es ergibt sich unter Voraussetzung der totalen Differenzierbarkeit von Φ

$$\begin{aligned} F'(\lambda) &= D\Phi(\Psi(\lambda))D\Psi(\lambda) = & (8.8) \\ & D_1\Phi(A(\lambda), B(\lambda), \lambda) A'(\lambda) + D_2\Phi(A(\lambda), B(\lambda), \lambda) B'(\lambda) + D_3\Phi(A(\lambda), B(\lambda), \lambda). \end{aligned}$$

- Die ersten beiden Summanden ergeben sich nach dem Hauptsatz der Integralrechnung: Die partiellen Ableitungen

$$D_1\Phi(a, b, \lambda) = -f(a, \lambda) \quad \text{und} \quad D_2\Phi(a, b, \lambda) = f(b, \lambda)$$

sind nach Voraussetzung stetig.

• Beim dritten Summanden betrachten wir für vorgegebene Grenzen $a, b \in \mathbb{R}$ die Hilfsfunktion $\tilde{\Phi} : [\tilde{\lambda}_1, \tilde{\lambda}_2] \rightarrow \mathbb{R}$, $\tilde{\Phi}(\lambda) := \Phi(a, b, \lambda)$, wir konzentrieren uns also auf den hier einzig wesentlichen Parameter λ . Dabei setzen wir (um ein *kompaktes* Intervall $[\tilde{\lambda}_1, \tilde{\lambda}_2]$ zu bekommen) $\tilde{\lambda}_1 := \lambda_1 + \Delta\lambda$, $\tilde{\lambda}_2 := \lambda_2 - \Delta\lambda$ mit einem kleinen $\Delta\lambda > 0$. O.B.d.A. setzen wir $b > a$ voraus. Die Einschränkung der stetigen Funktion $D_2f = \frac{\partial f}{\partial \lambda}$ auf das kompakte Rechteck $R := [a, b] \times [\tilde{\lambda}_1, \tilde{\lambda}_2]$ ist (nach Satz 9.21 von [An1]) gleichmäßig stetig, d.h. für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$ mit

$$|D_2f(x, \lambda) - D_2f(x', \lambda')| < \frac{\varepsilon}{b-a} \quad \left((x, \lambda), (x', \lambda') \in R, \|(x, \lambda) - (x', \lambda')\| < \delta \right). \quad (8.9)$$

Wir wollen zeigen, dass für $D\tilde{\Phi}(\lambda) = D \int_a^b f(x, \lambda) dx$ gilt:

$$D \int_a^b f(x, \lambda) dx = \int_a^b D_2f(x, \lambda) dx, \quad (8.10)$$

wir also die Differentiation nach dem Parameter λ unter das Integralzeichen ziehen können. Wir wissen aber noch nicht, ob $\tilde{\Phi}$ überhaupt differenzierbar ist (also ob die linke Seite von (8.10) existiert). Betrachten wir also den Differenzenquotienten, vermindert um die rechte Seite von (8.10):

$$\begin{aligned} \frac{\tilde{\Phi}(\lambda) - \tilde{\Phi}(\lambda')}{\lambda - \lambda'} - \int_a^b D_2f(x, \lambda) dx &= \int_a^b \left[\frac{f(x, \lambda) - f(x, \lambda')}{\lambda - \lambda'} - D_2f(x, \lambda) \right] dx \\ &= \int_a^b \left[D_2f(x, \tilde{\lambda}(x)) - D_2f(x, \lambda) \right] dx \end{aligned} \quad (8.11)$$

für ein $\tilde{\lambda}(x)$ zwischen λ und λ' , nach dem Mittelwertsatz der Differentialrechnung. Ist nun $|\lambda - \lambda'| < \delta$, dann ist nach (8.9) der Betrag von (8.11) kleiner als $\int_a^b \frac{\varepsilon}{b-a} dx = \varepsilon$. Da $\varepsilon > 0$ beliebig ist, haben wir damit (8.10) gezeigt.

Gleichzeitig haben wir auch die Stetigkeit aller partiellen Ableitungen von Φ bewiesen, was nach Satz 8.14 die totale Differenzierbarkeit von Φ impliziert und damit die Formel (8.8) rechtfertigt.

• Die Stetigkeit von F' folgt aus der stetigen Abhängigkeit aller Summanden in (8.7) von λ .
□

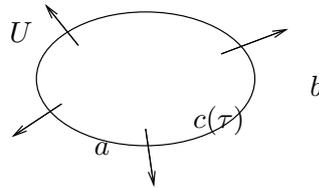
In Kapitel 20 von BLATTER [B] findet man eine etwas allgemeinere Version des Satzes.

8.30 Satz (Mittelwertsatz der Differentialrechnung)

Es sei $f : U \rightarrow \mathbb{R}$ auf der offenen Menge $U \subset \mathbb{R}^m$ differenzierbar, und U enthalte die Strecke S mit den Endpunkten a und b . Dann gibt es $\xi \in S$ mit

$$f(b) - f(a) = \text{grad}f(\xi) \cdot (b - a). \quad (8.12)$$

Beweis: Betrachte $h := f \circ c$ mit der Kurve $c : [0, 1] \rightarrow U$, $t \mapsto (1-t)a + tb$ von a nach b .



Es ist $c'(t) = b - a$ und damit nach der Kettenregel

$$h'(t) = \text{grad}f(c(t)) \cdot (b - a).$$

Anwendung des (eindimensionalen) Mittelwertsatzes ergibt andererseits die Existenz eines $\tau \in [0, 1]$ mit $h(1) - h(0) = h'(\tau)$. Setze $\xi := c(\tau)$. \square

Die Gleichung (8.12) läßt sich als mehrdimensionale Taylor-Formel nullter Ordnung auffassen, was deutlich wird, wenn man sie in der Form $f(b) = f(a) + \text{grad}f(\xi) \cdot (b - a)$ schreibt und den zweiten Summanden als Restglied auffasst.

Das folgende Beispiel zeigt dagegen, dass der Mittelwertsatz nicht auf Abbildungen $f : U \rightarrow \mathbb{R}^n$ mit $n > 1$ verallgemeinert werden kann.

8.31 Beispiel Für $f : \mathbb{R} \rightarrow \mathbb{C}$, $f(x) := e^{ix} = \cos(x) + i \sin(x)$ ist $(2\pi)^{-1} \int_0^{2\pi} f(x) dx = 0$, aber $0 \notin f([0, 2\pi])$. \diamond

9 Der Satz von Taylor im \mathbb{R}^n

9.1 Die Taylor-Formel

In der Taylor-Formel wird eine reelle Funktion $f : D \rightarrow \mathbb{R}$ bei $a \in D \subseteq \mathbb{R}$ durch ein geeignetes Polynom p approximiert und die Differenz $f - p$ abgeschätzt.

Als Vorübung betrachten wir den Fall, dass f selbst ein reelles Polynom, also Element des Vektorraumes $\mathbb{R}[x]$ ist. Dieser besitzt die Basis $(b_k)_{k \in \mathbb{N}_0}$ mit $b_k(x) := x^k$. Wir können also $f(x) = \sum_{k=0}^n a_k x^k$ mit dem Koeffizienten $a_k \in \mathbb{R}$ in der Form $f = \sum_{k=0}^n a_k b_k$ schreiben.

Stattdessen wollen wir für $a \in \mathbb{R}$ die neue Basis $(b_{a,k})_{k \in \mathbb{N}_0}$ von $\mathbb{R}[x]$ mit $b_{a,k}(x) := (x - a)^k$ verwenden, also f in der Form

$$f = \sum_{k=0}^m c_k b_{a,k} \quad \text{mit Koeffizienten } c_k \in \mathbb{R} \quad (9.1)$$

schreiben. Wie bestimmen wir diese? Dazu stellen wir fest, dass die ℓ -te Ableitung des Polynoms $b_{a,k}$ die Form $b_{a,k}^{(\ell)}(x) = \prod_{i=0}^{\ell-1} (k - i) \cdot (x - a)^{k-\ell}$ besitzt, also insbesondere $b_{a,k}^{(\ell)} = 0$ ist für $\ell > k$. Außerdem ist $b_{a,k}^{(\ell)}(a) = \delta_{k,\ell} k!$ mit der Kroneckerschen Deltafunktion δ . Bilden wir also die ℓ -te Ableitung von (9.1), dann ergibt sich

$$f^{(\ell)}(a) = \ell! c_\ell.$$

Daraus folgt für die Koeffizienten in (9.1)

$$c_k = \frac{f^{(k)}(a)}{k!} \quad (k \in \mathbb{N}_0).$$

Dies verallgemeinert sich folgendermaßen:

9.1 Satz (Taylor-Formel) *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $n \in \mathbb{N}_0$ und $a \in I$.*

1. Dann gilt für alle $f \in C^n(I)$

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + o((x-a)^n) \quad (x \rightarrow a). \quad (9.2)$$

2. Für alle $f \in C^{n+1}(I)$ ist das **Restglied**

$$R_{n+1} : I \rightarrow \mathbb{R}, \quad R_{n+1}(x) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$$

von der Form $R_{n+1}(x) = \frac{1}{n!} \int_a^x (x-y)^n f^{(n+1)}(y) dy$.

3. In diesem Fall gibt es für alle $x \in I$ ein $\xi = \xi(x)$ zwischen a und x mit der sog. **Lagrange-Form** des Restgliedes

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1}. \quad (9.3)$$

Beweis:

- Wir beweisen zunächst für $\ell \in \mathbb{N}$ und $f \in C^\ell(I)$ die Restgliedformel

$$R_\ell(x) = \frac{1}{(\ell-1)!} \int_a^x (x-y)^{\ell-1} f^{(\ell)}(y) dy, \quad (9.4)$$

aus der für $\ell := n+1$ sofort Teil 2. des Satzes folgt.

(9.4) reduziert sich für $\ell = 1$ auf die Aussage

$$f(x) - f(a) = \int_a^x f'(y) dy,$$

also den Fundamentalsatz der Differential- und Integralrechnung.

Für $m \in \{1, \dots, \ell-1\}$ ist $f \in C^m(I) \subset C^\ell(I)$. Also gilt nach Induktionsvoraussetzung

$$R_m(x) = \frac{1}{(m-1)!} \int_a^x (x-y)^{m-1} f^{(m)}(y) dy.$$

Mit partieller Integration ergibt sich daraus

$$\begin{aligned} R_m(x) &= -\frac{1}{m!}(x-y)^m f^{(m)}(y)|_a^x + \frac{1}{m!} \int_a^x (x-y)^m f^{(m+1)}(y) dy \\ &= \frac{f^{(m)}(a)}{m!}(x-a)^m + \frac{1}{m!} \int_a^x (x-y)^m f^{(m+1)}(y) dy. \end{aligned}$$

Also ist $R_{m+1}(x) = \frac{1}{m!} \int_a^x (x-y)^m f^{(m+1)}(y) dy$, woraus sich durch Induktion nach m (9.4) ergibt.

- Für $x \geq a$ und

$$f_{\min}^{(n+1)} := \inf\{f^{(n+1)}(y) \mid y \in [a, x]\}, \quad f_{\max}^{(n+1)} := \sup\{f^{(n+1)}(y) \mid y \in [a, x]\}$$

gilt

$$\frac{1}{n!} \int_a^x (x-y)^n f_{\min}^{(n+1)} dy \leq R_{n+1}(x) \leq \frac{1}{n!} \int_a^x (x-y)^n f_{\max}^{(n+1)} dy,$$

oder nach Auswertung der Integrale

$$\frac{f_{\min}^{(n+1)}}{(n+1)!}(x-a)^{n+1} \leq R_{n+1}(x) \leq \frac{f_{\max}^{(n+1)}}{(n+1)!}(x-a)^{n+1}.$$

Da $f^{(n+1)}$ stetig ist, gibt es ein $\xi \in [a, x]$, das (9.3) erfüllt. Der Fall $x \leq a$ lässt sich analog behandeln. Damit ist Teil 3. des Satzes bewiesen.

- Unter der Voraussetzung $f \in C^n(I)$ von 1. ist damit

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

Wegen der Stetigkeit von $f^{(n)} : I \rightarrow \mathbb{R}$ und $\lim_{x \rightarrow a} \xi(x) = a$ ist aber

$$f^{(n)}(\xi(x)) = f^{(n)}(a) + o(1) \quad (x \rightarrow a),$$

also $R_n(x) = \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^n)$. Daher ist $R_{n+1}(x) = o((x-a)^n)$, was Teil 1. des Satzes beweist. \square

9.2 Bemerkungen

1. Manchmal nennt man Teil 1. des Satzes 9.1 die **qualitative**, Teil 2. die **quantitative** Taylor-Formel.
2. Aus Letzterer folgt die Aussage

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!}(x-a)^k + \mathcal{O}((x-a)^{n+1}).$$

Diese ist präziser als (9.2), denn

$$(x-a)^{n+1} = o((x-a)^n) \quad (x \rightarrow a),$$

aber nicht umgekehrt. Man muss aber bedenken, dass in Teil 2. $(n+1)$ -malige stetige Differenzierbarkeit von f verlangt wird, während Teil 1. nur n -fache stetige Differenzierbarkeit voraussetzt. \diamond

Setzen wir für $f \in C^n(I)$ und $a \in I$

$$T_{a,n}f \in \mathbb{R}[x] \quad , \quad T_{a,n}f(x) := \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k,$$

dann ist dieses *Taylor-Polynom n -ter Ordnung von f* ein Polynom vom Grad

$$\deg(T_{a,n}f) \leq n.$$

Wir erhalten so eine lineare Abbildung

$$T_{a,n} : C^n(I) \rightarrow \mathbb{R}[x].$$

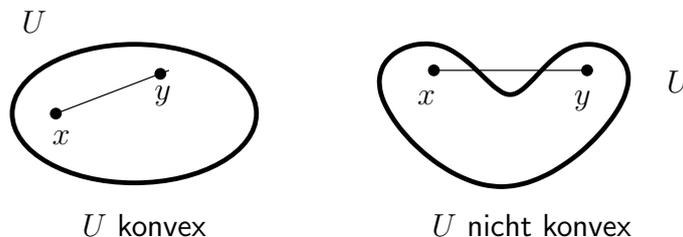
9.2 Die Taylor-Formel in m Dimensionen

Wir führen die mehrdimensionale Situation auf die eindimensionale zurück, indem wir für $x \in U$ die Funktion auf der Strecke zwischen a und x untersuchen, d.h.

$$h(t) := f(a + t(x-a)) \quad (t \in [0, 1]).$$

Dazu muss der Definitionsbereich U von f natürlich das Bild $h([0, 1])$ dieser Kurve enthalten.

9.3 Definition $U \subset \mathbb{R}^m$ heißt **konvex**, wenn für alle $x, y \in U$ auch die Strecke $\{(1-t)x + ty \mid t \in [0, 1]\}$ zwischen x und y in U enthalten ist.



9.4 Satz (Taylor)

Für $f \in C^{k+1}(U)$, $U \subset \mathbb{R}^m$ offen und konvex, $a, x \in U$ und $v := x - a$ gilt

$$f(x) = f(a) + \partial_v f(a) + \frac{\partial_v^{(2)} f(a)}{2!} + \dots + \frac{\partial_v^{(k)} f(a)}{k!} + R_{k+1}$$

mit $\partial_v^{(l)} := \underbrace{\partial_v \dots \partial_v}_{l\text{-mal}}$ und dem **Lagrange-Restglied**

$$R_{k+1} := \frac{\partial_v^{(k+1)} f(a + \theta v)}{(k+1)!} \quad , \quad \theta \in [0, 1] \quad \text{geeignet} \quad .$$

Beweis: Für $l = 1, \dots, k + 1$ ist

$$\frac{d^l}{dt^l} h(t) = \frac{d^l}{dt^l} f(a + tv) = \partial_v^{(l)} f(a + tv).$$

Daher folgt die Formel aus der (eindimensionalen) Taylor-Formel für h :

$$h(1) = \sum_{l=0}^k \frac{h^{(l)}(0)}{l!} 1^l + R_{k+1}. \quad \diamond$$

In dieser Form des Satzes, der die Richtungsableitung ∂_v in Richtung v benutzt, sieht man nicht direkt, dass f durch ein Polynom vom Grad $\leq k$ in x_1, \dots, x_m angenähert wird. Um dies zu sehen, muss man die *Binomische Formel* verallgemeinern:

9.5 Definition (Multiindices)

- Ein Element $p = (p_1, \dots, p_m) \in \mathbb{N}_0^m$ heißt **Multiindex der Ordnung**

$$|p| := p_1 + \dots + p_m.$$

- Die **Fakultät** des Multiindex ist $p! := p_1! p_2! \dots p_m!$.
- Für $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ schreibt man abkürzend die Monome in der Form

$$x^p := x_1^{p_1} \cdot \dots \cdot x_m^{p_m}.$$

9.6 Beispiel (Multiindex-Schreibweise)

Für $p := (2, 0, 1, 4)$ ist $|p| = 7$, $p! = 2!0!1!4! = 48$ und $x^p = x_1^2 x_3 x_4^4$. \diamond

Die Ordnung $|p|$ des Multiindex ergibt den *Grad* des Monoms x^p . Unter einem reellen Polynom q in m Variablen versteht man eine Funktion

$$q(x) = \sum_{p \in \mathbb{N}_0^m} c_p x^p$$

mit *Koeffizienten* $c_p \in \mathbb{R}$, die bis auf endlich viele gleich Null sind. Der *Grad* des Polynoms q ist der höchste Grad $|p|$ seiner Monome x^p mit $c_p \neq 0$. Dem *Nullpolynom* in m Variablen ordnet man Grad $-\infty$ zu.

9.7 Satz (Polynomial-Formel) Für $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ und $k \in \mathbb{N}$ ist

$$(x_1 + \dots + x_m)^k = k! \sum_{|p|=k} \frac{x^p}{p!}.$$

Beweis: Durch Induktion in der Zahl m der Variablen.

- $m = 1$: trivial
- $m = 2$: Binomische Formel:

$$(x_1 + x_2)^k = \sum_{p_1=0}^k \frac{k!}{p_1!(k-p_1)!} x_1^{p_1} x_2^{k-p_1} = k! \sum_{p_1, p_2, p_1+p_2=k} \frac{x_1^{p_1} x_2^{p_2}}{p_1! p_2!}.$$

- $m \mapsto m + 1$: $x := (x_1, \dots, x_m)$, $y := (x_1, \dots, x_m, x_{m+1})$

$$\begin{aligned} \underbrace{(x_1 + \dots + x_m)}_s + x_{m+1})^k &= (s + x_{m+1})^k = \sum_{l=0}^k \frac{k!}{l!(k-l)!} s^l x_{m+1}^{k-l} \\ &= k! \sum_{l=0}^k \frac{1}{l!(k-l)!} \left(l! \sum_{|p|=l} \frac{x^p}{p!} \right) x_{m+1}^{k-l} \\ &= k! \sum_{l=0}^k \sum_{|p|=l} \frac{x^p x_{m+1}^{k-l}}{p!(k-l)!} = k! \sum_{|q|=k} \frac{y^q}{q!} \end{aligned}$$

mit dem Multiindex $q := (p_1, \dots, p_m, l)$ der Länge $m + 1$. □

Mit der Schreibweise $D_l = \frac{\partial}{\partial x_l}$, $l = 1, \dots, m$ für die l -te partielle Ableitung und $D^p := \frac{\partial^{|p|}}{\partial x_1^{p_1} \dots \partial x_m^{p_m}}$ für den Multiindex $p = (p_1, \dots, p_m)$ ist die Richtungsableitung in Richtung $v \in \mathbb{R}^m$ gleich $\partial_v = v_1 D_1 + \dots + v_m D_m$ und nach Satz²⁴ 9.7

$$\partial_v^{(l)} = (v_1 D_1 + \dots + v_m D_m)^l = l! \sum_{|p|=l} \frac{v^p D^p}{p!}.$$

Damit ergibt sich die Gestalt

$$f(a + v) = \sum_{|p| \leq k} \frac{D^p f(a)}{p!} v^p + R_{k+1}$$

der Taylor-Formel, mit $R_{k+1} = \sum_{|p|=k+1} \frac{D^p f(a+\theta v)}{p!} v^p$ für ein $0 \leq \theta \leq 1$.

Wie im Eindimensionalen bezeichnet man mit $T_{a,k}(v) = \sum_{|p| \leq k} \frac{D^p f(a)}{p!} v^p$ das *Taylor-Polynom k -ter Ordnung am Entwicklungspunkt a* .

9.8 Beispiel (Taylor-Polynom)

Für $m = 2$ Variablen und $k = 3$ Ableitungen ergibt sich

$$\begin{aligned} f(a + v) &= f(a) + \\ & f_{x_1}(a)v_1 + f_{x_2}(a)v_2 + \\ & \frac{1}{2} f_{x_1 x_1}(a)v_1^2 + f_{x_1 x_2}(a)v_1 v_2 + \frac{1}{2} f_{x_2 x_2}(a)v_2^2 + \\ & \frac{1}{6} f_{x_1 x_1 x_1}(a)v_1^3 + \frac{1}{2} f_{x_1 x_1 x_2}(a)v_1^2 v_2 + \frac{1}{2} f_{x_1 x_2 x_2}(a)v_1 v_2^2 + \frac{1}{6} f_{x_2 x_2 x_2}(a)v_2^3 \\ & + R_4. \end{aligned}$$

²⁴**Vorsicht:** Ist dagegen v nicht ein fester Vektor, sondern ein *Vektorfeld*, d.h. $v : U \rightarrow \mathbb{R}^m$, dann gilt diese Formel nicht, weil dann i.A. $D_k(v_l f) = v_l D_k f + f D_k v_l \neq v_l D_k f$ ist!

9.3 Die Hesse-Matrix einer reellen Funktion

9.9 Definition

Für $U \subset \mathbb{R}^m$ offen, $f \in C^2(U)$ und $x \in U$ heißt die symmetrische Matrix

$$H_f(x) := \begin{pmatrix} f_{x_1 x_1}(x) & \dots & f_{x_1 x_m}(x) \\ \vdots & & \vdots \\ f_{x_m x_1}(x) & \dots & f_{x_m x_m}(x) \end{pmatrix}$$

die **Hesse-Matrix** von f am Punkt x . Man schreibt auch $\text{Hess } f := H_f$.

Damit ist für $f \in C^2(U)$ nach der qualitativen Taylor-Formel

$$\boxed{f(a + v) = f(a) + v \cdot \nabla f(a) + \frac{1}{2}v \cdot H_f(a)v + R_3}$$

mit $R_3(v) = o(\|v\|^2)$, denn $f(a + v) = f(a) + v \cdot \nabla f(a) + \frac{1}{2}v \cdot H_f(a + \theta v)v$ und $H_f(a + \theta v) = H_f(a) + o(\|v\|)$.

9.10 Beispiel (Hesse-Matrix) Für $f(x) := \cos(x_1) \cos(x_2)$ (siehe Beispiel 8.18) ist

$$\nabla f(x) = \begin{pmatrix} -\sin(x_1) \cos(x_2) \\ -\cos(x_1) \sin(x_2) \end{pmatrix} \quad \text{und} \quad H_f(x) = \begin{pmatrix} -\cos(x_1) \cos(x_2) & \sin(x_1) \sin(x_2) \\ \sin(x_1) \sin(x_2) & -\cos(x_1) \cos(x_2) \end{pmatrix}.$$

Für den Entwicklungspunkt $a := \begin{pmatrix} \pi/2 \\ -\pi/2 \end{pmatrix}$, also $\nabla f(a) = 0$, ergibt sich als Taylor-Polynom zweiter Ordnung von f bei a

$$T_{a,2}(v) = -v_1 v_2 = -(x_1 - \pi/2) \cdot (x_2 + \pi/2).$$

An diesem Punkt besitzt der Graph von f die Form eines Sattels. ◇

Allgemein wird der Graph von f bei a durch die Fläche zweiter Ordnung

$$\{(x, f(a) + (x - a) \cdot \nabla f(a) + \frac{1}{2}(x - a) \cdot H_f(a)(x - a)) \mid x \in U\}$$

genauer als durch die Tangentialebene

$$\{(x, f(a) + (x - a) \cdot \nabla f(a)) \mid x \in \mathbb{R}^m\},$$

approximiert. Dies ist besonders wichtig, um festzustellen, ob ein Extremwert der Funktion vorliegt.

9.4 Extremalstellen

9.11 Definition

Eine Funktion $f : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^m$ (nicht notwendig offen), besitzt

- ein **lokales Maximum** bzw. ein **lokales Minimum** $x \in U$, wenn eine Umgebung $V \subset U$ von x in U existiert, für die

$$f(x) \geq f(y) \quad \text{bzw.} \quad f(x) \leq f(y) \quad (y \in V).$$

- Ein **lokales Extremum** ist ein lokales Maximum oder Minimum.
- ein **(globales) Maximum** bzw. **(globales) Minimum** $x \in U$, wenn

$$f(x) \geq f(y) \quad \text{bzw.} \quad f(x) \leq f(y) \quad (y \in U).$$

- Ist sogar $f(y) < f(x)$ für alle $y \in V \setminus \{x\}$, dann heißt x **isoliertes lokales Maximum**, etc.

Genauer spricht man statt von Minima bzw. Maxima x von Minimal- bzw. Maximalstellen, um sie von den den Minimal- bzw. Maximalwerten $f(x)$ zu unterscheiden.

Wir erinnern uns, dass schon in einer Dimension das Verschwinden der ersten Ableitung weder notwendig noch hinreichend für das Vorliegen eines lokalen Extremums ist:

9.12 Beispiel (Beziehung zwischen Extremalstellen und Ableitung)

1. $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$ besitzt bei 0 ein lokales Minimum und $f'(0) = 0$.
2. $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^3$ besitzt bei 0 kein lokales Extremum, aber $f'(0) = 0$.
3. $f : [0, 1] \rightarrow \mathbb{R}$, $x \mapsto x^2$, besitzt bei $x = 1$ ein lokales Maximum, aber $f'(1) \neq 0$. ◇

Aus der reellen Kurvendiskussion wissen wir aus der Schule, dass eine auf einem offenen Intervall U definierte Funktion f in $x \in U$ sicher eine Extremalstelle hat, wenn $f'(x) = 0$ und $f''(x) \neq 0$ gilt. Dabei ist das Verschwinden der 1. Ableitung notwendig für das Vorliegen eines Extremums.

Dies werden wir nun auf Funktionen von m Variablen verallgemeinern. Ganz allgemein, also für vektorwertige Funktionen ist der folgende Begriff die geeignete Verallgemeinerung der Bedingung (nicht) verschwindender Ableitung:

9.13 Definition

Für $U \subseteq \mathbb{R}^m$ offen und $f \in C^1(U, \mathbb{R}^n)$ mit $n \leq m$ heißt $x \in U$ **regulärer Punkt**, wenn der Rang der Jacobi-Matrix von f maximal ist, d.h. wenn

$$\text{Rang}(Df(x)) = n.$$

Sonst heißt x **kritischer** (oder **singulärer**) **Punkt**.

9.14 Satz Ist $U \subseteq \mathbb{R}^m$ offen und hat $f \in C^1(U)$ bei $x \in U$ eine lokale Extremalstelle, dann ist $\text{grad}f(x) = 0$.

Beweis: Sei statt dessen $v := \text{grad}f(x) \neq 0$. Dann ist nach der Taylor-Formel

$$f(x + \lambda v) = f(x) + \lambda Df(x)(v) + o(\|\lambda v\|) = f(x) + \lambda \langle v, v \rangle + o(\lambda),$$

also wegen $\langle v, v \rangle > 0$ für $\varepsilon > 0$ geeignet:

$$\begin{aligned} f(x + \lambda v) &> f(x) && (0 < \lambda < \varepsilon) \\ f(x + \lambda v) &< f(x) && (-\varepsilon < \lambda < 0). \end{aligned}$$

Widerspruch! □

Satz 9.14 besagt also, dass $x \in U$ nur dann eine lokale Extremalstelle sein kann, wenn x kritischer Punkt ist (also $\nabla f(x) = 0$ ist).

Wie das Beispiel 9.12.3 zeigt, wird Satz 9.14 falsch, wenn man nicht die Offenheit des Definitionsbereiches U verlangt.

Wie Beispiel 9.12.2 zeigt, ist die Aussage des Satzes 9.14 auch nicht umkehrbar. Allerdings können wir mit Hilfe der Hesse-Matrix ein genaueres notwendiges Kriterium und auch hinreichende Kriterien für das Vorliegen eines lokalen Extremums formulieren.

9.15 Definition Eine symmetrische Matrix $A = A^T \in \text{Mat}(m, \mathbb{R})$ heißt

- **positiv definit** ($A > 0$), wenn $\langle x, Ax \rangle > 0$ ($x \in \mathbb{R}^m \setminus \{0\}$),
- **negativ definit** ($A < 0$), wenn $\langle x, Ax \rangle < 0$ ($x \in \mathbb{R}^m \setminus \{0\}$),
- **definit**, wenn A positiv definit oder negativ definit ist,
- **indefinit**, wenn es $x, y \in \mathbb{R}^m$ gibt mit $\langle x, Ax \rangle > 0$ und $\langle y, Ay \rangle < 0$,
- **positiv semidefinit** ($A \geq 0$), wenn $\langle x, Ax \rangle \geq 0$ ($x \in \mathbb{R}^m$),
- **negativ semidefinit** ($A \leq 0$), wenn $\langle x, Ax \rangle \leq 0$ ($x \in \mathbb{R}^m$),
- **semidefinit**, wenn A positiv semidefinit oder negativ semidefinit ist.

Der Gegensatz zu *indefinit* ist also *semidefinit*.

Für eine indefinite 2×2 -Matrix A ist der Graph der quadratischen Form $x \mapsto \langle x, Ax \rangle$ von der Form eines *Sattels*, siehe Abbildung 8.3.

9.16 Beispiel (Definitheit symmetrischer Matrizen)

1. Die Diagonalmatrix $A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_m \end{pmatrix}$ ist genau dann positiv (negativ) definit, wenn alle Eigenwerte $a_i > 0$ ($a_i < 0$) sind, und indefinit, wenn es Eigenwerte $a_i > 0 > a_j$ unterschiedlichen Vorzeichens gibt. A ist genau dann positiv (bzw. negativ) semidefinit, wenn für alle i gilt: $a_i \geq 0$ bzw. ($a_i \leq 0$).

2. Eine 2×2 -Matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ ist genau dann positiv definit, wenn $a > 0$ und $\det A = ac - b^2 > 0$ sind.

Beweis: Für $x := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist $\langle x, Ax \rangle = a$, sodass $a > 0$ eine notwendige Bedingung ist.

Ist $a > 0$, brauchen wir nur noch $x \in \mathbb{R}^2$ mit $x_2 \neq 0$ zu betrachten. Mit $y := x_1/x_2$ ist dann

$$\langle x, Ax \rangle = \underbrace{x_2^2}_{>0} \underbrace{(ay^2 + 2by + c)}_{>0 \text{ für } |y| \text{ groß.}}$$

Der Klammerausdruck hat genau dann reelle Nullstellen $y_{1/2} = -\frac{b}{a} \pm \sqrt{\frac{b^2}{a^2} - \frac{c}{a}}$, wenn $ac - b^2 \leq 0$. \square

So ist $A := \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ positiv definit, denn $a = 2$ und $\det A = 3$.

Dagegen ist $A := \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ indefinit, denn für $x := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $y := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist $\langle x, Ax \rangle = 6$ und $\langle y, Ay \rangle = -2$. \diamond

Wie im Beispiel 9.16.1 kommt es auch bei beliebigen symmetrischen Matrizen nur auf das Vorzeichen der Eigenwerte an. Zwar ist die exakte Berechnung der Eigenwerte für große Matrixdimension aufwendig oder unmöglich. Wie in Beispiel 9.16.2 kann man aber ihre Vorzeichen bestimmen, ohne sie zu berechnen. Allgemein gilt nämlich der

9.17 Satz (Sylvester) Eine symmetrische Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mm} \end{pmatrix} \in \text{Mat}(m, \mathbb{R})$$

ist genau dann positiv definit, wenn gilt:

$$\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0 \quad (k = 1, \dots, m).$$

Beweis: Wie aus der Linearen Algebra bekannt, lässt sich A auf Diagonalform $D := O^{-1}AO$ (mit O orthogonal) bringen. Man zeigt nun, dass die Diagonalmatrix $D = \text{diag}(d_1, \dots, d_m)$ genau dann nur positive d_i besitzt, wenn obige Bedingung an A erfüllt ist. Genauer in Lehrbüchern der Linearen Algebra, z.B. FISCHER [Fi], Kapitel 5.7. \square

Hinreichend ist nun folgender Extremstellentest:

9.18 Satz *Es sei $U \subseteq \mathbb{R}^m$ offen, $f \in C^2(U)$ und ein $x \in U$ kritischer Punkt. Dann ist x für*

- $H_f(x)$ positiv definit eine isolierte Minimalstelle,
- $H_f(x)$ negativ definit eine isolierte Maximalstelle,
- $H_f(x)$ indefinit keine Extremalstelle.

Beweis:

- Es sei $H_f(x)$ positiv definit, und

$$c := \frac{1}{2} \min_{\substack{v \in \mathbb{R}^m \\ \|v\|=1}} \langle v, H_f(x)v \rangle, \text{ also } c > 0.$$

Dann ist $\frac{1}{2} \langle y, H_f(x)y \rangle \geq c \|y\|^2$.

Nach der Taylor-Formel ist wegen $\text{grad}f(x) = 0$

$$\begin{aligned} f(x+y) &= f(x) + \frac{1}{2} \langle y, H_f(x)y \rangle + o(\|y\|^2) \\ &\geq f(x) + c\|y\|^2 + o(\|y\|^2) \\ &\geq f(x) + \frac{c}{2}\|y\|^2 \quad (\|y\| < \varepsilon) \end{aligned}$$

für ein geeignetes $\varepsilon > 0$. Also ist x eine isolierte Minimalstelle.

- Ist $H_f(x)$ negativ definit, dann ist $H_g(x)$ für $g := -f$ positiv definit. Beweis analog.
- Ist $H_f(x)$ indefinit, dann existieren $c > 0$ und Richtungen $v, w \in \mathbb{R}^m$, $\|v\| = \|w\| = 1$ mit $\langle v, H_f(x)v \rangle \geq c$ und $\langle w, H_f(x)w \rangle \leq -c$, also

$$f(x + \lambda v) \geq f(x) + \frac{c}{2}|\lambda|^2 \quad \text{bzw.} \quad f(x + \lambda w) \leq f(x) - \frac{c}{2}|\lambda|^2$$

falls $|\lambda| < \varepsilon$. Damit ist x weder Minimum noch Maximum. \square

9.19 Beispiel Es sollen die kritischen Punkte und Extrema von $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = \cos(x_1) \cos(x_2)$ aus Bsp. 8.18 bestimmt werden.

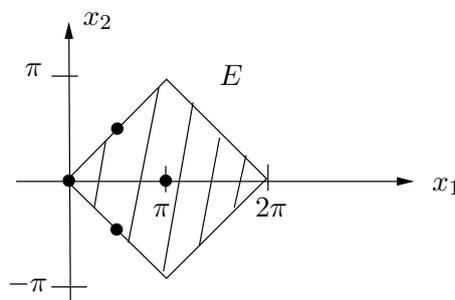
Dazu kann die Periodizität von f ausgenutzt werden. Da der Cosinus 2π -periodisch ist, besitzt f in beiden Variablen diese Periode. Es gilt aber mehr, nämlich dass $f(x + \ell) = f(x)$ für $\ell \in G := \{n_1(\frac{\pi}{\pi}) + n_2(\frac{\pi}{-\pi}) \mid n_i \in \mathbb{Z}\}$, denn $\cos(x \pm \pi) = -\cos(x)$.

Wir müssen uns also nur um die kritischen Punkte im von den Basisvektoren $\begin{pmatrix} \pi \\ \pi \end{pmatrix}$ und $\begin{pmatrix} \pi \\ -\pi/2 \end{pmatrix}$ des Gitters G aufgespannten Elementarbereiches

$$E := \left\{ c_1 \begin{pmatrix} \pi \\ \pi \end{pmatrix} + c_2 \begin{pmatrix} \pi \\ -\pi \end{pmatrix} \mid 0 \leq c_1, c_2 < 1 \right\}$$

kümmern. Die Bedingung $\nabla f(x) = 0$ ist gleichbedeutend mit

$$\sin(x_1) \cos(x_2) = \cos(x_1) \sin(x_2) = 0.$$



In E finden wir genau vier kritische Punkte:

1. Für $a := \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in E$ ist $\nabla f(a) = 0$ und $H_f(a) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ negativ definit, a ist also isoliertes Maximum, mit $f(a) = 1$.
2. $a := \begin{pmatrix} \pi \\ 0 \end{pmatrix} \in E$: $\nabla f(a) = 0$, $H_f(a) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ positiv definit, a ist also isoliertes Minimum, mit $f(a) = -1$.
3. $a_{\pm} = \begin{pmatrix} \pi/2 \\ \pm\pi/2 \end{pmatrix}$, $\nabla f(a_{\pm}) = 0$, $H_f(a) = \begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}$ indefinit, a_{\pm} ist also kein Extremum, sondern ein Sattel.

Die isolierten Maximalstellen von f bilden also das Gitter G , die isolierten Minimalstellen das verschobene Gitter $G + \begin{pmatrix} \pi \\ 0 \end{pmatrix}$. Sattelpunkte finden wir in $G + \begin{pmatrix} \pi/2 \\ \pi/2 \end{pmatrix}$ und in $G + \begin{pmatrix} \pi/2 \\ -\pi/2 \end{pmatrix}$. \diamond

10 Implizite Funktionen

Man kann Kurven und Flächen²⁵ im Raum *implizit* als Nullstellenmenge einer Funktion oder *explizit* durch Parametrisierung angeben.

10.1 Beispiele

²⁵Allgemeiner: so genannte *Untermannigfaltigkeiten* des \mathbb{R}^m , siehe Definition 10.18

- Die p -Sphäre $S^p \subset \mathbb{R}^m$ mit $p := m - 1$ ist implizit durch $f(x) = 0$ mit $f : \mathbb{R}^m \rightarrow \mathbb{R}, x \mapsto \|x\|^2 - 1$ gegeben.
- Für $S^2 \hookrightarrow \mathbb{R}^3$ ist eine explizite Parametrisierung durch

$$g : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times [0, 2\pi) \rightarrow \mathbb{R}^3, \quad g(\theta, \varphi) := \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ \sin \theta \end{pmatrix}$$

gegeben. Wir können aber auch die "Nordhalbkugel" $\{x \in S^2 \mid x_3 > 0\}$ durch

$$g : U \rightarrow \mathbb{R}^3, \quad (x_1, x_2) \mapsto \left(x_1, x_2, \sqrt{1 - x_1^2 - x_2^2}\right)$$

mit $U := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 < 1\}$ parametrisieren, also die ersten beiden Koordinaten des Punktes $x \in S^2$ benutzen.

- Beide Darstellungen haben ihre Vor- und Nachteile. Will man beispielsweise die Sphäre von einem Computer zeichnen lassen, bietet sich die explizite Parameterdarstellung an. Andererseits ist der geometrische Gehalt (Menge der Punkte, die Abstand 1 zum Ursprung haben) in der impliziten Form offensichtlicher. \diamond

Wie gehen wir nun vor, wenn wir eine implizit gegebene Kurve oder Fläche parametrisch darstellen wollen? Es sei z.B. $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ mit $U \subseteq \mathbb{R}^m$ offen und $n \leq m$, und wir wollen die *Nullstellenmenge*

$$f^{-1}(0) \subseteq U$$

parametrisieren. Heuristisch wird man erwarten, dass dazu $m - n$ Parameter nötig sind, denn die Nullstellenmenge erfüllt ja die n Gleichungen

$$f_1(x) = 0, \dots, f_n(x) = 0.$$

Dies stimmt schon im Fall einer linearen Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ nur, wenn die f_1, \dots, f_n linear unabhängig sind.

10.2 Beispiel $f : \mathbb{R}^m \rightarrow \mathbb{R}^n, f(x) = Ax - b$ mit $A \in \text{Mat}(n \times m, \mathbb{R}), b \in \mathbb{R}^n$. Die Nullstellenmenge $f^{-1}(0)$ ist hier gleich der Lösungsmenge des inhomogenen linearen Gleichungssystems

$$Ax = b. \tag{10.1}$$

Nun ist das Bild $f(\mathbb{R}^m)$ der affine Unterraum von \mathbb{R}^n , der den Punkt $-b$ enthält und von den Spaltenvektoren von A aufgespannt wird. Die affine Dimension von $f(\mathbb{R}^m)$ ist also gleich $\text{rang}(A) \leq \min(m, n)$.

- Ist nun $n > \text{rang}(A)$ (was insbesondere für $n > m$ gilt), dann besitzt (10.1) für typische $b \in \mathbb{R}^n$ gar keine Lösung, f also keine Nullstelle.
- Ist dagegen $n \leq m$ und $\text{rang}(A)$ maximal (also gleich n), dann ist die Lösungsmenge von (10.1) ein affiner Unterraum des \mathbb{R}^m von der Dimension $p := m - n$. Wir benötigen also p Parameter zur Parametrisierung von $f^{-1}(0)$. \diamond

Beispiel 10.1 bestärkt uns in der Vermutung, dass auch für allgemeine C^1 -Abbildungen f unter der Voraussetzung maximalen Ranges von $Df(x)$ (also der Linearisierung von f bei x) eine solche Parametrisierung durch $m - n$ der Koordinaten x_i lokal, d.h. in einer Umgebung von $x \in f^{-1}(0)$ schon möglich ist. Man muss nur die richtigen x_i wählen. Z.B. wäre es ungeschickt, in der Nähe des "Ostpols" $x := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in S^2$ zu versuchen, durch x_1 und x_2 zu parametrisieren, denn dort wird das Argument von $x_3 = \pm\sqrt{1 - x_1^2 - x_2^2}$ Null, und die Ableitung divergiert. Auf Vektorräumen M sind Nullstellen- und Fixpunktprobleme äquivalent, denn die Nullstellenmenge $f^{-1}(0) \subseteq M$ der Abbildung $f : M \rightarrow M$ ist gleich der Fixpunktmenge

$$\{x \in M \mid g(x) = x\} \quad \text{von} \quad g : M \rightarrow M \quad \text{mit} \quad g(x) := f(x) + x.$$

Zwei Näherungsverfahren werden in den verschiedensten Situationen verwandt:

- Die Iteration einer kontrahierenden Abbildung g . Damit soll ihr Fixpunkt gefunden werden.
- Das Newton-Verfahren. Dabei wird die Nullstelle der Tangente von f bei x als Näherung für die Nullstelle von f benutzt.

10.1 Das Newton-Verfahren

Manchmal kann man bei der Nullstellensuche aber die *Newton-Iteration*

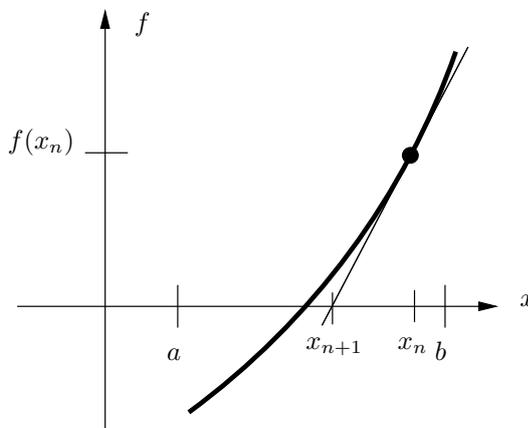
benutzen, die dann in m *quadratische Konvergenzgeschwindigkeit* liefert:

Es sei $f \in C^2([a, b], \mathbb{R})$, also eine zweimal stetig differenzierbare Funktion auf dem Intervall $[a, b]$, und

$$f'(x) \neq 0 \quad (x \in [a, b]).$$

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

ist der Schnittpunkt der Tangente an den Graphen von f bei $(x_n, f(x_n))$ mit der x -Achse, siehe nebenstehende Abbildung.



10.3 Satz (Newtoniteration) Ist $f(x^*) = 0$ für $x^* \in [a, b]$, dann gilt

$$|x_{n+1} - x^*| \leq |x_n - x^*|^2 \cdot M \quad \text{mit} \quad M := \frac{\max_{x \in [a, b]} |f''(x)|}{\min_{x \in [a, b]} |f'(x)|}. \quad (10.2)$$

Beweis: Für ein ξ im Intervall zwischen x_n und x^* gilt

$$\begin{aligned} |x_{n+1} - x^*| &= \left| x_n - \frac{f(x_n) - \overbrace{f(x^*)}^0}{f'(x_n)} - x^* \right| \\ &= \left| (x_n - x^*) \cdot \left(1 - \underbrace{\frac{f(x_n) - f(x^*)}{x_n - x^*}}_{=f'(\xi) \text{ (Mittelwertsatz)}} \cdot \frac{1}{f'(x_n)} \right) \right| \\ &= |x_n - x^*| \cdot \left| \frac{f'(x_n) - f'(\xi)}{f'(x_n)} \right| \\ &\leq |x_n - x^*|^2 \cdot \left| \frac{f'(x_n) - f'(\xi)}{x_n - \xi} \cdot \frac{1}{f'(x_n)} \right| \\ &= |x_n - x^*|^2 \cdot \frac{|f''(\eta)|}{|f'(x_n)|}, \end{aligned}$$

wobei nach dem Mittelwertsatz η geeignet zwischen x_n und ξ gewählt wird. \square

10.4 Beispiel Für $f(x) := x^2 - 2$ konvergiert das Newton-Verfahren $x_{n+1} := \frac{x_n}{2} + \frac{1}{x_n}$ viel schneller gegen die Nullstelle $\sqrt{2}$ als das Fixpunktverfahren von

Beispiel 7.16. Da f stetig und $f(1) < 0 < f(2)$ ist, befindet sich im Intervall $[a, b] := [1, 2]$ eine Nullstelle. Dort ist die Konstante M aus Satz 10.3 gleich 1.

In der folgenden gerundeten Tabelle ist als Fehlerschranke für x_{n+1} der Wert $(x_n - x^*)^2$ mit $x^* = \sqrt{2}$ eingetragen. In numerischen Anwendungen kann dagegen die Fehlerschranke (10.2) des Newton-Verfahrens nicht so einfach gewonnen werden wie beim Fixpunktverfahren, da die Nullstelle x^* nicht bekannt ist.

n	x_n	Fehler	Fehlerschranke
0	2	0.585786	
1	1.5000000000	0.085786	0.343145
2	1.4166666666	0.002453	0.007359
3	1.4142156863	0.000002	0.000006
4	1.4142135624	0.000000	0.000000

Ist der Startwert x_0 allerdings zu weit von der gesuchten Nullstelle x^* entfernt, braucht das Newton-Verfahren nicht zu konvergieren. Insbesondere ist Vorsicht geboten, wenn zwischen x_0 und x^* ein Extremum von f liegt.

10.5 Beispiel $f(x) := (x^2 - 1)(x^2 + (0.4)^2)$

Dieses Polynom besitzt die reellen Nullstellen ± 1 und die imaginären Nullstellen $\pm \frac{2}{5}i$. Es besitzt Minima bei $\pm\sqrt{0.42}$ und ein lokales Maximum bei 0. Newtoniteration:

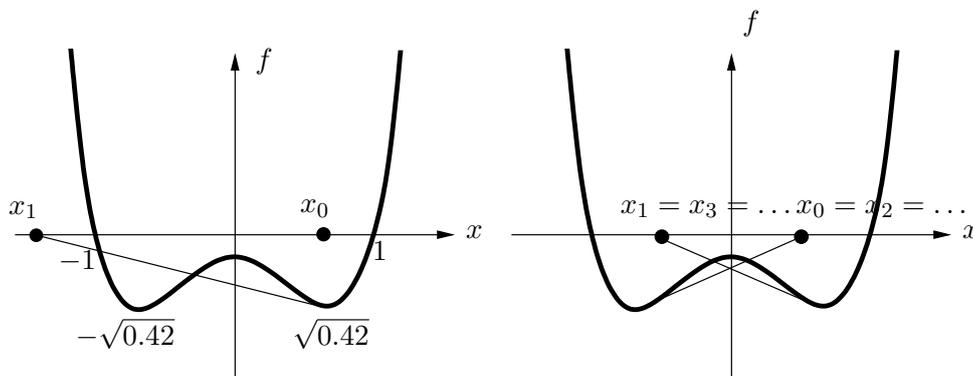
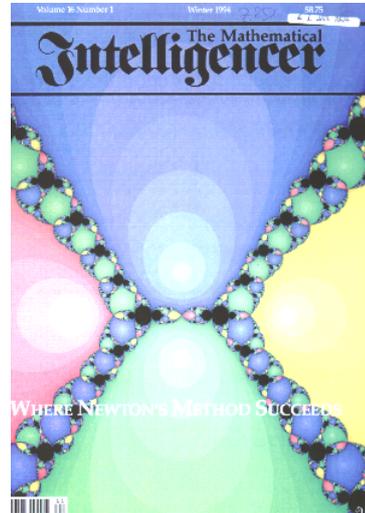


Abbildung 10.1: Newtoniteration eines Polynoms 4. Grades

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^4 - 0.84x_n^2 - 0.16}{4x_n^3 - 1.68x_n}$$

- Für $x_0 \gtrsim \sqrt{0.42}$ ist $\lim_{n \rightarrow \infty} x_n = 1$.

- Für $x_0 > \sqrt{0.42}$ ist $\lim_{n \rightarrow \infty} x_n = -1$, siehe Abb. 10.1 (links).
- Aber auch der Wechsel zwischen zwei Werten ist möglich, siehe Abb. 10.1 (rechts).
- In der komplexen Ebene besitzen die vier Nullstellen ihre *Bassins*, Bereiche des Anfangswertes, die zur Konvergenz führen. Diese sind in der nebenstehenden Abbildung mit verschiedenen Farben dargestellt. In den schwarzen Bereichen besitzt die Newtoniteration keinen Limes [aus: *Mathematical Intelligencer*, Vol. 16, No. 1].
◇



10.2 Das vereinfachte Newton-Verfahren

Wir nähern uns der Eingangsfrage nach den Nullstellen einer Funktion $f : U \rightarrow \mathbb{R}^n$ auf der offenen Teilmenge $U \subseteq \mathbb{R}^m$, indem wir den Spezialfall $n = m$ untersuchen, in dem wir $m - n = 0$ Parameter erwarten. Wir suchen also Nullstellen eines *Vektorfeldes* $f : U \rightarrow \mathbb{R}^m$ und erinnern uns an das *Newton-Verfahren*, das wir in Kapitel 10.1 für den eindimensionalen Fall benutzt haben.

Dort haben wir die Abbildung $x \mapsto F(x) := x - \frac{f(x)}{f'(x)}$ iteriert; geometrisch entspricht $F(x)$ dem Schnittpunkt der Tangente von f am Punkt $(x, f(x))$ mit der x -Achse.

Wir könnten dieses Verfahren nun direkt auf m Dimensionen verallgemeinern, indem wir $\frac{1}{f'(x)}$ durch die $m \times m$ -Matrix $(Df(x))^{-1}$ ersetzen. Dies ist zwar numerisch hervorragend, führt aber zu beweistechnischen Komplikationen, weswegen wir das *vereinfachte Newton-Verfahren* mit

$$x_{n+1} := F(x_n) \quad \text{für} \quad F(x) := x - A^{-1}f(x)$$

und x -*unabhängig* gewählter $m \times m$ -Matrix $A \approx Df(x)$ verwenden.

In einer Dimension liegt es nahe, dass für vernünftige A unser Verfahren immer noch zum Ziel führt, wenn auch nicht so schnell²⁶, siehe Abb. 10.2. Wir

²⁶Wenden wir das Fixpunktverfahren aus Kapitel 7.2 auf die Nullstellensuche für f an, dann wird $g(x) = f(x) + x$ iteriert, es ist also $A = -\mathbb{1}$.

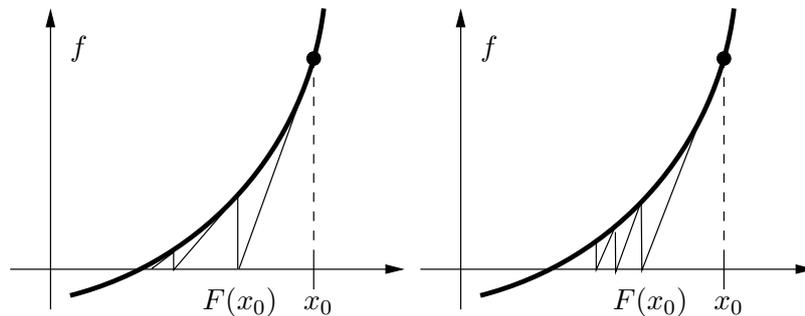


Abbildung 10.2: Newton-Verfahren (links) und vereinfachtes Newton-Verfahren (rechts), mit $A := f'(x_0)$

zeigen dies, indem wir F als kontrahierende Abbildung entlarven, also ein $\theta < 1$ finden mit $\|F(y) - F(x)\| \leq \theta \|y - x\|$.

Im eindimensionalen Fall können wir als Lipschitz-Konstante θ das Supremum des Betrags der Ableitung wählen. Nach dem folgenden Satz ist dies auch in m Dimensionen möglich.²⁷

10.6 Satz Ist $U \subseteq \mathbb{R}^m$ offen und konvex und $F \in C^1(U, \mathbb{R}^n)$, und ist $L := \sup_{x \in U} \|DF(x)\| < \infty$, dann gilt

$$\|F(y) - F(x)\| \leq L \|y - x\| \quad (x, y \in U).$$

Beweis: Es sei $v := F(y) - F(x) \neq 0$ (sonst ist nichts zu zeigen), und

$$V : \mathbb{R}^n \rightarrow \mathbb{R}, \quad V(z) := v \cdot z$$

der zu v duale Vektor. Die Abbildung $g := V \circ F : U \rightarrow \mathbb{R}$ hat nach Satz 8.30 (Mittelwertsatz) die Eigenschaft

$$g(y) - g(x) = Dg(\xi)(y - x)$$

für eine geeignete Konvexkombination $\xi = (1 - t)x + ty \in U$ mit $t \in [0, 1]$, sodass

$$|g(y) - g(x)| \leq \|Dg(\xi)\| \|y - x\|. \quad (10.3)$$

Nun ist $g(y) - g(x) = \|v\|^2$, andererseits $Dg(\xi) = v \cdot DF(\xi)$, also $\|Dg(\xi)\| \leq \|v\| \|DF(\xi)\|$. Eingesetzt in (10.3) ergibt dies $\|v\|^2 \leq \|v\| \|DF(\xi)\| \|y - x\|$. Daraus folgt nach Division durch $\|v\| > 0$ die Aussage. \square

Es bezeichnet

$$\text{GL}(n, \mathbb{K}) := \{M \in \text{Mat}(n, \mathbb{K}) \mid \det M \neq 0\}$$

die Gruppe der *invertierbaren* $n \times n$ -Matrizen über dem Körper \mathbb{K} . Damit können wir die Konvergenz des vereinfachten Newton-Verfahrens beweisen:

²⁷Für $A \in \text{Mat}(n \times m, \mathbb{R})$ ist die *Matrixnorm* durch $\|A\| := \sup_{v \in S^{m-1}} \|Av\|$ definiert.

10.7 Satz Es sei $U \subseteq \mathbb{R}^m$ offen, $f \in C^1(U, \mathbb{R}^m)$, und die abgeschlossene r -Umgebung von x_0 in U enthalten, d.h. $\overline{U_r(x_0)} \subseteq U$.

Ferner sei $A \in \text{GL}(m, \mathbb{R})$ so gewählt, dass

$$\|A^{-1}Df(x) - \mathbb{1}\| \leq \frac{1}{2} \quad \left(x \in \overline{U_r(x_0)}\right), \quad (10.4)$$

und es sei $\|A^{-1}f(x_0)\| \leq \frac{r}{2}$. Dann konvergiert das vereinfachte Newton-Verfahren

$$x_{k+1} := F(x_k) \quad \text{mit} \quad F(x) := x - A^{-1}f(x)$$

gegen die einzige Nullstelle ξ des Vektorfeldes f in $U_r(x_0)$.

Beweis:

- Aus (10.4) folgt nach Satz 10.6 für $\tilde{U} := \overline{U_r(x_0)} \subseteq U$

$$\|F(y) - F(x)\| \leq \sup_{z \in \tilde{U}} \|DF(z)\| \cdot \|y - x\| \leq \frac{1}{2} \|y - x\| \quad (x, y \in \tilde{U}),$$

denn $DF(z) = \mathbb{1} - A^{-1}Df(z)$. F ist also auf \tilde{U} kontrahierend, und

$$\|F(x_0) - x_0\| = \|A^{-1}f(x_0)\| \leq \frac{r}{2}.$$

- Daher ist für $x \in \tilde{U}$

$$\|F(x) - x_0\| \leq \|F(x) - F(x_0)\| + \|F(x_0) - x_0\| \leq \frac{1}{2} \|x - x_0\| + \frac{r}{2} \leq r,$$

sodass $F(\tilde{U}) \subseteq \tilde{U}$.

- Die Aussage folgt damit aus dem Banachschen Fixpunktsatz (Satz 7.14). □

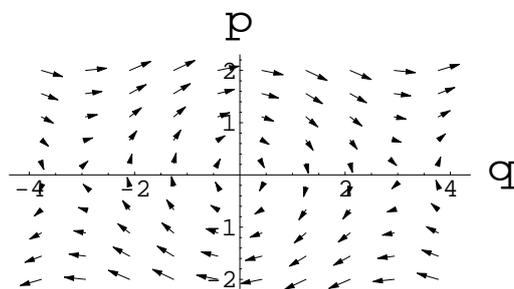
10.8 Bemerkungen

1. Setzt man z.B. $A := Df(x_0)$, und ist $A \in \text{GL}(m, \mathbb{R})$, dann ist Bedingung (10.4) für genügend kleines r erfüllt. Andererseits sollte r nicht zu klein sein, damit die zweite Bedingung $\|A^{-1}f(x_0)\| \leq \frac{r}{2}$ des Satzes erfüllt ist.

Beide Bedingungen zusammen lassen sich nur erfüllen, wenn der Startpunkt x_0 der Iteration schon nahe an einer isolierten Nullstelle von f liegt.

2. Andererseits wird für ein affines Vektorfeld der Form $f(x) = Ax + b$ (mit $A \in \text{GL}(m, \mathbb{R})$ und $b \in \mathbb{R}^m$) an einem beliebigen Anfangspunkt $x_0 \in \mathbb{R}^m$ die Ableitungsmatrix $Df(x_0)$ gleich A , sodass schon die erste Iterierte $x_1 = -A^{-1}b$ die Nullstelle von f ist.

10.9 Beispiel Wir betrachten das Vektorfeld $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f \begin{pmatrix} q \\ p \end{pmatrix} := \begin{pmatrix} p \\ -\sin q \end{pmatrix}$, in der $r := 1$ -Umgebung von $x_0 = \begin{pmatrix} q_0 \\ p_0 \end{pmatrix} := \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}$.



f ist das Vektorfeld der Pendel-Differentialgleichung

$$\dot{q} = p \quad , \quad \dot{p} = -\sin q$$

in der durch Ort q und Impuls p parametrisierten Ebene, siehe Abbildung.

Offensichtlich ist $\xi := \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in U_1(x_0)$ die einzige Nullstelle von f in $U_1(x_0)$.

Wir finden $Df(q, p) = \begin{pmatrix} 0 & 1 \\ -\cos q & 0 \end{pmatrix}$ und setzen $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, also $A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Damit ist

$$\|A^{-1}f(x_0)\| = \left\| \begin{pmatrix} \sin q_0 \\ p_0 \end{pmatrix} \right\| = \frac{1}{2}$$

und für $x \in \overline{U_1(x_0)}$

$$\begin{aligned} \|A^{-1}Df(x) - \mathbb{1}\| &= \left\| \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -\cos q & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} \cos(q)-1 & 0 \\ 0 & 0 \end{pmatrix} \right\| \\ &= 1 - \cos q \leq 1 - \left(1 - \frac{1}{2}q^2\right) = \frac{1}{2}q^2 \leq \frac{1}{2}. \end{aligned}$$

Die Bedingungen des Satzes 10.7 sind also erfüllt. Wir iterieren:

$$x_0 = \begin{pmatrix} q_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} \quad , \quad x_1 = \begin{pmatrix} q_1 \\ p_1 \end{pmatrix} = x_0 - \begin{pmatrix} \sin q_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

und landen mit Glück schon nach der ersten Iteration im kritischen Punkt.

Für allgemeine Anfangsbedingungen x_0 in der Nähe des Nullpunktes ist

$$x_1 = \begin{pmatrix} q_0 \\ p_0 \end{pmatrix} - \begin{pmatrix} \sin q_0 \\ p_0 \end{pmatrix} = \mathcal{O}(|q_0|^3). \quad \diamond$$

10.3 Konstruktion der impliziten Funktionen

Wir kommen jetzt zum allgemeinen Fall der Nullstellenmenge einer Funktion $f \in C^1(U, \mathbb{R}^n)$ mit $U \subseteq \mathbb{R}^m$ offen und $n \leq m$.

Wir nehmen an, dass wir schon eine Nullstelle kennen und wollen die Nullstellenmenge in der Nähe dieser Nullstelle parametrisieren. Wir benötigen dazu $p := m - n$ Parameter und wählen dazu p der Koordinaten. Ohne Beschränkung der Allgemeinheit sei also $x \in U$ in der Form

$$x = (y, z) \quad \text{mit} \quad y = (y_1, \dots, y_p) \quad \text{und} \quad z = (z_1, \dots, z_n)$$

mit $p := m - n$ geschrieben, und wir schreiben Df als $Df = (D_y f, D_z f)$ mit

$$D_y f := \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_p} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_p} \end{pmatrix}, \quad D_z f := \begin{pmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial z_1} & \cdots & \frac{\partial f_n}{\partial z_n} \end{pmatrix}.$$

Die zweite Matrix ist also quadratisch.

Unser Ziel ist es *lokal*, d.h. in der Nähe eines Punktes X der Nullstellenmenge von f , die implizit durch die Gleichungen

$$f_1(y_1, \dots, y_p, z_1, \dots, z_n) = 0, \dots, f_n(y_1, \dots, y_p, z_1, \dots, z_n) = 0$$

gegebenen z_1, \dots, z_n als Funktionen von y_1, \dots, y_p darzustellen.

10.10 Satz (Satz über implizite Funktionen) *Es gelte $f(X) = 0$ für $X = (Y, Z) = (Y_1, \dots, Y_p, Z_1, \dots, Z_n) \in U$ und $D_z f(X) \in \text{GL}(n, \mathbb{R})$.*

• *Dann gibt es offene Umgebungen $V := U_{r_y}(Y) \subset \mathbb{R}^p$, $W := U_{r_z}(Z) \subset \mathbb{R}^n$, $r_y, r_z > 0$ und eine Funktion $g \in C^1(V, W)$ mit*

$$\text{graph}(g) = f^{-1}(0) \cap (V \times W).$$

• *Ist $f \in C^k(U, \mathbb{R}^n)$ mit $1 \leq k \leq \infty$, dann ist $g \in C^k(V, W)$.*

10.11 Bemerkung g ist die gesuchte implizit definierte Funktion, die die Nullstellenmenge lokal parametrisiert, d.h.

$$f(y, g(y)) = 0 \quad (y \in V).$$

Ist z.B. $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = \|x\|^2 - 1$, und $X = (Y, Z) = (0, 1)$, dann ist für ein $r_y \in (0, 1)$, also $V = (-r_y, r_y)$ die implizite Funktion von der Form

$$g: V \rightarrow W, \quad y \mapsto \sqrt{1 - y^2}$$

mit dem Intervall $W = (1 - r_z, 1 + r_z)$, siehe Abb. 10.3. ◇

Beweis: • Im vereinfachten Newton-Verfahren setzen wir

$$A := D_z f(X) \quad \text{und} \quad F_y(z) := z - A^{-1} f(y, z).$$

Wegen der stetigen Differenzierbarkeit von f ist für ein geeignetes $r > 0$

$$\|D_z F_y(z)\| \leq \frac{1}{2} \quad \text{für alle} \quad (y, z) \in \bar{U}_r(Y) \times \bar{U}_r(Z) \subset U,$$

denn im Mittelpunkt dieses Gebietes ist ja

$$D_z F_y(z) = \mathbb{1} - A^{-1} D_z f(y, z)$$

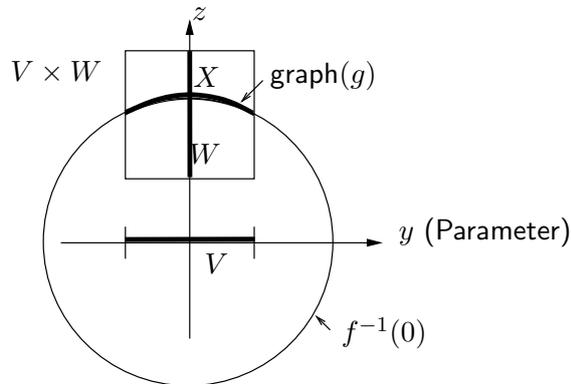


Abbildung 10.3: Lokale Parametrisierung des Kreises $S^1 = f^{-1}(0)$ durch g

gleich Null.

Ebenso gilt für kleine $r_y > 0$ und $V := U_{r_y}(Y)$

$$\|A^{-1}f(y, z)\| \leq \frac{r}{2} \quad (y \in V),$$

denn es ist ja $f(Y, Z) = 0$. Damit sind die Voraussetzungen von Satz 10.7 für alle Parameterwerte y in dieser Umgebung erfüllt, sodass das vereinfachte Newton-Verfahren für $y \in V$ Nullstellen $g(y) \in W := U_r(Z)$ liefert, die sogar eindeutig sind.

• Nun braucht $g : V \rightarrow W$ a priori noch nicht einmal stetig zu sein. Um dies zu zeigen, betrachten wir eine Nullstelle $x = (y, z)$ in der Nähe von $X = (Y, Z)$ und schreiben

$$\begin{aligned} 0 &= f(x) - f(X) = (f(y, z) - f(Y, z)) + (f(Y, z) - f(Y, Z)) \\ &= B_y(y - Y) + A_y(z - Z) \end{aligned}$$

mit

$$B_y := \begin{pmatrix} D_y f_1(\xi_1, z) \\ \vdots \\ D_y f_n(\xi_n, z) \end{pmatrix} \quad \text{und} \quad A_y := \begin{pmatrix} D_z f_1(Y, \eta_1) \\ \vdots \\ D_z f_n(Y, \eta_n) \end{pmatrix},$$

nach dem Mittelwertsatz für reelle Funktionen. Dabei liegen die ξ_i auf der Strecke zwischen y und Y , und die η_i auf der zwischen z und Z .

Auflösung nach $z - Z = g(y) - g(Y)$ ergibt

$$g(y) - g(Y) = M(y)(y - Y) \quad \text{mit} \quad M(y) := -A_y^{-1}B_y.$$

Da $\|M(y)\| \leq c$ für $y \in V$, ist zunächst $\lim_{y \rightarrow Y} g(y) = g(Y)$, g also stetig bei Y .

• Weiter ist in diesem Limes $M(y) = M(Y) + \mathcal{O}(\|y - Y\|)$, sodass g bei Y differenzierbar ist. Das gleiche Argument kann aber für alle $y \in V$ angewandt werden, sodass g in V differenzierbar ist.

• Stetig differenzierbar ist g , da aus $f(y, g(y)) = 0$ und der Kettenregel

$$\boxed{Dg(y) = -(D_z f(y, g(y)))^{-1} D_y f(y, g(y))}$$

folgt, und die rechte Seite in y stetig ist.²⁸

- Iteration des Argumentes für $f \in C^k(U, \mathbb{R}^n)$ liefert $g \in C^k(V, W)$. □

Das war nun ein langer Beweis und zur Abwechslung wenden wir ihn auf ein einfaches Beispiel an.

10.12 Beispiel Die Kreislinie $S^1 \subset \mathbb{R}^2$ ist die Nullstellenmenge von $f(x_1, x_2) = x_1^2 + x_2^2 - 1$. Wir wählen $X = (Y, Z) := (0, 1) \in S^1$ und sehen für $x = (y, z)$, dass $D_y f(x) = 2y$ und $D_z f(x) = 2z$, sodass $D_z f(X) = 2 \neq 0$. Wir können also den Satz über die implizite Funktion anwenden und eine lokal definierte Funktion g mit $f(y, g(y)) = 0$, $g(0) = 1$ aufsuchen.

Mit $A := D_z f(X) = 2$ ist

$$F_y(z) = z - \frac{1}{2}(y^2 + z^2) + \frac{1}{2}.$$

Wir starten mit der konstanten Funktion $g_0(y) := 1$, und erhalten

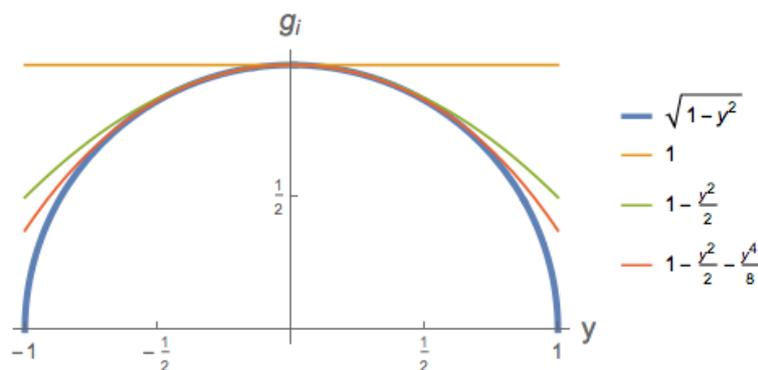


Abbildung 10.4: Newton-Approximation der Kreislinie

$$g_1(y) := F_y(g_0(y)) \quad , \quad \text{also} \quad g_1(y) = 1 - y^2/2.$$

Die nächste Iteration $g_2(y) := F_y(g_1(y))$ liefert

$$g_2(y) = 1 - \frac{y^2}{2} - \frac{1}{2}(y^2 + (1 - \frac{y^2}{2})^2 - 1) = 1 - \frac{y^2}{2} - \frac{1}{8}y^4,$$

und wir approximieren mit den g_i die Funktion

$$g(y) = \sqrt{1 - y^2} = 1 - \frac{y^2}{2} - \frac{1}{2 \cdot 4}y^4 - \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}y^6 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8}y^8 \dots$$

Es ist im Allgemeinen nicht möglich, aber auch nicht notwendig wie in Satz 10.10, die ersten p Koordinaten als Parameter zu benutzen:

²⁸Diese Formel ist nützlich, denn sie erlaubt die Berechnung von Dg ohne explizite Berechnung von g .

10.13 Satz Es sei $f \in C^1(U, \mathbb{R}^n)$ mit $U \subseteq \mathbb{R}^m$ offen und $n \leq m$. Ist $X \in f^{-1}(0)$ regulärer Punkt von f , also $\text{Rang}(Df(X)) = n$, dann lassen sich Koordinatenindizes $1 \leq i_1 < \dots < i_n \leq m$ auswählen, sodass

$$\begin{pmatrix} D_{i_1} f_1(X) & \dots & D_{i_n} f_1(X) \\ \vdots & & \vdots \\ D_{i_1} f_n(X) & \dots & D_{i_n} f_n(X) \end{pmatrix} \in \text{GL}(n, \mathbb{R})$$

gilt, diese Matrix also invertierbar ist.

10.14 Bemerkung Dies ist zunächst eine rein algebraische Aussage über die Matrix $Df(X)$. Setzt man aber $z_k := x_{i_k}$, $k = 1, \dots, n$ und nennt die restlichen $p = m - n$ Koordinaten y_1, \dots, y_p , dann lässt sich nach Satz 10.10 die Nullstellenmenge in der Nähe von X durch die Parameter y_k parametrisieren. \diamond

Beweis: Der Rang einer Matrix ist die Maximalzahl unabhängiger Spaltenvektoren (oder auch Zeilenvektoren). Wir können also durch Auswahl der Spalten i_1, \dots, i_n eine quadratische Untermatrix der Größe n von $Df(X) \in \text{Mat}(n \times m, \mathbb{R})$ erzeugen, die invertierbar ist. \square

10.15 Beispiel $f : \mathbb{R}^m \rightarrow \mathbb{R}$, $f(x) = \|x\|^2 - 1$. $Df(X) = 2X \neq 0$ für alle $X \in f^{-1}(0) = S^p$, $p := m - 1$. Also ist $\text{Rang}(Df(X)) = 1$ und wir finden in der Nähe jedes Punktes X der Sphäre p lokale Koordinaten für S^p .

Allerdings ist die Wahl abhängig von X : Genau wenn nämlich $X_i = \pm 1$ ist, darf x_i nicht als lokale Koordinate gewählt werden. \diamond

10.16 Definition Für $U \subseteq \mathbb{R}^m$ offen und $f \in C^1(U, \mathbb{R}^n)$ mit $n \leq m$ heißt $y \in \mathbb{R}^n$ **regulärer Wert** von f , wenn alle $x \in f^{-1}(y)$ reguläre Punkte von f sind, sonst **kritischer** (oder **singulärer**) **Wert**.

10.17 Beispiel Im Beispiel 10.15 sind alle $y \in \mathbb{R} \setminus \{-1\}$ reguläre Werte von f , während -1 kritischer Wert ist. Da die Zahlen $y < -1$ gar nicht im Bild von f enthalten sind, sind sie trivialerweise reguläre Werte. All die Urbilder $f^{-1}(y)$ regulärer Werte $y > -1$ sind Sphären um den Ursprung (mit Radius $\sqrt{y+1}$). \diamond

10.18 Definition Für $p \in \{0, \dots, m\}$ heißt eine Teilmenge $M \subseteq \mathbb{R}^m$ **p -dimensionale Untermannigfaltigkeit** des \mathbb{R}^m , wenn jeder Punkt $x \in M$ eine Umgebung $V_x \subseteq \mathbb{R}^m$ besitzt, so dass für eine geeignete Abbildung $f \in C^1(V_x, \mathbb{R}^{m-p})$ mit regulärem Wert 0 gilt:

$$M \cap V_x = f^{-1}(0).$$

Im einfachsten Fall ist $M = f^{-1}(0)$, aber man möchte auch Mengen wie im folgenden Beispiel als Mannigfaltigkeiten bezeichnen.

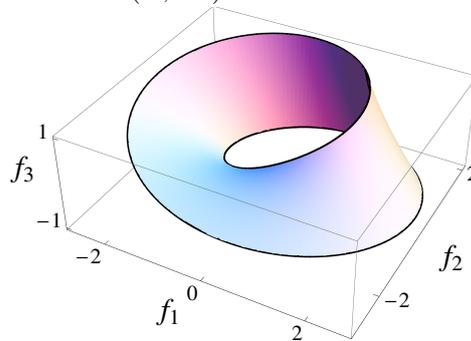
10.19 Beispiel Für $U := \mathbb{R} \times (-1, 1)$ und $F \in C^\infty(U, \mathbb{R}^3)$,

$$F(x, y) := \begin{pmatrix} (2-y \sin \frac{x}{2}) \sin x \\ (2-y \sin \frac{x}{2}) \cos x \\ y \cos \frac{x}{2} \end{pmatrix}$$

ist $M := F(U) \subset \mathbb{R}^3$ das so genannte **Möbius-Band**.

Für die Parametrisierung würde der Winkelbereich $x \in [0, 2\pi)$ ausreichen, denn zwar kommen Winkel $x/2$ in F vor, aber $F(x + 2\pi, y) = F(x, -y)$. $F(\mathbb{R} \times \{0\})$ ist eine Kreislinie vom Radius 2.

Da die Fläche M nur *eine* Seite besitzt, kann sie nicht Niveaumenge $f^{-1}(0)$ für einen regulären Wert 0 sein, denn sonst würde $\nabla f(x) \neq 0$ am Punkt $x \in M$ senkrecht auf der Fläche stehen und damit eine von zwei Seiten auszeichnen. \diamond



10.20 Definition Es sei $U \subseteq \mathbb{R}^n$ offen und $f \in C^1(U, \mathbb{R}^n)$.

- f heißt **Diffeomorphismus auf das Bild** $V := f(U) \subseteq \mathbb{R}^n$, wenn V offen, $f : U \rightarrow V$ bijektiv und $f^{-1} : V \rightarrow U$ stetig differenzierbar ist.
- f heißt **lokaler Diffeomorphismus**, wenn jeder Punkt $x \in U$ eine offene Umgebung $U_x \subseteq U$ besitzt, für die $f|_{U_x}$ ein Diffeomorphismus auf das Bild $f(U_x)$ ist.

Man kann Diffeomorphismen als Koordinatenwechsel ansehen, und da man gerne dem jeweiligen Problem angepasste Koordinaten verwendet, sind Diffeomorphismen eine häufig verwendete Klasse von Abbildungen.

10.21 Beispiel Eine affine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ besitzt die Form $f(x) = Ax + b$ mit $A \in \text{Mat}(n, \mathbb{R})$ und $b \in \mathbb{R}^n$. Sie ist genau dann ein Diffeomorphismus, wenn sie bijektiv ist, d.h. wenn $A \in \text{GL}(n, \mathbb{R})$ ist. \diamond

Aus diesem Beispiel liest man ab, dass die Regularität der Jacobi-Matrix Df Einfluss auf die Invertierbarkeit der Abbildung f hat, denn hier ist $Df = A$.

10.22 Satz Für $U \subseteq \mathbb{R}^n$ offen ist $f \in C^1(U, \mathbb{R}^n)$ genau dann ein lokaler Diffeomorphismus, wenn für alle $x \in U$ gilt: $Df(x) \in \text{GL}(n, \mathbb{R})$.

Beweis:

- Es sei f ein lokaler Diffeomorphismus und $g : V_x \rightarrow U_x$ die Umkehrfunktion von $f|_{U_x} : U_x \rightarrow V_x$. Dann gilt nach der Kettenregel

$$(Dg)(f(x))Df(x) = D(g \circ f)(x) = D\text{Id}_{U_x}(x) = \mathbb{1},$$

also $Df(x) \in \text{GL}(n, \mathbb{R})$.

- Es gelte umgekehrt $Df(z) \in \text{GL}(n, \mathbb{R})$. Um die lokale Inverse von f bei $z \in U$ zu finden, wenden wir den Satz über die implizite Funktion auf

$$F : \mathbb{R}^n \times U \rightarrow \mathbb{R}^n, \quad (y, z) \mapsto -y + f(z)$$

an. Nach Voraussetzung ist für $X := (f(z), z)$

$$D_z F(X) = Df(z) \in \text{GL}(n, \mathbb{R}),$$

und $F(X) = 0$. Anwendung von Satz 10.10 ergibt die Existenz einer auf der Umgebung $V_z := U_{r_y}(f(z))$ des Bildpunktes definierten Abbildung $g \in C^1(V_z, W)$ mit $W = U_{r_z}(z)$, für die $F(y, g(y)) = f(g(y)) - y = 0$ ist. Wir setzen $U_z := g(V_z) \subseteq W$. Sowohl g als auch $f|_{U_z}$ sind injektiv, denn sonst könnte nicht $f \circ g = \text{Id}_{V_z}$ gelten. Damit ist auch $g \circ f|_{U_z} = \text{Id}_{V_z}$ und nach der Kettenregel $Dg(y) \in \text{GL}(n, \mathbb{R})$ für alle $y \in V_z$. Damit ist U_z nach dem folgenden Satz eine offene Umgebung von z . \square

10.23 Satz *Es sei $U \subset \mathbb{R}^n$ offen und $f \in C^1(U, \mathbb{R}^n)$. Ist f regulär, d.h. gilt $Df(x) \in \text{GL}(n, \mathbb{R})$ für alle $x \in U$, dann ist $f(V)$ offen, falls $V \subseteq U$ offen ist.*

Beweis: Siehe z.B. Hildebrandt [Hi], Band 2, Kapitel 1.9. \square

11 Extrema mit Nebenbedingungen

Auch bei glatten Funktionen sind die Extremalstellen nicht notwendig kritische Punkte.

11.1 Beispiel Die Funktion $f : D \rightarrow \mathbb{R}$, $f(x) := x_1$ auf der Kreisscheibe

$$D := \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$$

besitzt eine Maximalstelle bei $a = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \partial D$. Ihr Gradient $\nabla f(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist bei a senkrecht auf dem Rand $\partial D = S^1$.

Allgemein können am Rand ∂D des Definitionsbereiches D lokale Extrema auftreten, ohne dass dort der Gradient verschwindet. \diamond

11.2 Beispiel Es sind $n \geq 3$ Punkte p_1, \dots, p_n so im positiven Umlaufsinn auf eine Kreislinie zu legen, dass das Polygon mit diesen Ecken maximale Fläche f hat. Für Radius 1 des Kreises ist $f = \sum_{i=1}^n f_i = \sum_{i=1}^n \frac{1}{2} \sin(x_i)$, wenn x_i der Zentriwinkel des i -ten Dreiecks ist, und f_i dessen Fläche, siehe Abbildung 11.1.

Es ist also die Funktion $f(x_1, \dots, x_n) := \sum_{i=1}^n \frac{1}{2} \sin(x_i)$ unter den Nebenbedingungen $0 \leq x_i < 2\pi$, $x_1 + \dots + x_n = 2\pi$ zu maximieren (diese Formel ist auch dann richtig, wenn ein Winkel größer als π ist). Jede Anordnung der Punkte p_i mit diesen Relativwinkeln führt dann zu einem Polygon maximaler Fläche. \diamond

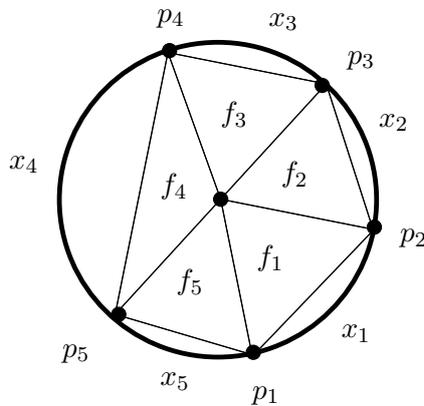


Abbildung 11.1: Fläche $f = f_1 + \dots + f_n$ eines einem Kreis einbeschriebenen Polygons mit n Ecken

11.1 1. Methode: Parametrisierung der Nebenbedingungen

Beispiel 11.1 Die Randpunkte $x \in \partial D$ lassen sich durch $x(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$, $t \in [0, 2\pi)$ parametrisieren; es ist $f(x(t)) = \cos t$, also maximal für $t = 0$, d.h. $x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. \diamond

Beispiel 11.2 Wir betrachten die Winkel x_1, \dots, x_{n-1} als unabhängige Variable und setzen $x_n := 2\pi - (x_1 + \dots + x_{n-1})$. Dann ist

$$f(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) = \tilde{f}(x_1, \dots, x_{n-1}) := \frac{1}{2} \sum_{i=1}^{n-1} \sin x_i + \frac{1}{2} \sin(2\pi - (x_1 + \dots + x_{n-1})).$$

$$\text{Damit ist } \nabla \tilde{f}(x_1, \dots, x_{n-1}) = \frac{1}{2} \begin{pmatrix} \cos(x_1) - \cos(x_1 + \dots + x_{n-1}) \\ \vdots \\ \cos(x_{n-1}) - \cos(x_1 + \dots + x_{n-1}) \end{pmatrix}.$$

Genau für $\cos(x_1) = \dots = \cos(x_{n-1}) = \cos(x_n)$ ist $\nabla \tilde{f}(x_1, \dots, x_{n-1}) = 0$. Nun kann $\cos(x_i) = \cos(x_j)$ nur gelten, wenn die Winkel $x_i = x_j$ sind, oder wenn $x_i + x_j = 2\pi$. Letzteres würde aber wegen unserer Nebenbedingung $x_1 + \dots + x_n = 2\pi$ bedeuten, dass alle Winkel ausser einem gleich Null sind, entsprechend der Fläche $f = 0$. Also ist das Maximum nur für die Winkel

$$x_1 = \dots = x_n = \frac{2\pi}{n} \quad \text{d.h. Fläche} \quad f = \frac{n}{2} \sin(2\pi/n)$$

realisiert, nur das gleichseitige Polygon besitzt also die maximale Fläche. \diamond

In den Beispielen traten als Nebenbedingungen

- Ungleichungen ($\sqrt{x_1^2 + x_2^2} \leq 1$) und
- Gleichungen ($x_1 + \dots + x_n = 2\pi$) auf.

Wir gehen also folgendermaßen vor:

- Im Fall von *Ungleichungen* $g(x) \leq 0$ (und in diese Form können wir jede Ungleichung $\tilde{g}(x) \leq c$ durch Übergang zu $g := \tilde{g} - c$ bringen) können wir zunächst untersuchen, ob im Gebiet $\{x \mid g(x) < 0\}$ kritische Punkte auftreten, denn dieses Gebiet ist für stetige g ja offen, sodass nach Satz 9.14 Extrema kritische Punkte sein müssen.
- Danach betrachten wir die Extrema von f auf der Nullstellenmenge $g^{-1}(0)$, d.h. auf dem Rand des Gebietes.

Problem: Gesucht sind die Extrema der Funktion $f \in C^1(U)$, $U \subseteq \mathbb{R}^m$ offen unter den Nebenbedingungen $g_1(x) = \dots = g_n(x) = 0$ für $g_i \in C^1(U)$, $n < m$.

Lösung: Wir fassen die Nebenbedingungen zu einer vektorwertigen Funktion $G = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} : U \rightarrow \mathbb{R}^n$ zusammen. Wenn für $\xi \in G^{-1}(0)$ $\text{Rang}(DG(\xi)) = n$ ist, können wir nach Satz 10.13 eine lokale Parametrisierung $h : V \rightarrow W$ der Nullstellenmenge $G^{-1}(0)$ mit $G(y, h(y))$ finden, und f auf dieser Untermannigfaltigkeit des \mathbb{R}^m als Funktion $y \mapsto f(y, h(y))$ diskutieren.

Dieses Verfahren ist aber oft zu mühsam.

11.3 Beispiel

Gesucht sind kritische Punkte und Extrema der quadratischen Form

$$f : \mathbb{R}^m \rightarrow \mathbb{R} \quad , \quad f(x) := \langle x, Ax \rangle$$

für eine Matrix $A \in \text{Mat}(m, \mathbb{R})$ unter der Nebenbedingung $g(x) := \|x\|^2 - 1 = 0$, also auf der Sphäre $S^{m-1} \subset \mathbb{R}^m$.

Eine Parametrisierung der Nebenbedingung bedeutet hier die Einführung von Koordinaten auf S^{m-1} . Dies ist vor allem für hohe Dimensionen m umständlich und unübersichtlich. Oft ist in vergleichbaren Fällen die Methode der Lagrange-Multiplikatoren vorzuziehen. \diamond

11.2 2. Methode: Lagrange–Multiplikatoren

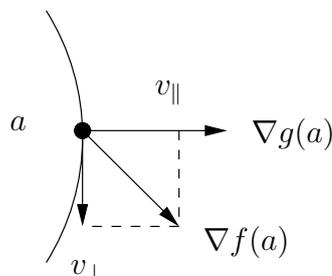
11.4 Satz Zu jeder Lösung a des Extremalproblems für $f \in C^1(U)$ unter der Nebenbedingung $g(x) = 0$, $g \in C^1(U)$ mit $\nabla g(a) \neq 0$ gibt es ein $\lambda_0 \in \mathbb{R}$ mit

$$\nabla f(a) + \lambda_0 \nabla g(a) = 0.$$

Die beiden Gradienten sind also parallel.

Beweis: $\nabla g(a)$ steht senkrecht auf der Nullstellenmenge $g^{-1}(0)$ durch a .

Wir spalten $\nabla f(a) = v_{\parallel} + v_{\perp}$ eindeutig in die zu $\nabla g(a)$ parallele bzw. senkrechte Komponente v_{\parallel} bzw. v_{\perp} auf.



Ist nun $v_{\perp} \neq 0$, dann ist die Richtungsableitung $\partial_{v_{\perp}} f(a) = \langle v_{\perp}, \nabla f(a) \rangle = \langle v_{\perp}, v_{\perp} \rangle > 0$, d.h. wir können f vergrößern, indem wir auf der Fläche $g^{-1}(0)$ in Richtung v_{\perp} gehen und in umgekehrter Richtung verkleinern. Widerspruch zur Extremalität! \square

Frage: Wie können wir den Satz benutzen, um die Extrema unter Nebenbedingungen zu finden?

Antwort: Mit einem *Trick*: Betrachten wir

$$F : U \times \mathbb{R} \rightarrow \mathbb{R} \quad , \quad F(x, \lambda) := f(x) + \lambda \cdot g(x).$$

Dann ist $\nabla F(x, \lambda) = \begin{pmatrix} \nabla f(x) + \lambda \nabla g(x) \\ g(x) \end{pmatrix}$, ist also $\nabla F(a, \lambda_0) = 0$, dann ist $g(a) = 0$ und $\nabla f(a) + \lambda_0 \nabla g(a) = 0$, wir haben also einen kritischen Punkt unter der Nebenbedingung gefunden!

11.5 Bemerkung Bei mehreren Nebenbedingungen

$$g_1(x) = \dots = g_n(x) = 0$$

betrachtet man analog die Funktion mit den *Lagrange-Multiplikatoren* $\lambda_1, \dots, \lambda_n$

$$F : U \times \mathbb{R}^n \rightarrow \mathbb{R} \quad , \quad F(x, \lambda_1, \dots, \lambda_n) := f(x) + \sum_{i=1}^n \lambda_i g_i(x). \quad \diamond$$

11.6 Beispiel Für die in Beispiel 11.3 betrachtete quadratische Form

$$f : \mathbb{R}^m \rightarrow \mathbb{R} \quad , \quad f(x) = \langle x, Ax \rangle$$

mit²⁹ $A = A^T$ und die Nebenbedingung $g(x) = 1 - \|x\|^2 \stackrel{!}{=} 0$ ist

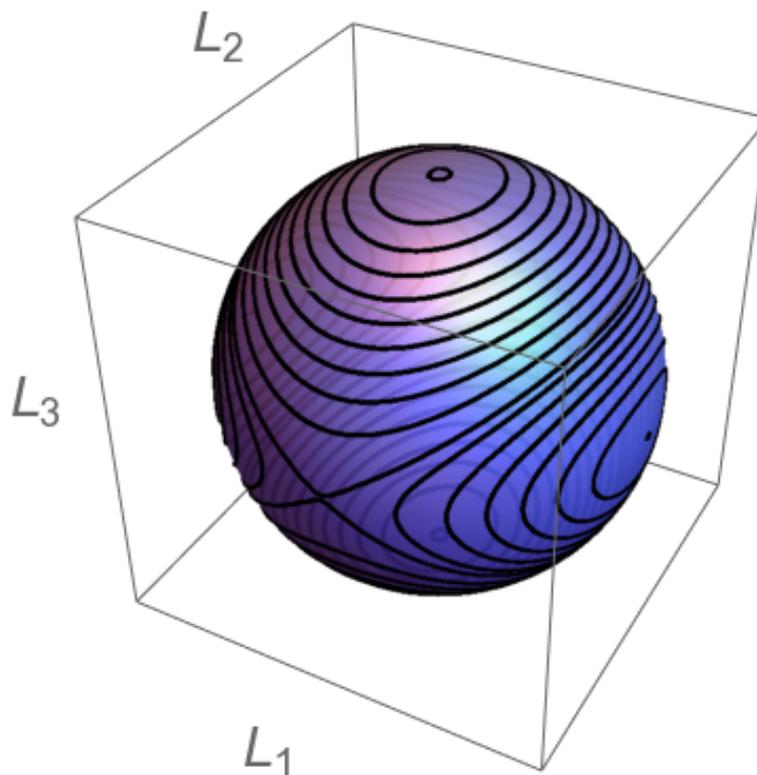
$$F : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}, \quad F(x, \lambda) = \langle x, Ax \rangle - \lambda(\langle x, x \rangle - 1) = \langle x, (A - \lambda \mathbb{1})x \rangle + \lambda,$$

²⁹Falls dagegen A nicht symmetrisch ist, geht man zur symmetrischen Matrix $\tilde{A} := \frac{A+A^T}{2}$ über, denn es gilt ja $\langle x, Ax \rangle = \langle x, \tilde{A}x \rangle$.

also $\nabla F(x, \lambda) = \begin{pmatrix} 2(A - \lambda \mathbb{1})x \\ 1 - \|x\|^2 \end{pmatrix}$. Wir müssen also die Eigenwerte $\lambda_i \in \mathbb{R}$ von A und die zugehörigen normierten Eigenvektoren $x_i \in S^{m-1}$, $(A - \lambda_i \mathbb{1})x_i = 0$ finden, denn (genau) für diese gilt $\nabla F(x_i, \lambda_i) = 0$.

Sind die Eigenwerte voneinander verschieden, dann ist mit x_i nur $-x_i \in S^{m-1}$ Eigenvektor zu λ_i .

Für $m = 3$ und $A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}$, $a_1 < a_2 < a_3$, ergibt sich folgendes Bild: Die Eigenvektoren $\pm e_1$ zum Eigenwert a_1 sind Minima der auf S^2 restringierten Funktion f , die Eigenvektoren $\pm e_3$ zum Eigenwert a_3 Maxima. Bei den Eigenvektoren $\pm e_2 \in S^2$ besitzt $f|_{S^2}$ eine indefinite Hesse-Matrix (Sattel), siehe Abbildung.



In beliebigen Dimensionen m gilt $f(-x) = f(x)$, die kritischen Punkte treten also paarweise an Antipoden von S^{m-1} auf. \diamond

Die Methode der Lagrange-Multiplikatoren hat den Vorteil, dass alle Variablen gleichbehandelt werden. Dadurch wird die Rechnung oft übersichtlicher.

11.7 Beispiel (Polygon im Kreis) Zu maximierende Funktion:

$$f(x) = \frac{1}{2} \sum_{i=1}^n \sin(x_i) \quad (\text{Fläche})$$

$$g(x) := -2\pi + \sum_{i=1}^n x_i \quad (\text{Nebenbedingung})$$

Für $F(x, \lambda) := f(x) + \lambda g(x)$ ist $\nabla F(x, \lambda) = \begin{pmatrix} \frac{1}{2} \cos(x_1) + \lambda \\ \vdots \\ \frac{1}{2} \cos(x_n) + \lambda \\ g(x) \end{pmatrix}$.

Aus der Bedingung $\nabla F(a, \lambda_0) = 0$ liest man ab:

- $\cos(a_1) = \dots = \cos(a_n) \quad (= -2\lambda_0)$.
- Unter der Bedingung $0 \leq x_i \leq \pi$ folgt $a_1 = \dots = a_n$. Falls ein Winkel $a_i \in [\pi, 2\pi)$ ist, kann die Bedingung $\cos(a_1) = \dots = \cos(a_n)$ nicht erfüllt sein.
- Wegen $g(a_1, \dots, a_n) = 0$ folgt $a_1 = \dots = a_n = \frac{2\pi}{n}$. ◇

12 * Ergänzungen

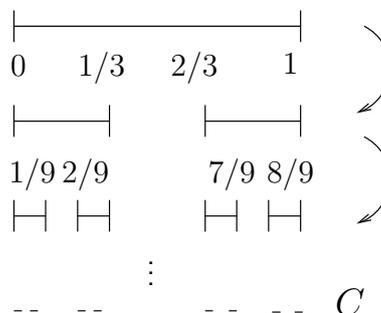
12.1 Physikalische Bedeutung der Fraktale

Manche Teilräume, wie die Cantor-Menge, waren ursprünglich als abstrakte Beispiele für Räume mit besonderen Eigenschaften ersonnen worden, ergeben sich aber natürlich als invariante Mengen so genannter dynamischer Systeme. Dieses und vergleichbare Phänomene werden *deterministisches Chaos* genannt.

12.1 Beispiel Die *Cantor-Menge* $C \subset [0, 1]$

entsteht, indem man das mittlere Drittel $(\frac{1}{3}, \frac{2}{3})$ des Intervalls entfernt, und Gleiches induktiv für alle Teilintervalle tut, siehe Abbildung.

Es ist $\partial C = C$, und die Menge der Häufungspunkte von C ist gleich C .



Definiert man nun $f : \mathbb{R} \rightarrow \mathbb{R}$ durch $f(x) := \frac{3}{2} \cdot \left(1 - 2 \left|x - \frac{1}{2}\right|\right)$ und $f^{(n)} : \mathbb{R} \rightarrow \mathbb{R}$ iterativ durch $f^{(0)}(x) := x$, $f^{(n+1)}(x) := f \circ f^{(n)}(x)$, so ist

$$C = \{x \in [0, 1] \mid \forall n \in \mathbb{N} : f^{(n)}(x) \in [0, 1]\}.$$

Denn es ist das mittlere Drittel gleich der Menge $\{x \in [0, 1] \mid f(x) \notin [0, 1]\}$

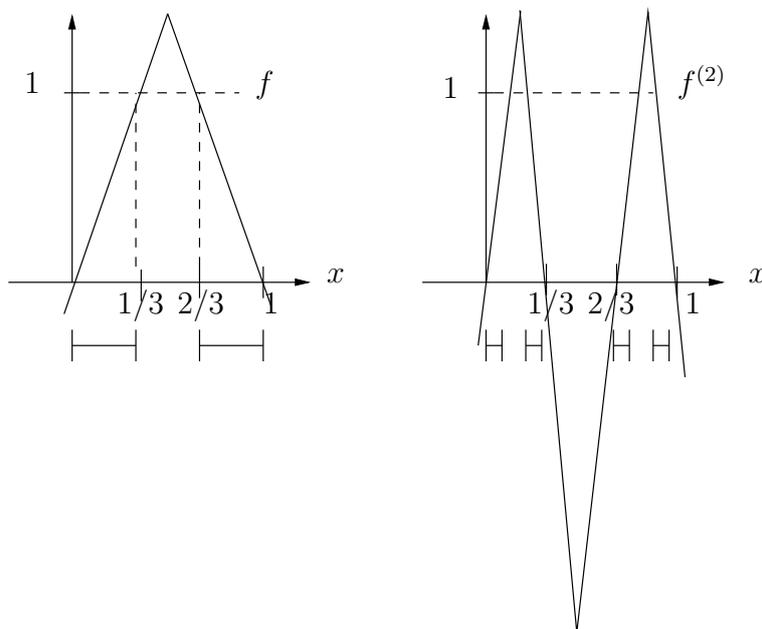


Abbildung 12.1: Zur Konstruktion der Cantor-Menge C mittels Iteration der Abbildung f

und die äußeren Drittel $[0, \frac{1}{3}]$, $[\frac{2}{3}, 1]$ werden durch f affin auf $[0, 1]$ abgebildet, siehe Abb. 12.1. \diamond

Statt iterierter Funktionen $f^{(1)}, f^{(2)}, \dots, f^{(n)}, \dots$ untersucht man in der Physik oft Lösungen von Differentialgleichungen wie z.B. der Newtonschen Kraftgleichung (3.1).

12.2 Beispiel Als invariante Mengen z.B. der Differentialgleichung für die

Planetenbewegung (sog. n -Körper-Problem der Himmelsmechanik) treten Mengen auf, die als topologische Räume der Cantor-Menge gleichen. Denn startet man einen Satelliten mit genügend hoher Geschwindigkeit, dann kann man im Prinzip durch Wahl des Startwinkels die Reihenfolge der Besuche der Planeten beliebig vorgeben.



Dies ist analog zur Vorgabe der Reihenfolge, in der die Iterierten $f^i(x)$, $i \in \mathbb{N}$ in Beispiel 12.1 die beiden Intervalle $[0, \frac{1}{3}]$, $[\frac{2}{3}, 1]$ besuchen

Bei nahen Besuchen lenken die Planeten den Satelliten ab; ähnlich wie in Beispiel 12.1 verläßt die Bahn andernfalls den beobachteten Bereich (hier das Sonnensystem, dort das Intervall $[0, 1]$). \diamond

In der *mengentheoretischen* Topologie werden Begriffe wie Rand, Häufungspunkt oder Stetigkeit untersucht. In der Physik wird sie benötigt um Eigenschaften von topologischen Räumen wie der Cantor-Menge zu klären, aber auch die der Funktionenräume, die z.B. in der Quantenmechanik auftreten.

12.2 Physikalische Bedeutung der Topologie

Auch topologische Räume die *lokal* ganz einfach aussehen, können *global* interessante Eigenschaften besitzen. Solche Phänomene werden in dem Untergebiet der sog. *algebraischen* Topologie untersucht:

12.3 Beispiel Der *Tangententialraum* am Punkt $x \in S^2$ der 2-Sphäre ist durch

$$T_x S^2 := \{y \in \mathbb{R}^3 \mid (x, y) = 0\}$$

definiert, besteht also aus der Ebene der zu S^2 bei x tangentialen Vektoren y .

Ein *Tangententialvektorfeld* an S^2 ist eine stetige Abbildung $f : S^2 \rightarrow \mathbb{R}^3$ mit $f(x) \in T_x S^2$. Modelliert etwa S^2 die Erdoberfläche, dann wird die an einem Ort x herrschende (horizontale) Windgeschwindigkeit durch den Wert $f(x)$ eines solchen Vektorfeldes beschrieben.

Es lässt sich nun beweisen, dass jedes Tangentialvektorfeld f mindestens eine Nullstelle x besitzt (d.h. $f(x) = 0$). Irgendwo auf der Erdoberfläche herrscht also immer Windstille³⁰.

³⁰Der Beweis findet sich z.B. in meinem Skript *Mathematische Physik 1 (Klassische Mechanik)*

Wäre dagegen (wie früher vermutet) die Erde von der Form einer (offenen) Kreisscheibe

$$\overset{\circ}{D}^2 := \{y \in \mathbb{R}^2 \mid \|y\| < 1\},$$

dann müsste das nicht so sein, siehe Abbildung 12.2.

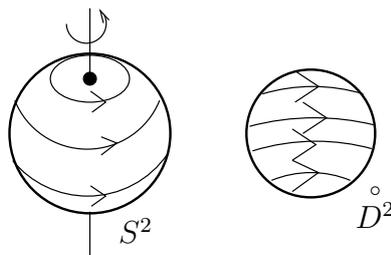


Abbildung 12.2: Geschwindigkeitsfelder. Links: auf der 2-Sphäre, mit Ruhelagen an Nord- und Südpol, rechts: auf der offenen Kreisscheibe, ohne Ruhelage

Das Phänomen ist wirklich globaler Natur, denn die Sphäre außer dem Südpol $S^2 \setminus \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right\}$ können wir mittels stereographischer Projektion bijektiv³¹ auf den \mathbb{R}^2 abbilden (siehe Abbildung 12.3), und den \mathbb{R}^2 dann mittels der Abbildung

$$\mathbb{R}^2 \rightarrow \overset{\circ}{D}^2 \quad , \quad x = (x_1, x_2) \mapsto \begin{cases} \frac{\tanh(\|x\|)}{\|x\|} \cdot (x_1, x_2) & , x \neq 0 \\ 0 & , x = 0 \end{cases}$$

bijektiv auf die offene Kreisscheibe $\overset{\circ}{D}^2$ abbilden. Wir können statt dem Südpol einen beliebigen Punkt $x \in S^2$ aussparen, und daher auf dem Rest $S^2 \setminus \{x\}$ der Sphäre ein Tangentialvektorfeld ohne Nullstelle finden. \diamond

Ein zweites Beispiel: Geschlossenen regulären Kurven in der Ebene kann man mit dem Rotationsindex eine *topologische Invariante* zuordnen, d.h. eine Zahl, die sich bei kleinen Deformationen der Kurve nicht ändert:

12.4 Satz *Es sei $c \in C^2([a, b], \mathbb{R}^2)$ eine nach Bogenlänge parametrisierte geschlossene Kurve, d.h.*

$$c(b) = c(a) \quad , \quad c'(b) = c'(a) \quad \text{und} \quad c''(b) = c''(a).$$

Dann ist der **Rotationsindex**

$$I(c) := (2\pi)^{-1} \int_a^b k(t) dt$$

³¹Sogar mit einem sog. Diffeomorphismus, siehe Definition 6.6

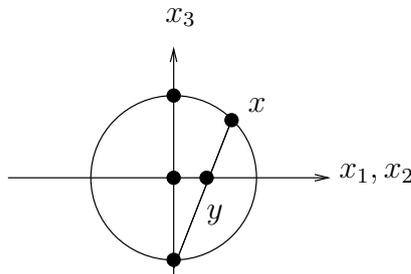


Abbildung 12.3: Stereographische Projektion des vom Südpol verschiedenen Punktes x der Sphäre S^2 auf den Punkt y der Ebene

eine ganze Zahl.

Beweis: Nach Definition ist $\int_a^b k(t) dt = \int_a^b \varphi'(t) dt = \varphi(b) - \varphi(a)$. Da (mod 2π) diese Winkel übereinstimmen (man setzt hier φ stetig fort!), ist $I(c) \in \mathbb{Z}$. \square

Für beliebig regulär parametrisierte geschlossene Kurven setzt man $I(c) := \frac{1}{2\pi} \int_a^b k(t) \|c'(t)\| dt$, was die Definition nach Bemerkung 6.12.2 invariant bezüglich orientierungserhaltender Transformationen macht.

12.5 Beispiel

1. Für den im mathematisch positiven bzw. negativen Sinn $n \in \mathbb{N}$ -mal durchlaufenen Kreis $c_{\pm}(t) := \begin{pmatrix} R \cos t \\ \pm R \sin t \end{pmatrix}$, $t \in [0, 2\pi n]$ vom Radius $R > 0$ ist $c'_{\pm}(t) = \begin{pmatrix} -R \sin t \\ \pm R \cos t \end{pmatrix}$ und $c''_{\pm}(t) = -c(t)$, also die Krümmung gleich $k_{\pm}(t) = \frac{\pm R^2}{R^3} = \frac{\pm 1}{R}$. Damit ist der Rotationsindex

$$I(c_{\pm}) = (2\pi)^{-1} \int_0^{2\pi n} k_{\pm}(t) \|c'_{\pm}(t)\| dt = \pm n.$$

2. Ist die Spur $c(I)$ der (bis auf Anfangs- und Endpunkt) injektiven Kurve $c : I \rightarrow \mathbb{R}^2$ von der Form ∞ , dann ist der Rotationsindex $I(c) = 0$. \diamond

In der modernen Festkörper- und der Teilchenphysik spielt die Topologie in Form so genannter topologischer Defekte und Quasiteilchen eine wichtige Rolle³².

³²Siehe z.B. P. Chaikin, T. Lubenski: Principles of condensed matter physics. Cambridge University Press, 1995, oder A. Schwarz: Quantum Field Theory and Topology. Springer, 1993

Literatur

- [An1] Knauf, A.: Vorlesung *Analysis I. Skript*, Erlangen 2019
- [BS] Bamberg, P.; Sternberg, S.: *Course in mathematics for students of physics*, Bd. 1 und 2. Cambridge University Press, 1998
- [Bl] Blatter, Ch.: *Analysis I-III*. Springer, 1974
- [Bri] Brieskorn, E.: *Lineare Algebra und analytische Geometrie*. Vieweg 1982
- [Br] Bröcker, T.: *Analysis*, Bd. 1 und 2. Spektrum der Wissenschaft, 1995
- [Ca] do Carmo, M.P.: *Differentialgeometrie von Kurven und Flächen*. Vieweg, 1998
- [DH] Deuffhard, P.; Hohmann, A. *Numerische Mathematik I. Eine algorithmisch orientierte Einführung*. de Gruyter, 2002
- [DB] Deuffhard, P., Bornemann, F.: *Numerische Mathematik II: Gewöhnliche Differentialgleichungen*. de Gruyter, 2002
- [Fi] Fischer, G.: *Lineare Algebra*. Vieweg 2000
- [Fo] Forster, O.: *Analysis*, Bd. 1-3. Vieweg, 1999
- [HW] Hellwig, K.-E.; Wegner, B.: *Mathematik und theoretische Physik*, Bd. 1 und 2. de Gruyter, 1992
- [Hi] Hildebrandt, S.: *Analysis 1*. Springer, 2002
- [Ja] Jänich, K.: *Topologie*. Springer, 1999
- [Kn] Knörrer, H.: *Geometrie*. Vieweg 1996
- [KW] H. Kerner, W. von Wahl: *Mathematik für Physiker*, Springer, 2006
- [MV] Meyberg, K., Vachnauer, P.: *Höhere Mathematik*, Bd 1 und 2. Springer, 1999
- [vW1] Waerden, B.L. van der: *Algebra Bd 1*. Springer 1971
- [Wu] Wüst, R.: *Höhere Mathematik für Physiker*, Bd. 1 und 2. de Gruyter, 1995

Index

- Abbildung
 - adjungierte 34
 - affine 63
 - iterierte 104
 - kontrahierende 104
 - normale 37
 - reguläre 163
- abgeschlossen 78
- abgeschlossene Hülle 78
- Ableitung
 - partielle 120
 - totale 120
- Affinität 63
- ähnlich 12
- algebraisch abgeschlossen 13
- Anfangsbedingung 107
- Anfangswertproblem 49, 107
- antihermesch 30
- antiselbstadjungiert 44
- Banach-Raum 73
- Banachscher Fixpunktsatz 104
- Basis (einer Topologie) 79
- Beschleunigung 96
- beschränkt 87
- Bit 70
- Bogenlänge 95
- Cantor-Menge 168
- Cauchy-Folge 73
- Cauchy–Schwarz–Ungleichung 31
- Contour-Plot 123
- darstellende Matrix 10
- definit 146
- diagonalisierbar 13,
- Diffeomorphismus 94, 162
- Differentialgleichung 19, 49, 99
 - gewöhnliche 99
 - homogene 100
 - lineare 100
 - Ordnung der 100
 - partielle 99
- Differentialoperator 132
- Differentiationsregeln 125
- differenzierbar
 - stetig 128
 - total 120
- Divergenz 132
- Doppelgerade 67
- Dreiecksungleichung 70
- duale Basis 35
- Eigenraum
 - verallgemeinerter 21
- Eigenvektor 37
 - verallgemeinerter 21
- Ellipse 68
- Euklidische Norm 50
- Extremum 145
- Faktorisierungssatz 4
- Fakultät 142
- Fitting-Index 22
- Fixpunkt 104
- Fixpunktmenge 151
- Folge
 - Cauchy- 73
 - Grenzwert 79
- folgenstetig 83
- Fourier-Transformation 43
- Fundamentalsatz der Algebra 4
- Funktion
 - elementare 1
 - ganzrationale 5
 - rationale 5, 9
- Funktionaldeterminante 131
- Geschwindigkeit 50
- Grad 3, 142
- Gradient 124
- Gram-Schmidt-Verfahren 33

Gruppe
 der Affinitäten 63
 orthogonale 40
 speziell orthogonale 40
 speziell unitäre 40
 unitäre 40
 Hamming-Abstand 70
 Häufungspunkt
 einer Folge 75, 79
 einer Menge 78
 Hauptachsentransformation 66
 Hausdorff-Raum 79
 Helix 98
 hermitesch 30
 Hesse-Matrix 144
 Höhenlinie 123
 Homöomorphismus 82
 Hyperbel 67
 Hyperboloid
 einschaliges 69
 zweischaliges 70
 Inneres 78
 Integration
 rationaler Funktionen 7
 Isometrie 39
 Jacobi-Matrix 120
 Jordan-Basis 28
 Jordan-Matrix 28
 Jordan-Zerlegung 25
 kanonisches Skalarprodukt 32
 Kegel 69
 kompakt 85
 Kongruenz 64
 konjugiert-linear 35
 konvex 141
 konvergent 73
 Koordinaten, homogene 100
 kritischer Punkt 146
 kritischer Wert 161
 Krümmung 96, 98
 Kurve 49
 geschlossene 171
 Länge einer 93
 Peano- 93
 Lagrange-Multiplikatoren 165
 Lagrange-Restglied 141
 Laplace-Operator 133
 Leitkoeffizient 4
 lineare Differentialgleichung 19
 Linearfaktor 4
 Lipschitz-Bedingung 107
 lipschitz-stetig
 global 104
 lokal 104
 Lösung einer DGL 107
 allgemeine 101
 singuläre 102
 spezielle 101
 vollständige 101
 Matrizen
 ähnliche 12
 äquivalente 11
 Maximum
 isoliertes 145
 lokales 145
 Metrik 70
 euklidische auf \mathbb{C} 71
 euklidische auf \mathbb{R}^n 71
 metrischer Raum 70
 Minimum 145
 Möbius-Band 162
 Multiindex 142
 Nebenbedingung 165
 Newton-Verfahren 151
 vereinfachtes 154
 nilpotent 13
 Nilpotenzindex 13
 Norm
 äquivalente Normen 89
 Euklidische Norm 50
 Operatornorm 50, 84
 Normalform 11

Nullstelle 4
 Nullstellenmenge 150
 offene Überdeckung 85
 Operatornorm 50, 84
 Ordnung
 einer Differentialgleichung 100
 orientierungserhaltend 95
 orthogonal 39
 Orthogonalbasis 32
 Orthonormalsystem 32
 Parabel 67
 Paraboloid 70
 Parametertransformation 94
 Partialbruchzerlegung 6
 Partition 19
 Pendel 157
 Phasenraum 107
 Picard–Abbildung 110
 Polarisationsidentität 31
 Polarkoordinaten 128, 131
 Polynom 3, 142
 Projektion
 Orthogonal- 33
 quadratische Form 166
 Quadrik 63
 Rang einer quadratischen Form 66
 Rand 78
 Raum
 metrischer 70
 topologischer 77
 vollständiger 73
 reguläre Kurve 91
 regulärer Punkt 146
 regulärer Wert 161
 Richtung 96
 Richtungsableitung 125
 Ring 3
 Rotation 135
 Rotationsindex 171
 Sattel 70, 123, 147
 Satz
 Banachscher Fixpunktsatz 104
 Bolzano–Weierstraß 87
 Heine–Borel 87
 Faktorisierungssatz 4
 Fundamentalsatz der Algebra 4
 Mittelwertsatz 137
 über implizite Funktionen 158
 Picard–Lindelöf 109
 Sylvester 147
 Trägheitssatz von Sylvester 66
 Taylor 141
 Weierstraß 92
 Schnittwinkel 95
 Schrödinger-Operator 48
 selbstadjungiert 44
 Signatur 66
 Skalarprodukt 30
 Sphäre 165
 Spur
 einer Kurve 49
 der Jacobi-Matrix 132
 stetig 82
 Taylor-Formel 141
 Teilraum 81
 Topologie 77
 diskrete 78
 indiskrete 78
 induzierte 81
 metrische 77
 Torus 81
 Umgebung 78
 unitär 39
 Untermannigfaltigkeit 161
 Vektor
 erzeugender 15
 primitiver 17
 Vektorfeld 115
 zeitabhängiges 107
 zyklisch 15
 Zykloide 94
 Zylinder 68